

IntechOpen

Fluid Dynamics, Computational Modeling and Applications

Edited by L. Hector Juarez



FLUID DYNAMICS, COMPUTATIONAL MODELING AND APPLICATIONS

Edited by **L. Hector Juarez**

Fluid Dynamics, Computational Modeling and Applications

<http://dx.doi.org/10.5772/2403>

Edited by L. Hector Juarez

Contributors

Xing-Qian Peng, Daniel Marinho, Tiago M. Barbosa, Vishveshwar Rajendra Mantha, Abel Rouboa, Antonio Silva, Sujudran Balachandran, Francesco Battista, Misuri Tommaso, Mariano Andrenucci, L. Hector Juarez, Maria Luisa Sandoval, Jorge Lopez, Rafael Resendiz, Hwataik Han, Katsuya Nagayama, Renat Sultanov, Dennis Guster, Kleiber Bessa, Daniel Legendre, Akash Prakasan, Habib Alehossein, Alfredo Iranzo, Felipe Rosa, Antonio Salva, Norshah H. Shuaib, Prem Gunnasegaran, Muhammad Fairuz Abdul Jalal, Hussein Mohammed, Eri Sandhita, Joao Carvalho, Peter Vidmar, Stojan Petelin, Marko Perkovič, Floriana La Marca, Monica Moroni, Antonio Cenedese, Stanislas Lopata, Pawel Oclon, Kazui Fukumoto, Claudio Bruno, Yoshifumi Kuriyama, Kenichi Yano, Seishi Nishido, Timothy Madden, Takahisa Yamamoto, Seiichi Nakata, Tsuyoshi Yamamoto, Tsutomu Nakashima, Marco Marcon, Marius Stoia-Djeska, Carmen Anca Safta, Jyh-Tong Teng, Sun Ze, Guimin Lu, Xingfu Song, Shuying Sun, Yuzhu Sun, Jin Wang, Jianguo Yu, Yusuke Tahara, Yutaka Masuyama, Toichi Fukasawa, Masanobu Katori, Vinod Kumar, Victor Udoewa, Hector Perez-De-Tejada, Cosmas Mafusire, Andrew Forbes

© The Editor(s) and the Author(s) 2012

The moral rights of the and the author(s) have been asserted.

All rights to the book as a whole are reserved by INTECH. The book as a whole (compilation) cannot be reproduced, distributed or used for commercial or non-commercial purposes without INTECH's written permission.

Enquiries concerning the use of the book should be directed to INTECH rights and permissions department (permissions@intechopen.com).

Violations are liable to prosecution under the governing Copyright Law.



Individual chapters of this publication are distributed under the terms of the Creative Commons Attribution 3.0 Unported License which permits commercial use, distribution and reproduction of the individual chapters, provided the original author(s) and source publication are appropriately acknowledged. If so indicated, certain images may not be included under the Creative Commons license. In such cases users will need to obtain permission from the license holder to reproduce the material. More details and guidelines concerning content reuse and adaptation can be found at <http://www.intechopen.com/copyright-policy.html>.

Notice

Statements and opinions expressed in the chapters are those of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

First published in Croatia, 2012 by INTECH d.o.o.

eBook (PDF) Published by IN TECH d.o.o.

Place and year of publication of eBook (PDF): Rijeka, 2019.

IntechOpen is the global imprint of IN TECH d.o.o.

Printed in Croatia

Legal deposit, Croatia: National and University Library in Zagreb

Additional hard and PDF copies can be obtained from orders@intechopen.com

Fluid Dynamics, Computational Modeling and Applications

Edited by L. Hector Juarez

p. cm.

ISBN 978-953-51-0052-2

eBook (PDF) ISBN 978-953-51-5650-5

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,100+

Open access books available

116,000+

International authors and editors

120M+

Downloads

151

Countries delivered to

Our authors are among the
Top 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

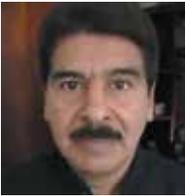
Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Meet the editor



Professor L. Hector Juarez received his Ph.D. in mathematics from the University of Houston (USA), in 1996. He also was a postdoctoral fellow and assistant visiting professor of the same institution during the period 1999-2002. Since 2002 he is full professor of applied mathematics at the Universidad Autonoma Metropolitana-Iztapalapa (UAM-I) in Mexico City, and currently is a Fellow of the National Research System in Mexico, level 2. He has participated in numerous research projects and published several research articles, reports, conference proceedings and book chapters, mainly in CFD, numerical solution of PDE, mathematical modeling and computer simulation. He has been supervisor of several graduate students, and has participated in numerous academic committees. It has also helped organize numerous national and international meetings in Mexico, and he has participated in many others abroad. Professor Juárez was recently elected member of the Main Board of the Mexican Mathematical Society for the period 2012-2014.

Contents

Preface XIII

Part 1 Winds, Building, and Risk Prevention 1

- Chapter 1 **Study of Wind-Induced Interference Effects on the Fujian Earth-Buildings 3**
Peng Xingqian, Liu Chunyan and Chen Yanhong
- Chapter 2 **Mass-Consistent Wind Field Models: Numerical Techniques by L2-Projection Methods 23**
L. Héctor Juárez, María Luisa Sandoval, Jorge López and Rafael Reséndiz
- Chapter 3 **Ventilation Effectiveness Measurements Using Tracer Gas Technique 41**
Hwataik Han
- Chapter 4 **Fluid Dynamic Models Application in Risk Assessment 67**
Peter Vidmar, Stojan Petelin and Marko Perkovič
- Chapter 5 **Sail Performance Analysis of Sailing Yachts by Numerical Calculations and Experiments 91**
Y. Tahara, Y. Masuyama, T. Fukasawa and M. Katori

Part 2 Multiphase Flow, Structures, and Gases 119

- Chapter 6 **A Magneto-Fluid-Dynamic Model and Computational Solving Methodologies for Aerospace Applications 121**
Francesco Battista, Tommaso Misuri and Mariano Andrenucci
- Chapter 7 **Mechanics of Multi-Phase Frictional Visco-Plastic, Non-Newtonian, Depositing Fluid Flow in Pipes, Disks and Channels 151**
Habib Alehossein

- Chapter 8 **Three Dimensional Simulation of Gas-Radiation Interactions in Gas Lasers** 175
Timothy J. Madden
- Chapter 9 **Fluid-Structure Interaction** 195
Stoia-Djeska Marius and Safta Carmen-Anca
- Chapter 10 **Study on Multi-Phase Flow Field in Electrolysis Magnesium Industry** 217
Ze Sun, Guimin Lu, Xingfu Song, Shuying Sun, Yuzhu Sun, Jin Wang and Jianguo Yu
- Chapter 11 **Fluid-Structure Interaction Techniques for Parachute** 239
Vinod Kumar and Victor Udoewa
- Part 3 Heat Transfer, Combustion, and Energy** 263
- Chapter 12 **Fluid Flow in Polymer Electrolyte Membrane Fuel Cells** 265
Alfredo Iranzo, Antonio Salva and Felipe Rosa
- Chapter 13 **Heat Transfer Enhancement in Microchannel Heat Sink Using Nanofluids** 287
P. Gunnasegaran, N.H. Shuaib, H.A. Mohammed, M.F. Abdul Jalal and E. Sandhita
- Chapter 14 **Modelling and Optimizing Operating Conditions of Heat Exchanger with Finned Elliptical Tubes** 327
Stanisław Łopata and Paweł Ocioń
- Chapter 15 **Simulation of H₂-Air Non-Premixed Flame Using Combustion Simulation Technique to Reduce Chemical Mechanisms** 357
Kazui Fukumoto and Yoshifumi Ogami
- Chapter 16 **Nuclear Propulsion** 381
Claudio Bruno
- Chapter 17 **Fluid Dynamics in Microchannels** 403
Jyh-tong Teng, Jiann-Cherng Chu, Chao Liu, Tingting Xu, Yih-Fu Lien, Jin-Hung Cheng, Suyi Huang, Shiping Jin, Thanhtrung Dang, Chunping Zhang, Xiangfei Yu, Ming-Tsang Lee, and Ralph Greif
- Part 4 Medical and Biomechanical Applications** 437
- Chapter 18 **Modelling Propelling Force in Swimming Using Numerical Simulations** 439
Daniel A. Marinho, Tiago M. Barbosa, Vishveshwar R. Mantha, Abel I. Rouboa and António J. Silva

- Chapter 19 **Surfactant Analysis of Thin Liquid Film in the Human Trachea via Application of Volume of Fluid (VOF)** 449
Sujudran Balachandran
- Chapter 20 **3D Particle Simulations of Deformation of Red Blood Cells in Micro-Capillary Vessel** 463
Katsuya Nagayama and Keisuke Honda
- Chapter 21 **Numerical Modeling and Simulations of Pulsatile Human Blood Flow in Different 3D-Geometries** 475
Renat A. Sultanov and Dennis Guster
- Chapter 22 **Biomechanical Factors Analysis in Aneurysm** 493
Kleiber Bessa, Daniel Legendre and Akash Prakashan
- Chapter 23 **Assessment of Carotid Flow Using Magnetic Resonance Imaging and Computational Fluid Dynamics** 513
Vinicius C. Rispoli, Joao L. A. Carvalho,
Jon F. Nielsen and Krishna S. Nayak
- Chapter 24 **Numerical Simulation for Intranasal Transport Phenomena** 537
Takahisa Yamamoto, Seiichi Nakata,
Tutomu Nakashima and Tsuyoshi Yamamoto
- Part 5 Additional Important Themes** 555
- Chapter 25 **Fluid-Dynamic Characterization and Efficiency Analysis in Plastic Separation of the Hydraulic Separator Multidune** 557
Floriana La Marca, Monica Moroni and Antonio Cenedese
- Chapter 26 **Optimization of Pouring Velocity for Aluminium Gravity Casting** 575
Y. Kuriyama, K. Yano and S. Nishido
- Chapter 27 **Fluid Dynamics Without Fluids** 589
Marco Marcon
- Chapter 28 **Fluid Dynamics in Space Sciences** 611
H. Pérez-de-Tejada
- Chapter 29 **Aero - Optics: Controlling Light with Air** 631
Cosmas Mafusire and Andrew Forbes

Preface

The content of this book covers several up-to-date topics in fluid dynamics, computational modeling and its applications, and it is intended to serve as a general reference for scientists, engineers, and graduate students. The book is comprised of 30 chapters divided into 5 parts, which include: winds, building and risk prevention; multiphase flow, structures and gases; heat transfer, combustion and energy; medical and biomechanical applications; and other important themes. This book also provides a comprehensive overview of computational fluid dynamics and applications, without excluding experimental and theoretical aspects.

The edition of this book was made possible thanks to the contribution of many scientists, and researchers in the field of fluid dynamics, and also thanks to the initiative of InTech, and the outstanding professional work of its staff and editors. This book covers a wide range of topics related to fluid mechanics, such as: meteorology, energy, aerospace, heat transfer, civil engineering, environmental, medicine, physiology, micro-fluids, and industry. In particular, the reader will find some specific chapters about ventilation, building, sailing yachts, heating, cooling, combustion, swimming, blood flow, arterial diseases, breathing and intranasal flow, fuel cells, casting, concrete slurries, parachutes, magnesium production, and plastic separation, among others. Some other specific topics available are: nuclear propulsion, fluid structure interaction, solar winds, aero-optics, gases, chemical lasers, and wind field recovery. There is also an interesting chapter about how to apply CFD techniques to solve problems, which are not directly related to fluid dynamics.

Dr. L. Hector Juarez
Department of Mathematics
U.A.M.-I., Mexico City
México
University of Houston
Department of Mathematics, Houston, Texas
USA

Part 1

Winds, Building, and Risk Prevention

Study of Wind-Induced Interference Effects on the Fujian Earth-Buildings

Peng Xingqian, Liu Chunyan and Chen Yanhong
*College of Civil Engineering, Huaqiao University, Quanzhou
China*

1. Introduction

As the only large-scale rammed-earth dwelling worldwide, Fujian earth-building gets much attention for its unique style, grand scale, ingenious structure, abundant cultural connotation, reasonable layout and the concept of keeping harmony with nature. In July 2008, Chuxi earth-building cluster, Hongkeng earth-building cluster, Gaobei earth-building cluster, Yangxiang Lou and Zhenfu Lou were listed among world heritage. They are important parts of Fujian earth-building with a long history, vast distribution, various types and rich connotation. Earth-building culture roots in oriental ethical relations and provides specific historical witness to traditional style of living by clansman, It is a unique achievement by employing rammed raw earth in large scale with "outstanding universal value".

Because of the high frequency of typhoon between summer and autumn in mountainous areas of the western Fujian, buildings in high and open areas often get serious damages, as shown in figure 1. In 2006, the 4th cyclone "BiLiSi" brought heavy damage to Daoyun Lou which is 400 years old. Six rooms in it collapsed, several tiles were blew off and the total number of damaged rooms reached more than 10. As one of the world cultural heritages, the protection, utilization and development of Fujian earth-building is the major issue to be deal with. Presently, the theory study for wind-resistant of low buildings is still not enough, the failure mechanism hasn't been studied thoroughly. For low buildings often appear in the form of groups, the related studies are even less. So research of wind interference effect in earth-building groups can not only fill the blank of research studies but also put forward some corresponding measures for protection of the world cultural heritage.

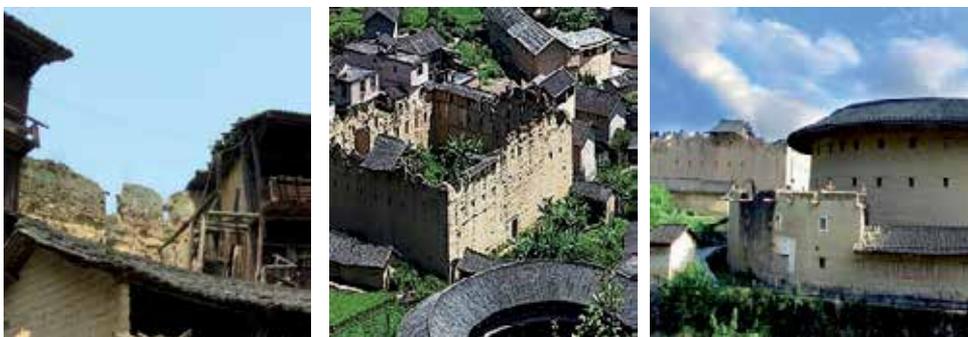


Fig. 1. Storm damage to the roof of earth-buildings

2. The influencing factors of wind interference effect

Fujian earth-buildings are often located in the form of groups, as shown in figure 2. Surface wind load is heavily influenced by the surrounding buildings and the main influencing factors include the height of the building, the relative position between buildings, section size and shape, the wind speed and wind direction, the type of wind field, etc.



Fig. 2. Tianluokeng earth-building cluster

2.1 The influence of landscape

Roughness of the landscape has a great influence on the structure wind loads, And under different wind, the interference effects between the buildings are quite different from the wind. Compared to the isolated building at the open area, Walker and Roy^[1] found that the average load, peak load and bending moment are increased in the urban area of wind load. Under the open countryside and suburban areas of different topography, Case. P.C ^[2] study on the transient external pressure of the gable roof building experimental. He pointed out that wind load at buildings in the city suburbs is lower than that in the open landscape. And the arrangement of groups help reduced the load on a single. Blessmann^[3] studied variety of landscape effects of wind interference, The results show that the moderating effect of the open landscape is most evident. Because of the turbulence is relatively low in open landscape, The pulse of wake in the upstream building has a strong correlation, Therefore, wind loads on downstream buildings caused by increased.

2.2 The influence of building's width and height

The width of the windward side of the housing has great influence on eddy size behind the leeward side, And the size of the upstream building construction also affect the downstream response of wind interference. Taniike^[4] study the Wind-induced interference effects under low turbulence contour and different section size in square columns, He pointed out that the average wind load to the along wind will decline with the increase size of upper building's section, and that dynamic response to the along wind will increase with increasing section width. Under normal circumstances, when the height of adjacent buildings is equal to or greater more than half of the height of the building, we should take into account the mutual

interference effects between groups, and ignoring the interference of the building which less than half the height of buildings^[5].

2.3 The influence of number of buildings

In previous tests of wind interference, we remain in the interference effect between two buildings for a long time, and rarely consider the interference effects between more than three buildings. Professor Xiezhuanngning^[6] studied the wind-induced disturbance response between the three buildings, and analysis the interference of the characteristics and mechanism by neural networks, spectral analysis and statistics. The results show that the combined effects of the two buildings would be stronger than a single building. Under the landscape of Class B, Interference factor of two buildings would be increased more than 79% of a single building.

2.4 The influence of spacing of building group

Holmes^[7] study the wind characteristics of the street on both sides of the building, and found upstream of the shadowing effect and the building construction the distance between a great relationship. Zhao qingchun^[8] have studied the low gable roof wind-induced interference effect, found group effect on the windward roof pressure front and rear degree of influence. When the workshop's distance was $2b$ the interference obviously. The wind tunnel experiments show that: when buildings adjacent cross-wind side by side, the gap flow effect presence in the region when $S / D \leq 2$, (S for the building spacing, D for the side of building); When Buildings are along the windward, shielding effect exists in $S / D \leq 3$ regions.

2.5 The impact of the wind stream

Tsutsumi, J.^[9] conducted a model test in different wind direction of wind load characteristics of the group, received the average wind pressure coefficient of the windward and leeward of buildings' surface. Compare and analyze the model's average wind pressure coefficient under different architectural layout, Get the average wind pressure coefficient varies with the change of wind direction. Generally speaking, the flow separation zone will increased by the skew wind, air disturbance will more severe, and the flow will become more complex.

3. Analysis of wind interference effect between two Earth-buildings

3.1 Calculation model

Fujian earth-building is ring-shaped with one-ring building or more. Here we simplify the model with ignoring the hallway, ancestral temple and such subsidiary structures. In the study of wind-induced interference, we select the typical circular earth-building to do the numerical simulation. The diameter of the biggest circular earth-building is 82 meters and the smallest is 17 meters while the common number of stories is $2 \sim 4$ ^[10]. The research model this article selects is 28 meters in diameter, 3 layers, 11.2 meters high and under conditions as shown in fig. 3. This chapter mainly studies the characteristics of wind load and the air flowing field of circular earth-buildings, the change rule under different spacing of which are also explored.

Here we major change the windward spacing between two earth-buildings and wind direction, as shown in fig.03. S stands for spacing between two earth-buildings, D stands for horizontal scale (the diameter of the bigger circular earth-building), n is valued respectively by 0.15, 0.5, 0.75, 0.25, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5 and 4.0. The characteristics of Fujian earth-buildings are: clay wall, general 1-m \sim 1.5 m of wall thick, chines-style tile roof and big pick eaves. When

typhoon comes, the big pick eaves are most easily swept away, leading to the damage of whole roof structure. Therefore, the roof zoning plan of the research object (i.e. the disturbed body) is shown in figure 4. The dividing is in a counterclockwise direction and the roof is divided into eaves part and ridge part. The upper surface of outside carry eaves are signed respectively by WTS1 ~ WTS8, the lower surface by WTX1 ~ WTX8. The upper surface of inside carry eaves are signed respectively by NTS1~ NTS8, the lower surface by NTX1 ~ NTX8. The ridge part is divided into the inside part and outside part, and they are signed respectively by NJ1~ NJ8 and WJ1~ WJ8.

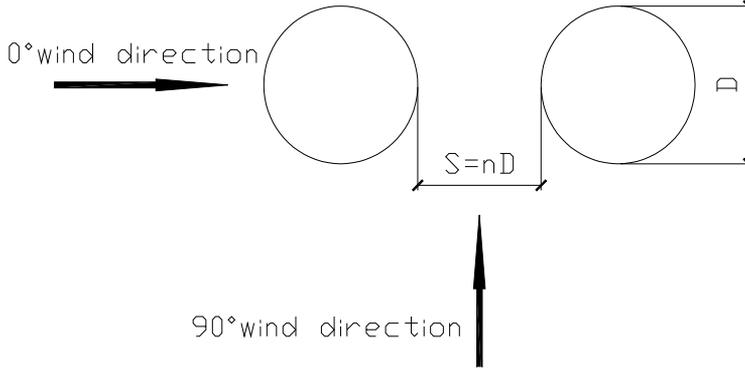


Fig. 3. Plan of Earth-building and wind direction

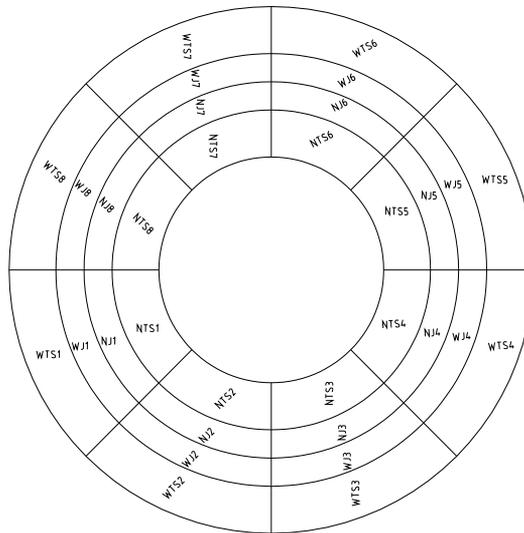


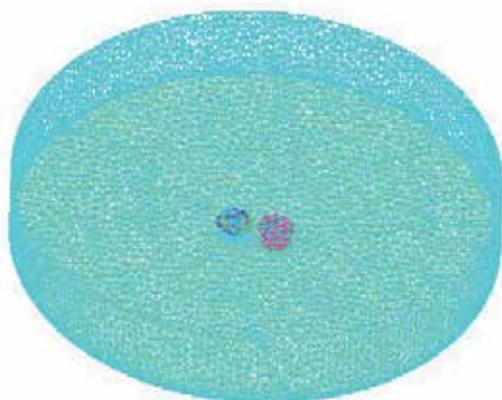
Fig. 4. Roofing zoning

Settings of basic parameters in numerical simulation: according to reference literatures^[11], domains' size can be set as following: $B \times L \times H = 600\text{m} \times 500\text{m} \times 100\text{m}$, its blocking rate is 0.6%, meets requirements. As shown in figure 5, the whole calculation domain is divided into two parts: internal area and external area. The cylinder with 380m diameter is the internal area-domain 1, the other part is external area-domain2. Domain 1 use the tetrahedron meshes

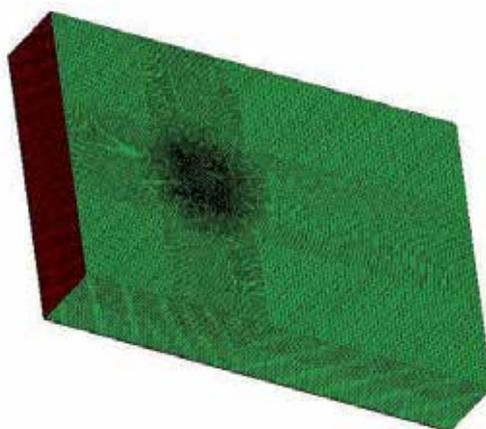
while domain 2 adopts convergence higher structured hexahedral meshes. Fujian Earth-building is located in rural mountain areas, $h_0 = 10\text{m}$, $v_0 = 5.35\text{m/s}$, Fujian Earth-building area belongs to the class B landform, roughness index $\alpha = 0.16$. Turbulence intensity $I(z) = 0.194$, turbulence integral scale $L_u = 60.55\text{m}$, kinetic energy $k(z)$ and dissipation rate $\varepsilon(z)$ are adopted as the following form:

$$\begin{cases} k(z) = 0.5 \times [I(z) \times \bar{u}(z)]^2 \\ \varepsilon(z) = \frac{4C_\mu^{3/4} k(z)^{3/2}}{KL_u} \end{cases} \quad (1)$$

The surface of buildings use non-slip wall, the two sides and top surface of the numerical wind tunnel use free gliding wall, the outlet of the numerical wind tunnel use open pressure export. This paper argues turbulence is fully development (The static pressure is zero). Turbulence model adopt shear stress transport model (SST $k - \omega$ model).



(a) Meshing of internal domain 1



(b) Meshing of internal domain 2

Fig. 5. Meshing of domain

3.2 Analysis of wind characteristic

This paper adopt every 45° wind direction to do wind interference analysis, for symmetry of the structure, three conditions were simulated in this paper under the same spacing. This paper analyzed wind pressure coefficient at local wind vector at Earth-building 2/3 highly level profile and centre vertical profile, and contrasted wind pressure coefficient between monomer Earth-building and group Earth-building

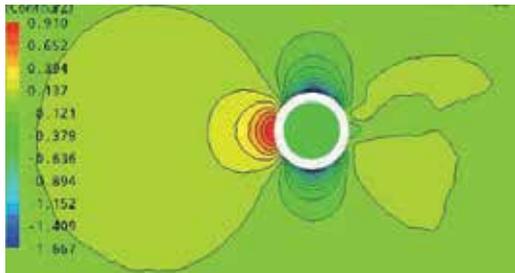
3.2.1 0° wind direction

At the windward area, flow has a positive stagnation point at 2/3 highly level profile, from the stagnation point airflow radiate outward^[9]. At the area above the point, the current rise upward and beyond Earth-building roof top; at the area below the point, airflow downward and flow to the ground. So this paper choose 2/3 highly level profile to discuss the wind field characteristics. Meanwhile, this paper select of center vertical profile as features

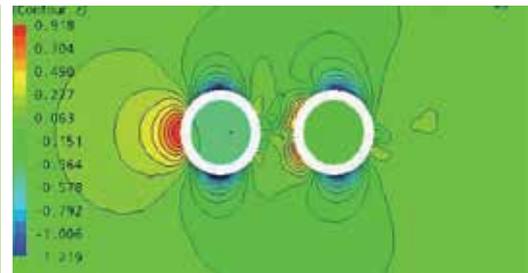
surface, analyze flow field characteristics between Earth groups Building through the observation of wind pressure coefficient graph of this vertical profile.

(1) Level cross section at 0° wind direction

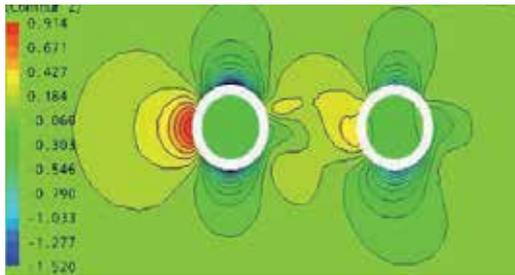
Fig. 6 shows the isocline of the air pressure coefficient at 2/3 highly level profile. From Fig 6(a) we can see the isocline wind pressure coefficient is very plump at the windward area and the two sides of single building. Wind pressure coefficient is positive in the windward,



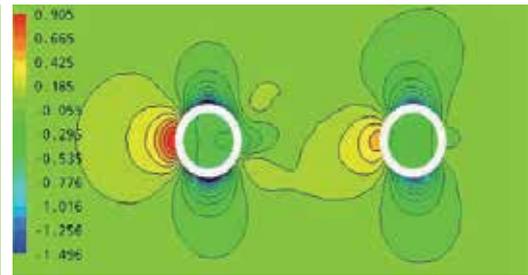
(a) Wind pressure coefficient of cross section of single building



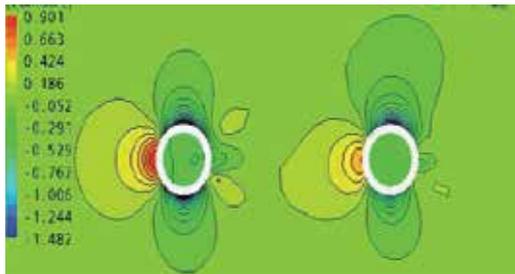
(b) Wind pressure coefficient of cross when $S=0.15D$



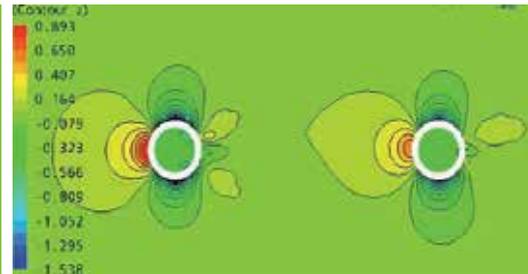
(c) Wind pressure coefficient of cross section when $S=0.75D$



(d) Wind pressure coefficient of when $S=1.5D$



(e) Wind pressure coefficient of cross section when $S=2.0D$



(f) Wind pressure coefficient of cross when $S=3.0D$

Fig. 6. Wind pressure coefficient of 2/3 highly level profile at 0° wind direction

and the closer to the building, the bigger it is. While this coefficient is negative in the side, and the closer to the building, the bigger the absolute value is. But in the leeward surface, we can see two air pressure coefficient equivalent envelope for the two vortexes formed at leeward. Figure 6 (b) is air pressure coefficient graph of two spacing is $0.15D$ circular Earth-building. Due to the distance between the two buildings smaller, flow between the two

buildings is more complex. Air pressure coefficient isocline mutual surrounded relatively intense, and the value has reduce trend. Which is especially noteworthy is the wind pressure coefficient isocline of perturbation building is quite different to monomer at windward direction; it appears two isocline large regions. The interfered building is affected by two vortices at the tail of the front Earth-building. When the spacing is $0.75D$, vortex is gradually developed, air pressure coefficient isocline between two Earth-buildings is linked together, and mutual interference is still evident. When the spacing is $1.5D$, the whirlpool basically develops fully and the isocline is tending to independence. When the spacing continues to increase to $3.0D$, development of whirlpool is fully, air pressure coefficient isocline around two Earth-buildings is full independence and tend to monomer conditions. At 0° wind direction, generally speaking, flow field of downstream Earth-building changes greatly, downstream Earth-building under the more obvious influence.

(2) Center vertical profile of 0° wind direction

Figure 7 gives wind pressure coefficients isocline of center vertical profile in different spacing.



(a) Wind pressure coefficient of center vertical vertical profile



(b) Wind pressure coefficient of center vertical profile when $S = 0.15D$



(c) Wind pressure coefficient of center vertical profile when $S = 0.75D$



(d) Wind pressure coefficient of center vertical profile when $S = 0.75D$



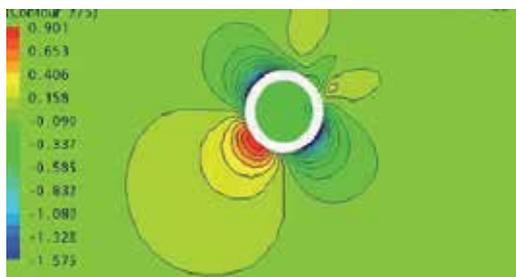
(e) Wind pressure coefficient of center vertical profile when $S = 2D$



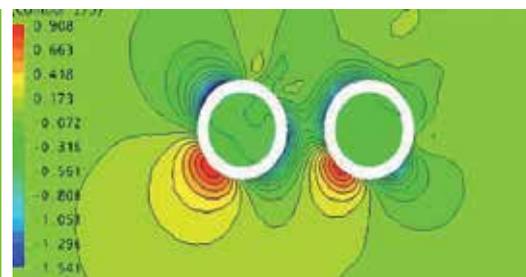
(f) Wind pressure coefficient of center vertical profile when $S = 3D$

Fig. 7. Wind pressure coefficient of central vertical profile of 0° wind direction

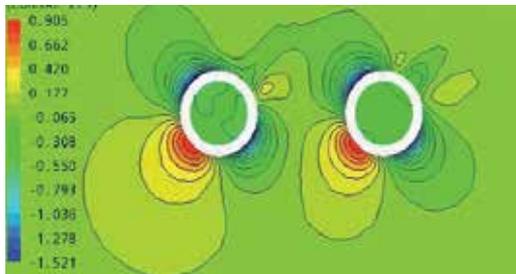
From figure 7 we can see that wind pressure coefficients of Earth-buildings center vertical profile are similar with one, when spacing for 3.0 D. Wind pressure coefficients isocline appear separation phenomenon is quite serious in the external roofs, where the separation point expose many isocline. Under the surface of wind pressure coefficients significantly greater than upper one, which above is negative, the other is positive in the external roofs. Wind pressure coefficients of upper and under surface is close in the external roofs, wind pressure coefficients of upper and under surface almost to zero in the internal roofs, when they are in the leeward flow fields. When both ones spacing is 0.15 D, The prevailing wind direction of wind field and leeward are significantly different, the prevailing flow fields is not affected, and drafting leeward surface whirlpool didn't develop completely Because of the stop function behind Earth-buildings. Whirlpool gradually development, Earth-buildings mutual interference slowly reduce, as spacing is increasing, finally wind field becomes into a monomer.



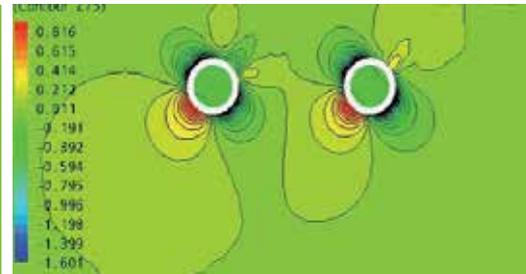
(a) Wind pressure coefficient of cross section of single building



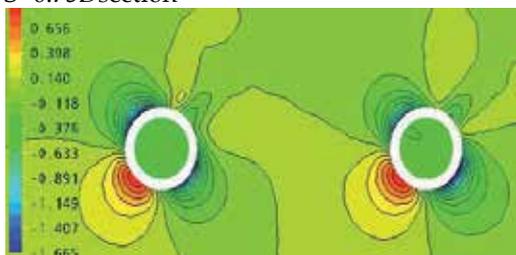
(b) Wind pressure coefficient of cross section when $S=0.15D$



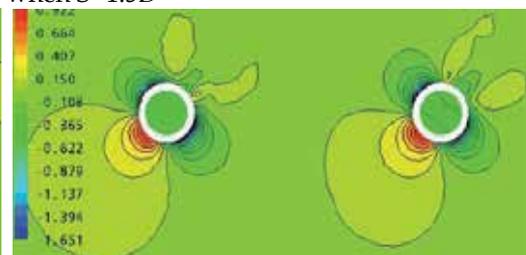
(c) Wind pressure coefficient of cross when $S=0.75D$ section



(d) Wind pressure coefficient of cross section when $S=1.5D$



(e) Wind pressure coefficient of cross section when $S=2.0D$



(f) Wind pressure coefficient of cross section when $S=2.0D$

Fig. 8. Wind pressure coefficient of 2/3highly level profile at 45° wind direction

3.2.2 45° wind direction

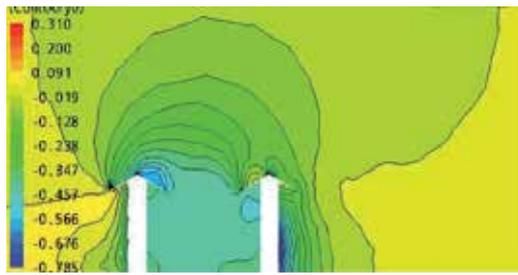
(1) Level cross section of 45° wind direction

Figure 8 45° wind direction is given level of the wind pressure coefficients cross section in 2/3 housing height place, we can see that Earth-buildings wind field changes significantly around in different wind direction for monomer Earth-building, the situation is similar with above, here is not to say much. For two Earth-buildings speaking, when spacing is 0.15 D, oblique flow fields makes Earth-buildings both sides have larger wind speed, but it is affected behind Earth-buildings, wind pressure coefficients isocline have inter-permeation by each other, and is very strong between two Earth-buildings wind pressure isocline with monomer markedly different in two Earth-buildings adjacent area and leeward surface for the front of Earth-buildings, drafting produces whirlpool is impeded, which leading to wind pressure coefficients reduce, and most regional present negative, interference phenomenon is seriously in the leeward surface Earth-buildings also wind pressure isocline with monomer markedly different in two Earth-buildings adjacent area and leeward surface for Behind Earth-buildings, but wind pressure changes very little in lateral area.

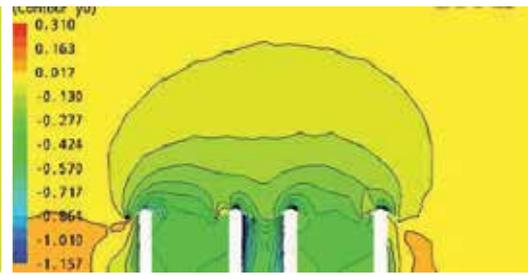
When the spacing becomes larger between two Earth-buildings, and from spacing 0.75D to 2.0D, with drafting place whirlpool developed slowly in the front of Earth-buildings, wind pressure coefficients isocline tend to be independent, interference become weak. When spacing for 3.0 D, interference has not obvious, the flow fields around the Earth-buildings is similar with monomer.

(2) Center vertical profile of 45° wind direction

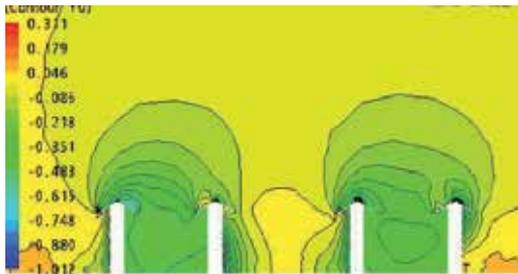
In figure 9, 45° wind direction are given under different spacing vertical section center air pressure coefficient isocline map, From figure 9 (a) which can be seen, air pressure coefficient value of Earth-buildings in the windward side is lesser and negative, near the Earth-buildings metopic air pressure coefficient absolute value increases, in the outside carry eaves, separated phenomenon of air pressure coefficient appeared. It's all negative value in fluctuation pick eaves. Both internal and external roof are affected by negative pressure, and the internal roof endure a bigger negative pressure. The leeward side is in negative pressure area, because of the blocking by Earth-buildings windward surface, wind pressure reduced. Fluctuation pick eaves pressure coefficients of the inside carry eaves are all negative value which offset each other. Outside carry eaves fluctuation surface wind pressure coefficient size differ not quite, the most air pressure coefficient negative value appeared in the leeward side metopic place. If both earth-buildings exist together, the mutual influence is obvious. In figure 9 (b) the spacing is 0.15D, Between two Earth-buildings regional wind pressure isocline showed great difference when monomer, between two Earth-buildings it has even been inter-permeation phenomenon, Behind Earth-buildings air pressure coefficient value in the windward side is negative and its absolute value increases of the monomer Earth-building. It shows that when two Earth-buildings interact with each other, the buildings in the downstream are in the area of architectural drafting upstream effect. The influence of its surface is opposite bigger. As spacing increase, center profile around two earth-buildings distributions of air pressure mutual interference gradually decreased, and change contour tend to monomer condition. When spacing is 3.0D, wind pressure coefficient changes contour line in center vertical of Earth-buildings is similar with monomer condition.



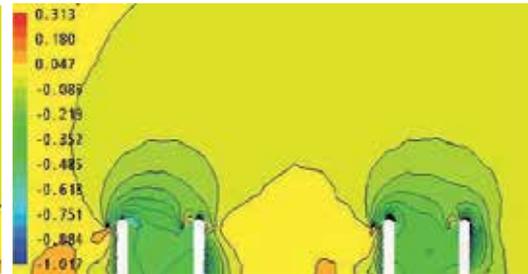
(a) Wind pressure coefficient of center profile of single building



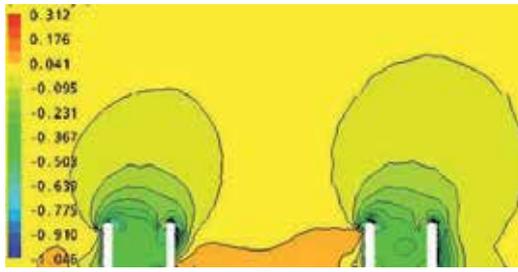
(b) Wind pressure coefficient of center vertical profile when $S = 0.15D$



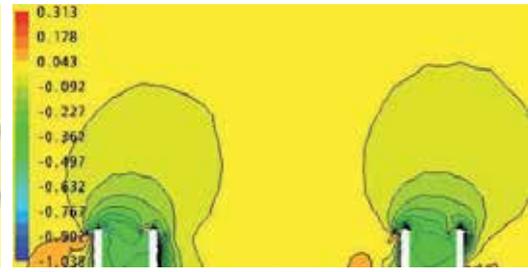
(c) Wind pressure coefficient of center vertical profile when $S = 0.75D$



(d) Wind pressure coefficient of center vertical profile when $S = 1.5D$



(e) Wind pressure coefficient of center vertical profile when $S = 2D$



(f) T wind pressure coefficient of center vertical profile when $S = 3D$

Fig. 9. 45° direction Angle of wind pressure coefficient vertical profile central figure

3.2.3 90° direction angle

Suppose define 0° direction angle as serial, adobe layout 45° direction angle is inclined column, then 90° direction angle, think adobe arrangement as coordination. Through the previous analysis, we know that , with the increasing distance, mutual interference will gradually decrease, buildings' field which surround by also tend to flow around the single Earth-building conditions.

(1) Level cross section of 90°wind direction

From Figure 10 (a),we can see that in 2/3 single adobe houses at the height of the level of cross-section of the wind pressure coefficient contour maps is same with 0° wind direction, on the windward side and side lines are full, the wind pressure coefficient is positive for integrity, and the more close from the adobe metopic walls ,the greater the pressure

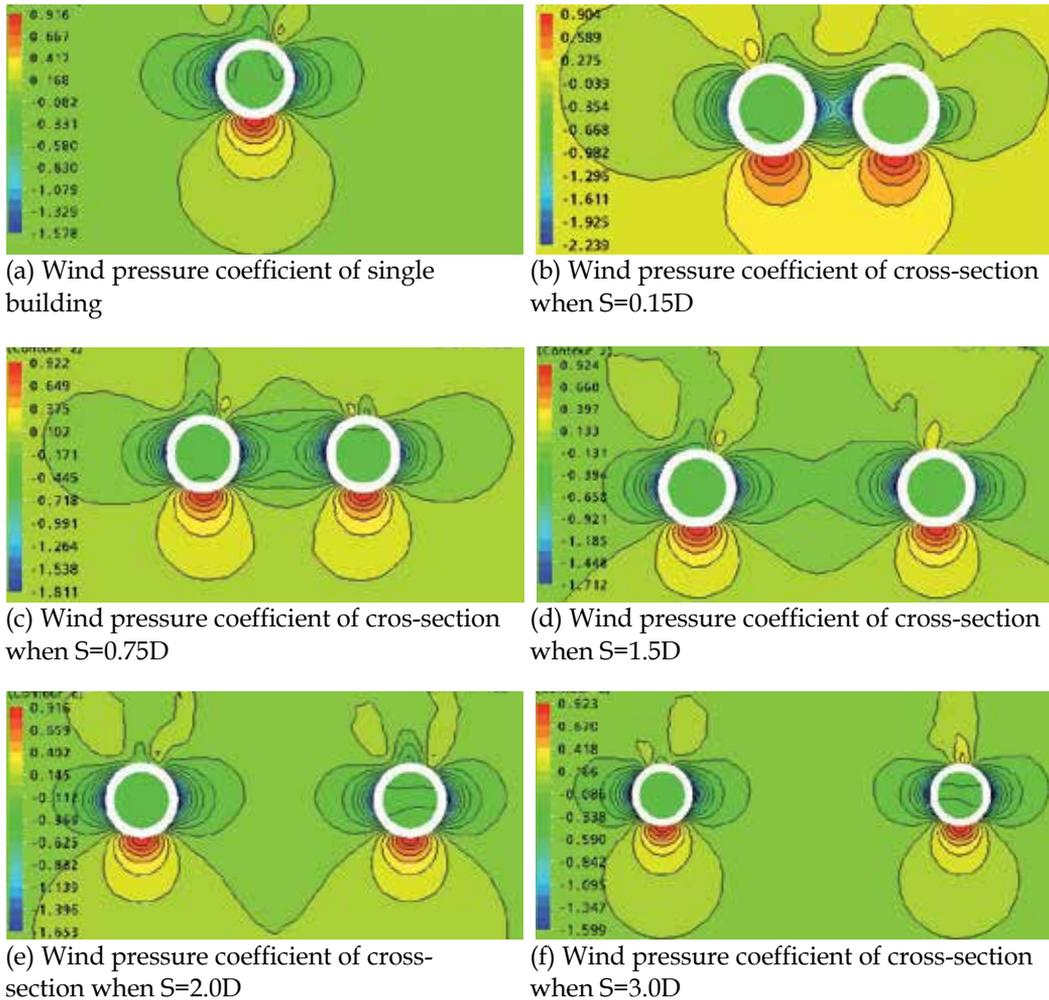
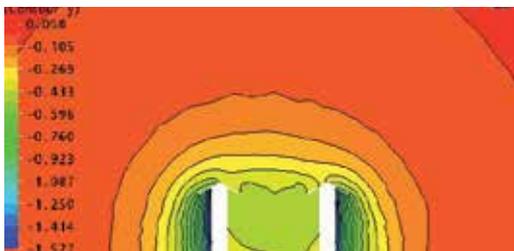


Fig. 10. Wind pressure coefficient of 2/3highly level profile at 45° wind direction

coefficient in the adobe negative side, the greater the closer the absolute value of the earth-building wall, or even 1.5. But in the leeward surface, due to the drafting place formed two swirls, we can see two wind pressure coefficient equivalent envelope. When there are two circular Earth-buildings, the spacing is $0.15D$ according to figure 10 (b), air flow are the prevailing wind direction, due to shunt bypass side collision between the smaller ones, adobe air spacing interaction, air pressure coefficient negative, and absolute 2.239, at maximum achieve isocline wind pressure coefficient, and mutual surrounded relatively intense numerical more monomer adobe has the tendency of increase. Along with the increasing of the spacing distance, two $0.75D$ adobe air pressure coefficient between each other 1.811, to an absolute value of interference still obvious, spacing for $1.5D$, isocline tend to independence, air pressure coefficient absolute 1.712, when spacing continue to increase to $3.0D$, two adobe air pressure coefficient isocline around almost completely independent, air pressure coefficient for 1.599, with monomer absolute adobe air pressure coefficient conditions are 1.578 already smaller maximum absolute value.

(2) Center vertical profile of 90° wind direction

Figure 11 is a different spacing center vertical section, air pressure coefficient isocline can see from figure 11, vertical center section on either side of the air pressure coefficient about isocline obvious symmetry, airflow around side of the surface wind pressure coefficient for exterior wall lateral negative, and the farther from metopic, the small wind pressure coefficient absolute YanXia outside carry with external surface wind pressure coefficient of the measured wind pressure coefficient on the side, almost the same, pick up within the surface wind pressure coefficients were coping negative, but under the surface for the absolute value is opposite bigger. Adobe, adobe has two air around the isocline except in two adjacent area changes remarkably, adobe big changes in other areas. When spacing is 0.15 D, air pressure coefficient isocline surrounded very intense, adjacent area outside carry eaves surface wind pressure coefficient negative, and more monomer when absolute value change, adobe air next pick eaves coefficient is bigger, near the Earth-buildings absolute value change big trend, maximum achieve 2.24, metopic isocline under the changing trends



(a) Wind pressure coefficient of center vertical of single building

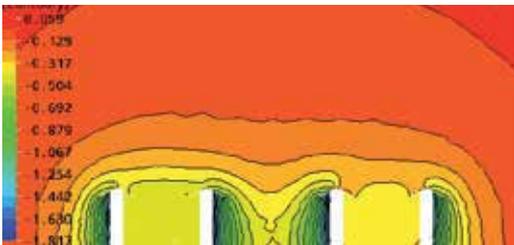
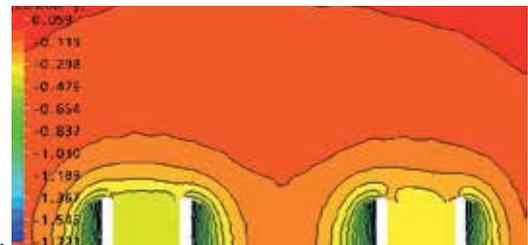
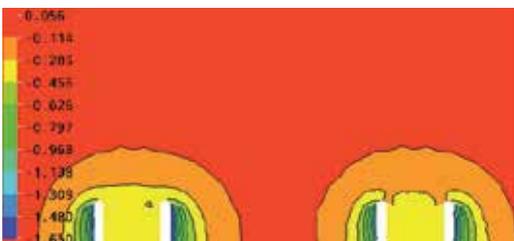
(b) Wind pressure coefficient of profile of center vertical profile when $S = 0.15D$ (c) Wind pressure coefficient of center vertical profile when $S = 0.75D$ (d) Wind pressure coefficient of center vertical profile when $S = 0.75D$ (e) Wind pressure coefficient of center profile when $S = 2D$ (f) Wind pressure coefficient of center vertical profile when $S = 3D$

Fig. 11. Wind pressure coefficient of 90° wind direction of central vertical profile

and pick eaves is just alike. With adobe spacing 0.75 D increases, spacing, adobe air pressure coefficient between areas surrounded by abate, and absolute phenomenon of wind pressure has reduce and decrease. With increased, when spacing distance, two Earth-buildings center 2.0 D air pressure changes around the isocline section together with monomer Earth-building working outline similar.

3.3 The change rule of average wind pressure coefficient disturbances

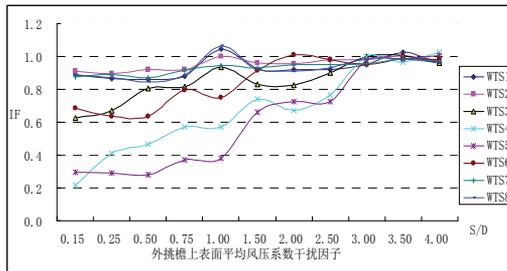
This paper through interference factor to quantitative description of interference effect adobe residences groups:

$$IF = \frac{C_{pI}}{C_{pA}} \tag{2}$$

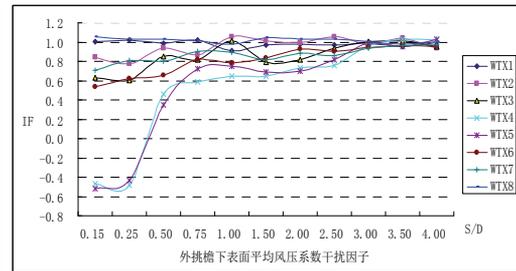
C_{pI} And C_{pA} are separately average wind pressure coefficients after and without wind interference.

3.3.1 0°wind direction

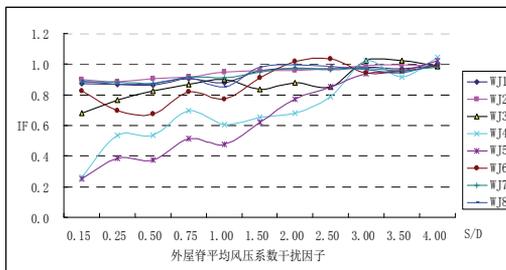
Figure 12 shows that under wind direction 0°, the average wind pressure coefficient interference factors of each zone of the interfered Earth-building roof is changed with the change of distance. In the Figure 12, abscissa S denotes for distance, D denotes for diameter.



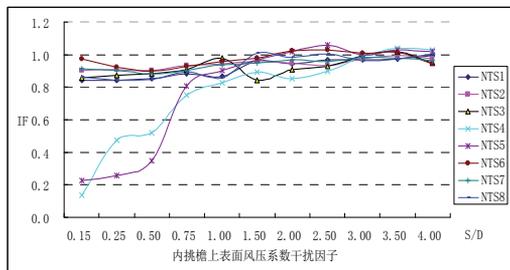
(a) Interference factors of average wind pressure coefficients of on upper surface outside carry eaves



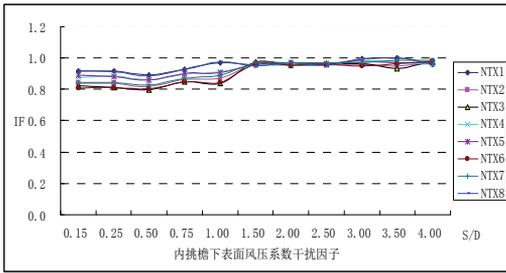
(b) Interference factors of average wind pressure coefficients on under surface of outside carry eaves



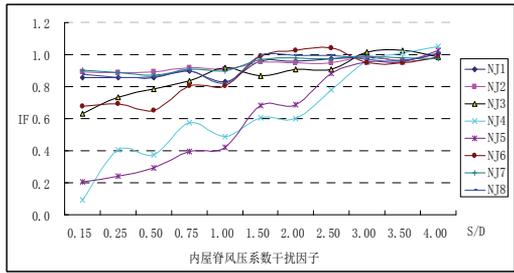
(c) Interference factors of average wind pressure coefficients of on external roof ridge



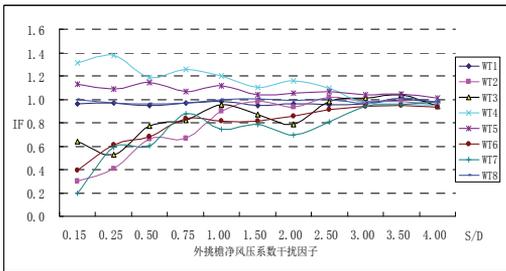
(d) Interference factors of average wind pressure coefficients on upper surface of inside carry eaves



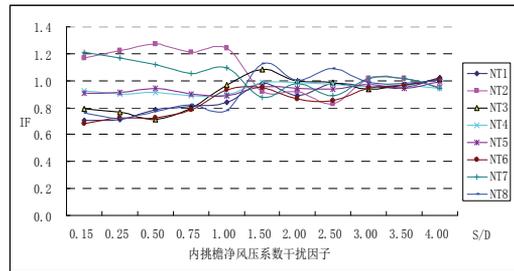
(e) Interference factors of average wind pressure coefficients on under surface of inside carry eaves



(f) Interference factors of average wind pressure coefficients on inner roof ridge



(g) Interference factors of net wind pressure coefficient on outside carry eaves



(h) Interference factors of net wind pressure coefficient on inside carry eaves

Fig. 12. The interference factors of wind pressure coefficient on each roof partition of 0° wind direction

What can be obtained by figure 12:

1. Upper surface of outside carry eaves: to sum up , the wind pressure coefficient in upper surface of outside carry eaves is reduced compared with single Earth-building, along with the increase of distance between, this trend weakened gradually, and when the distance reached 3D, interference factor approached to 1.0, interference affect basically can be ignored. Interference factor WTS1 and WTS8 have minimum amplitude, about 10%, the two surfaces are the farthest from the interfered Earth-building. Interference factor WTS4 and WTS5 have maximum amplitude, it reached 70 %, it has great influence with scrambling Earth-building, to the benefit of wind-resistant. Interference factor WTS2 and WTS6 decrease amplitude is about 30%, and interference factor WTS2 and WTS7 change amplitude is less ,is basically similar to WTS1 and WTS8.
2. Under surface of outside carry eaves: Roofing partition is symmetrical, the wind pressure coefficient interference factors under wind direction 0°change have obvious symmetry, and wind pressure coefficient are obviously reduce compared with single Earth-building. Interference factor WTS1 and WTS8 change amplitude is less changed with distance increases, in the 1.0 external floating up and down 5%, surface wind pressure coefficient interference factor WTS2 and WTS7, WTS3 and WTS6 increase with instance increase, among them , interference factor WTS2 and WTS7 are reduce mostly 20%, interference factor WTS3 and WTS6 reduce 40%. It's worth noting that WTX4 and

WTX5 at the leeward side, wind pressure coefficient changing, reducing up to 150%, when the distance increases, the pressure coefficient become negative from positive, interference factor become positive from negative, and when the spacing 3.0D, it is close to 1.0.

3. External roof ridge: The change regulation of the wind pressure coefficient interference factors in external roof ridge is basic same as the wind pressure coefficient interference factors in upper surface of inside carry eaves. The range ability of interference factor WJ1 and WJ8 is minimum, which is about 10%. The range ability of interference factor WJ4 and WJ5 is maximum, which is up to 70%, WJ3 and WJ6 interference factors reduced margin around 30%, WJ2 and WJ7 are same as WJ1 and WJ8 changes.
4. Upper surface of inside carry eaves: From figure 12 (d), in addition to see surface wind pressure coefficient interference factor NTS4 and NTS5 have obvious change, other various surface marked change are small coefficient of wind pressure reduction. Judging from the numerical simulation results, surface wind pressure coefficient NTS4 and NTS5 in smaller values, on change, although magnitude of 0.01 small changes, but the wind pressure coefficient embodied in the disturbances have changed greatly. Overall, wind pressure coefficient interference factor upper surface of inside carry eaves does not change significantly disturbances, the maximum 16%.
5. Under surface of inside carry eaves: from figure 12 (e), the each zoning of the eaves changes consistently, the wind pressure coefficient decreases, it is favorable to stand up the wind. The amplitude of NTX3 and NTX6 is relatively larger, 20% , it is the premises where backflow happens inside the Earthen ring. Although it is obstructed by the windward, the wind pressure has decreased; however, airflow returns violently, the wind pressure coefficient of the premises changes more than other premises. When the distance of the Earth-building is 1.5D, the influential factors have approached to 1.0, we can neglect the influences.
6. Inner roof ridge: The wind pressure coefficient disturbances of inner roof ridge and inside carry eaves have the similar variation tendency, but the variation amplitude of inner roof ridge is larger. The variation amplitude of NJ4 and NJ5 are still the biggest, they reach the 80%, the disturbances of NJ3 and NJ6 decrease about 30%, other each surface has the small variation amplitude of about 10%.
7. Bare wind pressure coefficient interference factors in outside carry eaves: For characteristic of Earth-building suction's 2.5 meter large carry eaves, we consider the up and down surface wind pressure coefficient of carry eaves respectively, then compose them, so we can obtain the bare wind pressure coefficient value that it will be used in design. Increase the interval, the disturbances in previous analysis will increase from the value that less than 1 to 1.0, in considering the bare wind pressure coefficient disturbances, the surfaces disturbances of WT4 and WT5 are both more than 1.0, the reason is that surface wind pressure absorb up and press down, it is disadvantage of structure to withstand wind. With the change of interval, the surface disturbances of WT1 and Wt8 are almost no impact, the reason is that up and down surface offset each other. The surface wind pressure of WT2 and WT7 decrease more than 70%, it is advantageous to withstand wind. The surface wind pressure of WT3 and WT6 also decrease to a certain degree, it is about 40%~50%.
8. Bare wind pressure coefficient interference factors in inside carry eaves: The surface wind pressure coefficient disturbances of NT2 and NT7 increase to a little range, it is about

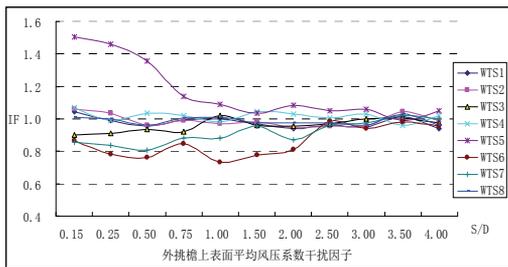
20%. Because of airflow in the Earth-building is backflow, the wind pressure coefficient disturbances of NT1、NT8、NT3、NT6 decrease to a little range, it is about 30%.

Overall, under 0° wind direction, the disturbed is in the upstream, then the disturbing effect is not so obvious, in addition to the individual; each partition of roof is advantageous to withstand wind. If downstream Earth-building have closer distance, the outer cornice of WT4 and WT5 are disadvantage to withstand wind, but value of its surface wind pressure coefficient is less, so it has little influence on withstand wind design. When interval of Earth-building reaches 3D, disturbances will approach to 1.0, the interference effect almost can be ignored.

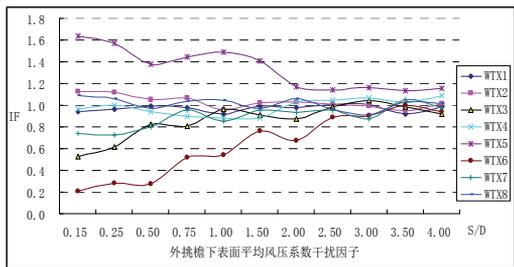
The disturbances of outer roof ridge and outer carry eaves have the similar change rules, the disturbances variation amplitude of inner roof are all smaller than outer roof, and outer carry eaves' absolute value of wind pressure coefficient is bigger than inner carry eaves, in real life situation, destroy the roof mainly begins from lifting tile of outer cornice roof, outer carry eaves is in a very disadvantage condition, so we mainly consider the outer cornice's change rule of wind pressure coefficient disturbances below.

3.3.2 45° wind direction

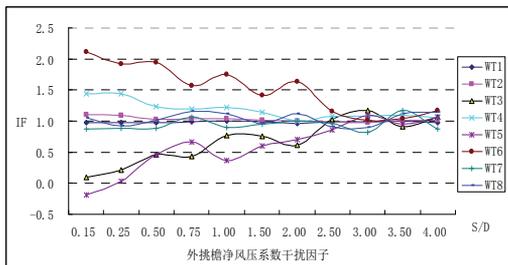
Figure 13 is a 45° wind direction under the pressure of the partition coefficient of confounding factors change with the pitch curve.



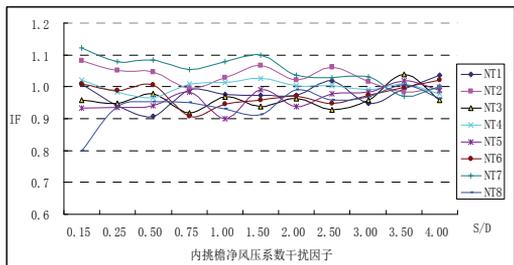
(a) Interference factors of average wind pressure coefficients on the upper surface of outside overhangs.



(b) Interference factors of average wind pressure coefficients on the lower surface of outside overhangs.



(c) Interference factors of net wind pressure coefficients of outside overhangs.



(d) Interference factors of net wind pressure coefficients of inside overhangs.

Fig. 13. The fact interference ors of average wind pressure coefficient on each roof partition of 45° wind direction

What can be drawn from figure 13:

1. External overhangs on the surface: WTS5 in the leeward surface, and adjacent interference Earth-building, where up to 1.5. There are increasing rapidly wind pressure and very adverse wind resistance. WTS6 interference factor is less than 1.0. WTS5 in a relatively inclined all along flow position, and wind pressure coefficient disturbances is asymmetry, because of the other Earth-building influence. There are WTS1, WTS2 WTS4 and WTS8 about 1.0, which interference is not obvious. WTS3, WTS6 and WTS7 interference factor has a small decrease. When the spacing reach to 2.5 D, interference factor affect will be ignore.
2. The lower surface of outer overhangs: WTX5 maximum interference factor of 1.6, WTX3 and WTX7 interference factor of Leeward surface decreases above 30%, Range of other surface pressure coefficients interference factors varies by less, remain in the vicinity of 1.0.
3. The interference factor of net air pressure coefficient of outside carry eaves: it reaches 2.0 or above on WT6, interference effect is serious, but the wind pressure coefficient of this surface is numerical small, 0.1 orders of magnitude, the wind resistant design does not control the surface. The interference factor of WT4 is 1.4, which is adverse for wind resistance. The interference factor of WT5 is negative in small spacing, and for WT5 is in the leeward surface where air pressure coefficient is small, the interference effect is beneficial although it is serious. The minimum interference factor of WT3 is 0.1, and it is getting bigger with increases of spacing (the interference factor turns to 1.0 when the spacing is 2.5D). Other interference factors are floating near 1.0; the interference phenomenon is not obvious.
4. The interference factor of net wind pressure coefficient of inside pick eaves: the variation interference factors of each roof partition are small, interference phenomenon is not obvious.

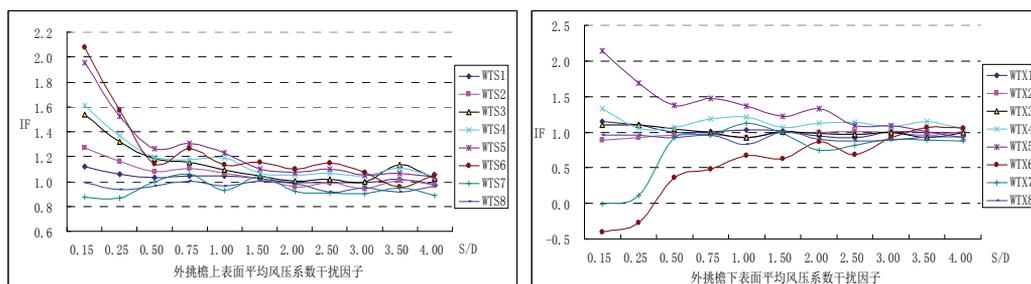
Overall, under the 45 °wind direction, except the interference factors of some roof partitions are near 1.0, the variation of interference factors of WT4, WT5 and WT6 are large, which should be taken into consideration of design.

3.3.3 90° wind direction

Figure 14 is a 90° wind direction under the pressure of the partition coefficient of confounding factors is changed with the pitch curve.

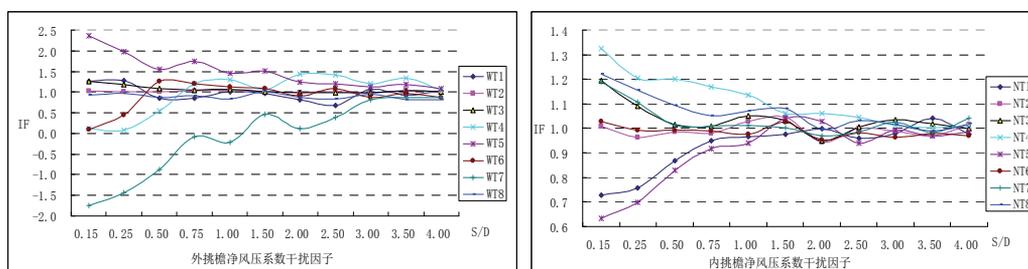
What can be drawn from figure 14:

1. External overhangs on the surface: WTS6 interference factor reached a maximum of 2.1, WTS5 interference factor is close to 2.0, WTS4 interference factor 1.6, WTS3 interference factor is also 1.53, with control effect in the gorge, the wind pressure coefficient increases more, is not conducive to Wind. Range of other surface confounding factors varies by less, WTS2 surface disturbance factor of 1.3, WTS7 interference factor of 0.9, WTS1 and WTS8 fluctuations are around 1.0. When the distance increased to 1.5D, external overhangs are reaching the district interference factor of 1.0, as the spacing increases, interference effects gradually weakened.
2. The lower surface of outer overhangs: WTX5 maximum interference factor of 2.1, when wind pressure coefficient of -1.23, air flow around the earth-buildings in the region seriously affected with each other. WTX4 interference factor is also higher, at 1.4. WTX6 and WTX7 interference factor decreases above 100%, large amplitude, but at the leeward wind pressure coefficient values smaller. Range of other surface pressure coefficients interference factors varies by less, remain in the vicinity of 1.0.



(a) Interference factors of average wind pressure coefficients on the upper surface of outside overhangs

(b) Interference factors of average wind pressure coefficients on the lower surface of outside overhangs



(c) Interference factors of net wind pressure coefficients of outside overhangs.

(d) Interference factors of net wind pressure coefficients of inside overhangs.

Fig. 14. The interference factors of average wind pressure coefficient on each roof partition of 90° wind direction

- Net wind pressure coefficients interference factor outside the overhangs: WT5 surface disturbance factors up to 2.5, the pressure coefficient of 0.53, interference is obvious. WT7 interference factor of -1.7, but the surface is in the leeward wake region, pressure coefficient value is small. WT4 and WT6 interference factor decreases by about 80%, favorable wind. The surface of the other confounding factors the district did not change significantly, at 1.0 fluctuate. When the spacing of 1.5D in the Earth-building. District confounding factors close to 1.0, interference effects can be ignored.
- Net wind pressure coefficients interference factor in the overhangs: the overhangs at the wind pressure coefficient absolute value of the partition is generally small, between 0.2 and 0.4. NT4 interference factor 1.32, NT5 minimum interference factor of 0.64, Earthen Ring back airflow significantly. NT8 interference factor of 1.22, NT7 interference factor of 1.18, a result of disturbed earth-building wake interference earthen interference facilities was greatly changed.

Overall, in 90° wind direction, the interference is obvious, considering the surface pressure coefficients, most of the partition surface disturbance factor greater than 1.0, but considering the superimposed effect of the upper and lower surfaces, the interference is not so prominent, but WT5, NT4 disadvantaged status of wind, earth-building roof to the attention of conservation measures.

4. Conclusion

Through the exhaustive numerical simulation about two typical circular earth-buildings in different wind directions and different spacing, we obtained the distribution characteristics of wind pressure both on the single earth-building and the groups under 0° , 45° , 90° wind direction, the interference effects of wind flow and the change rules that interference factor of average wind pressure coefficient changing with spacing on each roof partition. From the research above we can draw the conclusion that:

1. Numerical simulation can accurately simulate the air flowing field in different wind directions and wind pressure distribution on each roof partition of low buildings. Through the study of wind pressure coefficient and wind velocity contour on two feature faces, we can make qualitative analyses about the wind characteristics in buildings. Under the same wind direction, air flowing field around the earth-building in groups gets close to that of single building along with the increase of building spacing. Under different wind directions, the air flowing field is not the same when building spacing is different.
2. Because of the characteristics of earth-buildings: thick walls, long cornices and tile roof, we mainly analyses the variation of interference factors on each roof partition. The actual situation is that damage of tile roof starts with the lifting of roof tiles from the outside carry eaves. After analysis of the variation of interference factors under different wind directions, we found the change laws of interference factors on the ridge and outside carry eaves are consistent and there is less change of interference factors on the inside roof than the outside. This paper focuses on analyzing the variation of interference factors of wind pressure coefficient on the outside carry eaves under various conditions.
3. Under different wind directions, the variation range of interference factors of average wind pressure coefficient on both the upper and lower surface are smaller than that of net wind pressure coefficient of the outside carry eaves. Interference factors on the upper surface and that on the lower surface mutually reinforce the effects on the windward side and mutually reduce the effects on the leeward side.
4. Compared with a single earth-building, when two earth-buildings are in a line the variation of interference factors is not obvious. When the spacing between two earth-buildings reach 3 times of the bigger radius, the interference factor is close to 1.0 and the interference affect can be ignored basically.
5. when the wind direction is 45° with the line of two earth-buildings, the wind interference is very different to the situation when the wind direction is 0° . When the spacing between two earth-buildings reach 2.5 times of the bigger radius, the interference factor is close to 1.0 and the interference affect can be ignored in the wind resistant design. Under the 45° wind direction, maximum interference factor reach 2.1 on part WT6 which is unfavorable disturbance, but interference factors decrease on part WT5 which is positive.
6. Under the 90° wind direction, most roof partitions are severely disturbed. It shows obvious effect of narrow and wind pressure coefficient increases greatly. The interference factors on part WT5 reach up to 2.5 which are extremely unfavorable to wind resistant and should be paid attention to. When the spacing between two earth-buildings reach 1.5 times of the bigger radius, the interference effect is too weak to be considered in the wind resistant design.

5. References

- [1] Walker.G; Roy.R. Wind loads on houses in an urban environment[R]. University of Roorkee · India: Asia Pacific Symposium on wind engineering, 1985.
- [2] Case.P.C; Isyumov.N. Wind loads on low buildings with 4:12 gable roofs in open country and suburban exposures [J]. Journal of Wind Engineering and Industrial Aerodynamics. 1998(77-78): 107-118.
- [3] Blessmann.J. Buffeting effects of neighboring tall buildings [J]. Journal of Journal of wind engineering and industrial aerodynamics. 1985, 18(1): 100-105.
- [4] Taniike,Yoshihito. Turbulence effect on mutual interference of buildings [J]. Journal of Engineering Mechanics. 1991, 117(3): 443-456.
- [5] Zhang Xiangting. Engineering wind resistance design and calculation manual [M]. China architecture &building press, 1998.(in Chinese)
- [6] Xie Zhuangning. Research of Interference Effects of Wind Loads of a Cluster of Tall Buildings[D]. shanghai, Tongji university , 2003. (in Chinese)
- [7] Holmes.J.D. Wind pressures on tropical housing[J]. Journal of Wind Engineering and Industrial Aerodynamics. 1994,53(1-2): 105-123.
- [8] Zhao Qingchun; Peng Xingqian; Zhou Xianpeng; Qiao Changgui. Numerical simulation analysis of wind interference effects on the roof of low-rise gable-roofed buildings [J]. Journal of Fuzhou university.2008, 36(6): 863-867. (in Chinese)
- [9] Tsutsumi.J; Katayama.T; Nishida.M. Wind tunnel tests of wind pressure on regularly aligned buildings [J]. Journal of Wind Engineering and Industrial Aerodynamics. 1992, 43(3): 1799-1810.
- [10] Huang Hanmin. Fujian Earth-building [M]. Beijing: Sanglian Bookstore, 2003. (in Chinese)
- [11] Shao Kun; Peng Xingqian; Liu Chunyan; Xu Gang. Computational Domain Setting About Numerical Wind Tunnel Simulation of Earth-building [J].Journal of ZhengZhou institute of light industry. 2010, 25(4):55-58. (in Chinese)

Mass–Consistent Wind Field Models: Numerical Techniques by L2–Projection Methods

L. Héctor Juárez¹, María Luisa Sandoval¹, Jorge López²
and Rafael Reséndiz¹

¹*Departamento de Matemáticas, Universidad Autónoma Metropolitana Iztapalapa,
México City*

²*División de Ciencias Básicas, Universidad Juárez Autónoma de Tabasco
México*

1. Introduction

For several meteorological problems and a large number of applications, the knowledge of the 3–D wind field over a region is required. Examples include prediction of the transport, diffusion and dispersion of air pollutants in the atmosphere (Finardi et al., 2010; Sherman, 1978), realization of wind maps for the design of different urban and general projects (Castino et al., 2003), and the effect of wind on structures and fire spreading (Potter & Butler, 2009), among others. Moreover, meteorological wind fields are also required inputs for air quality models. In practice, usually limited horizontal wind field measurements are available, and therefore the calculation of the vertical motion must be predicted or calculated. Several methods and strategies, with various levels of complexity, have been proposed to address this problem. They can be included into two general model types: prognostic models and diagnostic models. Prognostic models are complex time–dependent hydrodynamic models governing air flow, including thermal effects, density variation and turbulent interaction. While these models are “realistic”, they are expensive to operate, need extensive computer facilities, and require specialized training for their operation. On the other hand, diagnostic wind models do not require the integration of the non–linear hydrodynamic equations. Instead, available interpolated data is used to generate wind fields, which satisfy some physical or dynamical constraints. For instance, to assure mass conservation, a simplified steady–state version of the continuity equation is imposed, and the resulting model is then called a mass–consistent model. A review of these models is available in Ratto et al. (1994) and Ratto (1996).

We focus in a variational mass–consistent model which is based in the original formulation by Sasaki (Sasaki, 1958). This approach has been used for a variety of meteorological problems (Castino et al., 2003; Pennel, 1983; Sherman, 1978; Wang et al., 2005). Mass–consistent models are attractive because of their simplicity, and because they are easy and economical to operate. In some applications, these models outperform the more sophisticated and expensive dynamical models (Ratto et al., 1994). However, mass–consistent models have some disadvantages, because they are based on incomplete or idealized models and have difficulty

representing flows accurately in data-sparse regions as mountains or oceans. Despite these limitations, mass-consistent models are a valuable tool for air quality applications and consequently several developments have taken place over last decades (Ferragut et al., 2010; Ratto, 1996; Ratto et al., 1994; Ross et al., 1988; Wang et al., 2005). Most of the results presented in this chapter has been published in the last few years (Flores et al., 2010; Núñez et al., 2007; 2006), but we also include some additional ideas and recent results.

The variational method proposed by Sasaki uses the continuity equation $\nabla \cdot \mathbf{u} = 0$, where \mathbf{u} is the wind velocity vector field on a given domain Ω . The method is based on the minimization of the functional L defined by

$$L(\mathbf{u}, \lambda) = \frac{1}{2} \int_{\Omega} \left\{ S(\mathbf{u} - \mathbf{u}^I) \cdot (\mathbf{u} - \mathbf{u}^I) + \lambda [\nabla \cdot \mathbf{u}] \right\} dV, \quad (1)$$

where \mathbf{u}^I is an initial observed wind field, λ is a Lagrange multiplier and S is a diagonal matrix with weighting parameters $\alpha_i > 0$, $i = 1, 2, 3$, called Gaussian precision moduli, related to the scales of the respective components of the velocity field. The vertical component of the initial wind field is taken as zero because meteorological stations usually do not measure this component. The Euler-Lagrange equations of (1) are:

$$\mathbf{u} = \mathbf{u}^I + S^{-1} \nabla \lambda, \quad (2)$$

Usually \mathbf{u} is obtained from (2), after λ is computed. Since $\nabla \cdot \mathbf{u} = 0$, then from (2) we obtain the elliptic equation $-\nabla \cdot (S^{-1} \nabla \lambda) = \nabla \cdot \mathbf{u}^I$, from which λ is obtained. To complement (close) this equation, two types of boundary conditions are commonly used: homogeneous Dirichlet boundary conditions, $\lambda = 0$, for open or “flow through” boundaries (like truncated boundaries), and Neumann boundary conditions, $\partial \lambda / \partial \mathbf{n} = 0$, for closed or “no flow through” boundaries (like the surface terrain or topography). Many authors have been used and recommend these boundary conditions (Kitada et al., 1986; 1983; Ratto et al., 1994; Sherman, 1978). However they are physically and mathematically inconsistent as we will show in this work. Even though, there have been several sophisticated developments in the numerical simulations of this model as, for instance, the application of multigrid methods (Wang et al., 2005), and the application of genetic algorithms to estimate parameters (Montero et al., 2005), it seems that the analysis of boundary conditions has not attracted the attention of the community in meteorology.

In this work we study how boundary conditions affect solutions of the elliptic equation for λ . We show that the application of incorrect boundary conditions may degrade the solutions several orders of magnitude, and we propose some strategies to overcome this problem. In particular, we introduce a new approach based on the saddle-point formulation of the constrained least squares formulation of the problem, which allows the introduction of successful techniques from computational fluid dynamics. This new approach does not require boundary conditions for the multiplier. It produces much better results, and it also helps us to establish more consistent boundary conditions on truncated nonphysical boundaries. We also explore other boundary conditions for the multiplier better suited for artificial truncated boundaries. Furthermore, we present some preliminary numerical results using a meshfree method based on a radial basis function collocation method.

2. Mathematical formulation of the problem

Let Ω be an open, simply connected and bounded region in \mathbb{R}^d ($d = 2$ or 3) with Lipschitz boundary $\partial\Omega = \Gamma_N \cup \Gamma_D$, where Γ_N is the part of the boundary associated to the surface terrain (topography), and Γ_D is the rest of the boundary (artificial vertical boundaries and top boundary), as shown in Figure 1. Given an initial vector field \mathbf{u}^I in Ω (which can be obtained

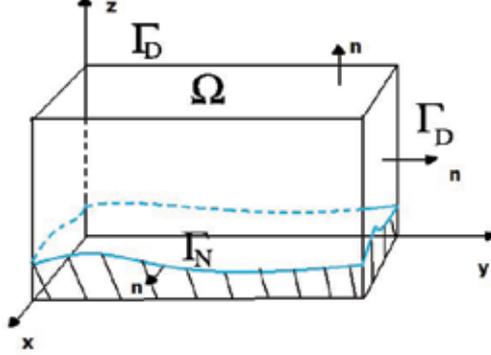


Fig. 1. Bounded region Ω .

by interpolating atmospheric data, or by other means), our goal is to generate a solenoidal field \mathbf{u} –called adjusted field– as close to \mathbf{u}^I as possible in a sense that will be clarified below, such that $\mathbf{u} \cdot \mathbf{n} = 0$ on Γ_N .

We define the following vector function spaces: $\mathbf{L}_2(\Omega) = L_2(\Omega)^d$ and $\mathbf{H}(\text{div}; \Omega) = \{ \mathbf{v} \in \mathbf{L}_2(\Omega) : \nabla \cdot \mathbf{v} \in L_2(\Omega) \}$. Then, the adjusted wind field \mathbf{u} must belong to the normed closed space

$$\mathbf{V} = \{ \mathbf{v} \in \mathbf{H}(\text{div}; \Omega) : \nabla \cdot \mathbf{v} = 0 \text{ and } \mathbf{v} \cdot \mathbf{n} = 0 \text{ on } \Gamma_N \}, \quad (3)$$

with the norm $\|\cdot\|_{S, \Omega}$ associated to the inner product $\langle \mathbf{u}, \mathbf{v} \rangle_S = \int_{\Omega} S \mathbf{u} \cdot \mathbf{v} \, d\mathbf{x}$, where $\mathbf{v} \cdot \mathbf{w} = \sum_1^d v_i w_i$ is the usual scalar product in \mathbb{R}^d . We can now formulate the problem as a least squares projection problem. For this purpose, we define a convex quadratic functional $J : \mathbf{V} \rightarrow \mathbb{R}$ as

$$J(\mathbf{v}) = \frac{1}{2} \|\mathbf{v} - \mathbf{u}^I\|_{S, \Omega}^2 = \frac{1}{2} \int_{\Omega} S(\mathbf{v} - \mathbf{u}^I) \cdot (\mathbf{v} - \mathbf{u}^I) \, d\mathbf{x}. \quad (4)$$

Then, our problem can be stated as follows:

$$\text{Given } \mathbf{u}^I \in \mathbf{L}_2(\Omega), \text{ find } \mathbf{u} \in \mathbf{V} \text{ such that } J(\mathbf{u}) \leq J(\mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}. \quad (5)$$

Due to the properties of this functional, $\mathbf{u} \in \mathbf{V}$ is a minimizer of J if and only if it is a stationary point of J :

$$\frac{\partial}{\partial \epsilon} J(\mathbf{u} + \epsilon \mathbf{v})|_{\epsilon=0} = \int_{\Omega} S(\mathbf{u} - \mathbf{u}^I) \cdot \mathbf{v} \, d\mathbf{x} = 0, \quad \forall \mathbf{v} \in \mathbf{V}. \quad (6)$$

The Lax–Milgram theorem guaranties that this equation has a unique solution.

3. The traditional approach. Advantages and difficulties

3.1 Derivation of the elliptic problem

The first approach is based on a Helmholtz–type decomposition of the Hilbert vector space $\mathbf{L}_2(\Omega)$, and it reduces to the traditional approach used by meteorologists.

Proposition 1 *The orthogonal complement in $\mathbf{L}_2(\Omega)$ of the closed subspace \mathbf{V} is*

$$\mathbf{V}^\perp = \{ \nabla q : q \in H^1(\Omega), \quad q = 0 \quad \text{on} \quad \Gamma_D \}.$$

An argument very similar to that given by Girault and Raviart (Girault & Raviart, 1986), shows that this decomposition is valid (details are given in (Núñez et al., 2007)). Therefore we get from (6) that $S(\mathbf{u} - \mathbf{u}^I) = \nabla \lambda$, with λ in

$$H_D^1(\Omega) \equiv \{ q \in H^1(\Omega) : q = 0 \text{ on } \Gamma_D \}. \quad (7)$$

With the above properties, we obtain a saddle–point problem for \mathbf{u} and λ (left), as well as the correspondent elliptic problem for λ (right):

$$S\mathbf{u} - \nabla \lambda = S\mathbf{u}^I, \quad \text{and} \quad \nabla \cdot \mathbf{u} = 0 \quad \text{in} \quad \Omega, \quad -\nabla \cdot (S^{-1}\nabla \lambda) = \nabla \cdot \mathbf{u}^I \quad \text{in} \quad \Omega, \quad (8)$$

$$\lambda = 0 \quad \text{on} \quad \Gamma_D, \quad \lambda = 0 \quad \text{on} \quad \Gamma_D, \quad (9)$$

$$\mathbf{u} \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_N, \quad -S^{-1}\nabla \lambda \cdot \mathbf{n} = \mathbf{u}^I \cdot \mathbf{n} \quad \text{on} \quad \Gamma_N. \quad (10)$$

To obtain the elliptic problem, we eliminate \mathbf{u} from the saddle–point problem using that \mathbf{u} belongs to \mathbf{V} . Once λ is calculated from (8)–(10), the adjusted field is recovered from (2).

Equation (8) has traditionally been used by meteorologists. However, this equation is generally introduced from a discussion in which it is not clear how to establish the proper boundary conditions for λ . The crucial argument in our study is the decomposition of $\mathbf{L}_2(\Omega)$ in orthogonal subspaces \mathbf{V} and \mathbf{V}^\perp , from which the boundary conditions for λ arises in a natural way, from the mathematical point of view. We would like to mention that the boundary condition (10) has already been used in recent research (Ferragut et al., 2010).

3.2 Finite element solution of the elliptic problem

The variational formulation of the elliptic problem (8)–(10) is

$$\int_{\Omega} S^{-1}\nabla \lambda \cdot \nabla q \, d\mathbf{x} = - \int_{\Omega} \mathbf{u}^I \cdot \nabla q \, d\mathbf{x}, \quad \forall q \in H_D^1(\Omega). \quad (11)$$

Here, we consider the two–dimensional case. Let \mathcal{T}_h be a finite element triangulation of $\bar{\Omega} \subset \mathbb{R}^2$ (Ciarlet, 2002), where h is taken as the space discretization step. Let's denote by P_1 the space of polynomials of degree less or equal than 1. Then, $\mathbf{L}_2(\Omega)$ and $H_D^1(\Omega)$ are approximated by the finite dimensional spaces

$$\mathbf{L}_h = \left\{ \mathbf{v}_h \in \mathcal{C}^0(\bar{\Omega})^2 : \mathbf{v}_h|_T \in P_1 \times P_1, \forall T \in \mathcal{T}_h \right\}, \quad (12)$$

$$H_h = \left\{ q \in \mathcal{C}^0(\bar{\Omega}) : q|_T \in P_1, \forall T \in \mathcal{T}_h, q = 0 \text{ on } \Gamma_D \right\}, \quad (13)$$

respectively. Thus, the finite element algorithm is: Given $\mathbf{u}_h^I \in \mathbf{L}_h$, find $\lambda_h \in H_h$ such that

$$\int_{\Omega} S^{-1} \nabla \lambda_h \cdot \nabla q \, d\mathbf{x} = - \int_{\Omega} \mathbf{u}_h^I \cdot \nabla q \, d\mathbf{x}, \quad \forall q \in H_h, \quad (14)$$

where $\mathbf{u}_h^I \in \mathbf{L}_h$ is the interpolant of the given initial velocity field \mathbf{u}^I . We obtain λ_h after solving the resulting system of linear equations, and the numerical approximation \mathbf{u}_h of \mathbf{u} is computed by the weak version of (2) as follows: Find $\mathbf{u}_h \in \mathbf{L}_h$ with $\mathbf{u}_h \cdot \mathbf{n} = 0$ on Γ_N such that

$$\int_{\Omega} (S\mathbf{u}_h) \cdot \mathbf{v} \, d\mathbf{x} = \int_{\Omega} (S\mathbf{u}_h^I) \cdot \mathbf{v} \, d\mathbf{x} - \int_{\Omega} \lambda_h \nabla \cdot \mathbf{v} \, d\mathbf{x}, \quad \forall \mathbf{v} \in \mathbf{L}_h, \quad \mathbf{v} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_N. \quad (15)$$

From now on, we identify the algorithm (14)–(15) as the *E1-algorithm*.

Example 1. We consider the solenoidal vector field $\mathbf{u}(x, z) = (x, -z)$ defined in $\Omega = (1, 2) \times (0, 1)$, so that $\mathbf{u} \in \mathbf{V}$. Assuming that we have $\mathbf{u}^I(x, z) = (x, 0)$ as an initial horizontal wind field, we want to apply the *E1-algorithm* to see how much we can recover of the vertical component of \mathbf{u} . For this numerical calculation, Ω is divided into a 80×80 triangular mesh, and we choose the following values for the Gaussian Precision moduli: $\alpha_1 = 1$ and $\alpha_3 = 0.001$. Figure 2 shows the exact field in red and the computed adjusted field in blue. Both fields agree fairly well almost everywhere, except on the vertical artificial boundaries $x = 1$ and $x = 2$.

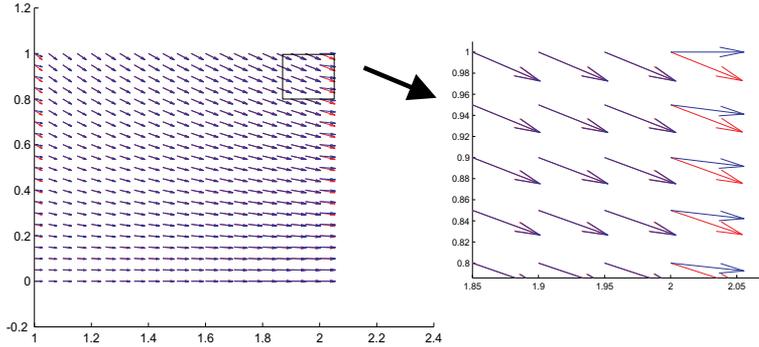


Fig. 2. Exact field $\mathbf{u} = (x, -z)$ in red, adjusted field obtained by the *E1-algorithm* in blue.

The relative error and the mean divergence of the computed solution are defined as

$$e_r = \frac{\|\mathbf{u} - \mathbf{u}_h\|_2}{\|\mathbf{u}\|_2}, \quad \text{and} \quad mdiv = \text{mean}_{\mathbf{x}_i} \{ \nabla \cdot \mathbf{u}_h(\mathbf{x}_i) \mid \mathbf{x}_i \text{ is a interior vertex} \}, \quad (16)$$

respectively. The point-wise divergence is computed in a weak sense, as follows

$$\nabla \cdot \mathbf{u}_h(\mathbf{x}_i) = - \int_{\Omega} \mathbf{u}_h \cdot \nabla \phi_i \, d\mathbf{x}, \quad (17)$$

where ϕ_i is the piece-wise linear base function associated to vertex node \mathbf{x}_i . For the present example, we obtain $e_r = 1.9 \times 10^{-2}$ and $mdiv = 4.1 \times 10^{-2}$.

The values for the Gaussian precision moduli were chosen based on numerical performance. Table 1 shows the behavior of e_r and $mdiv$ for different values of α_3 when α_1 is kept constant

and equal to one. Clearly the best results were obtained with $\alpha_3 = 0.001$. We will explain this behavior later on. But, for the moment, we want to emphasize that this algorithm produces satisfactory results almost everywhere, except on the boundary Γ_D , where homogeneous Dirichlet boundary conditions were imposed.

α_3	0.001	0.01	0.1	1	100	1000
e_r	1.9×10^{-2}	9.6×10^{-2}	1.4×10^{-1}	5.2×10^{-1}	6.4×10^{-1}	9.8×10^{-1}
$mdiv$	4.1×10^{-2}	-6.1×10^{-2}	2.9×10^{-1}	5.4×10^{-1}	7.8×10^{-1}	9.8×10^{-1}

Table 1. Numerical performance of *E1-algorithm* for different values of α_3 .

We can say that the main advantage of this traditional way to solve the problem is its simplicity, since it only involves the solutions of an elliptic partial differential equation (PDE). On the other hand, one of its major drawbacks is that inconsistent or incorrect boundary conditions, on truncated artificial boundaries, degrade the accuracy of the solution. In the rest of the chapter, we introduce some alternatives to overcome these problems.

4. A saddle-point formulation and the conjugate gradient algorithm

4.1 Derivation of the formulation

The second approach to solve the problem (5), or equivalently problem (6), is based on the usual methodology to solve constrained optimization problems. That is, we introduce the space of vector functions

$$\mathbf{V}_N = \{ \mathbf{v} \in \mathbf{H}(div; \Omega) : \mathbf{v} \cdot \mathbf{n} = 0 \text{ on } \Gamma_N \}, \quad (18)$$

together with the Lagrangian L defined on $\mathbf{V}_N \times L_2(\Omega)$ as

$$L(\mathbf{v}, q) \equiv J(\mathbf{v}) + \langle q, \nabla \cdot \mathbf{v} \rangle = \frac{1}{2} \int_{\Omega} S(\mathbf{v} - \mathbf{u}^I) \cdot (\mathbf{v} - \mathbf{u}^I) \, d\mathbf{x} + \int_{\Omega} q \nabla \cdot \mathbf{v} \, d\mathbf{x}.$$

A stationary point (\mathbf{u}, λ) of L solves the following saddle-point problem

$$\int_{\Omega} S\mathbf{u} \cdot \mathbf{v} \, d\mathbf{x} + \int_{\Omega} \lambda \nabla \cdot \mathbf{v} \, d\mathbf{x} = \int_{\Omega} S\mathbf{u}^I \cdot \mathbf{v} \, d\mathbf{x}, \quad \forall \mathbf{v} \in \mathbf{V}_N, \quad (19)$$

$$\int_{\Omega} q \nabla \cdot \mathbf{u} \, d\mathbf{x} = 0, \quad \forall q \in L_2(\Omega), \quad (20)$$

where λ **need not satisfy boundary conditions**. The solution \mathbf{u} is the minimizer of J , and now it is obtained from the enlarged space \mathbf{V}_N where free divergence is not required. Instead, the condition $\nabla \cdot \mathbf{u} = 0$ is relaxed by the introduction of the Lagrange multiplier λ so that \mathbf{u} must satisfy the weaker condition (20). To solve (19)–(20) we introduce a method which has shown to be very effective for solving Stokes problems in computational fluid dynamics (Glowinski, 2003). The idea is as follows: assuming that (\mathbf{u}, λ) is a solution of the problem (19)–(20), the vector field \mathbf{u} is decomposed as $\mathbf{u} = \mathbf{u}^I + \mathbf{u}_\lambda$, where \mathbf{u}^I is the given initial vector field, and $\mathbf{u}_\lambda \in \mathbf{V}_N$ solves

$$\int_{\Omega} S\mathbf{u}_\lambda \cdot \mathbf{v} \, d\mathbf{x} = - \int_{\Omega} \lambda \nabla \cdot \mathbf{v} \, d\mathbf{x}, \quad \forall \mathbf{v} \in \mathbf{V}_N. \quad (21)$$

Furthermore, \mathbf{u}_λ must satisfy (20) which has the following equivalent strong version

$$-\nabla \cdot \mathbf{u}_\lambda = \nabla \cdot \mathbf{u}^I. \quad (22)$$

A key point is that problem (21)–(22) can be formulated as a functional equation. For this we introduce the linear operator A from $L_2(\Omega)$ into $L_2(\Omega)$ defined by

$$Aq = -\nabla \cdot \mathbf{u}_q, \quad (23)$$

where $\mathbf{u}_q \in \mathbf{V}_N$ is the solution of

$$\int_{\Omega} S \mathbf{u}_q \cdot \mathbf{v} \, d\mathbf{x} = - \int_{\Omega} q \nabla \cdot \mathbf{v} \, d\mathbf{x}, \quad \forall \mathbf{v} \in \mathbf{V}_N. \quad (24)$$

With this definition, it is clear, from (21)–(22), that the multiplier λ satisfies the functional equation

$$A\lambda = \nabla \cdot \mathbf{u}^I. \quad (25)$$

4.2 Conjugate gradient algorithm

Operator A is selfadjoint, and strongly elliptic, since from (23) and (24) we have

$$\begin{aligned} \int_{\Omega} q' Aq \, d\mathbf{x} &= - \int_{\Omega} q' \nabla \cdot \mathbf{u}_q = \int_{\Omega} S \mathbf{u}_{q'} \cdot \mathbf{u}_q \, d\mathbf{x} \quad \forall q, q' \in L_2(\Omega), \\ \int_{\Omega} q Aq \, d\mathbf{x} &= \int_{\Omega} S \mathbf{u}_q \cdot \mathbf{u}_q > c \|\mathbf{u}_q\|_{L_2(\Omega)}^2 \quad \forall q \neq 0 \quad (0 < c < \min\{\alpha_i\}) \end{aligned}$$

Therefore, the following iterative conjugate gradient algorithm may be used to solve the infinite dimensional problem (25):

1. Given $\lambda^0 \in L_2(\Omega)$, solve for $\mathbf{u}^0 \in \mathbf{V}_N$

$$\int_{\Omega} S \mathbf{u}^0 \cdot \mathbf{v} \, d\mathbf{x} = \int_{\Omega} S \mathbf{u}^I \cdot \mathbf{v} \, d\mathbf{x} - \int_{\Omega} \lambda^0 \nabla \cdot \mathbf{v} \, d\mathbf{x}, \quad \forall \mathbf{v} \in \mathbf{V}_N.$$

$$\text{Set } g^0 = -\nabla \cdot \mathbf{u}^0 \quad \text{and} \quad d^0 = -g^0.$$

2. For $k \geq 0$, assuming we know $\lambda^k, g^k, d^k, \mathbf{u}^k$, find $\lambda^{k+1}, g^{k+1}, d^{k+1}, \mathbf{u}^{k+1}$, doing the following: Solve for $\underline{\mathbf{u}}^k \in \mathbf{V}_N$

$$\int_{\Omega} S \underline{\mathbf{u}}^k \cdot \mathbf{v} \, d\mathbf{x} = - \int_{\Omega} d^k \nabla \cdot \mathbf{v} \, d\mathbf{x}, \quad \forall \mathbf{v} \in \mathbf{V}_N.$$

$$\text{Set } w^k = -\nabla \cdot \underline{\mathbf{u}}^k \quad \text{and compute} \quad \alpha_k = \frac{\langle g^k, g^k \rangle}{\langle d^k, w^k \rangle}.$$

$$\text{Compute } \lambda^{k+1} = \lambda^k + \alpha_k d^k, \quad \mathbf{u}^{k+1} = \mathbf{u}^k + \alpha_k \underline{\mathbf{u}}^k, \quad g^{k+1} = g^k + \alpha_k w^k.$$

3. If $\langle g^k, g^k \rangle \leq \varepsilon \langle g^0, g^0 \rangle$, take $\lambda = \lambda^{k+1}$ and $\mathbf{u} = \mathbf{u}^{k+1}$ and stop. Otherwise, compute

$$d^{k+1} = -g^{k+1} + \beta_k d^k \quad \text{where} \quad \beta_k = \frac{\langle g^{k+1}, g^{k+1} \rangle}{\langle g^k, g^k \rangle}.$$

Do $k = k + 1$ and return to 2.

Above, $\langle \cdot, \cdot \rangle$ indicates the usual scalar product in $L_2(\Omega)$. Observe that the adjusted field \mathbf{u} is also computed as an intermediate step in the algorithm. In this algorithm, no boundary conditions are imposed on λ , contrary to what it was done in the first approach. This fact has a very important effect in the numerical calculation.

4.3 A mixed finite element method

To approximate the functions in \mathbf{V}_N and $L_2(\Omega)$, considered in the previous algorithm, we use the *Bercovier–Pironneau* finite element approximation (Bercovier & Pironneau, 1979). Functions in $L_2(\Omega)$ are approximated by continuous piecewise linear polynomials over a triangulation \mathcal{T}_h of Ω , while the elements in \mathbf{V}_N are also approximated by linear polynomials but now over a twice finer triangulation $\mathcal{T}_{h/2}$ of Ω . The fine triangulation $\mathcal{T}_{h/2}$ is obtained from a regular subdivision of each triangle $T \in \mathcal{T}_h$. Then, the functional spaces \mathbf{V}_N and $L_2(\Omega)$ will be approximated, respectively, by the finite dimensional spaces

$$\mathbf{V}_{Nh} = \left\{ \mathbf{v}_h \in C^0(\bar{\Omega})^2 : \mathbf{v}_h|_T \in P_1 \times P_1, \quad \forall T \in \mathcal{T}_{h/2}, \quad \mathbf{v}_h \cdot \mathbf{n} = 0 \text{ on } \Gamma_N \right\},$$

$$L_h = \left\{ q_h \in C^0(\bar{\Omega}) : q_h|_T \in P_1, \quad \forall T \in \mathcal{T}_h \right\},$$

We apply this mixed method, particularly to solve the integral equations in steps 1 and 2, as well as for the calculation of the weak divergence to obtain g^0 in step 1 and w^k in step 2. Those calculations require this mixed method, or any other stable finite element pair, to avoid instabilities in the numerical solution. Actually, the main cost of this algorithm is the solution at each iteration of the integral equation to get $\underline{\mathbf{u}}^k$ and the calculation of w^k . However, if the trapezoidal rule is applied to approximate the left hand side of the integral equations, we obtain a system of algebraic equations with diagonal matrix, and the cost to solve them is just a vector multiplication. We call this new algorithm the *CG–algorithm*.

Example 2. We consider again the initial horizontal field $\mathbf{u}^1 = (x, 0)$, as in Example 1 to test the performance of the *CG–algorithm*. In order to compare the numerical results with those obtained with the *E1–algorithm*, we chose $h = 1/40$ and $h/2 = 1/80$ in this case. To stop the iterations we choose $\varepsilon = 10^{-8}$ at step 3. Figure 3 shows the exact and the adjusted wind fields. The agreement is excellent this time, even at the vertical boundaries $x = 1$ and $x = 2$. The relative error and the average divergence are $e_r = 5.9 \times 10^{-4}$ and $mdiv = -5.3 \times 10^{-12}$, respectively. Note that we got a significant improvement: nearly two orders of magnitude better on the relative error, and about ten orders of magnitude better on the average divergence. The improvement of the relative error is mainly due to the reduction of the error on truncated boundaries, while the enhancing of average divergence is mainly due to the iterative method, because it stops when it reaches the tolerance (i.e. when the norm of the divergence is small enough).

To test further the *CG–algorithm* we consider two, more “realistic”, additional examples. The first one includes a domain with a topography of a cosine–shape, and the second one includes a domain with a real topography. In both cases, the “exact” wind field was obtained with a Stokes solver using the methodology described in (Glowinski, 2003). The initial wind field \mathbf{u}^1 was obtained dropping the vertical component of the “exact” one in both cases. Then, the vector wind field is recovered using the same discretization parameters as in example 2.

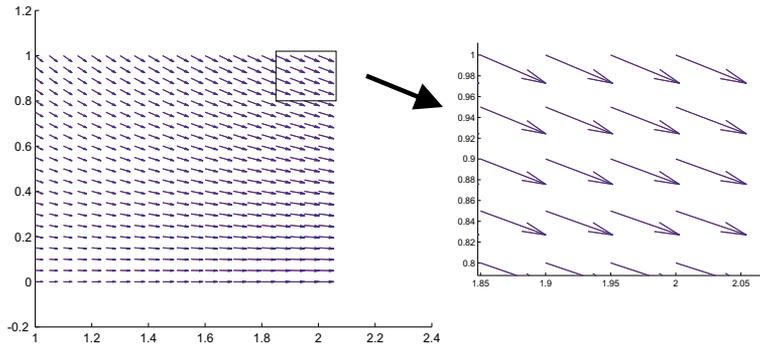


Fig. 3. Exact field $\mathbf{u} = (x, -z)$ in red, adjusted field obtained by the *CG-algorithm* in blue.

Example 3. Cosine-shape topography. In this case, we define the domain as follows

$$\Omega = \left\{ (x, y) \in \mathbb{R}^2 : 0 < x < 10, \frac{1}{2} \cos \frac{3\pi x}{10} + \frac{1}{2} < y < 10 \right\}.$$

The “exact” wind field satisfies $\nabla \cdot \mathbf{u} = 1.2 \times 10^{-16}$. Figure 4 shows the adjusted and “exact” wind fields.

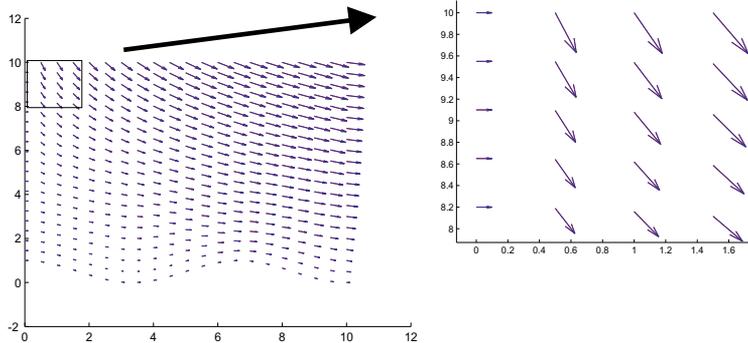


Fig. 4. “Exact” field for cosine topography in red, adjusted field obtained by the *CG-algorithm* in blue.

Example 4. Terrain elevation from real data. In this case, the domain is defined as

$$\Omega = \left\{ (x, y) \in \mathbb{R}^2 : 0 < x < 10, h(x) < y < 10 \right\},$$

where $h(x)$ is a function constructed via cubic splines, which interpolate discrete data over 10 Km of real topography of a certain region in Mexico, contained in a database (GTOPO, 1997). The “exact” wind field satisfies $\nabla \cdot \mathbf{u} = 6.1 \times 10^{-16}$. Figure 5 shows the adjusted and “exact” wind fields. We have an excellent agreement in all cases, even on truncated artificial boundaries. The relative error and the computed mean divergence are about the same order as in example 2. Table 2 shows a summary of the numerical results obtained with

the *CG-algorithm*. All numerical calculations were performed in a DELL Latitude D610 2.13 GHz laptop with an Intel Pentium M processor and 2 GB of RAM.

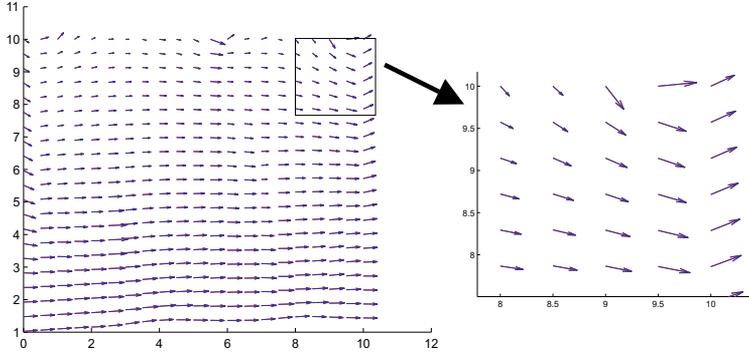


Fig. 5. “Exact” field for real topography in red, adjusted field obtained by the *CG-algorithm* in blue.

Example	Case with	e_r	$mdiv$	No. iters.	CPU-time (sec.)
2	$u(x,z) = (x,-z)$	5.9×10^{-4}	-5.3×10^{-12}	1214	3.9
3	cosine topography	3.6×10^{-6}	9.8×10^{-9}	955	3.3
4	real topography	4.1×10^{-6}	5.7×10^{-11}	1000	3.7

Table 2. Performance of the *CG-algorithm* for three different cases.

4.4 Preconditioned conjugate gradient method

The CPU time to solve the problem with the *CG-algorithm*, at the level of accuracy shown in Table 2, is about twice the CPU time needed to solve the problem with the *E1-algorithm*. In order to make this algorithm more reliable we need to speed up the iterative algorithm to get at least a comparable computational efficiency. Fortunately, we have found a good preconditioner for the iterative algorithm. This preconditioner is an optimal one, and we are presently working in its computer implementation, so we only describe here the main ideas without presenting numerical results yet.

Let $B : L_2(\Omega) \rightarrow L_2(\Omega)$ be an operator defined by

$$Bq = \phi_q, \quad \text{where } \phi_q \text{ solves : } \int_{\Omega} (S^{-1} \nabla \phi_q) \cdot \nabla \psi \, d\mathbf{x} = \int_{\Omega} q \psi \, d\mathbf{x} \quad \forall \psi \in H^1(\Omega) \quad (26)$$

Operator B is self-adjoint and elliptic, and satisfies $A(Bq) = q$, inside Ω , for every $q \in L^2(\Omega)$. An easy way to see these properties is considering the differential form of operators A and B :

$$Aq = -\nabla \cdot \mathbf{u}_q = -\nabla \cdot (S^{-1} \nabla q), \quad \text{since } S \mathbf{u}_q = \nabla q \text{ in } \Omega, \quad (27)$$

$$Bq = \phi_q = -[\nabla \cdot (S^{-1} \nabla)]^{-1} q, \quad \text{since } -\nabla \cdot (S^{-1} \nabla \phi_q) = q \text{ in } \Omega. \quad (28)$$

Then, from (27)–(28) we obtain

$$A(Bq) = A\phi_q = -\nabla \cdot (S^{-1}\nabla\phi_q) = q. \quad (29)$$

This shows that B can be used as an optimal preconditioner. Therefore, the additional cost of the preconditioned conjugate gradient algorithm is the solution of an elliptic problem at each iteration. However, this additional cost is offsetted by two nice properties: a) the preconditioning must reduce drastically the number of iterations (from about 1000 to less than 20, based on previous experience in CFD); b) there is a significant reduction of degrees of freedom in the discrete version of the elliptic problem associated to operator B . This elliptic problem is solved in a coarse mesh, and it is four times smaller than the elliptic problem for the multiplier λ in the 2-D case, and about eight times smaller in 3-D problems.

5. Some extensions and future research

In this section, we present some additional alternatives to look at the problem. We first consider a different set of boundary conditions on vertical truncated boundaries for the multiplier λ , and we show that it produces better results than the traditional approach. We also show that if we introduce ghost nodes on the truncated artificial boundaries, we get even a better improvement. Finally, we introduce radial basis functions to solve the elliptic problems for the multiplier, and show that this is a promising alternative for 3-D wind fields.

5.1 Alternative boundary conditions for the elliptic problem

From equations (19)–(20), we obtain

$$\int_{\Omega} (S\mathbf{u} - \nabla\lambda - S\mathbf{u}^I) \cdot \mathbf{v} \, d\mathbf{x} = \int_{\Gamma \setminus \Gamma_N} \lambda \mathbf{v} \cdot \mathbf{n} \, d\Gamma, \quad \forall \mathbf{v} \in \mathbf{V}_N, \quad (30)$$

$$\int_{\Omega} q \nabla \cdot \mathbf{u} \, d\mathbf{x} = 0, \quad \forall q \in L_2(\Omega). \quad (31)$$

The boundary integral in (30) vanishes in two cases, namely: when $\mathbf{v} \cdot \mathbf{n} = 0$ or when $\lambda = 0$ on $\Gamma \setminus \Gamma_N$. The first case is not possible since it holds only on Γ_N , and the second case is not a good choice on vertical boundaries as we have already seen in Section 3. However, there is a possibility: decompose Γ_D as the union of the vertical boundaries, Γ_V , and the top boundary, Γ_T . Now, at Γ_T we still impose $\lambda = 0$, and on Γ_V we impose the new boundary condition $\mathbf{u} \cdot \mathbf{n} = \mathbf{u}^I \cdot \mathbf{n}$. This new boundary condition is reasonable, since we assume that \mathbf{u}^I is the horizontal part of \mathbf{u} . Therefore, with this choice, we obtain the saddle-point problem (left) and its corresponding elliptic problem (right):

$$S\mathbf{u} - \nabla\lambda = S\mathbf{u}^I, \quad \text{and} \quad \nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega, \quad -\nabla \cdot (S^{-1}\nabla\lambda) = \nabla \cdot \mathbf{u}^I \quad \text{in } \Omega, \quad (32)$$

$$\lambda = 0 \quad \text{on } \Gamma_T, \quad \lambda = 0 \quad \text{on } \Gamma_T, \quad (33)$$

$$\mathbf{u} \cdot \mathbf{n} = \mathbf{u}^I \cdot \mathbf{n} \quad \text{on } \Gamma_V, \quad -S^{-1}\nabla\lambda \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_V, \quad (34)$$

$$\mathbf{u} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_N, \quad -S^{-1}\nabla\lambda \cdot \mathbf{n} = \mathbf{u}^I \cdot \mathbf{n} \quad \text{on } \Gamma_N. \quad (35)$$

The finite element algorithm for the elliptic problem is: Given $\mathbf{u}_h^I \in \mathbf{L}_h$, find $\lambda_h \in H_h$ such that

$$\int_{\Omega} S^{-1} \nabla \lambda_h \cdot \nabla q \, d\mathbf{x} = - \int_{\Omega} \mathbf{u}_h^I \cdot \nabla q \, d\mathbf{x} + \int_{\Gamma_V} q \mathbf{u}_h^I \cdot \mathbf{n} \, d\Gamma, \quad \forall q \in H_h, \quad (36)$$

where H_h is defined as in (13), but with $q = 0$ on Γ_T instead of Γ_D . Equation (36) differs from equation (14) only by the boundary integral on Γ_V . We call (36), together with (15), the *E2-algorithm*.

Example 5. Let us consider one more time the problem introduced in example 1 and in example 3, with the same discretization parameter, $h = 1/80$. The recovered wind field for both cases is better than the one obtained with the *E1-algorithm*, since the vertical component is recovered fairly well, not only in the interior of the domain but also at the vertical boundaries. Figure 6 shows the exact and recovered wind fields. Table 3 shows a summary of the results obtained in examples 1, 2, 3 and 5. For the case with exact wind field $\mathbf{u} = (x, -z)$ the immediate effect of this improvement is the reduction of the relative error by two orders of magnitude. However, we do not obtain a comparable reduction of the mean divergence. For the problems with cosine topography occurs the opposite. Actually, the numerical results show that the most effective algorithm to reduce both, the relative error and the average divergence is the *CG-algorithm*.

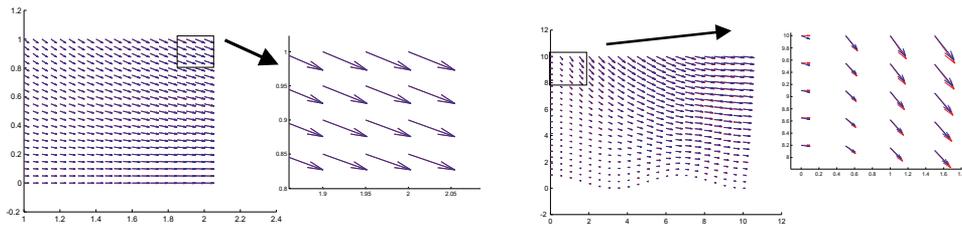


Fig. 6. Exact field (red) and adjusted field obtained by the *E2-algorithm* (blue). Left: case with exact field $\mathbf{u} = (x, -z)$. Right: case with cosine topography.

Ex.	Case with	Algorithm	ϵ_r	$mdiv$	No. iters.	CPU time(s)
1	$\mathbf{u} = (x, -z)$	<i>E1-algorithm</i>	1.9×10^{-2}	4.1×10^{-2}	—	1.78
2	$\mathbf{u} = (x, -z)$	<i>CG-algorithm</i>	5.9×10^{-4}	-5.3×10^{-12}	1214	3.90
5	$\mathbf{u} = (x, -z)$	<i>E2-algorithm</i>	4.0×10^{-4}	1.8×10^{-2}	—	1.78
3	cosine topography	<i>CG-algorithm</i>	3.6×10^{-6}	9.8×10^{-9}	955	3.30
5	cosine topography	<i>E2-algorithm</i>	9.2×10^{-2}	3.6×10^{-4}	—	2.08

Table 3. Summary of the numerical results obtained in examples 1, 2, 3 and 5.

5.2 Ghost nodes

Given that \mathbf{u} belongs to \mathbf{V} and satisfies (2), then λ satisfies the equations

$$-\nabla \cdot (S^{-1} \nabla \lambda) = \nabla \cdot \mathbf{u}^I, \quad \text{in } \Omega, \quad (37)$$

$$-(S^{-1} \nabla \lambda) \cdot \mathbf{n} = \mathbf{u}^I \cdot \mathbf{n}, \quad \text{on } \Gamma_N. \quad (38)$$

Instead of looking for boundary conditions on Γ_D , we may enforce mass conservation by asking any solution of (37)–(38) to satisfy

$$\int_{\Gamma} \mathbf{u} \cdot \mathbf{n} \, d\Gamma = \int_{\Gamma_D} (\mathbf{u}^I + S^{-1} \nabla \lambda) \cdot \mathbf{n} \, d\Gamma = 0. \quad (39)$$

Equations (37)–(39) imply the identity $\int_{\Omega} \nabla \cdot \mathbf{u}^I \, dx = \int_{\Gamma} \mathbf{u}^I \cdot \mathbf{n} \, d\Gamma$. Actually, this is the compatibility condition associated to the the above Poisson–Neumann–like problem. Therefore, this problem has a unique solution $\lambda \in H^1(\Omega)/\mathbb{R}$, and its variational formulation is: Given $\mathbf{u}^I \in \mathbf{L}_2(\Omega)$, find $\lambda \in H^1(\Omega)/\mathbb{R}$ such that

$$\int_{\Omega} (S^{-1} \nabla \lambda) \cdot \nabla q \, dx = - \int_{\Omega} \mathbf{u}^I \cdot \nabla q \, dx + \int_{\Gamma_D} q (\mathbf{u}^I + S^{-1} \nabla \lambda) \cdot \mathbf{n} \, d\Gamma, \quad \forall q \in H^1(\Omega)/\mathbb{R}. \quad (40)$$

Observe that when $q = 1$ in $H^1(\Omega)/\mathbb{R}$, we recover (39). However, the computational solution of this problem is not trivial, since the matrix associated to the discrete version is semidefinite. On the other hand, the symmetry of the matrix is lost because the boundary integral in the right-hand side has the unknown λ . A way to overcome this computational problem is to introduce “ghost nodes”, around and beyond of the nonphysical truncated boundary Γ_D . Then, we may impose $\lambda_h = 0$ and/or $\partial \lambda_h / \partial \mathbf{n} = 0$ on the outer layer of those ghost nodes. At the end, we discard the solution on the ghost nodes, and we only keep the solution values on the actual nodes. Actually, this is a well-known way to deal with differential equations in domains with truncated boundaries.

Example 6. We consider one more time the problem from example 1 with the same discretization parameters. We incorporate two layers of **ghost nodes** and impose $\lambda = 0$ on the outer layer. The recovered wind field obtained is such that the relative error and average weak divergence are $e_r = 2.1 \times 10^{-5}$ and $mdiv = 1.6 \times 10^{-6}$, respectively. The figure with the comparison of the adjusted wind field and the exact wind field is not shown, because it is very similar to Figure 5. Instead, we summarize in Table 4 the results for this example with the different algorithms.

Case	<i>E1-algorithm</i>	<i>E2-algorithm</i>	<i>Ghost-Nodes</i>	<i>CG-algorithm</i>
e_r	1.9×10^{-2}	4.0×10^{-4}	2.1×10^{-5}	5.4×10^{-4}
$mdiv$	4.1×10^{-2}	1.8×10^{-2}	1.6×10^{-6}	-5.2×10^{-12}

Table 4. Comparison of numerical solutions obtained with different algorithms.

Table 4 shows how boundary conditions degrade numerical calculations. It is observed that the solution improves each time the Dirichlet boundary condition $\lambda = 0$ is applied to a smaller section of the non-physical boundary. This is not surprising, since this boundary condition introduces a large artificial gradient, mainly on vertical truncated boundaries, when calculating the term $\nabla \lambda$ in order to get $\mathbf{u} = \mathbf{u}^I + S^{-1} \nabla \lambda$ at the corresponding boundary nodes. At this time, and taking in account the performance of every algorithm, we may recommend to use either the classical approach with ghost nodes or the saddle point problem approach with the conjugate gradient algorithm, specially if we do not have enough information at truncated boundaries.

5.3 Approximation with radial basis functions

A function $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ is called a radial basis function (RBF) if $\Phi(\mathbf{x}) = \phi(\|\mathbf{x}\|)$ where the kernel ϕ is a scalar function $\phi : \mathbb{R}^+ \rightarrow \mathbb{R}$, and d the spatial dimension. Usually $\|\mathbf{x}\|$ is denoted by r , and typical functions used in applications are, among others:

1. Multiquadrics, $\phi(r) = \sqrt{c^2 + r^2}$.
2. Gaussians, $\phi(r) = e^{-cr^2}$.
3. Thin plate splines, $\phi(r) = r^2 \ln(r)$.
4. Inverse multiquadrics, $\phi(r) = 1/\sqrt{c^2 + r^2}$.

The constant value c is called the *shape parameter*. The radial basis function method was first introduced in the 1970s for multivariate scattered data approximation, (Hardy, 1971). This interpolation problem is defined as follows:

Given a set of points $\{\mathbf{x}_j\}_{j=1}^n \subset \Omega \subset \mathbb{R}^d$, approximate the function $f(\mathbf{x})$ from the set of values $f_j = f(\mathbf{x}_j)$. A simple form is to define

$$s(\mathbf{x}) = \sum_{j=1}^k \lambda_j \phi(\|\mathbf{x} - \mathbf{x}_j\|) + p(\mathbf{x}). \quad (41)$$

where p is a polynomial which depends on the specific RBF. Then the interpolation condition

$$s(\mathbf{x}_i) = \sum_{j=1}^k \lambda_j \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) + p(\mathbf{x}_i) = f_i, \quad i = 1, \dots, n, \quad (42)$$

gives an algebraic system of equations for $\lambda = \{\lambda_j\}_{j=1}^k$. However, the corresponding matrix could be ill conditioned and, in some cases, even rank deficient, and special techniques are needed, like preconditioning and least squares (Buhmann, 2003), (Wendland, 2005).

In the last two decades, the main focus of the applications seems to have slowly shifted from scattered data approximation to the numerical solution of PDE. Radial basis function collocation methods for solving PDE are truly meshfree algorithms, in the sense that collocation points can be chosen freely and no connectivity between the points is needed or used (Kansa, 1990), (Narcowich & Ward, 1994). The main attraction of RBF collocation method to solve PDE is that it can be extended directly to solve 3-D problems. Moreover, due to the absence of a grid, these techniques are better suited than classical methods to cope with problems having complex boundaries. So, we think that the RBF collocation method is a good choice to study our problem, and we want to explore this alternative.

Let us denote by \mathcal{L} the linear elliptic differential operator for the multiplier λ , and \mathcal{B} the boundary operator. Suppose that we want to solve the problem

$$\mathcal{L} \lambda = -\nabla \cdot (S^{-1} \nabla \lambda) = \nabla \cdot \mathbf{u}^I \quad \text{in } \Omega, \quad (43)$$

$$\mathcal{B} \lambda = g \quad \text{on } \Gamma_N. \quad (44)$$

We consider a set of collocation points $\{\mathbf{x}_i\}_{i=1}^n$, with n_i points in the interior and n_b points on the boundary, so that $n = n_i + n_b$. We look for an approximate solution $\lambda_h(\mathbf{x}) =$

$\sum_{j=1}^n \lambda_j \phi(\|\mathbf{x} - \mathbf{x}_j\|)$, where the unknown vector $\{\lambda_i\}_{i=1}^n$ satisfies the system of equations

$$\mathcal{L} \lambda_h(\mathbf{x}_i) = \sum_{j=1}^n \lambda_j \mathcal{L} \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) = \nabla \cdot \mathbf{u}^I(\mathbf{x}_i), \quad i = 1, \dots, n_i, \quad (45)$$

$$\mathcal{B} \lambda_h(\mathbf{x}_i) = \sum_{j=1}^n \lambda_j \mathcal{B} \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) = g(\mathbf{x}_i), \quad i = n_i + 1, \dots, n. \quad (46)$$

Therefore the recovered wind field \mathbf{u}_h is given by

$$\mathbf{u}_h(\mathbf{x}) = \mathbf{u}^I(\mathbf{x}) + \sum_{j=1}^n \lambda_j S^{-1} \nabla \phi(\|\mathbf{x} - \mathbf{x}_j\|). \quad (47)$$

Example 8. As a last example we include a 3-D numerical calculation using radial basis functions. The exact synthetic wind field for this example is $\mathbf{u} = (x, y, -2z)$. We dropped the vertical component so that $\mathbf{u}^I = (x, y, 0)$. A multiquadric kernel with $c = 11.33$ was used, and we chose $\alpha_1 = \alpha_2 = \alpha_3 = 1$. The collocation points were obtained by a $5 \times 5 \times 5$ regular subdivision of $\Omega = (1, 2) \times (0, 1) \times (0, 1)$. Figure 7 shows the collocation points and the comparison between the exact and recovered wind field. The agreement is excellent, we obtained a relative error $e_r = 1.98 \times 10^{-4}$ and mean divergence $mdiv = -5.59 \times 10^{-6}$.

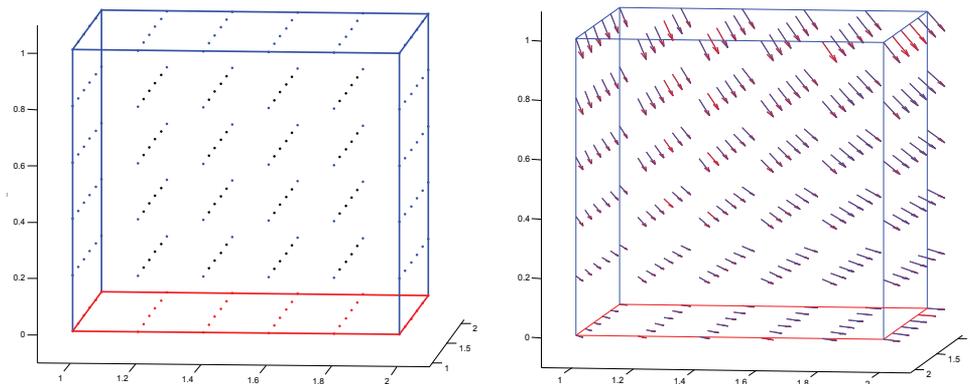


Fig. 7. Collocation points (left), and comparison of exact and recovered wind fields (right).

6. Conclusions

We studied the problem of generating an adjusted wind field from horizontal wind data by different numerical techniques. We have shown that boundary conditions can significantly affect numerical solutions depending of how we treat artificial truncated boundaries. The usual methodology (*E1-algorithm*) does not produce satisfactory results close to the vertical boundaries due to the high gradients introduced by the term $S^{-1} \nabla \lambda$ in (2), when homogeneous Dirichlet boundary conditions are imposed there. The formulation of the problem as a saddle-point one, with a functional equation that has a self-adjoint and strongly elliptic operator, allows the use of an iterative conjugate gradient algorithm (the *CG-algorithm*). This new methodology, in the context of mass consistent models, produces

much better results, it does not involve the solution of differential equations, and it does not require boundary conditions for the multiplier. An optimal preconditioner for the conjugate gradient algorithm was introduced, and we hope, based on previous experience with the solution of the Stokes problem, to reduce the number of iterations of nearly a thousand to a few tens.

In an attempt to improve the numerical results obtained with the traditional approach, we introduced new boundary conditions for the multiplier on vertical boundaries. These boundary conditions are demonstrated to be physically and mathematically consistent. The numerical reconstruction of the wind field was improved by two orders of magnitude, but a comparable reduction on the weak divergence is not always obtained. However, the introduction of “ghost nodes” produces more satisfactory results, reducing both the relative error and the mean weak divergence by two or more orders of magnitude. On the other hand, we have just started to explore meshfree methods. In particular, radial basis collocation methods seem to be a very simple reliably alternative to the reconstruction of three-dimensional wind fields, according to the preliminary numerical results shown in this work.

The application of the different alternatives and methodologies presented here to the more realistic three-dimensional case is a continuation of the present work. Another interesting issue is the potential extension and application of these methodologies to other experimental fields, such as fluid dynamics and computer vision. In particular, the reconstruction of solenoidal velocity fields from experimental data, obtained through the *particle image velocimetry* technique, is an important issue, (Adrian, 2005). Its relation with computer vision is established by *optical flow* estimation, (Ruhnau & Schnorr, 2007).

7. Acknowledgements

Authors wish to express our deep appreciation to InTech for the kind invitation to contribute with this chapter. No doubt this was a great motivation. Special thanks to Ms. Tajana Jevtic and Jana Sertic, InTech process managers, for their guidance, patience and kindness throughout the editorial process.

8. References

- Adrian, R. J. (2005). Twenty years of particle image velocimetry, *Exp. Fluids* 39(2): 159–169.
- Bercovier, M. & Pironneau, O. (1979). Error estimates for the finite element method solution of the stokes problem in the primitive variables, *Numer. Math.* 33: 211–224.
- Buhmann, M. D. (2003). *Radial basis functions: theory and implementations*, Cambridge University Press, United Kingdom.
- Castino, F., Rusca, L. & Solari, G. (2003). Wind climate micro-zoning: a pilot application to liguria region (north-western italy), *J. Wind. Eng. Ind. Aerodyn.* 91(11): 1353–1375.
- Ciarlet, P. G. (2002). *The Finite Element Method for Elliptic Problems. Re-edited as Vol. 40, SIAM, North-Holland Amsterdam (1970)*, Philadelphia, PA.
- Ferragut, L., Montenegro, R., Montero, G., Rodríguez, E., Asensio, M. L. & Escobar, J. M. (2010). Comparison between 2.5-d and 3-d realistic models for wind field adjustmen, *J. Wind. Eng. Ind. Aerodyn.* 98(10–11): 548–558.

- Finardi, S., Tinarelli, G., Nanni, A., Brusasca, G. & Carboni, G. (2010). Evaluation of a 3-d flow and pollutant dispersion modelling system to estimate climatological ground level concentrations in complex coastal sites, *Int. J. Environ. Pollut.* 16(1-6): 472-482.
- Flores, C. F., Juárez, L. H., Nuñez, M. A. & Sandoval, M. L. (2010). Algorithms for vector field generation in mass consistent models, *J. Numer. Methods for PDE* 26(4): 826-842.
- Girault, V. & Raviart, P. A. (1986). *Finite Element Methods for the Navier-Stokes Equations: Theory and Algorithms*, Springer-Verlag, Berlin.
- Glowinski, R. (2003). *Numerical Methods for Fluids (Part 3), Handbook of Numerical Analysis, volume IX*, North-Holland, Amsterdam.
- GTOPO, . (1997). Gtopo30 documentation, section 7, *U. S. Geological Survey* pp. 211-224. URL: www.scd.ucar.edu/dss/datasests/ds758.0html
- Hardy, R. L. (1971). Multiquadric equations of topography and other irregular surfaces, *J. Geophys. Res.* 76(8): 1905-1915.
- Kansa, E. J. (1990). Multiquadrics a scattered data approximation scheme with applications to computational fluid dynamics ii: Solutions to parabolic, hyperbolic and elliptic partial differential equations, *Comput. Math. Appl.* 19(8/9): 147-161.
- Kitada, T., Igarashi, K. & Owada, M. (1986). Numerical analysis of air pollution in a combined field of land/sea breeze and mountain/valley wind, *J. Clim. Appl. Meteorol.* 25(6): 767-784.
- Kitada, T., Kaki, A., H., U. & K., P. L. (1983). Estimation of the vertical air motion from limited horizontal wind data—a numerical experiment, *Atmos. Environ* 17(11): 181-2192.
- Montero, G., Rodríguez, E., Montenegro, R., Escobar, J. M. & González-Yuste, J. M. (2005). Genetic algorithms for an improved parameter estimation with local refinement of tetrahedral meshes in a wind model, *Adv. Eng. Softw.* 36: 3-10.
- Narcowich, F. J. & Ward, J. D. (1994). Generalized hermite interpolation via matrix-valued conditionally positive definite functions, *Math. Comp.* 63: 661-687.
- Núñez, M. A., Flores, C. & Juárez, H. (2007). Interpolation of hydrodynamic velocity data with the continuity equation, *J. Comput. Meth. Sci. Eng.* 7(1): 21-42.
- Núñez, M. A., Flores, C. & Juárez, L. H. (2006). A study of hydrodynamic mass-consistent models, *J. Comput. Meth. Sci. Eng.* 6: 1078-1089.
- Pennel, W. T. (1983). *An evaluation of the role of numerical wind field models in wind turbine siting. Technical Report PNL-SA-11129*, Battelle Memorial Institute, Pacific Northwest Laboratory, Richland, Washington.
- Potter, B. & Butler, B. (2009). Using wind models to more effectively manage wildfire, *Fire management today* 69(2): 40-46.
- Ratto, C. F. (1996). An overview of mass-consistent models, in D. P. Lalas & C. F. Ratto (eds), *Modeling of Atmosphere Flow Fields*, World Scientific Publications, Place of publication, pp. 379-400.
- Ratto, C. F., Festa, R., Romeo, C., Frumento, O. A. & Galluzzi, M. (1994). Mass-consistent models for wind fields over complex terrain: The state of the art, *Environ. Software* 9(4): 247-268.
- Ross, D. G., Smith, I. N., Manins, P. C. & Fox, D. G. (1988). Diagnostic wind field modeling for complex terrain: Model development and testing, *J. Appl. Meteor.* 27: 785-796.
- Ruhnau, P. & Schnorr, C. (2007). Optical stokes flow estimation: an imaging-based control approach, *Exp. Fluids* 42(1): 61-78.

- Sasaki, Y. (1958). An objective analysis based on the variational method, *Journal Met. Soc. Japan* 36: 77–88.
- Sherman, C. A. (1978). A mass-consistent model for wind fields over complex terrain, *J. Appl. Meteor.* Vol. 17(3): 312–319.
- Wang, Y., Williamson, C., Garvey, D., Chang, S. & Cogan, J. (2005). Application of a multigrid method to a mass-consistent diagnostic wind model, *J. Appl. Meteor.* 44(7): 1078–1089.
- Wendland, H. (2005). *Scattered data approximation*, Cambridge University Press, United Kingdom.

Ventilation Effectiveness Measurements Using Tracer Gas Technique

Hwataik Han
Kookmin University
Korea

1. Introduction

Ventilation effectiveness has been defined in various ways by many investigators. The term *ventilation efficiency* was first used by Yaglou and Witheridge (1937). They defined it as the ratio of the carbon dioxide concentration in a room to that in the extract duct. The ventilation was considered to be effective if the air contaminants in high concentration level are captured by the exhaust before it spreads out into the room. This definition has been the cornerstone of various definitions of ventilation efficiency ever since.

The mathematical concepts of age and residence time were introduced in investigations of mixing characteristics in reactors by chemical engineers such as Danckwerts (1958) and Spalding (1958). They mentioned the similarity between the mixing of gases in reactors and the mixing of air in ventilated rooms. Sandberg (1981) first applied the concept of *age of air* to ventilation studies. He summarized various definitions of ventilation efficiency including relative efficiency, absolute efficiency, steady state efficiency, and transient efficiency. The sooner the supply air reaches a particular point in the room, the greater the air change efficiency at that point. This concept has been widely accepted by many researchers and organizations throughout the world including ASHRAE and AIVC.

The ASHRAE Handbook (2009) states that *ventilation effectiveness* is a description of an air distribution system's ability to remove internally generated pollutants from a building, zone, or space, whereas *air change effectiveness* is defined as a description of a system's ability to deliver ventilation air to a building, zone, or space. Thus, ventilation effectiveness indicates the effectiveness of exhaust, whereas the air change effectiveness indicates the effectiveness of supply. However, the terminology *ventilation effectiveness* commonly includes both supply and exhaust characteristics.

In this chapter, we provide a one-to-one analogy between exhaust effectiveness and supply effectiveness using the concept of the age of air. The meanings of local and overall values of supply and exhaust effectiveness need be understood appropriately in conjunction with the aforementioned definitions of ventilation effectiveness.

We also extend the theory of the local mean age of air and local mean residual lifetime of contaminant to a space with multiple inlets and outlets. Theoretical considerations are given to derive the relations between the LMAs from individual inlets and the combined LMA of total supply air. In addition, the relations between the LMRs toward individual outlets and the combined LMR of total exhaust air are considered. These relations can be used to investigate the effect of each supply inlet and/or the contribution of each exhaust outlet in a

space with multiple inlets and outlets. Three examples of tracer gas applications are included in this chapter.

2. Definitions of ventilation effectiveness

2.1 Age and residual-life-time

Consider a point P in a room with one supply air inlet and one return air exhaust. The *age of air* is the length of time required for the supply air to reach the point. As air can reach the point through various paths, the mean value of the ages at the point is called the *local mean age* (LMA) of the air at P . Likewise, the length of time required for the contaminant located at P to reach an exhaust is called the *residual lifetime of the contaminant* at P . The mean value through various paths is the *local mean residual lifetime* (LMR).

Local mean age represents the un-freshness of supply air so that it can be used as a local supply index at the point. Local mean residual lifetime represents the slowness of removal of the contaminant generated at the point, and can be used to represent a local exhaust index. The LMA and LMR represent the local supply and exhaust effectiveness, respectively, at the point in the room. We note that they depend on the room airflow pattern only, and should not be dependent on the source distribution of a contaminant in the space unless the contaminant concentration alters the airflow characteristics of the room.

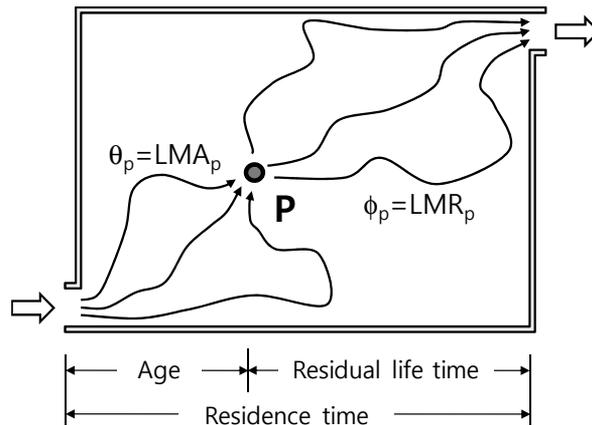


Fig. 1. Concept of age and residual lifetime of indoor air.

2.2 Supply and exhaust effectiveness

A complete mixing condition is considered to be a reference condition we can use to define ventilation effectiveness. The *air change rate* is the number of room volumes of air supplied in one hour, and is defined as Q/V where Q is the volumetric flow rate of air into the room and V is the room volume. The room *nominal time constant* τ_n is the inverse of the air change rate.

The local supply and exhaust indices are defined as the ratios of the local mean age and the local mean residual lifetime compared to the nominal time constant, respectively. These local indices can exceed 100% and can be as large as infinity.

Notice that the LMA at the exhaust means the total residence time of supply air in the space, which is the same as the LMR at the supply. We note that these values are equal to the nominal time constant.

$$\theta_{ex} = \phi_{sup} = \tau_n \quad (1)$$

Therefore, the *local supply index* can be understood as the ratio of the LMA at the exhaust to that at the point, and the *local exhaust index* as the ratio of the LMR at the supply to that at the point. The overall room effectiveness can be defined similarly. The definitions of supply and exhaust effectiveness are shown in Table 1. The subscript P is the location of interest, and $\langle \rangle$ indicates the spatial average over the entire space.

It will be proved later in this chapter that the room averages of LMA and LMR are identical. Therefore, the overall room supply effectiveness and exhaust effectiveness should be the same. We do not need to distinguish the overall values, but we call this the *room ventilation effectiveness*. Note that supply effectiveness and exhaust effectiveness are meaningful only for local values.

SUPPLY EFFECTIVENESS	EXHAUST EFFECTIVENESS
Age of Air $\theta_p =$ Local Mean Age at P $\langle \theta \rangle =$ Room Average of LMA	Residual Life Time of Air $\phi_p =$ Local Mean Residual Lifetime at P $\langle \phi \rangle =$ Room Average of LMR
Local Supply Index $\alpha_p = \frac{\tau_n}{\theta_p} = \frac{\theta_{ex}}{\theta_p}$ = LMA at exhaust/LMA at P	Local Exhaust Index $\varepsilon_p = \frac{\tau_n}{\phi_p} = \frac{\phi_{sup}}{\phi_p}$ = LMR at supply/LMR at P
Overall Room Supply Effectiveness $\langle \alpha \rangle = \frac{\tau_n}{\langle \theta \rangle}$	Overall Room Exhaust Effectiveness $\langle \varepsilon \rangle = \frac{\tau_n}{\langle \phi \rangle}$

Table 1. Definitions of supply and exhaust effectiveness using LMA and LMR

3. Tracer gas technique

3.1 Tracer gases

Tracer gas techniques have been widely used to measure air change rates and the air change effectiveness in a ventilated zone. Any measurable gas can be used as a tracer gas. It is desirable to follow air movements faithfully and for the gas to be nonreactive with other materials. Etheridge & Sandberg (1996) suggested that an ideal tracer gas should have the following characteristics:

- Not a normal constituent of the environment to be investigated.
- Easily measurable, preferably at low concentrations.
- Non-toxic and non-allergic to permit its use in occupied spaces.
- Nonreactive and non-flammable.
- Environmentally friendly.
- Economical.

A wide variety of gases have been employed as tracers. The characteristics of the most commonly used tracer gasses are given in Table 2. Carbon dioxide is a good tracer gas since it has a molecular weight similar to air and is mixed well with air. However, it has a background concentration of approximately 350 ppm, and it is produced by people and the

combustion of fuels in occupied spaces. The effect of the production should be compensated. Hydrogen gas and water vapor have also been used as tracer gases by Dufton and Marley (1935). They pointed out problems related to phase change and adhesion on surfaces when using water vapor. Sulfur hexafluoride is also a common tracer gas used in various ventilated spaces. It is not present in normal ambient air and can be used at very low concentrations. This minimizes the amount of tracer gas needed for a test. However, it has a molecular weight approximately five times that of air, and it should be diluted and/or well mixed with the surrounding air during injection.

Gas	Molecular weight	Boiling point (°C)	Density (15°) (kg/m ³)	Analytical method	Detection range (ppm)	Background concentration	Toxicity
Carbon dioxide	44	-56.6	1.98	IR	0.05-2000	Variable	Slight
Freon12	121	-29.8	5.13	IR GC-ECD	0.05-2000 0.001-0.05		
Helium	4	-268.9	0.17	MS		5.24	
Nitrous oxide	44	-88.5	1.85	IR	0.05-2000	0.03	
Sulphur hexafluoride	146	-50.8	6.18	IR GC-ECD	0.05-2000 0.00002-0.5		
Perfluron hexane	338	57.0		GC-ECD	10 ⁻⁸		

Table 2. Characteristics of commonly used tracer gases (Sandberg, 1981)

3.2 Tracer gas system

A general tracer gas system is composed of an injection and distribution system, a sampling and monitoring system, and a data acquisition and control system. An example of a typical experimental setup is shown in Fig. 2 (ASTM, 1993).

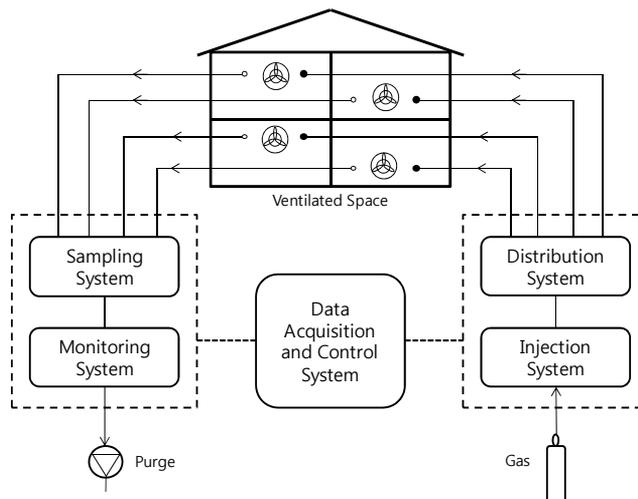


Fig. 2. Typical tracer gas experimental system.

3.2.1 Injection and distribution

The injection and distribution system releases an appropriate amount of tracer gas and distributes it into the zones. There are several means of releasing tracer gas, either manually or automatically. A graduated syringe or other containers of known volume may be used for simple manual injections. For automated injection systems, a compressed tracer gas supply is connected to a gas line with an electronic mass flow controller, or other tracer gas flow rate measurement and control devices.

An automatic distributing system includes a tubing network that dispenses a tracer gas via manifolds and automated valves, and pressure-operated valves that stop the flow from entering the tubing network when the tubing is not pressurized. There should be no leaks in the tubing. A mixing fan is frequently used for good mixing of tracer gases within a zone.

3.2.2 Sampling and monitoring

Air sampling can be achieved either manually or automatically. Manual samplers may include syringes, flexible bottles, or sampling bags with a capacity of at least three times the minimum sampler size of the gas analyzer used. Automatic samplers may utilize either a sampling network or automated samplers. Sampling networks consist of tubing, a manifold or selection switch that is typically solenoid-driven, and a pump that draws air samples through the network. Tracer gas molecules should not adhere to the tubing or manifold surfaces. Materials that absorb tracer gas may cause major inaccuracies in the measurement. There are various types of gas analyzers based on principles such as infrared spectroscopy, gas chromatography, or mass spectroscopy. A gas analyzer should be suited to the tracer gas used, and the concentration range studied.

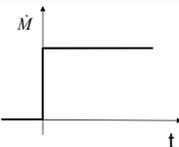
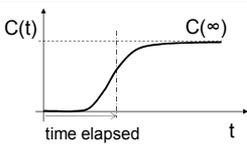
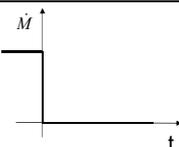
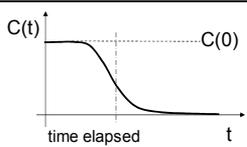
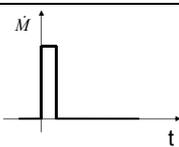
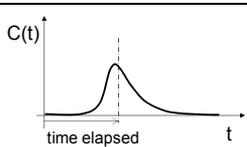
	INJECTION	MONITORING
Step-up method		
Step-down method		
Pulse method		

Table 3. Tracer injection methods and the corresponding concentration responses.

3.3 Tracer injection methods

There are three commonly used methods of injecting a tracer gas: step-up, step-down, and pulse methods. The *step-up method* introduces a tracer gas at a given time and onward until

it reaches a steady state. The concentration response at a monitoring point is observed continuously. As a steady state is reached, the concentration is maintained at the steady state value. The *step-down method* is the opposite of the step-up method. Tracer injection is stopped abruptly and the concentration decay is monitored at a monitoring point. The concentration decays exponentially and approaches a background concentration. The concentration decay method is frequently used to measure air change rate starting from a uniform mixing of room air. Finally, the *pulse method* introduces a certain amount of tracer gas in a short period of time. A peak concentration response is detected at a monitoring point with a time delay. The concentration decays down to an initial concentration after the peak. Table 3 shows concentration responses according to the three injection methods.

4. Measurements of ventilation effectiveness

4.1 LMA measurements

In order to measure the local mean age at point P, the tracer injection point should be at a supply diffuser and the monitoring point is at point P as shown in Fig. 3. LMA can be obtained by integrating the area above the concentration curve (shaded area) divided by the steady state concentration after a step-up tracer injection. Similarly, it is the area under the concentration curve for a step-down method. In the case of a pulse method, it can be calculated using the first moment of the area under the concentration curve. The equations used to calculate LMAs are shown in Table 4 for three injection procedures. The equations are different from one injection method to another, but the result should be the same. The superscripts and the subscripts of the concentrations indicate the injection and the monitoring points, respectively.

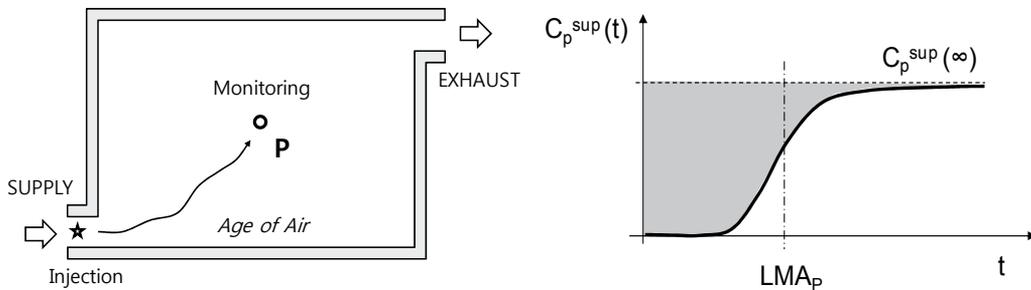


Fig. 3. Injection and monitoring points for LMA and transient step-up response.

It is known that the LMA distribution in a space is equivalent to the steady concentration distribution with uniformly-distributed sources in the space (Han, 1992). The proof is given in the appendix. Thus,

$$\theta_p = \frac{\bar{C}_p(\infty)}{\bar{m}} \quad (2)$$

where \bar{m} is the tracer generation rate per unit volume. In Eq. (2), C has an over-bar rather than a superscript, which represents a uniform tracer injection throughout the entire space. The local supply index, which is the ratio of the LMAs at the exhaust and at P, is calculated using the ratio of the steady concentrations with over-bars at those points. The steady

concentration at the exhaust can be obtained from the total tracer generation rate in the space, which is the product of the generation rate per volume times the space volume. Thus,

$$\alpha_p = \frac{\theta_{ex}}{\theta_p} = \frac{\bar{C}_{ex}(\infty)}{C_p(\infty)} \quad (3)$$

	LOCAL MEAN AGE	LOCAL MEAN RESIDUAL-LIFE-TIME
Step-up	$\theta_p = \int_0^{\infty} \left(1 - \frac{C_p^{\text{sup}}(t)}{C_p^{\text{sup}}(\infty)} \right) dt$	$\phi_p = \int_0^{\infty} \left(1 - \frac{C_{ex}^P(t)}{C_{ex}^P(\infty)} \right) dt$
Step-down	$\theta_p = \int_0^{\infty} \frac{C_p^{\text{sup}}(t)}{C_p^{\text{sup}}(0)} dt$	$\phi_p = \int_0^{\infty} \frac{C_{ex}^P(t)}{C_{ex}^P(0)} dt$
Pulse	$\theta_p = \frac{\int_0^{\infty} t \cdot C_p^{\text{sup}}(t) dt}{\int_0^{\infty} C_p^{\text{sup}}(t) dt}$	$\phi_p = \frac{\int_0^{\infty} t \cdot C_{ex}^P(t) dt}{\int_0^{\infty} C_{ex}^P(t) dt}$

Table 4. Equations to calculate LMA and LMR for three injection methods

4.2 LMR measurements

In order to measure the local mean residual lifetime at P, the injection point should be at P and the monitoring point should be at the exhaust. The LMR can be obtained using the equations in Table 4 similar to LMA equations.

In a step-up method, the exhaust concentration reaches a steady state value $C_{ex}^P(\infty)$ as time goes to infinity. The mass balance should be satisfied; thus, the steady concentration at the exhaust should be equal to the total mass generation divided by the airflow rate, \dot{M}/Q . Therefore, the LMR using a step-up method can be written as

$$\begin{aligned} \phi_p &= \int_0^{\infty} 1 - \frac{C_{ex}^P(t)}{\dot{M}/Q} dt \\ &= \frac{1}{\dot{M}} \int_0^{\infty} \dot{M} - Q \cdot C_{ex}^P(t) dt \end{aligned} \quad (4)$$

where \dot{M} is the contaminant generation rate at P. The first term in the integral is the total generation rate, and the second term is the rate of contaminant leaving the room through the extract duct. The integration of the difference up to the steady state results in the amount of contaminant left inside the room, which is called the *internal hold-up*. This is the product of the average room concentration times the room volume (Sandberg, 1981). Then, Eq. (4) can be written as

$$\begin{aligned} \phi_p &= \frac{\langle C^P(\infty) \rangle V}{\dot{M}} \\ &= \frac{\langle C^P(\infty) \rangle \tau_n}{C_{ex}^P(\infty)} \end{aligned} \quad (5)$$

The local exhaust index can be obtained either from the definition of the LMR ratio, or by the ratio of the room average concentration to the exhaust concentration when a source is located at P. Thus,

$$\varepsilon_p = \frac{C_{ex}^P(\infty)}{\langle C^P(\infty) \rangle} \quad (6)$$

Equation (6) looks quite similar to the classical definition by Yaglou and Witheridge (1937). They also defined ventilation efficiency as the ratio of the room average concentration to exhaust concentration for a given contaminant source. They understood this quantity as the overall efficiency of the room, not as a local efficiency at the given source location, though. The ratio shown in Eq. (6) is not the room exhaust index, but the local exhaust index for a given source located at P. Various definitions have been proposed for removal effectiveness by several authors (Sandberg and Sjoberg, 1983; Skaaret, 1986). Although there have been many studies on the measurement of LMA (Shaw *et al.*, 1992; Han *et al.*, 1999; Xing *et al.*, 2001), the distributions of LMR have rarely been measured experimentally (Han *et al.*, 2002).

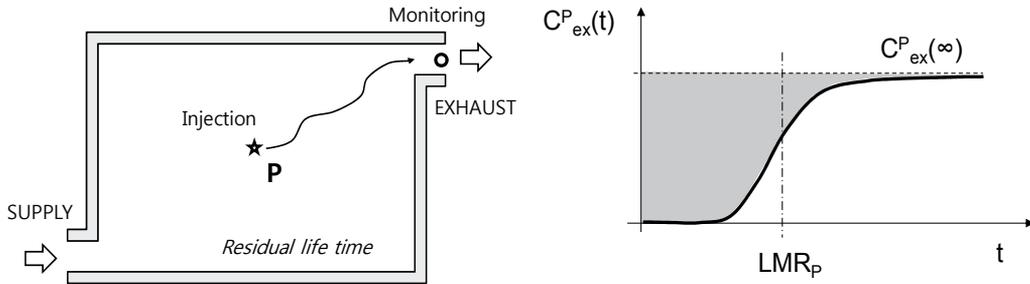


Fig. 4. Injection and monitoring points for LMR and transient step-up response.

4.3 Overall ventilation effectiveness

The overall room effectiveness is the spatial average of local values over the entire space. As previously discussed, LMA can be obtained using transient and steady approaches. The steady method indicates that the spatial average of LMA is the spatial average of the steady concentration distribution with uniformly distributed tracer sources of unit strength, as follows:

$$\langle \theta \rangle = \frac{\langle \bar{C}(\infty) \rangle}{in} \quad (7)$$

Therefore, the overall supply effectiveness; i.e., the ratio of LMA at exhaust to the room average LMA, equals the ratio of the concentration at exhaust to the spatial average of the steady concentration as follows:

$$\langle \alpha \rangle = \frac{\bar{C}_{ex}(\infty)}{\langle \bar{C}(\infty) \rangle} \quad (8)$$

On the other hand, to obtain the overall exhaust effectiveness, LMR should be obtained at every internal point to calculate its spatial average over the entire space. Unlike the method

used for LMA measurements, a monitoring point should be fixed at the exhaust, and a tracer should be injected at every point in the space repeatedly. The concentration response by simultaneous tracer injections can be obtained by superimposing every injection source present over the entire space, since the concentration equation is linear. Therefore, the room average exhaust effectiveness is the ratio of the exhaust concentration to the room average concentration with a uniformly distributed source superimposed in the space, which is identical to the steady method used to determine overall supply effectiveness. This concludes the proof that supply effectiveness equals the overall exhaust effectiveness of a given space, and that the room mean age of air is identical to the room mean residual lifetime:

$$\langle \varepsilon \rangle = \langle \alpha \rangle \quad (9)$$

The room mean age or the room mean residual lifetime can also be obtained from the transient concentration responses at exhaust according to Table 5 for different tracer injection methods (Kuehn *et al.*, 1998).

	ROOM MEAN AGE = ROOM MEAN RESIDUAL-LIFE-TIME
Step-up method	$\langle \theta \rangle = \langle \phi \rangle = \frac{Q}{V} \int_0^{\infty} t \cdot \left(1 - \frac{C_{ex}^{sup}(t)}{C(\infty)} \right) dt$
Step-down method	$\langle \theta \rangle = \langle \phi \rangle = \frac{Q}{V} \int_0^{\infty} t \frac{C_{ex}^{sup}(t)}{C(0)} dt$
Pulse method	$\langle \theta \rangle = \langle \phi \rangle = \frac{Q}{2V} \frac{\int_0^{\infty} t^2 \cdot C_{ex}^{sup}(t) dt}{\int_0^{\infty} C_{ex}^{sup}(t) dt}$

Table 5. Equations used to calculate RMA and RMR for three injection methods

5. Multiple inlets and outlets

5.1 LMA from multiple inlets

When there are multiple supply inlets, the LMA from one inlet is different than those from the other inlets. Consider a ventilated space configuration with two supply inlets as shown in Fig. 5. The airflow rates through the inlets are Q_a and Q_b , respectively, and room air is exhausted through an outlet on the other side of the space.

Suppose we inject a tracer gas only at inlet *a* by a step-up method. The supply concentration is assumed to be 1.0 at inlet *a* and 0.0 at inlet *b*. The concentration response at P, $C_p^a(t)$, is shown in Fig. 5. LMA_p^a is the area above the curve (left-hatched area). Subscript P represents a monitoring point, and superscript *a* represents an injection location. The steady concentration $C_p^a(\infty)$ has a value ranging between zero and unity because the supply concentration at inlet *a* is non-dimensionalized.

The response after a step-up injection at inlet *b* can be characterized similarly by LMA_p^b and $C_p^b(\infty)$. In this case, the non-dimensional supply concentration is 0.0 at inlet *a* and 1.0 at

inlet b . We note that the steady state concentration $C_p^b(\infty)$ is complementary to $C_p^a(\infty)$, since the inlet concentration boundary conditions are switched.

In the case of simultaneous tracer injections at both inlets, the concentration response at point P is given by the addition of the concentration responses from individual injections, as shown in Fig. 5.

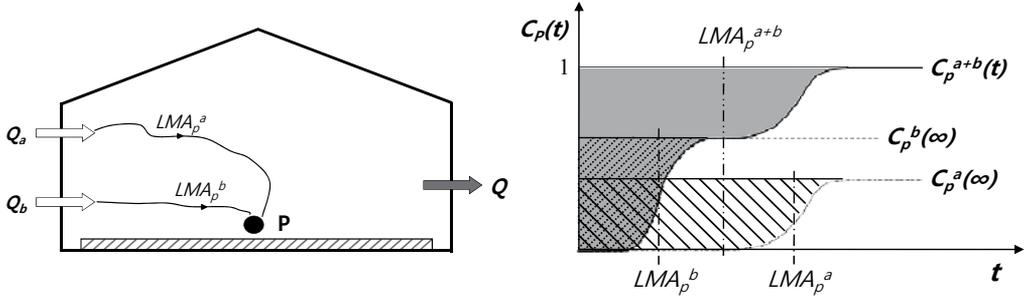


Fig. 5. Local mean age from individual supply inlets and concentration responses at P.

$$C_p^{a+b}(t) = C_p^a(t) + C_p^b(t) \quad (10)$$

This is because the indoor airflow pattern remains unchanged and the governing equation is linear with respect to concentration. Concentrations reach 1.0 at all internal points as a steady state is reached. Thus,

$$1 = C_p^a(\infty) + C_p^b(\infty) \quad (11)$$

The combined LMA is the area above the combined concentration curve, which is the shaded area in Fig. 5. The relations between the LMAs can be derived as follows (Han *et al.*, 2010):

$$LMA_p = C_p^a(\infty) \cdot LMA_p^a + C_p^b(\infty) \cdot LMA_p^b \quad (12)$$

Therefore, the combined LMA is the weighted average of the LMAs from each individual inlet, and the weighting factors for calculating the average are the corresponding steady state concentrations at the point. The steady state concentrations can be considered to be the contribution factors of the corresponding inlets for characterizing the supply air conditions at the point.

5.2 LMR to multiple outlets

If there are multiple outlets, the contribution of each outlet is different with respect to eliminating contaminants generated in a space, depending on the relative source locations. Consider a case with two outlets with exhaust flow rates of Q_a and Q_b as shown in Fig. 6. The time for the contaminant generated at P to reach one exhaust, LMR_p^a , is different from the time to reach the other, LMR_p^b . The total amount of contaminants exhausted by one outlet is different from that exhausted by the other. Figure 6 shows concentration responses at the exhausts according to a step-up injection at point P. The combined exhaust

concentration is the average of individual exhaust concentrations weighted by the airflow rates through the outlets, as follows:

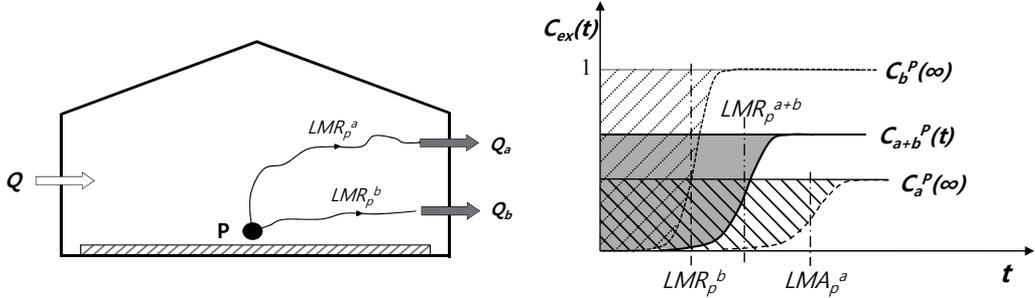


Fig. 6. Local mean residual lifetime at P and concentration responses at the exhausts.

$$C_{ex}(t) = \frac{Q_a}{Q} C_a^P(t) + \frac{Q_b}{Q} C_b^P(t) \quad (13)$$

The individual LMRs to the outlets can be obtained by integrating the areas above the corresponding concentration curves:

$$LMR_p^a = \int_0^{\infty} 1 - \frac{C_a^P(t)}{C_a^P(\infty)} dt \quad (14)$$

$$LMR_p^b = \int_0^{\infty} 1 - \frac{C_b^P(t)}{C_b^P(\infty)} dt$$

Similarly, the combined LMR can be obtained by the area above the average exhaust concentration curve. The combined LMR can be rearranged using Eq. (13), and can be expressed with the individual LMRs as follows:

$$LMR_p = \int_0^{\infty} 1 - \frac{C_{ex}(t)}{C_{ex}(\infty)} dt$$

$$= \int_0^{\infty} \left(\frac{\frac{Q_a}{Q} C_a^P(\infty) + \frac{Q_b}{Q} C_b^P(\infty)}{C_{ex}(\infty)} \right) - \left(\frac{\frac{Q_a}{Q} C_a^P(t) + \frac{Q_b}{Q} C_b^P(t)}{C_{ex}(\infty)} \right) dt \quad (15)$$

$$= \frac{C_a^P(\infty) Q_a}{C_{ex}(\infty) Q} \cdot LMR_p^a + \frac{C_b^P(\infty) Q_b}{C_{ex}(\infty) Q} \cdot LMR_p^b$$

$$= \frac{\dot{M}_a}{\dot{M}} \cdot LMR_p^a + \frac{\dot{M}_b}{\dot{M}} \cdot LMR_p^b$$

Therefore, the combined LMR is the weighted average of the individual LMRs. The weighting factors are the percentages of the contaminant removal rates through the corresponding exhaust outlets. They can be understood as the contribution factors of the individual outlets for a given tracer source at P.

6. Examples of tracer gas applications

6.1 Effect of supply air temperature on LMA distributions

6.1.1 Problem description

It is often observed that fresh air supplied to a space is bypassed directly to an exhaust without contributing to effective room ventilation. Bypass affects the ventilation effectiveness of the room significantly. It is quite common in office buildings, especially when warm air is supplied from ceiling diffusers in the winter season. The following example considers the effect of supply air temperature on LMA distribution in a rectangular space with a diffuser and a return grill on the ceiling.

6.1.2 Experimental setup

A schematic of the experimental chamber is shown in Fig. 7. The chamber measures 1.95 m \times 1.95 m \times 1.45 m. The height of 1.45 m is about one-half of a full-scale office room. The interior surfaces (walls and floors) are made of aluminum panels. By circulating temperature-controlled fluid through the passages embedded in each panel, the temperatures of the walls and floors are precisely controlled. The ceiling is insulated with polystyrene insulation boards of 50 mm thickness. Air is supplied to the chamber through three linear sections that measures 0.635 m in length and 0.0508 m in width each. The three sections are aligned to form a 0.0508 m \times 1.905 m straight slot inlet in the ceiling. The return slot is identical to the inlet and is also placed in the ceiling. This configuration produces a two-dimensional (2-D) flow in the chamber. A detailed description of the physical structure and the control system of the chamber is given by Corpron (1992).

6.1.3 Similitude

In this study, The Reynolds number and Archimedes number are considered important in simulating the full-scale conditions. These dimensionless numbers are defined as follows:

$$\text{Re} = \frac{\rho u L}{\mu} \propto \frac{\text{inertia force}}{\text{viscous force}} \quad (16)$$

$$\text{Ar} = \frac{\beta g L \Delta T}{u^2} \propto \frac{\text{buoyancy force}}{\text{inertia force}} \quad (17)$$

For a half-scale model, the characteristic length is related as $L_m = N L_f$, where N equals 0.5. Subscript m stands for the model and f represents the full scale. As the thermodynamic properties ρ , μ , and β are assumed to be constant for both, the characteristic velocity of the model needs to be increased by a factor of $1/N$. Also, the temperature difference needs to be increased by a factor of $1/N^3$ for similarity. Thus,

$$u_m = \frac{1}{N} u_f \quad (18)$$

$$(\Delta T)_m = \frac{1}{N^3} (\Delta T)_f \quad (19)$$

The air change per hour (ACH) is the ratio of volumetric flow rate to the volume of the room. By a simple mathematical manipulation, the relation of ACH between the model and the prototype becomes

$$(ACH)_m = \frac{1}{N^2}(ACH)_f \quad (20)$$

6.1.4 Experimental procedure

The airflow and temperature conditions of the chamber were adjusted and checked until the steady state was reached. The sampling tube was positioned at a monitoring point using the three-dimensional (3-D) traversing system. Using a syringe, 3 mL of SF₆ gas was injected into the supply duct. The gas monitor started to take data at the same time as the gas injection, which works on the principle of electron capture gas chromatography. Concentration data were recorded every 70 s until the concentration fell within 1% of the maximum concentration. The same measurement was repeated with a delayed injection by 35 s to double the number of data points. The sampling port was then moved to the next position, and the aforementioned procedure was repeated to cover the entire cross-section at the center of the chamber.

In order to investigate the effect of thermal buoyancy, three different temperature conditions were tested: isothermal, cooling, and heating. The experimental conditions and measurements are summarized in Table 6. The values of the corresponding full scale situation are shown in parentheses.

	Isothermal	Cooling	Heating
Pressure drop across nozzle [mmH ₂ O]	17.5	11.2	22.5
Supply velocity at diffuser [m/s]	1.032	0.813	1.209
Supply airflow rate [m ³ /h]	345	272	404
Air change per hour [ACH]	62.6 (15.6)	49.4 (12.4)	73.3 (18.3)
Supply air temperature [°C]	24.4 (24.4)	-1.6 (19.2)	57.1 (30.5)
STD of supply air temperature [°C]	0.1	1.2	1.9
Exhaust air temperature [°C]	24.4 (24.4)	27.6 (22.9)	41.4 (28.5)
Wall temperature [°C]	24.4 (24.4)	46.0 (25.2)	-3.7 (22.9)
STD of wall temperature [°C]	0.1	4.4	2.2
Ceiling temperature [°C]	24.4 (24.4)	24.3 (22.5)	23.4 (26.3)
Mean temperature [°C]	24.4	22.2	26.7
T _{wall} - T _{supply} [°C]	0 (0)	47.6 (6.0)	-60.8 (-7.6)
Reynolds number	3374	3128	3293
Archimedes number	0	0.1215	-0.0690

*Numbers in () indicate the corresponding values in full-scale situations.

Table 6. Experimental conditions and measurements (Han, 1999)

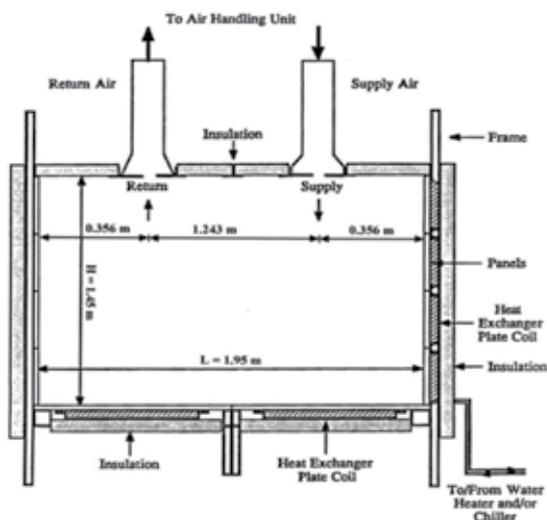


Fig. 7. Cross-section of a thermal chamber (Han, 1999).

6.1.5 Results and discussion

Figure 8 shows LMA contours superimposed with the LMA data measured at 36 locations for an isothermal condition. The LMA distribution is nearly uniform over the entire cross-section except at corners. The maximum LMA was observed at the center, which indicates there is a large recirculation in the middle of the chamber. A velocity vector drawing by Liang (1994) is also shown in Fig. 8. The air jet from the supply inlet moves downward and leaves the chamber through the exhaust after making a large clockwise circulation in the space. The supply jet is attached to the right wall, and then separates before it hits the floor. This tendency for flows to attach to walls is known as the Coanda effect. It is interesting to note that the distribution of local mean age in the space shows a good overall picture of the airflow pattern in the space.

For a cooling condition, the supply air temperature is lower than the room temperature and the buoyant force acts downward, which is the same as the direction of the supply air. Figure 9(a) shows the LMA distribution in the chamber. Assisted by buoyancy, the mixing of the flow is enhanced and the local mean age values are more uniformly distributed compared to the isothermal condition. The location of maximum LMA is shifted downward and to the right in comparison with the isothermal condition. The maximum LMA value is less than that in the isothermal case.

In a heating condition, the thermal buoyancy opposes the inertial effect. The local mean age distribution is shown in Fig. 9(b). A large variation in LMA can be observed in the chamber because of thermal stratification. The variation is small at the upper part and large at the lower part of the chamber. The air jet from the supply port does not seem to penetrate into the space effectively; rather, it short circuits to the exhaust duct. Liang (1994) observed that the flow field under the heating condition was unstable and the supply jet oscillated slowly within the chamber. Because of the oscillatory behavior, velocity vectors could not be measured in the experiment, and only the frequency of the oscillatory motion was reported. The room mean ages obtained by integrating the local mean age values over the entire space give 118 s, 120 s, and 234 s for the isothermal, cooling, and heating conditions, respectively.

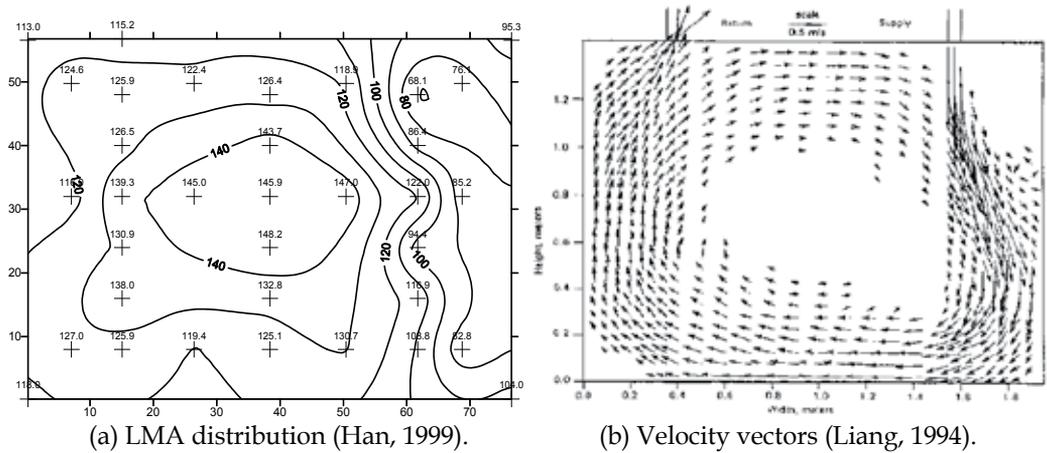


Fig. 8. LMA distribution and velocity vector fields for isothermal condition.

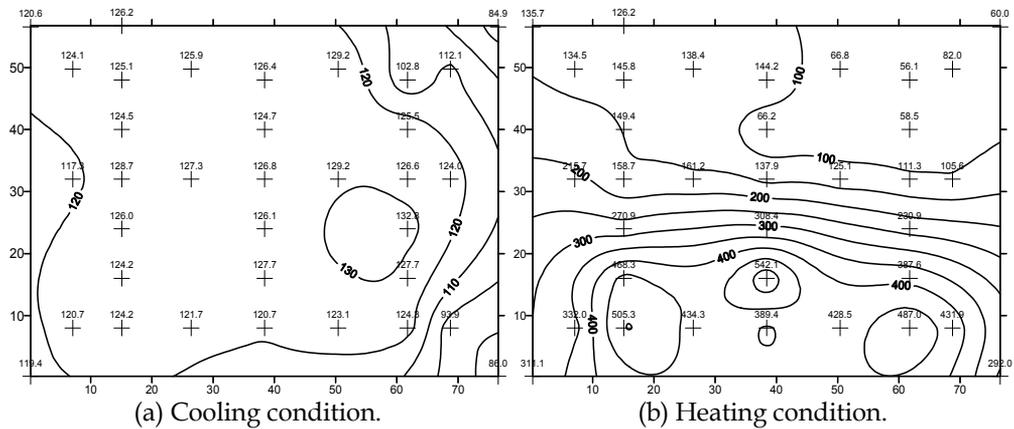


Fig. 9. Local mean age distributions for cooling and heating conditions (Han, 1999).

6.1.6 Concluding remarks

Using a pulsed injection method using SF₆ tracer gas, LMA distributions were measured in a half-scale thermal chamber. Boundary conditions were applied that simulated isothermal, heating, and cooling conditions by controlling the supply air temperature and the wall temperature.

1. The LMA distribution was found to be closely related to the velocity distribution in the chamber. The results for LMA distributions are in good agreement qualitatively with the velocity patterns obtained by Liang (1994).
2. For an isothermal condition, the largest LMA occurred at the center and not at the corners, which indicates that there was a large recirculating zone at the center. During a cooling operation, supply air penetrated deeply into the chamber, and mixing was enhanced compared to the isothermal condition.
3. For a heating condition, there was a large variation of local mean ages due to thermal stratification in the chamber. It can be concluded that local ventilation effectiveness in the lower part of a room can be very poor under heating operations.

Further research needs to be done to improve the tracer gas technique and to apply the technique to various applications.

6.2 Effect of inlet/outlet configurations on LMA and LMR distributions

6.2.1 Problem description

The airflow pattern in a ventilated space varies according to the locations of supply inlets and exhaust outlets. In this example, LMA and LMR distributions are measured and compared in a rectangular enclosure with three different inlet/outlet configurations. A supply slot is fixed at the top of a right wall, and an exhaust slot is varied at the bottom-left (Case 1), bottom-right (Case 2), and top-left (Case 3) locations.

6.2.2 Experimental setup

The experimental chamber has dimensions of 1.8 m x 1.2 m x 0.9 m. There is a supply slot on the top of the right wall, and an exhaust slot at one of the three locations. The supply and exhaust slots are 0.025 m in width, and supply air was discharged horizontally. The airflow rate ranged from 4 to 76 ACH. The pressure inside the chamber was maintained neutral by an exhaust fan in order to minimize infiltration through the envelope. Sulfur hexafluoride at 30% concentration was used as a tracer gas. Using a syringe, 10 mL of SF₆ was injected into a polystyrene tube, and the gas was mixed with a continuous stream of nitrogen. The diluted tracer gas was discharged at a point in the chamber through a porous sphere 40 mm in diameter connected at the end of the injection tube. A tracer gas detector is a multi-gas monitor based on the non-disperse infrared (NDIR) absorption principle. To visualize airflow patterns in the chamber, helium bubbles were discharged into a supply air duct, and a sheet of light was illuminated through a glass window along the center of the chamber. A schematic diagram of the experimental setup is shown in Fig. 10.

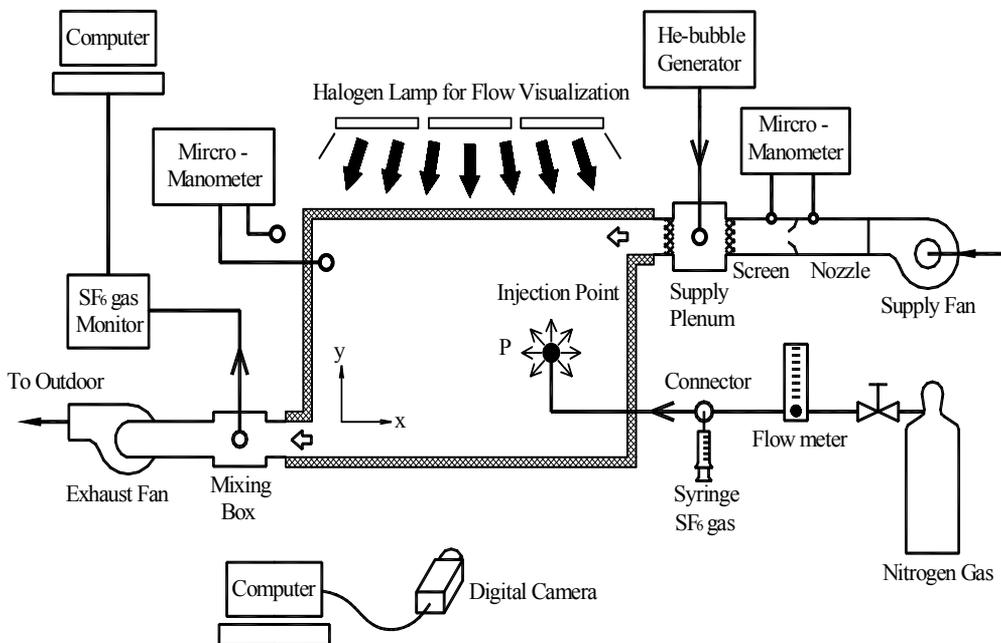


Fig. 10. Schematic diagram of experimental setup (Han *et al.*, 2002).

6.2.3 Experimental procedure

For the LMR measurements, tracer gas was injected through a porous sphere at a point in the chamber and the transient tracer gas concentration variation was measured at the exhaust. After the tracer gas was exhausted completely from the chamber, the injection sphere was moved to another position. The procedure was repeated for other internal points. There are 15 injection points equally spaced in the center plane of the chamber.

For LMA measurements, all the experimental conditions were identical to the LMR case, but tracer gas was injected at a supply duct. Then, transient tracer gas concentration variation was measured at the internal points. Experiments were conducted for three different exhaust locations under isothermal room temperature conditions.

6.2.4 Results and discussion

Flow visualization results are shown in Fig. 11 for three different exhaust locations. The air change rates are 12ACHs. For Case 1, the air supplied in the horizontal direction moved toward the lower-left exhaust in the diagonal direction. The room air formed two large recirculating flows at the upper-left and lower-right corners. For Case 2, the air supplied along the ceiling changed its direction by the opposite wall and made a large counter-clockwise circulation in the chamber. We note that the airflow pattern was quite similar to a complete mixing condition. For Case 3, supplied air faced directly toward the exhaust. The room air was mostly stagnant, but with a slow recirculation due to the viscous action of the bypassing flow along the ceiling.

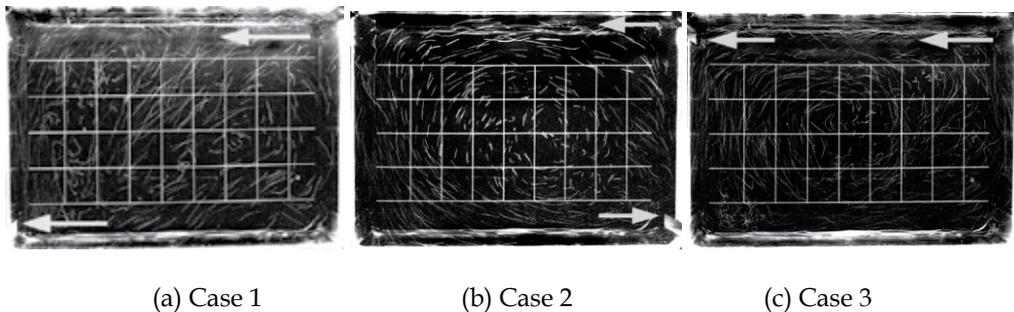


Fig. 11. Flow visualization results for three inlet-outlet configurations.

Contours of LMA and LMR are plotted in Fig. 12 for Case 1. It can be seen that both of the distributions are closely related to the airflow pattern shown in Fig. 11(a). LMA and LMR values are large within recirculating zones. LMA is small adjacent to the supply inlet and large adjacent to the exhaust, whereas LMR is small adjacent to the exhaust and large adjacent to the supply inlet. Figure 13 shows LMA and LMR distributions for Case 2. The LMA near ceiling is relatively small, whereas the LMR near floor is small. Both have large values within a large recirculation zone at the center. Figure 14 shows the results for Case 3. The LMA and LMR are small near ceiling, and large adjacent to the floor. The airflow pattern in Fig. 11(c) indicates there was large stagnant recirculation in the lower part of the space. The tracer gas diffused out into the lower part could not be exhausted effectively.

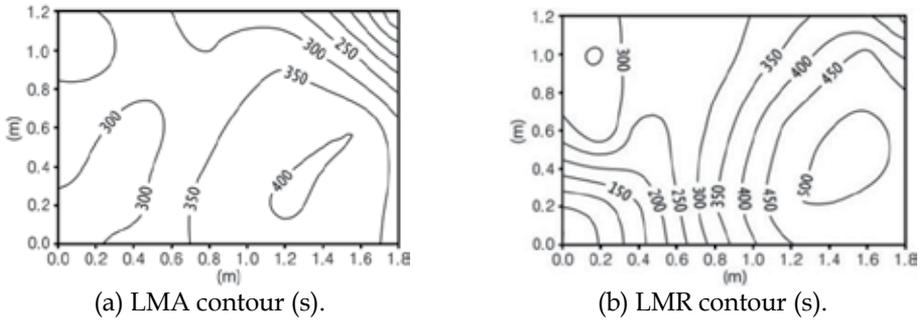


Fig. 12. LMA and LMR distributions for Case 1 (Han *et al.*, 2002).

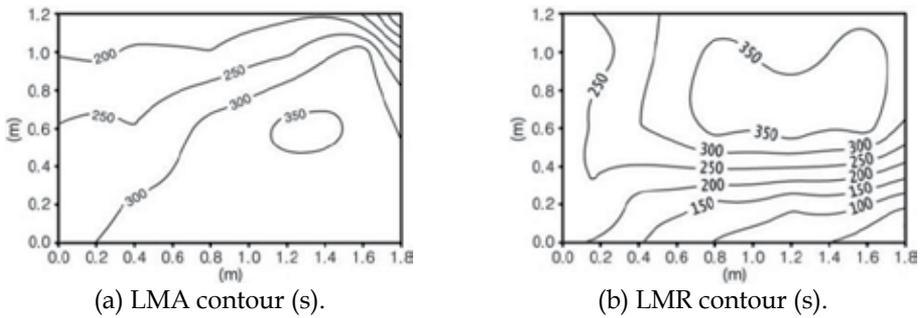


Fig. 13. LMA and LMR distributions for Case 2 (Han *et al.*, 2002).

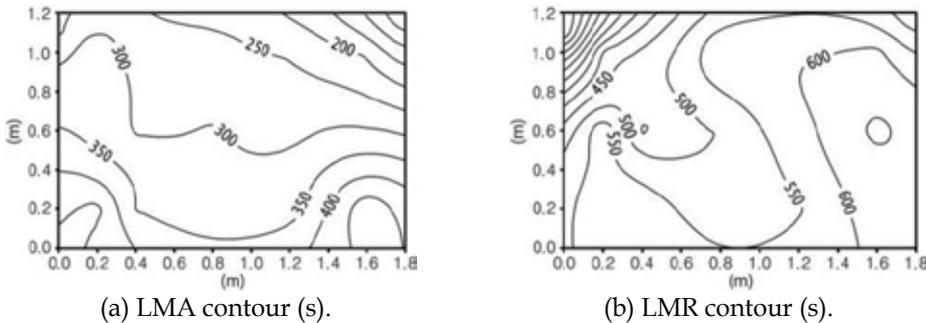


Fig. 14. LMA and LMR distributions for Case 3 (Han *et al.*, 2002).

Figure 15 shows room mean ventilation effectiveness for various air change rates. For Case 1, ventilation effectiveness decreased as the air change rate increased, but remained nearly constant for large air change rates over 20. It varies between 0.8 and 1.0, which is similar to a complete mixing condition. Note that the room ventilation effectiveness is 1 for complete mixing conditions, and 2 for perfect piston flow conditions. For Case 2, as the air change rate increased, the effectiveness increased initially and decreased slowly afterward. The effectiveness remained nearly constant for ACH over 20, similar to Case 1. However, the ventilation effectiveness in Case 3 is significantly lower compared to Cases 1 and 2, especially when the air change rate was low. This is due to the fact that the supply jet was not mixed well with the air in the chamber.

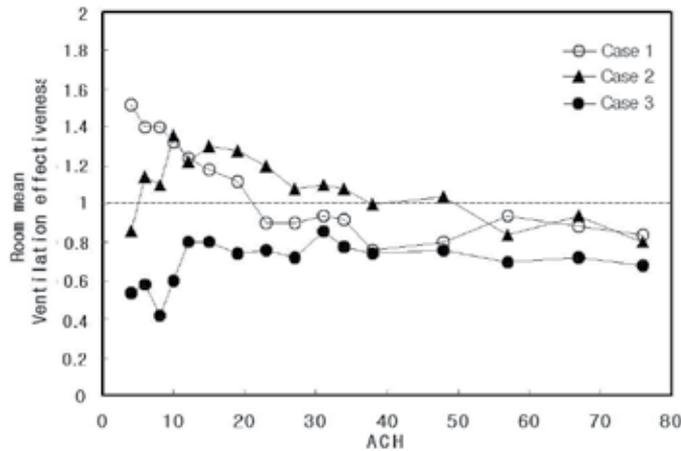


Fig. 15. Effect of air change rate on room mean ventilation effectiveness for three cases.

6.2.5 Concluding remarks

The distributions of LMA and LMR were obtained in a rectangular space with three different inlet and outlet configurations, and the corresponding airflow patterns were visualized.

1. The distributions of LMA and LMR show different characteristics, but both are closely related to the airflow pattern in the space.
2. LMA values are small adjacent to supply inlets, and large adjacent to return-air exhausts. LMR values are small adjacent to exhausts, and large adjacent to supply air inlets, as expected.
3. Compared to Cases 1 and 2, Case 3 shows poor overall room ventilation effectiveness, since the supply air jet is directed toward the exhaust outlet located at the opposite side.
4. The overall ventilation effectiveness depends not only on supply-exhaust configurations, but also on the air change rate.

The concept of local mean residual lifetime of the contaminant can be used in designing the layouts of exhausts and contaminant sources in a building such as a smoking zone, whereas concept of local mean age can be used in designing a proper distribution of fresh supply air into an occupied zone.

6.3 LMA distributions in a space with multiple inlets

6.3.1 Problem description

A space with multiple inlets is considered. It has a pentagonal shape with two inlets and a single outlet, which models a simplified livestock building. The LMAs from individual inlets are obtained by injecting a tracer gas at each inlet separately, and the combined LMA is obtained by injecting a tracer gas at both inlets simultaneously. This example is intended to verify the relation previously derived theoretically between the LMAs.

6.3.2 Experimental setup

The experimental chamber is pentagonal in shape with a height of 1.4 m and a width of 3.0 m. It is roughly a one-third scale model of a livestock building. The chamber has a length of 0.15

m, and thus can be considered 2-D. There are five openings in the model, and three are used for our experiment. Two openings on the left wall (vents *a* and *b*) are used as supply inlets, and one opening on the opposite wall (vent *d*) is used as an exhaust. Vents that are not used for the experiment have been carefully sealed. The sizes of all openings are 0.05 m x 0.15 m. A schematic diagram of the experimental setup is shown in Fig. 16.

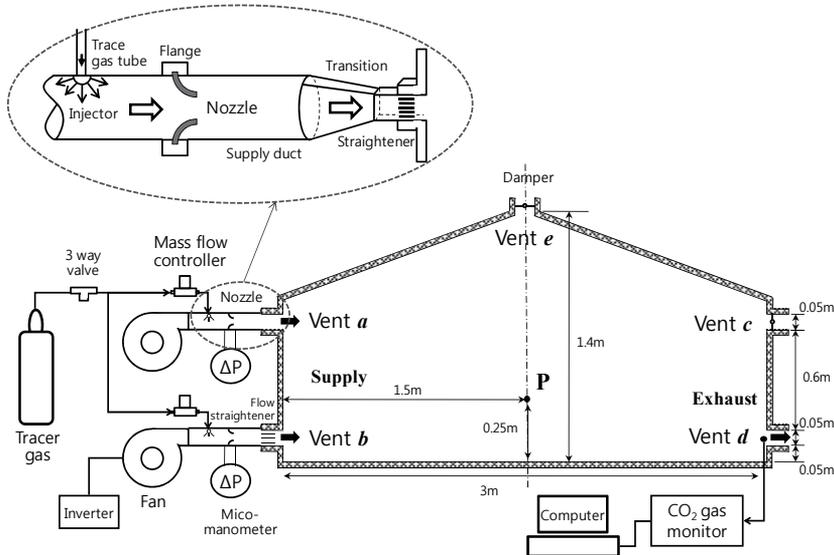


Fig. 16. Schematic diagram of experimental setup (Han *et al.*, 2011).

Carbon dioxide was used as a tracer gas. Injection ports were installed in both supply ducts upstream of the flow nozzles to ensure that the tracer gas was well mixed with incoming air streams. The amount of tracer gas was controlled by mass flow controllers (MFCs). A MFC contains a thermal mass flow meter that measures the air temperature rise across an internal heater. The range is between 0 and 10 L/min, and the error is reported to be below 1% of the measured values. A step-up method was adopted for tracer injection using a MFC. The tracer gas injection rate was held constant until a steady state condition was reached. The gas detector was an infrared single gas analyzer, and the sampling interval was 1.6 s. The range of the monitor was 20,000 ppm maximum, and the accuracy is 1% of the range.

6.3.3 Experimental procedure

Three cases of tracer injections were applied: injection at vent *a*, injection at vent *b*, and injection at vents *a* and *b* simultaneously. In order to obtain local mean age distributions in the space, tracer concentration responses were measured at 19 internal points evenly distributed in the space, including point P. The airflow rate was varied from 16.2 to 54 CMH. The airflow rates of vents *a* and *b* are maintained to be the same.

6.3.4 Results and discussion

Figure 17 shows concentration responses measured at point P and at the exhaust. The concentrations have been obtained by subtracting the background concentration, which is the average concentration measured before a tracer injection is applied. The total airflow

rate was 27 CMH. Each figure shows three injection cases: injection at vent *a* (Case *a*), at vent *b* (Case *b*), and at vents *a* and *b* (Case *c*).

The concentrations increased rapidly initially, and reached a constant steady state. The steady concentrations indicate the effective supply airflow rates contributing to the ventilation at the point by each supply inlet in a relative sense. The steady concentration in Case *b* is greater than that in Case *a*, which means the ventilation performance at point P was influenced more by the supply air from vent *b* than by the supply air from vent *a*. At the exhaust, the steady concentrations in Cases *a* and *b* are nearly the same, since the airflow rates of the two inlets are the same. We note that the non-dimensional steady concentrations at the exhaust could be determined by the relative airflow rates from the two supply inlets.

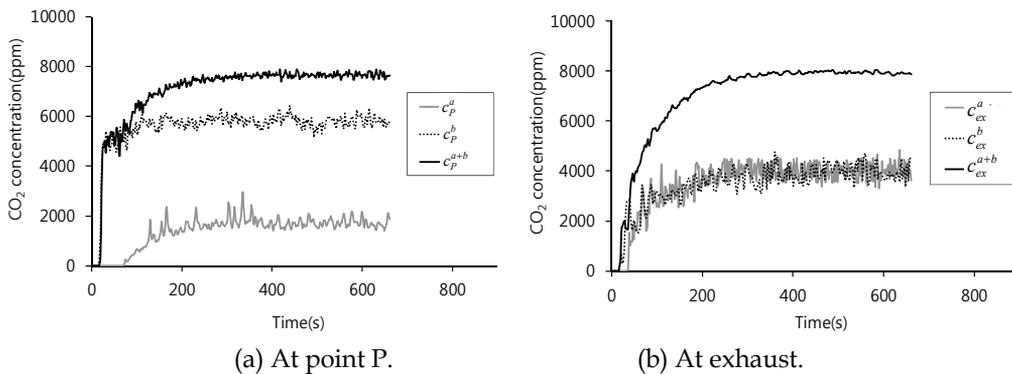


Fig. 17. Concentration responses at P and at exhaust after step-up injections at the inlets (Han *et al.*, 2011).

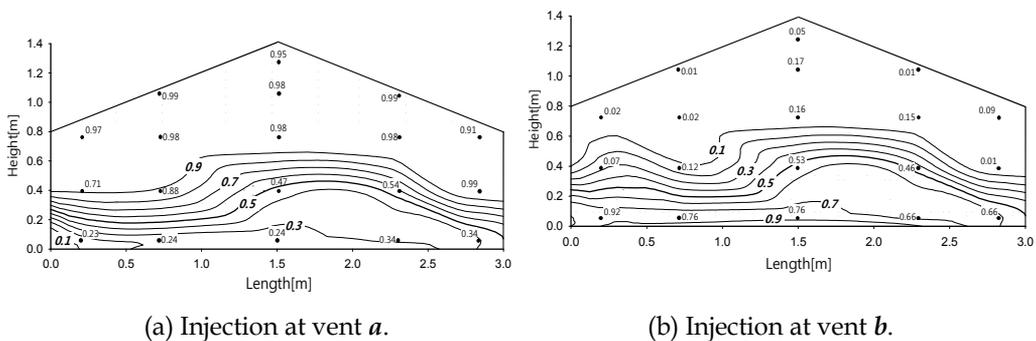


Fig. 18. Spatial distributions of steady concentrations (Han *et al.*, 2011).

The steady concentrations were obtained by taking the averages of the fluctuating concentrations for a certain period of time after reaching the steady state. Figure 18 shows the spatial distributions of the steady concentrations measured at internal points. The concentration values have been made dimensionless by dividing those by the steady concentrations obtained in Case *c*. Iso-concentration contours were drawn based on the numerical values measured at the grid points. In Fig. 18(a), non-dimensional steady concentrations by vent *a* are greater than 0.5 at an upper part of the space, and less than 0.5

at a lower part of the space. The concentration distributions by vent **b** are the opposite, as shown in Fig. 18(b). The non-dimensional steady concentrations are complimentary to each other; i.e., the sum of the concentrations is unity at any point in the space.

LMA contours from individual inlets are shown in Fig. 19. Figure 19(a) shows small LMAs in the vicinity of vent **a**, and large LMAs near vent **b** and at the upper-right corner. Figure 19(b) shows small values starting from vent **b** along the floor up to the exit on the right, and large values at three corners in the upper part of the space. By following the contour lines, we can visualize the approximate airflow pattern in the space directed toward the exit. Figure 19(c) shows the combined LMA by simultaneous injections at both supply inlets. The LMAs are small along the floor near vent **a**, and along the left part of the roof near vent **b**. The combined LMA can be calculated from the individual LMAs according to Eq. (12). The distribution is shown in Fig. 19(d) and can be compared to Fig. 19(c). The overall patterns are in good agreement.

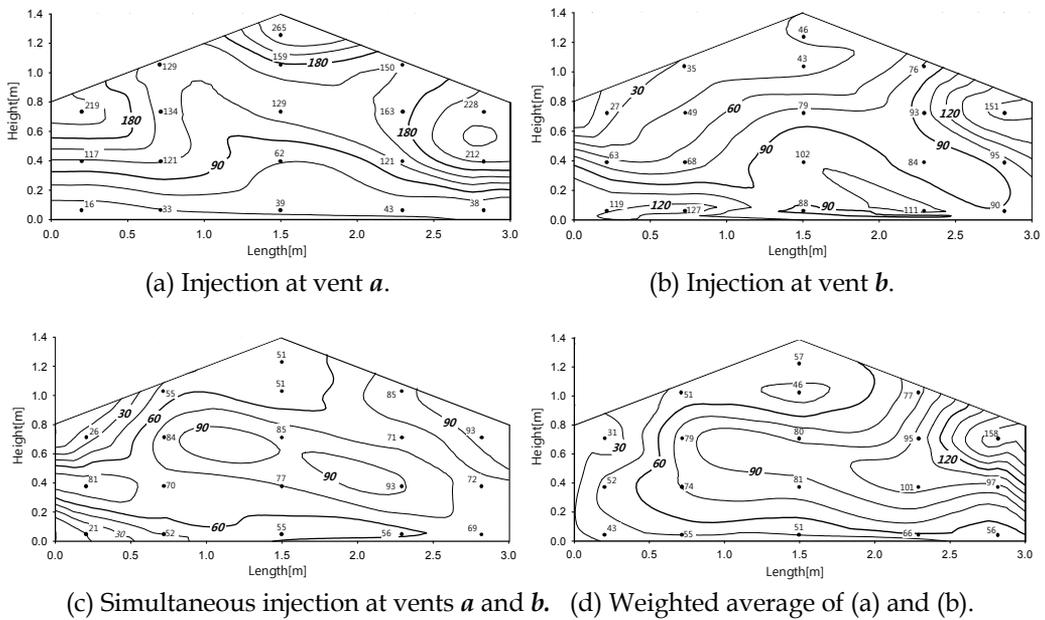


Fig. 19. Spatial distributions of local mean ages for Cases **a**, **b**, and **c**, with weighted averages (Han *et al.*, 2011).

The local mean ages at P are shown in Fig. 20(a) for various airflow rates. The airflow rate is expressed with the nominal time constant, which is the inverse of the air change rate. As the nominal time increased, both LMA_p^a and LMA_p^b increased linearly. The slope of LMA_p^a is greater than that of LMA_p^b . The combined LMA by total supply air, LMA_p^c , falls between the two sets. The figure also shows the LMA data calculated from the individual LMAs using Eq. (12).

The LMA values at the exhaust are expressed with respect to the nominal time constant in Fig. 20(b). The individual LMAs at exhaust indicate the residence time of the air supplied

through the corresponding inlets. The longer the individual LMA, the longer the corresponding supply air resides in the space. The combined LMAs are in the midst of individual LMAs, all of which vary linearly with respect to the nominal time constant. The weighted averages calculated from the individual LMAs are also shown in the figure. Notice that the weighting factors at exhaust are both 0.5 in this case, since the airflow rates are the same for both inlets. Theoretically, the combined LMA should be the same with the nominal time constant regardless of the airflow rates, which is shown by a solid line in the figure. We note that the combined LMAs appeared within the 10% error band.

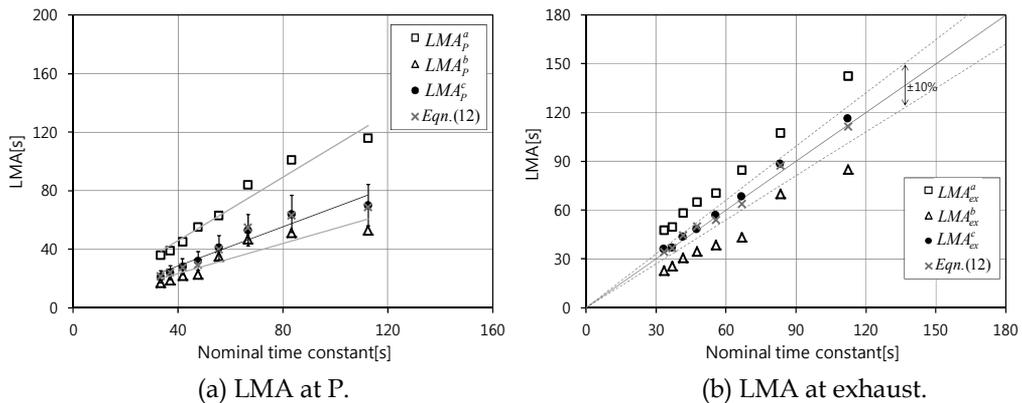


Fig. 20. Local mean ages of supply air at point P and at the exhaust as a function of nominal time constant.

6.3.5 Concluding remarks

In this example, a case of multiple inlets was considered. The relations between LMA values from individual inlets and the combined LMA were obtained experimentally in a simplified model space simulating livestock applications. The following conclusions are drawn from these results.

1. Our experimental results confirmed the theoretical relation between the individual LMAs and the combined LMA of the total supply air. The weighting factors are the steady concentrations obtained with a continuous step-up tracer injection at the corresponding supply inlets.
2. At every point in the space, the non-dimensional steady concentrations are complimentary to each other. The non-dimensional steady concentration at a point can be considered as a relative contribution factor of an individual inlet to the supply characteristics at the point.
3. The spatial distribution of an individual LMA indicates how fast the supply air from the corresponding inlet can reach the space, and it is closely related to the airflow pattern in the space.
4. These experimental procedures were verified by the fact that the overall local mean ages at the exhaust are in good agreement with the nominal time constants.

The concepts and the relations developed in this study can be applied to various applications to quantify supply characteristics of individual inlets.

7. Conclusion

The purpose of ventilation is to supply fresh air to an occupied space and to effectively remove contaminants generated within the space. Ventilation performance is determined not only by the air change rate, but also by the ventilation effectiveness.

This study dealt with ventilation effectiveness based on the concept of the age of air. Ventilation effectiveness was categorized into supply effectiveness and exhaust effectiveness. The local supply index was represented by the local mean age of supply air; similarly, the local exhaust index was represented by local mean residual-life-time of contaminant. Overall room ventilation effectiveness was expressed as one value, regardless of supply and exhaust, because the room average of the local supply index was found to be identical to that of the local exhaust index.

The age concept has been extended to a space with multiple inlet and outlet openings. Theoretical derivations were made to obtain the relations between the LMAs from individual inlets and the combined LMA of total supply air, as well as the relations between the LMRs toward individual outlets and the overall LMR of the total exhaust air. Those relations can be used to investigate the effect of each supply inlet among many inlets, and the contribution of each exhaust outlet among many outlets in a space with multiple inlets and outlets.

The tracer gas technique provided a powerful tool in our ventilation studies for measuring the ventilation effectiveness of a conditioned space as well as to evaluate the performance of diffusers and exhaust grills. The ventilation theories provided in this chapter can be applied to various applications to provide good indoor air quality and to save ventilation energy use in buildings.

8. Appendix

It can be easily proved that the local mean age distribution in a space is equivalent to the steady concentration distribution with uniformly distributed sources of unit strength in the space.

The general equation that governs the transient concentration distribution can be expressed as

$$\frac{\partial C}{\partial t} + \bar{v} \cdot \nabla C = \nabla \cdot (D \nabla C) + \dot{m} \quad (\text{A1})$$

where D is the diffusion coefficient of the contaminant in air. Consider the case of a step-down procedure with no contaminant source in the space. By integrating Eq. (A1) from zero to infinity with \dot{m} equal to zero, we obtain

$$C(\infty) - C(0) + \bar{v} \cdot \nabla \int_0^{\infty} C dt = \nabla \cdot (D \nabla \int_0^{\infty} C dt) \quad (\text{A2})$$

The steady concentration is zero; thus, Eq. (A2) can be rewritten as

$$\bar{v} \cdot \nabla \left[\int_0^{\infty} \frac{C}{C(0)} dt \right] = \nabla \cdot (D \nabla \left[\int_0^{\infty} \frac{C}{C(0)} dt \right]) + 1 \quad (\text{A3})$$

The expression in the bracket is the local mean age under a step-down procedure.

On the other hand, Eq. (A1) can be simplified for steady concentration with uniformly distributed sources. As \dot{m} is constant through the space, the equation can be simplified as

$$\bar{v} \cdot \nabla \frac{C}{\dot{m}} = \nabla \cdot (D \nabla \frac{C}{\dot{m}}) + 1 \quad (\text{A4})$$

Therefore, the steady concentration divided by the source strength equals the local mean age in the space:

$$\therefore \int_0^{\infty} \frac{C}{C(0)} dt = \frac{\bar{C}(\infty)}{\dot{m}} \quad (\text{A5})$$

9. References

- AIVC, (1990). A Guide to Air Change Efficiency, Technical Note AIVC28, Air Infiltration and Ventilation Centre, Coventry, United Kingdom.
- ASHRAE, (2009). *ASHRAE Handbook-Fundamentals*, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, USA.
- ASTM, (1993). Standard Test Methods for Determining Air Change in a Single Zone by Means of a Tracer Gas Dilution, E741-93, American Society for Testing and Materials, USA.
- Corpron, M. H. (1992). *Design and Characterization of a Ventilation Chamber*, M.S. Thesis. Department of Mechanical Engineering, University of Minnesota, Minneapolis, USA.
- Danckwerts, P. V. (1958). Local Residence-Times in Continuous-Flow Systems, *Chemical Engineering Science*, Vol. 9, pp. 78-79.
- Dufton, A. F. & Marley, W. G. (1935). Measurement of Rate of Air Change, Institution of Heating and Ventilating Engineers, Vol. 1, p. 645.
- Etheridge, D. & Sandberg, M. (1996). *Building Ventilation: Theory and Measurement*, John Wiley & Sons, New York, USA.
- Han, H. (1992). Calculation of Ventilation Effectiveness Using Steady-State Concentration Distributions and Turbulent Airflow Patterns in a Half Scale Office Building, Proc. of Int.l Symp. on Room Air Convection and Ventilation Effectiveness, pp. 187-191, Tokyo, Japan.
- Han, H.; Kuehn, T. H. & Kim, Y. I. (1999). Local Mean Age Measurements for Heating, Cooling, and Isothermal Supply Air Conditions, *ASHRAE Trans.*, Vol. 105, Pt. 2, pp. 275-282.
- Han, H.; Choi, S. H. & Lee, W. W. (2002). Distribution of Local Supply and Exhaust Effectiveness according to Room Airflow Patterns, *International Journal of Air-conditioning and Refrigeration*, Vol. 10, No. 4, pp. 177-183.
- Han, H.; Shin, C. Y.; Lee, I.B. & Kwon, K. S. (2010). Local Mean Ages of Air in a Room with Multiple Inlets, *Int. J. of Air-Conditioning and Refrigeration*, Vol. 18, No. 1, pp. 15-21.
- Han, H.; Shin, C. Y.; Lee, I.B. & Kwon, K. S. (2011). Tracer Gas Experiment for Local Mean Ages of Air from Individual Supply Inlets in a Space with Multiple Inlets, *Building and Environment*, Vol. 46, pp. 2462-2471.
- Kuehn, T. H.; Ramsey, J. W. & Threlkeld, J. L. (1998). *Thermal Environmental Engineering*, 3rd ed., Prentice Hall, London, United Kingdom.
- Liang, H. (1994). Room Air Movement and Contaminant Transport, *Ph.D. Thesis*. Department of Mechanical Engineering, University of Minnesota, Minneapolis, USA.
- Sandberg, M. (1981). What is Ventilation Efficiency, *Building and Environment*, Vol. 16, No. 2, pp. 123-135.

- Sandberg, M. & Sjöberg, M. (1983). The Use of Moments for Assessing Air Quality in Ventilated Rooms, *Building and Environment*, Vol. 18, No. 4, pp. 181-197.
- Shaw, C. Y., Zhang, J. S., Said, M. N. A., Vaculik, F. & Magee, R. J. (1992). Effect of Air Diffuser Layout on the Ventilation Conditions of a Workstation-Part II: Air Change Efficiency and Ventilation Efficiency, *ASHRAE Trans.*, Vol. 99, Pt. 2, pp. 133-143.
- Skaaret, E. (1986). Contaminant Removal Performance in Terms of Ventilation Effectiveness, *Environmental International*, Vol. 12, Issues 1-4, pp. 419-427.
- Spalding, D. B. (1958). A Note on Mean Residence-Times in Steady Flows of Arbitrary Complexity, *Chemical Engineering Science*, Vol. 9, pp. 74-77.
- Yaglou, C. P. & Witheridge W. N. (1937). Ventilation Requirements, *ASHVE Trans.*, Vol. 42, pp. 423-436.
- Xing, H., Hatton, A. & Awbi, H. B. (2001). A Study of the Air Quality in the Breathing Zone in a Room with Displacement Ventilation, *Building and Environment*, Vol. 36, pp. 809-820.

Fluid Dynamic Models Application in Risk Assessment

Peter Vidmar, Stojan Petelin and Marko Perkovič
*University of Ljubljana, Faculty of Maritime Studies and Transport
Slovenia*

1. Introduction

Risk is a common name for the probability of a hazard turning into a disaster. Vulnerability and hazard are not dangerous in and of themselves, but if they come together, they generate a risk. However, risk can be reduced and managed. If we are careful about how we treat the environment, and if we are aware of our weaknesses and vulnerabilities to existing hazards, then we can take measures to make sure that hazards do not turn into disasters.

Hazard from LNG (Liquefied Natural Gas) cargo begins in the first processing stage of natural gas liquefaction and loading the substance into LNG tankers. The transport itself over the sea is the safest part of the distribution process, as is demonstrated by the statistic of nautical accidents in the past 40 years (DNV, 2007, Perkovic et al., 2010 & Gucma, 2007).

A review of a Rand Corporation document (Murray et al.) published in 1976 indicates a high level of safety for LNG tankers. The document indicates that in the initial 16-year history (from 1959 up to 1974) there had been no significant accidents. It should be noted, though, that in 1974 the world LNG fleet included only 14 vessels; by November, 2009, there were 327 vessels, a figure expected to increase to 350 vessels sometime in 2010 (LNG Journal, 2008).

The DNV (Det Norske Veritas) counts 185 nautical accidents involving LNG tankers, all without severe consequences for the crew. The frequency of LNG tanker accidents is therefore 5.6×10^{-2} per ship year. The findings of the DNV (2007) furthermore demonstrate that the potential loss of life for the LNG crew member is 9.74×10^{-3} or less, considering the occupational fatality rate onboard gas tankers is 4.9×10^{-4} . The analysis of the northern Adriatic Sea (Petelin et al. 2009), or, precisely, the gulf of Trieste, demonstrates that nautical accidents should occur with a frequency of 1.25×10^{-2} per year, assuming the current traffic density, and increases to 2.62×10^{-2} if the ship traffic increases by 100%.

The hazard associated with LNG is mainly in its potential to cause severe fires resulting in heat radiation. If a large quantity of LNG is spilled into a pool, the cloud that is formed as it evaporates is a mixture of natural gas, water vapour, and air. Initially the cloud is heavier than air (due to its low storage temperature) and remains close to the ground. The buoyancy moves the natural gas upward at a gas temperature of around 170 K (-103°C), as experimentally demonstrated by ioMosaic (2007). The major influences on natural gas diffusion are environmental conditions. The cloud moves in the direction of the wind and the wind causes the cloud to mix with more air. If the concentration of gas in the air is between 5% and 15% it is flammable and burns if it contacts any ignition source. A concentration of gas smaller than 5% will not ignite and if the concentration is over 15% the air becomes saturated. The explosion of natural gas is not possible in open spaces because

the low velocity of flame spread, around 0.4 m/s, is not enough to produce a pressure wave (Fells et al, 1969). The burnout of gas/air mixture in open air could result only in a flash fire. Explosions could only result in enclosed spaces where gas is going to accumulate.

2. Risk definition and acceptance criteria

2.1 Individual risk

The individual risk in regard to LNG is calculated as location specific, to a person exposed outside 24 hours per day. In several countries the authorities have defined criteria which have to be met in order to assess a level of societal risk as acceptable. Criteria for some countries were analysed and discussed by Trbojević (2005). Some of these are used internationally. In most countries an individual risk of $1E10^{-3}$ per year is taken as the upper bound criteria to assess the acceptability of an activity for employers working in an industrial installation. The upper risk criterion for the public is therefore $1*10^{-4}$ per year. Also the risk at locations where vulnerable objects are situated is taken into account. Vulnerable objects are those where people are present who react in a different way to a threat posed upon them as ordinary people. This difference can be caused by differences in state of health or the possibility of evacuating the location in case of danger. The calculation of the individual risk for a specific failure event is influenced by three main parameters: the two coordinate location x,y and weather conditions (wind speed and stability). The individual risk $IR_{x,y/w}$ is therefore the function of the frequency F_{fe} of the failure event occurring in a year, the probability of the event in a particular direction (influenced by weather) and the probability of people dying due to exposure (DNV, 2008).

$$IR_{x,y/w} = F_{fe} \int_{\theta_1}^{\theta_2} [P_{\theta/w} P_{d/\theta w}] d\theta \quad (1)$$

θ is the direction of the release,

θ_1 is the lower value of θ that influences the computation point,

θ_2 is the upper value of θ that influences the computation point,

$P_{\theta/w}$ is the probability of the release occurring in the direction of the wind and

$P_{d/\theta w}$ is the probability of death considering the direction of the release and weather.

This is the contribution to the individual risk from a single event under specific weather conditions. To obtain the overall individual risk at a given point all possible events must be taken into account. However in the sense of the order of magnitude the worst events (low probability/high consequences) have the most influence on individual risk. In the calculation process, low consequence events increase the size of high risk isolines, but usually do not influence the low risk isolines, risk lower than $1*10^{-6}$ that reaches greatest distances. The strong dependence of the individual risk on weather conditions influences the total risk, calculated using the following equations:

$$IR_{point_x,y} = \sum_{Weathers} P_w \cdot IR_{x,y/w} \quad (2)$$

and

$$IR_{total_x,y} = \sum_{all_x,y} IR_{point_x,y} \quad (3)$$

where

$IR_{point_{x,y}}$ is the individual risk over all weather conditions for a specific event

P_w is the probability of specific weather and

$IR_{total_{x,y}}$ is the individual risk as a sum of all events under all weather conditions.

2.2 Societal risk

Societal risk posed by an LNG terminal facility or hazardous activity is measured by the probability that a group of persons would be exposed to a hazardous level of harm (fatality) due to all types of accidents at the facility or its hazardous activity. The societal risk is dependent on both the density of people in the vicinity of an LNG terminal and the location of the population with respect to the facility. The societal risk is generally presented in the form of a curve, expressing the relationship between the annual probability (F) of exceeding a given number of fatalities and the number N (Trbojević, 2005). In most countries the risk assessment is performed on the basis of potential fatalities to the exposed population. Different countries use slightly different criteria for risk acceptability. In the UK, the Health and Safety Executive (HSE) guidelines are to use the individual risk as the principal measure, but also to use the societal risk criteria for land use planning. The acceptability criteria levels for risks for facilities in the UK are specified by HSE (1989). Facilities are permitted only when these (published) criteria are met. In the Netherlands, however, both the individual risk criteria and the societal risk criteria have to be met when considering (in risk assessment) those events whose hazardous effects extend to such distances at which the conditional probability for lethality is higher than 1% (Bottelberghs, 2000 & Raj et al., 2009). The risk tolerability criteria for fatalities established in various countries for societal risks are summarized in Fig. 1 (Trbojević, 2005).

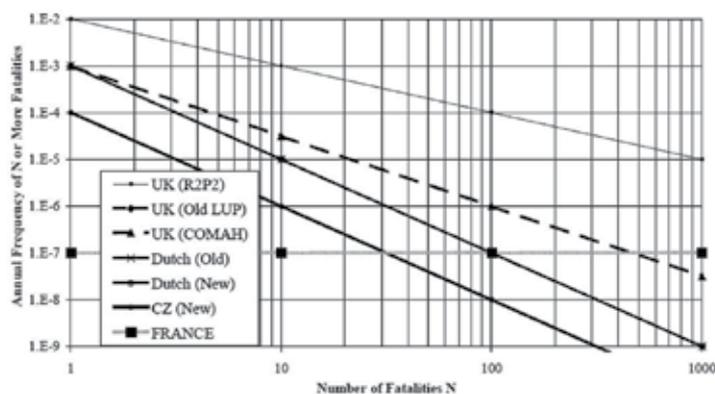


Fig. 1. Comparison of F-N Criteria (Trbojević, 2005)

The risk calculations have to determine the number of people killed by a particular event and attach this to the associated frequency of the event to form the F-N curve. Probability of death is calculated from the consequence model and so it only remains for the risk model to integrate the probability of death for each event over the specified population. N_{f/f_e} represents the number of people killed by a given event, particular considering the weather category/wind direction combination. It is calculated for the population on a grid according to Equation (4) (DNV, 2007).

$$N_{f/fe} = \iint_{area} P_{d,xy} n_{xy} dx dy \quad (4)$$

3. Consequences analyses background

Different scenarios of LNG release on water surface are possible, depending on where a rupture occurred. Accident scenarios like a collision between ships, grounding, and collision at the coast could all lead to a rupture of primary cargo containment that could spill LNG into the environment. The collision of two ships at 90°, in which the side collapsed ship is the LNG carrier is the worst scenario. The HAZID model of this is presented by (Pitblado, 2004). The scenarios of grounding and collision with the shore are less critical because the impact force is distributed over a wider area of the ship and the penetration is not as deep in the vessel's structure. During the unloading operation the transfer arm could break because of an unpredictable ship movement, even it is moored. The shutdown of the transfer operation occurs in 1 to 10 minutes depending on the emergency system used. Transfer lines with EMS (Emergency Shut Down System) could interrupt the transfer in 1 minute. During the calculation a tolerance of an additional minute is added.

The investigation conducted indicates that the consequence approach should be based on some conservative initial conditions. We found that the best idea is to assume the greatest possible rupture size on primary containment or the transfer pipe, resulting from an accident. In this way a separate investigation could be conducted. The first is the calculation or assumption of possible rupture sizes and the second is the analyses of the resulting spills dynamics. In the paper the rupture size is assumed and taken from the literature (Pitblado, 2004), but the spill dynamics are computed.

3.1 Analyses approaches

The computation of potential consequences is the first step toward understanding the severity of an accident. Several commercial computer programs apply a lumped parameter modelling technique. Empirical equations derived from physical equations consider different release scenarios, meteorological conditions, obstacle density, etc. A zone modelling technique divides the physical space into zones. Within each zone the uniform physical phenomena are computed using physical and empirical equations. With this approach several approximations and simplifications within models are assumed to reduce the complexity of formulations and to reduce the computation time. In some applications, like gas dispersion dynamics, results from empirical models could lead to underestimated results. For this reason the empirical models usually apply corrective parameters or factors to compute more conservative, corrective, results. Overestimated consequences results could lead to overestimated risk results and falsely estimate the risk level to exceed allowed limits (Fig. 1). When this is suspected, further consequences analysis should be conducted, applying more accurate methods. The case of gas dispersion is computed using a CFD (Computational Fluid Dynamic) code and analysed for improved results.

3.2 Rupture size

The estimation of the rupture size caused by an accident involving an LNG carrier is not easy, because of the variables: a complex structure, the type of accident and the location of the primary rupture. The LNG carrier has four to five barriers that would have to fail before the LNG cargo is spilled. The shield of the LNG containment allows deformation

before it ruptures. The characteristic of this material is that it remains elastic at a cargo storage temperature of -162°C . The second fact is that the reservoir is filled up to 96-97% and the rest is gas, about 800 m^3 . Because of this gas space the containment structure allows slow deformations that occur during an accident. The accident of the El Paso Paul Kayser (Bubbico et al., 2009) is a good example. The accident caused significant deformations of the ship's hull and ribbed construction yet the deformation of the cargo containment was limited to one meter and there was no release. Pitpablo (2004) suggests that the maximum possible rupture size is 250 mm in diameter when caused by grounding or collision with the shore.

Research conducted by DNV based on accident statistics and ship damage are based on the IMO (International Maritime Organisation) Guideline for Alternative Tanker Design (IMO, 1995). DNV found that an LNG tanker could resist a hull deformation of 3 meters and the reservoirs could resist 2 additional meters of deformation. Considered and analyzed ruptures on LNG vessels are only caused by accidents from the traffic point of view. Terroristic attacks and attacks with weapons are not considered. However, the interpretation of DNV (2008) and Sandia National Laboratory (Hightower et al., 2006) is widely acceptable. Major damage to LNG tankers by weapons is less probable because the vessel's structural stiffness is much greater than any building, bridge or other land structure. In case a projectile breaks through the primary reservoir, there is a high probability of immediate ignition, a local fire or even the destruction of the ship, but not the formation of a flammable gas cloud and subsequent flash fire that is a danger to the neighbour population. On this basis DNV suggests the consideration of the largest rupture size of 1500 mm for a terroristic attack.

Therefore the further scenarios analysis would consider the following LNG reservoir rupture size:

- 250 mm Maximum possible rupture size caused by grounding
- 750 mm Maximum possible rupture size caused by collision
- 1500 mm Maximum possible rupture size caused by terrorist attack
- 7,000 m³/h Maximum possible leakage rate for 10 min
- 10,000 m³/h Maximum possible leakage rate caused by sabotage for 60 min

3.3 Atmospheric dispersion of a vapour cloud

As mentioned above, the spreading of the LNG pool is a dynamic process that is alimented by the spilling from the tank slowed down by evaporation or even vapours burning. Research conducted by Hissong (2007) and ioMosaic (Melhem, 2007) presented in Table 1 explain that the evaporation rate could vary depending on water conditions.

	Evaporation rate Max value kg/m ² s	Evaporation rate 20 sec. avg. kg/m ² s
Water	0.303	0.147
Water	0.245	0.191
Water	0.230	0.196
Ice	0.513	0.191
Ice	0.333	0.171
Brine	0.254	0.186

Table 1. Evaporation rates depending on water conditions

Leak rates from a tank rupture were calculated; initial gas, "flashing" and an aerosol are formed. An aerosol is a cloud of tiny liquids droplets or fine solid droplets suspended in air. Calculating the droplet evaporation along the cloud trajectory, the overall vapour generation rate is obtained. In this section the dispersion model of the vapour cloud is described. The model considers different types of release: instantaneous, steady continuous and transient for dense (active) and lean (passive) gases.

Our case considers a ground level instantaneous release from a tank rupture. That is the most catastrophic scenario, one which can lead to fast vapour cloud formation near ground level and with a high gas concentration. Because of a very fast transient and changes of variables, it is difficult to predict the course of events, especially close to the source of dispersion. Once the cloud, modelled as a cylinder, is formed, it begins to slump under the effect of gravity. The velocity of the edge of the cloud can be described as (Safer SysteM, 1996):

$$\frac{dR}{dt} = k_1 \left[\left(\frac{\rho_{\text{CLOUD}} - \rho_A}{\rho_{\text{CLOUD}}} \right) \cdot g \cdot h \right]^{1/2} \quad (5)$$

k_1 is a slumping constant that depends on the characteristic of released gases and weather conditions. The most important factor is wind speed. It is important to note that the dispersion model does not assume turbulent flow. Dispersion of the cloud is therefore a suitable element to be computed with the CFD model. The rate at which liquid fuel evaporates when burning is a function of the liquid temperature and the concentration of fuel vapour above the pool surface. According to the Clausius-Clapeyron relation, the volume fraction of the fuel vapour above the surface is a function of the liquid boiling temperature;

$$V_f = \exp \left[- \frac{h_v W_f}{R_f} \left(\frac{1}{T_s} - \frac{1}{T_b} \right) \right] \quad (6)$$

where h_v is the heat of vaporization, W_f is the molecular weight, T_s is the surface temperature, and T_b is the boiling temperature of the fuel.

The cloud transport due to wind is modelled as (Vidmar et al., 2003 & Safer System, 1996):

$$\frac{dx}{dt} = U_{\text{CLOUD}}(z = 0.4 h_{\text{CLOUD}}), \quad (7)$$

where x is the downwind distance variable and U_{CLOUD} is the cloud speed, assumed to be equal to the wind speed. The value 0.4 times the cloud height is assumed to be a reference height or the centre of gravity of wind force in the direction of cloud movement.

In most dense gas releases, it is expected that there will be a central core region of uniform concentration along with edges at which the concentration decreases. It is expected that close to the source, the edges will be sharp and as the cloud disperses downwind the edges will become less steep. The concentration field is calculated considering a Gaussian distribution. That means that the isopleth limits, or the edge of observed concentration limits, take on a typical Gaussian distribution. Gaussian correlations for the atmospheric cloud dispersion are not proper for any initial conditions and release type, but are used as a conservative model (McGrattan, 1997).

4. Scenarios and results

The first simulations conducted are used to compare results obtained from the lumped model (fast computation) and CFD model like FDS (Fire Dynamics Simulator). The scenarios include a simple geometry with a domain size 500m x 200m x 100 m in x , y and z directions. The diameter of the spilled pool is related with the diameter of the hole on a reservoir shield. Possible or applied scenarios are refined in a previous section, the resulting pool sizes are further obtained from functions explained in (ioMosaic, 2007). The following scenarios assume the hole at the bottom of the reservoir and 9000 kg of LNG content is spilled in 10 seconds and forms the pool. The evaporation rate of natural gas from the pool surface is calculated by the lumped model and assumed to be 0.16 kg/m²s for the CFD model in a stable sea and stable weather conditions.

1. D= 50 m Calculated pool diameter for the hole diameter 250 mm
2. D= 200 m Calculated pool diameter for the hole diameter 750 mm
3. D= 450 m Calculated pool diameter for the hole diameter 1500 mm
4. D= 196 m Calculated pool diameter for leaking 7,000 m³/h for 10 min
5. D = 400 m Calculated pool diameter for leaking 10,000 m³/h for 60 min

Because of the very similar pool sizes scenarios 2 and 4 are calculated once as are scenarios 3 and 5.

4.1 Pool size related to rupture size

The course of events after an accident could follow different directions before the course of the accident is definitive. Depending on the accident dynamics, and the time and place of occurrence it could lead to different levels of risk. An example is the immediate ignition of a vapour cloud. If this occurs in a populated area, it represents a high risk event, but if it occurs outside the populated area it would be a low risk event. The reason is in the consequences of the event. The second event mentioned would be high risk if the ignition of the cloud would be delayed and the cloud would be transported downwind into a populated area.

The first scenario assumes the evaporation of natural gas from the LNG pool on a water surface of an area of 400 m². The evaporating gas is moved downwind at 2 m/s; the surrounding temperature is 20°C. The model does not include complex geometry, buildings, and relief that would slightly change the flow dynamics of the vapour cloud.

Material ID	Methane
Specific heat	2.2 kJ/kg K
Conductivity	0.03281 W/m K
Reference temperature	-162 °C
Absorption coefficient	40 1/m
Boiling temperature	-162 °C
Fuel rate	1.0
Density (gas)	1.8 kg/m ³
Gas rate	1.0
Mass flux	max: 0.16 kg/m ² s
Gas mass fraction	1.0

Table 2. Mandatory input data for Scenario 1:

The table above explains how the cloud formation is modelled with FDS. There is no phase transition modelled from liquid to gas. Instead of this the gas release with a specified mass flux dynamics or evaporation rate (Table 1) is applied.

The first scenario treats a minor leak and consequently a smaller spilling pool. The scenario is primary used for model testing and for the selection of adequate boundary and initial conditions. It is also used for the comparison of results with further models where the differences of consequences could be compared for different releases.

4.2 Results

The comparison of results from the lumped model and CFD model explain the adequacy of the analysis. Fig. 2 shows the results of a Methane concentration field. The bold polyline represent the concentration 3000 ppm and the shape of the cloud downwind. The overlaid curves (continuous and dashed) represent the cloud height, calculated by the lumped model described above. Instead of the concentration field the dotted red line represents the cloud height obtained from averaging local heights of the Methane cloud.

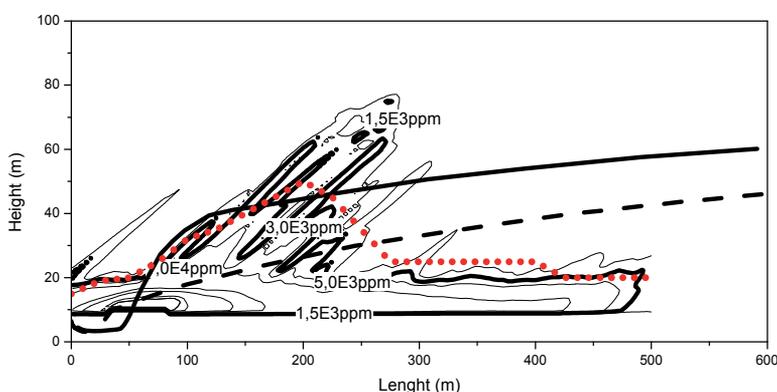


Fig. 2. Methane cloud height 8 min after release, concentration CH₄ [ppm]

The following in Fig. 3 is the comparison of evaporated gas concentration between two models. 5% concentration is selected because it represents the LFL (lower flammable limit) for methane. The time scale, presented on a graph requires a second axis only for the lumped model results. However, the comparison of CFD fields and the cloud heights of the lumped model show close results and time dynamics.

The obtained results are a good guideline for the setup of further model scenarios. One finding is that the FDS program does not allow for a satisfying definition of the pool evaporation process; therefore the dispersion is overestimated at a later time. The pool evaporation process does not have equal dynamics in both models. The solution found is the combination, where the pool evaporation process is modelled by a lumped model and the cloud dispersion by CFD. This technique is also applied in further scenarios.

A small release of LNG on a water surface reaches the thermal equilibrium fairly rapidly and the evaporated gas quickly dilutes to concentrations below the LFL (low flammable limit) that is 5% in the air. The area at risk of fire is about 150 to 200 m downwind.

The model does not include complex geometry, buildings, or relief that would slightly change the flow dynamics of the vapour cloud. In real situations local gas accumulations could occur in bounded spaces where the gas could exceed 5% concentrations. The literature

(Fells et al., 1969, ioMosaic, 2007 & Hissong, 2007) does not devote special attention to this phenomenon because it would require overly complex stochastic machinations.

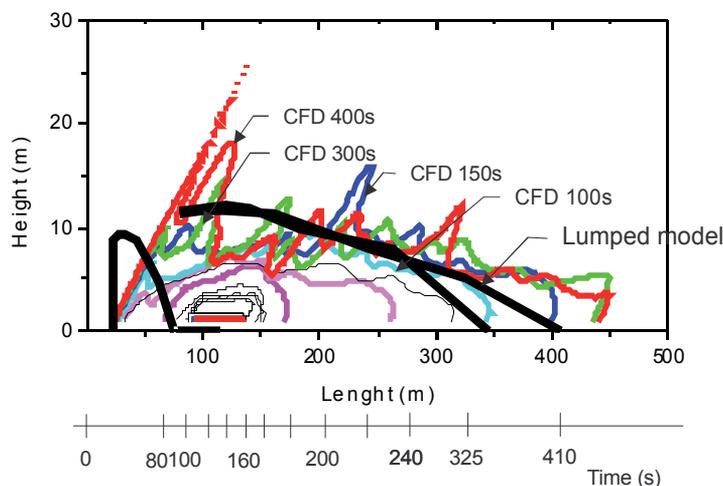


Fig. 3. Methane concentration of 5% computed with a lumped model and CFD model for Scenario 1

5. Risk assessment of an LNG terminal

The assessment of the installation presented is focused on the accidental events of the LNGC approaching territorial waters through the northern Adriatic separation zone, entering the port and the unloading operation in the sense of individual risk and social risk. Therefore the appropriate approach is to discover and analyse processes and installation elements that could lead to undesired consequences. There are several widely approved methods, like HAZID (Hazard Identification) and HAZOP (Hazard and Operability), described in Macdonald (2004) and in the Guidelines for the management of change for process safety (2008) that are used for structured and systematic examination of a planned or existing process or operation in order to identify and evaluate problems that may represent risks to personnel or equipment. The HAZOP study and report are not presented in the paper, but the evaluated accident events frequencies are used for the evaluation of individual and social risk.

5.1 Scenario based models

The population is exposed to risk depending on the reliability of the installation and its components. This is done according to recommendation of the Manual Risk Calculations (2008), Guidelines for the management and change for process safety (2008), DNV reports (2008) and ioMosaic reports (2007). In general the following scenarios are modelled for vessels:

- catastrophic rupture with instantaneous release of full inventory;
- a large leak resulting in a continuous release with a fixed duration (10 minutes);
- a small leak (10 mm)

The failure frequencies for the scenarios are derived from the Dutch guidelines (2008). The study leads to several accident scenarios with elevated risk and takes into account accidents

that occur due to technical failures. The catastrophic rupture with full lost of inventory is left out of possible technical failure scenarios. This is an important assumption because technical failures could be identified and evaluated by the above methods. The conducted QRA (Quantitative Risk Assessment) takes into account scenarios listed in Table 3.

Scenario	Diameter	Volume (m ³) or release rate (kg/s) - Release time	Failure frequency (per year)
P1- Loading arm - rupture with ESD (Emergency Shut Down)	16"	700 kg/s - 120 sec	7.5E-05
P2- Loading arm - rupture without ESD	16"	705 kg/s - 1800 sec	7.5E-08
P3- Loading arm - 10% leak (max. 50mm)	1.6"	16 kg/s	7.5E-04
P4- Vapour return arm - rupture with ESD	16"	9 kg/s	2.5E-08
P5- Vapour return arm - 10% leak (max. 50mm)	1.6"	0.1 kg/s	2.5E-04
P6- Jetty drum -catastrophic rupture		30 m ³	4.8E-07
P7- Jetty drum - release in 10 min		30 m ³	4.8E-07
P8- Jetty drum -10 mm leak	10 mm	30m ³	9.6E-06

Table 3. Scenarios modelled in QRA (DNV, 2008)

The worst consequences would result from the largest releases. The rupture of a loading arm during the unloading operation could lead to large spills and influence the risk level. The difference between the first and second scenario is the installation of the emergency shutdown valve on the transfer line. This valve restricts the release time for ruptures of the arms. A value of 120 seconds as maximum release duration is chosen as it is widely accepted to be a conservative value for pipe isolation. Also the scenario that the ESD valves can fail has been taken into account (probability of 0.001 is taken into account). In that case a maximum release time of 30 min is applied.

The defined scenarios are assessed for their impact on external safety by using criteria for individual risk and for societal risk. Fig. 4 shows curves of the individual risk for scenarios P1 to P8 from Table 3, which includes unloading operations on moored tankers. The criterion for individual risk is $1E-6$ / avg year for population, a commonly accepted risk throughout the world (Trbojević, 2005). The area with this risk is contained within the installation boundary. The personnel of the installation and of the terminal are the only ones exposed to risk higher than $1E-6$ /avg year. However, the employed personnel are adequately educated and trained for possible threats and have defined procedures to overcome possible emergencies.

Scenarios with a major influence on risk are larger spills (scenarios P1 and P2) like rupture of a loading arm at lower wind speeds. The reason is that the vapour cloud remains compact for a longer time and moves downwind, possibly into the direction of a populated area.

The populated neighbourhoods located south and north of the port are exposed to a much lower risk than $1E-6$ /avg year. In this case the LNG unloading operation, considering the eight most likely accident events, is safe for people living or passing through the LNG terminal neighbourhood. Personnel working on the LNG terminal on pier 2 (Fig. 4) and on neighbouring terminals on pier 1 are covered by the risk zone $1E-6$ / avg year.



Fig. 4. Individual risk for LNG terminal

The societal risk is further presented in Fig. 5 with an F-N curve. The maximum expected number of fatalities is 20 with the probability of 4×10^{-7} that is found at the end of F-N curve.

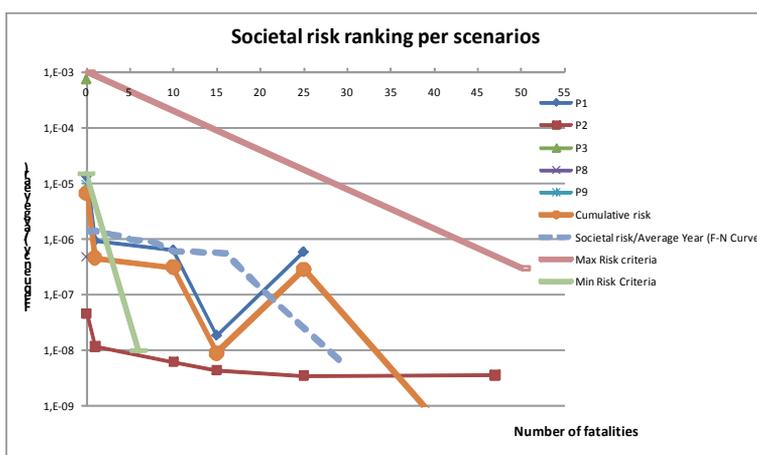


Fig. 5. Social risk of LNG terminal

The entire societal risk curve (F-N) is located within the ALARP (As Low as Reasonably Practicable) area delimited by the Max and Min Risk Criteria straight lines. This could be taken as an argument that the LNG terminal operations are not harmful to neighbouring populations. Employed personnel on the jetty and pier 2 are the only people who contribute to societal risk, excluding them, the F-N curve would be lower than the Min Risk Criteria. The summary of risk assessment shows that the $1E-6$ / avg year individual risk contours do not reach any vulnerable location outside the port area and therefore the level of societal risk is not unacceptable, since the presence of people in the direct vicinity of the jetty is limited and the distance to populated areas like the town centre is too far to pose a societal risk. The most important factor in risk assessment is determining the level of assessment, which then

determines which scenarios and types of accident are investigated. Here we dealt only with technical failures during the unloading operation with major predictable consequences. Modern industrial installations are no longer problematic regarding safety aspects mainly because of the application of state of the art standards for construction, materials, and operations and so on. During last decade a lot of attention has been focused on unpredictable risks that are human and societal related; that is to say, terrorism, which has not yet been well defined. A lot of work is done by developing countries to overcome this risk and to include it in different standards and procedures. But up to now no one standard regarding construction or materials of industrial installations has been changed because of the risk posed by terrorism, though several new operational procedures have been applied in practice.

6. Conclusion on LNG risk assessment

The investigation conducted is focused on the discovery of an appropriate approach to risk assessment for an LNG delivery terminal located close to a populated area. The assessment should focus mainly on the identification of accident scenarios which results in individual risk higher than $1E-3$ / avg year for the neighbouring population. Because the risk calculations are most influenced by large accident consequences, these accidents should be analysed in detail to avoid over or under estimation of fast computing consequence (dispersion) models, which are usually used in risk assessment software. The comparison of a lumped model and CFD model is therefore conducted and the differences are analysed and discussed. After obtaining satisfying results the individual and societal risk is computed for the specified number of accident scenarios. The lumped model approach for consequence calculation is the best choice in the initial phase of risk analyses when several simulations need to be done. Additional detailed analyses are required when the risk is near the limit of that ALARP region. The numerical simulation of a problematic scenario could explain whether the consequences and therefore the risk is really too high, or whether it should fall within an acceptable area. Usually this is not a satisfying solution, so regarding this some modifications of the project are commonly proposed.

A possibility for further development is on a risk model based on CFD that relies on several overlaying CFD scenarios in combination with risk criteria functions and obtained individual risk levels. Depending on the accuracy of analysed scenarios the risk range could be much more realistic than in lumped models where the range is intentionally more conservative. In any case, the method would be very time consuming yet could be quite useful in the last stage of risk assessment.

7. Fluid dynamic models for road tunnels risk assessment

The definition of the deterministic approach in safety analyses arises from the need to understand the conditions that emerge during a fire accident in a road tunnel. The key factor of the tunnel operations during the fire is the ventilation, which during the initial phases of the fire have a strong impact on the evacuation of people and later on the access of the intervention units in the tunnel. The text presents the use of the CFD model in the tunnel safety assessment process. The set-up of the initial and boundary conditions and the requirement for grid density found from validation tests of an FDS (Fire Dynamics Simulator) is used to prepare different kinds of fire scenarios in different ventilation conditions; natural, semi transverse, transverse and longitudinal ventilation. The observed

variables, soot density and temperature, are presented in minutes time steps through the entire tunnel length. Comparing the obtained data in a table allows the analyses of the ventilation conditions for different heat releases from fires. The second step is to add additional criteria of human behaviour inside the tunnel (evacuation) and human resistance to the elevated gas concentrations and temperature (Haack, 1998 & 2002).

7.1 Methodological approach on tunnel safety

In order to identify the interactive and uniting relationships in a system, analysis is necessary to replace the apparent structure of individual statements on the components of a system and their relationships with their underlying common logical structure (system analysis). For example, if we are dealing with a system which we call "a chemical process plant", we get at its various components successively, by means of deductive analysis: the buildings, the operators, the storage tanks, the control systems, the operating procedures, etc.. Each such component is thrown into the modelling reality by a distinct act of noticing, and is steadily held together with those components already segregated. The aim of system analysis is to investigate the system's behaviour (i.e. the succession of its states over time) on the basis of its components' changes with time. The results of system analysis can be expressed in qualitative and quantitative terms (statements resulting from "qualitative analysis" and numbers resulting from "quantitative analysis").

The deterministic approach breathes into the analysis of the greater part of physical events like fire source characteristic and its dynamics, the operation of the ventilation system and other conditions as well as their reciprocal interactions. The approach leads also to the definition of the technical system "safety efficiency" in the range of possibilities that exist in a real world and are functionally descriptive. When the approach is used in practice, we should define a number of "safety categories" base on events probability and consequences for the individual risk. The example is presented in a Table 4.

(a) Likelihood categories Severity category (frequencies)	Qualitative definition	Underlying quantitative definition (times per year)
A	Probability once in a year	0.3 – 3
B	Possible but not likely	0.03 – 0.3
C	Unlikely	0.003 – 0.03
D	Very unlikely	0.0003 – 0.003
E	Remote	0.00003 – 0.0003

(b) Consequence categories Severity category (consequences)	Qualitative definition	Underlying semi-quantitative definition
1	Catastrophic	Multiple fatalities
2	Major	Single fatality, multiple injuries
3	Very serious	Permanently disable injuries
4	Serious	Serious injury, full recovery
5	Minor	Lost time injury, short absence from work

Table 4. Deterministic safety analysis – supposed safety categories (Kirchsteiger, 1999)

Note that, in these schemes, a quantitative definition is often given in addition to the qualitative definition, mainly to ensure consistency in the course of the analysis and provide benchmarks ("semi-quantitative analysis"). In schemes of this type, the assessment team, usually comprising members of line management, safety engineers and operations personnel, will first identify all hazards, using HAZOP or similar approaches, and then assigns a severity category to each of these, for both likelihood and consequences (PIARC, 2003 & Brussaard, 2004).

Following the assumptions of (Kirchsteiger, 1999), a "risk matrix" would then be defined as a 5 x 5 matrix with each side corresponding to one severity category.

"Likelihood" "Severity category "	Consequences "Severity category"				
	5	4	3	2	1
A					
B					
C					
D					
E					

Table 5. Deterministic safety analysis – example of risk matrix (Kirchsteiger, 1999)

Different shading in a table indicates different risk levels. Hazards with high assessments, such as A1, B1 and A2 in the black squares, are thought of as being very severe and requiring immediate action to reduce. Hazards with low assessments, such as E5, E4 and D5 in the white squares, are considered to require no further action. Hazards between these two (grey squares) are considered worthy of some improvement if a cost-effective solution can be found.

7.2 Computer models and simulations

Deterministic models that would consider all physical parameters are almost unfeasible in practice and if feasible would require very complex and time consuming computations. The application of deterministic analyses results in practice is conditioned by the simplification of some physical phenomena (like turbulence) (Gasser et al., 2002 & Floyd et al., 2001).

The fluid flow is modelled by solving the basic conservation equations. Those are conservation of mass (8), conservation of mixture fraction (9), conservation of momentum (10) and conservation of energy (11) using a form for low Mach number (McGrattan, 2001 & Fletcher, 1991). The approximation involves the filtering out of acoustic waves.

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0 \quad (8)$$

$$\frac{\partial \rho}{\partial t} (\rho Z) + \nabla \cdot \rho Z \mathbf{u} = \nabla \cdot \rho D \nabla Z \quad (9)$$

$$\rho \left(\frac{\partial \mathbf{u}}{\partial t} + \frac{1}{2} \nabla |\mathbf{u}|^2 - \mathbf{u} \times \boldsymbol{\omega} \right) + \nabla \tilde{p} = (\rho - \rho_\infty) \mathbf{g} + \nabla \cdot \boldsymbol{\tau} \quad (10)$$

$$\rho c_p \left(\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T \right) = \dot{q}_c'' - \nabla \cdot \mathbf{q}_R + \nabla \cdot k \nabla T \quad (11)$$

Where ρ is a density, \mathbf{u} is a velocity vector, Z is the mixture fraction, T the temperature and D is a molecular diffusivity. \tilde{p} is the perturbation pressure caused by pressure differences, τ the viscosity stress tensor and k the thermal conductivity. \dot{q}_c'' and $\nabla \mathbf{q}_R$ are the source terms of chemical reaction and radiation, respectively. The radiation term has a negative sign because it represents a heat sink.

The effect of the flow field turbulence is modelled using LES (Large Eddy Simulation), in which the large scale eddies are computed directly and the sub-grid scale dissipative processes are modelled (Sagaut, 2002). The unknown sub-grid stress tensor τ is modelled by Smagorinsky model (Lesieur, 1997).

Further the combustion model is based on the assumption that the combustion is mixing-controlled. This implies that all species of interest can be described in terms of the mixture fraction Z . Heat from the reaction of fuel and oxygen is released along an infinitely thin sheet where Z takes on its stoichiometric value as determined by the solution of the transport equation for Z . The state relations are calculated for a stoichiometric reaction of C_7H_{16} (Oil), which is proposed by (McGrattan, 2001, Heskestad, 1995 & Mingchung, 1999) and called a Crude oil reaction.

8. Tunnel fire analysis

The idea is based on the creation of a deterministic risk matrix as it is showed in the Table 5. The safety category is represented by the power of the fire and the type of ventilation at different strengths and on other side the consequences are evaluated in the time during the progress of the fire. The risk criteria are defined as a relation between the hot smoke layer height, the distance from the fire position and the evacuation time of the users. In case the speed of the smoke is higher than the speed of the evacuation and in case the height of the hot layer is higher than the speed of the evacuation, the risk is high.

8.1 Tunnel fire scenario

All together 12 tunnel fire scenarios are presented. Three levels of fire force are simulated, each with four different types of ventilation. The span of the fire force is between 20 MW, 50 MW and 100 MW whilst the ventilation is sorted from the less to the more effective: 1 - natural, 2 - longitudinal, 3 - semi transverse and 4 - transverse or improved transverse ventilation.

The whole section of the simulated tunnel is 650 m long, the other dimensions are width 10 m and height 8 m or 6 m when the roof is lowered. Though the dimensions and shape of the tunnel tube partly differentiate among them that does not influence what happens during the fire. That is why ordinary skeleton measurements are chosen. The geometry of the tunnel model, the type of ventilation and the location of the fire are shown on Fig. 6.

The fire is placed on a distance of 300 m in all the models, it differs only in the size of the burning area. The focus point is defined as the heat source to which the combusted model calculates the mass transfer on the base of the accorded combusting reaction and the oxygen consumption. The ignition point is shown in Fig. 6. When we define the base igniting temperature, heat conductivity, calorific value, etc. (depending on the models demands) it is treated in the model as combustible substance and it cooperates with the generation of heat

in the combusting model. In case of the described scenarios, the base is relatively small or of small volume, that is why the heat contribution of the burning base is only a few percent of the defined freed heat of the boundary condition.

8.2 Initial, boundary conditions and discretization

The definition of the initial and boundary conditions is a peculiarity of each model. Four elementary ways of ventilation are discussed: natural, longitudinal, semi transverse and transverse. In definition of the geometry of the tunnel, the natural and longitudinal ventilation are discussed together and the semi transverse and transverse ones also in the same way (Woodburn, 1996). The comparison is shown on the lower picture:

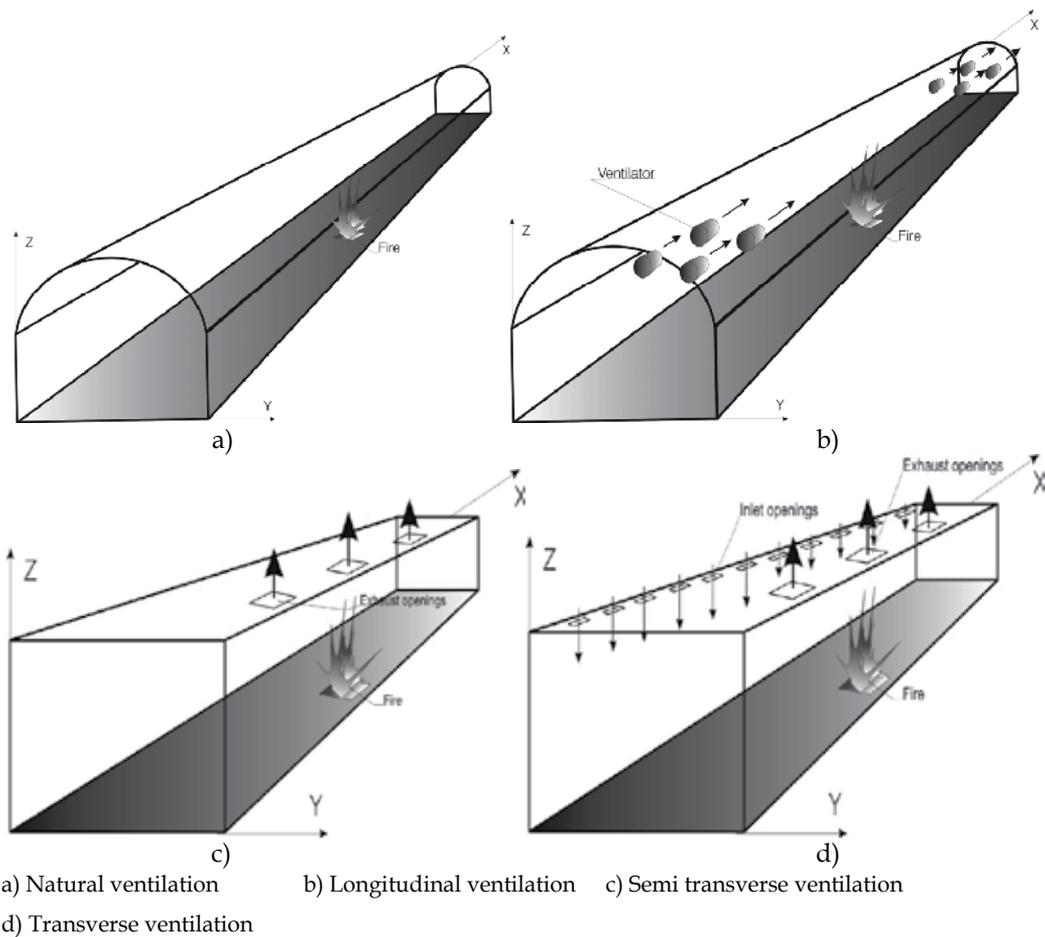


Fig. 6. Model geometry in different ways of ventilation

It is clear that the tunnel models with natural and longitudinal ventilation take the whole section of the tunnel, however the tunnel models with semi transverse and transverse ventilation consider only the light section of the tunnel (without the ventilation drains). The suction flaps are defined with the speed margin condition on the limit of the calculating domain (Cheng et al., 2001).

8.3 Parameters and approach to the result analysis

The simulation results are presented on levels of fire force and types of tunnel ventilation. The consequences of the distance of the smoke and the temperature are qualitatively evaluated from the current and temperature field. With this, it must be noted that mistakes are possible in calculating the average value in different time and space steps, which are limited with the unified way of average calculating. With that, it is true that the risk of exposure to smoke is that the participant is exposed in the moment when the smoke reaches him. The most risky examples are the ones when the participant does not start with the immediate self-rescue procedure after the start of the fire and the second when the spreading speed of the smoke is higher than the self-rescue procedure speed of the participants in the tunnel. The other risk criterion is high temperature that usually has a lower contribution to the risk than smoke. In most cases this depends on the way of ventilation.

The limit value of the concentration of smoke particles (PM10 heavy particles with the diameter up to 10 μm) is 1000 mg/m^3 (Vidmar et al., 2003) and the limit temperature is 50°C. Though the smoke particles are less problematic from a poisonous point of view, than other combustible products (CO_2 -carbon dioxide, CO-carbon monoxide, HCN-hydrogen cyanide, HCl-hydrogen chloride, etc.) their relation to the concentration is conditional and often very similar. From different experiments in the Memorial Tunnel (1996) it can be found for example concentrations of smoke particles and CO in relation around 10:1. A similar relation can be also found on toxic levels of these products. LC_{50} (Lethal Concentration 50 %) for soot particles is 30 g/m^3 in a 30 min exposure or 1-3 $\text{g}/\text{min}/\text{m}^3$ LC_{50} , for CO is 2000-3500 ppm, which is 2300-4000 mg/m^3 in a 30-60 min exposure. The limit temperature values of human endurance are according to Gann (1994) 100°C for 30 min and 75°C for 60 min of exposure. Because this information is true for an adult man it is the most optimal. But within the same research there are difficulties in breathing already at 65°C of air temperature. Taking this into account there are two values that are used in the result analysis. The chosen limit concentration of smoke particles is 1000 mg/m^3 and the limit temperature is 50°C.

The risk or consequences are divided in five categories that are shown in the Table 6, these are:

1. LR – low risk: smaller injury
2. MR – medium risk: serious injury with full recovery
3. SR – serious risk: permanent injury
4. VHR – very high risk: low casualty number (1-3), numerous injured
5. EHR – extremely high risk: numerous casualties

In the result analysis each category matches a logical inscription and it conditions with the time from the start of the simulation, any distance from the fire area, fire force, way of ventilation, limit value of the concentration of smoke particles and the limit temperature. Then follow the conditional clauses of each category:

LR: $\text{ASD} < 500$.

MR: $\text{ASDL} > 500 \wedge \text{SLH} > \text{ASLH}$

SR: $\text{ASD} > 500$.

VHR: $\text{ASDL} > 500 \wedge \text{SLH} < \text{ASLH}$

EHR: $((\text{TH} \vee \text{VHT}) \wedge \text{AT} > 50) \vee \text{ATL} > 50$.

Where the abbreviations mean:

ASD - Average smoke density value in profile [mg/m^3]

- ASDL - Average smoke density value in layer [mg/m³]
- SLH - Smoke layer height [m]
- ASLH - Allowed smoke layer height [m]
- AT - Average temperature in profile [°C]
- ATL - Average temperature in layer [°C]
- TLH - Temperature layer height [m]
- ATLH - Allowed temperature layer height [m]

8.4 Evacuation model

The easier discussing of results is enforced with the understanding of people behaviour during a fire in the tunnel which after the spotted fire begins with the self-rescue procedure. This is a withdrawal from the tunnel or to the first transverse passage in two tube tunnel scenarios. The movement of the people in similar conditions is very unpredictable, some become immediately aware of the danger and begin with the self-rescue procedure others do not perceive the danger in time and start with the self-rescue procedure too late. On self-rescue there is a simplified model of people movement in the tunnel. The model takes into consideration the elementary movement parameters as: start of the self-rescue, walking speed, tunnel length and logical curiosity that in the initial location north or south arranges the movement direction north or south. With this the possibility of a tunnel user approaching the fire during the self-rescue procedure is excluded in the model. With a program the self-rescue procedure is defined with the following conditional note:

```

IF (XSt-1 = 'S' ∧ XL0 - vH · (t - tSR) ≤ 0) then 'OUT'
else
    IF (XSt-1 = 'S') then XL0 - vH · (t - tSR)
    else
        IF (XSt-1 = 'J' ∧ XL0 + vH · (t - tSR) ≥ Lp) then 'OUT'
        else
            XL0 + vH · (t - tSR)
        END IF
    END IF
END IF
    
```

IF (XL_t = 'OUT') then 'OUT'
else
ABS(X_{poz} - XL_t)
END IF

Users' location N-North or
S-South from the fire

60		120		180		240	
Oddaljenost N→S	Oddaljenost od požara	Oddaljenost N→S	Oddaljenost od požara	Oddaljeno st. N→S	Oddaljenost od požara	Oddaljeno st. N→S	Oddaljenost od požara
Evaquation:							
Initial location							
380.00	28.00 S	515.00	163.00 S	605.00	253.00 S	OUT	OUT S

Fig. 7. Evacuation model

The marks in the note represent:

$X_{SJ_{t-1}}$ - Position of the user regarding the location of the fire (N or S) in precedent time step [m]

XL_0 - Starting location of the self-rescue observed from a starting portal [m]

XL_t - Users location in the observed time period [m]

X_{poz} - Locations of the fire observed from a starting portal [m]

v_H - Walking speed [m/s]

t - Momentarily observed time period [s]

t_{SR} - Delay of the self-rescue after the start of the fire [s]

With the presented model the possibilities of the analysis or the following of the movement of the users in the tunnel increase additionally. The calculated locations are then used for checking the temperature and the smoke concentration on the ground in these places and consequently the level of risk.

8.5 Results

The first level of risk is presented by the presence of smoke that includes the first four risk stages, the presence of high temperature contributes additional (the highest) risk stage. The Table 6 presents a deterministic register of risk for a constant location in a tunnel during a fire that is 252 m north of the fire. The picture that we get whit this is very representative because it confirms the theory on safety analyses. From the table the safety categories can be seen and appropriate consequences can be allocated. Table 6 has especially a comparative purpose for finding the influence of different ways of ventilation on the fire dynamic, smoke and temperature development. We can logically assume that the risk in low fire force is lower in comparison with bigger fires. Following the same logic along with the consideration of different ways of ventilation and manner of management it soon becomes difficult. One of the noticeable differences is the level of the calculated risk (MR – medium risk) in longitudinal ventilation of a 20 MW fire. It is expected that the increased risk also appears in the 50 MW fire but it is not so. The search for a cause is difficult because this is hidden in the fluid dynamics during the fire, taking into account that the geometry, the discreteness and the initial and boundary conditions (except the force of the fire) are unaltered.

The second important result in the table is the possibility of analysing the influence of turning on the ventilators on the forming of the smoke curtain. It is especially noticeable in the transverse ventilation of 50 MW and 100MW fires where on the turning on, local increased temperatures and smoke concentrations occur. Table 6 is made as a functionally dependent dynamic register, which chooses the calculated values from the data base with the changing of the observed location and on the basis of conditions from the chapter above, calculates the risk.

With the evacuation model is possible to define the tunnel user's location in time intervals of one minute on the bases of the starting user's position, delay with the self-rescue procedure and the walking speed. In this way it is possible to predict the tunnel users movement for the following 15 min and check the smoke concentration and the temperature height to which they will be exposed or establish the risk level. The table represents a conceptual

Safety Category	Contingent											
	40	45	50	55	60	65	70	75	80	85	90	100
Power	2200 M	2300 M	2400 M	2500 M	2600 M	2700 M	2800 M	2900 M	3000 M	3100 M	3200 M	3300 M
Vibrations	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000	2100
Frequency	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000	2100
Amplitude	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000	2100
Direction	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000	2100
Phase	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000	2100
Period	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000	2100
Waveform	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000	2100
Envelope	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900	2000	2100
Transverse	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Longitudinal	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Normal	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Transverse	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Longitudinal	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Normal	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Transverse	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Longitudinal	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Normal	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Transverse	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Longitudinal	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Normal	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Transverse	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Longitudinal	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR
Normal	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR	LR

Table 6. The deterministic risk register for the chosen observer location

model for a general presentation of the risk in tunnels with different types of ventilation and different fire forces.

9. Conclusion on tunnel risk assessment

For a register of a fire in a tunnel and a safety evaluation, the probability accession is too general because a greater event number of physical legality is shown with a statistical probability. A relatively accurate fire dynamics register, which is possible with mathematical models, is often meant only for science. That is why the dissertation is ideally oriented in the use of mathematical CFD models for the making of a system of scenarios that can be further used for developing an effective fire plan or fire management, fire drills, etc. A complex of fire scenarios in different tunnel ventilations and fire forces is presented in the work. The work includes a qualitative analysis of the current circumstances in four different ventilation conditions; natural, longitudinal, transverse and semi transverse. In this way a comparison of individual types of ventilation systems, ventilation plans and their effectiveness in assuring sufficient evacuation times is possible. Also a possibility of usage on a singular tunnel is presented, for which a deterministic safety analysis within a selected number of scenarios would be made. Such an accession requires a lot of calculating time but it is changeable in the development of the safety analysis and fire plan. The geometry and some ventilation plans are "constants" in this case and only the fire location can be changed.

Apart from the number and way of setting the scenarios, the simulation results are values of the selected variables. The discussed variables are mostly the smoke density and the temperature which define the different risk levels on the basis of the human endurance in increased values and conditional interacting dependence. On this basis the deterministic risk register is made which is the key element of the dissertation. The register presents a passage between a practical way of using the CFD model and the user who needs clear and fast accessible data of the situation during a fire in a tunnel.

10. References

- Brussaard, L.A., Kruiskamp, M.M., Oude Essink M.P., The Dutch model for the Quantitative Risk Analyses of roadd tunnel, Ministry of Transport, The Netherlands 2004
- Cheng, L.H., T.H. Ueng, C.W. Liu, Simulation of ventilation and fire in the underground facilities, *Fire Safety Journal* 36 (2001) 597-619.
- Commision of the European Communities, Directive of the European Parliament and of the Council on minimum safety requirements for tunnels in the Trans-European Road Network, Brussels, 30.04.2004.
- Fletcher, C.A.J., *Computational Techniques of Fluid Dynamics-Volume 2*, Second edition, Springer-VerlagBerlin Heidelberg, 1991.
- Floyd, J.E., Wiczorek, C.J., Vandsburger, U., 2001. Simulation of the Virginia tech fire research laboratory using Large Eddy Simulations with mixture fraction chemistry and finite volume radiative heat transfer, *INTERFLAM 2001; Proc. intern. Symp., Edinburgh*, 17-19 September 2001:767-778.

- Gann, R.G., Hall, J.R., Fire Conditions for Smoke Toxicity Measurement, Fire and Materials, vol.18, 193-199, 1994
- Gasser, I., Struckmeier, J., Modelling and simulation of fires in vehicle tunnels, Computational fluid dynamics and data analyses, November 2002
- Haack A., Fire Protection in Traffic Tunnels: General and Results of the EUREKA Project, Tunnelling and Underground Space Technology, Vol. 13, No. 4, pp. 377-381.1998.
- Haack, A., Current safety issues in traffic tunnels, Tunnelling and Underground Space Technology 17 p.p. 117-127 - 2002.
- Heskestad, G., SFPE Handbook, chapter Fire Plumes. National Fire Protection Association, Quincy, Massachusetts, 2nd edition, 1995.
- Jang, H.M., Chen, F. A novel approach to the transient ventilation of road tunnels, Journal of Wind Engineering and Industrial Aerodynamics 86 (2000) 15-36.
- Kirchsteiger, C., On the use of probabilistic and deterministic methods in risk analysis, Journal of Loss Prevention in the Process Industries 12 (1999) 399-419
- Kunsch J.P., Critical velocity and range of a pre-gas plume in a ventilated tunnel, Atmospheric Environment 33 (1999) pp.: 13-24.
- Lesieur, M., Turbulence in Fluids, Kluwer Academic Publisher, Netherlands 1997, ISBN: 0-7923-4415-4(HB).
- McGrattan, K., Baum, H., Rehm, R., Hamins, A., Forney, G.P., Floyd, J.E. and Hostikka, S., 2001. Fire Dynamics Simulator - Technical reference guide, National Institute of Standard and Technology, NISTIR 6783, 2001.
- Memorial Tunnel Fire Ventilation Test Program, Massachusetts Highway Department, Central Artery/Tunnel Project, USA-1996
- Mingchung, L., Vaughan, B., Stoichiometric combustion model with oxygen threshold improved predictions for fire simulation using CFD model, pp. 559570, 1999.
- Persson, M., Quantitative Risk Analysis, Procedure for the Fire Evacuation of a Road Tunnel, Department of Fire Safety Engineering, Lund University, Sweden, Lund 2002
- PIARC Technical Committee on Tunnel Operation, Fire and Smoke Control in Road Tunnels, France 2003
- Sagaut, P., Large Eddy Simulations for Incompressible Flows, second edition, Springer Berlin Heidelberg 2002.
- Vidmar, P., Petelin, S., An analysis of a fire resulting from a traffic accident, Journal of Mechanical Engineering 49(2003), ISSN 0039-2480, pp. 1-13
- Woodburn P. J., Britter R. E., CFD Simulations of a Tunnel Fire, Fire Safety Journal 16 (1996) 35-62.

References on LNG

- Bottelberghs, P. H. (2000). Risk analysis and safety policy developments in the Netherlands. Journal of Hazardous Materials, 71
- Bubbico, R., Di Cave, S., Mazzarotta, B., Preliminary risk analysis for LNG tankers approaching a maritime terminal, Journal of Loss Prevention in the Process Industries, Volume 22, Issue 5, September 2009, Pages 634-638.

- Center for Chemical Process Safety, Guidelines for the management of change for process safety, John Wiley & Sons, Inc., Hoboken, New Jersey, 2008
- D.W. Hissong, Keys to modeling LNG spills on water, *Journal of Hazardous Materials* 140 (2007) 465-477
- Det Norske Veritas, MPACT Theory manual, Internal publication, DNV software, June 2007
- Det Norske Veritas, Risk Assessment LNG import Koper: nautical and unloading operations, Report no/DNV Reg No.: / 124UI0A-4, 2008
- Drysdale Dougal, An introduction to Fire dynamics, John Wiley & Sons Ltd - 1998
- Fells, I., Rutherford, A. G., Burning velocity of methane-air flames, *Combustion and Flame*, Volume 13, Issue 2, April 1969, Pages 130-138
- Gucma L. "Evaluation of oil spills in the Baltic Sea by means of simulation model and statistical data. *International Maritime Association of Mediterranean*", Kolev and Soares editors), Balkema 2007.
- Handleiding risicoberekening Bevi: inleiding, Module A/B/C - Versie 3.0", RIVM, January 2008
<http://www.hse.gov.uk/>
- International Maritime Organisation, IMO Guideline for alternative tanker design, 1995.
- ioMosaic, Modelling LNG Spreading and Vaporisation, ioMosaic Corporation 2007
- McGrattan, K., Baum, H., Rehm, R., Hamins, A., Forney, G.P., Floyd, J.E. and Hostikka, S., 2001. *Fire Dynamics Simulator - Technical reference guide*, National Institute of Standard and Technology, NISTIR 6783, 2001.
- Melhem, G. A., Kalelkar, A. S., *Understanding LNG Fire Hazards*, ioMosaic Corporation 2007
- Michael Hightower, M., Luketa-Hanlin, A., Gritzo, L. A., Covan, J. M., Sandia National Laboratories, Review of the independent risk assessment of the proposed Cabrillo liquefied natural gas deepwater port project, 2006
- Nautical Risk assessment LNG transport Rostock", DNV Energy, December 2007
- Perkovic, M., Gućma, L., Przywarty, M., Gućma, M., Petelin, S., Vidmar, P., Nautical risk assessment for LNG operations at the Port of Koper, International conference on traffic science, Portorož 2010
- Petelin S., Vidmar P., Perkovič M., Luin B., Kožuh M., Predlog prometno-varnostnih analiz za plinski terminal (Slovenian only), Portorož 2009
- Pitblado, R. M., Baik, J., Hughes, G. J., Ferro, C., Shaw, S. J., *Consequences of LNG Marine Incidents*, Det Norske Veritas (USA) Inc., Houston, 2004.
- Raj, P. K. , Lemoff, T., Risk analysis based LNG facility siting standard in NFPA 59A, *Journal of Loss Prevention in the Process Industries* 22 (2009) 820-829
- Safer System LLC, User's Guide Trace 8-Description of modelling algorithms, Westlake Village, California , USA-1996
- SFPE Handbook, Fire protection engineering, 2nd edition, National Fire Protection Association 1995
- Trbojevic, V. M., Risk criteria in EU, ESREL'05, Poland, 27-30 June 2005
- Vidmar, P., Petelin, S., An analysis of a fire resulting from a traffic accident, *Journal of Mechanical Engineering* 49(2003), ISSN 0039-2480, pp. 1-13

Vidmar, P., Petelin, S., Analysis of the effect of an external fire on the safety operation of a power plant, *Fire Safety Journal* 41 (2006) 486-490
World LNG Carrier Fleet, *LNG Journal*, Maritime Content Ltd, July/August 2008

Sail Performance Analysis of Sailing Yachts by Numerical Calculations and Experiments

Y. Tahara, Y. Masuyama, T. Fukasawa and M. Katori
*National Maritime Research Institute, Kanazawa Institute of Technology
Osaka Prefecture University, North Sails
Japan*

1. Introduction

Sails of a sailing yacht can be considered as multiple soft thin wings (membrane wings) with relative large cambers, and are often used at large attack angles. The shape of sail is determined as an equilibrium state of both aerodynamic force and tension acting on the sail surface. In particular, a spinnaker used for the running condition is a very soft membrane like a parachute, and the shape is simply formed by self-generated aerodynamic forces which are strongly affected by the sail shape itself. These facts lead to new challenges in the present problem, i.e., in the measurements the sail shape must be accurately measured in the flying condition, and in numerical simulation of flow and forces the sail flying shape is correctly given or predicted as a part of solution. The present study concerns the authors' ongoing effort on analyses of sail performance of sailing yachts by numerical calculations and experiments, and in this paper, the focus of discussions is more on the former. Two computational fluid dynamics (CFD) methods are used in the present study, and the results are validated through detailed comparison with experimental data. The data are obtained in onboard full-scale measurements by using a sail dynamometer boat. Our study concerns both the upwind and downwind sailing conditions; however, we focus on the former in the present chapter due to the limitation of space in this book. More detailed background of the present work is well described in Masuyama et al. (2009).

One of the two CFD methods is a Vortex Lattice method (VLM). Although the VLM is a potential flow calculation, it is well known the results agree well with the measured data at the upwind condition of small attack angle. The VLM is used as the sail design and making tool due to the quick convergence ability for the parametric survey of sail shape to obtain the desired sail performance, and also due to good compatibility with the finite element method (FEM) for the strength analysis. In this paper, a method to shed wake vortices step-by-step developed by Fukasawa was adopted in the Vortex Lattice method (Fukasawa, 1993; Fukasawa & Katori, 1993).

Another CFD method is a Multiblock Reynolds-Averaged Navier-Stokes (RANS)-based CFD named "*FLOWPACK*". This code was developed by Tahara specifically for CFD education and research, and design applications for ship hydrodynamics, aerodynamics, and fluid engineering (Tahara, 2008). As part of the developments for application to design problems, a complete multiblock domain decomposition feature was included. The numerical method of *FLOWPACK* solves the unsteady RANS and continuity equations for mean velocity and pressure. Either a zero or a two-equation turbulence model can be used for turbulence flow calculation, and in the present study the former was used. The *FLOWPACK* was included as a

solver in a sail performance analyzer named “*Advanced Aero Flow (AAF)*” developed by Katori (Katori, 2009). The AAF is a specialized package for the calculation of sail performance of sailing yachts, and composed of both mesh generator and post analyzer.

The sail shapes and performance were measured using a sail dynamometer boat *Fujin* under sailing condition on the sea (Masuyama et al., 1997a, 1997b). *Fujin* is a 34-foot LOA boat, in which load cells and CCD cameras were installed to simultaneously measure the sail forces and shapes. At the same time, the sailing conditions of the boat, e.g., boat speed, heel angle, wind speed, and wind angle, were measured. The shapes and 3D coordinates of the sails were used for the input data of the numerical calculations, and the calculated results were compared with the measured data. The sail coordinates with aerodynamic coefficients are tabulated for some sailing conditions in order to provide benchmark data for the CFD validation.

In this paper, overview of the above-mentioned CFD methods and experiments are described. As the aforementioned, sail flying shapes are considered in the present CFD so that the accurate prediction of flow and aerodynamic forces is possible. Discussion of the results is based on detailed comparison with the measurements. The discussion also includes the current capability of the CFD methods in the present problem, and prognosis for the enhancement of the capability in future work for higher accuracy and/or more complicated flow simulation. It will be noteworthy that the overall trends of the flow and the aerodynamic forces measured in the experiments are fairly well predicted by the present computations; and at the same time, experimental techniques originally implemented and used in the present study are shown very promising and capable to provide very detailed benchmark data for CFD validation.

2. Sail plan for the analysis

In this study the experiments and numerical calculations were performed for the upwind sailing condition. The sail shapes and performance were measured using a sail dynamometer boat *Fujin*. The sail plan of the *Fujin* and the coordinate system are shown in Fig.1. The principal dimensions of the boat and the detailed measurements of the sails are also shown in Table 1. The measurement system of the boat and testing conditions are described in section 5, and the measured and calculated results are compared and discussed in section 6.

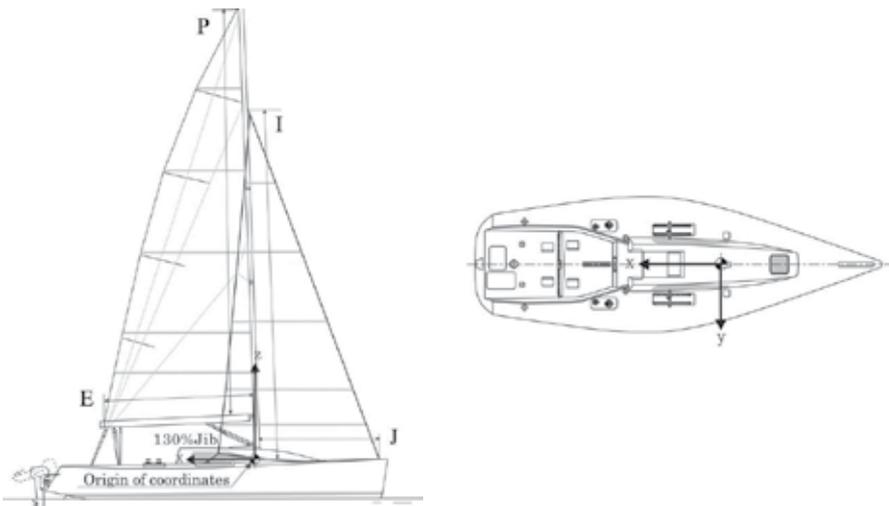


Fig. 1. Schematic showing the sail plan of *Fujin* with 130% jib and the coordinate system

HULL		SAIL DIMENSIONS		
Length Over All [m]	10.35		Mainsail	130% Jib
Length Water Line [m]	8.80	Peak Height [m]	13.82	10.70
Breadth Maximum [m]	3.37	Luff Length [m]	12.50	11.45
Breadth Water Line [m]	2.64	Foot Length [m]	4.44	4.89
Displacement [ton]	3.86	Sail Area [m ²]	33.20	26.10
SAIL		Height [%]	Chord Length [m]	
I [m]	11.00	0	4.44	4.89
J [m]	3.61	10	4.13	4.44
P [m]	12.55	20	3.85	3.94
E [m]	4.51	40	3.23	2.94
I, J, P, E of Sail are defined in Fig. 1.		60	2.43	1.97
		80	1.39	0.98
		100	0.15	0.10

Table 1. Principal dimensions of *Fujin* and detailed measurements of sails

3. Overview of Vortex Lattice method (VLM)

3.1 Basic concept of Vortex Lattice method

The Vortex Lattice method is a branch of CFD, and it is often used at the early stage of yacht sail design because of the comparatively less computational time. This method is based on the potential theory, similar to the panel method, and the flow around the sail is expressed by discrete vortices. The Vortex Lattice method has its root in the lifting line theory formulated by Prandtl in 1918. A wing is represented by a single vortex line in the lifting line theory, and the force acting on the wing is approximated by the force acting on the vortex line.

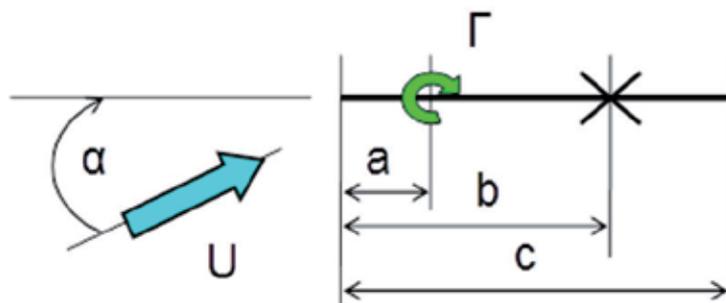


Fig. 2. 2-Dimensional flat plate wing and a vortex filament

Firstly, 2-dimensional flow around a flat plate wing is considered. A vortex filament is located at a distance "a" from the leading edge of the wing as shown in Fig.2. Although the onset flow U is constant, the flow over the wing is accelerated, while it is decelerated below the wing, because of the flow induced by the vortex filament. This leads to the pressure decrease on the back surface and the pressure increase on the front surface of the wing accordingly to Bernoulli's theorem. This means that the flow around the wing can be realized by a vortex filament in the flow. The strength of vortex filament, or sometimes called circulation, is determined by a boundary condition on the wing; that is, there is no cross flow through the

wing. This boundary condition is usually satisfied at a certain point called control point. Assuming that the control point is located at a distance “ b ” from the leading edge of the wing shown in Fig.2, and satisfying the boundary condition at this point, the strength of the vortex filament can be determined. Once the strength of vortex filament is determined, the lift acting on the vortex filament can be calculated according to Kutta-Joukowski theorem, that is,

$$L = \rho U \Gamma \quad (1)$$

where ρ is the density of the fluid and Γ is the strength of vortex filament. If the calculated lift is assumed to equal that generated in a 2-dimensional thin parabolic shape airfoil, the locations of the vortex filament and the control point are determined to be $a = c/4$ and $b = 3c/4$, where “ c ” is the chord length of the wing. This is called 1/4-3/4 rule, which was shown by Pistoiesi (Pistoiesi, 1937). This rule is used as the basis of the present Vortex Lattice method.

3.2 Application to sail configuration

As the yacht sail is a 3-dimension shape body, attention should be paid to the treatment of the end of vortex line. According to the Helmholtz's theorem on vortex, the vortex line should expand from the boundary to the boundary of the flow or shuts oneself and makes vortex ring. Accordingly, in the 3-dimensional body, the vortex line should expand infinity from the edge of the body. In the lifting line theory, or the Vortex Lattice method, the vortex line is assumed to be a horseshoe type shown in Fig.3(a), and the vortex line changes its direction at the edge of the body to extend to infinity as trailing free vortices.

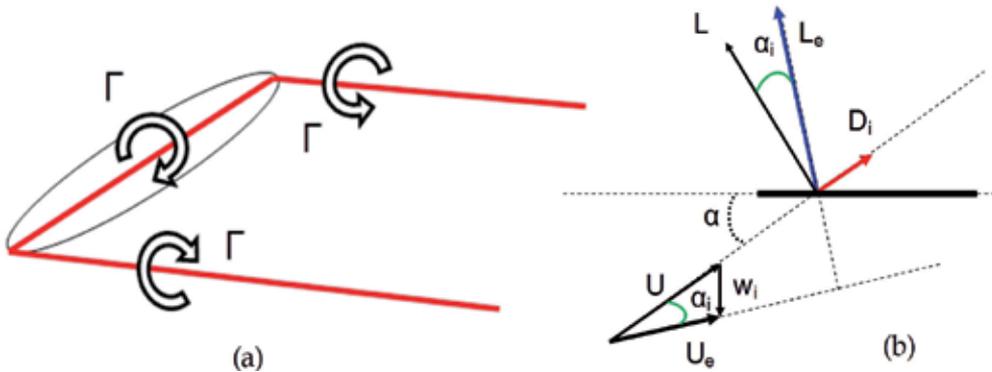


Fig. 3. Horseshoe vortex and downwash effect

In case of the horseshoe vortex shown in Fig.3(a), the flow induced by each vortex line affects the onset flow in the magnitude and the direction. According to Biot-Savart law, the velocity vector induced by a slight part of the vortex line $d\ell$ is given by

$$\vec{v} = \frac{\Gamma}{4\pi} \oint \frac{d\vec{\ell} \times \vec{r}}{|\vec{r}|^3} \quad (2)$$

where \vec{r} is a position vector from the vortex part to the point concerned. The downward velocity called downwash w_i is calculated by using Equation (2), which affects the onset flow. This causes the reduction of the attack angle of total inflow by α_i as shown in Fig.3(b),

and the total velocity of inflow into the wing changes to U_e . If the lift acting on the wing is defined as the force perpendicular to the onset flow direction, it is given by the following formula according to Kutta-Joukowski theorem.

$$L = L_e \cos \alpha_i \approx L_e = \rho U_e \Gamma \approx \rho U \Gamma \quad (3)$$

In this case, the force in the onset flow direction is generated, which is the apparent drag called induced drag given by

$$D_i = L_e \sin \alpha_i \approx L_e \alpha_i = \frac{L_e}{U} w_i \approx \frac{L}{U} w_i \quad (4)$$

The induced drag is a distinctive drag in a 3-dimensional wing, and does not appear in a 2-dimensional wing.

In the Vortex Lattice method, the lift, induced drag, and center of pressure are calculated by arranging horseshoe vortices of different strength on the surface of sail. By placing a number of horseshoe vortices, the sail of complex shape with twist, camber, or two or more sails, can be analyzed. Falkner used the name "Vortex Lattice" firstly in his report, in which a wing was covered with a grid of straight horseshoe vortices (Falkner, 1943, 1946). In 1950's, only the analysis where the trailing vortices are placed in the straight line was able to be carried out because of the computer capability, and the accuracy was questionable. It was 1965 when the Vortex Lattice method started to demonstrate its ability along with the development of computer, and the method came to be used for the performance prediction of yacht sail. The yacht sail is one of the most suitable objects for applying the Vortex Lattice method because of its thickness, if the viscous effect of fluid can be disregarded. An application of the Vortex Lattice method to the performance prediction of yacht sail will be explained in the following paragraph with the use of a step-by-step procedure to estimate the trailing vortex deformations.

Discretized horseshoe vortices are located on the sail plane in the Vortex Lattice method. It is usual to divide the sail plane into quadrilateral panels as shown in Fig.4, and the

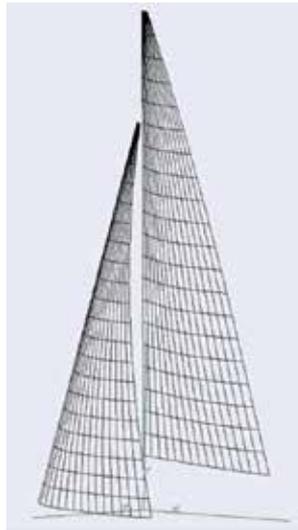


Fig. 4. Panel discretization of sails

horseshoe vortices are placed at 1/4 length of panel from the front end edge of the panel so as to trail the trailing vortices rearwards. The strengths of the horseshoe vortices are determined by satisfying the boundary condition on the sail; that is, the total flow of the onset flow and the induced wake by vortices is parallel to the sail surface at control points. The control point is taken to be the point 3/4 of length of panel from the front end edge of the panel according to the 1/4-3/4 rule.

According to Biot-Savart law, the velocity vector at the control point of i -th panel induced by other vortices are given by

$$\vec{v}_i = \frac{1}{4\pi} \int \frac{\Gamma (\vec{d}\ell \times \vec{r})}{|\vec{r}|^3} = \sum_j \frac{\Gamma_j}{4\pi} \sum_{k=1}^3 \frac{\cos\alpha_{kji} + \cos\beta_{kji}}{h_{kji}} \vec{e}_{kji} \quad (5)$$

where α_{kji} , β_{kji} , h_{kji} , are the angles and the distance of a k -th filament of a j -th horseshoe vortex and i -th control point shown in Fig.5, which shows a plane containing the vortex filament and the control point. \vec{e}_{kji} is a unit vector perpendicular to the plane shown in Fig.5. $k=1, 2, 3$ in Equation (5) denotes each vortex filament of a horseshoe vortex. With the use of Equation (5), the boundary condition on the control point can be given by

$$\sum_i^{NB} \vec{v}_i \cdot \vec{n} = \vec{U} \cdot \vec{n}_i - \sum_i^{NW} \vec{v}_i \cdot \vec{n} \quad (6)$$

where \vec{U} is an onset flow velocity vector and \vec{n} is the unit normal vector at the control points. NB is the number of bound horseshoe vortices on the sail plane and NW is the number of trailing horseshoe vortices in the wake. Equation (6) can be written in the vector matrix form by

$$[\Lambda]\{\Gamma\} = \{\mathbf{u}_n\} + \{\mathbf{v}_n\} \quad (7)$$

Solving Equation (7), the strength of bound vortices can be obtained.

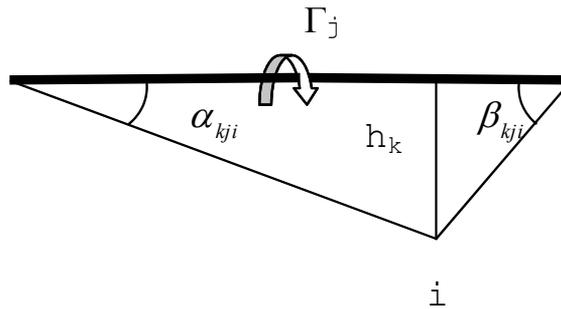


Fig. 5. K -th filament of j -th horseshoe vortex and i -th control point

3.3 Step-by-step vortex shedding technique

The important point in the Vortex Lattice approach to yacht sail is the handling of the wake of the sail. The wake vortices proceed downstream from the trailing edge, or leech/foot of sail, in the Vortex Lattice method. The location of wake vortices are determined by the condition that they are free vortices; that is, the stream line of wake vortices should be

parallel to the velocity field induced by total vortex system. A step-by-step procedure developed by Fukasawa was adopted in this paper to determine the strength of the bound vortices and the location of wake vortices (Fukasawa, 1993; Fukasawa & Katori, 1993; Masuyama et al., 1997a, 1997b). The wake vortices are shed from the trailing edge in each time step according to Helmholtz's theorem; that is,

$$\Gamma_B + \Gamma_W = 0 \quad (8)$$

$$\frac{D\Gamma_B}{Dt} = 0 \quad (9)$$

where Γ_B and Γ_W are the total strength of bound vortices, or the circulation around the sail, and the strength of wake vortices, respectively. From Equation (9), we have

$$\frac{\partial \Gamma_B}{\partial t} = \tilde{U} \frac{\partial \Gamma_B}{\partial \xi} \quad (10)$$

and substituting Equation (10) into (8), we have

$$\frac{\partial \Gamma_B}{\partial t} = -\tilde{U} \frac{\partial \Gamma_W}{\partial \xi} \quad (11)$$

where ξ is taken to the downstream direction, and \tilde{U} is the local velocity at the wave vortex. Assuming that the wake vortices proceeds $\Delta\xi$ downstream in a time step Δt with the velocity \tilde{U} , the strength of wake vortex shed at time step k , can be given by integrating Equation (11), that is,

$$\Delta\Gamma_W^k = -\frac{1}{\tilde{U}} \int_0^{\Delta\xi} \frac{\partial \Gamma_B}{\partial t} d\xi = -\frac{\Delta\xi}{\tilde{U} \Delta t} (\Gamma_B^k - \Gamma_B^{k-1}) = \Gamma_B^{k-1} - \Gamma_B^k \quad (12)$$

Equation (12) means that the strength of wake vortex shed at time step k is the increase of the strength of bound vortex from time step $k-1$ to time step k . Once a vortex filament is shed at time step k , it proceeds downstream with a constant strength according to the local field velocity, i.e., each horseshoe wake vortex moves in the direction of field velocity in each time step. The field velocity is updated in every time step. The calculation is carried forward until the the calculated lift and drag forces converges. The forces vector and the moment acting on the sail are calculated accordingly to Kutta-Joukowski theorem.

$$\vec{F} = \rho \int \Gamma \tilde{U} \times d\vec{s} \quad (13)$$

with the use of the vortex strengths of the wake vortices and the bound vortices determined by solving Equation (7).

Finally, the overall numerical solution procedure of the present Vortex Lattice method is summarized as follows:

- Step 1. Divide the sail planes into quadrilateral panels, and allocate horseshoe vortices on the sail plane.
- Step 2. Input the mast rake angle, heel angle of the yacht, apparent wind speed and apparent wind angle.

- Step 3. Solve the strength of bound horseshoe vortices on the sail plane with non-deformed trailing vortices.
- Step 4. Compute the total circulation around the sail caused by horseshoe vortices.
- Step 5. Compute the increment of the total circulation and shed free horseshoe vortices according to Equation (12).
- Step 6. Calculate the local velocities along the wake, and deform the trailing vortices.
- Step 7. Compute the force vector and moment acting on the sails.
- Step 8. Solve the strength of bound horseshoe vortices on the sail plane with trailing free vortices.
- Step 9. Repeat Step 4 through Step 8 until the force is converged.

Fig.6 shows the example calculation results. In the present study, the mast and rigging were not considered for the series calculations, and the mirror image was taken into account about the deck plane of the boat. Since the vortex lattice methods do not predict viscous drag, the viscous drag acting on the sails and rigging was calculated empirically using a drag coefficient C_{Dp} . The value of C_{Dp} was obtained from the measured data in the previous papers and formulated for the upwind condition as follows:

$$C_{Dp} = 0.0026 \gamma_A + 0.005 \quad (14)$$

where γ_A is apparent wind angle in degrees.

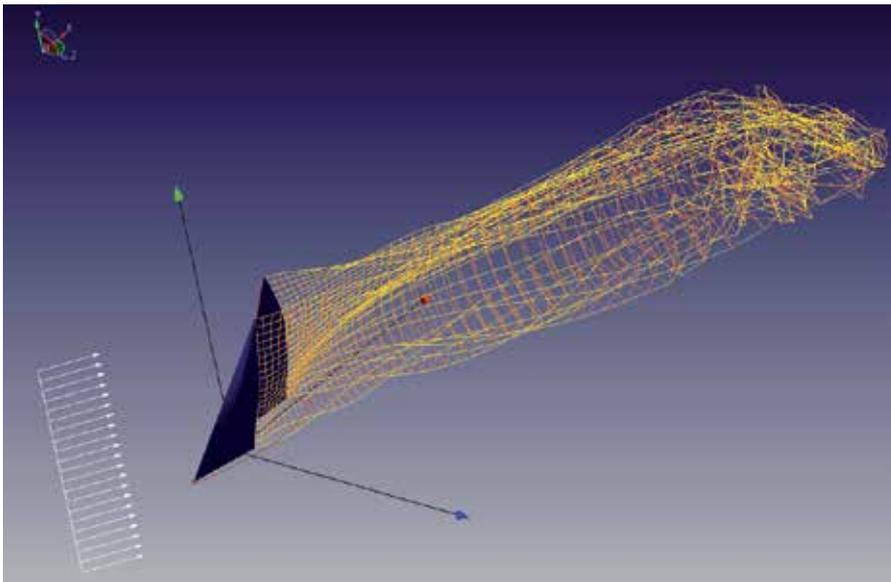


Fig. 6. Calculated wake by Vortex Lattice method

4. Overview of Reynolds-averaged Navier-Stokes equation method

The RANS-based CFD method used in the present study was *FLOWPACK*. The code was developed by Tahara specifically for CFD education and research and for design applications for ship hydrodynamics, aerodynamics, and fluid engineering. As part of the developments for application to design problems, a complete multiblock domain

decomposition feature was included. At present, *FLOWPACK* has a good interface with the authors' inhouse automatic grid generator as well as with commercial grid generation software. For a complete documentation of the method is available in Tahara (2008). In the following, an overview of the numerical method is given.

4.1 Governing equations

Let us consider a sail system fixed in the uniform onset flow (see Fig.1 for the basic coordinate system). The non-dimensional RANS equations for unsteady, three-dimensional incompressible flow can be written in Cartesian tensor notation as

$$\frac{\partial U_i}{\partial t} + U_j \frac{\partial U_i}{\partial x^j} + \frac{\partial \overline{u_i u_j}}{\partial x^j} + \frac{\partial p}{\partial x^i} - \frac{1}{\text{Re}} \nabla^2 U_i = 0 \quad (15)$$

$$\frac{\partial U_i}{\partial x^i} = 0 \quad (16)$$

where U_i ($i=1,2,3$) = (U, V, W) and u_i ($i=1,2,3$) = (u, v, w) are the Cartesian components of mean and fluctuating velocities, respectively, normalized by the reference velocity U_0 , x^i ($i=1,2,3$) = (X, Y, Z) is the dimensionless coordinates normalized by a characteristic length L , $\text{Re} = U_0 L / \nu$ is the Reynolds number, ν is the kinematic viscosity, the barred quantities $\overline{u_i u_j}$ are the Reynolds stresses normalized by U_0^2 , and p is the pressure normalized by ρU_0^2 . If $\overline{u_i u_j}$ are related to the corresponding mean rate of strain through an isotropic eddy viscosity ν_t , i.e.,

$$\overline{u_i u_j} = \nu_t \left(\frac{\partial U_i}{\partial x^j} + \frac{\partial U_j}{\partial x^i} \right) - \frac{2}{3} \delta_{ij} k \quad (17)$$

where $k = (\overline{uu} + \overline{vv} + \overline{ww}) / 2$ is the turbulent kinetic energy, Equation (15) becomes

$$\frac{\partial U_i}{\partial t} + \left(U_j - \frac{\partial \nu_t}{\partial x^j} \right) \frac{\partial U_i}{\partial x^j} - \frac{\partial \nu_t}{\partial x^j} \frac{\partial U_j}{\partial x^i} + \frac{\partial p}{\partial x^i} \left(p + \frac{2}{3} k \right) - \frac{1}{R_\phi} \nabla^2 U_i = 0 \quad (18)$$

where $1/R_\phi = 1/\text{Re} + \nu_t$, and $\phi = U_i$ ($i=1,2,3$). Equations (16) and (18) can be solved for U_i and p when a suitable turbulence model is employed to calculate the eddy-viscosity distribution. Either a zero or a two-equation turbulence model can be used for turbulent flow calculation, and a model used for the present study is the former, i.e., Baldwin-Lomax model (Baldwin & Lomax, 1978), which is an algebraic scheme that makes use of a two-layer isotropic eddy-viscosity formulation. Detailed validation study of this model for boundary layer flows around three-dimensional bodies was done by the author (Tahara, 1995; Tahara & Stern, 1996). In this model, the eddy viscosity is evaluated as follows:

$$\begin{cases} (\nu_t)_{inner} & y \leq y_c \\ (\nu_t)_{outer} & y > y_c \end{cases} \quad (19)$$

where y is the distance normal to the wall surface and y_c is the minimum value of y where both the inner and outer viscosities match. The inner viscosity follows the Prandtl-Van

Driest formula, i.e., $(\nu_t)_{\text{inner}} = \lambda^2 |\omega|$, where $\lambda = \kappa \gamma [1 - \exp(-y^+/A^+)]$ is the turbulent length scale for the inner region, κ and A^+ are model constants, $|\omega|$ is the vorticity magnitude, and y^+ is the dimensionless distance to the wall. In the outer region, eddy viscosity is given by $(\nu_t)_{\text{outer}} = K C_{\text{cp}} F_{\text{wake}} F_{\text{Kleb}}$, where K and C_{cp} are model constants, $F_{\text{wake}} = \min(y_{\text{max}} F_{\text{max}}, C_{\text{wk}} y_{\text{max}} U_{\text{dif}}^2 / F_{\text{max}})$, and $F_{\text{Kleb}} = [1 + 5.5(C_{\text{Kleb}} y / y_{\text{max}})^6]^{-1}$. The F_{max} and y_{max} are determined by the value and corresponding location, respectively, of the maximum of $F = y |\omega| [1 - \exp(-y^+/A^+)]$. The quantity U_{dif} is the difference between maximum and minimum velocity magnitudes in the profile and is expressed as $U_{\text{dif}} = (U^2 + V^2 + W^2)_{\text{max}}^{1/2} - (U^2 + V^2 + W^2)_{\text{min}}^{1/2}$. C_{Kleb} and C_{wk} are additional model constants. Numerical values for the model constants are $A^+ = 26$, $\kappa = 0.4$, $K = 0.0168$, $C_{\text{cp}} = 1.6$, $C_{\text{wk}} = 1.0$, and $C_{\text{Kleb}} = 0.3$.

4.2 Discretization and velocity-pressure coupling

In the following, discretization and velocity-pressure coupling of the present RANS method are described. First, it is convenient to rewrite the transport equations for momentum (U_i) in the following general form:

$$\nabla^2 \phi = R_\phi \left[\sum_{j=1}^3 \left(U_j - \frac{1}{\sigma_\phi} \frac{\partial v_t}{\partial x^j} \right) \frac{\partial \phi}{\partial x^j} + \frac{\partial \phi}{\partial t} \right] + s_\phi \quad (20)$$

where ϕ again represents any one of the convective transport quantities (U_i), and s_ϕ is the source function for the corresponding quantity. We transform the physical space (x^i, t) into a rectangular region in the computational space (ξ^i, τ) using the following coordinate transformations:

$$t = \tau, \quad x^i = x^i(\xi^j), \quad e_i \cdot \nabla \phi = \frac{1}{J} \sum_{j=1}^3 b_i^j \frac{\partial \phi}{\partial \xi^j}$$

$$\nabla^2 \phi = \sum_{i=1}^3 \sum_{j=1}^3 g^{ij} \frac{\partial^2 \phi}{\partial \xi^i \partial \xi^j} + \sum_{j=1}^3 f^j \frac{\partial \phi}{\partial \xi^j}, \quad \frac{\partial \phi}{\partial t} = \frac{\partial \phi}{\partial \tau} - \frac{1}{J} \sum_{i=1}^3 \sum_{j=1}^3 b_i^j \frac{\partial x^i}{\partial \tau} \frac{\partial \phi}{\partial \xi^j}$$

Then the continuity equation (16) and the transport equations (20) for momentum parameters can be written as

$$\frac{1}{J} \sum_{i=1}^3 \sum_{j=1}^3 \frac{\partial}{\partial \xi^i} (b_j^i U_j) = 0 \quad (21)$$

$$\sum_{j=1}^3 \left(g^{ij} \frac{\partial^2 \phi}{\partial \xi^i \partial \xi^j} - 2a_\phi^j \frac{\partial \phi}{\partial \xi^j} \right) = R_\phi \frac{\partial \phi}{\partial \tau} + S_\phi \quad (22)$$

where

$$2a_\phi^j = \frac{R_\phi}{J} \sum_{n=1}^3 b_n^j \left(U_n - \frac{\partial x^n}{\partial \tau} - \frac{1}{J \sigma_\phi} \sum_{m=1}^3 b_n^m \frac{\partial v_t}{\partial \xi^m} \right) - f^j$$

$$S_\phi = s_\phi - 2 \left(g^{12} \frac{\partial^2 \phi}{\partial \xi^1 \partial \xi^2} + g^{13} \frac{\partial^2 \phi}{\partial \xi^1 \partial \xi^3} + g^{23} \frac{\partial^2 \phi}{\partial \xi^2 \partial \xi^3} \right)$$

The geometric coefficients b_i^j , g^{ij} , and f^j appearing in the above equations are defined by Thompson et al. (1985). The transport equations (22) on a computational cell (shown in Fig.7(a)) are linearized and evaluating coefficients and source term at the center node P of the element yields

$$\sum_{j=1}^3 \left(g_P^{jj} \frac{\partial^2 \phi}{\partial \xi^j \partial \xi^j} - 2(a_\phi^j)_P \frac{\partial \phi}{\partial \xi^j} \right) = (R_\phi)_P \frac{\partial \phi}{\partial \tau} + (S_\phi)_P \quad (23)$$

or

$$g_P^{11} \phi_{\xi^1 \xi^1} + g_P^{22} \phi_{\xi^2 \xi^2} + g_P^{33} \phi_{\xi^3 \xi^3} = 2(C_\phi)_P \phi_{\xi^1} + 2(B_\phi)_P \phi_{\xi^2} + 2(A_\phi)_P \phi_{\xi^3} + (R_\phi)_P \phi_t + (S_\phi)_P \quad (24)$$

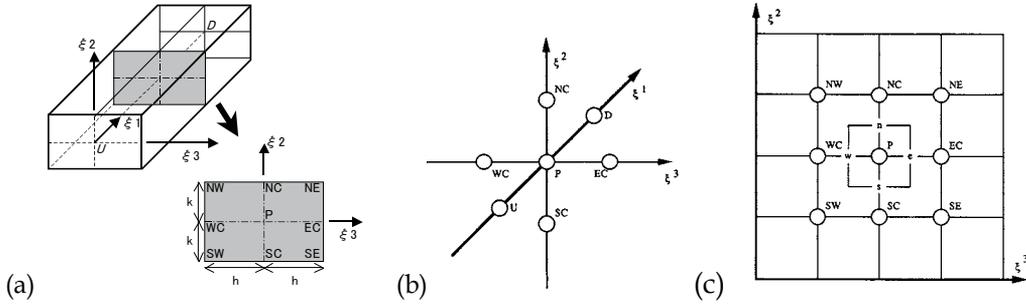


Fig. 7. Definition sketch of a computational cell (a), and nodes in regular grid (b) and continuity cell (c).

The dimensions of the computational cell are $2l \times 2k \times 2h$, where $l = 1/\sqrt{g_P^{11}}$, $k = 1/\sqrt{g_P^{22}}$, and $h = 1/\sqrt{g_P^{33}}$. The above equation is discretized by the finite-analytic scheme. Solution dependent coefficients are analytically derived by solving the above linearized transport equation using a hybrid method which combines a two-dimensional analytic solution in $\xi^2 \xi^3$ -plane with one dimensional analytic solution in the ξ^1 direction. By specifying boundary conditions on the faces of the cell as a combination of exponential and linear functions, which are the natural solutions for the linearized transport equation, Equation (24) can be solved by the method of separation of variables. When the solution is evaluated at the center node P of the element, the following twelve-point finite analytic formula is obtained:

$$\phi_P^n = \frac{1}{1 + C_P \left(C_U + C_D + \frac{R\phi}{\Delta \tau} \right)} \left[\begin{array}{l} C_{NE} \phi_{NE}^n + C_{NW} \phi_{NW}^n + C_{SE} \phi_{SE}^n + C_{SW} \phi_{SW}^n \\ + C_{EC} \phi_{EC}^n + C_{WC} \phi_{WC}^n + C_{NC} \phi_{NC}^n + C_{SC} \phi_{SC}^n \\ + C_P \left(C_U \phi_U^n + C_D \phi_D^n + \frac{R\phi}{\Delta \tau} \phi_P^{n-1} \right) - C_P (S_\phi)_P \end{array} \right] \quad (25)$$

where

$$\begin{aligned}
C_U &= \frac{Ce^{Cl}}{l \sinh Cl}, C_D = \frac{Ce^{-Cl}}{l \sinh Cl}, C_{SC} = \frac{e^{Bk}}{2 \cosh Bk} P_A, C_{NC} = e^{-2Bk} C_{SC} \\
C_{WC} &= \frac{e^{Ah}}{2 \cosh Ah} P_B, C_{EC} = e^{-2Ah} C_{WC}, C_{SW} = \frac{e^{Ah+Bk}}{4 \cosh Ah \cosh Bk} (1 - P_A - P_B) \\
C_{SE} &= e^{-2Ah} C_{SW}, C_{NW} = e^{-2Bk} C_{SW}, C_{NE} = e^{-2Ah-2Bk} C_{SW} \\
C_P &= \frac{h \tanh Ah}{2A} (1 - P_A) = \frac{k \tanh Bk}{2B} (1 - P_B), P_A = 4E_2 Ah \cosh Ah \cosh Bk \coth Ah \\
P_B &= 1 + \frac{Bh \coth Bk}{Ak \coth Ah} (P_A - 1), E_2 = \sum_{m=1}^{\infty} \frac{-(-1)^m (\lambda_m h)}{\left[(Ah)^2 + (\lambda_m h)^2 \right]^2 \cos \sqrt{A^2 + B^2 + \lambda_m^2} k} \\
\lambda_m h &= \left(m - \frac{1}{2} \right) \pi
\end{aligned}$$

The subscripts P, U and D denote the center, upstream and downstream nodes, respectively, and NC, NW, WC, etc. denote the nodes in the $\xi^2 \xi^3$ -plane in terms of compass directions. The superscripts (n) and (n-1) refer to the current and previous time levels, and $\Delta \tau$ is the time step. The solution of the complete flow equations involves a global iteration process, in which the velocity-pressure coupling is effected by PISO-type predictor-corrector steps. The pressure equation is derived by introducing pseudo-velocities at staggered locations while maintaining the regular grid arrangement for all the transport equations. Fig.7(b) and (c) show the locations of nodes in the regular grid in the $\xi^2 \xi^3$ -plane. All transport quantities and pressure are evaluated at the regular nodes. In deriving the pressure equation, a control volume is employed as a continuity cell, to establish the coupling between the velocity and pressure fields. The pressure equation used in this study is written as

$$\begin{aligned}
& \left(E_d^{11} + E_u^{11} + E_n^{22} + E_s^{22} + E_e^{33} + E_w^{33} \right) p_P \\
& = E_d^{11} p_D + E_u^{11} p_U + E_n^{22} p_{NC} + E_s^{22} p_{SC} + E_e^{33} p_{EC} + E_w^{33} p_{WC} - \hat{D}
\end{aligned} \tag{26}$$

with

$$\hat{D} = \hat{U}_d^1 - \hat{U}_u^1 + \hat{U}_n^2 - \hat{U}_s^2 + \hat{U}_e^3 - \hat{U}_w^3$$

Here E^{ij} and a modified pseudovelocity \hat{U}^i at the regular node are

$$E^{ij} = \frac{R C_P}{J \left[1 + C_P \left(C_U + C_D + \frac{R}{\Delta \tau} \right) \right]} \sum_{m=1}^3 b_m^i b_m^j, \hat{U}^i = \sum_{n=1}^3 b_n^i \hat{U}_n - E^{ij} \frac{\partial p}{\partial \xi^j} - E^{ik} \frac{\partial p}{\partial \xi^k} \tag{27}$$

where \hat{U}_i is a pseudovelocity given by the decomposition of Equation (25) for U_i into \hat{U}_i plus the pressure gradient terms, such that

$$U_i = \hat{U}_i - \frac{R C_P}{J \left[1 + C_P \left(C_U + C_D + \frac{R}{\Delta \tau} \right) \right]} \sum_{j=1}^3 b_i^j \frac{\partial p}{\partial \xi^j} \quad (28)$$

The coefficients and the modified pseudovelocities in the above equations are defined at the staggered node, and obtained from those at the regular node by the one-dimensional linear interpolation. The solution of the complete flow equations involves a global iteration process, in which the velocity-pressure coupling is effected by predictor-corrector steps. In the predictor step, the pressure field at the previous time step is used in the solution of the implicit equations (25) to obtain the corresponding velocity field. Since the velocity field generally does not satisfy mass conservation, a corrector step is needed. In the corrector step, the explicit momentum equations (28) and the implicit pressure equation (26) are solved iteratively to ensure the satisfaction of the continuity equation.

4.3 Multiblock (domain decomposition) capability, and overall numerical solution procedure

As mentioned earlier, the multiblock (domain decomposition) capability is facilitated in the present RANS method. This capability is essential for simulation of flow around complex geometry, e.g., multiple sail system for sailing yacht as focused in the present study. Fig.8 shows overview of the present multiblock computational grid, while the grid is generated by using an automatic gridding scheme developed by the present author (Masuyama et al., 2009). Note that the gridding engine together with the present RANS method was recently implemented into a comprehensive sail performance prediction software “*Advanced Aero Flow*” (Katori, 2009). See the reference for more details of the scheme. Total number of grids is around a half million, and the number of multiblock is 48. Free-stream, symmetry, and wall-surface (no slip) boundary conditions are imposed on outer and top boundaries, bottom boundary, and sail surface boundary, respectively. For the results shown in this paper, the mast and rigging are not considered in the series calculations, and the bottom boundary is located at the same height as that of deck plane of the boat (see Masuyama et al., 2009, for the results for which mast influences in computation are considered).

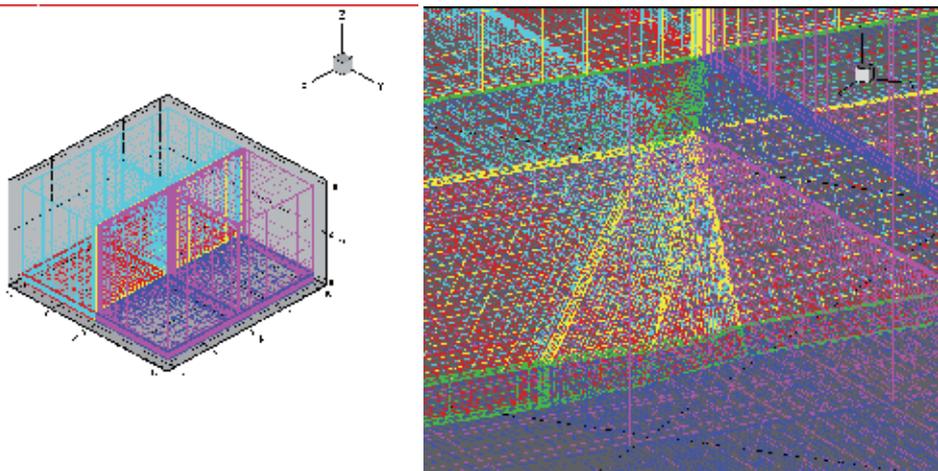


Fig. 8. Overview of the present multiblock computational grid.

The basic strategy to handle the multiblock follows domain decomposition technique to solve the elliptic PDE by using several subdomains. After adequate discretization is applied and a simple preconditioner is introduced, the discrete alternating Schwarz’s method to solve the PDE is used for boundary matching. Finally, Fig.9 shows the code structure of the present RANS method, and the overall numerical solution procedure of the present RANS method is summarized as follows:

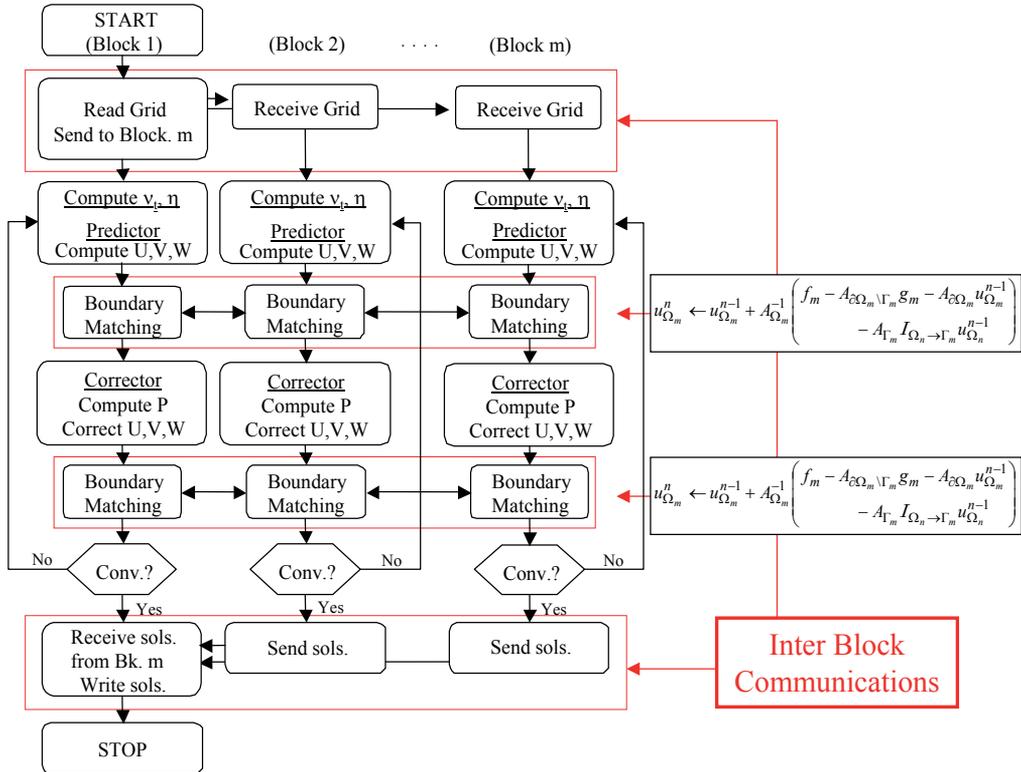


Fig. 9. PISO type solution algorithm for the present multiblock RANS method.

- Step 1. Input the computational grid, setup parameters, and boundary condition information.
- Step 2. Specify the initial conditions for the velocity, pressure and turbulence fields.
- Step 3. Compute the geometric coefficients.
- Step 4. Compute the finite-analytic coefficients for the transport equation.
- Step 5. Compute eddy viscosity distribution.
- Step 6. Solve transport equation for velocities (U, V, W) using the previous pressure field (predictor stage for the velocity field).
- Step 7. Compute the coefficients of pressure equation.
- Step 8. Solve pressure equation.
- Step 9. Using the newly obtained pressure, calculate the new velocity field explicitly (corrector stage for the velocity field).
- Step 10. Update the finite-analytic coefficients for the transport equation for velocities (U, V, W).

Step 11. Repeat Step 8 through Step 10 for the specified number of times.

Step 12. Return to Step 4 for the next time step, until the time step reaches the given maximum value.

More details of the present RANS method are described in Tahara et al. (2006a, 2006b) in addition to the above-cited references.

5. Measurements of upwind sail performance in full-scale condition using sail dynamometer boat *Fujin*

5.1 Full-scale measurements

Full-scale onboard measurements are free from scale-effect problem by wind tunnel tests and appear more promising, but the challenge becomes how to accurately measure forces acting on the sail. Such studies on sail force measurements were performed by Milgram et al. (1993), Masuyama et al. (1997a, 1997b), and Hochkirch et al. (1999), who built full-scale boats with onboard sail dynamometer systems.

Milgram (1993) showed in his pioneering work that the sail dynamometer boat, *Amphitrette*, is quite capable. This measurement system consists of a 35-foot boat with an internal frame connected to the hull by six load cells, which were configured to measure all forces and moments acting on the sails. In his work, the sail shapes were also measured and used for CFD analyses; however unfortunately, details of the sail shape and performance data were not presented. Hochkirch et al. (1999) also built a 33-foot dynamometer boat *DYNA*. The aerodynamic forces acting on the sail were measured and compared with the results from wind tunnel tests (Hansen et al. 2003). The measured data were also used as input to the CFD calculation and a parametric survey was carried out (Krebber et al. 2006). Masuyama and Fukasawa were encouraged by Milgram's work, and built a sail dynamometer boat, *Fujin*. The measurement system installed in the *Fujin* and the results of calibration test and sailing test were reported by Masuyama et al. (1997a and 1997b).

5.2 Measurements by sail dynamometer boat *Fujin*

The *Fujin* was originally built for conducting tests on sails for the Japanese America's Cup entry in 1994. *Fujin* is a 10.3m-long ocean cruiser with a sail dynamometer system in the hull which can directly measure sail forces and moments. Fig. 1 shows the general arrangement of the *Fujin*. The test sails were made to correspond to a typical sail plan for an International Measurement System (IMS) class boat. The rigging of the *Fujin* was originally designed for testing sails for the International America's Cup Class (IACC) boat. The jib of IACC boat is relatively small. Therefore, the longitudinal position of the jib rail track of the *Fujin* was located further forward than that of the typical IMS boat. For this reason, the tests were performed using a fully batten mainsail and a 130% jib instead of a 150% jib. The sails were made by North Sails Japan. The axes system is also shown in Fig. 1. The origin is located on the vessel's centerline at the aft face of the mast (x -direction), and the height of deck level at the base of the forestay (z -direction). Table 1 shows the principal dimensions of the boat and the detailed measurements of the sails, where "I", "J", "P" and "E" are the measurement lengths of sail dimensions for the IMS rule as defined in Fig.1.

The aerodynamic coefficients and the coordinates of the center of effort of the sails are defined as follows:

$$C_X = \frac{X_S}{\frac{1}{2}\rho_a U_A^2 S_A}, \quad C_Y = \frac{Y_S}{\frac{1}{2}\rho_a U_A^2 S_A}, \quad x_{CE} = \frac{N_S}{Y_S}, \quad z_{CE} = \frac{K_S}{Y_S} \quad (29)$$

where X_S and Y_S are the force components along the x and y axes of the boat respectively, and K_S and N_S are the moments around the x and z axes. x_{CE} and z_{CE} are the x and z coordinates of the center of effort of the sails (CE). The thrust force coefficient C_X is expressed as positive for the forward direction and the side force coefficient C_Y is positive for both port and starboard directions. It should be noted that the coordinates are given in the body axes system. Therefore, when the boat heels the Y_S force component is not in the horizontal plane but is normal to the mast. The aerodynamic forces acting on the mast and rigging are included in the measured sail forces.

5.2.1 Measurement system of aerodynamic performance and sail shape

The sail dynamometer system is composed of a rigid aluminum frame and four load cells. The frame is separated structurally from the hull and connected to it by the load cells. The general arrangement of the dynamometer frame is given in Fig.10. The load cells are numbered in the figure. Two of these are 1-component load cells and the others are 2-component ones. Hence, these load cells form a 6-component dynamometer system, and their outputs can be transformed to the forces and moments about the boat axes using a calibration matrix. All rig components such as mast, chain plates, winches, lead blocks, etc. are attached to the aluminum frame. The under deck portion of the mast is held by the frame, and the other rig components are attached to the frame through the deck holes. The data acquisition system and calibration method for the *Fujin* were described by Masuyama et al. (1997a and 1997b).

The sail shape was recorded using pairs of CCD cameras. The lower part of the mainsail was photographed using the CCD camera pair designated A in Fig.11. These were located at the mast top, 50 cm transversely from each side of the mast. The upper part of the mainsail was photographed using a portable video camera from below the boom. The lower part of the jib was photographed using the camera pair designated B in Fig.11, which were located at the intersection point of the forestay and the mast, 10 cm transversely from each side of the mast. The upper part of the jib was photographed using a portable video camera from inside the bow hatch. For measuring convenience, horizontal stripes were drawn on the mainsail and jib at heights of 10, 20, 40, 60 and 80% of each sail. The sail shape images were analyzed using the sail shape analyzing software, SSA-2D, developed by Armonicos Co. Fig.12(a) shows an example of processed image of the mainsail using the SSA-2D. This software calculates the curvature of the sail section by marking several points of the sail stripe and the reference line on the PC display, and indicates the parameters such as chord length, maximum draft, maximum draft position, entry angle at the luff, i.e., leading edge, and exit angle at the leech, i.e., trailing edge, as shown in Fig.12(b). The apparent wind speed (AWS) and apparent wind angle (AWA) are measured by an anemometer attached on the "Bow unit" as shown in Fig.11. This unit post can rotate freely to maintain its vertical attitude when the boat heels in order to measure the wind data in the horizontal plane. The height of the anemometer coincides with the geometric center of effort (GCE) of the sail plan. The wind speed and wind angle sensors were calibrated by wind tunnel tests in advance and the calibration equations were obtained. The *Fujin* also has motion measuring instruments such

as an Optical Fiber Gyroscope (roll and pitch angles), a Flux Gate Compass (heading angle), a Differential type GPS receiver, a speedometer (velocity in the x direction) and a potentiometer for rudder angle. These data are recorded by an onboard computer simultaneously with the data from the load cells.

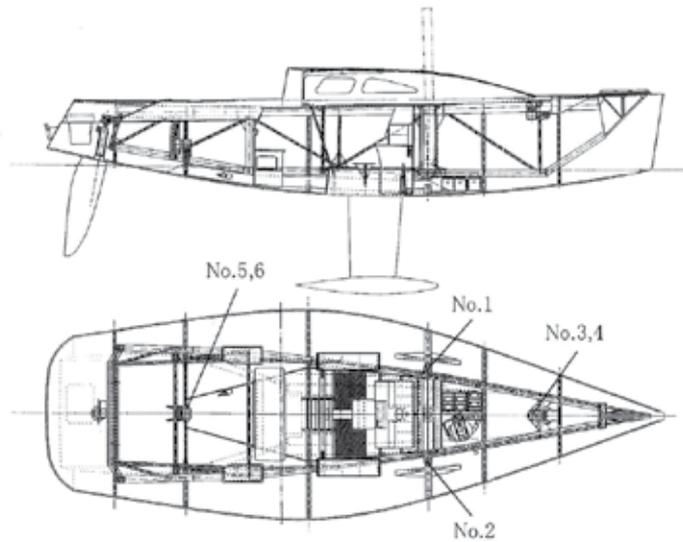


Fig. 10. General arrangement of dynamometer frame in *Fujin*.

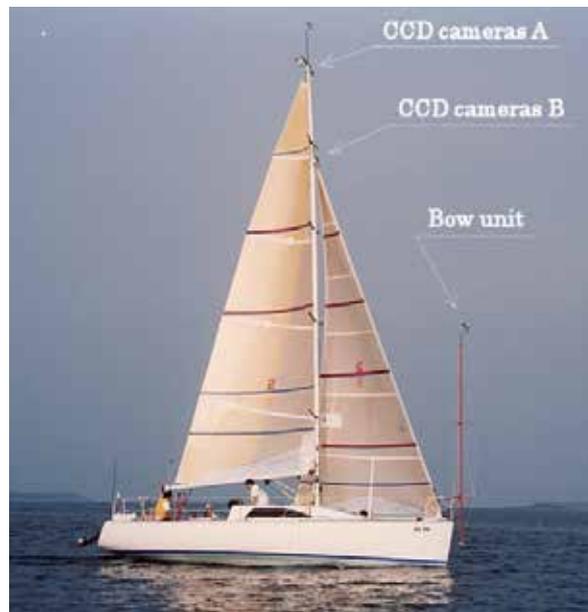


Fig. 11. Sea trial condition in light wind with 130% jib.

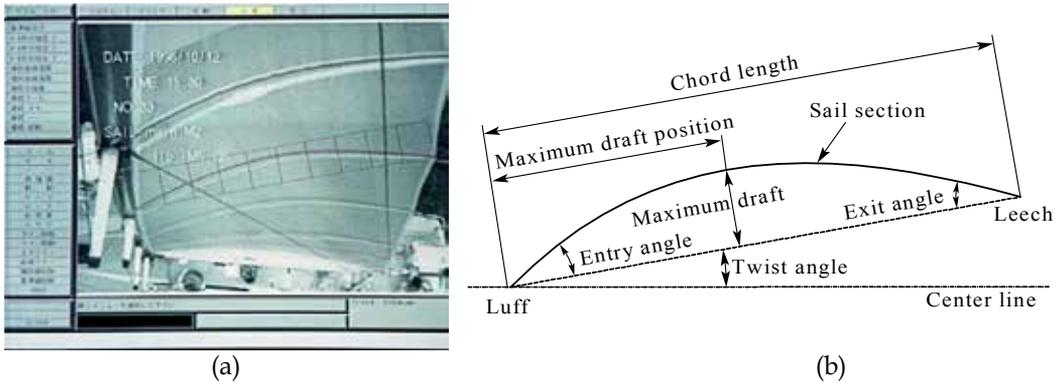


Fig. 12. (a) Example of a processed image of the mainsail using SSA-2D. (b) Measured sail shape parameters.

5.2.2 Test condition and error analysis

The sea tests were performed on Nanao Bay off the Noto Peninsula. The bay is approximately eight nautical miles from east to west and five from north to south. The bay is surrounded by low hills, and the mouth connecting it to the Japan Sea is narrow. Therefore, there is little tidal current in the bay, and the wave heights are relatively low even though the wind can be strong. The close-hauled tests were conducted over the apparent wind angle (AWA) range of 20 to 40 degrees, and the apparent wind speed (AWS) range of 5 to 11m/s. The effect of the AWA, and the draft and twist of the mainsail on the sail performance were measured. Data sampling was started when the sailing condition was considered to be in steady state. The sampling rate for the data acquisition system was set at 10Hz. Data sampling was continued for 90 seconds, and during this time the sail shapes were recorded using the CCD cameras. The boat was steered carefully during this time. However, the measured data contained some variation due to wind fluctuation and wave reflection on the hull. Therefore the steady state values for the aerodynamic coefficients were obtained by averaging the data over a 30 to 60 seconds period, in which the AWA was closer to the target value than during the whole 90 second period. For these tests if the range of deviation of AWA exceeded ± 5 degrees, the results were discarded. All of the measured coefficients are plotted with error bars indicating the range of deviation over the averaging period.

5.3 Comparison between experimental and calculated results

In this chapter, the experimental results and the calculated results for the following cases will be compared:

- a. Variation with apparent wind angle
- b. Variation with mainsail twist angle

For each series, first the sail coefficients: C_L , C_D , C_X and C_Y , and the coordinates of x_{CE} and z_{CE} are given. Then, the calculated the sail surface pressure and streamlines using the RANS-based CFD are presented for two typical cases in each series. Finally, the shapes and three-dimensional coordinates of the sails are tabulated for each case corresponding to those where the RANS-based CFD results are given.

5.3.1 Variation with apparent wind angle

Fig.13 shows the performance variation for the mainsail and 130% jib configuration as a function of AWA. In the figure the solid symbols indicate the experimental results and the open symbols indicate the calculated results using the VLM and the RANS-based CFD. For the experimental results, both data from the starboard (Stbd) and port tack (Port) are shown. All of the measured coefficients are plotted with error bars indicating the range of deviation over the averaging period. There are some discrepancies between the data from each tack. During the experiments, efforts were made to remove this asymmetrical performance. However, the boat speed actually differed on each tack. It can be concluded that there was a slight asymmetry in the combination of the hull, keel, rudder and dynamometer frame. The numerical calculations were performed using the measured shape data. In order to avoid confusion when interpreting the figure, the calculated results are indicated only for the port tack. Therefore, the calculated and experimental points for the port tack correspond to each other.

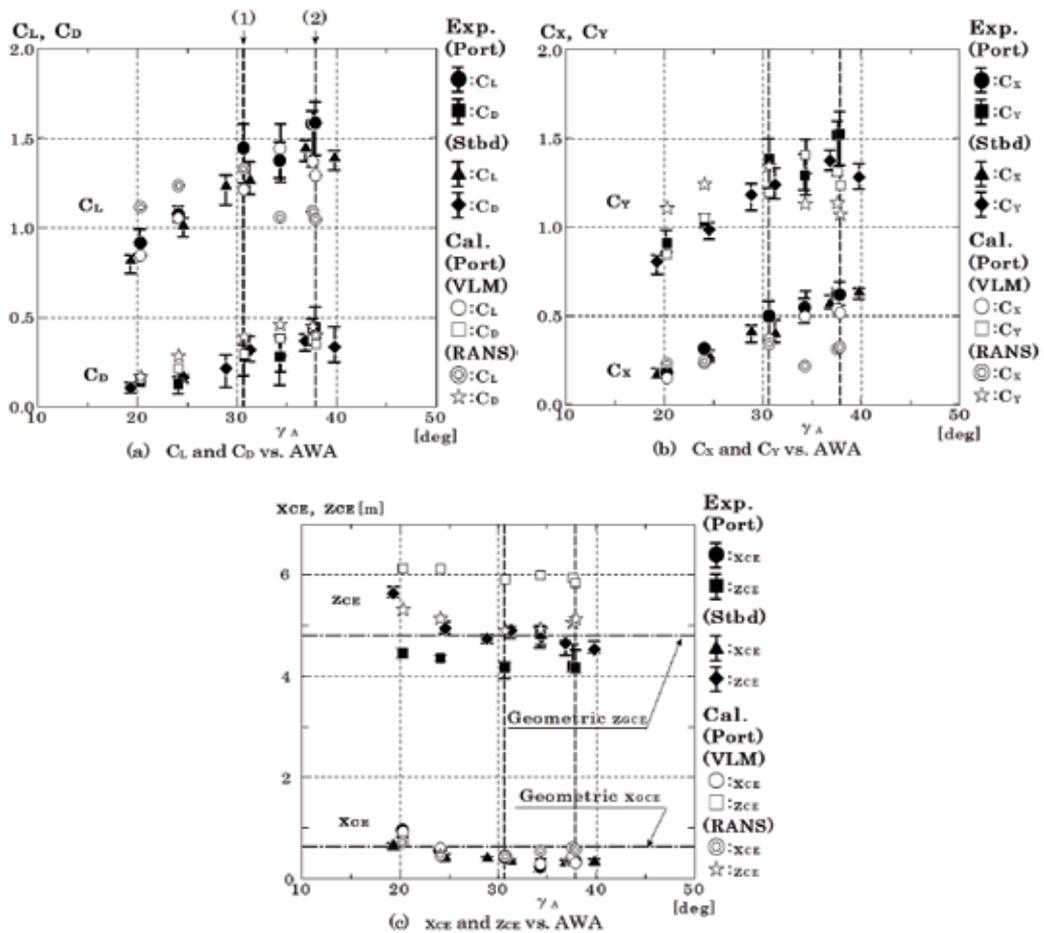


Fig. 13. Performance variation as a function of apparent wind angle (AWA) for mainsail and 130% jib.

In this figure, AWA ranges from 20.3 degrees to 37.9 degrees for the port tack. The former is the closest angle to the wind that was achieved, and the latter is typical of a close reaching condition, where the sail is eased for the power down mode. There is some scatter in the experimental data because this is made up from measurements taken with the sails trimmed in slightly different ways. The experimental value of C_L in Fig.13(a) varies with AWA from 0.91 to 1.58. For the close reaching condition, the sails were not well trimmed to satisfy the power down mode. A sample of measured sail sections at this condition is shown in a figure attached to Table 2(2). From the figure, it can be seen that both the mainsail and the jib are not eased sufficiently to correspond to the large AWA. This is the reason for the decrement in the measured lift curve slope of C_L at the range of AWA angles over about 35 degrees.

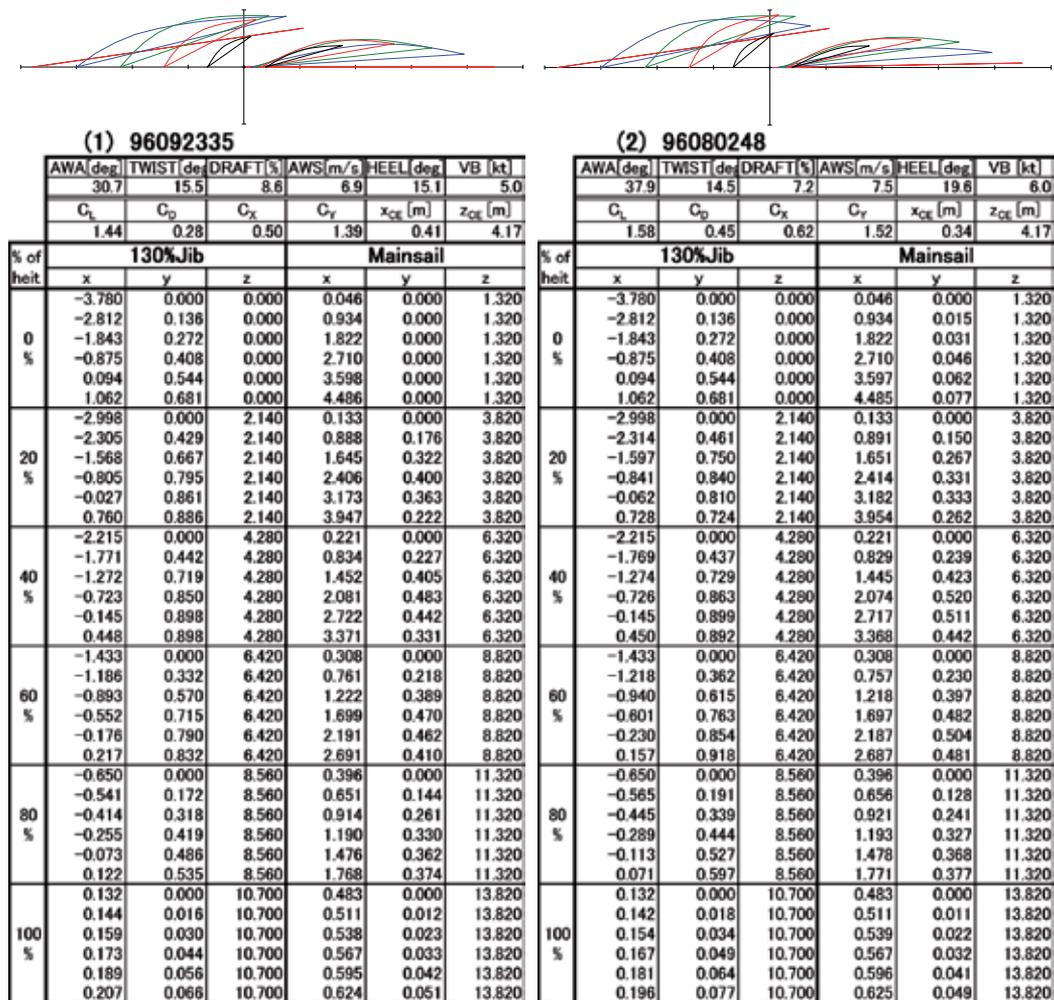
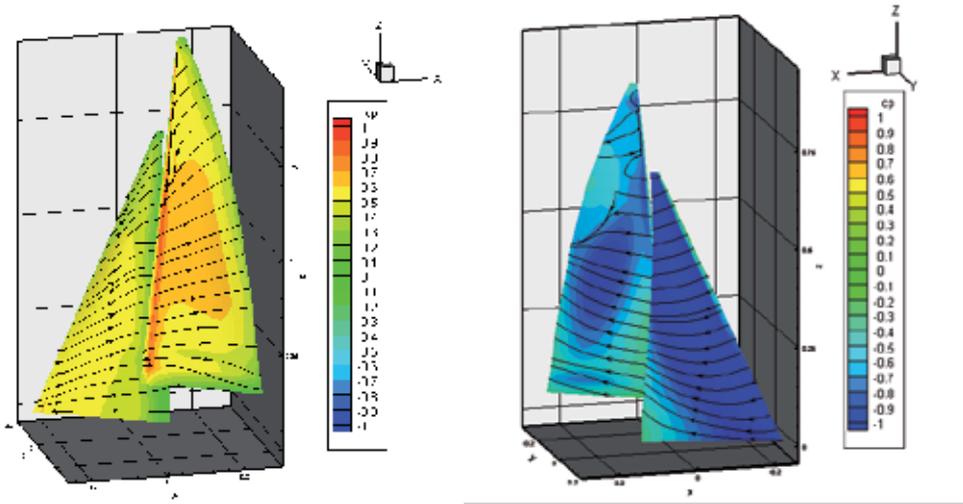


Table 2. Sail shapes, measured experimental data and three-dimensional coordinates of the sails for the cases of (1) 96092335 and (2) 96080248.

The calculated results for C_L using the VLM show good agreement with the experiments at AWA angles less than about 35 degrees. Over about 35 degrees, the calculated results are lower than the measured ones. This shows that the calculated results strongly indicate the effect of incorrect sail trimming. The results for C_L using the RANS-based CFD show the same trends with the experiments, but are slight higher than those from the experiments for AWA between 20 degrees to 30 degrees and lower for AWA greater than 30 degrees. In particular, the decrease in C_L for AWA values greater than 30 degrees is considerably large. This will be discussed later with the calculated sail surface pressure and streamlines. The calculated results for C_D slightly over predict those from the experiments. Fig.13(c) shows the coordinates of the center of effort of the sails. The x and z coordinates of the geometric center of effort (x_{GCE} and z_{GCE}) are 0.63m aft and 4.80m above the origin, which are indicated by alternate long and short dashed lines in the figure. It is seen that both the experimental and the calculated coordinates of x_{CE} are near x_{GCE} and move slightly forward with increasing AWA. Unfortunately, there is a wide scatter in the experimental values of z_{CE} . This is thought to be because the measured K_s moment contains a large component from the mass of the dynamometer frame and rigging (659kg). This moment was subtracted from the measurement, taking into account the measured heel angle. If there is a slight error in the position of center of gravity of the dynamometer frame, or in the measured heel angle, the error in the calculated moment will be large. However, though there is a scatter in the measured data, it can be seen that z_{CE} is decreasing as AWA increases. The trends in the movement of both x_{CE} and z_{CE} as functions of AWA might be caused by the decrement of force acting on the aft and upper parts of the sails due to the loosening of main and jib sheets with increasing AWA. The calculated results for z_{CE} obtained using the RANS-based CFD show the same trend as the experiments. On the other hand, the calculated results using VLM are considerably higher than the experimental ones. This might be caused by over estimation of the force acting on the upper portion of the mainsail. In this area, since the jib is not overlapping, flow separation may occur easily. However, the VLM does not take flow separation into account.

Figures 14(1) and 14(2) show the calculated results of the sail surface pressure and streamlines using RANS-based CFD. Fig.14(1) indicates the case of experiment ID 96092335 (AWA= 30.7deg.), and 14(2) indicates ID 96080248 (AWA= 37.9deg.). These data correspond to the plotted points on the vertical dotted lines (1) and (2) in Fig.13. In Fig.14, the left and right diagrams correspond to the port and starboard sides, i.e., pressure and suction sides, respectively. In 14(1), although slight flow separation on the suction side of mainsail is seen, the streamlines of both sides run smoothly. On the other hand, in 14(2), considerable flow separation is occurring, in particular, on the suction side of jib. This is the main reason for the reduction of C_L value in the RANS-based CFD calculation at (2) in Fig.13(a). The shapes and three-dimensional coordinates of the sails are given in Table 2. The numbered (1) and (2) tables correspond to the cases of experiment ID 96092335 and ID 96080248, respectively. These also correspond to the calculated results shown in Fig.14. The figures described above the tables show the sail section profiles at 0, 20, 40, 60 and 80% of the sail height. The dimensions of these three-dimensional coordinates are given in the tables including 100% height section data. The positive direction of the x coordinate is aft. The four lines at the top of the tables are the measured values for the wind and sail trim conditions, the boat attitude and the sail performance coefficients.

(1)



(2)

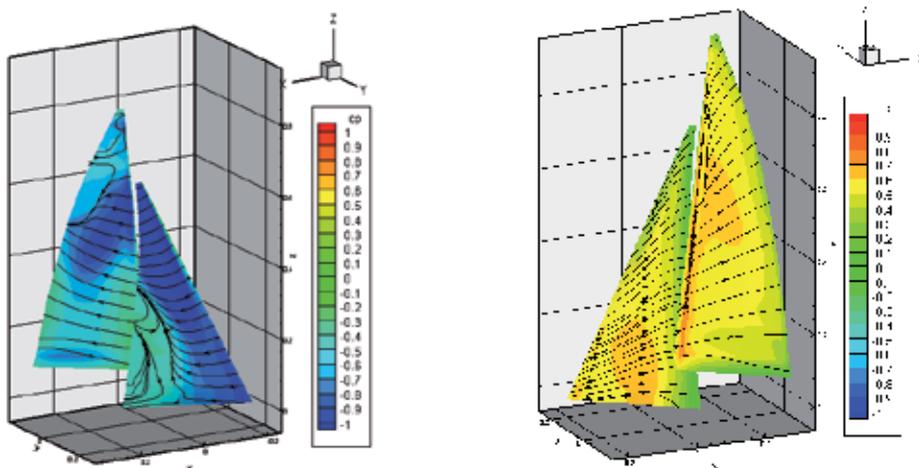


Fig. 14. (1) Surface pressure and streamlines obtained by RANS-based CFD at experimental ID 96092335 (AWA=30.7 deg.) and (2) ID 96080248 (AWA=37.9 deg.).

5.3.2 Variation with mainsail twist angle

Fig.15 shows the performance variation for the mainsail and 130% jib configuration as a function of mainsail twist angle. The mainsail twist was changed by varying the main sheet tension. The boom angle was kept parallel with the boat centerline by moving the main sheet traveler. The experiment was performed for an average value of AWA of 30 ± 2 degrees and mean draft at around 10%. The jib shape was fixed. The twist angle is defined as the angle between the boom line and section chord line at 80% height. In the figure, the

twist angle ranges from 4.5 degrees to 24.9 degrees for the port tack. Varying the twist angle by 20.4 degrees, results in the value of C_X in Fig. 15(b) changing from 0.33 to 0.39 (18%), and the value of C_Y changing from 1.16 to 1.39 (20%). It can be seen that the maximum C_X occurs at a twist angle of around 15 degrees. The considerable decrease in C_Y with increasing twist angle is also worth noticing. In this case, the calculated results of both VLM and RANS-based CFD for C_X and C_Y , and C_L and C_D correspond to the measured values very well.

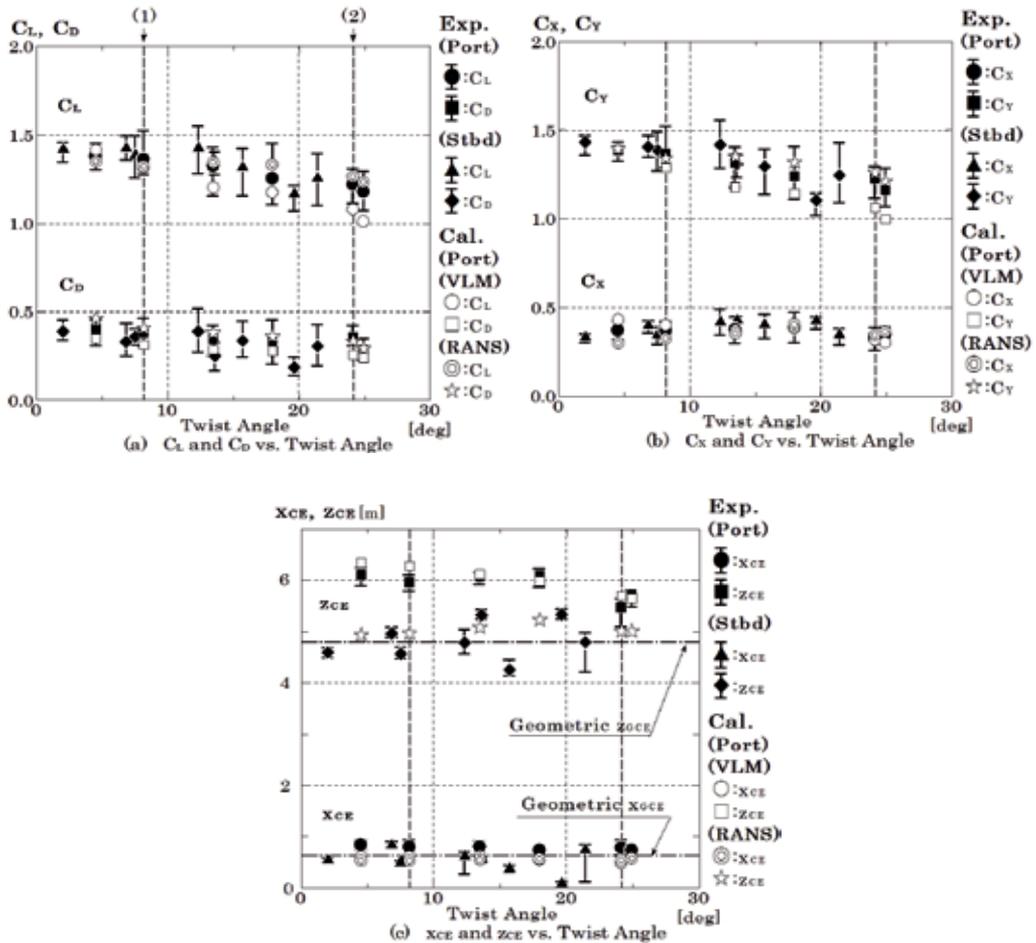


Fig. 15. Performance variation as a function of mainsail twist angle for mainsail and 130% jib.

Figures 16(1) and 16(2) show the calculated results using RANS-based CFD. Fig.16(1) corresponds to ID 97072213 (twist angle = 8.2 deg.), and 16(2) corresponds to ID 97072218 (twist angle = 24.1 deg.). It can be seen in Fig.16(1) that the streamlines on the upper part of the suction side of the mainsail for the smaller twist angle, show considerable flow

separation. This is caused by the large angle of attack at the upper part of the sail due to the small twist angle. On the other hand, for the higher twist angle shown in Fig.16(2), there is a low negative pressure area at the luff on the suction side of mainsail due to the small angle of attack. This is what causes the considerable reduction in the calculated value for C_x at (2) in Fig.15(b). Table 3 shows the shapes and three-dimensional coordinates of the sails for cases (1) and (2), which correspond to the calculated results shown in Fig.16. Further measured data and comparison with the numerical calculations are described by Masuyama et al. (2007 and 2009)

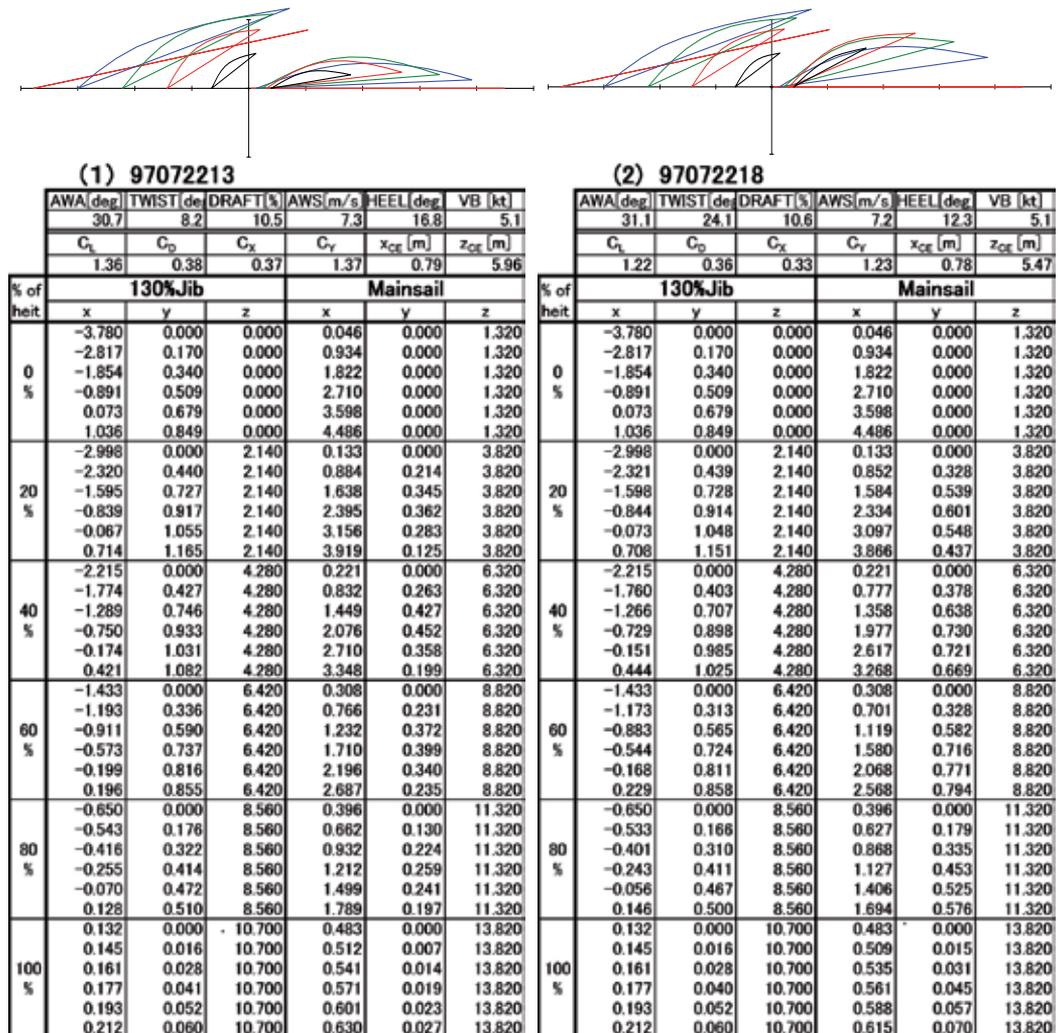
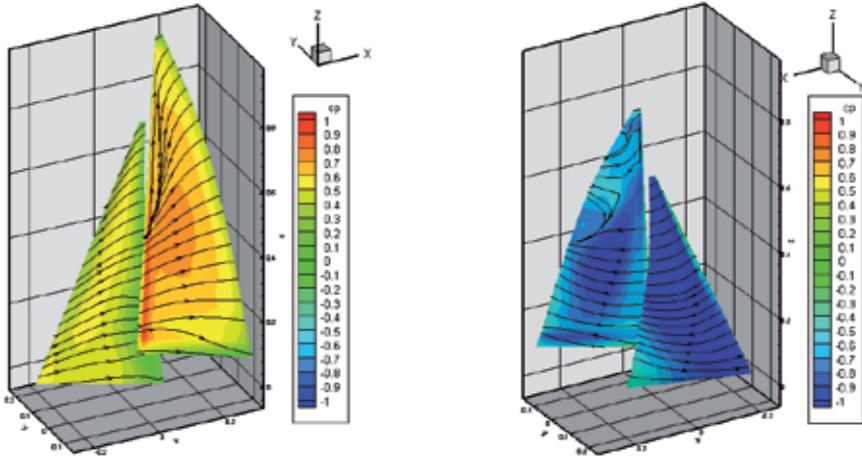


Table 3. Sail shapes, measured experimental data and three-dimensional coordinates of the sails for the cases of (1) 97072213 and (2) 97072218.

(1)



(2)

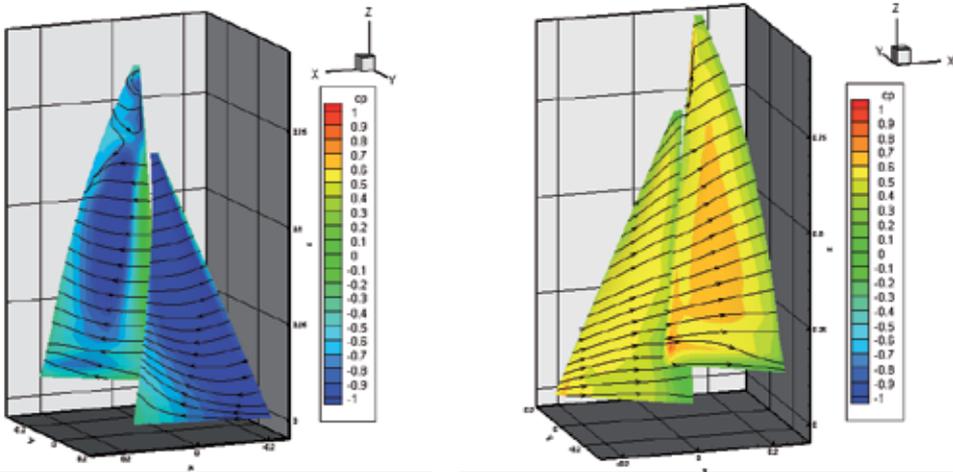


Fig. 16. (1) Surface pressure and streamlines obtained by RANS-based CFD at experimental ID 97072213 (twist angle= 8.2 deg.) and (2) ID 97072218 (twist angle= 24.1 deg.).

6. Discussion of numerical calculation methods

The flow is dominated by multiple-lifting-surface aerodynamic interactions. For larger AWA values, in particular, a large-scale flow separation exists on the leeward side of the sails. In general, there is complex vortex generation in the wake, especially near the top and bottom of the sails, i.e., tip vortices are generated and are influenced by the boundary layer flows on the sails. The resultant aerodynamic forces are mostly dominated by the pressure component, whereas the contribution of the frictional component is generally small. The accurate prediction of the boundary layer flows on the sails and the three-dimensional flow separation, associated with the abovementioned vortex generation, are big challenges for

RANS-based CFD. The geometrical complexity is also another significant challenge to RANS-based CFD. The accuracy in the prediction of the CE is of great interest, in association with the correct prediction of the above-mentioned three-dimensional flow separation.

Through the analyses of the multiblock RANS-based CFD, it appears that the overall trends of the flow and the aerodynamic forces measured in the experiments are fairly well predicted by the present computations. It is also seen that the multiblock domain decomposition considered here is very effective for the present mainsail and jib configurations. The automatic gridding scheme used successfully generates high-quality structured grids for the various sail geometries, AWA, and heel angles considered in the present study. Although there are advantages to a structured grid system for high-resolution in the boundary layer flow, building a grid in this fashion is difficult to apply to complex geometries. This problem appears to be resolved by the present scheme.

The Vortex Lattice method is, on the other hand, a convenient tool to predict the lift and induced drag acting on the sail accurately at the apparent wind angles less than about 35 degrees. The computational time of the method is about a few minutes for one calculation condition. The longitudinal coordinates, or x coordinate, of the center of effort of the sail can also be calculated with accuracy by the Vortex Lattice method, however, the estimated vertical coordinate, or z coordinate, of the center of effort by the Vortex Lattice method is considerably higher than the experimental ones. This may be caused by the fact that the flow at the upper portion of mainsail is easily separated because of the absence of jib overlapping, while the flow separation cannot be taken into account in the Vortex Lattice method

7. Conclusion

The sail performance analysis of sailing yacht was carried out by using numerical calculations and experiments. Focus in the present manuscript is especially on the upwind sailing condition. The sails considered here are IMS type, and the shapes and performance were measured by using the sail dynamometer boat *Fujin*. The measured sail flying shapes were used by numerical analysis, where two CFD methods developed by the authors were used, i.e. a multiblock RANS-based CFD method by Tahara and a VLM-based CFD method by Fukasawa. It appears that the overall trends of the flow and the aerodynamic forces measured in the experiments are fairly well predicted by the present computations; and at the same time, the present sail performance database based on the full scale onboard measurements are very useful for validation study of numerical methods. As compared to maturity of VLM, that for RANS-based CFD is still in underway but the future prospect is shown promising, especially for capability in predicting separation flow field where viscous effects of fluid are significant. The authors believe that our sail performance database associated with accurate sail flying shape measurements will be able to contribute to the further development of more advanced CFD methods.

Although details are not described in the present manuscript, our current effort is directed toward the more challenging problem, i.e., extension of the present work for the downwind sailing condition. Since the onboard sail shape measurement system of *Fujin* is incompetent for the spinnaker measurement due to its balloon shape, the sail shapes and performance are measured using wind tunnel equipment, and such activities are already in progress. The sail shapes are recorded using digital cameras and processed to obtain 3D coordinates using solid shape analyzer software, which provides 3D coordinates from digital photographs taken from several different directions. Importantly, the sail shapes and the sail forces and

moment acting on the model are measured simultaneously. The numerical simulation by using a RANS-based CFD method is also in progress. Along with integration with the aforementioned sail design and performance prediction software *AAF*, more details of our work will be reported in our future publications.

8. Acknowledgments

The sail dynamometer boat *Fujin* was built for sail tests for the Japanese America's Cup entry by a Grant-in-Aid from the Nippon Foundation and the authors would like to thank the Nippon Foundation for providing them this invaluable tool. The authors wish to express their thanks to Yamaha Motor Co. Ltd. for building *Fujin* and to North Sails Japan Co. for making the sails. We would like to thank Dr. Martin Renilson for his valuable discussions and comments on this article. We would also like to thank Mr. H. Mitsui, the harbormaster of the Anamizu Bay Seminar House of the Kanazawa Institute of Technology, for his assistance with the sea trials. Help with the sea trials given by graduate and undergraduate students of the Kanazawa Institute of Technology is also gratefully acknowledged. The graduate students were Masaya Miyagawa, Takashi Hasegawa, and Munehiko Ogihara. Finally, we would like to acknowledge the effort by Mr. Naotoshi Maeda who carried out most of the RANS-CFD simulations as a part of his work on M.S. thesis at Osaka Prefecture University.

9. References

- Baldwin, B.S. & Lomax, H. (1978). Thin Layer Approximation and Algebraic Model for Separated Turbulent Flows, *AIAA Paper*, 78-257, *AIAA 16th Aerospace Science Meeting*, Reno, NV., USA, 1978
- Falkner, V.M. (1943). The Calculation of Aerodynamic Loading on Surface of Any Shape, *Aeronautical Research Committee Reports and Memoranda No.1910*, Ministry of Aircraft Production, London, UK
- Falkner, V.M. (1946). The Accuracy of Calculations Based on Vortex Lattice Theory, *British Aeronautical Research Council*, No.9621
- Fukasawa, T. & Katori, M. (1993). Numerical Approach to Aeroelastic Responses of Three-Dimensional Flexible Sails, *Proceedings of 11th Chesapeake Sailing Yacht Symposium*, pp.87-105, Annapolis, USA, January 29-30, 1993
- Fukasawa, T. (1993). Aeroelastic Transient Response of 3-Dimensional Flexible Sail, *Proceedings of International Conference on Aero-Hydroelasticity*, pp.57-62, Beijing, China, October 18-21, 1993,
- Hansen, H., Jackson, P. & Hochkirch, K. (2003). Comparison of Wind Tunnel and Full-scale Aerodynamic Sail Force, *International Journal of Small Craft Technology (IJSCT)*, Vol. 145 Part B1, (2003), pp. 23-31
- Hochkirch, K. & Brandt, H. (1999). Fullscale Hydrodynamic Force Measurement on the Berlin Sailing Dynamometer, *Proceedings of SNAME 14th Chesapeake Sailing Yacht Symposium*, pp.33-44, Annapolis, Maryland, USA, January, 1999
- Katori, M. (2009). Advanced Aero Flow Software Manual, *North Sails Japan*, December 2009.
- Krebber, B. & Hochkirch, K. (2006). Numerical Investigation on the Effects of Trim for a Yacht Rig, *Proceedings of 2nd High Performance Yacht Design Conference*, pp.13-21, Auckland, New Zealand, February, 2006

- Masuyama, Y. & Fukasawa T. (1997a). Full Scale Measurement of Sail Force and the Validation of Numerical Calculation Method, *Proceedings of SNAME 13th Chesapeake Sailing Yacht Symposium*, pp. 23-36, Annapolis, Maryland, USA, January, 1997
- Masuyama, Y., Fukasawa T. & Kitasaki T. (1997b). Investigations on Sail Forces by Full Scale Measurement and Numerical Calculation (Part 1: Steady Sailing Performance), *Journal of Society of Naval Architects of Japan*, Vol. 181, (June 1997), pp.1-13, ISSN 0514-8499, (in Japanese)
- Masuyama, Y., Tahara, Y., Fukasawa, T. & Maeda, N. (2007). Database of Sail Shapes vs. Sail Performance and Validation of Numerical Calculation for Upwind Condition, *Proceedings of SNAME 18th Chesapeake Sailing Yacht Symposium*, pp.11-31, Annapolis, Maryland, USA, March, 2007
- Masuyama, Y., Tahara, Y., Fukasawa, T. & Maeda, N. (2009). Database of Sail Shapes versus Sail Performance and Validation of Numerical Calculation for the Upwind Condition, *Journal of Marine Science and Technology*, JASNAOE, Vol. 14, No. 2, (2009), pp.137-160, ISSN 0948-4280
- Milgram, J. H., Peters, D. B. & Eckhouse, D.N., N. (1993). Modeling IACC Sail Forces by Combining Measurements with CFD, *Proceedings of SNAME 11th Chesapeake Sailing Yacht Symposium*, pp.65-73, Annapolis, Maryland, USA, January, 1993
- Pistolesi, E. (1937). Betrachtungen über die gegenseitige Beeinflussung von Tragflügelsystemen. *Gesammelte Vorträge der Hauptversammlung 1937 der LilienthalGesellschaft*, pp.214-219
- Tahara, Y. & Stern, F. (1996). A Large-Domain Approach for Calculating Ship Boundary Layers and Wakes for Nonzero Froude Number, *Journal of Computational Physics*, Vol. 127, (1996), pp. 398-411
- Tahara, Y. (1995). An Application of Two-Layer k- ϵ Model to Ship Flow Computation, *Journal of Society of Naval Architects of Japan*, Vol. 177, (1995), pp. 161-176
- Tahara, Y. (2008). A Reynolds-Averaged Navier-Stokes Equation Solver for Prediction of Ship Viscous Flow with Free Surface Effects, *Proceedings of NAPA User Meeting 2008*, Helsinki, Finland, May, 2008.
- Tahara, Y., Tohyama, S. & Katsui, T. (2006a). CFD-Based Multi-Objective Optimization Method for Ship Design, *International Journal of Numerical Methods in Fluids*, Vol. 52, (2006), pp. 449-527
- Tahara, Y., Wilson, R., Carrica, P. & Stern, F. (2006b). RANS Simulation of a Container Ship Using a Single-Phase Level Set Method with Overset Grids and Prognosis for Extension to Self-Propulsion Simulator, *Journal of Marine Science and Technology*, Vol. 11, No. 4, (2006), pp. 209-228
- Thompson, J.F., Warsi, Z.U.A., & Mastin, C.W. (1985). Numerical Grid Generation, *Elsevier Science Publishing Co., Inc.*, NY.

Part 2

Multiphase Flow, Structures, and Gases

A Magneto-Fluid-Dynamic Model and Computational Solving Methodologies for Aerospace Applications

Francesco Battista¹, Tommaso Misuri² and Mariano Andrenucci²

¹*CIRA SCpA Italian Aerospace Research Centre*

²*ALTA SpA
Italy*

1. Introduction

Computational plasma physics is concerned primarily with the study of the evolution of plasma by means of computer simulation. The main task of this computational branch is to develop methods able to obtain a better understanding of plasma physics. Therefore a close contact to theoretical plasma physics and numerical methods is necessary. Ideally computational plasma physics acts as a pathfinder to guide scientific and technical development and to connect experiment and theory. To build a valid computer simulation program means to devise a model which is sufficiently detailed to reproduce faithfully the most important physical effects, with a computational effort sustainable by modern computers in reasonable time (Dandy, 1993).

Computational models have played an important role in the development of plasma physics since the beginning of the computer age. Advances in our understanding of many plasma phenomena like magnetohydrodynamic instabilities, micro-instabilities, transport, wave propagation, etc. have gone hand-in-hand with the increased computational power available to researchers. Several trends are evident in how computer modelling is carried out: the models are becoming increasingly complex, for example, by coupling separate computer codes together. This allows for more realistic modelling of the plasma. Presently several efforts are carried out in different countries to develop plasma numerical tools for several applications such as fusion, electric propulsion, active control over hypersonic vehicles: these efforts lead to a growing experience in CMFD field (see Park et al. 1999, Kenneth et al 1998, Taku and Atsushi 2004, Cristofolini et al 2007, Miura and Groth 2007, MacCormack 2007, Yalim 2001, Giordano and D'Ambrosio 2004, Battista 2009).

The chapter presented was carried out in the context of a research activity motivated by renewed interest in investigating the influence that electromagnetic fields can exert on the thermal and pressure loads imposed on a body invested by a high energetic flow. In this regard, spacecraft thermal protection and the opportunity to use active control surfaces during planetary (re)entry represent the driving engineering applications. The contents of the study should be considered, to a certain extent, a systematic re-examination of past work complemented with somewhat innovative ideas.

So, in this chapter, methodologies for plasma modelling have been developed and then implemented and tested into a numerical code EMC3NS, developed in the frame of this

work. This one has been used and tested primarily to support the design of experiments in problems of active flow control over bodies in an electromagnetic field in order to modify shock waves position, and consequently the aerothermal environment.

These kind of experiments are not so common in the scientific community especially using air as working gas, thus the lack of experimental data and lesson learned in this field make numerical computation play an important role. So the use of valid computer simulation is crucial for a correct design of a flow control experiment set up diagnostics and to understand main physical phenomena related to this kind of problems.

The use of magnetic fields to control external flows is not new (for instance see Bush (1958), Resler & Sears (1958)); recent computational and experimental technologies have moved this approach from mere possibility to real application, highlighting the considerable advantage that may be gained in re-entry operation.

The physical ideas are incorporated into the so-called magneto-fluid-dynamic interaction concept: global body forces can be applied to a weakly ionized plasma using electromagnetic devices embedded in the vehicle. Therefore, there is a growing interest in using weakly ionized gases (plasmas) and electric and magnetic fields in high-speed aerodynamics. Wave and viscous drag reduction, thrust vectoring, reduction of heat fluxes, sonic boom mitigation, boundary-layer and turbulent transition control, flow turning and compression, on-board power generation, and scramjet inlet control are among plasma and MHD technologies that can potentially enhance performance and significantly change the design of supersonic and hypersonic vehicles and thrusters.

Meanwhile, despite many studies devoted to these new technologies, a number of fundamental issues have not been adequately addressed. Any plasma created in gas flow and interacting with electric and magnetic fields would result in gas heating. This heating can certainly have an effect on the flow and, in some cases, can be used advantageously. However, a more challenging issue is whether significant non-thermal effects of plasma interaction with electric and magnetic fields can be used for high-speed flow control. In conventional MHD of highly conducting fluid, electric and magnetic effects give rise to ponderomotive force terms, which can be interpreted as gradients of electric and magnetic field pressures. These ponderomotive forces are successfully used for plasma containment in fusion devices and also play an important role in astrophysics. One might hope that these forces can also be used for control of high-speed flow of ionized air. However, the great importance of ponderomotive forces in fusion and astrophysical plasmas is due to the fact that those plasmas are fully, or almost fully, ionized and, therefore, are highly conductive. In contrast, high speed air encountered in aerodynamics is not naturally ionized, even in boundary layers and behind shocks if the flight Mach number is below about 8, due to the low static temperature. So in this case studies are necessary to set up technologies and methodologies to control the flow in such weak ionized regime.

2. Governing equations

In the present section the set of equations to be solved is exposed; as previously described, the choice of a model, and thus the acceptance of all the underlying hypotheses, determines the nature of the results that numerical computations provide. The analysis has been carried out in order to provide a set of equations suitable for the problematic exposed in the introduction, with a limited computational cost and that have a structure easily adaptable

with classical aero-thermo-dynamic codes (e.g. H3NS developed by CIRA) that are nearly in the same field of application. For these scopes three set of equations have been analyzed: multi-fluid equations (Wagner 1998), MHD equations (Helander 2006), and MFD equations for a single fluid composed by more than one species (Giordano 2002).

The first set of equations (three fluid equations), where each species (charged and not) is considered as a single fluid interacting with the other through collisions, is one of the more appropriate way to describe plasma flows. However, even if the collision process is well described, it requires a large computational cost because of the large number of equations to be solved (especially in presence of more chemical species (Giordano 2002)). The MHD set of equations describes the motion and the electrodynamics of an electrically neutral but electrically conducting fluid. These assumptions do not permit to have any information about the species present in the flowfield; moreover diffusive transport and charge separation are excluded by this model. These drawbacks are critical in the exclusion of this approach for our purposes. The MFD equations for a single multispecies fluid is the set of equations that more than the others is suitable for our scope since it does not loose information about species (charged and not), and furthermore the collisions process can be modelled by transport coefficients. Besides this model is not computationally heavy and the equation structure is easily adaptable to the one used in typical finite volume codes for aerothermodynamics application. The model that will be used and described further is able to consider a multi-temperature gas with vibrational temperature and electronic temperature in non equilibrium with the translational temperature.

The assumptions used to write down the system of equations are the following:

1. the velocity is not relativistic $u \ll c$;
2. the phenomena under consideration are slow enough that $\frac{\partial}{\partial t} \sim v/L \ll \omega_{ce} = \sqrt{\frac{ne^2}{m_e \epsilon_0}}$;
3. the plasma is sufficiently collisional $v_{ii} \gg \sqrt{\left(\frac{m_i}{m_e}\right)} \frac{\partial}{\partial t}$,
4. the Debye length and the ion gyro radius are small $\lambda_D \ll L, \rho_i \ll L$.

The equations are written in the following; mass balance equations of the components in chemical non equilibrium conditions yield,

$$\frac{\partial}{\partial t} \int_V \rho_i dV + \int_S (\rho_i \underline{V} + \underline{J}_{m_i}) \cdot \underline{ds} = \int_{dV} \Omega_i dV \quad (2.1)$$

where \underline{J}_{m_i} ¹ are the components mass diffusive fluxes and the Ω_i ² are the production rate of the i-th specie,

¹ $\underline{J}_{m_i} = \text{diffusive flux } i\text{-component} = \frac{\rho}{p} \frac{M_i}{M^2} \sum_{j=1}^n M_j \underline{D}_{ij} \cdot \nabla c_j + \underline{D}_i^{me} \cdot (\underline{E} + \underline{V} \times \underline{B}) - \frac{1}{T} \underline{D}_i^T \cdot \nabla T$

² $\Omega_i = \text{prod_term} = \sum_{k=1}^r \xi_k \nu_{ik} M_i$

Pointing out the charge density as $\rho_C = -eN_A \sum_{i=1}^n \frac{\rho_i}{M_i} \sigma_{iS}$ (where σ_{iS} is equal to (+1) for electrons, to (-1) for ions and to (0) for neutral components) and summing up for all the charged species, then the electric charge balance equation is obtained

$$\frac{\partial}{\partial t} \int_V \rho_C dV + \int_S (\underline{J}_q + \rho_C \underline{V}) \cdot d\underline{S} = 0 \quad (2.2)$$

where the current is the sum of convection and conduction current $\underline{J} = \rho_C \underline{V} + \underline{J}_Q$ ³; the momentum equation is

$$\frac{\partial}{\partial t} \int_V \rho \underline{V} dV + \int_S (\rho \underline{V} \underline{V} + p \underline{I} - \underline{\tau}) \cdot d\underline{S} = \int_V (\rho_C \underline{E} + \underline{J} \times \underline{B}) dV \quad (2.3)$$

where $\bar{\tau}_{i,j} = \mu \left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i} \right) - \frac{2}{3} (\nabla \cdot \bar{V}) \delta_{ij}$ is the viscous stress tensor and the total energy equation reads:

$$\frac{\partial}{\partial t} \int_V \rho e_m dV + \int_S ((\rho e_m + p) \underline{V} + \underline{J}_Q - \underline{\tau} \cdot \underline{V}) \cdot d\underline{S} = \int_V (\underline{J} \cdot \underline{E}) dV + \int_{\partial V} \underline{q}_r \cdot d\underline{S} \quad (2.4)$$

where the total energy is $e_m = \frac{k_B T}{\gamma - 1} + \frac{mv^2}{2} + e_{vibr}$, the heat flux due to radiation is

$$\underline{q}_r \cdot \underline{n} = \int_0^\infty \int_0^{2\pi} \int_0^{2\pi} I_\nu(\vartheta, \phi) \cos(\vartheta) \sin(\vartheta) d\vartheta d\phi d\nu.$$

The vibrational energy equations:

$$\frac{\partial}{\partial t} \int_V \rho e_{vj} dV + \int_{\partial V} \underline{n} \cdot (\rho e_{vj} \underline{V}) ds = \int_V \rho \dot{e}_{vj} dV \quad (2.5)$$

provides the terms in the Eqn.(2.4).

The opportunity of considering electronic temperature in the vibrational energy equations is currently object of study. The production terms of the vibrational energy equation are evaluated through the classical Landau Teller non equilibrium equation (Vincenti & Krouger 1967).

2.1 Chemical model

Given a generic set of N_r chemical reactions,

$$\sum_{i=1}^{N_s} \nu_{ir}' \chi_i \Leftrightarrow \sum_{i=1}^{N_s} \nu_{ir}'' \chi_i \quad (2.6)$$

³ $\underline{J}_Q = \sum_{j=1}^n \bar{\lambda}_{ej}^p \cdot \nabla c_j + \bar{\lambda}_e \cdot (\underline{E} + \underline{V} \times \underline{B})$

where ν'_{ir} and ν''_{ir} are the stoichiometric coefficients for the r-th reaction and χ_i are the chemical species involved in the reactions, the production rate Ω_i of species i can be written as

$$\Omega_i = \rho M_i \sum_{r=1}^{N_r} (\nu''_{ir} - \nu'_{ir}) \left(k_{fr} \prod_{i=1}^{N_s} q^{v'_{ir}} - k_{br} \prod_{i=1}^{N_s} q^{v''_{ir}} \right) \quad (i = 1, \dots, N_s - 1) \quad (2.7)$$

where k_{fr} and k_{br} are the forward and backward rate constants of the reaction r, which are assumed to follow the Arrhenius temperature law:

$$k_{fr} = A_{fr} T^{\beta_{fr}} \cdot e^{\frac{-E_{fr}}{RT}} \quad (2.8)$$

$$k_{br} = A_{br} T^{\beta_{br}} \cdot e^{\frac{-E_{br}}{RT}} \quad (2.9)$$

where the pre-exponential factors A_i , the temperature exponents β_r , and the activation energies E_i depend upon the adopted kinetic scheme.

The backward rate constants are related to the forward rate constants through the equilibrium constants $K_{eq,r}$ i.e.:

$$K_{eq,r} = \frac{k_{fr}}{k_{br}} \quad (2.10)$$

In some cases the equilibrium coefficients can be given as a function of the temperature, and the backward reaction rate can be calculated directly from the Eqn. (2.9).

For argon the reaction coefficients that have been considered in order to account for ionization are:

n.	Reaction	A [mol-cm-s-K]	β	Ea [J/mol]
1	AR \leftrightarrow AR+ + e-	1.52e18	0.50	1520000
2	AR + e \leftrightarrow AR+ + 2e-	2.50e34	-3.8	1510793

where the reaction rate coefficients of the first reaction have been derived from:

$$k = A T^{\beta} \cdot e^{\frac{-E_a}{3RT}} = P d_{12}^2 N_A^2 \sqrt{\frac{8\pi k_B T}{m_{12}}} e^{\frac{-E_a}{3RT}}$$

this formula is consequent from the expression of binary collision rate where m_{12} is the reduced mass, d_{12} is the mean diameter, k_B is the Boltzmann constant and P is the steric factor. The second reaction rate coefficient has been taken from (Gokcet 2004).

For air ionization one of the most important reactions governing the distribution of free electrons present in high temperature air plasmas is (Dunn & Lordi 1969):



In this work, in order to model air ionization several 7-species models and 11 species models have been used (Gupta et al. 1989, Park 1993, Kang et al.1973) and compared also with a

model generated using (Park 1993) and (Kang et al.1973) derived in order to better fit experimental data w.r.t. the original 11 species models. The scheme adopted is reported hereinafter specifying for each reaction the respective source.

n.	Reaction	A [mol-cm-s-K]	B	Ea [J/mol]	Ref
1	$O_2 \leftrightarrow 2 O$	1.0e22	-1.5	494706.8	P
2	$N_2 \leftrightarrow 2 N$	3.0e22	-1.6	941190.08	P
3	$NO \leftrightarrow N + O$	1.1e17	0.0	627737.2	P
4	$NO + O \leftrightarrow N + O_2$	8.4e12	0.0	161715.08	P
5	$N_2 + O \leftrightarrow N + NO$	6.4e17	1.0	319257.96	P
6	$N + O \Rightarrow NO^+ + e^-$	1.6e06	1.5	265892.8	K
7	$N + O \Leftarrow NO^+ + e^-$	6.7e21	-1.5	0	K
8	$O + O \Rightarrow O_2^+ + e^-$	1.6e17	-0.98	671803.52	K
9	$O_2^+ + e^- \Leftarrow O + O$	8.0e21	-1.5	0	K
10	$N + N \Rightarrow N_2^+ + e^-$	1.4e13	0	671803.52	K
11	$N_2^+ + e^- \Leftarrow N + N$	8.0e21	-1.5	0	K
10	$O_2 + N_2 \Rightarrow NO + NO^+ + e^-$	1.38e20	-1.84	1172330.4	K
12	$NO + NO^+ + e^- \Leftarrow O_2 + N_2$	8.0e21	-2.5	0	K
13	$NO + N_2 \Rightarrow NO^+ + N_2 + e^-$	2.2e15	-0.35	897955.2	K
14	$NO^+ + N_2 + e^- \Leftarrow NO + N_2$	2.2e26	-2.5	0.0	K
15	$NO + O_2 \Rightarrow NO^+ + O_2 + e^-$	8.8e15	-0.35	897955.2	K
16	$NO^+ + O_2 + e^- \Leftarrow NO + O_2$	8.8e26	-2.5	0	K
17	$O + e^- \leftrightarrow O^+ + e^- + e^-$	3.9e33	-3.78	1317832.4	P
18	$N + e^- \leftrightarrow N^+ + e^- + e^-$	8.8e26	-3.82	1401807.84	P
19	$O + O_2^+ \Rightarrow O^+ + O_2$	2.92e18	-1.11	232803.2	P
20	$O^+ + O_2 \Leftarrow O + O_2^+$	7.80e11	0.5	0.0	P
21	$N_2 + N^+ \Rightarrow N + N_2^+$	2.02e18	0.81	108087.2	P
22	$N + N_2^+ \Leftarrow N_2 + N^+$	7.80e11	0.5	0.0	P
23	$O + NO^+ \Rightarrow NO + O^+$	3.63e15	-0.6	422371.52	P
24	$NO + O^+ \Leftarrow O + NO^+$	1.50e13	0.0	0.0	P
25	$N + NO^+ \Rightarrow NO + N^+$	1.00e19	-0.93	507178.4	P
26	$NO + N^+ \Leftarrow N + NO^+$	7.80e11	0.0	0.0	P
27	$O_2 + NO^+ \Rightarrow NO + O_2^+$	1.80e15	-0.6	274375.2	P
28	$NO + O_2^+ \Leftarrow O_2 + NO^+$	1.50e13	0.0	0.0	P
29	$O + NO^+ \Rightarrow N^+ + O_2$	1.34e13	0.31	642453.69	P
30	$N^+ + O_2 \Leftarrow O + NO^+$	1.00e14	0.0	0.0	P

Table 1. Kinetic scheme obtained using (Park 1993) and (Kang et al.1973)

2.2 Thermodynamic model

By definition a thermally perfect gas is characterized by specific heats that depend only on temperature.

Hence, for a mixture of thermally perfect gases:

$$dh = c_p(T)dT \tag{2.11}$$

$$p = \rho RT \tag{2.12}$$

where R and c_p are, respectively, the gas constant and the specific heat at constant pressure of the mixture. They are obtained as weighted average of the gas constants and specific heats of the single species i -th:

$$R = \sum_i Y_i R_i \quad c_p(T) = \sum_i Y_i c_{pi}(T) \tag{2.13}$$

The single species specific heats and enthalpies are computed by using the (Gordon and McBride 1971) polynomial fits:

$$\frac{c_p}{R} = a_1 + a_2 T + a_3 T^2 + a_4 T^3 + a_5 T^4 \tag{2.14}$$

$$\frac{h}{RT} = a_1 + \frac{a_2}{2} T + \frac{a_3}{3} T^2 + \frac{a_4}{4} T^3 + \frac{a_5}{5} T^4 + \frac{a_6}{T} \tag{2.15}$$

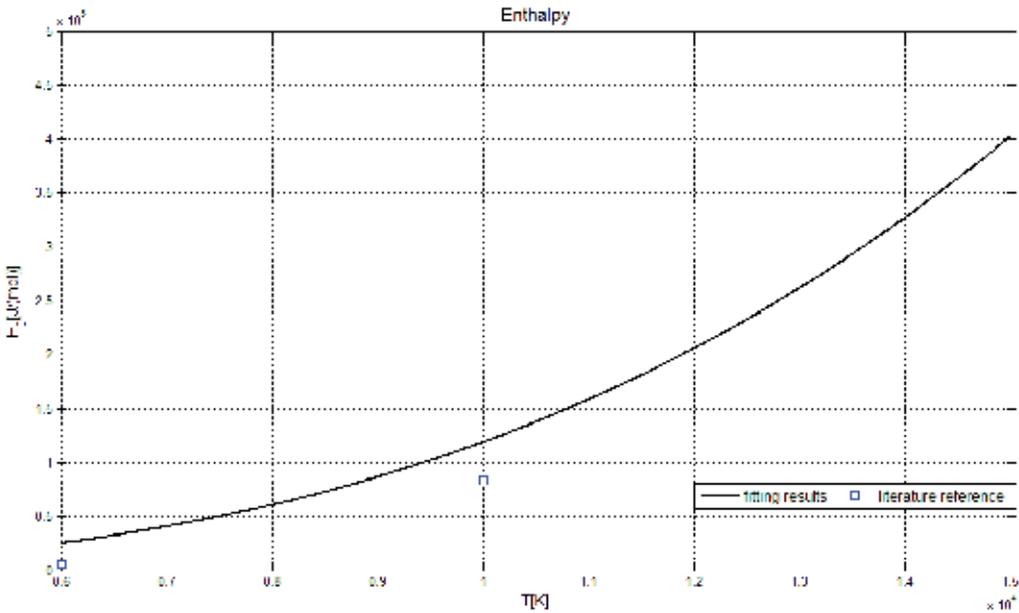


Fig. 1. Enthalpy fit extension for temperatures ranging from 6000K and 15000K.

These fits are written for gases with the maximum temperature of 6000K, so it has been necessary to extend the enthalpy fits to values of temperature typical of ionization problem (up to 15000K). Here is briefly reported the fitting procedure for argon: from literature (Drellishak et al. 1963) thermodynamic properties are derived for high temperatures then the coefficients are found using a least squares fitting procedure. The results for Argon enthalpy are shown in Fig. 1. Thermodynamic properties for high temperature air, used to find fit coefficients, have been found in (Hans & Heims 1968).

2.3 Transport model

Another critical point in the simulation of high enthalpy flows is the determination of the transport coefficients. In fact, the widely used Sutherland law is suitable only at low temperatures, while more complex models must be used when temperature exceeds 1000 K. An interesting way to compute transport coefficients is based on the calculation of the collision integral starting from intermolecular potentials knowledge (Hirshfield et al. 1954) and it leads to the following expressions for the diffusivities between the i -th and the j -th species and also for conductivity and viscosity with different sets of constants given in Gupta et al. 1958, Hans & Heims 1968. In this way computational efficiency is maximized since transport coefficients are computed only once at the start of each calculation.

$$\ln D_{ij} = \sum_{n=1}^N d_{n,ij} (\ln T)^{n-1} \quad \ln \lambda_i = \sum_{n=1}^N b_{n,i} (\ln T)^{n-1} \quad \ln \mu_i = \sum_{n=1}^N a_{n,i} (\ln T)^{n-1}$$

Once derived the single species properties, total conductivity and viscosity are calculated using Wilke (Hirshfielder 1954) formulas. The mixture diffusion coefficient for species i -th is obtained as reported in Bird (1954).

Polynomial coefficients for ionized species have been found fitting data from (Capitelli et al. 2000); the results of the fitting procedure for NO^+ , N^+ and O^+ are shown for viscosity in Fig. 2.

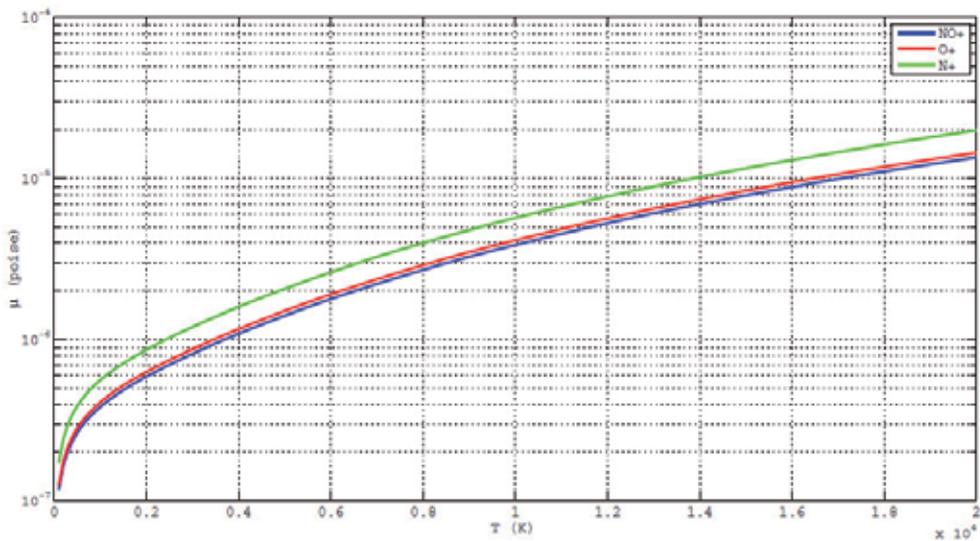


Fig. 2. Viscosity in function of temperature for charged species NO^+ , N^+ and O^+

2.4 Electrical conductivity

In order to evaluate electrical conductivity its classical expression has been written down:

$$\sigma_0 = \frac{n_e e^2}{m_e \nu_{eH}}$$

where ν_{eH} is the collision frequency between electrons and heavy particles measured in Hz. By substituting the constant values the expression becomes:

$$\sigma = 2.81794 \cdot 10^{-8} \frac{n_e}{\nu_{eH}}$$

Considering collision frequency as $\nu_{eH} = n_H Q_{eH} \langle u_e \rangle$ and substituting the electron mean velocity (under Maxwellian distribution hypothesis) the following expression yield:

$$\sigma = 4.535909 \cdot 10^{-12} \alpha \frac{T_e^{-1/2}}{Q_{eH}} \quad (2.16)$$

where $\alpha \approx n_e/n_H$ is the degree of ionization. For what concerns Q_{eH} an expression for air could be found in (Baum 1965).

2.5 Maxwell equations

Maxwell equations have to be solved together with the fluid equations in order to solve the full MFD system: equations are the following:

$$\int_V (\nabla \times \underline{E}) dV = - \frac{\partial}{\partial t} \int_V \underline{B} dV \quad (2.17)$$

$$\int_V (\varepsilon_0 c^2 \nabla \times \underline{B}) dV = \int_V \left(\underline{J} + \varepsilon_0 \frac{\partial \underline{E}}{\partial t} \right) dV \quad (2.18)$$

they can be rewritten in conservative form,

$$\frac{\partial}{\partial t} \int_V \underline{B} dV + \int_{\delta V} (\underline{n} \times \underline{E}) dS = 0 \quad (2.19)$$

$$\int_V \left(\frac{\partial \underline{E}}{\partial t} \right) dV - \int_{\delta V} (c^2 \underline{n} \times \underline{B}) dS + \int_V \frac{\underline{J}}{\varepsilon_0} dV = 0 \quad (2.20)$$

2.6 Collision modeling

In order to correctly describe collision processes that in this kind of equations determine the transport properties of the fluid, a momentum equation could be written down for each spatial direction, neglecting time derivatives as:

$$\begin{cases} u_x = \mu E_x - \frac{D}{n} \frac{\partial n}{\partial x} + \frac{(u_y \omega_{cz} - u_z \omega_{cy})}{\nu} \\ u_y = \mu E_y - \frac{D}{n} \frac{\partial n}{\partial y} + \frac{(u_z \omega_{cx} - u_x \omega_{cz})}{\nu} \\ u_z = \mu E_z - \frac{D}{n} \frac{\partial n}{\partial z} + \frac{(u_x \omega_{cy} - u_y \omega_{cx})}{\nu} \end{cases} \quad (2.21)$$

In the previous equations the cyclotronic vector $\overline{\omega_c} = \frac{q}{m_{ie}} [B_x B_y B_z]^T$ has been considered that divided by ν gives the Hall parameter vector that represent the Hall parameter along all the space dimension: $[\beta_x \beta_y \beta_z]^T$.

Decoupling the equations multiplying by ne , and reminding from previous section that $\sigma_0 = \frac{ne^2}{m\nu}$, with a matrix pivoting technique it can be obtained the relation between electric field and current, i.e.

$$\begin{pmatrix} j_x \\ j_y \\ j_z \end{pmatrix} = \sigma_0 \begin{bmatrix} \frac{1+\beta_x^2}{1+\beta_x^2+\beta_y^2+\beta_z^2} & \frac{\beta_x\beta_y+\beta_z}{1+\beta_x^2+\beta_y^2+\beta_z^2} & \frac{\beta_x\beta_z-\beta_y}{1+\beta_x^2+\beta_y^2+\beta_z^2} \\ \frac{\beta_y\beta_x-\beta_z}{1+\beta_x^2+\beta_y^2+\beta_z^2} & \frac{1+\beta_y^2}{1+\beta_x^2+\beta_y^2+\beta_z^2} & \frac{\beta_y\beta_z+\beta_x}{1+\beta_x^2+\beta_y^2+\beta_z^2} \\ \frac{\beta_z\beta_x+\beta_y}{1+\beta_x^2+\beta_y^2+\beta_z^2} & \frac{\beta_z\beta_y-\beta_x}{1+\beta_x^2+\beta_y^2+\beta_z^2} & \frac{1+\beta_z^2}{1+\beta_x^2+\beta_y^2+\beta_z^2} \end{bmatrix} \begin{pmatrix} E_x - \frac{Dm\nu}{n^2e^2} \frac{\partial n}{\partial x} \\ E_y - \frac{Dm\nu}{n^2e^2} \frac{\partial n}{\partial y} \\ E_z - \frac{Dm\nu}{n^2e^2} \frac{\partial n}{\partial z} \end{pmatrix} \quad (2.22)$$

From Eqn. (2.22) the tensor $\overline{\lambda_e}$ is directly derived.

3. Numerical discretization of the full system and solving methodologies

In the previous parts the fundamental equations governing an MFD plasma flow have been described. They constitute a system of 11 scalar equations plus one scalar equation for each species considered and one scalar equation for each vibrating species considered. In the following the numerical strategy and methods adopted to solve the systems will be exposed. All the equations briefly reported in the previous section have the following conservative form:

$$\frac{\partial}{\partial t} \int_V (W) dV + \int_{\partial V} (\underline{F} \cdot \underline{n}) dS = \int_V \underline{\Omega} dV \quad (3.1)$$

so they can be simply discretized following a finite volume approach:

$$V \frac{dW}{dt} + \sum_{\beta=1,6} (\underline{F} \cdot \underline{n} \cdot \Delta S)_\beta = V \underline{\Omega} \quad (3.2)$$

The unknown vector (W), the flux vector (\underline{F}) and the source term vector ($\underline{\Omega}$) used in the finite volume discretization are reported below.

$$\underline{W} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ \rho e_m \\ B_x \\ B_y \\ B_z \\ E_x \\ E_y \\ E_z \\ \rho_1 \\ \dots \\ \dots \\ \rho_{n-1} \\ \rho e_{v,1} \\ \dots \\ \rho e_{v,Nv} \end{pmatrix} \quad \underline{F_{inv,x}} = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ (p + \rho e_m)u \\ 0 \\ -E_z \\ E_y \\ 0 \\ +c^2 B_z \\ -c^2 B_y \\ \rho_1 u \\ \dots \\ \dots \\ \rho_{n-1} u \\ \rho e_{v1} u \\ \dots \\ \rho e_{vNv} u \end{pmatrix} \quad \underline{F_{inv,y}} = \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ \rho vw \\ (p + \rho e_m)v \\ E_z \\ 0 \\ -E_x \\ -c^2 B_z \\ 0 \\ c^2 B_x \\ \rho_1 v \\ \dots \\ \dots \\ \rho_{n-1} v \\ \rho e_{v1} v \\ \dots \\ \rho e_{vNv} v \end{pmatrix} \quad \underline{F_{inv,z}} = \begin{pmatrix} \rho w \\ \rho vw \\ \rho w^2 + p \\ (p + \rho e_m)w \\ -E_y \\ E_x \\ 0 \\ c^2 B_y \\ -c^2 B_x \\ 0 \\ \rho_1 w \\ \dots \\ \dots \\ \rho_{n-1} w \\ \rho e_{v1} w \\ \dots \\ \rho e_{vNv} w \end{pmatrix}$$

$$\underline{\Omega} = \begin{pmatrix} 0 \\ \rho_c E_x + J_y B_z - J_z B_y \\ \rho_c E_y + J_z B_x - J_x B_z \\ \rho_c E_z + J_x B_y - J_y B_x \\ J_x E_x + J_y E_y + J_z E_z \\ 0 \\ 0 \\ 0 \\ J_x / \epsilon_0 \\ J_y / \epsilon_0 \\ J_z / \epsilon_0 \\ \Omega_1 \\ \dots \\ \dots \\ \Omega_{n-1} \\ \Omega_{v1} \\ \dots \\ \Omega_{Nv} \end{pmatrix} = \begin{pmatrix} 0 \\ \rho_c E_x + J_{qy} B_z + \rho_c v B_z - J_{qz} B_y - \rho_c w B_z \\ \rho_c E_y + J_{qz} B_x + \rho_c w B_y - J_{qx} B_z - \rho_c u B_z \\ \rho_c E_z + J_{qx} B_y + \rho_c u B_y - J_{qy} B_x - \rho_c v B_x \\ J_{qx} E_x + \rho_c u E_x + J_{qy} E_y + \rho_c v E_y + J_{qz} E_z + \rho_c w E_z \\ 0 \\ 0 \\ 0 \\ J_{qx} + \rho_c u / \epsilon_0 \\ J_{qy} + \rho_c v / \epsilon_0 \\ J_{qz} + \rho_c w / \epsilon_0 \\ \Omega_1 \\ \dots \\ \dots \\ \Omega_{n-1} \\ \Omega_{v1} \\ \dots \\ \Omega_{Nv} \end{pmatrix}$$

$$\begin{aligned}
 \underline{F}_{visc,x} = & \begin{pmatrix} 0 \\ \tau_{xx} \\ \tau_{xy} \\ \tau_{xz} \\ u\tau_{xx} + v\tau_{xy} + w\tau_{xz} + \sum_{i=1}^n J_{Ux,i} + q_{r,x} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ J_{x,1} \\ \dots \\ J_{x,n-1} \\ q_{vx,1} + e_{v1}J_{x,1} \\ \dots \\ q_{vx,N_v} + e_{v,N_v}J_{x,N_v} \end{pmatrix} & \underline{F}_{visc,y} = & \begin{pmatrix} 0 \\ \tau_{xy} \\ \tau_{yy} \\ \tau_{yz} \\ u\tau_{xy} + v\tau_{yy} + w\tau_{yz} + \sum_{i=1}^n J_{Uy,i} + q_{r,y} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ J_{y,1} \\ \dots \\ J_{y,n-1} \\ q_{vy,1} + e_{v1}J_{y,1} \\ \dots \\ q_{vy,N_v} + e_{v,N_v}J_{y,N_v} \end{pmatrix} & \underline{F}_{visc,z} = & \begin{pmatrix} 0 \\ \tau_{xz} \\ \tau_{yz} \\ \tau_{zz} \\ u\tau_{xz} + v\tau_{zy} + w\tau_{zz} + \sum_{i=1}^n J_{Uz,i} + q_{r,z} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ J_{z,1} \\ \dots \\ J_{z,n-1} \\ q_{vz,1} + e_{v1}J_{z,1} \\ \dots \\ q_{vz,N_v} + e_{v,N_v}J_{z,N_v} \end{pmatrix} & (3.3)
 \end{aligned}$$

3.1 Numerical solving strategy description: Weak coupling

From the vectorial form of the system of equations we have a synthetic view of what kind of system we have to solve. It is also clear that it includes terms of different physical nature, in particular it combines acoustic and electromagnetic terms. In the present analysis it has been chosen to split this system of equations into the two following sub-systems:

1. Fluid-Dynamics equations sub-system
2. Maxwell equations sub-system

This strategy, named from literature (Giordano and D'Ambrosio 2005) 'weak coupling', allows us to treat separately the Maxwell equations and then to "freeze" the resulting electromagnetic field and to solve the fluid-dynamic sub-system.

The main reason behind the choice of this procedure is the highly different characteristic speed of propagation between sound and light. The fluid-dynamic phenomena (accounted for in the first sub-system) propagate at the speed of sound, while the electromagnetic phenomena (described by the Maxwell equations) propagate at the speed of light. Since light travels at least one hundred thousand times faster than sound, it is advisable to split the main system. Moreover, the differences of signals propagation in space make crucial the adoption of different numerical methods to solve each sub-system. Since the equations are solved both in space and time, one of the most striking differences between these two numerical methods will be the choice of the integration time interval, dt . Much shorter intervals are requested when solving Maxwell equations, because of the higher propagation speed for the phenomena involved. The numerical method for the Maxwell equations solver will be investigated in the next section. Another issue to be faced with is how these two sub-systems communicate each other. It is apparent that they are coupled, because the electromagnetic fields appear on the right hand side of the momentum and energy conservation equations, while the fluid velocity and the electric transport properties are directly linked to the current density (which plays an important role in the second Maxwell equation). Therefore it is necessary to find a way to solve the sub-systems alternately. The algorithm to be followed is sketched in the figure below:

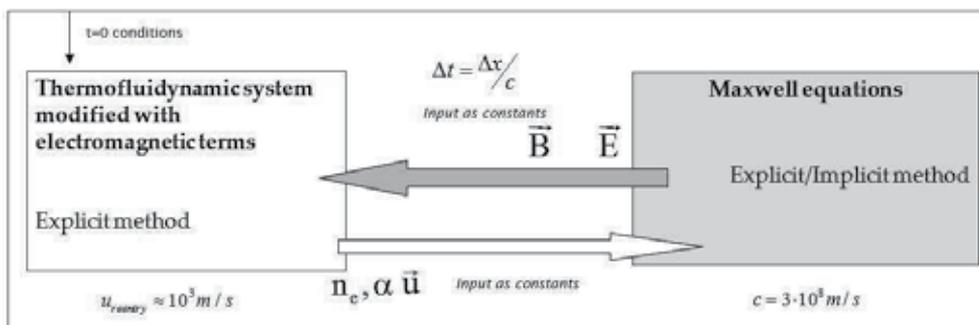


Fig. 3. System solving strategy

So, given the initial conditions (both for fluid-dynamic variables and electromagnetic fields) the solution procedure to solve the system is:

- The fluid-dynamic subsystem is solved.
- Maxwell equations are solved, using as input the value of current density which has been found at the end of step one. This leads to update the electromagnetic field intensity.
- The newly found values for E and B are sent back again to the fluid dynamic over to update the residual of the equations, which perform another step like the first one (keeping E and B constant).

These three steps make the solution march in time for an interval dt . They have to be repeated iteratively until the required total time of simulation is reached.

3.2 Fluid-dynamic subsystem: Starting platform description and upgrades through magneto-fluid-dynamic problems treatment

CIRA H3NS code numerical structure has been used as starting platform. H3NS is a RANS structured multi-block finite volume solver that allows for the treatment of a wide range of compressible fluid dynamics problems considering air as a working gas. The fluid can be treated considering air as a perfect gas or as a mixture of perfect gases in of thermo-chemical non equilibrium (5 species air and 3 vibrational temperatures). With respect to the numerical formulation, conservation equations are written in integral form, and discretized with a finite volume, cell centred technique. Eulerian fluxes are computed with a Flux Difference Splitting method (Borrelli and Pandolfi 1990). Second order formulation is obtained by means of an Essentially Non Oscillatory reconstruction of interface value. Viscous fluxes are computed with a classical centred scheme. Time integration is performed by employing an explicit multistage Runge-Kutta algorithm coupled with an implicit evaluation of the source terms.

More accurate description of the models implemented and validation tests in 2D and 3D hypersonic problems could be found in (Schettino et al. 2008), (Battista et al. 2007). In order to be able to treat magneto fluid dynamic problems the following upgrades have been implemented:

- Capability in the treatment of a generic multi-component reacting mixture of perfect gases including ionized species (Arrhenius formulation).

- Electromagnetic terms added at the equations implemented
 - Thermodynamic properties formulation in terms of Gordon-Mc Bride Polynomial fits. (Gordon and McBride 1971)
 - Extension of thermodynamic properties (specific heat, enthalpy and entropy) to higher temperatures.
 - Included ionized species transport properties.
 - Effects of electromagnetic field on transport accounted.
- Models considered for these upgrades have been widely discussed in previous Sections.

3.3 Numerical methodologies for Maxwell equations : Implicit and explicit Maxwell equation solver design

In this section the numerical methods used to solve Maxwell equations are discussed. So far, two methods have been employed, one of them is implicit, the other one explicit. Each method has its own strengths and its drawbacks. In the two paragraphs below, they are explained and compared. The equations to be discretized are the Maxwell equations in conservative form ((2.19), (2.20)), i.e.:

$$\frac{\partial}{\partial t} \int_V (\underline{W}^M) dV + \int_{\partial V} (\underline{F}^M) dS = \int_V \underline{\Omega}^M dV \quad (3.4)$$

With:

$$\underline{W}^M = \begin{pmatrix} B_x \\ B_y \\ B_z \\ E_x \\ E_y \\ E_z \end{pmatrix} \quad \underline{F}_x^M = \begin{pmatrix} 0 \\ -E_z \\ E_y \\ 0 \\ +c^2 B_z \\ -c^2 B_y \end{pmatrix} \quad \underline{F}_y^M = \begin{pmatrix} E_z \\ 0 \\ -E_x \\ -c^2 B_z \\ 0 \\ c^2 B_x \end{pmatrix} \quad \underline{F}_z^M = \begin{pmatrix} -E_y \\ E_x \\ 0 \\ c^2 B_y \\ -c^2 B_x \\ 0 \end{pmatrix} \quad \underline{\Omega}^M = \begin{bmatrix} 0 \\ 0 \\ 0 \\ J_{qx} + \rho_c u / \epsilon_0 \\ J_{qy} + \rho_c v / \epsilon_0 \\ J_{qz} + \rho_c w / \epsilon_0 \end{bmatrix} \quad (3.5)$$

The superscript “M” reminds that this method is being applied to the Maxwell equations.

3.3.1 Implicit methodology

The great strength of all implicit numerical schemes is their great stability. Using an implicit method it is possible to choose a larger integration time interval without compromising the stability of the method itself. This is most relevant when dealing with signals which travel so fast as the electromagnetic waves.

Explicit schemes make time advancement very simple, but often suffer severe restrictions on time step due to the loss of stability. On the other side, implicit schemes generally have much better numerical properties in term of stability and so allow a larger time step, but require the solution of a system of equations to perform time integration. When analyzing a 3D problem as in the present case, the matrix associated to the system is too large to be inverted in a reasonable computational time. This means that an iterative method has to be applied to solve the system. So, the time saved using a larger time step is actually lost in solving a large linear system of equations at each iteration. To introduce the implicit method, a 1D simple case is described. Then, the result obtained will be extended to the 3D

case. Discretizing equation (3.1), and looking to it is possible to write in a finite volume fashion:

$$\frac{\Delta W_N^M}{\Delta t}(A\Delta x) = \left((F_{IN}^M)^{K+1} - (F_{OUT}^M)^{K+1} \right) A + (\Omega_N^M)^{K+1/2} (A\Delta x) \quad (3.6)$$

Here the subscript “N” addresses the cell we are referring to, while the superscript “K” indicates the time interval considered.

F_{IN}^M and F_{OUT}^M can be written as:

$$F_{IN}^M = \frac{1}{2} (F_{N}^M + F_{N-1}^M) \quad F_{OUT}^M = \frac{1}{2} (F_{N}^M + F_{N+1}^M)$$

Considering again equation (3.6), expanding to the first order all the terms with superscript different than “K” (i.e. the terms that are not evaluated at the instant “k”) and dividing both sides by the surface area A, yields:

$$\begin{aligned} \frac{\Delta W_N}{\Delta t} \Delta x + \frac{1}{2} (F_{N+1} - F_{N-1}) + \left(\frac{1}{2} \left(\frac{\partial F_{N+1}}{\partial W_{N+1}} \right) \Delta W_{N+1} - \frac{1}{2} \left(\frac{\partial F_{N-1}}{\partial W_{N-1}} \right) \Delta W_{N-1} \right) = \\ (\Omega_N) \Delta x + \frac{1}{2} \left(\frac{\partial \Omega_N}{\partial W_N} \right) \Delta W_N \Delta x \end{aligned} \quad (3.7)$$

Here the superscript “k” has been dropped, since it is clear that each term in the above equation (3.7) is evaluated at the time “k”.

For a 3D analysis this procedure leads to the following equation (assuming to work on a “x,y,z” Cartesian grid, with cubic cells):

$$\begin{aligned} & \begin{matrix} [C] & [R] & [L] \\ \left[1 - \frac{\Delta t}{2} \left(\frac{\partial \Omega_N}{\partial W_N} \right) \right] \tilde{\Delta} W_N - \frac{1}{2} \left(\frac{\partial F_{N+1}^x}{\partial W_{N+1}} \right) \frac{\tilde{\Delta} t}{\Delta x} \tilde{\Delta} W_{N+1} + \frac{1}{2} \left(\frac{\partial F_{N-1}^x}{\partial W_{N-1}} \right) \frac{\tilde{\Delta} t}{\Delta x} \tilde{\Delta} W_{N-1} - \\ [R] & [LLL] & [RRR] \\ - \frac{1}{2} \left(\frac{\partial F_{N+n_x}^y}{\partial W_{N+n_x}} \right) \frac{\tilde{\Delta} t}{\Delta y} \tilde{\Delta} W_{N+1} + \frac{1}{2} \left(\frac{\partial F_{N-n_x}^y}{\partial W_{N-n_x}} \right) \frac{\tilde{\Delta} t}{\Delta y} \tilde{\Delta} W_{N-1} - \frac{1}{2} \left(\frac{\partial F_{N+n_x n_y}^z}{\partial W_{N+n_x n_y}} \right) \frac{\tilde{\Delta} t}{\Delta z} \tilde{\Delta} W_{N+1} + \\ [LLL] \\ + \frac{1}{2} \left(\frac{\partial F_{N-n_x n_y}^z}{\partial W_{N-n_x n_y}} \right) \frac{\tilde{\Delta} t}{\Delta z} \tilde{\Delta} W_{N-1} = - \frac{1}{2} (F_{N+1}^x - F_{N-1}^x) \frac{\tilde{\Delta} t}{\Delta x} - \\ - \frac{1}{2} (F_{N+n_x}^y - F_{N-n_x}^y) \frac{\tilde{\Delta} t}{\Delta y} - \frac{1}{2} (F_{N+n_x n_y}^z - F_{N-n_x n_y}^z) \frac{\tilde{\Delta} t}{\Delta z} + (\Omega_N) \tilde{\Delta} t \end{matrix} \end{aligned} \quad (3.8)$$

Δ : is used to indicate a spatial interval

$\tilde{\Delta}$: is used to indicate a time interval

n_x, n_y : are the numbers of cells along x and y apxis respectively

Although here the solution, $\tilde{\Delta}W$, can be easily obtained, the trouble with this method regards its stability. So the CFL (Courant-Friedrichs-Lewy) number must be chosen carefully in order to have an acceptable time step and to grant the method stability.

Considering the very high speed of signal propagation ($c \cong 3 \cdot 10^8$ m/s), this means that the time step must be extremely short. However, this second method turns out to be faster than the implicit one due to its simplicity. The second order explicit numerical method used in this work and reported in 1D formulation in Eqn.3.11 (again, N represent the space step and K the time step) has been compared with literature numerical methods like Lax-Friederichs and Lax-Wendroff methods (Chung 2010), for the solution of the classical advection equation test consisting in the propagation of a square and a smooth wave, using as conditions a wave speed of 60 m/s, a grid spacing (Δx) of 10^{-3} m and a time step (Δt) of 10^{-5} s.

$$q_N^{K+1} = \frac{1}{3} \left[q_N^K + q_{N-1}^K + q_{N+1}^K + \frac{1}{2} \frac{\Delta t}{\Delta x} (q_{N+1}^K + q_{N+2}^K - q_{N-1}^K - q_{N+2}^K) \right] \quad (3.12)$$

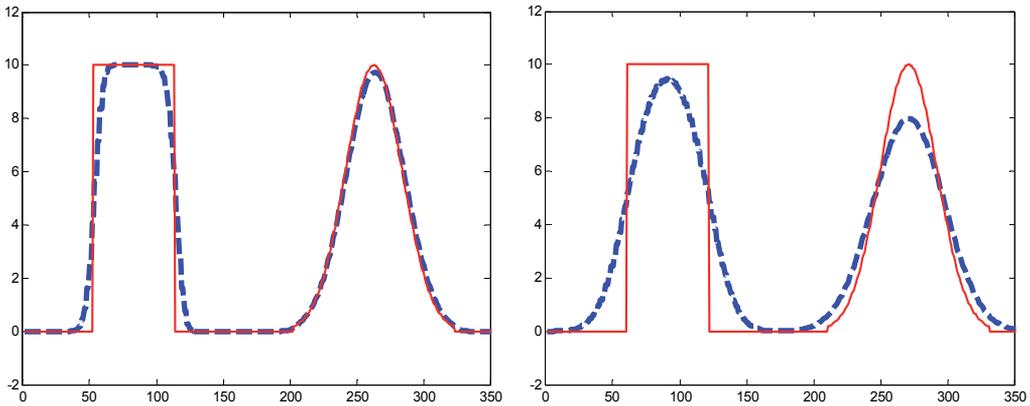


Fig. 4. Lax-Friederichs results after 40 timesteps (left) and 400 timesteps (right)

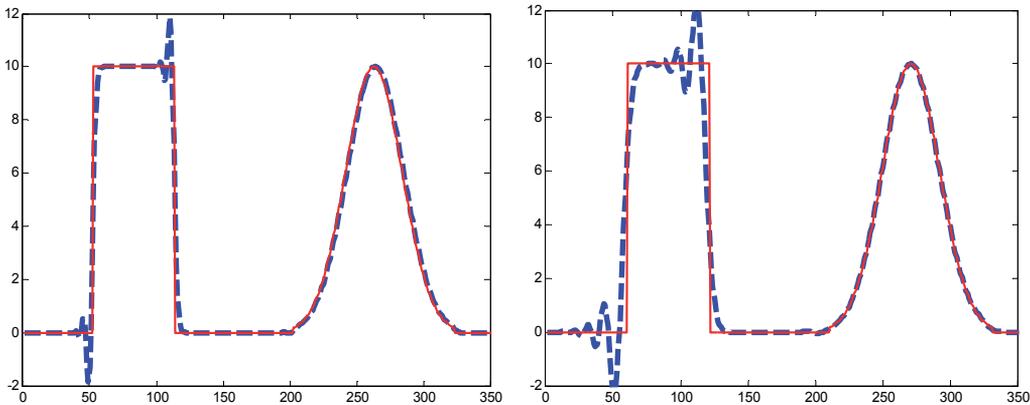


Fig. 5. Lax-Wendroff results after 40 timesteps (left) and 400 timesteps (right)

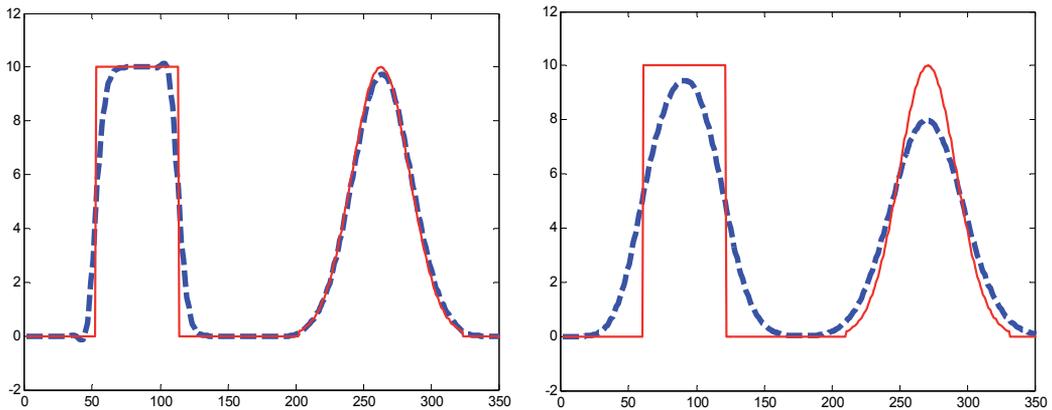


Fig. 6. Results with the method suggested in the present work results after 40 timesteps (left) and 400 timesteps (right)

The results show that the proposed numerical method is comparable with the literature ones in terms of quality of the result; moreover, due to its simplicity it is faster than other more complex methods available in literature (Chung 2010).

3.3.3 Orlanski condition for free boundaries

When dealing with electromagnetic waves in a finite domain, one of the most complicated problems is to set a proper free boundary condition in order to avoid wave reflections. An option could be to extend the computational domain far enough from the electromagnetic wave sources. This is not acceptable, because of the cost in terms of computational time and storage memory. A 3D code like the one developed in this work (HOPE) is already much demanding and it is important to keep the total number of cells as low as possible. Besides, if the time span simulated is not very short, the waves will reach the boundary anyway and they will be reflected back into the domain (moreover the grid has to be the same for the fluid dynamic part and it is impossible to consider to solve NS multi specie equation in a too much extended domain).

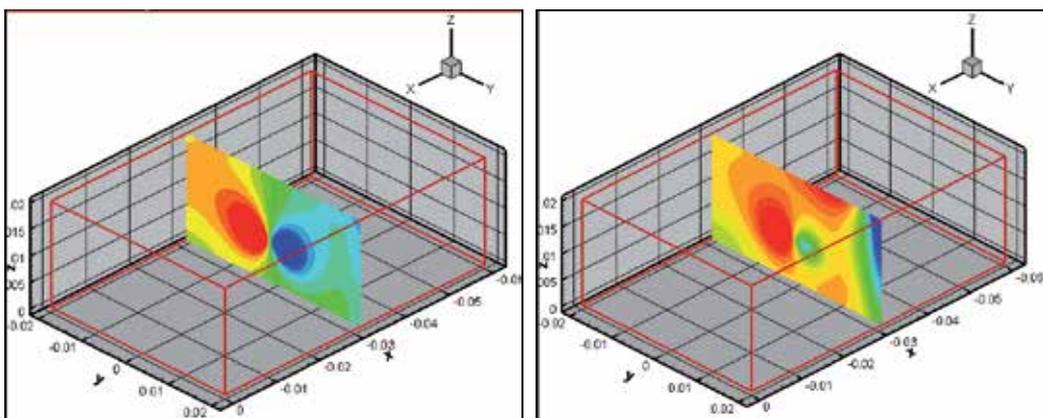


Fig. 7. Magnetic field B_z generated by a current flowing along the “ x ” direction, after 200 time steps (left after 1000 time steps (right)). The electromagnetic wave is reflected at the boundary and the solution is completely corrupted after 1000 time steps.

Initially the condition adopted on the fluxes crossing the boundary cells was simply the following one:

$F_{N+1}^K = F_N^K$, where the subscript N indicates the last cell of the domain and N+1 is the so called ghost cell; K indicates the time step considered.

So the flux associated to the ghost cell was exactly the one associated to the last cell of the domain. Of course this leads to wave reflection, as it is apparent in the following pictures where the magnetic field generated by a wire is represented. After 200 iterations the wave has already reached the boundary and starts to be reflected, after 1000 iterations the wave reflection has completely messed up the solution.

To avoid wave reflections, Orlandi (1976) conditions have been adopted. These are special conditions that define the flux through the boundaries considering what happened at previous time instants. So the boundary flux at instant t_N is dependent on the fluxes in the nearby cells at instants t_{N-1} and t_{N-2} . The condition can be written as:

$$F_{N+1}^K = \frac{\left(1 - \frac{\Delta t}{\Delta x} c\right)}{\left(1 + \frac{\Delta t}{\Delta x} c\right)} F_{N+1}^{K-2} + \frac{2 \frac{\Delta t}{\Delta x} c}{\left(1 + \frac{\Delta t}{\Delta x} c\right)} F_N^{K-1} \quad (3.13)$$

Δt , time step

Δx , grid spacing

Some complication can be found if the propagation velocity is unknown, but dealing with electromagnetic waves there is no such problem. The propagation velocity is constant and equal to c .

Again the previous case of a magnetic field generated by a wire is represented in the next picture. Only the boundary conditions have been changed and this time there is no reflection at all. The waves are absorbed and in both cases, after 200 and after 1000 iterations, the solution is clean.

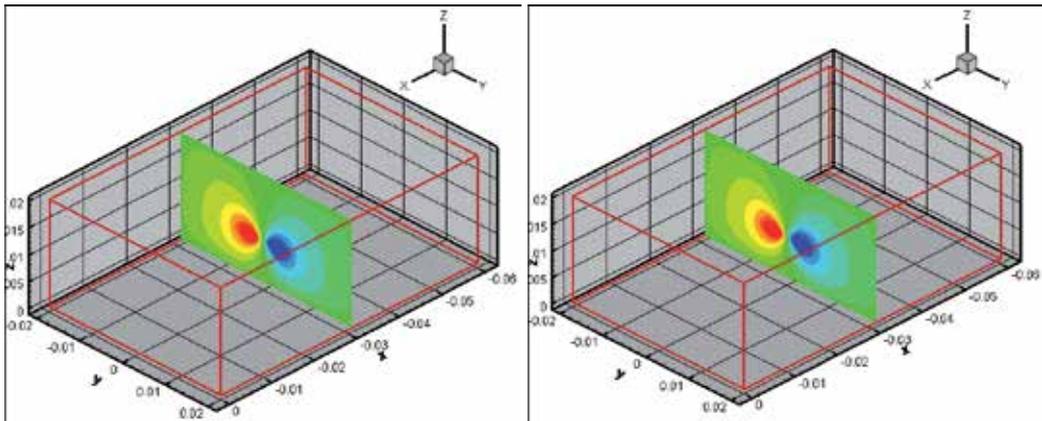


Fig. 8. Magnetic field B_z generated by a current flowing along the “x” direction, after 200 time steps (left) and after 1000 time steps (right). Applying Orlandi conditions there is no reflection and the situation is unchanged in the two cases.

4. Models validation numerical tests

A simple validation test for different models was presented in Battista et al. (2008, 2009). Hereinafter three main validation tests will be presented.

4.1 Validation test 1 - Ionization chemistry in re-entry experiment: RAM-C II

The aim of the RAM-C program was to obtain a better understanding of the factors that influence transmission of radio waves through plasmas and to search for methods to reduce or eliminate blackout (Schexnayder et al 1977). In this frame some experimental data have been collected about the electrons number density at different flight speed and altitudes. The RAM-C test has been considered here in order to compare numerical results in terms of electron number density with the experimental data. For our scopes, the trajectory point characterized by an altitude of 70 km and a re-entry speed of 7654 m/s, corresponding to a Mach number of about 26, has been considered.

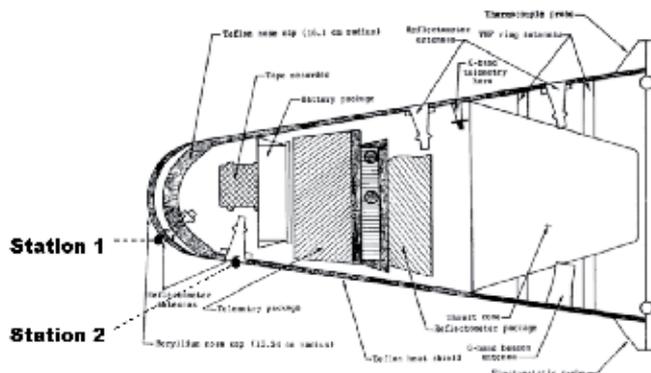


Fig. 9. Geometry and payload configuration of the RAM-C II test with highlighted electron density measurement stations considered in this work.

The results obtained with the 11 species model proposed in this work are shown in the following Fig. 10 and compared with experimental data in terms of electron number density.

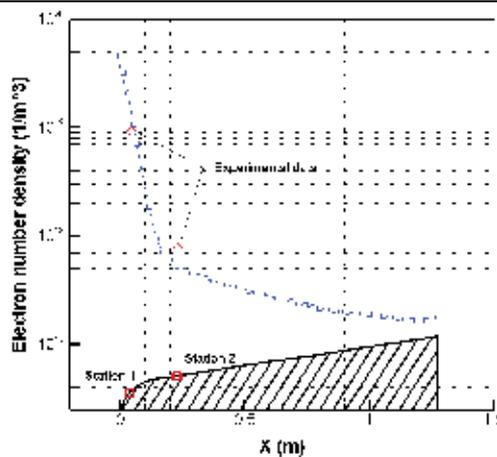


Fig. 10. Numerical and experimental data comparison at different RAM-C stations.

4.2 Validation test 2 - Ionization chemistry in an expansion experiment: Calspan nozzle

In the test campaign described in Dunn and Lordi (1969), electron temperatures and electron densities have been measured in the Calspan nozzle; the working gas was air and the equilibrium reservoir conditions were $T_0=6850$ K and $P_0=0.94738$ MPa. Geometry and grid used in this test is reported in Battista (2009). In Fig. 12 are reported the comparison of numerical data obtained with different kinetic schemes with experimental data. It is evident that the scheme proposed here is the one, among the 11 species schemes, that better reproduce experimental n_e behaviour.

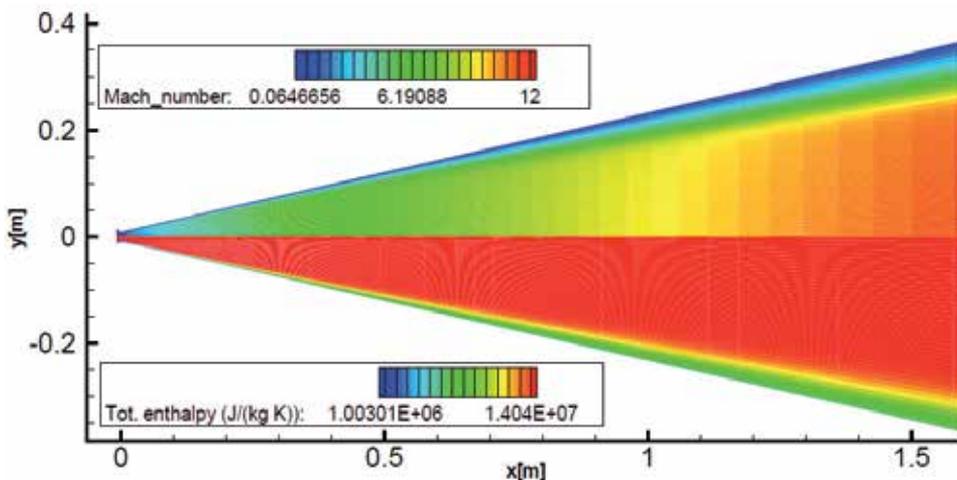


Fig. 11. Test 2: Mach number and total enthalpy distributions

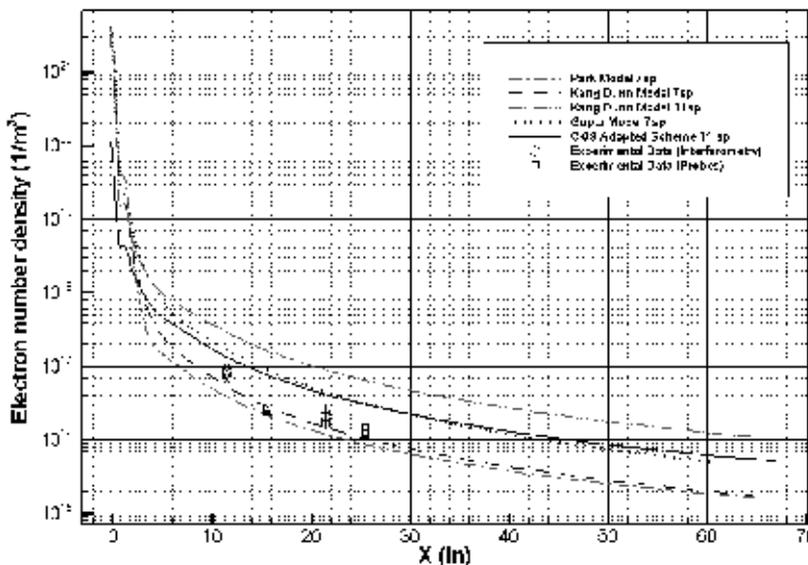


Fig. 12. Test 2 Electron number density comparison with experimental data

4.3 Validation test 3 - Coupled simulation in argon: ALTA test

The basic concepts of magneto fluid-dynamic interaction for plasma flow control in the shock layer of an hypersonic vehicle are here briefly described. Referring to the axis-symmetric blunt body configuration shown in Fig. 13, assuming that a part of the gas is ionized, and that some device embedded within the body generates the B field (amber lines), free charges are subject to a force per charge units $\underline{u} \times \underline{B}$.

$$\underline{F} = q(\underline{E} + \underline{u} \times \underline{B}) \quad (4.1)$$

This force exerted upon the charged particles in the shock layer tends to drive them in the direction $\underline{u} \times \underline{B}$, and (due to a different collisional behaviour between ions and electrons) on a macroscopic scale generates a current density flowing orthogonally to the velocity field direction (Faraday currents).

The current density can be expressed by the generalized Ohm law:

$$\underline{J} = \sigma(\underline{E} + \underline{u} \times \underline{B}) \quad (4.2)$$

The Faraday current density generates the $\underline{J} \times \underline{B}$ body force represented with red arrows. These forces act in a direction opposite to the flow. Furthermore, since conductivity in electromagnetic field assumes a tensorial form, Hall currents (due to Hall collisions) arise orthogonally to B and $\underline{u} \times \underline{B}$ directions and weaken Faraday currents as well as $\underline{J} \times \underline{B}$ force.

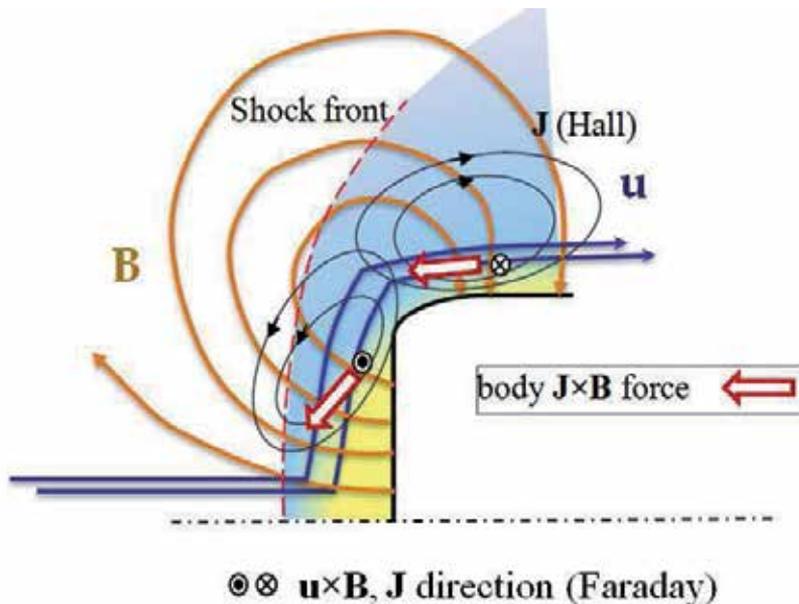


Fig. 13. Scheme of the interaction over a blunt body

Such concepts are used for active flow control in the experiment considered here and used for models validation.

The test considered here is the one described in Cristofolini et al. (2006) and numerically investigated. It consists in an array of the magnets that has been assembled to form a conical

test body with a half-vertex angle of 22.5 deg and the maximum diameter of 40 mm. A cylindrical, 40 mm diameter section is placed behind the cone. Two pressure sensors are embedded in the iron elements between the magnets, one on the cone surface and one in the cylindrical expansion region. The test geometry is reported in Fig. 14. The free stream conditions considered are directly obtained from the nozzle argon viscous test reported in Battista (2009), since the experiment considered here has been carried out in ALTA Heat facility using Mach 6 nozzle .

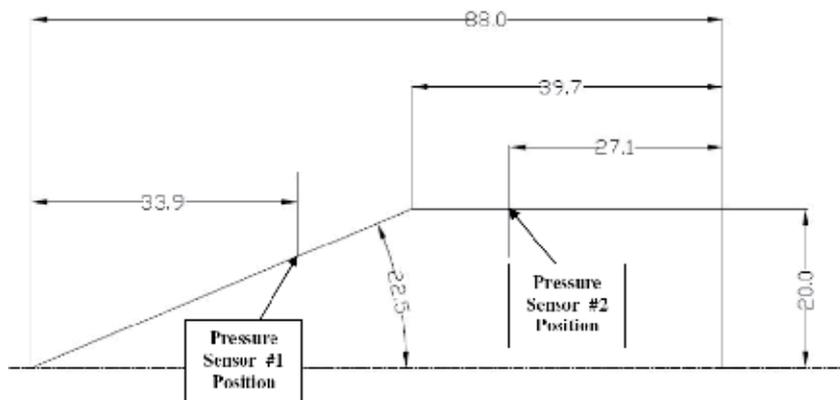


Fig. 14. Geometry of the test body

Calculations have been carried out on the domain shown in Fig. 15 where also the computational grid characterized by 78x60 cells is depicted.

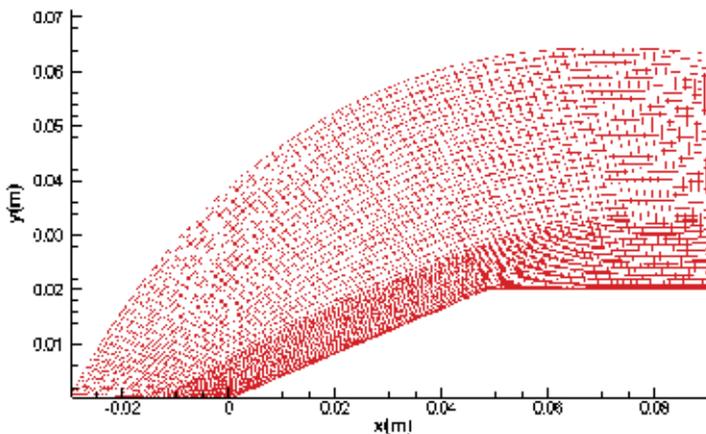


Fig. 15. Computational domain and grid

Magnetic field is given as boundary condition on the wall of the test article and has been deduced from Cristofolini (2006). For the outer boundary the Orlanski boundary condition has been used except for the face on the symmetry plane where the proper boundary condition has been applied. Results in terms magnetic field are reported in Fig. 16.

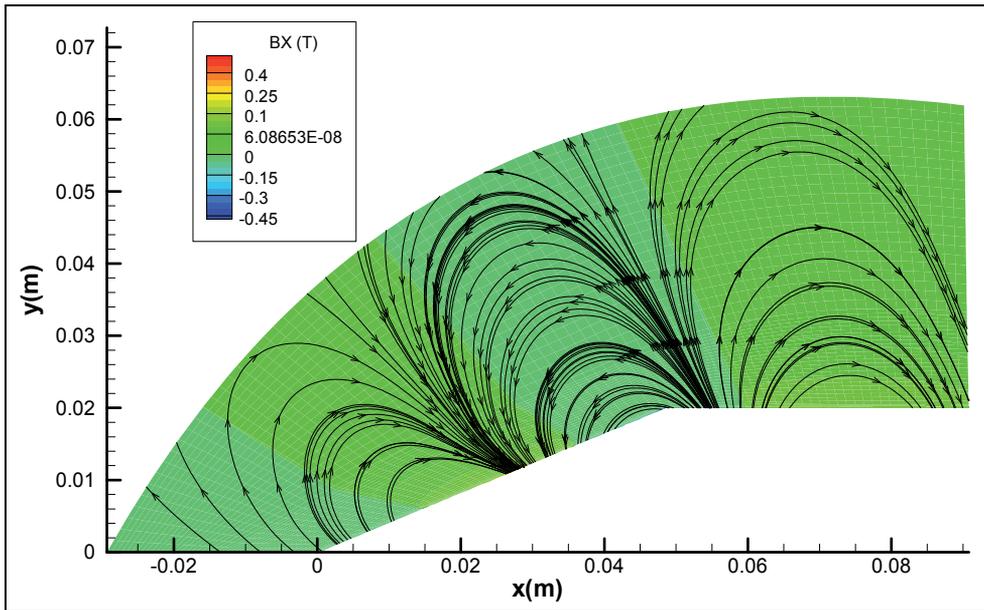


Fig. 16. Computed magnetic field lines.

In Fig. 17 has been shown the comparison in terms of pressure between magnetic and non-magnetic case, in Fig. 18 the comparison with experimental data in terms of pressure shows a quite good agreement; it has to be remarked that the MHD interaction has a strong effect on body pressure distribution, this could be used for an active control scopes.

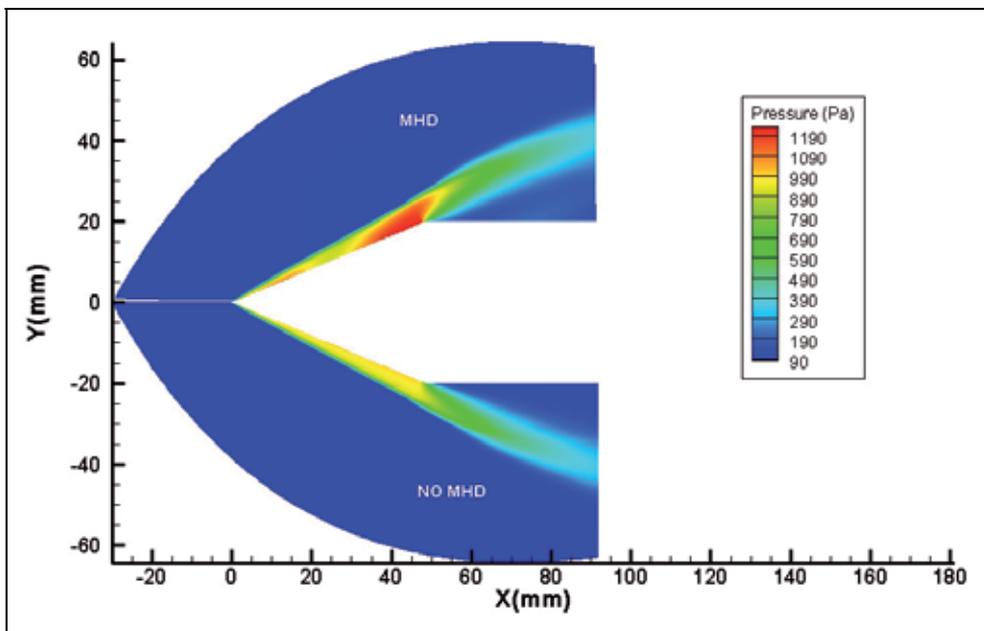


Fig. 17. Computed pressure field without and with MHD interaction

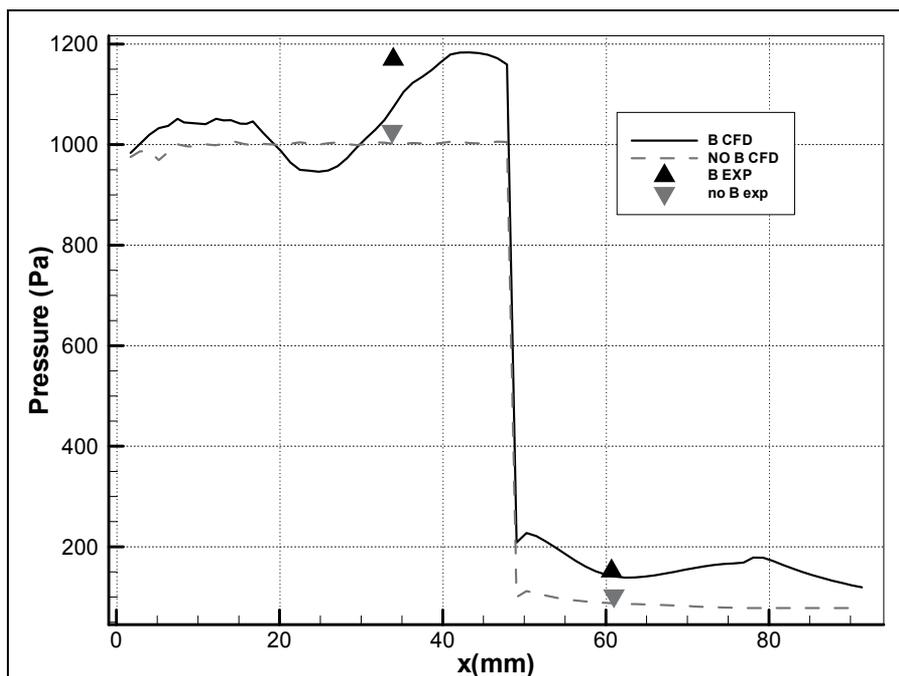


Fig. 18. Pressure along the body wall

5. Conclusions

A three-dimensional model for magnetofluiddynamic problems solution has been proposed considering plasma as a single fluid in thermo-chemical non equilibrium. For chemistry, transport and thermodynamic treatment a macroscopic approach has been followed that allows for the treatment of a general mixture of perfect gases including ionized species. This, in turn, permits not only to deal with non-equilibrium air problems, but gives the opportunity to consider seeded flows or extraterrestrial atmospheric flows in presence of ionization. Models for electrical conductivity have been proposed for air depending upon gas composition and temperature, accounting also the effect of charged particles collisions that could be important in some cases. For what concerns numerical solving strategy a loosely coupling technique to solve full system of equations (Navier-Stokes and Maxwell) has been extended to 3D problems. Both numerical methods for Maxwell equations have been developed and tested, in particular both time marching explicit and implicit methods have been considered and tested for the solution of Maxwell equations. Actually explicit methodology is more rapidly reaching convergence than the implicit one.

Test validation results are in good agreement with available experimental data, and shows that effect of magneto-fluid-dynamic interaction could be relevant for aerospace applications.

Currently CIRA and ALTA with Bologna University Electrical Department are working together in the design of MFD interaction experiments in air in order to understand the real opportunity of application of these technology to Earth re-entry (Cristofolini et al 2010).

6. Nomenclature

Latins

B	Magnetic field
D	Diffusion coefficients
e	energy, electron charge
e-	electron
E	energy
E	Electric field
h	Enthalpy per unit mass
k	Rate constant
J	Flux, Currents
m	mass
M	Mach number, Molecular Weight
n	Number density
N_A	Avogadro Number
P	Pressure
q̇	Heat flux
q	generic system variable
R	Radius
T	Temperature
t	Time
S	Surface
u_i	velocity components
<u>V</u>	Velocity
V	Volume
x	x-coordinate
y	y-coordinate
z	z-coordinate

Greeks

ε₀	Dielectric Vacuum constant
λ_D	Debye Length
<u>λ_e</u>	Dielectric tensor
μ	Mobility
Ω	Production terms
ρ	Density
<u>σ</u>	Conductivity
<u>τ</u>	Stress tensor
ν	Collision Frequency
ω	Frequency

Subscripts

a	activation
∞	freestream conditions
ad	adiabatic wall

ce	electron oscillation
e	electronic
i	ionic
r	radiative
v	vibrational
stag	stagnation point
tot	total
w	wall

7. Acronyms

BC	Boundary Condition
CFD	Computational Fluid Dynamics
CFL	Courant - Freidricks -Lewy Number
CMFD	Computational Magneto Fluid Dynamics
EMC3NS	Electro Magnetic Generalized Chemistry High Enthalpy 3D Navier-Stokes solver
ENO	Essentially Not Oscillatory
EU	Inviscid computation
FDM	Finite Difference Method
FVM	Finite Volume Method
FDS	Flux Difference Splitting
H3NS	Hypersonic high enthalpy 3D Navier-Stokes solver.
MFD	Magneto Fluid Dynamics
MHD	Magneto Hydro Dynamics
NE	Non equilibrium computation
NS	Navier Stokes computation
PG	Perfect Gas approximation computation
PIC	Particle In Cell

8. References

- Battista F. (2009). *Magneto-Fluid-Dynamics Methods for Hypersonic Aerothermodynamics and Space Propulsion Applications*, Ph.D. thesis, Pisa University, etd-11302009-122005.
- Battista, F., Clemente, M. D., and Rufolo, G. C. (2007). *Aerothermal Environment Definition for a Reusable Experimental Re-entry Vehicle Wing Leading edge*, 46th AIAA Thermophysics Conference Miami AIAA 2007-4048.
- Baum, C. E. (1965). *The calculation of Conduction Electron Parameters*, Tech. rep., AFWL.

- Bird R.B., Stewart W.E., and Lightfoot E.N. (1954). *Transport Phenomena*, John Wiley and Sons, New York.
- Borghì C.A., Cristofolini A., Carraro M.R., Gorse C., Colonna G., Passaro A., Paganucci F. (2006). *Non intrusive Characterization of the Ionized Flow Produced by Nozzle of an Hypersonic Wind Tunnel*, 14th AIAA/AHI Space Planes and Hypersonic Systems and Technologies Conference, AIAA 2006-8050.
- Borrelli, S. and Pandolfi, M. (1990). *An Upwind Formulation for the Numerical Prediction of Non Equilibrium Hypersonic Flows*, 12th International Conference on Numerical Methods in Fluid Dynamics, Oxford, UK.
- Bush, W. (1958). *Magnetohydrodynamic-hypersonic flow past a blunt body*, Journal of the AeroSpace Sciences, Vol. 25, pp. 685–690.
- Capitelli, M., Gorse, C., Longo, S., and Giordano., D. (2000). *Collision integrals of high-temperature air species*, Journal of Thermophysics and Heat Transfer , Vol. 14, pp. 259–268.
- Chung, T.J. (2010). *Computational Fluid Dynamics* , Cambridge University Press, Second edition.
- Cristofolini A. et al. (2006), *Experimental Investigation on the MHD Interaction around a Sharp Cone in an Ionized Argon Flow*, AIAA-2006-3075, 37th AIAA Plasmadynamics and Lasers Conference, San Francisco, California.
- Cristofolini A., Borghi C.A., Neretti G., Passaro A., Baccarella D., Schettino A., Battista F., (2010). *Experimental Activities on the MHD Interaction in a Hypersonic Air Flow around a Blunt Body*, 41st Plasmadynamics and Lasers Conference, AIAA 2010-4490.
- Cristofolini, A., Borghi, C. A., and Colonna, G. (2007). *Numerical Analysis of the Experimental Results of the MHD Interaction around a Sharp Cone,* 38th AIAA Plasmadynamics and Lasers Conference.
- D'Ambrosio, D. and Giordano, D. (2005). *A Numerical Method for Two-Dimensional Hypersonic Fully Coupled Electromagnetic Fluid Dynamics*, 36th AIAA Plasmadynamics and Lasers Conference AIAA 2005- 5374.
- Dandy, R. (1993). *Plasma Physic*, Springer-Praxis Series in Space Science and Technology, Cambridge Press, Cambridge.
- Drellishak K. S., Knopp C.F., Bulent Cambel A. (1963), *Partition Function and Thermodynamic properties of Argon Plasma*, The Physics of fluids Volume 6, N.9.
- Dunn, M.G., Lordi, J.A. (1969). *Measurement of Electron Temperature and Number Density in Shock Tunnel Flows. Part II: Dissociative Recombination rate in air*, AIAA Journal Vol.7, N.11,2099-2104.
- Giordano D. (2002). *Hypersonic Flow Governing Equation with electromagnetic field*, AIAA 2002-2165.
- Giordano, D. and D'Ambrosio, D.(2004). *Electromagnetic Fluid Dynamics for Aerospace applications. Part I*, 35rd AIAA Plasmadynamics and Lasers Conference, AIAA Paper 2004-2165.
- Gokcet T.(2004), *N₂-CH₄-Ar chemical kinetic model for simulation of Atmospheric Entry to Titan* AIAA 2004-2469.

- Gordon, S., McBride, P. (1971). *Computer Program for Calculation of Complex Chemical Equilibrium Compositions, Rocket performance, Incident and Reflected Shocks and Chapman-Jouguet Detonations*, NASA Report SP-273, 1971.
- Gupta, R., Yos, J., and Thompson, R. A. (1989). *A Review of Reaction Rates and Thermodynamic and Transport Properties for an 11-Species Air Model for Chemical and Thermal Nonequilibrium Calculations to 30000 K*, Tech. rep., NASA Technical Memorandum 101528.
- Hansen F., Heims P.S.(1968). *A review of the thermodynamic, transport and chemical reaction rate properties of high temperature air*, NACA Technical note 4359.
- Helander P. (2006). *Magnetohydrodynamics*, 43^o Culham Plasma Physic Summer School, Lectures notes.
- Hirshfelder, J. O., Curtiss, C. F., and Bird, R. B. (1954). *Molecular Theory of Gases and Liquids*, John Wiley and Sons.
- Kang, S.-W., Jones, W. L., and Dunn, M. G. (1973) *Theoretical and Measured Electron-Density Distributions at High Altitudes*, AIAA Journal, Vol. 11, pp. N.2.
- Kenneth, G. P. and Philip, L. R., Timur, J. L., Tamas, I. G., and Darren, L. D. Z. (1998). *A Solution-Adaptive Upwind Scheme for Ideal Magnetohydrodynamics*, Physics of Plasma, Vol. 154, pp. 204–309.
- MacCormack, R. W. (2007). *Numerical Simulation of Aerodynamic Flow within a Strong Magnetic Field with Hall Current and Ion Slip*, 38th AIAA Plasmadynamics and Lasers Conference, 2007.
- Mitchner, M. and Kruger, C. H. (1973), *Partially Ionized Gases*, chap. IV, John Wiley & Sons, New York.
- Miura, K. and Groth, C. (2007). *Development of Two-Fluid Magnetohydrodynamics Model for Non-Equilibrium Anisotropic Plasma Flows*, 38th AIAA Plasmadynamics and Lasers Conference.
- Orlanski, I. (1976). *A Simple Boundary Condition for Unbounded Hyperbolic Flows*, Journal of Computational Physic, Vol. 21, 1976, pp. 251–269.
- Park, W., Belova, E., Fu, G., and Tang, X. (1999). *Plasma simulation studies using multilevel physics models*, Physics of Plasma, Vol. 6, pp.1796–1803.
- Park, C. (1993). *Review of Chemical-Kinetic Problems of future NASA Missions, I: Earth Entries*, Journal of Thermophysics and Heat Transfer, Vol.7, No.3.
- Resler, E. and Sears, W. (1958). *The prospects for magneto-aerodynamics*, Journal of the Aeronautical Sciences, Vol. 25, pp. 235–245.
- Schettino, A., Battista, F., Ranuzzi, G., and D'Ambrosio, D. (2008). *Rebuilding of new experimental tests on a double cone at Mach 9*, The 6th European Symposium on Aerothermodynamics for Space Vehicles.
- Schexnayder C.J., Evans J.S. and Huber Paul W. (1977), *Comparison of theoretical and experimental electron density for RAM C flights from The Entry Plasma Sheath and its Effect on Space Vehicle Electromagnetic Systems*. NASA SP-252, 622 pages, published by NASA, Washington, D.C., p.277.
- Taku, A. and Atsushi, F. (2004). *Full Wave Analysis of MHD Modes Using Multi-Fluid Dielectric Tensor in an Inhomogeneous Plasma*, J. Plasma Fusion Res., Vol. 6, pp. 164–168.

- Vincenti W., and Krouger C., Jr. (1967). *Introduction to physical Gas Dynamics*, Wiley, New York, pp 21, 197-244.
- Wagner H.P., Kaeppler H.J. and Auweter-Kurts M. (1998). *Instabilities in MPD thruster flows: Investigation of drift and gradient driven instabilities using multi-fluid plasma models*, *Journal of Physics* 31 559-541.
- Yalim, J., (2001). *Implementation of An Upwind FVM Solver for Ideal MHD for Space Weather Applications*, Ph.D. thesis, University of Washington.

Mechanics of Multi-Phase Frictional Visco-Plastic, Non-Newtonian, Depositing Fluid Flow in Pipes, Disks and Channels

Habib Alehossein

*CSIRO Earth Science and Resource Engineering, University of Queensland
Australia*

1. Introduction

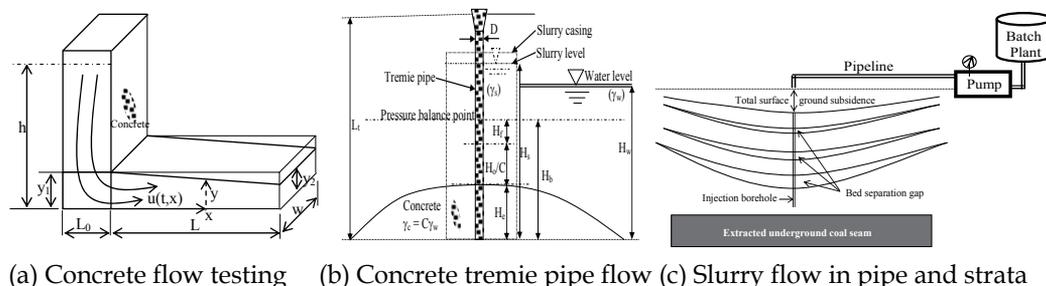
Background: Fresh concrete, fly ash and mining slurries are all frictional-visco-plastic fluids. Fresh concrete flow in Tremie pipes is used to control concrete flow rate and minimise bleeding and dilution when concrete is poured into deep submerged excavations for pile foundation construction. Slurries with very fine aggregates are used to backfill underground voids and mines to prevent subsidence and surface structural damage. Backfilling and injection of granular materials into mining induced voids, separated beddings and cracks, as either diluted slurry or concrete paste, is widely used to control subsidence. As a viable environmental solution, mine waste and rejected materials from underground coal seams are used in both backfilling and injection mine operations. For example, during longwall mining the grout slurry is pumped into the separated beds of the fractured rock mass through a pipeline connected to a central vertical borehole, which is drilled deep into the inter-burden rock strata above the coal mine seam. Either as dilute slurry or thick paste or cake, the fill material normally needs to travel a significant longitudinal distance either in a channel, a tremie pipe, a long pipeline, or radially in a disk shaped crack in the rock mass. An undesirable blockage can occur in the central borehole, in the crack or in the transportation channel or pipeline system when the slurry velocity falls below a certain critical threshold velocity, indicating a material phase change from cohesive-viscous to cohesive-frictional. This chapter presents complete analytical solutions of the required pump pressure versus fluid volume rate for such multi-phase fluids, which are categorised as frictional Bingham-Herschel-Bulkley fluids. The theory derived can be applied to flow of such fluids in pipes, disks and open channels. Furthermore, general analytical solutions have been developed for such fluids in terms of velocity and pressure gradients and velocity and pressure, as a function of flow length (e.g. pipe length, disk radial distance, or channel length) from which special and familiar equations for simpler fluids are derivable by mathematical reduction of the general equations. The formulation is distinct in considering many new aspects including: variable shear parameters rather than fixed values; inclusion of total nonlinear behaviour; and, implementation of a friction function to mimic behaviour of the depositing and consolidating stiff slurry or paste, which can cause a significant pressure rise as a result of the increased shear resistance.

Bingham-Herschel-Bulkley fluids: Recent laboratory and field experiments on mine-backfill fluids, slurries, cements, pastes and concretes proved their wide range of shear resistance

and complex behaviour in response to shearing necessitating development of a general, nonlinear, cohesive, viscous, frictional, nonlinear, non-Newtonian model of shear stress versus shear strain rate, as an extension to the classical Bingham-Herschel-Bulkley fluid [1-9]. Viscous plastic behaviour of such fluids are further simplified or idealised as a reduced or special case of the general nonlinear case. In practice, and for various engineering applications, this generic shear stress function is central in all mathematical formulations to describe fluid flow as a function of pressure gradient. Examples of such applications are flow of slurry, paste and concrete through pipes and tremie pipes and channels for fluid transportation and testing purposes, flow through disks, cracks, joints and rock fractures for injection and backfilling purposes. As a first approach, the shear stress vs. strain rate relation may be idealised by a simple linear function, the so called viscous-plastic Bingham line, which may be derived from a simple linear regression analysis of laboratory experimental data [4-5]. The value of the shear strength function at zero shear strain rate, i.e. plastic yield strength (also called cohesion), and slope of the linear curve (viscosity) are two important parameters of the fluid property in the simple linear idealised case.

Grouts and slurries: The large cavern created by an underground mine may eventually lead to failure of the overburden rock, propagating layer by layer to the surface, resulting in substantial ground surface subsidence [10-13], as schematically shown in Figure 1(c). As a major potential hazard, mining induced subsidence significantly affects mining costs where major surface structures and natural environment need to be protected, e.g. mining under river systems, gorges, cliffs, power lines, pipelines, communication cables, major roads and bridges, and other significant surface facilities [11]. Remedial measures to manage damage caused by subsidence can often be very costly with potentially damaging impacts and irreversible consequences. Backfilling and injection of granular materials into the mining induced voids, separated beddings and cracks, as either diluted slurry or concrete paste, is widely used to control mine subsidence. Grouts and slurries made of mine and power plant wastes and rejects are viable environmental backfill solutions to both ground stability and mine waste management problems [12]. Like concrete paste, the flowing slurry can be categorised as a generally nonlinear viscous cohesive (Bingham Herschel-Bulkley) fluid [5-11]. However, in mining applications, to reduce ground surface subsidence and control the propagation of the overburden movement to the surface, the solid particles in the injected slurry must deposit in the bed separation gaps of the coal seam over-burden strata, e.g. in longwall mining the grout slurry is pumped into the separated beds of the rock mass from a batching plant source through pipelines connected to a central vertical borehole, which is drilled deep into the over-burden rock above the coal seam (Figure 1(c)). Flow blockage can occur in the injection system, when the slurry velocity falls below a certain critical threshold velocity. The stiffening, consolidating non-flow slurry can generally be categorised as a frictional cohesive soil [14]. In other words, a change of material phase from cohesive-viscous to cohesive-frictional will occur. Using a smaller scale model, this field injection practice has been simulated at the QCAT laboratory of the Commonwealth Scientific & Industrial Research Organisation (CSIRO) in Brisbane, Australia, to study the influence of various grout injection parameters by pumping slurries through various pipes of different sizes and diameters and for different applications (Figures 1-2). As an important industrial application, grout injection into the inter-burden strata is used as a modern technology to control and reduce coal mine subsidence [10-13]. Slurry mixes of coal mine and power plant waste materials, e.g. fly ash or any other coal wash rejects, are injected back into the inter-burden rock strata during longwall mining [4-5]. To reduce subsidence and control inter-burden strata movement, the injected

slurry solid particles must deposit in the opened strata bed separation gaps or cracks before crack closure [10-13]. The mechanics of non-Newtonian fluids flowing between parallel disks is a classical fluid mechanics problem that has been studied by a number of researchers in the past for their specific problems of interest [3-4].



(a) Concrete flow testing (b) Concrete tremie pipe flow (c) Slurry flow in pipe and strata
 Fig. 1. Various applications of viscous slurry and paste fluids: (a) channel flow for workability and consistency testing of concrete; (b) Concrete termie pipe flow into submerged foundations; (c) multi-phase slurry flow in pipes and fractured rock strata for void backfilling.

Concrete: Fresh concrete flow through Tremie pipes is used to control concrete flow rate and minimise segregation, bleeding and dilution when poured or placed into deep submerged excavations for pile foundation construction. Slurries with very fine aggregates are used to backfill underground voids and mines to prevent subsidence and surface structural damage. Backfilling and injection of granular materials into the mining induced voids, separated beddings and cracks, as either diluted slurry or concrete paste, is widely used to control coal mine subsidence. As a viable environmental solution, mine waste and rejected materials from underground coal seams in the form of either cementitious or non-cementitious grout, are used in both backfilling and injection mine operations. For example, during longwall mining the grout slurry is pumped into the separated beds of the rock mass through a central vertical borehole, which is drilled deep into the inter-burden rock strata above the coal seam. Either as dilute slurry or thick paste or cake, the fill material normally needs to travel a significant distance in either a long pipeline, or radially in a disk type crack formation of the rock mass. An undesirable blockage can occur in the central borehole, in the disk gap or in the transportation channel or pipeline system, when the slurry velocity falls below a certain critical threshold velocity, indicating a material phase change from cohesive-viscous to cohesive-frictional. Indirect index measure of concrete viscosity and plastic yield is made via an L-box channel measuring workability and flowability of tremie pipe concrete. The L-box test, originally developed for super-workable concrete [6-9] is a relatively newly introduced concrete test to measure the consistency, workability and flowability of a tremie pipe concrete, hence, it is indirectly related to concrete plastic yield and viscosity [6]. Concrete is poured in the rectangular vertical chimney part of the L-box and is allowed to flow in the horizontal channel part, once a sliding gate is opened. The time and profile of the concrete flow in the horizontal channel is measured to compare viscous-plastic behaviour of different concretes. The variation of flow velocity with time and position from the time the sliding gate is opened until the flow reaches static equilibrium has been simulated and formulated by a representative dimensionless partial differential equation (PDE). Mathematically, the resulting equation is of the same form as a non-homogeneous heat-conduction equation [6].

In this chapter complete analytical solutions of the required pump pressure versus fluid volume rate are discussed for such multi-phase fluids, which are categorised as frictional Bingham-Herschel-Bulkley fluids. The discussed theory can be applied to flow of such fluids in pipes, disks and open channels. Furthermore, general analytical solutions have been developed for complex fluids in terms of velocity and pressure gradients and velocity and pressure, as a function of flow length (e.g. pipe length, disk radial distance, or channel length) from which special and familiar equations for simpler fluids are derivable by mathematical reduction of the general equations. The formulation is distinct in considering many new aspects including: variable shear parameters rather than fixed values; inclusion of total nonlinear behaviour; and, implementation of a friction function to mimic behaviour of the depositing and consolidating stiff slurry or paste, which can cause a significant pressure rise as a result of the increased shear resistance.

2. Frictional Bingham-Herschel-Bulkley fluid

The general constitutive equation, relating fluid shear stress to shear rate for such general nonlinear, non-Newtonian, viscous, plastic, frictional fluids, which can be applied to fresh concrete, mine backfill slurries and high frictional multiphase fluids, is as follows [4-9]

$$\boldsymbol{\tau}(t,\mathbf{x}) = \mu(t,\mathbf{x}) \left(-\frac{\partial \mathbf{u}(t,\mathbf{x})}{\partial \hat{\mathbf{x}}} \right) + \eta(t,\mathbf{x}) \left(-\frac{\partial \mathbf{u}(t,\mathbf{x})}{\partial \hat{\mathbf{x}}} \right)^n + \boldsymbol{\tau}_0(t,\mathbf{x}) + \xi(t,\mathbf{x})p(t,\mathbf{x}) \quad (1)$$

In Equation (1) $\boldsymbol{\tau}$ is shear stress tensor, \mathbf{u} is velocity vector, μ and η are linear and nonlinear viscosities, $\boldsymbol{\tau}_0$ is plastic yield, p is pipe pressure and ξ is concrete friction coefficient. The last term, involving the friction and pressure terms (ξp), is a frictional resistance term which can be applied only when a pipe blockage occurs due to the concrete granular material friction and needs to be reopened by a higher pressure flow, otherwise it can be ignored [4-9]. See Figure 2 for a visual definition of the different shear terms and parameters involved in Equation (1).

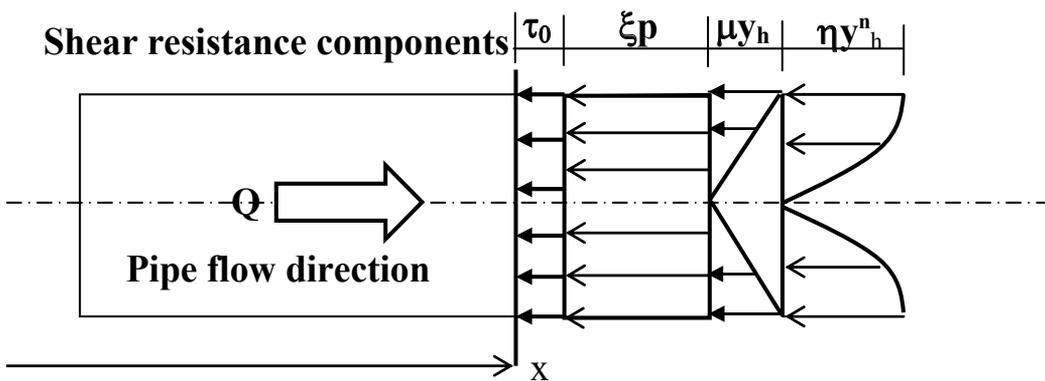


Fig. 2. Schematic diagrams showing various shear stress components in Equation (1). τ_0 is the constant uniform plastic yield component, with no viscosity; μ is the Newtonian linear viscosity coefficient of the linear velocity gradient y with a wall value y_h ; η is the non-linear viscosity; ξ is the friction coefficient of the fluid pressure p .

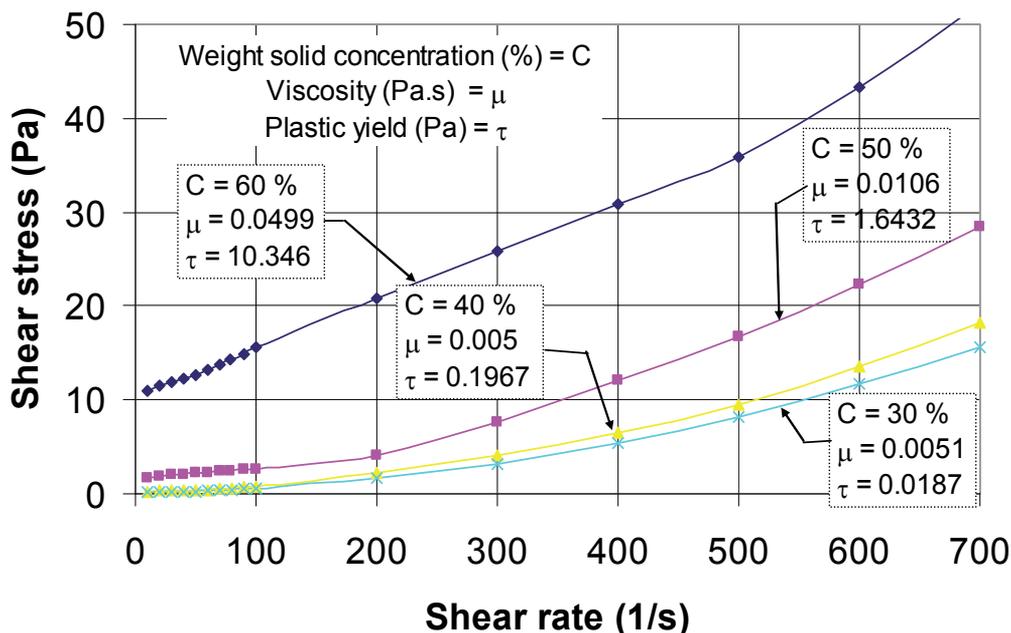


Fig. 3. Shear stress vs shear rate (range 0-700/s) for a sample at solid weight concentration range 30%-60%. Numbers in legend table show Bingham plastic linear fit model to these results in the range of 0-100/s. For instance, for 50% solid concentration in the shear rate range of 0-100/s of the sample, the linear viscosity (Bingham slope) is 0.0106 Pa.s and the yield (Bingham intercept) is 1.6432 Pa.

We can measure these parameters by a viscometer-testing device [4-5]. Sometimes for slurries of various solid particle concentrations, the viscometer test results can conveniently fit into one or two linear models (Bingham plastic) for the whole range of shear strain rate. Figure 3 shows examples of multi-linear or bilinear model for slurries of different concentrations for two distinct range of shear strain rate 0-100 /s and 100-700 /s. The Bingham models can be identified by its two main parameters (yield intercept and viscosity slope).

3. Governing equations

Governing equations of most fluid mechanics problems normally start with the general basic Reynolds transport theorem of continuum mechanics [15]. This is initially an integral relation stating that the sum of the changes of any intensive fluid property, such as mass, momentum and energy, defined over a control volume CV, denoted here by symbol ζ , must be equal to what is gained or lost through the boundaries of the volume, or control surface (CS), plus what is created or consumed by sources and sinks inside the control volume [15].

$$\frac{d}{dt} \int_{CV} \zeta dV + \int_{CS} \zeta \bar{\mathbf{u}} \cdot \bar{\mathbf{n}} dA + \int_{CV} Q dV = 0 \quad (2)$$

In Equation (2), $\bar{\mathbf{u}}$ is fluid velocity vector, $\bar{\mathbf{n}}$ is normal vector to the control surface dA , and Q is the fluid source or sink. Using integration by parts, the second integral can also be

transformed to a volume integral by the divergence theorem. Since the whole grouped volume integral must be zero for any arbitrary control volume CV , it implies that the integrand itself must be zero. Therefore, our theory can be started with the following general basic differential equation:

$$\frac{d\zeta}{dt} + \bar{\nabla} \cdot \zeta \bar{\mathbf{u}} + Q = 0 \quad (3)$$

Applying the general conservation Equation (3) to mass and momentum, the Navier-Stokes isothermal equations of continuity and momentum [15] are recovered:

$$\bar{\nabla} \cdot \rho \bar{\mathbf{u}} + \dot{\rho} = 0 \quad (4)$$

$$\bar{\nabla} \cdot \bar{\boldsymbol{\sigma}} + \rho \bar{\mathbf{g}} = \rho \frac{D\bar{\mathbf{u}}}{Dt} \quad (5)$$

where $\bar{\nabla} = \partial / \partial x_i \bar{\mathbf{e}}_i$ is the gradient vector, ρ is the fluid density, $\bar{\boldsymbol{\sigma}}$ is the stress tensor and $\bar{\mathbf{g}}$ is the body acceleration or gravity vector. The stress tensor depends on a mean fluid normal stress or pressure ($p = \delta_{ij}^{-1} \sigma_{ij}$) and a deviator stress representing shear stresses τ_{ij} , which depends on fluid viscosity and velocity gradients.

$$\sigma_{ij} = -p\delta_{ij} + \tau_{ij} \quad (6)$$

As shown by [3-4], the deviatoric shear stress in (6) for cohesive, frictional, viscous, non-Newtonian slurries depends not only on fluid velocity gradients and yield plastic shear strength, but also on fluid pressure causing frictional resistance to flow, particularly during a blockage ($Q \rightarrow 0$). As discussed earlier in Equation (1), on the basis of several laboratory experiments on soil like slurries, a general shear stress versus shear strain constitutive material law is proposed for viscous, cohesive, frictional, plastic slurries in which the fluid shear stress is a nonlinear function of shear rate and longitudinal distance. Equation (1) has the following general form when written tensor notation is applied:

$$\tau_{ij} = \mu(u_{i,j} + u_{j,i}) + \eta(u_{i,j} + u_{j,i})^n + \tau_{0ij} + \xi p \delta_{ij} \quad (7)$$

As discussed earlier and also shown in Figure 2, the first term on the right hand side is the familiar linear Newtonian component, the second term is the nonlinear pseudo-plastic component, the third term is the yield component and the fourth term is the pressure component, in which ξ is a coefficient of granular material friction which is the same tangent function of the material friction angle [14]. In the theoretical analysis discussed here, it is assumed that:

1. the flow is laminar and the fluid is incompressible, steady state, stationary, and isothermal and axisymmetric with no eddies and no gravity effects;
2. axially symmetric condition implies that the radial flow component (in a pipe) and circumferential flow (in a disk) must vanish. In other words $u_{x2} = u_{x3} = 0$;
3. when there is a full blockage ($Q = 0$), the friction term ξp is a dominant term;
4. the classical term "fluid" has loosely been used interchangeably with "slurry"; to refer to a "slurry", whenever an equivalent "fluid" model can represent the overall, average, mechanical behaviour of the "slurry".

We now introduce a new set of independent variables to represent coordinate axes of both radial disk and pipe flow systems, namely: x_1 axis is always in the direction of the main flow direction, i.e. either radial disk flow, or longitudinal pipe or channel flow, x_2 is the axis normal to the flow direction in the flow cross-section plane, and x_3 is the hoop or circumferential axis direction identical to hoop angle θ .

4. Reduced one dimensional equations

Consider now the simple problem of fluid flow through either (i) a uniform circular pipe of inside diameter $2h$, as shown in Figure 4 (left), or (ii) a radial disk of thickness $2h$, as in Figure 4 (right), or (iii) a channel, as part of an L-Box testing device shown in Figure 1 (a).

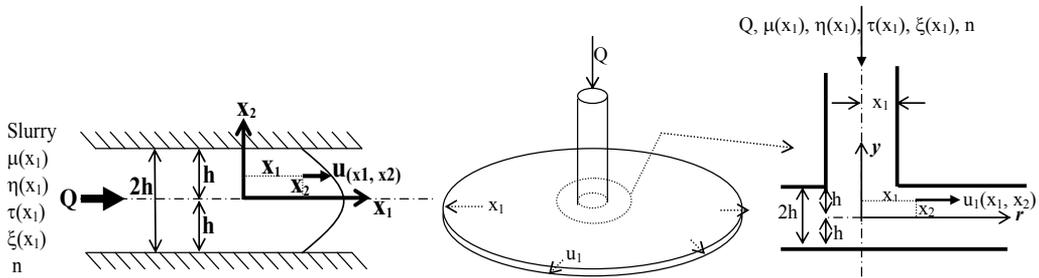


Fig. 4. Non-Newtonian viscous-plastic flow in a pipe (left) and in a radial disk (right) with fluid flow parameters

Fluid flow through a pipe of uniform, circular, cross-section is known as the Hagen-Poiseuille flow problem [5]. It is assumed that the circular pipe flow is symmetric around the pipe longitudinal x -axis, the normal stresses are simply the fluid pressure, p , the fluid is incompressible and non-Newtonian in a steady state condition, there is no velocity component in the pipe circular cross-sectional plane, i.e. the plane normal to the pipe length direction. Similarly, it is assumed that the flow in a disk is also non-Newtonian, steady state, incompressible and laminar and the radial disk flow is also cylindrically axi-symmetric [4]. Hence, implementing all these assumptions implies that in all flow cases both normal to flow velocity components (u_2 , u_3) are zero and there is no variation in velocity or pressure with time. In other words, we have

$$\mathbf{u} = \mathbf{u}_1 = \mathbf{u}(x_1, x_2) \quad (8)$$

$$u_2 = u_3 = u_0 = 0 \quad (9)$$

$$\partial u_i / \partial t = \partial u_i / \partial \theta = \partial p / \partial t = \partial p / \partial \theta = \partial u_i / \partial x_2 = \partial p / \partial x_2 = 0 \quad (10)$$

We now define two separate flow gradient functions [4-5], (a) gradient with respect to x_1 and (b) gradient with respect to x_2 , viz.

$$y = -\frac{\partial u_1}{\partial x_2} \quad (11a)$$

$$\Psi = -x_1 u_1 \quad (11b)$$

$$\psi = -x_1 \frac{\partial u_1}{\partial x_2} = x_1 y = \Psi' \quad (11c)$$

Hence, the general, basic equations of continuity (4) and momentum (5) reduce to Equations (12) for pipe flow and (13a,b) for radial flow.

$$\text{PF (Pipe flow):} \quad \frac{1}{x_2} \frac{\partial(x_2 \tau)}{\partial x_2} = \frac{\partial p}{\partial x_1} \quad (12)$$

$$\text{RF (Radial Flow):} \quad -\frac{1}{x_1} \frac{\partial \Psi}{\partial x_1} + \frac{\partial u_2}{\partial x_2} = 0 \quad (13a)$$

$$\frac{\partial p}{\partial x_1} + \frac{\partial \tau}{\partial x_2} = 0 \quad (13b)$$

It should be noticed that gravity effect can be incorporated into the pressure gradient term in pipe flow Equation (12), where γ is the fluid unit weight and β is the inclination angle of the pipe with respect to horizontal axis. In order to consider gravity in a pipe flow, $\frac{\partial p}{\partial x}$ in

(12) should be replaced with $\frac{\partial p}{\partial x} \pm \gamma \sin(\beta)$, as suggested in [2]. The shear stress $\tau(x_1, x_2)$ in (12) and (13) is a 2D version of the general case shown in (7), which is reproduced in Equation (14) below [4-5].

$$\tau(x_1, x_2) = \mu(x_1)y + \eta(r)y^n + \tau_1(x_1) + \xi(x_1)p(x_1) \quad (14)$$

No slip boundary conditions, i.e. no velocity at the pipe or disk walls, and a full axial or radial symmetry of the flow are assumed. Hence, Equations (11-13) must be solved subject to the following boundary conditions:

$$u = u_1 = u(x_2 = h) = 0 \quad (15a)$$

$$\frac{\partial u}{\partial x_2}(x_2 = 0) = 0 \quad (15b)$$

Substituting (14) in (12) and (13) and integrating over x_2 , will give us the following pressure-gradient equations

$$\text{PF:} \quad \frac{x_2}{2} \frac{dp}{dx_1}(x_1) = \tau = \mu(x_1)y(x_2) + \eta(x_1)y^n(x_2) + \tau_0(x_1) + \xi(x_1)p(x_1) \quad (16)$$

$$\text{RF:} \quad x_2 \frac{dp}{dx_1}(x_1) = \mu(x_1)y + \eta(r)y^n \quad (17)$$

The pressure gradient Equations (16) and (17) must be satisfied at all points, including the boundary point, h . Therefore, at the wall boundary point ($x_2=\pm h$) we have

$$\text{PF:} \quad \frac{h}{2} \frac{dp}{dx_1}(x_1) = \tau_h = \mu(x_1)y_h + \eta(x_1)y_h^n + \tau_0(x_1) + \xi(x_1)p(x_1) = F(y_h) \quad (18)$$

$$\text{RF:} \quad h \frac{dp}{dx_1} = \mu(x_1) \left(\frac{1}{x_1} \psi_h \right) + \eta(x_1) \left(\frac{1}{x_1} \psi_h \right)^n = f(\psi_h/x_1) = f(y_h) \quad (19)$$

where y_h , τ_h and ψ_h are the boundary values of y , τ and ψ , i.e. at the point $x_2 = h$. In other words,

$$\text{PF:} \quad y_h = y(x_2 = h) = - \left. \frac{\partial u}{\partial x_2} \right|_{x_2=h}, \quad \tau_h = \tau(x_2 = h), \quad \psi_h = \psi(x_2 = h) \quad (20)$$

If y_h in (18), or ψ_h in (19), are known, the pipe pressure p can be calculated by integrating these equations directly. The result can still be in integral forms depending on the complexity of the coefficient functions such as: viscosity $\mu(x_1)$, or $\eta(x_1)$, plasticity, $\tau(x_1)$, or friction $\xi(x_1)$ [4–5].

$$\text{PF:} \quad \Delta p = \frac{2}{h} \int_{x_0}^{x_1} \tau_h(x_1) dx_1 \quad (21a)$$

$$\text{RF:} \quad \Delta p = \frac{1}{h} \int_{x_0}^{x_1} f(\psi_h/x_1) dx_1 \quad (21b)$$

For example, the pipe pressure can be written in a general integral form (Equation (22)) in terms of the integral coefficients A_1 , B_1 , C_1 [5].

$$p = \left(p_0 v_0 + (A_1 y_h + B_1 y_h^n + C_1) \right) v^{-1} \quad (22)$$

In Equation (22), $v(x_1)$ is an exponential function of x_1 and h and A_1 , B_1 , C_1 are integral coefficients similar to those produced for radial flow [4], viz.

$$v(x) = e^{-\frac{2}{h} \int \xi dx} \quad (23)$$

$$A_1 = \frac{2}{h} \int_{x_0}^x v(x) \mu(x) dx \quad (24a)$$

$$B_1 = \frac{2}{h} \int_{x_0}^x v(x) \eta(x) dx \quad (24b)$$

$$C_1 = \frac{2}{h} \int_{x_0}^x v(x) \tau_0(x) dx \quad (24c)$$

As shown in [4-5], evaluations of these integrals become straight forward and generic, if we first find the normalised forms of the dimensional functions, $\mu(x)$, $\eta(x)$, $\tau(x)$ and $\xi(x)$ in the shear stress Equations (16) and (17). Using the same symbols, but in *Italic* fonts, let the *italic* symbols μ , η , τ and ξ , be the normalised counterparts of $\mu(x)$, $\eta(x)$, $\tau(x)$ and $\xi(x)$, respectively. As discussed in detail in [4-5], all the four functions μ , η , τ , ξ , can be represented by one symbolic generic function, α , i.e. $\alpha(\alpha_0, \alpha_s, r, n_\alpha)$. The normalised form of α is simply

$$\alpha(r) = 1 - \alpha_s (r - 1)^{n_\alpha} \quad (25)$$

in which α_s is the normalised slope and n_α is the general power factor for any nonlinear behaviour. In other words,

$$\alpha = \frac{\alpha}{\alpha_0}, \alpha_s = \left(1 - \frac{\alpha_\infty}{\alpha_0} \right) \left(\frac{r_\infty}{r_0} - 1 \right)^{-n_\alpha} \quad (26)$$

Therefore, the integrals (24) have the following general non-dimensional form which can be integrated numerically [4-5]:

$$\text{PF: } I(x) = \int_1^x v(x) \alpha(x) dx = \int_1^x e^{k \int \xi dx} \alpha(x) dx \quad (27a)$$

$$\text{RF: } I(x) = \int_1^x x^{-n} [1 - \alpha_s (x - 1)^{n_\alpha}] dx = \frac{x^{1-n} - 1}{1-n} - \frac{\alpha_s}{1-n} \left\{ x^{1-n} (x - 1)^{n_\alpha} - n_\alpha \int_1^x x^{1-n} (x - 1)^{n_\alpha - 1} dx \right\} \quad (27b)$$

However, for constant properties of slurry, the A_1 , B_1 , C_1 parameters reduce to either a simple exponential function of $x=x_1$, in the presence of a friction coefficient, i.e. $\xi \neq 0$, or a simple linear function in terms of pipe length ($L = x - x_0$), in the absence of friction coefficient, i.e. $\xi=0$. In other words, the values of these coefficients, as described by the integrals (24), can be calculated from relations (28) for the case where $\mu = \eta = \tau_0 = \text{constant}$, but $\xi=0$, and from relations (29) for the case where $\xi = \mu = \eta = \tau_0 = \text{constant}$, but $\xi \neq 0$.

$$A_1 = 2h^{-1} \mu (x - x_0) = 2h^{-1} \mu L, \quad B_1 = 2h^{-1} \eta L, \quad C_1 = 2h^{-1} \tau_0 L \quad (28)$$

$$A_1 = -\frac{\mu}{\xi} v_L, \quad B_1 = -\frac{\eta}{\xi} v_L, \quad C_1 = -\frac{\tau_0}{\xi} v_L, \quad v_L = e^{-\frac{2}{h} \xi L} \quad (29)$$

Figures 5 and 6 show some typical values of the function I for a range of viscosity and plasticity parameters in a radial flow, namely n_α from (0) to (2), n from (0) to (2), α_s from (0) to (2) and normalised $r=x$ from (1) to (100) [4].

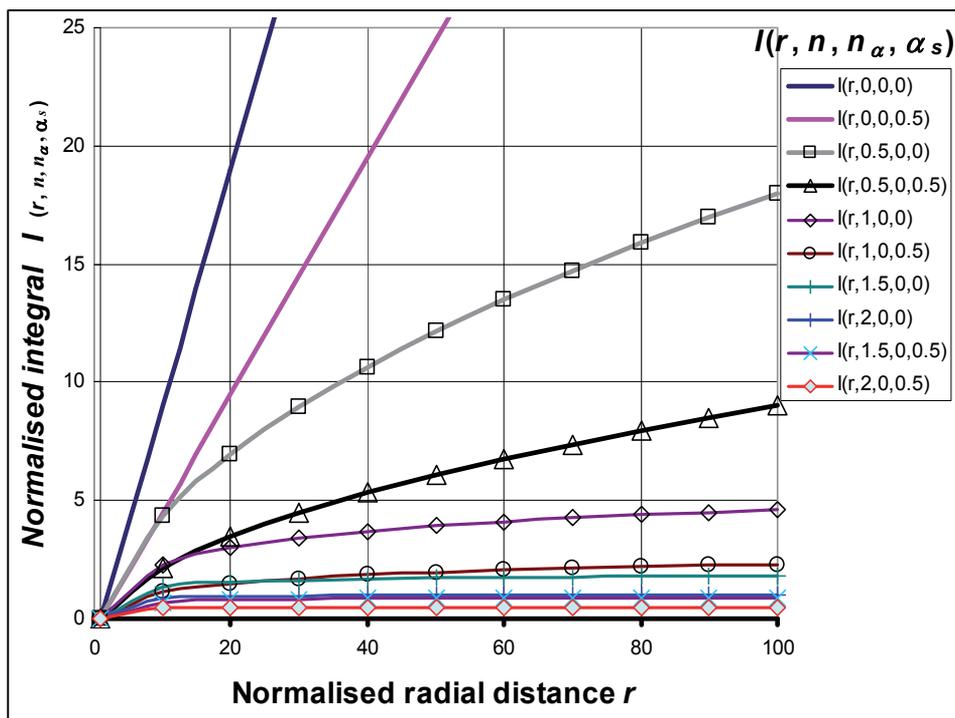


Fig. 5. Values of integral function I_α with radial distance r for some values of parameters n , n_α and α_s , as indicated in Equation (27b)

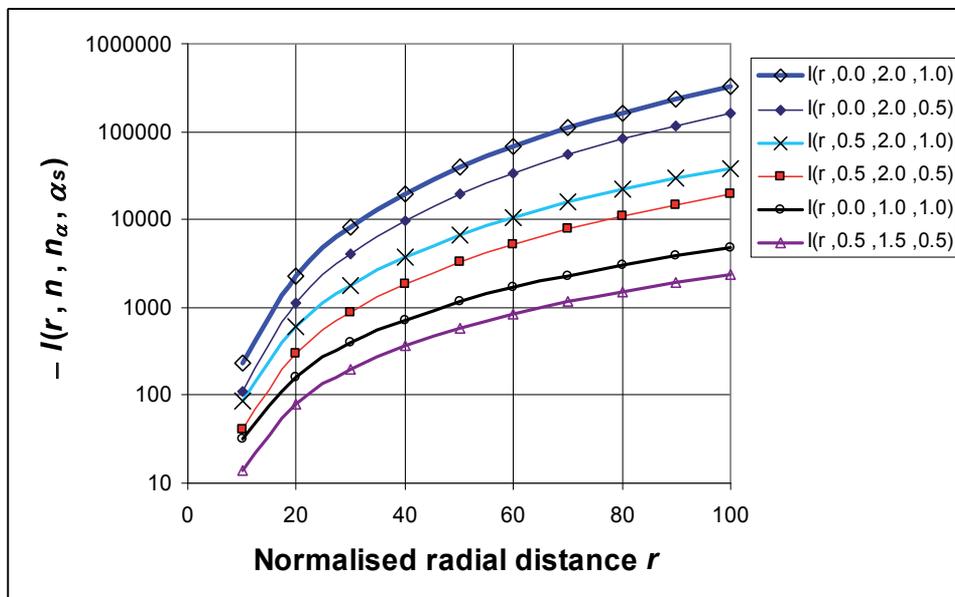


Fig. 6. Values of integral function I_α with radial distance r for some values of parameters n , n_α and α_s , as indicated in Equation (27b)

5. Solutions for pipe and radial flow

The pressure gradient dependency in all Equations (16)-(17) can be removed by dividing general functions or Equations (16)-(17) to their corresponding boundary values, i.e. Equations (18)-(19). Thus, we have a ratio of two polynomial functions, with a numerator that is a function of y , and a denominator that is a function of y_h , as shown in Equation (30a, b), in which a, b, A, B, C are functions of flow line distance x_1 only.

$$\text{PF:} \quad x_2 = h \frac{\tau}{\tau_h} = h \frac{Ay + By^n + C}{Ay_h + By_h^n + C} = h \frac{F(y)}{F(y_h)} = h \frac{F}{F_h} \quad (30a)$$

$$\text{RF:} \quad x_2 = h \frac{\tau}{\tau_h} = h \frac{ay + by^n}{ay_h + by_h^n} = h \frac{f(y)}{f(y_h)} = h \frac{f}{f_h} \quad (30b)$$

To solve (30a) or (30b) for our primary unknowns, either y or ψ , we need another equation in terms of the flow rate, Q , which must be conserved at any section normal to x_1 direction. The results are integral equations relating velocity gradient y , or y_h (or ψ , or ψ_h in the case of radial flow) to the flow rate Q [4-5].

$$\text{PF:} \quad Q = \int_A \bar{u} \cdot d\bar{A} = 2\pi \int_0^h x_2 u_1 dx_2 = \frac{\pi}{3} h^3 (y_h - F_h^{-3} G_h) \quad (31a)$$

$$\text{RF:} \quad Q = \int_A \bar{u} \cdot d\bar{A} = 2\pi \int_0^h x_2 u_1 dx_2 = 2\pi h^2 (\psi_h - f_h^{-2} g_h) \quad (31b)$$

Values of velocity gradient at the wall boundary, y_h , or the function $\psi_h = x_1 y_h$, needs be calculated generally by the Newton-Raphson iteration [16]. Hence solutions to (31) take the following general forms:

$$\text{PF:} \quad y_{h_{i+1}} \Leftrightarrow y_{h_i} - \frac{(Q_h - y_h)G'(y_h) + G(y_{h_i})}{(Q_h - y_h)G''(y_h)}, \quad Q_h = \frac{3}{\pi h^3} Q \quad (32a)$$

$$\text{RF:} \quad \psi_{h_{i+1}} \Leftrightarrow \psi_{h_i} - \frac{(Q_h - \psi_h)g'(\psi_h) + g(\psi_{h_i})}{(Q_h - \psi_h)g''(\psi_h)}, \quad Q_h = \frac{1}{2\pi h^2} Q \quad (32b)$$

In Equation (31)-(32) f, F, g and G are polynomial functions of the unknown variable y or ψ .

$$g(\psi) = \int f^2 d\psi = \frac{a^2}{3} \psi^3 + \frac{b^2}{2n+1} \psi^{2n+1} + \frac{2ab}{n+2} \psi^{n+2} \quad (33a)$$

$$g'(\psi) = f^2(\psi) = a^2 \psi^2 + b^2 \psi^{2n} + 2ab \psi^{n+1} \quad (33b)$$

$$g''(\psi) = 2a^2 \psi + 2nb^2 \psi^{2n-1} + 2(n+1)ab \psi^n \quad (33c)$$

$$f(\psi) = a\psi + b\psi^n \quad (33d)$$

$$f'(\psi) = a + nbx^{n-1} \quad (33e)$$

$$\begin{aligned} G(y) = \int F^3 dy = & \frac{A^3}{4}y^4 + A^2Cy^3 + \frac{3AC^2}{2}y^2 + C^3y + \\ & + \frac{B^3}{3n+1}y^{3n+1} + \frac{3AB^2}{2n+2}y^{2n+2} + \frac{3B^2C}{2n+1}y^{2n+1} + \\ & + \frac{3A^2B}{n+3}y^{n+3} + \frac{6ABC}{n+2}y^{n+2} + \frac{3BC^2}{n+1}y^{n+1} \end{aligned} \quad (34a)$$

$$\begin{aligned} G'(y) = F^3 = & A^3y^3 + 3A^2Cy^2 + 3AC^2y + C^3 + B^3y^{3n} + 3AB^2y^{2n+1} + \\ & + 3B^2Cy^{2n} + 3A^2By^{n+2} + 6ABCy^{n+1} + 3BC^2y^n \end{aligned} \quad (34b)$$

$$\begin{aligned} G''(y) = & 3A^3y^2 + 6A^2Cy + 3AC^2 + 3nB^3y^{3n-1} + 3(2n+1)AB^2y^{2n} + \\ & + 6nB^2Cy^{2n-1} + 3(n+2)A^2By^{n+1} + 6(n+1)ABCy^n + 3nBC^2y^{n-1} \end{aligned} \quad (34c)$$

$$F(y) = \tau = Ay + By^n + C \quad (34d)$$

$$F'(y) = A + nBy^{n-1} \quad (34e)$$

Determination of other parameters is rather straight forward. Once, y_h or ψ_h are determined, the wall shear stress, radial pressure gradient and pressure functions can be determined directly from Equations (16)-(19). For instance, we can calculate the fluid velocity by direct integration of the velocity gradient.

$$\text{PF:} \quad u(x_2) = u(0) + \int_0^{x_2} \frac{\partial u}{\partial x_2} dx_2 = u(0) - h \frac{(n+1)Ay^2 + 2nBy^{n+1}}{2(n+1)F(y_h)} \quad (35a)$$

$$\text{RF:} \quad u(x_2) = u(0) + \int_0^{x_2} \frac{\partial u}{\partial x_2} dx_2 = u(0) - \frac{(n+1)a\psi^2 + 2nb\psi^{n+1}}{2(n+1)x_2f(\psi_h)} \quad (35b)$$

where $u(0)$ is the maximum velocity at the flow centre line given by

$$\text{PF:} \quad u(0) = u_0 = u_{\max} = h \frac{(n+1)Ay_h^2 + 2nBy_h^{n+1}}{2(n+1)F_h} \quad (36a)$$

$$\text{RF:} \quad u(0) = u_0 = u_{\max} = h \frac{(n+1)a\psi_h^2 + 2nb\psi_h^{n+1}}{2(n+1)x_2f_h} \quad (36b)$$

Hence, the velocity profile across the flow cross-section is given by

$$\text{PF:} \quad u = u_{\max} \left(1 - \frac{(n+1)Ay^2 + 2nBy^{n+1}}{(n+1)Ay_h^2 + 2nBy_h^{n+1}} \right) \quad (37a)$$

$$\text{RF:} \quad u = u_{\max} \left(1 - \frac{(n+1)a\psi^2 + 2nb\psi^{n+1}}{(n+1)a\psi_h^2 + 2nb\psi_h^{n+1}} \right) \quad (37b)$$

The average flow velocity can also be determined in the usual manner by integrating (36) directly, namely

$$\bar{u} = \frac{1}{\int_0^h x_2 dx_2} \int_0^h u(x_2) x_2 dx_2 = u_{\max} (1 - \lambda_0) \quad (38)$$

Where λ_0 is a function of y_h , or ψ_h (in the case of radial flow [4]).

$$\text{PF:} \quad \lambda_0 = G_0(y_h) = 1 - \frac{y_h F_h - G_h F_h^{-2}}{\frac{3}{2} A y_h^2 + \frac{3n}{n+1} B y_h^{n+1}} \quad (39a)$$

$$\text{RF:} \quad \lambda_0 = g_0(x_h) = \frac{1 + n_1 a^{-1} b \psi_h^{n-1} + n_2 a^{-2} b^2 \psi_h^{2(n-1)}}{3 + n_3 a^{-1} b \psi_h^{n-1} + n_4 a^{-2} b^2 \psi_h^{2(n-1)}} \quad (39b)$$

In Equation (39b) n_i is a constant depending only on the power factor n , given by

$$n_1 = \frac{3n(n+3)}{(n+1)(n+2)}, \quad n_2 = \frac{6n^2}{(n+1)(2n+1)}, \quad n_3 = 3 \frac{(3n+1)}{(n+1)}, \quad n_4 = \frac{6n}{(n+1)} \quad (39c)$$

All the above solutions (e.g. Equations (32)) are also reducible to the classical solutions. For example, the average flow velocity becomes half of the maximum flow velocity for pipe flow, and 2/3 of the maximum flow velocity in the case of radial flow, for the case of pure Newtonian fluid [4-5].

Slurry flow may be assumed to stop in the case of a blockage ($Q \rightarrow 0$), which means the values of y_h and $g(y_h)$ are identically zero. This is due to the effects of the cohesive frictional terms (ξ and τ_0) introduced in the shear stress Equations (16-17), which now become dominant in blocking the slurry flow. In the slurry industry, a critical question always arises on what the minimum pump pressure is required for a given slurry flow rate either to transport it to a given distance, or be able to reopen a blockage in a specified pipe length. The minimum required pump pressure can be calculated from Equation (21), which depends on the wall shear resistance in the pipeline or the disk. The wall shear stress is a function of the longitudinal distance x and velocity gradient y_h . Practically, during the field slurry injection, the minimum required pump pressure to transport the slurry to a given distance is one of the most important questions that needs to be addressed [4-5].

6. Tremie pipe concrete flow

The flow theory developed for general viscous-plastic-frictional fluids can be applied to fresh wet concrete pastes and slurries as well. Again, an important question would be the relation between the flow rate, Q , and the pressure, p . Similar to the pipe flow discussed above, we may assume a fully developed laminar, one dimensional flow, where x_2 is the radial distance from the tremie pipe's axis of symmetry (Figure 1, middle) [6-9]. Gravity plays a major role as the main driving force for concrete flow in tremie pipes, thus, it cannot be ignored. Therefore, equation (12) becomes

$$\frac{\partial(x_2\tau)}{\partial x_2} = x_2 \left(\frac{\partial p}{\partial x_1} + \gamma_c \sin(\theta) \right) \quad (40)$$

It is assumed that the tremie pipe generally makes an angle θ with the horizontal x axis, where $\theta = 90^\circ$ indicates a vertical tremie. In Equation (40), x_2 is the coordinate radius or radial distance from the pipe's cross-sectional centre, p is pressure and γ_c is the effective concrete unit weight. As a special case, a specific analytical solution from the general solution (32a) can be derived for a linear Bingham-plastic model [5]. In this particular solution, the tremie pipe flow rate, Q , becomes a 4th order polynomial function of the tremie pipe diameter, D . Furthermore, the flow rate is inversely proportional to the viscosity, μ_0 , and (partially) proportional to the differential pressure at the two ends of the tremie pipe [6], Δp , of length L_t . In other words we have:

$$Q = \frac{\pi D^4}{128\mu} \left(\frac{\Delta p}{L_t} + \gamma_c \sin(\theta) \right) - \frac{\pi D^3}{24\mu} \tau_0 \quad (41)$$

The pressure differential between the two ends of the tremie, Δp , can in theory accept any arbitrary value; from negative to zero and positive numbers. In the case of a zero Δp , the driving pressure is simply the gravity term containing the concrete unit weight, γ_c , and inclination angle (θ).

7. Concrete flow in a rectangular channel

Concrete flow during pouring and flowing in channels, chutes and testing equipment for testing purposes are normally not at a steady state situation [6]. General time-dependent 2D and 3D differential equations governing flow of concrete in rectangular channels and chutes can be developed and solved numerically, as shown in [6]. However, for the sake of understanding, it is also possible to reduce these equations to a simple 1D form, based on an assumption that there is no significant independent variation in any variable or function in the normal directions x_2 and x_3 compared to the longitudinal main flow direction x_1 . In other words,

$$\tau_{,x_2}(t,x_1) - p_{,x_1}(t,x_1) - \rho u_{,t}(t,x_1) = 0 \quad (42)$$

which gives a solution in terms of Fourier coefficients [6]

$$u(t,x) = \sum_{n=0}^{\infty} e^{-a_n \kappa_n^2 t} (A_n \cos(\kappa_n x) + B_n \sin(\kappa_n x)) + mx^2 \quad (43)$$

In the solution (43) κ is an arbitrary constant satisfying both the differential equation and the boundary conditions, while A_n and B_n are Fourier coefficients to be determined from the boundary conditions [6]. Figure 7 shows a typical result for various values of n truncating the number of Fourier terms. It shows results of the Fourier analysis for the two cases of $u(0, x)$ and $u(0.5, x)$, and the increasing effects of the number of Fourier terms, namely from $n = 5, 10$ to 120 . The second line in the figure corresponds to velocity at time $t = 0.5$ for different profile points along the x line using $n = 120$. Notice that since continuity and differentiability is not a requirement at the end points of a Fourier series analysis, it doesn't converge to the numerical solution at point $x = 1$, as expected.

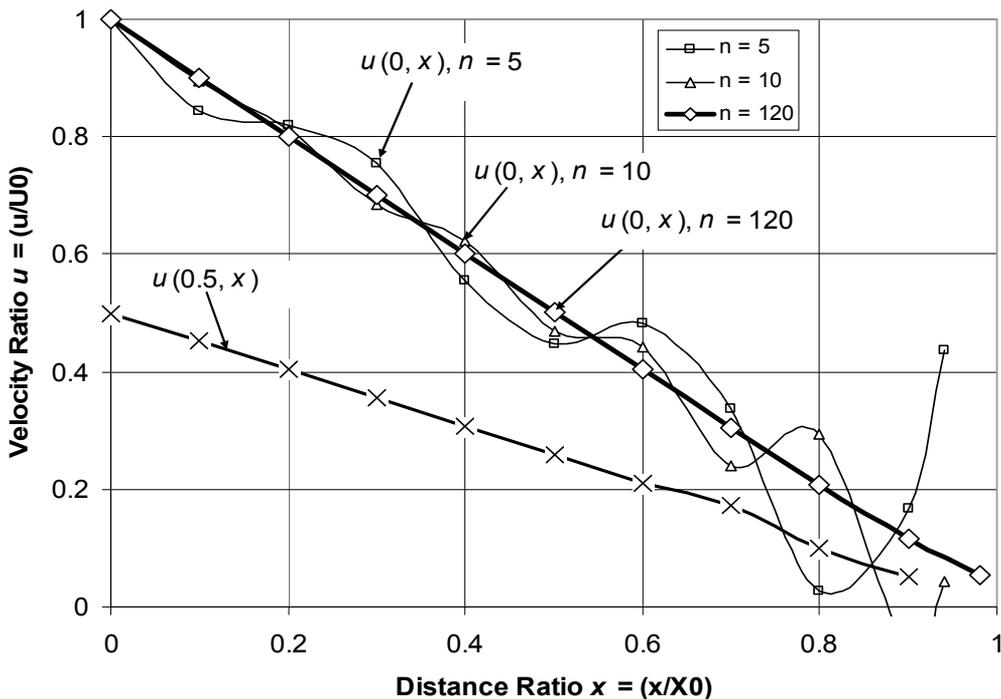


Fig. 7. The function $u(0, x)$ represented by a Fourier series with different number of Fourier coefficients ($n = 5, 10, 120$).

8. Discussion

The above general theory is certainly reducible to simpler classical Newtonian and Bingham models with appropriate parameter substitutions [4-5]. Classical special cases can be derived, e.g. (i) pure, uniform, viscous, Newtonian slurry; (ii) Pure, uniform, cohesive (plastic), non-Newtonian slurry; (iii) Linear Bingham viscous plastic slurry. In case (i), (ii) and (iii) the shear stress function (1) reduces to either (i) the simplest, classical, linear function of the shear strain multiplied by a constant viscosity number, i.e. $\tau(x_1, x_2) = \mu_0 y$; or (ii) just a pure plastic material with no viscosity, i.e. $\tau(x_1, x_2) = \tau_0$; or a linear Bingham viscoplastic model, i.e. $\tau(x_1, x_2) = \mu_0 y + \tau_0$. For instance, for a pipe flow in a pure viscous Newtonian fluid, we have

$$\tau_{rx}(r,x) = \mu(x) \left(-\frac{\partial u(r,x)}{\partial r} \right) = \mu y, \quad \mu(r) = \mu_0, \quad \eta = \tau = \xi = n = 0 \quad (44a)$$

Substituting these values in equations (16-32), we recover the well-known Newtonian solutions [15], as expected. It can be seen that for this case the function $g_0(y_h) = \frac{1}{2}$ confirming the classical result, $\bar{u}_r(r) = \frac{1}{2} u_{r,max}$. Furthermore, we have:

$$g_h = \frac{\mu_0^3}{4} y_h^4 \quad (44b)$$

$$g_0 = \frac{1}{2} \quad (44c)$$

$$y_h = \frac{4Q}{\pi h^3} \quad (44d)$$

$$\tau_{hx} = \frac{4\mu_0 Q}{\pi h^3} \quad (44e)$$

$$\left| \frac{dp}{dx} \right| = \frac{8\mu_0 Q}{\pi h^4} \quad (44f)$$

$$P_0 - P = \frac{8\mu_0 Q}{\pi h^4} L \quad (44g)$$

$$u(0,x) = u_{max} = \frac{2Q}{\pi h^2} \quad (44h)$$

$$u(r,x) = u_{max} \left(1 - \frac{y^2}{y_h^2} \right) = u_{max} \left(1 - \frac{r^2}{h^2} \right) \quad (44i)$$

$$\bar{u} = u_{max} (1 - g_0) = \frac{1}{2} u_{max} \quad (44j)$$

Figure 8 shows velocity profiles in normalised form for the “three special cases” discussed above. For the general Bingham fluid with constant non-zero values of τ_0 and μ_0 , the pipe velocity profile follows a parabolic curve close to the special pure viscous case (i), when fluid viscosity is dominant or the ratio τ_0/μ_0 is small, and moves towards the uniform profile of the special case (ii), when plastic yield or cohesion is dominant or the ratio τ_0/μ_0 is large, as shown in the figure. Figure 9 demonstrates an example of a Bingham plastic solution for radial disk flow, where the contribution of each of the two shear parameters is separately demonstrated.

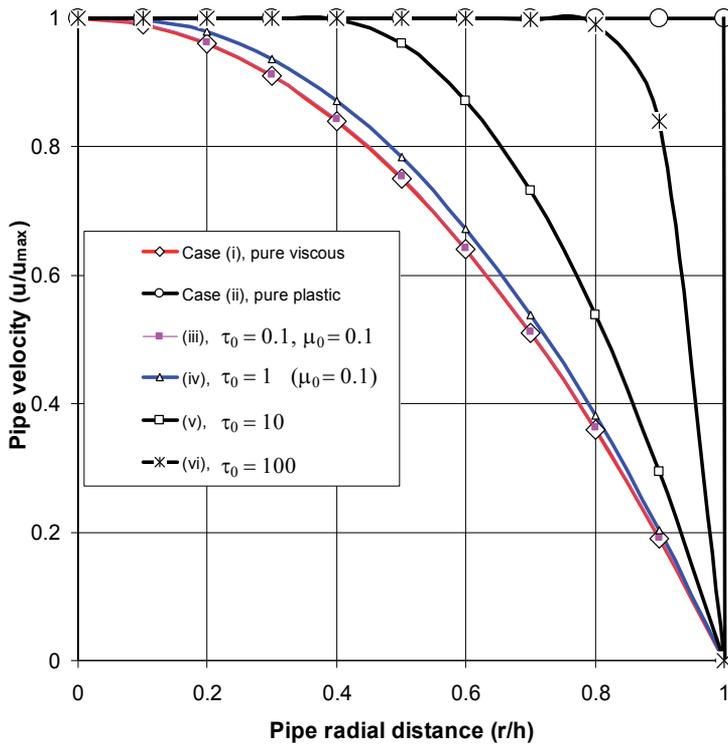


Fig. 8. Comparisons of normalised velocity profiles for different slurries of various viscosity (μ) and plasticity (τ) in a pipe flow

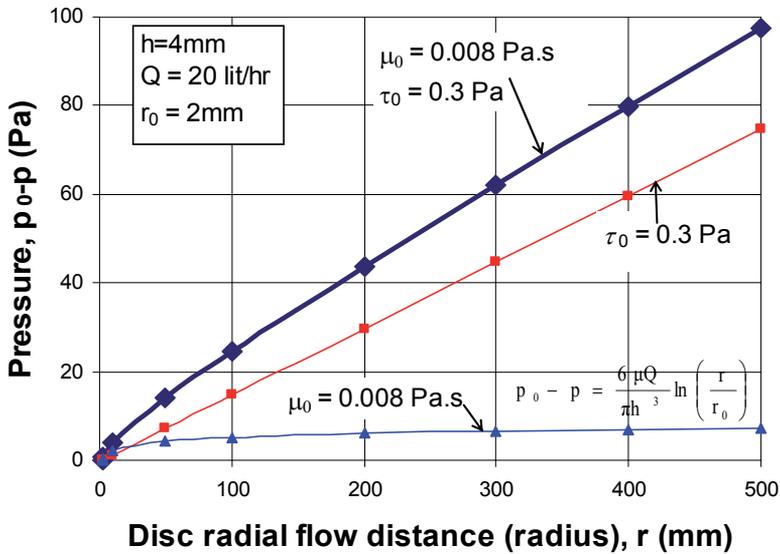


Fig. 9. Contribution of viscosity and cohesion to pressure drop for a Bingham plastic slurry in a radial disk flow

As a numerical pipe flow example, consider a slurry modelled by a linear Bingham plastic model, where $\mu_0 = 0.1\text{Pa}\cdot\text{s}$ and $\tau_0 = 0.1\text{Pa}$. As shown in Table 1, the maximum and average velocities in the pipe are 3.54 m/s and 1.76 m/s, respectively. The table also shows a list of values for several other variables and parameters used in the present theory.

Y_h (1/s)	$f(y_h)$	$g''(y_h)$	$g'(y_h)$	$g(y_h)$	$-dp/dx$ (Pa/m)	u_{\max} (m/s)	$g_0(y_h)$	u_{ave} (m/s)	τ_{hx} (Pa)	P_{min} (Pa)
141.804	14.280	61.179	2912.220	103969.296	571.217	3.504	0.498	1.760	14.280	571217.381

Table 1. Numerical example for cohesive-viscous slurry ($\mu = 0.1 \text{ Pa}\cdot\text{s}$, $\tau = 0.1$)

Slurry behaviour is controlled by its two distinct material components, i.e. the solid particles and the water. Depending on the velocity of the fluid and the terminal velocity and physical characteristics of the suspending solid particles, the slurry behaviour can evolve by two distinct characteristics, a uniform viscous fluid or a fluid with separated, submerged, sedimentation deposit; where the latter is the favourite mechanism in mining grout injection. The solid particle concentration or viscosity is constant in the former case and (increasingly) variable in the latter. The more the concentration of the particles, the greater is the effect of the frictional viscosity, as observed in our direct viscosity measurements and also consistent with the empirical equations.

When working with slurries made of particulate and granular materials for injection operations in the field, it is quite possible to encounter pipe blockage. This is when the last term in Equation (7) or (14) becomes non-zero and hence dominates the process due to high frictional shear resistance against the slurry flow. Several laboratory blockage tests have been carried out to confirm the role and effects of this frictional term in Equation (14). In these experiments, initially the pump pressure was reduced gradually during an injection process to reduce flow velocity causing settlement and sedimentation of the grains until full blockage has occurred. An attempt to reopen the same blockage was made by increasing the pump pressure. However, a much higher than the initial pump pressure was required to reopen the blockage, confirming the effect of the frictional term in Equation (14). Figure 10 demonstrates how the pressure can increase rapidly before or behind a blockage, resulting in a substantial head loss. This theoretical exponential trend agrees with similar experimental measurements reported in the literature [5].

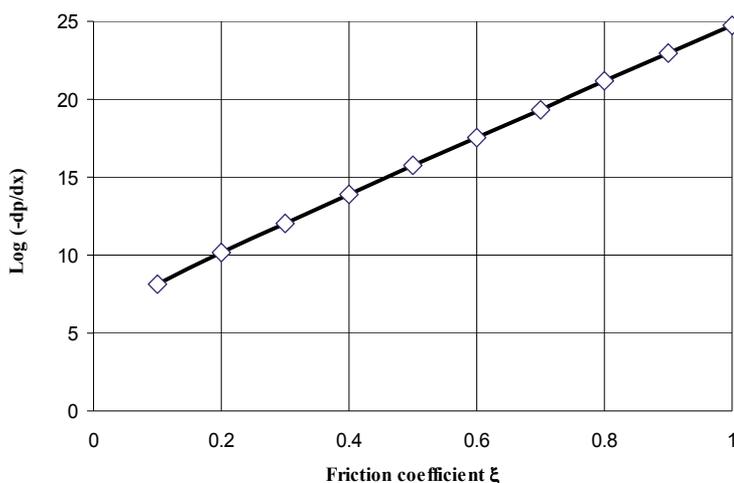


Fig. 10. Effect of frictional coefficient on pressure gradient for a given slurry

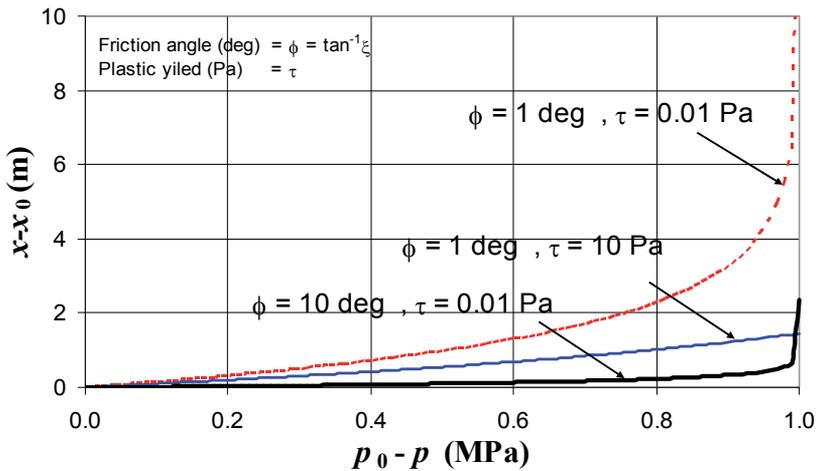


Fig. 11. Required pipe distance (from x_0 to x) is nonlinearly proportional to pressure (p_0 to p) for a linear uniform pipe. The slope of the relation depends on frictional-cohesive properties of the granular materials of the slurry. The higher the friction or cohesion is, the smaller the required distance at a given pressure differential is.

In practice, a blockage is usually reopened by pumping a less viscous fluid (e.g. water) at a very high pump pressure and minimum viscous shear resistance. The pump pressure required is a function of not only frictional properties of the deposited sediment, but also the size distribution of the aggregates (Figure 11).

9. Conclusions

On the basis of continuum equations of fluid and soil mechanics, a comprehensive, versatile, slurry shear model has been developed for transportation of grout, paste and fill materials used in the civil and mining industries, covering a wide range of material characteristics and behaviour, namely from the flowing fluid slurries to consolidated solid deposits in underground coal mining induced rock fractures. The theory has been specifically tailor made for grout flows through uniform pipes, discs and tremies, in order to transport material to designated injection or backfill targets. The theory can mimic both flow and blockage behaviour of the fill material. The tool can be used to predict variations of pressure and velocity and their gradients, as a function of flow rate, in the entire backfill-placement system from batching plant to the borehole cracks and foundation excavations.

The shear theory can mimic shear resistance of both: (i) a cohesive, viscous flow and (ii) a stationary, cohesive, pressure-dependent, frictional, plastic soil. The pressure dependent frictional term in the shear stress model determines the frictional resistance of the deposited fill material during a blockage. Consistent with laboratory and field experiments, the theoretical pump pressure required to open a blockage is orders of magnitude greater than the amount needed for pumping the same material when it is under a steady state flow. This explains why very high pump pressures are often needed to clean blockages compared with much lower pressures required during steady state slurry flows.

Concrete flow and placement into deep foundations is normally performed under several harsh environmental conditions of tightness, inaccessibility and deep submergence.

Therefore, it must be self compacting, self levelling and maintain its original quality, homogeneity and integrity all the way from the tremie pipe to the discharge point and then through the narrow paths between heavy reinforcements. Similar to a viscous-plastic slurry or paste, shear behaviour of a fresh tremie pipe concrete was explained by a linear Bingham plastic model. Traditional slump and spread tests together with the L-box tests are used as indirect index tests to measure physical visco-plastic properties of concrete. However, the concrete industry needs also to develop a large scale viscometer testing method to measure viscosity and plastic yield of tremie pipe concrete directly. Based on the Bingham parameters, the governing relation between the steady state concrete flow rate and the required pressure gradient was presented. To maintain a successful, uniform, steady state flow in the tremie pipe, a balance pressure height must be determined and controlled through the entire process of concrete pour or discharge.

Appendix

Notation

In the following derivations, italic symbols are used for normalised parameters, or quantities representing their counterparts denoted by the same non-italic symbols. For instance, $r = r/r_0$ represents the normalised form of the radial distance variable r with respect to a reference distance r_0 , i.e. the radius of the central vertical pipe.

Italic symbols indicate normalized, or dimensionless quantities, e.g. the fluid velocity, $u = u/U_0$ is the dimensionless form of the dimensional counterpart quantity, u .

a, b, c	Shear stress function coefficients
A, B, C	Integral function coefficients
A_i, B_i, C_i	Constants of Bingham plastic solution
C	Solid concentration by weight or mass = $C_{\text{weight}} = C_{\text{volume}} (\rho_{\text{solid}}/\rho_{\text{mix}})$
f	A polynomial function of fluid velocity gradient ($f(y) = ay + by^n + c$)
g	A polynomial function of fluid velocity gradient ($g(y) = \int f^3 dy$)
$()_h$	Index "h" denoting function value at either pipe or disk boundary walls
$()_0$	Index "0" denoting initial or constant value of a variable
L	Pipe length ($L = x - x_0$) from reference section x_0 to any section x .
n	Shear strain power factor, Function power factor
\mathbf{n}	Normal vector
p	Fluid pressure
Q	Flow rate
$Q_h = \frac{3}{nh^3} Q$	A flow rate related constant
r	Radial distance from pipe centre; polar r coordinate axis, disk radial distance from borehole or disk centre in a radial flow
r_0	Radius of central vertical pipe connected to disk
$x_1, x_2, x_3 = \theta$	Subscripts indicating longitudinal (radial in disk flow), normal to flow cross-section and circumferential coordinates, respectively
x	Longitudinal distance or coordinate axis along pipe length
x_0	Initial reference point in a pipeline section along x_1 direction

$u = u_1$	Longitudinal velocity (u_1) in both pipe and disk radial flow
\mathbf{u}	Velocity vector with velocity components ($u_1, u_2, u_3=u_0$)
u_t	Terminal velocity or free fall, submerged solid particle limit speed
u_D	Deposition velocity or particle speed at minimum pressure gradient
μ	Viscosity coefficient (linear term)
v	Volume concentration of solids in slurry mix
$\Psi = -u x_1$	Radial velocity times radial distance function
$\psi = \Psi' = y x_1$	Derivative of Ψ with respect to x_2
D/Dt	Total time derivative
$df/dt = \dot{f}$	Local time derivative of a function f
$d/d\zeta_j$	Derivative with respect to a coordinate axis ζ_j
$\partial/\partial\zeta_j$	Partial derivative with respect to a coordinate axis ζ_j
$\bar{\nabla}$	Gradient vector
w	Weight concentration of solids in slurry mix
ψ	-Ratio of disk radial velocity gradient and x_2 , dY/dx_2
Y	-Radial velocity times distance x_1 , integral of ψ
α	Generic symbol representing either one of functions: η, μ, τ, ξ
$\gamma = \partial u_1 / \partial x_2$	Shear strain rate (velocity gradient)
η	Viscosity coefficient (non-linear)
μ	Viscosity coefficient (linear)
ρ	Density, slurry density
v	Volume concentration of solids in slurry mix
θ	Circumferential (hoop) coordinate axis
τ	Shear stress, cohesion (yield stress), stress tensor
$\xi = \tan(\phi)$	Friction coefficient, ϕ = friction angle
<i>Subscripts</i>	
0	Initial value, reference value for normalisation
∞	Final far field value of a property
α	Value corresponding to property α
μ, η, τ, ξ	Value corresponding to properties μ, η, τ, ξ , respectively
$x = x/x_0$	Gradient velocity divided by radius, $x_0 = r_0 u_{r0} / h = X_0 / h$
$\mu = \mu/\mu_0$	Viscosity coefficient (linear)
$\eta = \eta/\eta_0$	Viscosity coefficient (nonlinear)
$\tau = \tau/\tau_0$	Shear stress, cohesion
$\xi = \xi/\xi_0$	Friction coefficient
p	= fluid pressure function
Q	= volume rate or fluid flow rate
τ_0	= plastic yield or cohesion intercept in linearised Bingham plastic model
τ	= shear stress function of viscosity, plastic yield and shear strain gradient
u	= fresh concrete or fluid velocity, x-component of velocity in 1D model
u_y	= y component of fluid velocity in 2D model
$u_{,y} = \frac{\partial u}{\partial y}$	= y gradient of velocity in 1D model, shear strain rate

$u_{,t} = \frac{\partial u}{\partial t}$	= fluid acceleration or velocity rate
U_0	= fresh concrete velocity at L-box entrance (reference velocity)
u or u^*	= dimensionless flow velocity function in L-box test
u_i^k	= finite difference velocity function u at time k and position i
v	= a dimensionless function representing fluid velocity
W	= width of L-box (in out of plane z direction)
x	= x axis and flow direction in tremie pipe and L-box test
x or x^*	= flow direction and distance in L-box test
\mathbf{x}	= position vector with components x, y, z
$\hat{\mathbf{x}}$	= vector normal to position vector for velocity gradient calculations
X_0	= maximum concrete flow distance in L-box test (reference length)
$\xi(t, \mathbf{x})$	= friction coefficient function when concrete blockage occurs
y	= y -axis coordinate; vertical position in L-box test; pipe flow radial direction
y_1, y_2	= two end coordinates of L-box horizontal open channel
Y_D	= height drop along L-box horizontal open channel ($y_1 - y_2$)

10. Acknowledgements

This work was conducted within the Subsidence Control Research Program of CSIRO. Industry support from BHP Billiton, ACARP (C16023; C12019) and CIA of Australia is gratefully acknowledged. The Author would like to thank CSIRO colleagues Dr Jane Hodgkinson and Dr Cameron Huddleston-Holmes for their review of the initial manuscript.

11. References

- [1] Govier, G.W., Aziz, K. 1972. *The flow of complex mixtures in pipes*, Van Nostrand Reinhold Co, New York.
- [2] Middleman, S., 1977. *Fundamentals of Polymer Processing*, McGraw-Hill, NY.
- [3] T. Yen Na and A.G. Hansen, 1967, "Radial flow of viscous non-Newtonian Fluids between disks", *Int. J. Non-linear Mechanics*. Vol. 2, pp. 261-273. Pergamon Press Ltd., Great Britain.
- [4] Alehossein H. 2009, "Viscous, cohesive, non-Newtonian, depositing, radial slurry flow", *International Journal of Mineral Processing*, V. 93 No. 1, 2009, pp. 11-19.
- [5] Alehossein H., Shen, B., Qin, Z. and Huddleston-Holmes, C.R. 2012, "Flow analysis, transportation and deposition of frictional, viscoplastic slurries and pastes in civil and mining engineering", *ASCE, Journal of Materials in Civil Engineering*.
- [6] Alehossein, H., Beckhaus, K. and Larisch, M. 2012, "Analysis of L-Box test for tremie concrete", *ACI, Journal of American Concrete Institute*.
- [7] Beckhaus, K., Larisch, M., Alehossein, H., Ney, P., Northey, S., Lucas, G., Dux, P., Buttling, S., Lucas, G., Vanderstaay, L. 2011. *Tremie Concrete for Deep Foundations*. Recommended Practice published by the Concrete Institute of Australia (CIA).

- [8] Beckhaus, K. and Larisch, M., Alehossein, H., Northey, S., Ney, P., Lucas, G., Vanderstaay L. 2011, "New approach for tremie concrete used for deep foundations", ICAGE conference, Perth 2011.
- [9] Beckhaus, K. and Larisch, M., Alehossein, H., Northey, S., Ney, P., Lucas, G., Vanderstaay L. 2011, "Performance-based specifications for concrete, 14 - 15 June 2011, Leipzig, Germany
- [10] Alehossein, H. 2009, "A triangular model of caving subsidence", *J. Applied Earth Science (Trans. Inst. Min. Metall. B)*, Vol 118, No 1, pp. 1-4.
- [11] Shen, B. and Alehossein, H. 2009. *ACARP Project C16023, Australia - Subsidence Control Using Coal Washery Waste*. An extension research project to ACARP C12019 in 2003 on "Subsidence control using overburden grout injection technology".
- [12] Alehossein, H., Poulsen, B. A. 2010, "Stress analysis of longwall top coal caving". *International Journal of Rock Mechanics and Mining Sciences*, 47(1), 30-1.
- [13] Alehossein, H. 2010, "Mechanics of slurries for rock fractures & voids", WCHFA-2010, August 2010, Xian, China
- [14] R.F. Craig, 1997. *Soil Mechanics*. E & FN SPON, Routledge. New York.
- [15] Fox, R.W., Pritchard, P.J. and McDonald, A.T. 2009. *Introduction to fluid mechanics*, Don Fowley (Wiley & Sons), New York.
- [16] Alehossein, H. and Hood, M. 1999, "Application of linearized dimensional analysis to rock cutting", *International Journal of Rock Mechanics and Mining Sciences*, V. 36, 1999, pp. 701-709.

Three Dimensional Simulation of Gas-Radiation Interactions in Gas Lasers

Timothy J. Madden

US Air Force Research Laboratory, Directed Energy Directorate

Kirtland Air Force Base, New Mexico

USA

1. Introduction

The spectroscopically measured lineshape of an atomic transition provides a wealth of useful information relative to diagnosing the state of a gas. The center of the lineshape is specific to a particular transition of a specific atom. The width of the lineshape indicates the amount of broadening of the transition, due to the effects of both collisions with other particles in the gas and Doppler shift due to the movement of the atom. Since the Doppler shift is proportional to velocity, the width of the transition can be used to estimate the degree of random molecular motion in the gas, expressed macroscopically as temperature. A Doppler shift to the frequencies in the transition can also occur through the bulk motion of the gas, and this can be used to examine the velocity field of the gas. The astronomy community was the first to recognize the utility of these concepts in practical applications, stemming to the early 1930s.^{1,2} In 1934 Stuve and Elvey³ showed that by including a bulk gas velocity Doppler broadening term in the Voigt equation for the transition lineshape in addition to the random thermal motion term, it was possible to estimate whether a stellar atmosphere was 'turbulent' or not based upon the fit of the Voigt equation to the measured transition lineshapes for that atmosphere. Taking the theory a step further, using estimated optical paths or length scales for the stellar atmospheres that they were measuring, they were able to estimate median gas velocities and correlate increasing velocity magnitude with increasing temperature of the stellar atmosphere.

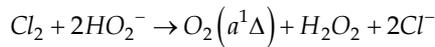
More recently, continuously tunable diode lasers have been applied to lineshape measurement of transitions within species in the chemical oxygen-iodine laser (COIL) flowfield as a means to determine number density and laser gain on the I $2P_{1/2} \rightarrow 2P_{3/2}$ transition. Davis and Allen et al^{4,5,6,7} applied lineshape measurement and Voigt fitting methods to various COIL species as a mechanism to determine concentration and translational temperature in the COIL flowfield as an experiment diagnostic. While the effect of the bulk gas velocity was not taken into account in these investigations, the diagnostics developed by these investigations provided the means to do so. Nikolaev et al,⁸ applied the same type of continuously tunable diode laser to the I $2P_{1/2} \rightarrow 2P_{3/2}$ transition to investigate the influence of mean gas velocity through a COIL device upon the laser gain. They showed that by varying the angle at which the diode laser beam passes through the COIL flowfield, the line center gain can be varied through bulk velocity induced Doppler broadening. Using this mechanism, they were able to determine the mean flow velocity for the COIL and convert the laser gain

diagnostic into a flow diagnostic. Similarly, they recognized the propensity for ‘turbulence’ to contribute to the broadening of the lineshape. Invoking an assumption of isotropic turbulence to characterize the bulk vorticity in the flows that they were measuring, they examined the lineshape data using basically the same theory developed by Stuve and Elvey and estimated a 10 K effect for a mean turbulent fluctuation velocity of 30 m/s. Given experimental evidence that temperatures determined from experimentally measured lineshapes for the $I^2P_{1/2} \rightarrow ^2P_{3/2}$ transition in COILs may be as much as 50 K higher than what is expected based upon thermodynamic predictions, the implication is that either the magnitude of the flow fluctuations is much larger than 30 m/s or some other process such as H_2O condensation is liberating heat and increasing the temperature.

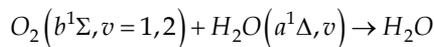
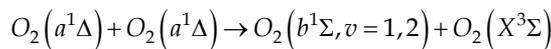
The work presented here examines the issue of the interplay between an optical field in a COIL and the bulk motion of the gas through simulation of the physics. Previous work by Madden and Miller⁹ presented data from COIL simulations indicating that the flowfield should be marked by substantial flow unsteadiness, as documented in the experiments of Fric and Roshko¹⁰ for fluid dynamically similar types of flows. These simulations provide the capability to examine the optical field to gas interaction in the lineshape directly, including the full physics of the fluid dynamics, chemistry, heat release, and flow instability without resorting to statistical turbulence theory. Direct prediction of the fluid dynamic physics allows for examination of the magnitude, and spatial and temporal characteristics of the flow fluctuations, an important consideration given the presence of both strong compressibility which stabilizes and heat release which may either stabilize or de-stabilize the flow. From this perspective, it is not only possible to characterize the optical field to gas interaction through the $I^2P_{1/2} \rightarrow ^2P_{3/2}$ transition lineshape, but also to assess whether the flow is actually turbulent or not.

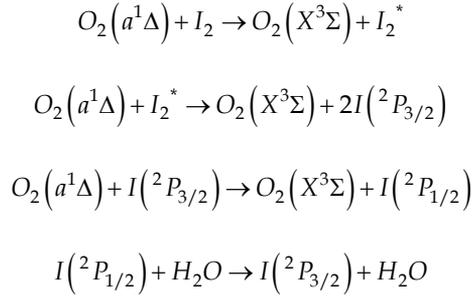
2. Problem and methodology

Gas lasers operate through a series of processes, often exothermic, that lead to a population inversion, or non-Boltzmann energy distribution, within the energy states of the species in the gas. In COILs, two separate but coupled non-Boltzmann energy distributions are created. The first is through the liquid phase reaction of chlorine with the hydroxyl ion generating a population inversion within oxygen with the electronically excited state of singlet delta oxygen:



The $O_2(a^1\Delta)$ subsequently diffuses out of the liquid phase into a carrier gas, He, and is transported toward the mixing and reaction region upstream of the resonator. During transport, energy pooling processes erode the $O_2(a^1\Delta)$ population through pooling and deactivation. Further downstream, a mixture of molecular iodine, I_2 , and a diluent gas, He, is injected into the $O_2(a^1\Delta)$ carrying flow and reacts with the singlet oxygen. The combination of these pooling and reaction processes can be briefly summarized as:





The final product of the COIL chemical kinetics processes is the $^2P_{1/2}$ electronically excited state of atomic iodine, a second population inversion in atomic iodine that results from the first in O_2 . The flow unsteadiness associated with the jet-in-crossflow interaction of the secondary He/ I_2 flow with the primary $O_2(a^1\Delta)$ carrying flow leads to the generation of vortices that are transported downstream into the supersonic region where the laser cavity is located. These vortices through their rotation have associated velocity components that are parallel to the transit of the optical field. The motion of the bulk flow in the direction of the optical field leads to Doppler broadening of the atomic iodine (I) $^2P_{1/2} \rightarrow ^2P_{3/2}$ transition lineshape. This effect is important in two ways: the first is the decrease in the line center gain of the $^2P_{1/2} \rightarrow ^2P_{3/2}$ transition and the second is in the breadth of the lineshape. The first effect is important relative to the performance of the laser, but the second effect is most important relative to the application of spectroscopic lineshape measurement. Understanding the influence of the vortices upon the breadth of the I $^2P_{1/2} \rightarrow ^2P_{3/2}$ transition lineshape is therefore important to understanding the accuracy of the temperature diagnostic in COIL experiments. Additionally, the magnitude of the flow fluctuations, their temporal variation, and their scale all are related to the question of whether the flow is fully turbulent or not. Relative to the modeling of the gas flows through COIL devices, these flows can best be described as the translation of particles of different chemical composition with collisional interactions occurring between the particles and between the particles and the photons within the radiation field. Mathematically, this flow of particles is treated as a continuum and is approximated by Navier-Stokes continuity equations for mass, momentum, and energy. In integral form, these equations are given by:

$$\frac{\partial}{\partial t} \iiint_{vol} \langle Q \rangle dvol + \iint_A (\vec{F} - \vec{F}_v) \cdot \hat{n} dA = \iiint_{vol} \langle S \rangle dvol$$

The production rates of the species due to chemical reactions are determined for each species continuity equation using the Arrhenius rate law for the forward rates, with the backward rates determined from the equilibrium and forward rates. The computational fluid dynamics (CFD) code GASP from AeroSoft, Inc. is employed to solve the Navier Stokes equations for the simulations performed here. The GASP COIL model utilizes 10 to 13 species mass conservation equations for the chemically reacting components of the COIL flow, depending upon the COIL model used, in addition to the base conservation equations for momentum and energy. An effective binary diffusion model is used to describe concentration and pressure contributions to mass diffusion, an important process in the low density COIL flowfield. 13 species, 52 reactions and 10 species, 20 reactions finite-rate chemistry mechanisms¹¹ are used to model the gas phase chemical kinetic processes that generate the population inversion in atomic iodine necessary for laser oscillation in

COILs. These models utilize an empirical construct for I_2 dissociation in the presence of $O_2(a^1\Delta)$ developed by Heidner¹² that has been utilized within the COIL community for 25 years. Recent work by Azyazov et al¹³ indicates that the I_2 dissociation process is somewhat more complex than the simplified construct developed by Heidner. However, work by Waichman¹⁴ indicates that the difference in power predictions from CFD simulations using the new mechanism from Azyazov and the original Heidner mechanism is of order 10% for what is termed ‘high pressure operation,’ i.e. essentially optimal operation with regard to the I_2/O_2 ratio, and similar to the conditions considered in this investigation. From this standpoint, the 10% difference between the I_2 dissociation mechanisms for the ‘high pressure,’ optimal COIL operational conditions simulated here is not expected to substantially alter interpretation of the results given here.

The laser gain and the associated lineshape for the $I^2P_{1/2} \rightarrow P_{3/2}$ transition is determined from the flowfield calculated variables using the equations:

$$\alpha(v_0) = \frac{7}{12} \left(\frac{A\lambda_0^2}{8\pi} \right) \phi(v_0) \left(N_{I(^2P_{1/2})} - \frac{1}{2} N_{I(^2P_{3/2})} \right)$$

$$\phi(\omega(v)) = \left[\left(\frac{a^2 4 \ln 2}{\pi^3} \right)^{1/2} \frac{1}{\Delta v_D} \int_{-\infty}^{+\infty} \frac{e^{-y^2}}{(\omega - y)^2 + a^2} dy \right]$$

$$\omega = \frac{2 \left(v - v_0 \left(1 + \frac{W}{c} \right) \right)}{\Delta v_D} \sqrt{\ln 2}$$

$$a = \frac{\Delta v_L}{\Delta v_D} \sqrt{\ln 2}$$

$$y = \frac{2v_z v_0}{c \Delta v_D} \sqrt{\ln 2}$$

$$\Delta v_D = \frac{2}{\lambda} \sqrt{\frac{2RT \ln 2}{m_i}}$$

$$\Delta v_L = \frac{T_{ref}}{T} P \sum_{i=1}^N \alpha_i \chi_i$$

A difference between the approach to including the effect of the bulk flow velocity components on the lineshape used here and that used by Stuve and Elvey and later by Nikolaev is that they are not assumed to be described by a Gaussian probability distribution. Here, the velocity distribution is directly determined through prediction of the velocity components including their scale, magnitude, and temporal variation.

The computational grid used for the COIL simulations consists of 29 blocks and 8.1 million grid cells. The computational domain which this grid discretizes represents the smallest geometrically similar element within the flowfield in the COIL experiment hardware, denoted a ‘unit-domain.’ The unit-domain consists of a supersonic ($M \sim 2.2$) converging-diverging nozzle

section with one large and two small injector orifices that issue reactants into the primary flow through this nozzle. It was previously shown¹⁵ that this particular configuration of the 'unit-domain' as opposed to one half as wide splitting the large injector orifice is necessary to properly capture unsteady fluctuations in the lateral or Z direction relative to Figure 1. Fig. 1 illustrates the unit-domain computational grid within the context of a surface rendering of the experiment's mixing nozzle. The orifices inject a sonic mixture of He and I₂ into a subsonic primary flow composed of He, O₂(a¹Δ), O₂(X³Σ), H₂O, and Cl₂ inducing the complex 3-D flow structure as the jet issuing from the orifice interacts with the crossflow. It is the combination of the interfacial area created by the complex 3-D flow structure and molecular diffusion that mixes the two flows. Boundary conditions accomplish the unit-domain approximation through the enforcement of planar symmetry at the nozzle centerline in the vertical direction and periodicity at the side boundaries in the lateral direction. No-slip constant temperature boundary conditions are used at the wetted surfaces of the nozzle and orifices, with the temperature fixed at 400 K at the orifice region walls and 300 K at the nozzle walls. The nozzle

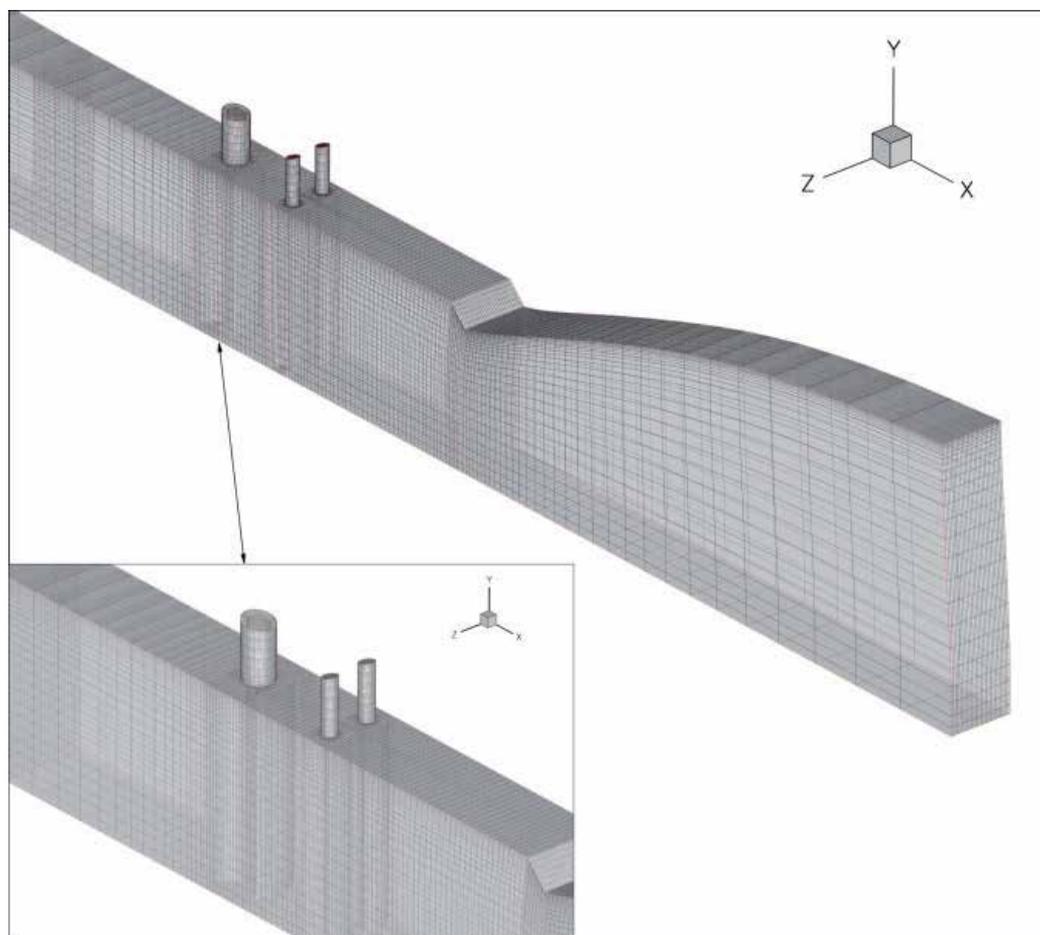


Fig. 1. Computational grid used in the 3-D GASP simulations of COIL hardware, shown in relation to a rendering of the wetted surfaces in the COIL laser hardware. Note that the resolution show here is 1/64th that used in the execution of the simulation.

and injector subsonic inflow boundary conditions fix the total pressure, total temperature, and the species fractions at constant values, while the derivative of the static pressure is set to 0. The nozzle outflow boundary condition sets the second derivative of the dependent variables to 0 as is appropriate for supersonic flows. For execution of this model, References providing additional details beyond the scope of this article may be found in Madden et al.^{16,17}

3. Results

The 3-D GASP model for the COIL flowfield was executed in time accurate mode utilizing 1st order accurate time integration and 3rd order spatial accuracy. The computation was advanced with a physical time step of 4.0×10^{-9} sec over 159,500 time steps to a physical time of 0.00064 sec. Analysis of this simulation indicated the need to increase the resolution of the grid in the streamwise direction within the nozzle expansion, and a second simulation was performed that quadrupled the number of grid cells in this region bringing the total number of grid cells to 14.7 million. This simulation was continued for an additional 124,000 time steps with a time step of 2.0×10^{-9} sec., and added 0.000247 sec of physical time to the simulation.

The time accurate execution of the GASP COIL model generated a prediction of flow unsteadiness that did not decay over the time intervals that the computations were advanced. The unsteadiness was found to extend from the jet/primary interaction region at the point of the He/I₂ transverse injection and continue downstream undiminished. Fig. 2 illustrates the resulting impact of the flow unsteadiness upon the structure of the flow within the He/I₂ jet. Here a volumetric rendering of the I₂ mole fraction is plotted within 3-D space. The rendering demonstrates the presence of complex, periodic structures associated with the unsteady vortex generation, and indicates that there is a strong interaction between the fluid from the large and small orifices. The break-up of the jets occurs as the fluid from the small orifices begins to interact with that from the large, illustrated by the change in the vortical structure from smooth, periodic fluctuations to

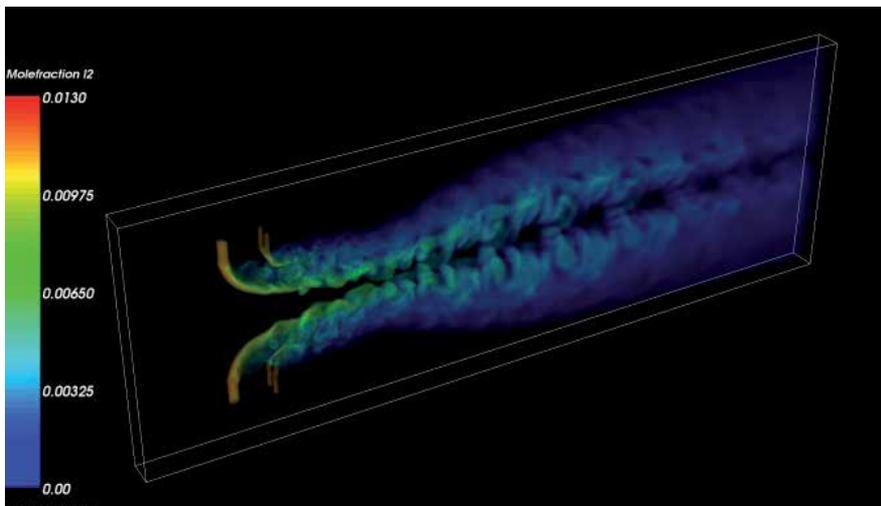


Fig. 2. Fixed time snap shot of the unsteady fluid dynamic structures manifested in the He/I₂ jet from the 10 species, 22 reactions 3-D GASP COIL simulation. Note that the view is reflected about the symmetry plane at the nozzle centerline to provide an enhanced perspective.

the more complex structure associated with the break-up. Since molecular diffusion and chemical reactions are strongly correlated with spatial gradients of reactant concentration, the presence of these unsteady flow structures has considerable impact upon the generation and variation of $I(^2P_{3/2})$ and $I(^2P_{1/2})$. Additionally, the gain is directly proportional to the number densities of $I(^2P_{3/2})$ and $I(^2P_{1/2})$ and the gain will correlate with the flow structure as illustrated in Fig. 3.

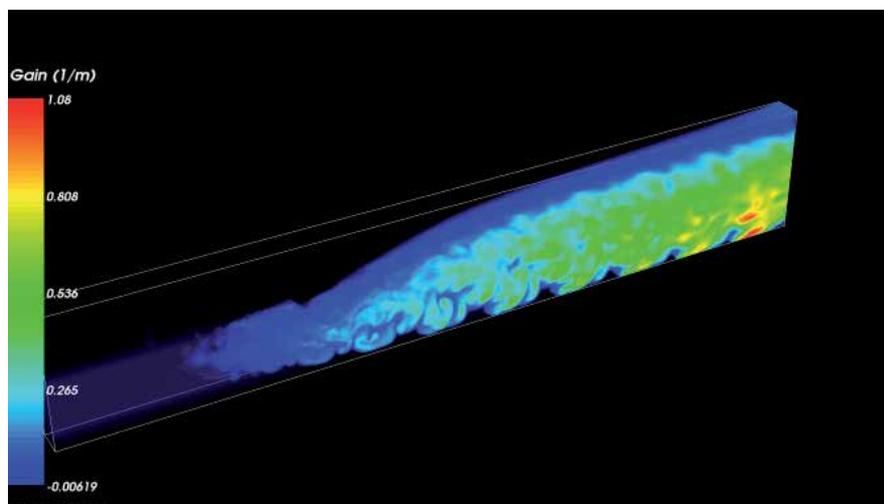


Fig. 3. Fixed time snap shot of the three-dimensional gain variation within the simulation volume.

The underlying sources of the flow unsteadiness are the 'wake vortices' in the flow behind the He/I₂ jets. These vortices have their origin in the primary flow boundary layer fluid that wraps around the orifice flow and projects away from the wall with rotation, subsequently interacting with the jet fluid. Fig. 4 illustrates this with plots of streamlines originating in the boundary layer fluid upstream of the jet shown in conjunction with vortex cores. The cores were extracted using the vortex core identification functions of Sujudi and Haines¹⁸ implemented in the Tecplot visualization software from Amtec Engineering. The red vortex core traces can be found at the center of the swirls in the streamlines, identifying the location of the vortices. Two wake vortices are seen projecting away from the wall immediately behind the large orifice, and a third projects downward from the wall in the center of the domain between the large and small injector orifices. Similarly, additional vortices are seen projecting away from the wall behind the small injector orifices. Temporal variation of the flow fluctuations and the Fourier analysis of these fluctuations for the two simulations is shown in Figs. 5, 6. Fourier analysis of the fluctuations from the 10 species, 22 reaction simulation yielded frequencies of 56, 94, 131, and 150 kHz, whereas the 13 species, 52 reaction simulation yielded frequencies of 12.1, 24.9, 42.3, 60.4, 96.6, 133.9, 151, and 292 kHz. The peak fluctuation amplitudes for the lateral velocities are 175 m/sec in both simulations. The presence of the unsteady fluctuations within the jet structure is consistent with the experimental literature on jet-in-crossflow interactions. In particular, the capture of the wake vortices as well as their location is entirely consistent with the experimental investigations, exemplified in the work of Frik and Rhosko.

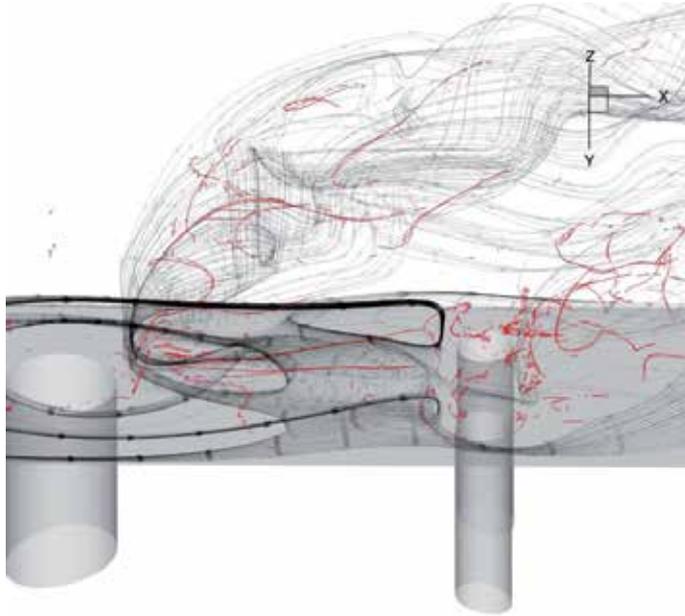


Fig. 4. 3-D streamtraces (in black) initiated upstream of the large injector orifice 2% of the channel height from the wall in conjunction with vortex cores (in red) extracted from the 10 species, 22 reaction 3-D GASP COIL simulation.

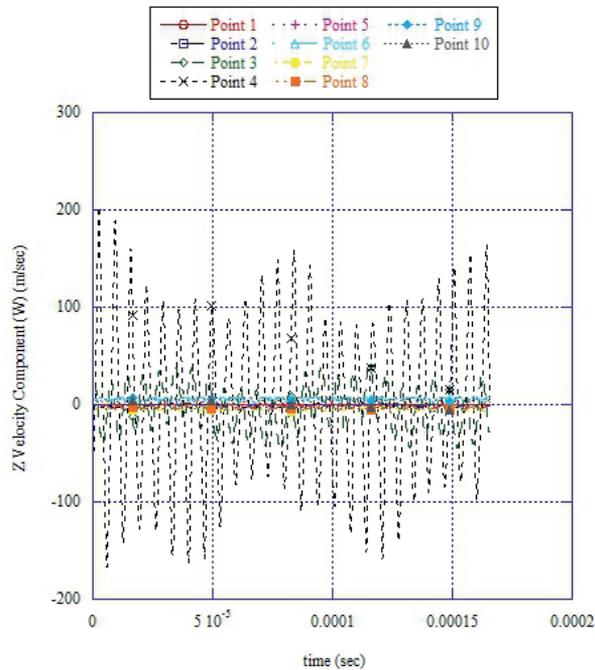


Fig. 5. Time variance of the lateral (Z) velocity component (W) at various monitoring points in the jet nearfield from the COIL simulation.

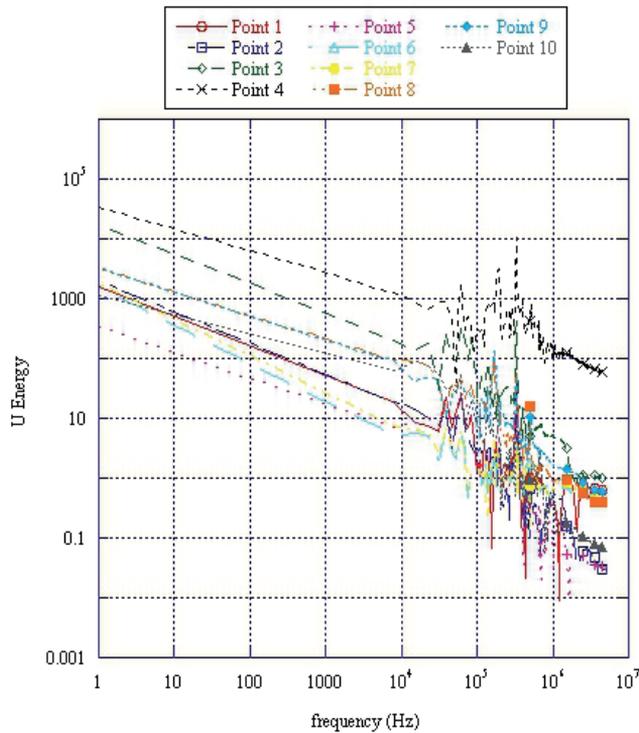


Fig. 6. Fourier analysis of temporal variation of the streamwise or Z direction velocity (W) component in Fig. 5.

While the predicted presence of the unsteady fluctuations within the jet structure is entirely consistent with the experimental literature on jet-in-crossflow interactions, the question of validation relative to COIL flow conditions is still relevant. This issue was addressed by conducting a non-reacting flow experiment in which a He/I₂ mixture was injected into a He/O₂ primary flow with the nozzle geometry and flow conditions matching the hardware simulated here. A single large and two small orifices were used on the upper surface of the nozzle, thus differing from the hardware simulated here which used an array of orifices on both the upper and lower surfaces. Planar laser induced fluorescence (PLIF) was used to image the jet fluid, providing the ability to visualize the jet structure. Fig. 7 shows the PLIF image resulting from a 600 μm thick sheet of 565 nm laser light passing through the centerline of the jet from the large orifice and causing fluorescence within the I₂ molecules, illuminating the jet structure; Fig. 8 shows a 2-D cut from the simulation at the large injector centerline as a point of comparison. The combination of the primary flow velocity, the spacing between vortices in the jet fluid, and the size of the vortices gives information to compare directly with the simulations. The size of the first 3 vortices in the PLIF image is 0.8, 1.9, and 2.3 mm which give an indication of the spatial growth rate of the vortices as they are shed. The size of the first 3 predicted vortices in the simulation are 0.8, 1.1, and 1.3 mm, a somewhat smaller growth rate than that in the PLIF image. Observing the difference in the jet trajectory between the experiment and the simulation, attributable to the presence of an opposing set of identical jets across the vertical centerline in the simulation, it is possible that the smaller growth rate in the simulation is due to interference from the opposite jets. However, the frequency determined

from the spacing between the vortices in the PLIF image and the primary flow velocity is 132 kHz, differing by less than 1% from the 131 and 133.9 kHz frequencies extracted from the simulation associated with the vortex shedding.

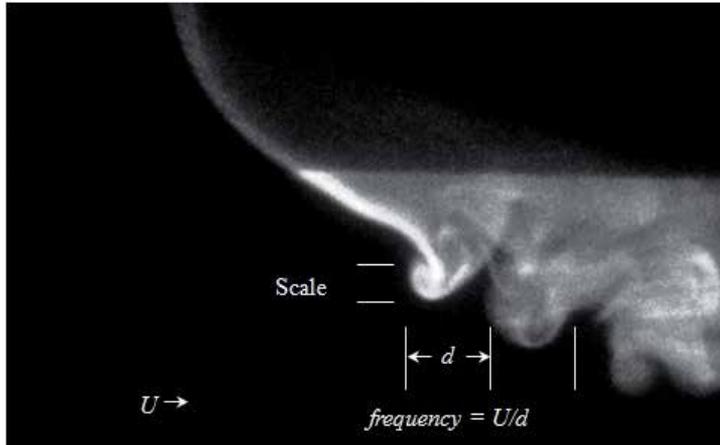


Fig. 7. PLIF experiment image of a jet-in-crossflow at flow conditions and with a configuration similar to that in the simulations performed here. This image was used to extract frequency and scale for validation of the models unsteady prediction characteristics. Note that the nozzle throat blocks the PLIF laser light, darkening the upper 1/3 of the image.



Fig. 8. Planar cut at the centerline of the large injector orifice from a single time step in the COIL simulation.

The end question to be answered here is ‘what effect does the unsteady flow structure have upon the predicted lineshape?’ This question is answered by accumulating a dataset containing the spatial and temporal variation of the quantities necessary to calculate the gain. This dataset is first spatially averaged in the Z axis direction and then temporally averaged, consistent with the behavior of a tunable diode small signal gain probe or lasing action over a period of time. Mathematically, this is expressed as:

$$\alpha_{ave}(v, y) = \frac{1}{t_{tot} z_{tot}} \iint \alpha(v, y, z, t) dz dt$$

where t_{tot} is the total time interval of the integration, z_{tot} is the total distance of integration in the Z axis direction, and $\alpha(v,y,z,t)$ is determined using the equations listed above. For the simulations performed here, the exit plane of the nozzle as shown in Fig. 1 was exported every $1.0e-07$ sec of physical time, and the Z integration was performed along lines of constant Y. Fig. 10 shows a composite gain lineshapes versus normalized frequency with the frequency axis normalized to the line center of the $I^2P_{1/2} \rightarrow ^2P_{3/2}$ transition at $1.31527\mu m$. Here t_{tot} is $6.4 \cdot 10^{-04}$ sec and z_{tot} is the full dimension in the z direction. This plot shows that the lineshape including the influence of the gas velocity is broadened somewhat and the line center gain is lower, indicating lower amplification for the regions of the flowfield where the broadening effects of the velocity field are significant.

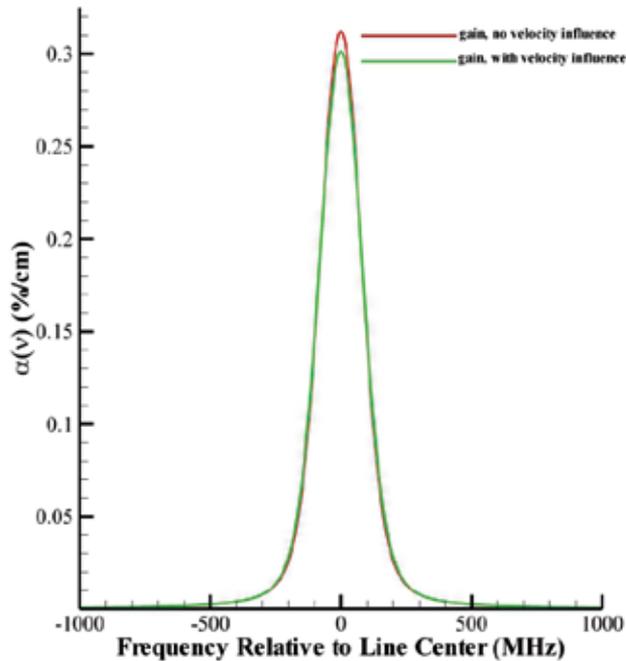


Fig. 9. Time, space averaged gain lineshape from the COIL simulation.

The transition lineshapes predicted with the above methodology can next be fit using the gain equations from above to characterize temperature extraction when strong flow fluctuations impact the lineshape. The spatially, temporally averaged lineshape is made using the assumption that all of the Doppler broadening be due to random, thermal motion, allowing the lineshape to be treated as synthetic representation of experimentally measured lineshapes. The absence and the presence of the bulk gas velocity influence in the lineshape provides an independent parameter for quantifying the effect of bulk gas velocity in terms of temperature. To fit the lineshape equation to the spatially, temporally averaged lineshapes reduced from the unsteady simulations, the MINPACK¹⁹ implementation of the Levenberg-Marquardt^{20,21} nonlinear fit algorithm is used²² to determine the best fit of to the data as a function of a given set of fit parameters. Observing the functional form of the lineshape equation, it is seen that it consists of a shape function $\phi(v)$ given by the Voigt function, and an amplitude given by:

$$\frac{7}{12} \left(\frac{A\lambda_0^2}{8\pi} \right) \left(N_{I(\nu^2 P_{1/2})} - \frac{1}{2} N_{I(\nu^2 P_{3/2})} \right)$$

Given that the Voigt represents the probability versus frequency, the infinite integral is 1.

$$\int_{-\infty}^{+\infty} \phi(\nu) d\nu = 1$$

Integrating allows for determination of and separation of the Voigt from the lineshape:

$$\begin{aligned} \int_{-\infty}^{+\infty} \alpha(\nu) d\nu &= \int_{-\infty}^{+\infty} \frac{7}{12} \left(\frac{A\lambda_0^2}{8\pi} \right) \phi(\nu) \left(N_{I(\nu^2 P_{1/2})} - \frac{1}{2} N_{I(\nu^2 P_{3/2})} \right) d\nu \\ &= \frac{7}{12} \left(\frac{A\lambda_0^2}{8\pi} \right) \left(N_{I(\nu^2 P_{1/2})} - \frac{1}{2} N_{I(\nu^2 P_{3/2})} \right) \int_{-\infty}^{+\infty} \phi(\nu) d\nu \end{aligned}$$

and

$$\alpha(\nu) = \phi(\nu) \int_{-\infty}^{+\infty} \alpha(\nu) d\nu$$

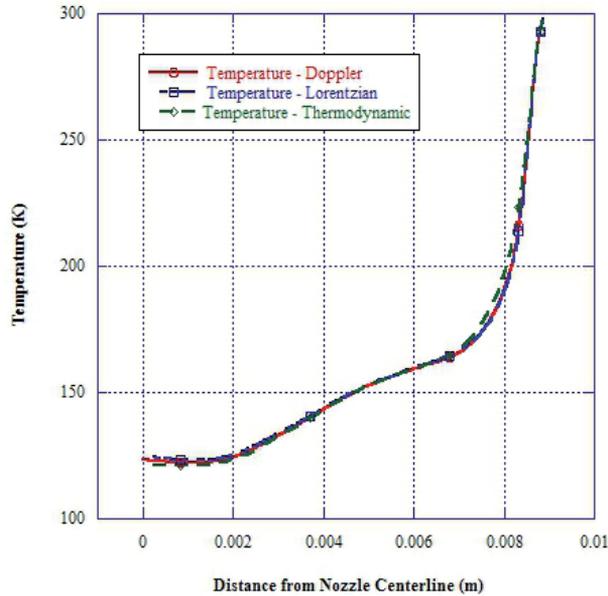


Fig. 10. Comparison of temperature values from the 13 species, 52 reaction simulation; Doppler and Lorentzian values were determined based upon the lineshape in Fig. 9 without the velocity dependence and the thermodynamic values averaged from the simulation data.

Thus the individual values of the gain $\alpha(\nu)$ at each ν can be divided by the infinite integral of $\alpha(\nu)$ to recover the Voigt. However, in practice it was found that fitting this normalized

lineshape introduced errors because the integral bounds are limited to the frequency domain of ± 1.75 GHz about the line center used to construct the lineshape numerically and are not infinite as required by the Voigt function. The effect of this error on the fit is such that evaluation of the infinite integral of $\phi(\nu)$ via is 0.99, not 1.0. This translated into a 5 K error in the temperature determination, and was subsequently determined to be avoidable if the infinite integral over $\alpha(\nu)$ were treated as a separate fitting parameter. In practical terms, the fitting function becomes:

$$\alpha(\nu, \Delta\nu_D, \Delta\nu_L) = \beta\phi(\nu, \Delta\nu_D, \Delta\nu_L)$$

where β , $\Delta\nu_D$, and $\Delta\nu_L$ are the fitting parameters used in the nonlinear fit. With the values for the Doppler and Lorentzian widths $\Delta\nu_D$ and $\Delta\nu_L$ determined, separate values for the temperature can be found using the gain equations from above in conjunction with an independent determination of the pressure and the mole fractions. Here, the simulation data were spatially and temporally averaged using the same procedure as for the gain determine average pressures and species mole fractions. When performing this analysis with experimentally measured lineshapes, wall pressures and mean mole fractions based upon the measured molar flow rates may be the only values available and may lead to errors in the extracted temperatures.

The degree of agreement between the Doppler and Lorentzian determined temperatures and the average thermodynamic temperature from the simulation provides a means to assess the lineshape fit and the interpretation of the fit. To perform this comparison, the thermodynamic temperatures were averaged from the simulation data using the same integration procedure as used for the gain. As a test of the procedure, a fit was performed to a lineshape that was calculated without the functional dependence upon the bulk velocity. The comparison between the Doppler, Lorentzian, and thermodynamic temperatures in this test is shown in Fig. 10 is very close, to within 1%. Next, the temperatures are determined for the lineshape in which the bulk velocity was included in the functional dependence of the lineshape equation. As shown in Fig. 11, the Doppler temperature diverges from the Lorentzian and thermodynamic temperatures in the flow regions closer to the nozzle centerline, with the maximum difference being 14 K. This trend is consistent with the observation that the unsteady eddies increase in size toward the center of the nozzle as illustrated in Fig. 2. Thus the Doppler temperature is registering the clear influence of the bulk gas velocity, whereas the Lorentzian temperature remains consistent with the thermodynamic temperature. Given the results illustrated here, a suggested method of fitting experiment lineshapes Where poor signal-to-noise can interfere with the nonlinear fitting would be to fix the Lorentzian width based upon an independent temperature measurement such as a rotational temperature, and then determine the Doppler width. Then the degree to which the Doppler temperature differs from the rotational or other temperature measurement would be an indicator of the magnitude of flow rotation. It should be noted that fitting procedures in which the Doppler and Lorentzian component temperatures are forced to be equal, as is commonly done, could very well be masking information regarding the magnitude of eddy strength and the degree of unsteadiness in the experiments. Another test would be to re-visit existing lineshape temperature determinations for experiments in which flow unsteadiness could potentially be present, and perform the nonlinear fit without fixing either the Doppler and Lorentzian component. The difference in temperatures could then be correlated against fluid dynamic parameters such as Reynolds number that are relevant to eddy generation to identify trends.

The temperature determination from the averaged lineshapes illustrates the sensitivity of the temperature determination process to the magnitude of the flow velocity components. Observing the decrease in grid density within the nozzle expansion in Fig. 1 and the dissipation of the vortices within Fig. 2, it is reasonable to question whether the increase in grid cell size with streamwise distance in the expansion is filtering the smaller eddies as well as decreasing the strength of the larger eddies through numerical dissipation. Fig. 12 illustrates this point through the plotting of a constant vorticity isosurface with a value of $2.0 \cdot 10^5 \text{ s}^{-1}$. The figure shows that this particular vorticity magnitude, while persistent within the freestream before the nozzle throat, disappears rapidly as the flow passes through the throat. This decrease is consistent not just with the decrease in grid density but also with the physical effect of the flow expansion and the commensurate stretching of the vortices. To separate the effect of flow dilation upon vorticity from that of grid dissipation, it is necessary to decrease grid dissipation by increasing the grid resolution.

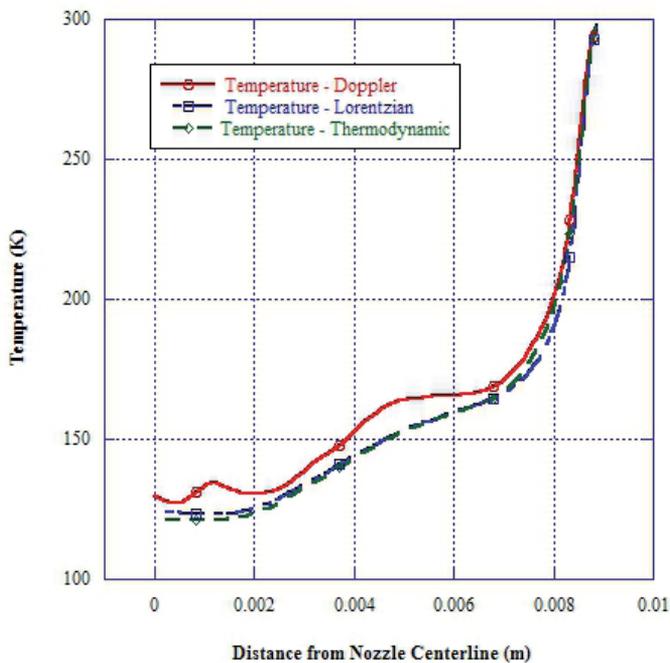


Fig. 11. Comparison of temperature values from the simulation; Doppler and Lorentzian values were determined based upon the lineshape in Fig. 9 with the velocity dependence and the thermodynamic values averaged from the simulation data.

The effect of grid dissipation upon the transport of vorticity in the nozzle expansion and the subsequent effect upon the lineshape is examined by quadrupling the number of grid cells in the streamwise direction within the nozzle expansion while keeping the number of cells in the transverse directions constant. This increases the number of cells in the streamwise direction from 128 to 512. This new grid is used to continue the simulation and accumulate sufficient temporal information regarding the influence of the unsteadiness upon the lineshape. Data necessary for the lineshape analysis were collected within the same plane transverse to the flow as in the first simulation.

The first point of examination of the effect of increased grid resolution is the vorticity. The isosurface of constant total vorticity at a value of $2.0 \cdot 10^5 \text{ s}^{-1}$ is generated for the new grid and plotted in Fig. 13. The first and most obvious feature distinguishing the two plots is that the vorticity loss within the downstream end of the nozzle expansion found in the simulation using the original grid does not appear in the simulation using the higher cell density grid, confirming that the grid is impacting the transport of the flow fluctuations in this region.

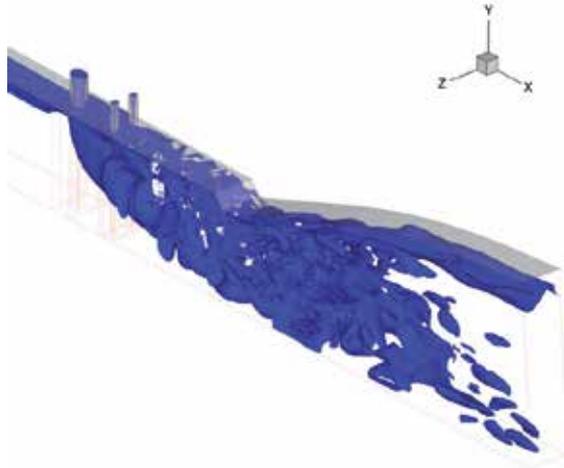


Fig. 12. An isosurface of constant total vorticity with a value of $2.0 \cdot 10^5 \text{ (1/s)}$. This simulation used 128 points in the streamwise direction in the grid within the nozzle expansion.

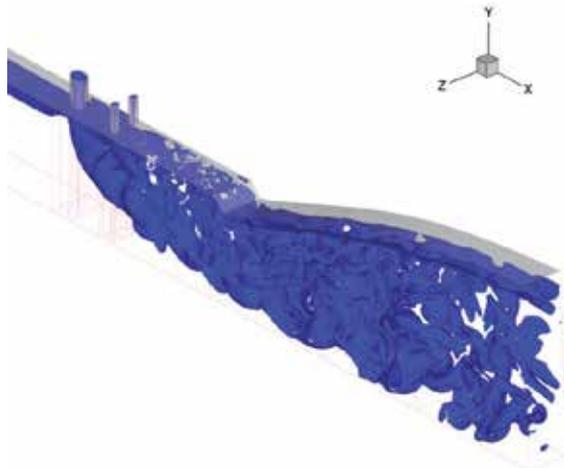


Fig. 13. An isosurface of constant total vorticity with a value of $2.0 \cdot 10^5 \text{ (1/s)}$. This simulation used 512 points in the streamwise direction in the grid within the nozzle expansion. Note the continued presence of the vorticity through the nozzle exit.

While vorticity is a measure of flow rotation and is associated with the velocity components in the lateral direction that influence the lineshape, it is also correlated with mixing and the grid resolution effect shown here would be expected to extend beyond the broadening to the

general prediction of the COIL flowfield. An examination of the Doppler and Lorentzian temperature components based upon fitting of the lineshapes for the original and higher resolution grid simulations provides the final point of comparison. Fig. 14 shows the Doppler, Lorentzian, and thermodynamic temperatures for the new grid simulation compared with the Doppler temperature from the original grid simulation. Comparing the Doppler temperatures, the values from the higher density grid simulation are somewhat higher over a greater section of the nozzle than the original grid, consistent with the analysis performed above. Additionally, the maximum difference between the Doppler and Lorentzian temperatures for the new grid simulation is 25 K, an increase from 14 K for the original grid simulation. This confirms the hypothesis that the grid resolution impacts the lineshape analysis and strengthens the argument that flow unsteadiness is a factor in temperature extraction from lineshapes in these flows.

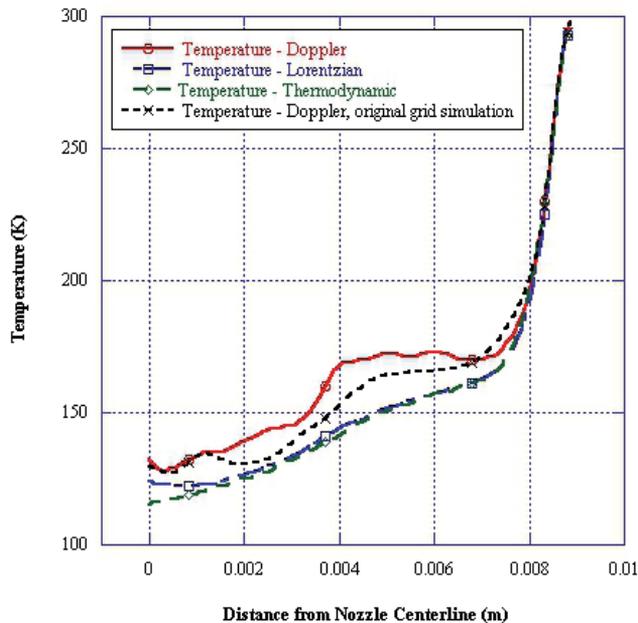


Fig. 14. Comparison of temperature values from both the original and high density grid simulations. Note that the Doppler temperature is somewhat higher with the high density grid simulation.

The result of a 25 K effect of the unsteady flow structures on the temperature extraction from the Doppler and Lorentzian components naturally leads to questions about the velocity distribution predicted here and the Gaussian distribution assumed by Stuve and Elvey and later by Nikolaev. To analyze this question, the time dependent W velocity component data from the 2D planes used above to generate lineshapes was separated into 1 m/sec bins ranging from -200 to 200 m/sec, and temporal and spatial frequency of occurrence was counted. The probability of occurrence for a particular 1 m/sec bin was calculated as the frequency of occurrence divided by the total frequency of occurrence for all velocity bins within the freestream portion of the 2D plane across the entire time of integration. The plot of this data is shown in Fig. 15. The distribution shows a symmetric and finite probability of occurrence of the W velocity from -100 to 100 m/sec. When this data

is fit with a Gaussian distribution function, shown with the solid blue line in Fig. 15, the Gaussian distribution shows a somewhat lower probability in the region from -40 to -100 m/sec and 40 to 100 m/sec. The reason for this difference can be explained in terms of the turbulent flow assumption underlying the use of the Gaussian velocity distribution. The assumption of turbulence implies a character to the flow fluctuations that is consistent with a Gaussian distribution, i.e. random and having an exponential range of energies and scales. Fig. 15 indicates that the predicted velocity probability distribution does not have this type of character, implying that the random, stochastic character of isotropic turbulence that allows an exponential range of energies and scales to develop is not present in this simulation. Simply said, the flow is not turbulent.

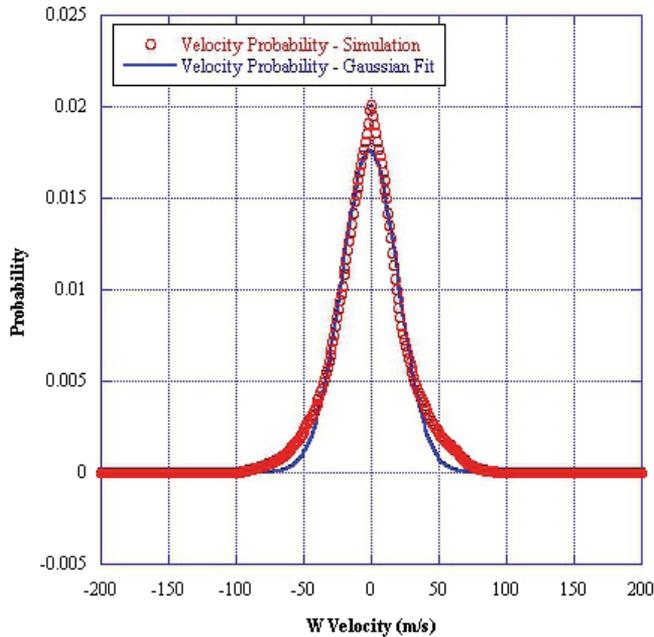


Fig. 15. Velocity probability as determined from statistics collected in the simulation compared with a Gaussian fit to this data.

The question of turbulence can be examined in a preliminary sense by examining the temporal variation of the lateral velocity component at a point within the 2D plane. The point chosen is within a region in this plane marked by high fluctuation magnitudes, is associated with the largest differences between the Doppler and Lorentzian temperatures, and is more likely to be turbulent. Fig. 16 shows the variation of the W velocity component with time at this point. As can be seen, the velocity variation is regular and sinusoidal in character, with some random fluctuation in the amplitude within the first 7×10^{-5} sec and near 2.4×10^{-4} sec. Comparing with Fig. 5 which plots the W velocity versus time in the region surrounding the injectors where the flow unsteadiness originates, it is seen that the sinusoidal character of the fluctuations is preserved from the injection region, through the nozzle throat, and downstream to the end of the isentropic expansion region where the data for Fig. 16 are extracted. If the flow were to transition to fully isotropic turbulence, the sinusoidal character of the fluctuations would be lost between Figs. 5 and 16 as the larger

scale vortices interact to generate smaller vortices, leading to the cascade of temporal and spatial scales that characterizes turbulence.

Returning to the issue of lineshape fitting and the use of a Gaussian function to describe the bulk velocity probability distribution, the results of this work indicate that for the low Reynolds number flows simulated here that are typical of COIL operation, a Gaussian velocity distribution function does not describe the bulk flow fluctuations because the flow is not turbulent. However, this statement has broader implications as statistical turbulence descriptions such as the k, ε and k, ω models are now frequently employed in COIL simulation. Since the COIL chemistry, particularly I_2 dissociation, is sensitive to the local concentrations of reactants and the concentrations are directly determined by mixing and chemistry, the predictions given by COIL models that use turbulence models to predict the mixing should be viewed cautiously. This statement holds particularly for efforts to develop and prove new COIL chemistry mechanisms, where accurate prediction of the reactant mixing characteristics provides a basis for accurate examination of the predictions from the chemistry model.

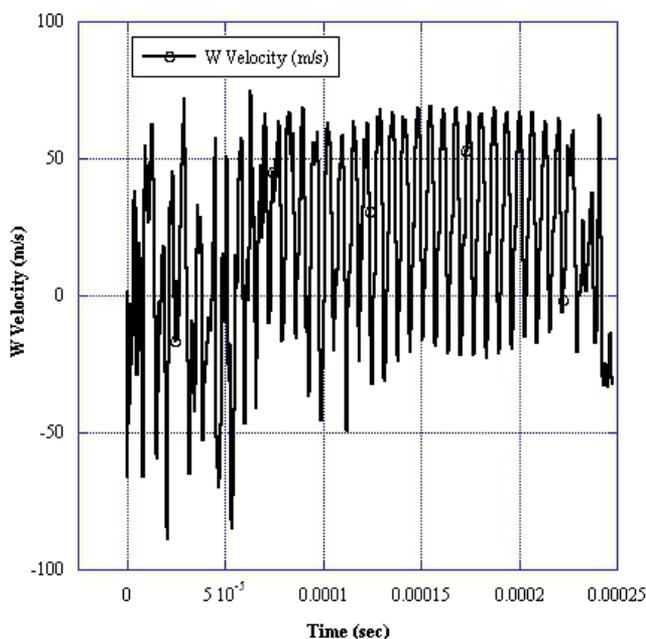


Fig. 16. Example of W velocity variation with time within the nozzle exit plane for the higher density grid over the period of the simulation.

4. Summary and conclusions

In conclusion, the results provided here predict that the bulk gas flow velocity components associated with vortex rotation will broaden the lineshape for quantum transitions of species within the gas flow. The studies performed here show that if the bulk gas velocity influence is ignored when fitting the lineshape and the vortices are present, the effect will be a divergence between the reduced Doppler and Lorentzian values of the temperature. The results presented here indicate that this difference will be as high as 25 K, with some

variation based upon location in the flow. Based upon the observed decrease in vorticity within the nozzle expansion in conjunction with decreased grid density, an additional simulation was performed that increased the grid in the streamwise direction in this region by a factor of 4. Comparing the results from the simulations from different grids, the higher density grid dramatically decreased dissipation of vorticity, captured higher magnitude velocity amplitudes, and showed a 11 K increase in the difference between the Doppler and Lorentzian temperatures over the original grid simulation, illustrating the impact of numerical considerations upon the predictions of the model. Based upon the observed trends in temperature prediction, it is suggested that the difference between the temperatures extracted from the fit values for the Doppler and Lorentzian widths could be correlated against changes in fluid dynamic parameters such as Reynolds number to determine trends.

The velocity probability distribution for the flow at the exit of the nozzle expansion was examined, in the location where the lineshape examination was performed. Analysis of this velocity distribution showed it to be non-Gaussian, with higher probabilities for higher magnitude velocity fluctuations. Further analysis illustrates that the fluctuations are sinusoidal in nature as is characteristic of laminar unsteady flows, as opposed to the random, stochastic fluctuations associated with turbulent flows. This indicates that a Gaussian probability distribution function, as is commonly used in the astronomical community for Voigt lineshape analysis of spectral lineshapes of atoms in stellar atmospheres that are assumed to have turbulent bulk flow velocities, is not strictly appropriate for these flows.

5. Acknowledgements

The author would like to thank Drs. Kevin Hewett and John McCord for fruitful discussions regarding the methods used in experimental lineshape measurement and temperature extraction for COIL devices. Mr. Jeremy Stanford generated the computational grids used in the simulations discussed here. Computational resources were provided by the DoD High Performance Computing Modernization Challenge Allocation at the Aeronautical Systems Center, Army Research Laboratory, Engineering Research and Development Center, and Naval Oceanographic Major Shared Resource Centers.

6. References

- [1] Struve, O. and Morgan, W. W., "On the Intensities of Stellar Absorption Lines," *Proceedings of the National Academy of Sciences of the United States of America*, 18, No. 9, pp. 590-594, 1932.
- [2] Elvey, C. T., "The Intensities of Some Multiplets of Fe II and Ti II in Stellar Spectra," *Astrophysical Journal*, 79, pp. 263-269, 1934.
- [3] Struve, O. and Elvey, C. T., "The Intensities of Stellar Absorption Lines," *Astrophysical Journal*, 79, pp. 409-440, 1934.
- [4] Allen, M.G., Carleton, K.L., Davis, S.J., Kessler, W. J., and McManus, K. R., "Diode laser-based measurements of water vapor and ground state oxygen in chemical oxygen iodine lasers," AIAA-1994-2433, 25th AIAA *Plasmadynamics and Lasers Conference*, Colorado Springs, CO, 1994.

- [5] Davis, S. J., Allen, M. G., Kessler, W. J., McManus, K. R., Miller, M. F., and Muihall, P. A., "Diode laser-based sensors for chemical oxygen iodine lasers," Paper 2702-18, *Proceedings of SPIE Conference on Gas and Chemical Lasers*, 2702, pp. 195-201, San Jose, CA, 1996.
- [6] Davis, S. J., Kessler, W. J., Bachmann, M., and Mulhall, P.A., "Collisional broadening coefficients for oxygen and water absorption lines used in COIL diagnostics," Paper 3268-80, *Proceedings of SPIE Conference on Gas and Chemical Lasers*, 3268, pp. 218-226, San Jose, CA, 1998.
- [7] Davis, S. J., Kessler, W. J., and Keating, P. B., "Progress in the development of sensors for COIL devices," *Proceedings of SPIE Conference on Gas, Chemical, and Electrical Lasers and Intense Beam Control and Applications*, 3931, pp. 156-161, 2000.
- [8] Nikolaev, V.D., Zagidullin, M. V., Svistun, M. I., Anderson, B. T., Tate, R. F., and Hager, G. D., "Results of Small-Signal Gain Measurements on a Supersonic Chemical Oxygen Iodine Laser with an Advanced Nozzle Bank." *IEEE Journal of Quantum Electronics*, 38, no. 5, May 2002.
- [9] Madden, T. J. and Miller, J. H., "Simulation of Flow Unsteadiness in Chemical Laser Flowfields," AIAA-2004-0805. *42nd AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, 5-8 Jan, 2004.
- [10] Fric, T. F., and Roshko, A., "Vortical structure in the wake of a transverse jet," *J. Fluid Mech*, 279, pp. 1-47, 1994.
- [11] Perram, G. P., *Int. J. Chem. Kinet.* 27, 817-28 (1995).
- [12] R. F. Heidner III, C. E. Gardner, G. I. Segal, and T. M. El-Sayed, "Chain-Reaction Mechanism for I₂ Dissociation In the O₂(1-Δ)-I Atom Laser," *Journal of Physical Chemistry*, 87, 2348 (1983).
- [13] Azyazov, V. N.; Pichugin, S. Yu.; Heaven, M. C., "On the dissociation of I₂ by O₂(a ¹Δ): Pathways involving the excited species I₂(A ³Π_{2u}, A ³Π_{1u}), I₂(X ¹Σ_g⁺), and O₂(a ¹Δ, v)," *Journal of Chemical Physics*, Volume 130, Issue 10, pp. 104306-104306-9 (2009).
- [14] Waichman, K., Barmashenko, B. D. and Rosenwaks, S. "Comparing modeling and measurements of the output power in chemical oxygen-iodine lasers: A stringent test of I₂ dissociation mechanisms," *Journal of Chemical Physics*, 133, (2010).
- [15] Madden, T. J. "An Analysis of Mechanisms Underlying Flow Unsteadiness in Chemical Oxygen-Iodine Laser Mixing Systems," AIAA-2005-5390, *36th AIAA Plasmadynamics and Lasers Conference*, Toronto, Ontario, Canada, June 6-9, 2005.
- [16] Madden, T. J. and Solomon, W. C., AIAA 97-2387, *28th Plasmadynamics and Lasers Conference*, Atlanta, GA, June 23-25, 1997.
- [17] Madden, T. J., *SPIE Proceedings of XIVth International Symposium On Gas Flow & Chemical Lasers and High Power Laser Conference*, Wroclaw, Poland, 25-30 August, 2002.
- [18] Sujudi, D. and Haines, R., "Identification of Swirling Flow in 3-D Vector Fields," AIAA Paper 95-1715, *12th AIAA Computational Fluid Dynamics Conference and Open Forum*, San Diego CA, 19-22 June 1995.
- [19] www.netlib.org/minpack.
- [20] Levenberg, K. "A Method for the Solution of Certain Problems in Least Squares," *Quart. Appl. Math.* 2, pp. 164-168, 1944.
- [21] Marquardt, D. "An Algorithm for Least-Squares Estimation of Nonlinear Parameters" *SIAM J. Appl. Math.* 11, pp. 431-441, 1963.
- [22] Keating, P. B., private communication, Sept. 1998.

Fluid-Structure Interaction

Stoia-Djeska Marius and Safta Carmen-Anca
*POLITEHNICA University of Bucharest, Bucharest
Romania*

1. Introduction

The fluid-structure interaction (FSI) phenomenon is the result of the interactions of multiple continuum fields. The fluid (gas or/and liquid) forces act on a neighboring elastic solid, which is deformed and thus influences the flow of the adhering fluid. Due to the deformation of the solid, both the fluid velocity and the fluid domain change. Usually, this interaction takes place in the presence of other mechanical fields and as important examples we mention here the external body forces and acoustic fields. The mathematical modeling of the FSI phenomenon requires mathematical models for each field implied in and also the coupling mechanisms between the fields. For instance, one can imagine a mathematical model of a FSI phenomenon where the fluid flow field is modeled with the unsteady Navier-Stokes equations, the solid elastic field with the Navier equations while for the acoustic field the inhomogeneous wave equation based on Lighthill's analogy can be used and the external body forces are due to gravity. The boundary conditions for the fluid, solid and acoustic models represent here the coupling mechanism between the fields. However, even everything looks simple and clear the accurate and efficient solution of FSI problems is still a highly complicated task and an open area of research. Furthermore, using complicated models as those enumerated above the only way to obtain a solution to a specific FSI problem is the numerical one. The Computational Fluid Dynamics (CFD), Computational Structural Dynamics (CSD), Finite Element Method (FEM) provide us with specific **mathematical models** and **numerical techniques** that can be coupled to build a numerical solver for a FSI problem. However, even for simple and without engineering relevance problems the computational effort required to solve a specific problem becomes huge if the **mathematical modeling of the physical phenomenon** is not carefully done. As in many other multiple field interaction problems the **level of accuracy** of the mathematical model is crucial for both the efficiency and quality of the results. Further, we notice also that the development of in-house codes for such purposes is a difficult task.

Today we know from the experience earned from a century of practice that different levels of approximation can be used successfully to build up a simplified mathematical model for a FSI problem. This can be done by simplifying the flow equations and by choosing engineering models for the elastic structure. With appropriate coupling mechanisms one can thus obtain mathematical models which are well-matched for practical purposes, starting from the design and up to advanced (flow and/or structural dynamics) control purposes. The use of different simplifying hypothesis (about the dimensionality of the flow and structural models, for instance) requires from the engineer a deep understanding of the multifield interaction phenomenon so that by neglecting different components of the

governing equations the simplified mathematical model still captures the essential of the physical phenomenon under concern. Based on this approach the engineers have now at disposal simple mathematical models and even formulae, which offer an insight into a FSI phenomenon or another and also a useful tool to solve a specific FSI problem. This is evident if one takes a look in the field of aeroelasticity, see (Bishplinghoff et al, 1955; Bishplinghoff & Ashley, 1956; Dowell & Ilgamov, 1988).

Historically, the aeroelasticity was the first branch of the FSI which occurred, (Fung, 1956; Dowell, 1975). Aeroelastic and/or hydroelastic vibrations are sustained by air/water forces induced by the moving structure itself. Due to this interaction the elastic structure may suffer large elastic deformations or even divergent oscillations causing failure. This evidence and the concern for very light-weight and thus flexible aircraft and propulsion systems structures have led to the occurrence of the aeroelasticity branch of science from the time when the aeronautical industry was at the beginning. The focus in aeroelasticity was and remains on the safety of the aircraft or civil structures only. However, in the last decades and in parallel with the progress in the computational fluid and structural dynamics and in the computer technology the aeroelasticity shifted forward in the direction of FSI. Currently we are interested in the behavior of both fluid and structural systems and perhaps the best examples come from the field of hydroelasticity and bioengineering applications.

Nowadays, the FSI phenomenon covers a large area of engineering applications and the dedicated literature means thousands of papers. In what follows we present a rather coarse than comprehensive survey of the research work done in different engineering domains to investigate experimentally and numerically different FSI type phenomena of actual interest. The reader is encouraged to search forward in the list of publications for a deeper investigation of the newest results in its domain of work.

A recent application of fluid structure interaction is the **energy harvested** by converting the mechanical strain of an elastic material, under the pressure of a fluid, into electric potential using piezoelectric materials, (Doaré, 2011). For example, a flexible plate fully coupled to a simple dissipative electrical circuit through piezoelectric layers could be an attractive candidate for flow energy harvesting if an axial flow through the plate will produce self-sustained periodic oscillations of the solid body. A global analysis of fluttering modes of a finite-length plate confirmed that waves or modes destabilized by piezoelectric coupling maximize the energy conversion efficiency. Another application is the hydroelastic analysis of very large floating structures used in ocean space (as hydromechanical equipment in the wave's energy conversion process), (Karmakar, 2009). These structures consist of many **articulated elastic plates**, called modules. The articulation of the elastic plates is done by the connectors which depend on the stiffness constants known as the vertical linear spring stiffness and flexural rotational spring stiffness. It is very important to have a stable behaviour of the entire floating structure in any case of water depth (finite depth, infinite depth or shallow water). Assuming that the fluid is inviscid and incompressible, and the motion is non-rotational and simple harmonic in time with angular frequency ω , the fluid structure interaction analysis showed that the number of zeros in the reflection coefficient is maximum in the case of infinite water depth and minimum in case of the shallow water approximation. It was observed that the multiple articulated plates were behaved like a single continuous plate if the vertical linear springs and the flexural rotation springs are operating simultaneously. Furthermore, in (Ohkusu, 2004), the fluid structure interaction of **a large and thin floating structure** exposed to the sea wave action is analyzed. An analytical approach to predict vibration of this class of structures was developed. The plate vibration is obtained in an

explicit analytical form and is useful in preparing the computational approach for the real structure of a floating airport.

Flow induced vibrations in heat exchanger tubes is another example of FSI, (Mitra, 2009). The induced vibration in these structures is caused by vortex shedding behind the bluff body of the heat exchanger tubes, the fluid-elastic instability and turbulent buffeting. The experiments shows that fully flexible arrays become unstable at a lower flow velocity and those tubes are more stable in steam-water flow as compared to air-water flow. It was found that the boiling water could have a stabilizing effect on fluid-structure instability. **Fluttering and galloping** of lightweight structures are typical results of the fluid structure interaction. The flow regime is important in the analysis of these phenomena, (Barrero-Gil, 2009). Numerical simulations of a transverse galloping phenomenon of a square cylinder at low Reynolds numbers confirm the possibility of galloping with no hysteresis for $Re > 159$.

Large domains of technical problems are dominated by **vortex-induced vibrations** (VIV) as an effect of fluid-structure interaction phenomenon, (Galvao, 2008). The controlling of the wake behind a bluff body means to eliminate unsteady transverse loads while reducing drag. Experimental procedure and numerical simulation demonstrated that it could be obtained a potential flow in the wake behind the bluff body if combinations of flow directing hydrofoils are attached behind a circular cylinder. Two and four symmetric hydrofoils, and a triangular fairing attachment configuration are studied as a **passive control** method with the consequences of no VIV and small drag forces on the cylinder. The stability behaviour of a wake flow pattern is influenced by Reynolds number, turbulence intensity, aspect ratio, end effect, wall proximity, (Kuo, 2009). **Active and passive control** methods could be used to control the flow in the wake behind a bluff body so that could be reduced the form drag, could be suppress vortex shedding or could be changed the heat transfer characteristics. For example, two small control cylinders, with diameter ratio $d/D = 0.25$ are placed symmetrically along the separating shear layers at various stream locations to control the wake behind a circular cylinder in uniform flow at $Re_D = 80$. Further, in (Pereira Gomes, 2011) the flow-structure interaction a reference test case involving the coupling of unsteady fluid flow and structure motion is studied. It was considered the structure of an aluminium front **cylindrical body** with an attached elastic thin metal plate including a rear mass at the trailing edge. The structure is fixed with one rotational degree of freedom located in the centre of the model front cylinder. The structure is such designed as to attain a self-exciting periodical swivelling movement when exposed to a uniform laminar flow with a Reynolds number up to 270. Good results in the time-phase space were obtained regarding the reproducibility of the coupled fluid-structure motion. An experimental approach of a **structure array of cantilever beams** is demonstrated that neighbouring beams interact through the fluid and their dynamic behaviour is modified, (Kimber, 2009). Aerodynamic interaction between neighbouring cantilever beams operating near their first resonance mode and vibrating at amplitudes was found to be comparable to their widths. Experimental correlations were found to be used to predict the aerodynamic damping in arrays of vibrating cantilevers. In (Facchinetti, 2004) the model of a VIV is described as a one degree of freedom system, elastically supporting a rigid circular cylinder constrained to oscillate transversally to a stationary and uniform flow of free stream velocity. The fluctuating behaviour of vortex street is modelled by a nonlinear oscillator satisfying the van der Pol equation. The wake oscillator was coupled with the structure oscillator and generic forms of coupling have been qualitatively and quantitatively analyzed. The van der Pol wake oscillator model may be extended to 3-D vortex induced vibrations.

Applications in **medicine** regarding the interaction between of a rigid, spherical cancer cell with a deformable white blood cell were simulated by developing a quasi-steady technique, (Hoskins, 2009). The six-degree-of-freedom motion, adhesion kinetics, structural mechanics, and fluid dynamics governing equations were all solved, in the context of an octree-based adaptive mesh. The Lagrangian approach was used in grid generation to simulate a cellular system. Weakly coupled fluid-structure interaction models were used for the analysis of the periodic unsteady incompressible flow inside compliant vessels and to simulate **the blood flow in arteries**, (Beulen, 2009). Here it was successfully applied a time-periodic method which proved to have a far better computational stability than the weakly coupled methods based on time step-wise coupling. The method was applied to straight, curved and bifurcating vessels geometries.

Commonly, the coupled fluid-structure dynamic analysis problem in **turbomachinery** could be performed by considering the two indirect coupling methods: a) one based on the cyclic symmetry properties of both structure and fluid, or b) the uncoupled approach which assumes that there is no aerodynamic coupling between the modes and aerodynamic forces, (Tran, 2009). A multi-parameter minimum state modelling method was developed using the spline approximation and the minimum state modelling applied to a numerical model of an aircraft engine compressor disk. In this manner, the number and the cost of the aerodynamic computations in the solutions of the aeroelastic systems were reduced. Further, in (Gnesin, 2004), a fluid structure interaction problem is analyzed as the interaction of aerodynamic, inertial and elastic forces acting on **the blades of the rotor of a turbine stage**. Mathematical model and numerical results for aeroelastic behaviour of a steam turbine last stage with rotor blades of 760 mm are presented. The algorithm proposed involves the coupled solution of 3D unsteady flow through a turbine stage and the dynamics problem for rotor-blade motion by the action of aerodynamic forces, without separating the outer and inner flow fluctuations. The unsteady Euler equations described the 3-D transonic gas flow through the stator and rotor blades in relative motion. An explicit monotonous finite volume difference scheme Godunov-Kolgan scheme was used. Modal approach and a 3-D finite element model of a blade are used in the structural analysis. It was obtained the high-frequency harmonics which are corresponding to the rotor moving past one stator blade pitch. Low-frequency harmonics are caused by blade oscillations and flow non-uniformity downstream from the blade row. Based on experimental **flutter stability** data acquired and using a new model of actuator disk approach, a **reduced-order model of a compressor** was analyzed in (Copeland, 2004). Here it was considered the actuator disk model for the blade rows. The control volume approach was used to determine the aerodynamic force on the flexible blades. Numerical simulations were used to analyze the dynamic behaviour of the compressor and the predicted flutter and stall stability boundaries were proposed.

A numerical method and a procedure to calculate the flow-induced noise in a **centrifugal pump** is presented in (Langthjem, 2004a). First a hydrodynamic incompressible analysis is done to obtain the "background-flow" and noise generating fluid forces. The flow-induced noise in a centrifugal pump was calculated using a computationally method to estimate the noise generating "background-flow". The analysis is restricted to a two-dimension formulation. The fluid structure interaction between the fluid, the rotating blades and the volute tongue, and the interaction between the fluid and the rotor alone are the causes of pressure fluctuations and correspond to dipole sources. Using an acoustic analogy the geometry of the pump was assimilated with a point source for the pump's inlet, the blades of the impeller are covered with vortex elements with discrete, bound vortices and the

casing is covered with source panels. A discrete vortex method was used to simulate the flow within a centrifugal pump. The unsteady impeller blade surface forces are estimated. The velocity and pressure fluctuations on the impeller blades are dominated by rotational frequency and its higher harmonics. Further, (Langthjem, 2004b), a hydroacoustic compressible analysis was considered to find estimate the noise in a centrifugal pump. A two-dimensional numerical method to estimate the acoustic pressure fluctuations in a centrifugal pump has been described. The unsteady surface forces which act on the rotor blade are considered to be acoustic dipoles. The discrete vortex method was used to estimate the strengths of the dipoles and to obtain the pressure distribution, and then to obtain the velocity field by applying the unsteady Bernoulli equation. A boundary element method was used to obtain the strengths of the dipoles. The frequency-domain solution is useful in design optimization so that the flow-noise is minimized.

In (Howell, 2009) it was developed a new computational model of the linear fluid-structure interaction of a **cantilevered-free flexible plate** with an ideal flow in a channel. The transient behaviour of the system is analysed by numerical simulations and a global linear-stability map of the system for the infinite-time limit is obtained. The effects of shed vorticity, channel walls, a rigid-inlet surface, temporally varying inlet flow-velocity, and variable plate stiffness were investigated and the flutter instability dependence upon system configuration was found. For example, a short flexible plate (in standard configuration) is destabilised by single-mode flutter caused by an irreversible energy transfer from fluid to structure that principally occurs over the middle part of the flexible plate. A long flexible plate is destabilised by a modal-coalescence flutter and the region of the plate where most destabilising energy transfer occurs mainly downstream half.

Aeroelastic design considerations related to **long-span bridges** and VIV are also described in (Frandsen, 2004). Governing design criteria for long-span bridges involve the aeroelastic phenomena of vortex-induced oscillations, flutter and buffeting. So, an acceptable flutter limit is one of the principal design criteria for long-span bridges. Scanlan's linearized theory assumes the prescribed motions and is used to estimate the flutter derivatives. Theodorsen's inviscid flat-plate theory is used to estimate the flutter derivatives, too. This study attempts to investigate the use of a coupled fluid-structure interaction finite element solver applied to a long-span bridge. Aerodynamics effects of bridge flutter are investigated using fluid and structural two-dimension finite elements on moving non adaptive grids. The moving interface between fluid and structure is modelled through the arbitrary Lagrangian-Eulerian formulation. It is verified that the flat-plate theory of Theodorsen gives comparatively accurate solutions despite of inviscid flow hypothesis. Also, the prediction of flutter instability for sharp edge bridge deck does not appear sensitive to turbulence and three-dimensional modelling.

In the field of aeronautical applications, the fundamental research in the field of FSI continues to go forward. For instance, in (Dessi, 2004) a three-degree-of-freedom of an airfoil with a **control surface** is the physical model which includes different. A 2-D incompressible potential flow has been considered in the model. A standard Runge-Kutta algorithm in conjunction with a 'shooting method' was used to numerically integrate the governing equation and a **stable and unstable limit cycle oscillation** was obtained. The amplitudes and frequencies of limit cycles dependences on the flow speed V_∞ are obtained. The terms, from the normal-form equations, which are essentially responsible for the nonlinear system behaviour are identified. Using a discrete gust model, Dessi identified and analyzed the damped or undamped **wing oscillations** for different gust's parameters, i.e., intensity and gradient, (Dessi, 2008). Stability analysis has been carried out on a simplified aeroelastic

model in two cases, without and with gust excitation. In the first case, an approximation of the basin of attraction of stable limit cycles in the space of initial conditions was discussed. In the second case, was identified a critical gust intensity for a given gust gradient.

Flow control techniques are currently investigated as a primary tool for the **control of aeroelastic vibrations**. **Active control** of flow separation over an air foil using synthetic jets is an already operational technique, see (You, 2008). An unstructured-grid finite-volume large-eddy solver was used to simulate the behaviour of flow inside the synthetic-jet actuator and the synthetic-jet/cross-flow interaction. Numerical simulations confirm 70% increase in lift coefficient if the flow separation is delayed with a synthetic-jet actuation. In (Levasseur, 2008) it was proposed the **control of cavity** flows with two passive acoustic oscillation suppression devices: the rod-in-cross flow and the flat-top spoiler located in the upstream boundary layer. The average reduction of the total pressure level was similar for both devices and was obtained 3–4 dB reduction. The selected test-case is a cavity of length/depth ratio equal to 5, at Mach $M_\infty = 0.85$ and Reynolds number of $Re_L = 7 \cdot 10^6$.

Finally, a comprehensive and utilitarian review of experimental and numerical modelling on **vortex-induced vibrations** in the last twenty years, (Sarpkaya, 2004) and on open FSI problems (Païdoussis, 2005), are useful references to understand the consequences of the fluid structures interaction in technical applications.

One of the conclusions coming off from the examples above is that the numerical solution of the FSI problems plays an essential role in any kind of work in this field: research, design, decision make.

2. Numerical solution of FSI problems

In the classical linear approach the behavior of a fluid-structure system is investigated in the frequency domain, (Bishplinghoff et al., 1955; Fung, 1956). In this case, the calculation of either stability boundaries (flutter problem) or dynamic response of the linear system is separated from the computation of linearized unsteady aerodynamic forces. However, in order to investigate nonlinear fluid-structure interactions the governing equations of structural and fluid motion have to be solved in the time domain, (Alonso & Jameson, 1994; Carstens et al., 2003; Liu et al., 2001). This can be done either with time-staggered algorithms (Piperno et al., 1995; Lesoinne & Farhat, 1996) or with coupled algorithms, (Alonso & Jameson, 1994). In the first case optimized numerical integration methods for each of the structural and fluid dynamics models can be used and thus the two sets of equations are treated separately and integrated in a leap-frog fashion. The drawback of this approach is that one can never have a fully converged fluid-structure system at any one time step. This may cause significant energy errors for large times and thus the prediction of the dynamic response and mainly of dynamic instabilities is uncertain. The full coupling of the flow equations with the structural model can be achieved efficiently using an implicit time advancement scheme based on the dual-time approach. The aeroelastic system is thus fully coupled and at each time step a fully converged solution is obtained. In order to avoid the grid generation at each time step the transpiration type boundary conditions are used for the flow equations.

2.1 The mathematical model and numerical solution of the fluid-structure Interaction problem

The numerical simulation of Fluid-Structure Interaction (FSI) phenomena is based on a mathematical model that describes the structural and fluid dynamics and their coupling

through the boundary conditions. There are different levels of accuracy in the modeling of both the structure behavior and the fluid flow.

2.1.1 The computational structural dynamics problem and its use in aeroelasticity

Usually the equations of structural dynamics are obtained using the Finite Element Method and represented by a system of second order differential equations, (Przemieniecki, 1968; Zienkiewics, 1971). The linear behaviour of a structure under the action of time dependent external loads is thus described using a finite element model which, for small deformations has the well-known linear form:

$$\underline{M}\ddot{\underline{q}} + \underline{C}\dot{\underline{q}} + \underline{K}\underline{q} = \underline{Q} \quad (1)$$

If natural coordinates are used, then the left-hand side of eq. (1) has a decoupled form. Starting from Eq. (1) and using appropriate initial conditions, the dynamic response of the structure can be calculated using dedicated numerical algorithms like those known as Newmark or Hughes-Hilbert-Taylor. However, for aeroelastic calculations the right-hand side of Eq. (1) represents the generalized aerodynamic forces that depend on the generalized coordinates and can be calculated only by solving the flow problem. The structural problem can be rewritten in the phase-state space as:

$$\frac{d\underline{x}}{dt} = \underline{A}\underline{x} + \underline{F}(\underline{U}_f(\underline{x}, \dot{\underline{x}})) \quad (2)$$

where the state variable vector is:

$$\underline{x} = \{\underline{q}, \dot{\underline{q}}\}^T \quad (3)$$

and with the matrix

$$\underline{A} = \begin{bmatrix} 0 & \underline{I} \\ -\underline{M}^{-1}\underline{K} & -\underline{M}^{-1}\underline{C} \end{bmatrix} \quad (4)$$

and the force term:

$$\underline{F} = \left\{ \begin{array}{c} 0 \\ \underline{M}^{-1}\underline{Q}(p(\underline{U}_f(\underline{x}, \dot{\underline{x}}))) \end{array} \right\} \quad (5)$$

Finally, the Eq. (2) is accompanied by the initial condition:

$$\underline{x}(t=0) = \underline{x}_0 \quad (6)$$

In classical aeroelastic investigation the force term given by Eq. (5) is linearised and, in some simple but useful cases it can be expressed analytically, (Bishplinghoff et al, 1955; Bishplinghoff & Ashley, 1956; Fung, 1956). Under additional assumptions like harmonic time dependence of the displacements and forces the solution of the classical aeroelastic problem of flutter boundaries calculation is simplified from Eq. (2) to an eigenvalue problem. This is the so-called flutter analysis in the frequency domain, which is a usual practice in aeroelasticity.

If the linearization is not possible, than Eq. (5) shows that the structural problem cannot be separated from the flow problem and the FSI has to be investigated in the time-domain. This means that the solution of Eq. (2) is obtained numerically, using an explicit or implicit solver that advances in time the structure state.

However, in the phase-space the numerical solution of Eq. (2) with (6) can be done using the so-called time staggered algorithms, (Piperno et al., 1995; Lesoinne & Farhat, 1996). These allow a formal decoupling of the structural and flow problems through a time step. The main advantage of the time staggered algorithms is that there are now available optimised numerical solvers for each of the two problems, taken individually. Their drawback is that the structure and fluid states are not determined exactly at the same time and thus the energy transfer from fluid to structure and reverse cannot be accurately predicted (Alonso & Jameson, 1994; Carstens et al., 2003). This has a negative impact on the accurate time-domain calculation of flutter like instabilities of the aeroelastic systems.

2.1.2 The computational fluid dynamics problem

In the last decades, unsteady Computational Fluid Dynamics (CFD) has emerged as the basic approach available for predicting the flow behaviour the case of FSI problems, (Anderson et al., 1984; Ferziger & Peric, 1999; Löhner, 2008).

In the present work the fluid is considered inviscid, in order to simplify the mathematical formulations. For compressible flows with shocks, the Euler equations of gas dynamics written in conservative form and with standard notations are:

$$\frac{\partial \underline{U}}{\partial t} + \frac{\partial \underline{f}(\underline{U})}{\partial x} + \frac{\partial \underline{g}(\underline{U})}{\partial y} + \frac{\partial \underline{h}(\underline{U})}{\partial z} = 0 \quad (7)$$

where the unknowns and fluxes are given by:

$$\begin{aligned} \underline{U} &= \{\rho \quad \rho u \quad \rho v \quad \rho w \quad \rho E\}^T, \\ \underline{f} &= \left\{ \rho u \quad \rho u^2 + p \quad \rho uv \quad \rho uw \quad \rho u \left(E + \frac{p}{\rho} \right) \right\}^T, \\ \underline{g} &= \left\{ \rho v \quad \rho uv \quad \rho v^2 + p \quad \rho vw \quad \rho v \left(E + \frac{p}{\rho} \right) \right\}^T, \\ \underline{h} &= \left\{ \rho w \quad \rho uw \quad \rho vw \quad \rho w^2 + p \quad \rho w \left(E + \frac{p}{\rho} \right) \right\}^T \end{aligned} \quad (8)$$

The pressure is determined through the state equation for a thermodynamic perfect gas:

$$p = (\kappa - 1) \left[\rho E - \frac{(\rho u)^2 + (\rho v)^2 + (\rho w)^2}{2\rho} \right] \quad (9)$$

The initial condition necessary for the solution of Eqs. (7, 8) states that in the entire computational domain the fluid state is known at the beginning of the simulations, i.e.:

$$\underline{U}(t=0) = \underline{U}_0 \quad (10)$$

It is known that for flows at low Mach numbers the Eqs. (7, 8) become difficult to be solved numerically, because of the ill-conditioning that imposes very small time steps to be used. For practical purposes it is therefore preferable to replace the inviscid compressible equations with the inviscid incompressible flow model in the Chorin's formulation:

$$\underline{\Theta} \frac{\partial \underline{V}}{\partial t} + \underline{\Phi} \left(\frac{\partial \underline{f}^i(\underline{V})}{\partial x} + \frac{\partial \underline{g}^i(\underline{V})}{\partial y} + \frac{\partial \underline{h}^i(\underline{V})}{\partial y} \right) = 0 \quad (11)$$

where the unknowns and the flux terms are:

$$\begin{aligned} \underline{V} &= \{P \quad u \quad v \quad w\}^T, \\ \underline{\Theta} &= \text{diag}[0, \quad 1, \quad 1, \quad 1], \underline{\Phi} = \text{diag}[\beta^2, \quad 1, \quad 1, \quad 1], \\ \underline{f}^i &= \{u \quad u^2 + P \quad uv \quad uw\}^T, \\ \underline{g}^i &= \{v \quad uv \quad v^2 + P \quad vw\}^T, \\ \underline{h}^i &= \{w \quad uw \quad vw \quad w^2 + P\}^T. \end{aligned} \quad (12)$$

In Eqs. (11, 12) the pressure is defined by $P = p/\rho$ and the compressibility factor β has to be provided. This conservative formulation of the incompressible flow equations allows a numerical treatment similar to that used for the compressible equations. The initial condition for the Eqs. (11, 12) is similar to Eq. (10).

For both inviscid flow models used in this work, **the coupling of the structural and fluid fields** is assured by the statement that the fluid normal velocity on the moving contact surface S equals the wall normal velocity $w_n(\underline{z}, t)$, $\underline{z} \in S(\underline{x}(t))$, i.e.:

$$u \cdot n_x + v \cdot n_y + w \cdot n_z = w_n(\underline{z}, t) \quad (13)$$

This physical boundary condition at the moving solid wall is used for all the inviscid flow equations, accompanied by other appropriate boundary conditions imposed on all the boundaries of the computational domain. **The pressure field on the solid boundary** is obtained from the solution of the flow equations and is used in Eq. (2) through the force term given by Eq. (5).

There are a lot of numerical methods that can be used to solve the flow equations. Starting from the conservative of the unsteady flow equations, the most common way is to use the finite volume method, (Batina, 1991; Ferziger & Peric, 1999; Frink, 1992; Nkonga & Guillard, 1994). The spatial discretization of the flow equations using a cell-centred Godunov type finite volume method leads to a system of differential equations written in semidiscrete form:

$$\frac{d\underline{U}_f}{dt} = R(\underline{U}_f(\underline{x}, \dot{\underline{x}})) \quad (14)$$

where the right hand side is the residual vector that contains the discretised fluxes and includes implicitly the boundary conditions given by Eq. (13). The left hand side contains the cell-averaged values of the fluid states on a current cell:

$$\begin{aligned}\underline{U}_f(t) &= \{\underline{U}_e(t) \mid e = 1, M\}^T, \text{ with} \\ \underline{U}_e(t) &= \frac{1}{\Omega_e} \int_{\Omega_e} \underline{U}(t) \cdot d\Omega\end{aligned}\quad (15)$$

CFD now makes available a variety of explicit and/or implicit algorithms dedicated to the solution of the unsteady flow problems modelled by Eqs. (14, 15) and with appropriate boundary and initial conditions, (Aftosmis et al., 1995; Alonso & Jameson, 1994; Ferziger & Peric, 1999; Frink, 1992; Singh et al., 1995; Trepanier et al., 1993).

2.1.3 The fluid-structure interaction model

From Eqs. (2) and (14) the coupled system of FSI equations in semidiscrete form is obtained as:

$$\frac{d}{dt} \begin{Bmatrix} \underline{x} \\ \underline{U}_f \end{Bmatrix} - \begin{Bmatrix} \underline{Ax} + \underline{F}(\underline{U}_f(\underline{x}, \dot{\underline{x}})) \\ R(\underline{U}_f(\underline{x}, \dot{\underline{x}})) \end{Bmatrix} = 0 \quad (16)$$

The coupling in Eq. (16) is generally nonlinear. This equation clearly shows that the fluid and structure states should be evaluated exactly at the same time. However, there are classical **decoupled** solvers using **time-staggered algorithms**. The most known example is due to (Piperno et al., 1995), who introduced a difference of $\Delta t / 2$ between the structure and fluid state calculations.

2.1.4 A fully-implicit numerical scheme

The direct solution of Eq. (16) by an explicit or implicit time advancement algorithm eliminates the already mentioned drawbacks of the time staggered algorithms.

The coupled FSI equations have at least two important features. The first is the huge number of unknowns to be solved simultaneously. It is therefore a necessity the development and implementation of efficient and less time consuming algorithms for the solution of the coupled equations. The second aspect to be taken into account is that the typical time integration steps imposed by accuracy and/or stability reasons usually differ by order of magnitude between the CSD and CFD problem. For instance, in aeroelastic problems the number of unknowns of the CFD problem considerably overrides the number of unknowns of the CSD problem. Further, the temporal resolution of the CSD problem, where usually only the low frequency vibration modes of the structure are investigated, is much greater than that required by the accurate solution of the CFD problem with a numerically stable method. Therefore, the best choice for the solution of Eqs. (16) seem to be the fully implicit time advancement, coupled with an efficient dual-time stepping for the solution of the nonlinear discrete equations, (Alonso & Jameson, 1994). One can rewrite the coupled fluid-structure system given by Eq. (16) in a more convenient form of a system of ordinary differential equations:

$$\frac{d\underline{w}}{dt} + \underline{R}_w(\underline{w}, \dot{\underline{w}}) = 0 \quad (17)$$

where the unknowns and residuals are:

$$\underline{w} = \begin{Bmatrix} \underline{x} \\ \underline{U}_f \end{Bmatrix}, \quad \underline{R}_w = - \begin{Bmatrix} \underline{Ax} + \underline{F}(\underline{U}_f(\underline{x}, \dot{\underline{x}})) \\ R(\underline{U}_f(\underline{x}, \dot{\underline{x}})) \end{Bmatrix} \quad (18)$$

A second order accurate fully implicit time discretization of Eq. (17) is given by:

$$\frac{3\bar{w}^{n+1} - 4\bar{w}^n + \bar{w}^{n-1}}{2\Delta t} + \underline{R}_w(\bar{w}^{n+1}) = 0 \quad (19)$$

For typical FSI problems Eq. (19) is a huge nonlinear algebraic system that must be solved at each time step. A better way to obtain the new values \bar{w}^{n+1} is to build up and solve a system of ordinary differential equations in the dual time t^* :

$$\begin{aligned} \frac{d\bar{w}^*}{dt^*} + \underline{R}^*(\bar{w}^*) &= 0, \quad t^* \rightarrow \infty \\ \bar{w}^*(t^* = 0) &= \bar{w}^n, \quad \text{where} \\ \bar{w}^{n+1} \stackrel{\text{not}}{=} \bar{w}^*, \quad \underline{R}^*(\bar{w}^*) &\stackrel{\text{not}}{=} \frac{3}{2}\bar{w}^* + R_w(\bar{w}^*) - \frac{4\bar{w}^n - \bar{w}^{n-1}}{2\Delta t} \end{aligned} \quad (20)$$

The residuals contain the boundary conditions and the force acting on the structure that effectively couple the two problems. The boundary conditions and the force are actualized at every internal step. Thus, the solution represents the fluid and structure states at the new physical time.

A four time steps Runge-Kutta scheme in the dual time provides the solution of Eq. (20) very efficiently for $t^* \rightarrow \infty$ for each physical time step. The dual time step Δt^* depends on physical time step Δt used in Eq. (19) and from numerical stability reasons should obey a condition like:

$$\Delta t^* \leq \frac{2}{3} \Delta t \quad (21)$$

Due to the fully implicit character of Eq. (19), which is also A-stable, the physical time step is determined from accuracy considerations only. The main advantage of this approach is the explicit advancement in the dual time space, which eliminates the solution of huge algebraic systems.

2.1.5 Problems related to moving grids and the geometric conservation law

This work deals with the computation of unsteady flows in moving geometries. Because the position of the solid boundaries, i.e. of the structure determines at least partially the fluid domain boundaries, it becomes necessary to perform the numerical calculation of the fluid flow on a moving mesh. The flow equations must therefore be re-written in an Arbitrary Lagrangian-Eulerian (ALE) formulation and the solution procedure of the FSI problem is coupled to the grid dynamics, (Batina, 1991; Ferziger & Peric, 1999; Lesoinne & Farhat, 1996; Trepanier et al., 1993).

It is generally recognized that it is difficult to generate a structured grid about complex configurations. Furthermore, this difficulty is magnified when the aeroelastic deformation of the aircraft or other displacements of the solid boundaries are considered, since the grid must move to conform to the instantaneous shape of the computational domain. As an alternative, algorithms have been developed that make use of unstructured grids. The unstructured grid methods, therefore, have the advantage (over structured grid methods in) that they can easily treat complex geometric configurations as well as complicated flow physics.

The grid management discussed here follows the work pioneered by (Batina, 1991), which has proposed a dynamic grid algorithm employable for problems where the solid body deforms or performs small displacements. The original unstructured grid is generated corresponding to the initial (undeformed) position of the boundaries. The edges of the grid are considered as a system of interconnected springs. In two dimensions, the unstructured grids typically are made up of triangles. Each edge of each triangle is modeled by a tension spring. The spring stiffness for a given edge $i-j$ is taken to be inversely proportional to the length of the edge raised to a power, which is written as

$$K_{ij} = \frac{1}{\left[(x_i - x_j)^2 + (y_i - y_j)^2 \right]^{p_{ij}/2}} \quad (22)$$

where p_{ij} is a parameter used to control the stiffness depending on its position relative to the moving solid wall. Usually, $p_{ij} \approx 1$, but we mention here the fact that we are interested to be able to maintain the grid quality by controlling the nodes displacements via this parameter. Then, by assembling the contributions of all edges, the unknown displacements $\delta \bar{x}$ of the free internal nodes and the known displacements $\delta \bar{x}_{mb}$ of the border corresponding to the moving body or other fixed boundaries are coupled by the finite-element type matrix homogeneous equation:

$$\begin{bmatrix} K & K_{coup} \\ K_{coup}^T & K_{mb} \end{bmatrix} \cdot \begin{Bmatrix} \delta \bar{x} \\ \delta \bar{x}_{mb} \end{Bmatrix} = 0 \quad (23)$$

The rigidity matrix in (23) is singular and the indices *coup* indicate coupling terms due to the springs connecting internal nodes with border nodes. The solution $\delta \bar{x}$ of the non-singular algebraic system obtained from (23) can be obtained iteratively, by using several Jacobi iterations. The initial guesses of the displacements are the displacements at the previous time level. An example of deforming grids is given in Figure 1.

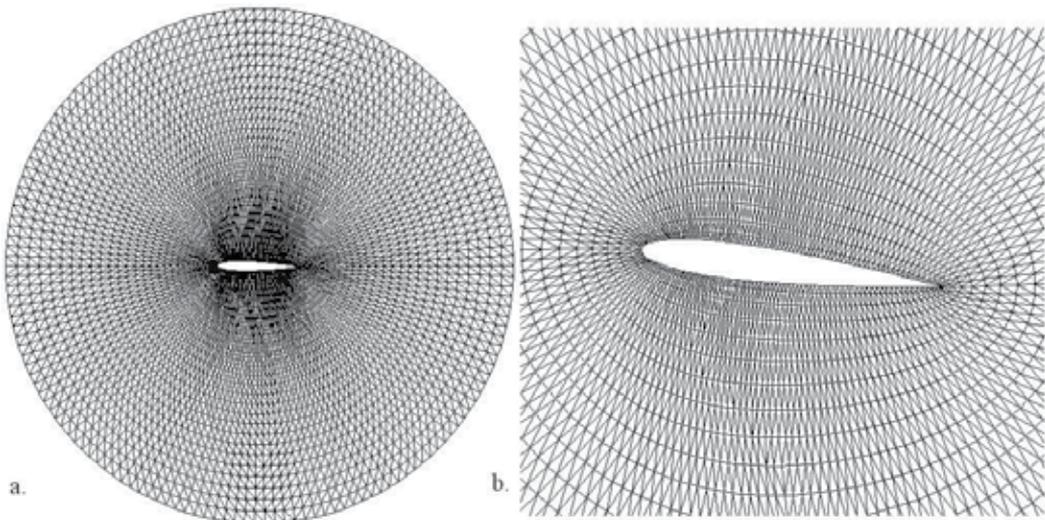


Fig. 1. The undeformed (a) and deformed (b) grid around a NACA 0012 airfoil.

The integral conservative form of the flow equations can be written in ALE formulation for a time-dependent bounded domain $\Omega(t)$ with a boundary $\partial\Omega(t)$ as

$$\frac{\partial}{\partial t} \iiint_{\Omega} \underline{U} dV + \iint_{\partial\Omega} \bar{F}(\underline{U}) \cdot \bar{n} dS = 0 \quad (24)$$

where the normal inviscid flux looks like:

$$\bar{F}(\underline{U}) \cdot \bar{n} = (V_n - \omega_n) \underline{U} + p \cdot \{0, n_x, n_y, n_z, V\}^T \quad (25)$$

Here $\bar{n} = (n_x, n_y, n_z)$ is the exterior unit normal on the boundary $\partial\Omega(t)$, the fluid velocity is \bar{V} and $\bar{\omega}$ is the velocity of the volume boundary. The projections of these velocities on the boundary normal are denoted with V_n and ω_n , respectively.

The time-dependent domain $\Omega(t)$ in (24) can be the entire computational domain as well as an arbitrarily specified control cell. When applied to a moving cell, the condition that (24) preserves the trivial solution of a uniform flow field leads to the integral condition called the **Geometric Conservation Law**:

$$\frac{\partial}{\partial t} \iiint_{\Omega} dV = \iint_{\partial\Omega} \bar{\omega} \cdot \bar{n} dS \quad (26)$$

The equation (26) describes a relationship where the time rate of change of the volume of any cell must be exactly balanced by the volumes swept by its boundaries. In order to avoid errors induced by the moving mesh, the condition (26) needs to be satisfied numerically, in addition to the flow equations. Furthermore, this condition must be solved numerically using the same scheme that is used to advance the conservation laws of the fluid to provide self-consistent solution for the local cell volumes. Thus, once the new positions of the moving nodes of the mesh have been calculated, the geometric conservation law is discretized and used to correlate the local cell volumes at the current time level and the normal velocity of the faces which is used in the normal numerical flux calculation. The GCL and the appropriate calculation of the normal velocities of the faces completely define the numerical problems related to the moving mesh.

2.1.6 Numerical implementation of the physical boundary condition for ALE formulations

On the solid moving walls, the boundary conditions like Eq. (13) are enforced by using the classical idea of image cells, (Frink, 1992). In these imaginary cells, a fictitious fluid state is defined by setting the flow variables according to the type of the reflective boundary. Then, the Riemann problem so created is solved with the approximate Riemann solver to compute the normal flux across these boundaries. For a moving solid wall, the relative normal velocity is reflected to ensure the impermeability condition. This gives for the normal velocity component in the imaginary cell:

$$V_n^{imag} = -V_n + 2 \cdot u_{wall} \quad (27)$$

The numerical boundary conditions for the pressure, density and tangent velocity fields are taken as those of the adjacent boundary cell. For a static solid wall, the normal velocity is simply negated.

The ALE formulations studied up to now has two practical drawbacks. First, at each time step anew grid must be built - in some way - on the fluid domain, and the associated grid velocity must be computed. Both grid deformations and velocities fields must follow the structure deformation and be smooth in time and space. Second, the flux vectors are modified by the ALE formulation and thus the corresponding solvers must be changed in depth.

The transpiration method is, quite simply, a means by which to trick the flow solver into seeing some sort of deflection in the mesh that is not actually there. If a change in surface normal is known, from structural dynamics solver for example, then this change in normal could be applied directly to the existing fixed grid through a slight modification of the existing surface normal. With transpiration, the nodes and faces affected by a surface deflection simply require a modification of its existing normal. Even though the surface is not actually deflected, all the flow sees is the normal at that particular nodal location, it does not matter what that normal is. Let's consider the case of a moving boundary represented in Figure 2.

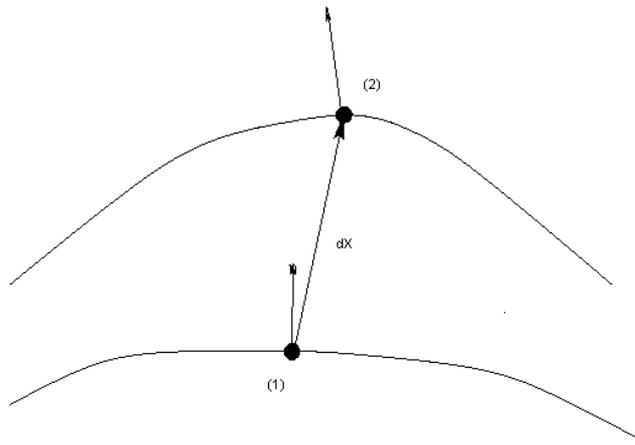


Fig. 2. The displacement of a surface point from the initial position (1) to the final position (2) during a time step

From the first order approximation of the normal velocity on the solid surface in the point (2) one obtains:

$$\bar{V}_2 \cdot \bar{n}_2 = \bar{V}_1 \cdot \bar{n}_1 + \frac{d(\bar{V}_1 \cdot \bar{n}_1)}{d\bar{X}} d\bar{X} + O(d\bar{X}^2) \quad (28)$$

Then, since \bar{n}_2 and thus the normal velocity $\bar{V}_2 \cdot \bar{n}_2$ are known from the structural problem, one can derive directly the transpiration velocity that modifies the normal velocity at the fixed point (1):

$$\bar{V}_1 \cdot \bar{n}_1 = V_T = \bar{V}_2 \cdot \bar{n}_2 - \frac{d(\bar{V}_1 \cdot \bar{n}_1)}{d\bar{X}} d\bar{X} \quad (29)$$

Usually, the right hand side term in Eq. (21) is approximated by:

$$V_T \approx \bar{V}_2 \cdot \bar{n}_2 \quad (30)$$

In order to overcome the difficulties due to the use of moving grids and to be able to solve at low costs fluid-structure interaction problems at moderate deformations, the transpiration method represents a useful engineering approach. Its most important advantage is that it does not require to update the computational grid or the flux solvers routines. It only involves modifications of the interface boundary conditions. However, this works well only for small deformations of the structure.

2.2 The sensitivity analysis for fluid-structure Interaction problems

The modeling procedure of a given fluid-structure system is required to analyze its dynamic behavior under different functional conditions. Furthermore, the mathematical modeling is a necessary step in the designing a control system with a given purpose. Usually, the FSI mathematical model is nonlinear and comprises independent and dependent variables, the dependent being related to the independent ones through a system of equations and relationships. By linearizing this system around an operating point a linear model can be obtained. This linear model describes the behavior of the FSI system around the equilibrium and thus can be used for stability as well as for dynamic response analysis.

Mathematical models include uncertainties, because of the parameters whose actual values are known only approximately or could vary around some reference values, (Cacuci et al., 1980; Cacuci, 1981, 1981). As an example taken from aeroelasticity, we enumerate here the flow speed, the air density, the elastic and inertial characteristics of the elasto-dynamic system and last but not least the aerodynamic derivatives.

The purpose of the local sensitivity analysis is to determine quantitatively the behavior of the system responses locally around a chosen point of the trajectory in the phase-space of parameters and dependent variables. The sensitivities are the derivatives of functional type responses with respect to the parameters. These sensitivities are useful for a better understanding of the behavior of the aeroelastic system. There are three methods that can be used to determine the sensitivities.

The first method is the most common. Running the code that calculates the system response a second time with an increment of a single parameter provides the sensitivity to that parameter. This method consumes large CPU times because for each sensitivity, one complete new run of the code is required. The second method that is available is called the forward sensitivity analysis method. It requires the construction of the linear forward sensitivity equations that may be derived directly from the original system in a consistent manner. Obviously, the sensitivity of the numerical solution to a parameter variable requires a complete solution of the linear forward sensitivity problem (FSP). The advantage is that this calculation is cheaper than the previous one, because of the linearity of the FSP. The disadvantage is that the conclusions are limited to small variations in parameters, (Appel & Gunzburger, 1997; Cacuci et al., 1980). The third method is called the adjoint sensitivity analysis method (ASAM). It requires the construction of the adjoint sensitivity equations, which may be obtained from the forward sensitivity equations and are also linear. Thus, it provides only linear answers. The advantage is that the adjoint function method is largely cheaper, because it provides the sensitivities to all parameters by solving only one time the adjoint equations. The adjoint method has been originally developed in (Cacuci, 1981), who introduced the basic concepts of a comprehensive sensitivity theory. This theory gives precise mathematical and physical meaning to the concept of the sensitivity of a response to parameters that are functions. The adjoint method has been

extensively used mainly in the field of internal flows simulations, (Cacuci, 1981, 1981; Ounsy et al., 1994; Toomarian et al., 1988).

Remark. In parallel with the sensitivity analysis of a response, for stability analysis purposes it is also useful to derive a computational frame for the calculation of the sensitivities of the eigenvalues resulting from eigenvalue problems.

In this paragraph we present the theoretical developments necessary for the application of both the forward and the adjoint sensitivity analysis methods for a typical FSI mathematical model. Starting from the state-space formulation we present the formalism that uses adjoint functions associated to state variables to determine sensitivities of the response to all the parameters. The response considered here is of functional-type and the parameters are represented by the initial conditions of the base-case problem and by the inertial, elastic and aerodynamic coefficients of the model. The adjoint system satisfied by the adjoint functions is determined and shown to be a well-posed linear system of ordinary differential equations subjected to specific initial.

2.2.1 The basics of the adjoint sensitivity analysis method

For the purposes of sensitivity analysis, let the spatial-discretized FSI mathematical model be represented in operator form as a system of nonlinear time-dependent differential equations:

$$\begin{aligned} O(X, \underline{G}) &= 0, t \in [0, t_f] \\ X(0) &= X_0 \end{aligned} \quad (31)$$

where the state variables are $X(t)$ and $\underline{G} = (G_1, \dots, G_M)$ is the vector containing all the problem parameters. We assume that the well-posed initial value problem (IVP) represented by Eqs. (31) is solved for a set of nominal (or base-case) parameter values \underline{G} and has a unique solution, denoted here by X . The response considered in this analysis is a nonlinear functional depending simultaneously on X and \underline{G} and generally defined as:

$$R(X, \underline{G}) = \int_0^{t_f} E[X(t), \underline{G}] dt \quad (32)$$

where t_f is some final time value and E is a nonlinear function depending on time and the system solution and parameters. The response of the aeroelastic system is the total mechanical energy at the final time, for instance. It is worthwhile to notice here that the integrals appearing in Eq. (32) may be considered as an inner product in the Hilbert spaces of all square integrable vector functions to whom the solution of Eq. (31) belongs, for instance:

$$\langle \alpha, \beta \rangle = \int_0^{t_f} \alpha(t) \beta(t) dt \quad (33)$$

It is obvious that small variations $\underline{g} = (g_1, \dots, g_M)$ in the system parameters induce variations in the solution, x , so that the perturbed solution $X + x$ is satisfying a system similar to Eq. (31). The objective of local sensitivity analysis is to analyze the behavior of the system responses locally around a chosen point (X, \underline{G}) of trajectory in the combined phase-space of state variables Q and parameters \underline{G} . The most general and fundamental concept for the

definition of the sensitivity of a response to variations in the system parameters is the Gateaux (G) differential. The G-differential $DP(\underline{S}, s)$ of an operator $P(S)$ at \underline{S} with an increment s is defined as

$$DP(\underline{S}, s) = \left[\frac{d}{d\varepsilon} P(\underline{S} + \varepsilon \cdot s) \right]_{\varepsilon=0} \quad (34)$$

Applying Eq. (34) to Eq. (32) yields the sensitivity $DR(X, \underline{G}; x, \underline{g})$ of the response as:

$$DR(X, \underline{G}; x, \underline{g}) = \int_0^{t_f} \left[\frac{\partial E}{\partial X} \right]^{(X, \underline{G})} \cdot x dt + \int_0^{t_f} \left[\frac{\partial E}{\partial \underline{G}} \right]^{(X, \underline{G})} \cdot \underline{g} dt \quad (35)$$

The first and the second terms of the Eq. (35) are customarily called the indirect effect term and the direct effect term, respectively. The exact value of the perturbed response is predicted by the sensitivity to first-order accuracy in parameters and solution variations, i.e.:

$$R(X + x, \underline{G} + \underline{g}) - R(X, \underline{G}) = DR(X, \underline{G}; x, \underline{g}) + \Theta \left(\|x\|^2 + \|\underline{g}\|^2 \right) \quad (36)$$

the indirect effect term and thus the desired sensitivity of the response functional can be evaluated only after determining the solution perturbations x . Up to first order, the relationship between x and \underline{g} is obtained by taking G-differentials of Eqs. (1). This leads to the FSP given by:

$$\begin{aligned} o(x, \underline{g}; X, \underline{G}) &= 0, \\ x(0) &= x_0, t \in [0, t_f] \end{aligned} \quad (37)$$

Once the solution of the base-case problem represented by Eq. (31) has been determined, then for a given vector of parameter changes the FSP could be solved to determine the variations in the flow problem solution, x . The solution of the FSP is advantageous to employ only if the number of different responses exceeds the number of system parameters. The alternative way, however, is to eliminate the explicit appearance of x in Eq. (35). This elimination process relies on the possibility to construct an adjoint operator corresponding to Eq. (37). Using the inner product given by Eq. (33) and integration by parts, one obtains from Eqs. (7) for the adjoint vector Φ of x :

$$\langle \Phi, o(x, \underline{g}; X, \underline{G}) \rangle = \langle x, o^*(\Phi; X, \underline{G}) \rangle + B(x, \Phi) \quad (38)$$

where $o^*(\Phi; X, \underline{G})$ is the adjoint operator to $o(x, \underline{g}; X, \underline{G})$ and $B(x, \Phi)$ is the associated bilinear form. Eqs. (37) and (38) yield:

$$\langle x, o^*(\Phi; X, \underline{G}) \rangle = -B(x, \Phi) \quad (39)$$

Further, **the explicit appearance** of the unknown x in Eq. (35) **is eliminated** if the adjoint vector function Φ is determined so that it satisfies the linear equation:

$$o^*(\Phi; X, \underline{G}) = \left[\frac{\partial E}{\partial X} \right]^{(X, \underline{G})} \quad (40)$$

The solution of the Eq. (40) can be done successfully only if one imposes appropriate initial conditions. These must be set so that all the terms involving unknown values of x disappear from Eq. (39). If this is done, then the sensitivity of the response is determined directly using the relationship:

$$DR(X, \underline{G}; x, \underline{g}) = \tilde{B}(x, \Phi) + \int_0^{t_f} \left[\frac{\partial E}{\partial \underline{G}} \right]^{(X, \underline{G})} \underline{g} dt \quad (41)$$

In Eq.(41) $\tilde{B}(x, \Phi)$ is the remaining part of the bilinear form $B(x, \Phi)$ after the elimination of the unknown terms $x(t_f)$ through the use of the final time conditions for Φ , i.e. $\Phi(t_f) \equiv 0$. This condition is transformed into an initial condition using the time variable change $\tau = t_f - t$. This leads to a new IVP for the adjoint function Φ :

$$o^*(\Phi; X, \underline{G}) = \left[\frac{\partial E}{\partial X} \right]^{(X, \underline{G})}, \quad \Phi(\tau) \equiv 0 \quad (42)$$

Thus, the adjoint functions are independent of parameter variations. The source term in Eq. (42) depends on the choice of the response. Finally, we notice that the IVP for the adjoint functions is linear and depends on the base-case solution X .

The local sensitivity analysis method offers therefore the possibility to quantify the influence of all the parameters on the response of an aeroelastic system. Compared with the direct and forward methods, the adjoint sensitivity analysis method is numerically efficient but requires important theoretical developments and these are summarized in this work. We notice that beyond the general form of the equations, the use of the adjoint sensitivity analysis method needs also a completely different and new computer code to be developed and tested.

2.2.2 The sensitivity of the eigenvalues in local stability analysis

For local stability analysis in the state-space form the forward linear system (37) or any other linear system has a homogeneous part written like:

$$\begin{cases} \dot{x} - A(X, \underline{G})x = 0, \\ x(0) = x_0 \end{cases} \quad (43)$$

where A is the state matrix. This depends on the base case solution and also on the problem's parameters, which are supposed to have small variations around the base-case values. Using the right eigenvectors problems for both the forward and adjoint homogeneous systems, one can proof that the sensitivity of the r^{th} eigenvalue Λ_r with respect the p^{th} parameter g_p can be determined using the following formula:

$$\frac{\partial \Lambda_r}{\partial g_p} = \Psi_r^T \frac{\partial A}{\partial g_p} \Phi_r \quad (44)$$

The reader is encouraged to obtain this formula as an exercise. The calculation of the derivative of the state matrix with respect to the parameters is not a simple task and for many situations it must be done numerically. However, the last equation is useful for understanding, analysis and design of the aeroelastic system.

3. Conclusion

In this chapter we first outline the general principles of FSI followed by some examples taken from aeronautical, civil and biomedical engineering. One of the conclusions coming off from these examples is that the numerical solution of the FSI problems plays an essential role in any kind of work in this field.

Then, we focused on the numerical simulation of the FSI problems using intensive computational techniques. These are based on a mathematical model that describes the coupling of structural and fluid dynamics models through the boundary conditions. There are different levels of accuracy in the modeling of both the structure behavior and the fluid flow. For simplicity reasons only we have chosen the case of inviscid flows interacting with linear structures. In the phase-space the numerical solution can be done using the so-called time staggered algorithms. These allow a formal decoupling of the structural and flow problems through a time step. The main advantage of the time staggered algorithms is that there are now available optimised numerical solvers for each of the two problems, taken individually. Their major drawback is that the structure and fluid states are not determined exactly at the same time and thus the energy transfer from fluid to structure and reverse cannot be accurately predicted. This is why a better way for the solution of the FSI equations is the fully implicit time advancement, coupled with an efficient dual-time stepping for the solution of the nonlinear discrete equations. The main advantage of this approach is the explicit advancement in the dual time space, which eliminates the solution of huge algebraic systems. Aspects related to the treatment of moving grids and the related Geometric Conservation Law and to the imposition of the physical boundary condition on the moving wall are also mentioned.

The last part of the chapter presents in some details the adjoint sensitivity analysis method, which is useful for a better understanding of FSI problems. The purpose of the local sensitivity analysis is to determine quantitatively the behavior of the system responses locally around a chosen point of the trajectory in the phase-space of parameters and dependent variables. The sensitivities are the derivatives of functional type responses with respect to the parameters. The local sensitivity analysis offers the possibility to quantify the role of the parameters on the response of an aeroelastic system. The adjoint sensitivity analysis method is numerically efficient but requires important theoretical developments and these are summarized in this work. We notice that beyond the general form of the equations, the use of the adjoint sensitivity analysis method needs also a completely different and new computer code to be developed and tested.

For stability analysis purposes the sensitivities of the eigenvalues of a state matrix to the problem's parameters are useful and we presented the way to follow to calculate them. The derivatives of an eigenvalue with respect to the problem's parameters show the importance of those parameters on the stability of the FSI system. The last equation is useful for understanding, analysis and design of the aeroelastic system.

4. Acknowledgment

This paper was partially supported by the Ministry of Education and Research, Romania, project CNCSIS ID_919/2007 and project 82-082/2008.

5. References

- Aftosmis, M., Gaitonde, D., Tavares, T.S. (1995). Behavior of Linear reconstruction Techniques on Unstructured Meshes, *AIAA Journal*, Vol. 33, No. 11, pp. 2038-2049
- Alonso, J.J., Jameson, A. (1994). Fully-Implicit Time-Marching Aeroelastic Solutions, *AIAA Paper 94-0056*, 32nd AIAA Aerospace Sciences Meeting and Exhibit, January 10-13, Reno, NV.
- Anderson, D. A., Tannehill, J.C., Pletcher, R.H. (1984). *Computational Fluid Mechanics and Heat Transfer*, McGraw-Hill Book Company, ISBN 0-07050328-1, USA
- Appel, J. R., Gunzburger, M. D. (1997). Difficulties in Sensitivity Calculations for Flows with Discontinuities, *AIAA Journal*, Vol. 35, pp. 842-848
- Batina, J.T. (1991). Unsteady Euler Algorithm with Unstructured Dynamic Mesh for Complex-Aircraft Aerodynamic Analysis, *AIAA Journal*, Vol. 29, No. 3, pp. 327-333
- Barrero-Gil, A., Sanz-Andrés, A., Roura, M. (2009). Transverse galloping at low Reynolds numbers. Brief Communication. *Journal of Fluids and Structures*, Vol. 25, pp. 1236-1242, ISSN 0889-9746
- Beulen, B.W.A.M.M., Rutten, M.C.M., Van de Vosse, F.N. (2009). A time-periodic approach for fluid-structure interaction in distensible vessels. *Journal of Fluids and Structures*, Vol. 25, pp. 954-966, ISSN 0889-9746
- Bishplinghoff, R.L., Ashley, H., Halfman, R.L. (1955). *Aeroelasticity*, Addison-Wesley Company, Cambridge, Massachusetts
- Bishplinghoff, R.L., Ashley, H. (1956). *Principles of Aeroelasticity*, John Wiley and Sons, New-York.
- Cacuci, D. G., Weber, C. F., Oblow, E. M., Marable, J. H. (1980). Sensitivity Theory for General Systems of Nonlinear Equations, *Nucl. Sci. Eng.*, Vol. 75, 88.
- Cacuci, D. G. (1981). Sensitivity Theory for Nonlinear Systems: I. Nonlinear Functional Analysis Approach, *J. Math. Phys.*, Vol. 22, 2794
- Cacuci, D. G. (1981) Sensitivity Theory for Nonlinear Systems: II. Extensions to Additional Classes of Responses, *J. Math. Phys.*, Vol. 22, 2803
- Carstens, V., Kemme, R., Schmitt, St. (2003). Coupled simulation of flow-structure interaction in turbomachinery, *Aerospace Science and Technology*, 7, pp. 298-306
- Copeland, G.S., Rey, G.J. (2004). Comparison of experiments and reduced-order models for turbomachinery high-incidence flutter. *Journal of Fluids and Structures*, Vol. 19, pp. 713-727, ISSN 0889-9746
- Dessi, D., Mastroddi, F. (2008). A nonlinear analysis of stability and gust response of aeroelastic systems. Brief Communication. *Journal of Fluids and Structures*, Vol. 24, pp. 436-445, ISSN 0889-9746
- Dessi, D., Mastroddi, F. (2004). Limit-cycle stability reversal via singular perturbation and wing-flap flutter. *Journal of Fluids and Structures*, Vol. 19, pp. 765-783, ISSN 0889-9746
- Doaré, O., Michelin, S. (2011). Piezoelectric coupling in energy-harvesting fluttering flexible plates: linear stability analysis and conversion efficiency. *Journal of Fluids and Structures*, doi: 10.1016/j.jfluidsstructs.2011.04.008

- Dowell, E.H. (1975). *Aeroelasticity of plates and shells*, Noordhoff International Publishing, Leyden
- Dowell, E.H., Ilgamov, M. (1988). *Studies in nonlinear aeroelasticity*, Springer-Verlag, New-York
- Facchinetti, M.L., de Langre, E., Biolley, F. (2004). Coupling of structure and wake oscillators in vortex-induced vibrations. *Journal of Fluids and Structures*, Vol. 19, pp, 123–140, ISSN 0889-9746
- Ferziger, J.H., Peric, M. (1999), *Computational Methods for Fluid Dynamics* (2nd edition), Springer, ISBN 3-540-65373-2, Germany
- Frandsen, J.B. (2004). Numerical bridge deck studies using finite elements. Part I: flutter. *Journal of Fluids and Structures*, Vol. 19, pp. 171–191, ISSN 0889-9746
- Frink, N.T. (1992). Upwind Scheme for Solving the Euler Equations on Unstructured Tetrahedral Meshes, *AIAA Journal*, Vol. 30, No. 1, pp. 70-77
- Fung, Y.C. (1956). *An introduction to the theory of aeroelasticity*, John Wiley and Sons, New-York
- Galvao, R., Lee, E., Farrell, D., Hover, F., Triantafyllou, M., Kitney, N., Beynet, P. (2008). Flow control in flow–structure interaction. *Journal of Fluids and Structures*, Vol. 24, pp. 1216–1226, ISSN 0889-9746
- Gnesin, V.I., Kolodyazhnaya, L.V., Rzadkowski, R. (2004). A numerical modelling of stator-rotor interaction in a turbine stage with oscillating blades. *Journal of Fluids and Structures*, Vol. 19, pp. 1141–1153, ISSN 0889-9746
- Hoskins, M.H., Kunz, R.F., Bistline, J.E., Dong, C. (2009). Coupled flow–structure–biochemistry simulations of dynamic systems of blood cells using an adaptive surface tracking method. *Journal of Fluids and Structures*, Vol. 25, pp. 936–953, ISSN 0889-9746
- Howell, R.M., Lucey, A.D., Carpenter, P.W., Pitman, M.W. (2009). Interaction between a cantilevered-free flexible plate and ideal flow. *Journal of Fluids and Structures*, Vol. 25, pp. 544–566, ISSN 0889-9746
- Karmakar, D., Bhattacharjee, J., Sahoo, T. (2009). Wave interaction with multiple articulated floating elastic plates. *Journal of Fluids and Structures*, Vol. 25, pp. 1065–1078, ISSN 0889-9746
- Kimber, M., Lonergan, R., Garimella, S.V. (2009). Experimental study of aerodynamic damping in arrays of vibrating cantilevers. *Journal of Fluids and Structures*, Vol. 25, pp. 1334–1347
- Kuo, C.-H., Chen, C.-C. (2009). Passive control of wake flow by two small control cylinders at Reynolds number 80. Brief Communication. *Journal of Fluids and Structures*, Vol. 25, pp. 1021–1028, ISSN 0889-9746
- Langthjem, M.A., Olhoff, N. (2004). A numerical study of flow-induced noise in a two-dimensional centrifugal pump. Part I. Hydrodynamics. *Journal of Fluids and Structures*, Vol. 19, pp. 349–368, ISSN 0889-9746
- Langthjem, M.A., Olhoff, N. (2004). A numerical study of flow-induced noise in a two-dimensional centrifugal pump. Part II. Hydroacoustics. *Journal of Fluids and Structures*, Vol. 19, pp. 369–386, ISSN 0889-9746
- Lesoinne, M., Farhat, C. (1996). Geometric conservation laws for flow problems with moving boundaries and deformable meshes, and their impact on aeroelastic computations. *Comput. Methods Appl. Mech. Engrg.*, 134, pp. 71-90
- Levasseur, V., Sagaut, P., Mallet, M., Chalot, F. (2008). Unstructured Large Eddy Simulation of the passive control of the flow in a weapon bay. *Journal of Fluids and Structures*, Vol. 24, pp. 1204–1215, ISSN 0889-9746

- Liu, F., Cai, J., Zhu, Y. (2001). Calculation of Wing Flutter by a Coupled Fluid-Structure Method, *Journal of Aircraft*, Vol. 38, No. 2
- Löhner, R. (2008). *Applied CFD Techniques. An Introduction Based on Finite Elements Methods* (2nd ed.), John Wiley & Sons, Ltd., ISBN 978-0-470-51907-3, GB
- Mitra, D., Dhir, V.K., Catton I. (2009). Fluid-elastic instability in tube arrays subjected to air-water and steam-water cross-flow. *Journal of Fluids and Structures*, Vol. 25, pp. 1213–1235, ISSN 0889-9746
- Nkonga, B., Guillard, H. (1994). Godunov type method on non-structured meshes for three-dimensional moving boundary problems, *Comput. Methods Appl. Mech. Engrg.* 113, pp. 183-204
- Ohkusu, M., Namba, Y., (2004), Hydro elastic analysis of a large floating structure. *Journal of Fluids and Structures*, Vol. 19, pp. 543–555, ISSN 0889-9746
- Ounsi, A., Brun, B., de Crecy, F. (1994) The adjoint sensitivity method, a contribution to the code uncertainty evaluation, *Nucl. Eng. Des.*, Vol. 149, 357
- Païdoussis, M.P. (2005). Some unresolved issues in fluid-structure interactions. *Journal of Fluids and Structures*, Vol. 20, pp. 871–890, ISSN 0889-9746
- Pereira Gomes, J., Yigit, S., Lienhart, H., Schäfer, M. (2011). Experimental and numerical study on a laminar fluid–structure interaction reference test case. *Journal of Fluids and Structures*, Vol. 27, pp. 43–61, ISSN 0889-9746
- Piperno, S., Farhat, C., Larrouturou, B. (1995). Partitioned procedures for the transient solution of coupled aeroelastic problems. Part I: Model problem theory and two-dimensional application, *Comput. Methods Appl. Mech. Engrg.* 124, pp. 79-112
- Przemieniecki J. S. (1968). *Theory of Matrix Structural Analysis*, McGraw-Hill, New York
- Sadeghi, M., Liu, F. (2001). Investigation of Non-Linear Flutter by a Coupled Aerodynamics and Structural Dynamics Method, *AIAA Paper 2001-0573*, 39th AIAA Aerospace Sciences Meeting and Exhibit, January 8-11, Reno, NV
- Sarpkaya, T. (2004). A critical review of the intrinsic nature of vortex-induced vibrations. *Journal of Fluids and Structures*, Vol. 19, pp. 389–447, ISSN 0889-9746
- Schmit, R.F., Glauser, M.N., Ahmadi, G. (2004). Flow and turbulence conditions in the wake of a H-section in cross flow. *Journal of Fluids and Structures*, Vol. 19, pp. 193–207, ISSN 0889-9746
- Singh, K.P., Newman, J.C., Baysal, O. (1995). Dynamic Unstructured Method for Flows Past Multiple Objects in Relative Motion, *AIAA Journal*, Vol. 33, No. 4, pp. 641-649
- Toomarian, N., Wacholder, E., Kaizerman, S. (1988). Sensitivity Analysis of Two-Phase Flow Problems, *Nucl. Sci. Eng.*, Vol. 99, 53, 1988
- Tran D.-M. (2009). Multi-parameter aerodynamic modeling for aeroelastic coupling in turbomachinery. *Journal of Fluids and Structures*, Vol. 25, pp. 519–534, ISSN 0889-9746
- Trepanier, J.Y., Reggio, M., Paraschivoiu, M., Camarero, R. (1993). Unsteady Euler Solutions for Arbitrarily Moving Bodies and Boundaries, *AIAA Journal*, Vol. 31, No. 10, pp. 1869-1875
- You, D., Moin, P. (2008). Active control of flow separation over an airfoil using synthetic jets. *Journal of Fluids and Structures*, Vol. 24, pp. 1349–1357, ISSN 0889-9746
- Zienkiewiecs O. C. (1971). *The Finite Element in Engineering Science* McGraw-Hill, London-New York

Study on Multi-Phase Flow Field in Electrolysis Magnesium Industry

Ze Sun, Guimin Lu, Xingfu Song, Shuying Sun,
Yuzhu Sun, Jin Wang and Jianguo Yu

*State Key Laboratory of Chemical Engineering
National Engineering Research Center for Integrated Utilization of Salt Lake Resources
East China University of Science and Technology, Shanghai
China*

1. Introduction

Magnesium, the 8th most abundant element in the earth's crust, was discovered and isolated by Sir Humphrey in 1808. Magnesium is classified as an alkaline earth metal. It is found in Group 3 of the periodic table. It thus has a similar electronic structure to Be, Ca, Sr, Ba and Rd. The density of magnesium at 20°C is 1.738 g/cm³. At the melting point of 650°C this is reduced to 1.65 g/cm³. On melting there is an expansion in volume of 4.2%. Magnesium is the lightest metal in large-scale commercial use. After World War II, the magnesium industry attempted to develop magnesium for a number of applications. Most of its peacetime uses take advantage of the light weight or other chemical properties. The uses of magnesium as a structural material were, however, very few. The bulk was used as an alloying element in aluminium alloys. Other uses, such as deoxidation of steel, chemical and pyrotechnics, outweighed structural uses, especially in energy-saving and environmental protection applications was wide because of its contribution to reduce energy consumption and greenhouse gas emissions. The most successful peacetime application for magnesium was in the original German Volkswagen car that was designed by Ferdinand Porsche. The VW Beetle used large magnesium alloy die castings for the crankcase and the transmission housing (both cast in halves) plus a number of smaller castings. Each Beetle contained more than 20 kg of magnesium alloy. Many of the other applications developed during the World War II, could not be quickly converted to civilian uses. Some of the uses such as aircraft wheels and aircraft engine castings and troop carrying buses were modified and then used for the basis of civilian industries. With more attention to energy and environment, magnesium will hold greater promise as a new weight-saving replacement for denser steel and aluminum alloys, and demands for magnesium will increase sharply in the future. In recent decades, demands for magnesium and its productivity increased sharply shown in Figure 1, the world productivity of magnesium reached 860,000 tons at 2007.^[1-3]

There are six sources of raw materials for the production of magnesium: magnesite, dolomite, bischofite, carnallite, serpentine and sea water. These sources differ in the magnesium content, in production methods, and in their origins. Some are mined from mines, some in open

mining, others originate in various processes carried out on sea water and salt lakes, and another material originates from the waste of the asbestos production process.

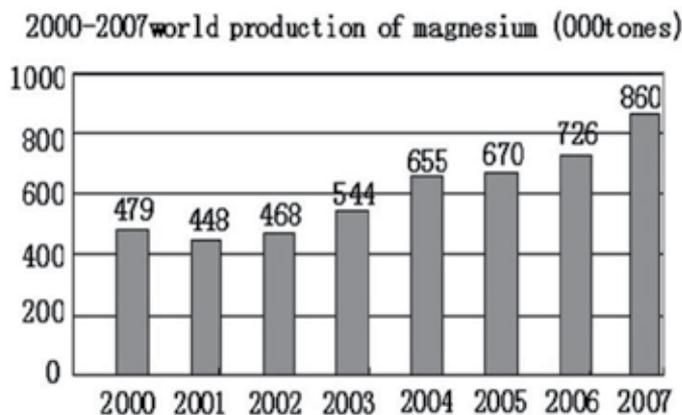


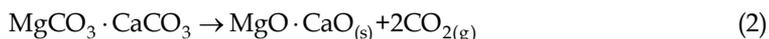
Fig. 1. 2000-2007 world production of magnesium

Material	Chemical formula
Magnesite	$MgCO_3$
Dolomite	$MgCO_3 \cdot CaCO_3$
Bischofite	$MgCl_2 \cdot 6H_2O$
Carnallite	$MgCl_2 \cdot KCl \cdot 6H_2O$
Serpentine	$3MgO \cdot 2SiO_2 \cdot 6H_2O$
Sea water	$Mg^{2+}_{(aq)}$

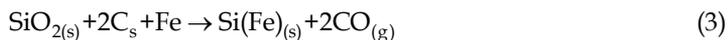
Table 1. Raw materials and their molecular formula

1.1 Thermal reduction methods

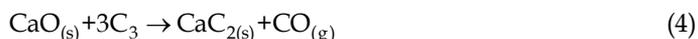
The only ores used in the production of magnesium with thermal reduction technology are dolomite and magnesite. These ores are extracted through customary mining methods, mainly through open mining. The ore extracted from the mine undergoes calcination at temperatures of 700-1,000°C. At this temperature the material releases CO_2 gas, according to Eqs. (1) and (2), for magnesite and dolomite, respectively [3,4]:



Following calcination the material is ground into a fine powder. Another main raw material that is used in thermal reduction is an alloy of silicon and iron called Ferrosilicon. The silicon content in this alloy is 65-85%, and at times the mix also contains small quantities of aluminum. The preparation of the ferrosilicon is carried out by reducing silica with coal, containing iron (scrap iron), at high temperatures [3,4]:



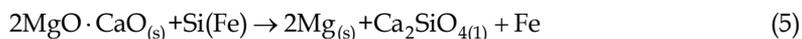
An additional material, which was used in the past as a thermal reduction agent, is Calcium carbide. The process of preparation of this material is relatively simple, but it requires high temperatures of about 1,800–2,000°C. Below is the reaction used in the production of this material:



In China, most thermal reduction magnesium plants adopted the traditional Pidgeon process.

1.2 The pidgeon process^[3,4]

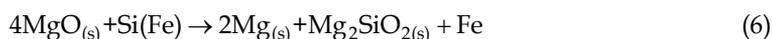
This process was developed in the 1940s in Ontario Canada, by Prof. Pidgeon and the Timminco company. Lately, this process has received new attention and constitutes a central process in the magnesium production at a large number of Chinese manufacturers. The reaction that takes place is eqs. (5).



$$P=1 \text{ at}'; T=1700^\circ\text{C} \text{ or } P=1 \text{ mm Hg}; T=1150\text{--}1200^\circ\text{C}$$

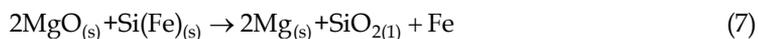
A retort used in this process. The dimensions of the retort are 2.7–3.3 m with a diameter of 28–35 cm. The capacity of the retort is about 120 kg. The source of energy in the process carried out in China is normally coal, while the calcination process and the heating furnaces require 14–20 t of coal for the production of one ton of magnesium.

On completion of the process a magnesium crown is obtained, weighing 12–20 kg, which is then extracted from the upper part of the retort. Due to the usually high temperature, the magnesium in this case will contain high concentrations of aluminum, manganese, iron and other impurities. The above process can also be carried out with magnesite as an alternative to dolomite; the working conditions are almost identical and the reaction in this case is Eq. (6).



$$P=10 \text{ mm Hg}; T=1220^\circ\text{C}$$

Another version of the above process is carried out at a higher temperature, according to the reaction appearing in Eq. (7).



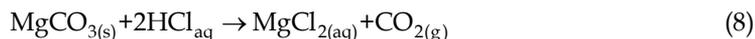
$$P=1 \text{ at}'; T=2300^\circ\text{C} \text{ or } P=1 \text{ mm Hg}; T=1500^\circ\text{C}$$

The advantage of this reaction is the higher output from each retort (up to about 80%), while the disadvantages in this case are many. The main one being the higher temperature required for the process, 1,500 °C, about 300 °C higher than with the regular process. Work at a higher temperature usually causes the evaporation of impurities and a lower quality of material. In addition, there is accelerated amortization of tools, and this in addition to the higher cost of energy and the need for accessibility to magnesite.

1.3 Electrochemical method

The basic raw materials for the production of magnesium with the electrochemical process are generally divided into two: salts containing chloride and raw materials that must be transformed into salts containing chloride. Eventually, all the materials will become either bischofite or carnallite prior to drying and feeding into the electrolysis cells.

In the hydro magnesium process used in Canada the raw material for the process is magnesite, which is mined in mines (mainly in China). The magnesite is dissolved in a hot HCl solution, thus arriving at a magnesium chloride-rich solution:



The acid solution must be hot for an efficient melting process. The basic ore of the magnesite is rich in various metallic impurities, sulfates and boron, which can be a problem at the electrolysis stage. Therefore, already at this stage a number of processes are carried out, whose purpose is to separate the impurities from the magnesium chloride.

1.4 Drying magnesium chloride salts

Carnallite, bischofite and their solutions contain large quantities of water. Carnallite and bischofite are hydrate salts containing six molecules of water per crystal and at times even additional water in attached form. The use of these materials in the electrolysis process requires the removal of all the water prior to the electrolysis.

Carnallite and bischofite tend to decompose in the hydrolysis reaction to magnesium oxide and HCl, at relatively low temperatures, and, therefore, the drying process is actually the most complicated and hardest stage in the electrolysis methods for the production of magnesium. In this matter it should be stated that most magnesium production R&D in the last decades has focused on the drying process for extraction of anhydrous magnesium chloride with low content of magnesium oxide.

Comparison with the thermal process, especially, the traditional Pidgeon process, the process of the magnesium production by electrolyzing the molten magnesium salt is more environmentally friendly. Moreover, in the electrolysis process, a high current efficiency (over 90%) can be reached with the application of the advanced diaphragmless electrolytic cell. Therefore, the research and development of the electrolysis technology of magnesium is more significant in future.

The electrolysis process of magnesium chloride is complicated, since there exist six important physical fields coupling with each other in the molten salt electrolyzer, namely electric field, magnetic field, flow field, temperature field, concentration field and stress field, all of which have the effects on the current efficiency, the direct current power consumption as well as the lifetime of the anode. Up to now, there are scarce researches on physical fields in the molten salt electrolyzer. So the optimization of the multi-fields is very significant for the optimum design of the large-scale magnesium electrolyzer. For example, in the electrolysis process of the molten magnesium salt, the current above 400kA, magnesium is produced from cathode as droplets, while the by-product chlorine gas is produced from anode as bubbles, both trying to rise the surface of the electrolyte as a result of the density difference with the electrolyte, so a three-phase flow under the strong electromagnetic field will be formed in the electrolytic cell^[5,6]. During the electrolysis process, the multi-phase molten electrolyte circulation is very important to the current efficiency, better circulation better current efficiency in electrolysis magnesium industry. Many scientists and engineers have

paid attention to the flow field. Two measurements, cold model experiments and numerical simulation, are frequently adopted to study the flow field in electrolyzer. Holliday^[7] used the similarity criteria to study electrolysis magnesium in low-density electrolytes. Burnakin used H₂O and argon gas to describe the hydrodynamics of a two-phase flow^[8]. He controlled gas flow rate corresponding to current density from 0.1–0.4 A/cm² to simulate working electrolyzer. Brunakin also built equations for two-phase flow characteristics for a magnesium electrolyzer^[9]. Korobov^[10] and Shcherbinin^[11-13] built mathematical model to simulate the temperature and electric fields in the magnesium electrolyzer. Agalakov^[14] investigated the flow field of the multiphase flow in the electrolyzer by the computational numerical simulation with the two-dimensional mathematical model under the following assumptions and simplify.

- Only focused on the flow phenomena
- Two-dimensional model
- Steady flow
- Bubbles with uniform radius
- No interaction between bubbles
- Flow field without electromagnetic force

He got two circulations in electrolyzer, the main circulation was the transportation circulation and the small circulation would be resulted in the damage of the back wall. Author's previous work established mathematical model of flow field and validated the math model using particle image velocimetry (PIV) data^[15]. Gökhan^[16] used new style electrolyzer to produce magnesium, and studied the fluid characteristics in the electrolyzer.

2. Method of study flow field in electrolysis magnesium

2.1 Cold model

In building a cold model relative to the original system, some similarity criteria are as follows:

1. Geometric Similarity

It is satisfied with Geometric Similarity requirements when the ratio of any length in model system to the corresponding length in the original system is the same everywhere.

2. Dynamic Similarity

A physical model is thought to be dynamic similarity with the original model when the magnitude of forces that cause movement at corresponding location in each model is in a fixed ratio. How to choose dynamic similarity condition is the key to successfully simulate the flow characteristic in the original system.

3. Kinematic Similarity

Kinematic similarity means similarity of motion. Velocity is kept in the close fixed ratio in the corresponding locations between the model system and original system.

4. Chemical Similarity

Chemical similarity concerns the establishment of the necessary conditions in the model so that the rate of chemical reaction at any location is proportional to the rate of the same reaction at the corresponding time and location in the prototype.

During establishing a cold model, it is difficult to satisfy all the similarity criteria simultaneously. But it is necessary to comply with important criteria that include some significant variables and try to follow the other similarity criteria.

According to the similarity criteria, such as geometric similarity, dynamic similarity, kinematic similarity, cold models of magnesium electrolysis cell were built^[15].

A laboratory scale cold model experiment was set up. A transparent vessel had the shape of a cube with $104 \times 4 \times 10$ cm in dimension. Lead anodes and graphite cathodes with $6 \times 2 \times 0.2$ cm were installed in parallel with each other, as shown as in Figure 2. The distance between anode and cathode was 80 mm. The electrolyte was zinc sulfate solution with 2 mol/L concentration. Direct current was added to anode and cathode with various current densities. In electrolyzing, zinc was deposited on cathodes and oxygen was released from anodes and moves upward the surface of the electrolyte.

The Particle Image Velocimetry system (PIV) is provided by Dantec Dynamics. It is a specifically novel non-intrusive visualization experimental technique, set up to provide instantaneous velocity vector measurements in a cross-section of the model. The available camera is part of the PIV system from Dantec Dynamics. Cameras use high-performance progressive-scan-interline CCD chips which includes 1018×1008 light-sensitive cells and an equal number of storage cells.

The target area was in the middle of the electrolytic cell perpendicular to the surface of electrode and limited one pair of electrodes. The area was located the mirror plane on electrodes. Usually, seeding particles are added to the water in order to observe the motion of the transparent water. In our experimental system, oxygen bubbles generated on anodes are small enough as tracking particles. The FlowMap software was used for post processing the images.



Fig. 2. Cold model experiment and PIV instrument

Figure 3 shows the typical streamlines of the flow field at a cross-section perpendicular to electrodes in cold model experiment system, measured by PIV instrument. In Figure 3, anode was at the left side of the picture, and the opposite side was cathode. The high density of streamlines represents the high flow velocity. A bigger eddy was formed nearby anode, mainly as a result of the upward movement of oxygen bubbles released from anode at electrolyzing. This phenomenon is explained as pumping effects on the flow field, that is, with oxygen bubbles leave to the surface of electrolyte, the space which had been occupied by oxygen bubbles became empty, and the nearby electrolyte would fill into the space under the fluid pressure.

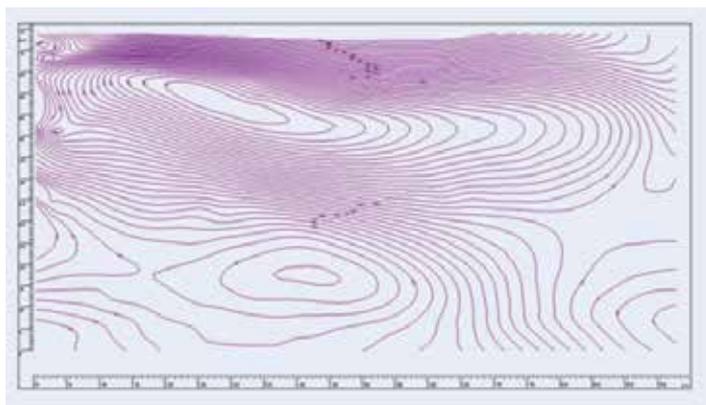


Fig. 3. Streamlines of the flow field in cold model experiment with 0.35 A/cm^2 current density measured by PIV (anode: left, cathode: right)

The flow velocity distribution between anode and cathode was measured by PIV instrument, shown in Figure 4. The target position was at 10% height of anode near the electrolyte surface. The flow velocities increase along the distance from both anode and cathode, respectively, and reach the maximal flow velocity in the middle of one pair of electrodes. Because of the oxygen bubbles continuously upward movement on anode, the position corresponding to the maximal flow velocity was slight close to cathode.

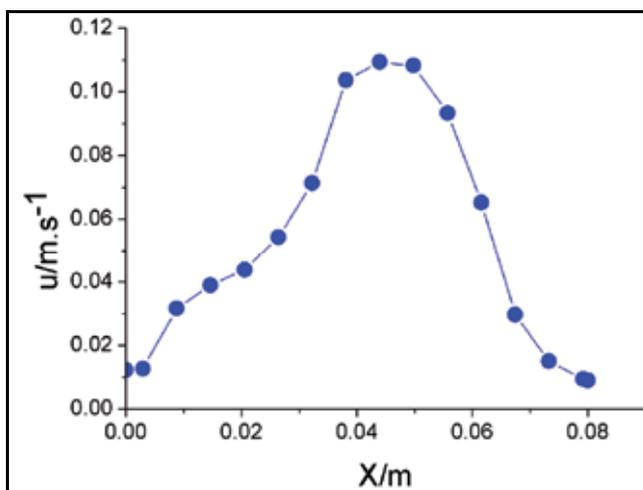


Fig. 4. Flow velocity distribution between experiment at 0.35 A/cm^2 current density. (anode:left, cathode: right)

2.2 Cold model 2

According to geometric similarity, systems are considered to be geometrically similar when the ratio of any length in one system to the corresponding length in the model system is the same everywhere.

In the experiment, the physical model is built according to 1:5 ratio to an original model, as shown in Table 2.

Parameters	Cell parameters for the plant (mm)			Parameter for the physical model (mm)		
	whole size	X=6165	Y= 1870	Z=2065	X=420	Y= 374
anode size	X=150	Y=1000	Z=2020	X=30	Y=200	Z=404
cathode size	X=50	Y=1000	Z=950	X=10	Y=200	Z=190
total cathode number		18			2	
total anode number		16			1	

Table 2. Size comparison between the cell and physical model

A physical model is thought to be dynamic similarity with the original model when the magnitude of forces that cause movement at corresponding location in each model is in a fixed ratio. How to choose dynamic similarity condition is the key to successfully simulate the flow characteristic in the electrolytic magnesium. For aluminum electrolysis cell, both the density differential ratio $((\rho_1 - \rho_2) / (\rho_1 + \rho_2))$ and dynamic interface tension $(T / (\rho_1 + \rho_2))^{[17]}$ were used as dynamic similarity conditions to simulate the aluminum flow characteristic. The modeling phases were chosen according to dynamic similarity conditions of close dynamic surface tension, close density differential ratio and the close kinematic viscosity to simulate the flow characteristic in electrolytic magnesium, the details are shown in Table 3.

Model \ Parameters	Electrolysis cell (700°C)			Physical model cell (20°C)		
	Mg (liquid)	electrolyte	chloride	Silicon oil	water	Chloride
density (g/cm ³)	1.58	1.7	0.000903	0.927	0.998	0.00016
dynamic viscosity (cst)	0.715	1	34	0.715	1	0.133
surface tension (mN/m)	0.547	108	—	0.022	72.88	—
dynamic surface tension $T / (\rho_1 + \rho_2)$	32.9	—	—	37.76	—	—
density differential ratio $(\rho_1 - \rho_2) / (\rho_1 + \rho_2)$	0.0366	—	—	0.0368	—	—

Table 3. Comparison between the cell and physical model

Kinematic similarity means that velocity is kept in the close fixed ratio in the corresponding locations between the physical model system and original electrolytic magnesium. In this cold model, argon gas velocity is kept close to the molar flow rate of chlorine released in anode at anodic current density of 0.314 A/cm². The initial gas velocity in the physical model can be figured out according to the following formula.

$$V_{\text{gas}} = \frac{0.286 \times [80 + 48 \times (29.5 - D)] \times 60}{2 \times 96500} \quad (9)$$

D: distance between anode and the cell bottom; $[80 + 48 \times (29.5 - D)]$: total area of the immersed anode; 96500: electrochemical equivalent

In the electrolytic magnesium system, the relationship between magnesium collection rates and electrolyser structure has not figured out quantitatively because of the quite complicated flow characteristic. The effects of multi-parameters such as anode-cathode distance (ACD), the gradient angle for silicon oil leading channel (θ), depths of electrode (D), width of the service area (W), on recovery rate of magnesium(ϕ) are very significant. The cold model 2 was shown in fig. 5.

Split the anode into 10 chamber, drill the surface of each chamber with the small hole (Φ 0.5 mm), then the gas can release from the anode equable. As the gas, oil also release equable from the small hole in the cathode surface.

After a series of regression experiment in the physical modeling, the results were obtained as follows: ACD: anode-cathode distance; θ : the gradient angle for silicon oil leading channel; D: depths of electrode; W: width of the service area; Φ : recovery rate of oil

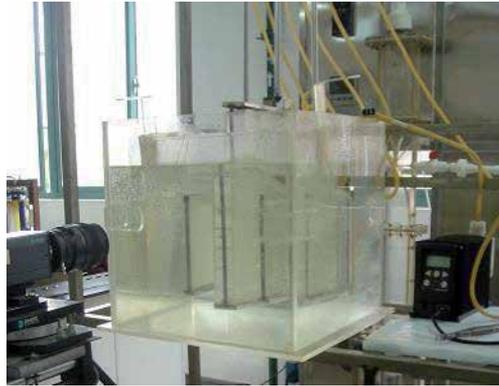


Fig. 5. Cold model of magnesium electrolyzer using silicon oil and argon

By variance analysis, the significance for the recovery rate of silicon oil is in the sequence of θ , ACD and D, and finally W. The recovery rate of silicon can reach the highest under the experimental conditions of 60 mm(ACD), 28 mm(D), 127 mm(W), 12° (θ) according to the experimental data processing.

The relationship between magnesium collection rates and electrolyser structure can be described by the following function formula. In this section , for the very first time we carried out physical model experiments and try to find out the relationship on the magnesium recovery and above variations, and build an quantitative formula between them, the equation as follows:

$$\begin{aligned} \Phi = & -75.6 + 1.42(\text{ACD}) - 0.051D + 1.385W - 0.37\theta + 0.006(\text{ACD}) \times \\ & D - 0.017(\text{ACD}) \times \theta - 0.003D \times W + 0.028W \times \theta - 0.013(\text{ACD})^2 - 0.007W^2 - 0.1\theta^2 \end{aligned} \quad (10)$$

The single parameter experiment was carried out near the best experimental conditions to check the effect of each parameter on the recovery rate of silicon.

No.	ACD	D	W	θ	Φ
1	80	75	120	12	84
2	80	75	80	4	73.5
3	80	35	120	4	68
4	80	35	80	12	78.8
5	40	75	120	4	64.55
6	40	75	80	12	79.7
7	40	35	120	12	84.3
8	40	35	80	4	67.5
9	87.06	55	100	8	80.3
10	27.94	55	100	8	71.6
11	60	82.06	100	8	80.1
12	60	27.94	100	8	89.5
13	60	55	127.06	8	79
14	60	55	72.94	8	80.9
15	60	55	100	13.4	89
16	60	55	100	2.6	75.5
17	60	55	100	8	83.5

Table 4. Regression experiment and result

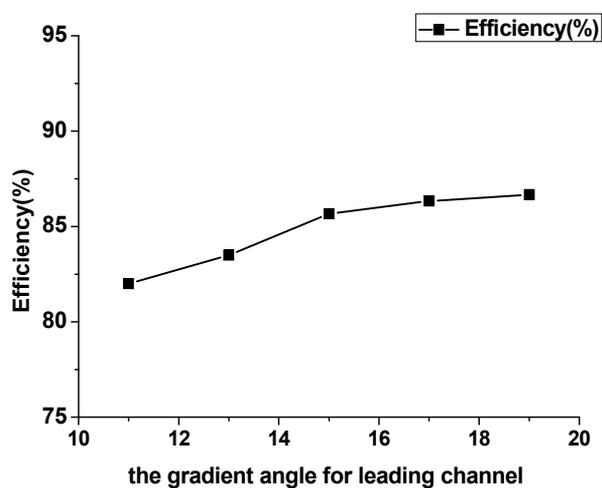
Fig. 6. The relationship between θ and φ

Fig. 6 showed the effects of the gradient angle for silicon oil leading channel (θ) on the recovery rate of oil. The recovery rate of silicon increase simultaneously with the increase of the gradient angle for silicon oil leading channel, and (φ) reach the highest recovery rate at (θ) of 12° and then maintained no change with increase of (θ).

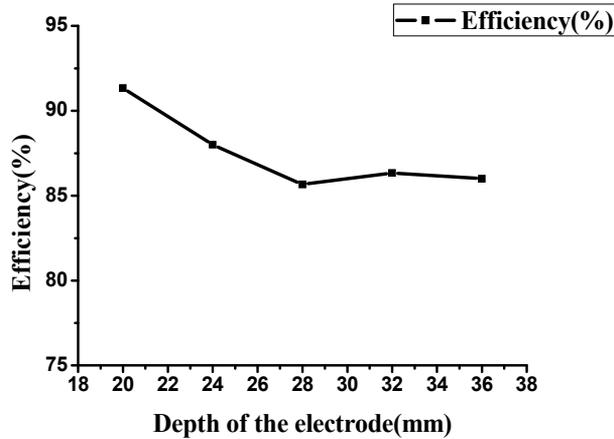


Fig. 7. The relation between D and φ

Fig. 7 gave the effects of depths of electrode (D) on recovery rate of silicon oil φ . It can be seen that the recovery rate increases with the increase of depth of electrode, and reaches the highest recovery rate of silicon oil at depth of electrode as 28mm. the largest recycle rate of silicon oil is 91.34 wt-% under the experimental conditions of 60 mm (ACD), 20 mm (D), 127 mm (W), 12° (θ). When the anode immerses in the electrolyte more deeply, more gas will releases into the electrolyte, which will cause the strong electrolyte circulation in the cell and more oil will be bring into the service area, therefore the recovery rate of silicon oil will be improved.

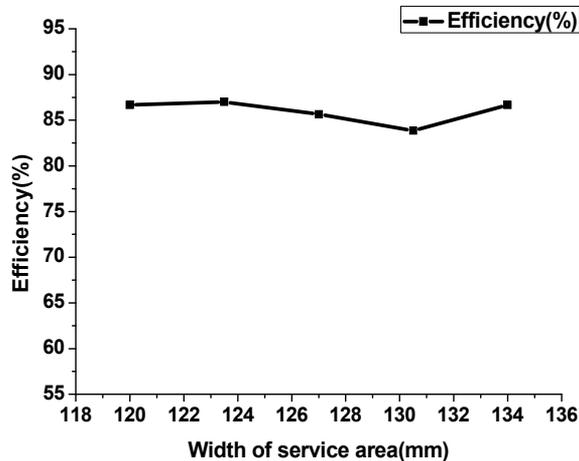


Fig. 8. The relation between W and φ

According to Fig. 8, near the best experiment conditions, the recovery rate of silicon was nearly the same. Fig. 9 displays the relationship between ACD and φ , as the W , near the best experiment conditions, in which the ACD is 6 cm, φ tent to be unrelated to ACD.

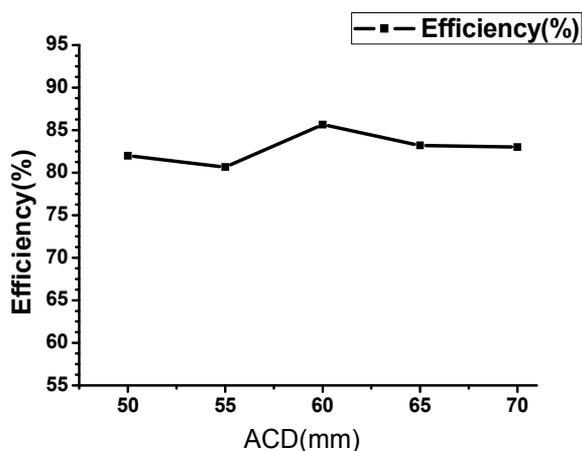


Fig. 9. The relation between ACD and φ

The effects of anodic shape on gas movement along the anode surface was investigated by the PIV, The spray angle for gas from anodic bottom to top maintains at 9° , and does not change with anodic shape and flow rate. The results were shown in fig. 10 and table 5.



Fig. 10. Spray angle for rectangle, positive and negative dovetail anode

shape flow rate	rectangle anode	positive dovetail anode	negative dovetail anode
5 L/min	8.6°	8.8°	8.5°
10 L/min	9°	9°	8.7°
15 L/min	9°	9°	9.2°

Table 5. Value of the spray angle

3. Simulation method

Numerical simulation of three-phase flow field in magnesium electrolyzer involves several subjects including electromagnetism and fluid dynamics. The model leans on the following basic hypotheses^[18]:

1. The electromagnetic fields satisfy the stationary Maxwell's equations, and Ohm's law is moreover supposed to be valid in the electrolyzer.
2. The fluids are immiscible, incompressible and Newtonian.
3. The chlorine gas is as another continuous phase.
The electrical conductivity γ and magnetic permeability λ are constant, respectively.
4. Fluid without chemical reactions.

Electromagnetism is based on Maxwell equations and Ohm's law. The differential form of Maxwell partial differential equations is given as follows:

$$\nabla \times H = J + \frac{\partial D}{\partial t} \quad (11)$$

$$\nabla \times E = -\frac{\partial B}{\partial t} \quad (12)$$

$$\nabla \cdot D = \phi \quad (13)$$

$$\nabla \cdot B = 0 \quad (14)$$

Where H is the magnetic intensity vector, J is the conduction current density, D is the electric displacement vector, E is the electric field intensity vector, B is the magnetic flux density vector and ϕ is the free charge volume density.

The electromagnetic constitutive equations of simple isotropic material mediums can be expressed as:

$$D = \beta E \quad (15)$$

$$B = \lambda H \quad (16)$$

$$J = \gamma E \quad (17)$$

Where $\beta = \beta_r \beta_0$ is the absolute electric constant, $\lambda = \lambda_r \lambda_0$ is the absolute permeability and γ is the electric conductivity.

Fluid dynamics is based Navier-Stokes equations, for multi-phase, volume of fluid (VOF) was an effective method for three-phase flow in magnesium electrolyzer¹². For the q^{th} phase, the equations have the following forms:

Continuity equation

$$\frac{1}{\rho_q} \left[\frac{\partial}{\partial t} (\alpha_q \rho_q) + \nabla \cdot (\alpha_q \rho_q \vec{v}_q) \right] = S_{\alpha q} + \sum_{p=1}^n \left(\dot{m}_{pq} - \dot{m}_{qp} \right) \quad (18)$$

Where α_q was the volume fraction of q^{th} phase in discrete volume; ρ_q was the density of q^{th} phase; \dot{m}_{pq} and \dot{m}_{qp} were mass exchange between q^{th} phase and p^{th} phase, in the electrolyzer, the mass exchange was zero; $S_{\alpha q}$ was the source term, $S_{\alpha q} = 0$.

The volume fraction equation will not be solved for the primary phase; the primary-phase volume fraction will be computed based on the following constraint:

$$\sum_{q=1}^n \alpha_q = 1 \quad (19)$$

Considering the effects of electromagnetic field on molten salt, the Lorentz force was as source term in momentum equation:

$$\frac{\partial}{\partial t}(\rho \vec{v}) + \nabla \cdot (\rho \vec{v} \vec{v}) = -\nabla_p + \nabla \cdot \left[\mu \left(\nabla \vec{v} + \nabla \vec{v}^T \right) \right] + \rho \vec{g} + \vec{F} \quad (20)$$

Where F is the Lorentz force, and $F = J \times B$.

A single momentum equation is solved throughout the domain, and the resulting velocity field is shared among the phases. In order to get the solution of Navier-Stokes equations, turbulence model is essential to close set of equations. Turbulence model divides three types: direct numerical simulation (DNS), large eddy simulation (LES) and Reynolds-Averaged Navier-Stokes (RANS) Models. The most used model among the turbulence models is the Reynolds-Averaged Navier-Stokes (RANS) Models. The k- ϵ model is the most suitable one. In the case of the k- ϵ model, two transport equations are considered; one for the computation of the turbulent kinetic energy and one for the turbulent dissipation rate.

$$\frac{\partial}{\partial t}(\rho k) + \frac{\partial}{\partial x_j}(\rho u_j k) = \frac{\partial}{\partial x_j} \left[\left(\frac{\mu_t}{\sigma_{k0}} \right) \frac{\partial k}{\partial x_j} \right] + G_k - \rho \epsilon \quad (21)$$

$$\frac{\partial}{\partial t}(\rho \epsilon) + \frac{\partial}{\partial x_j}(\rho u_j \epsilon) = \frac{\partial}{\partial x_j} \left[\left(\frac{\mu_t}{\sigma_{\epsilon 0}} \right) \frac{\partial \epsilon}{\partial x_j} \right] + \frac{\epsilon}{k} (C_{\epsilon 1} G_k - C_{\epsilon 2} \rho \epsilon) \quad (22)$$

Where $G_k = \mu_t \frac{\partial u_i}{\partial x_j} \left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j} \right)$, C_μ , $C_{\epsilon 1}$, $C_{\epsilon 2}$, $\sigma_{\epsilon 0}$, σ_{k0} are constants, They are $C_\mu = 0.09$,

$C_{\epsilon 1} = 1.44$, $C_{\epsilon 2} = 1.92$, $\sigma_{\epsilon 0} = 1.3$, $\sigma_{k0} = 1.0$ respectively.

The solution of electromagnetic field can be made by using magnetic vector potential (MVP) method in ANSYS 11.0. The finite elements type SOLID 69 is used to get electric field and SOLID97 to get electromagnetic field. The flow field distribution is obtained by using Fluent 6.3. To solve the governing partial differential equations, Navier-Stokes equations, for the conservation of mass, momentum, and scalars such as turbulence in integral form, a control-volume technique and SIMPLE procedure are used. The momentum source term, Lorentz forces, is read by user defined function (UDF). All simulations are performed on a personal computer equipped with two processors, Intel core 6300 and 2 GB main memory running under the Windows operating system.

In the whole calculation process, coupled fields calculation is the main characteristic. Firstly, in order to get the electromagnetic field distribution, coupling calculation of electric field and electromagnetic field are undertaken, where electric field is as electromagnetic field's input. Secondly, considering effects of electromagnetic field during flow field calculation, flow field coupled electromagnetic field through taking Lorentz forces as source term in momentum equation. The coupled calculation process is shown in Figure 11.

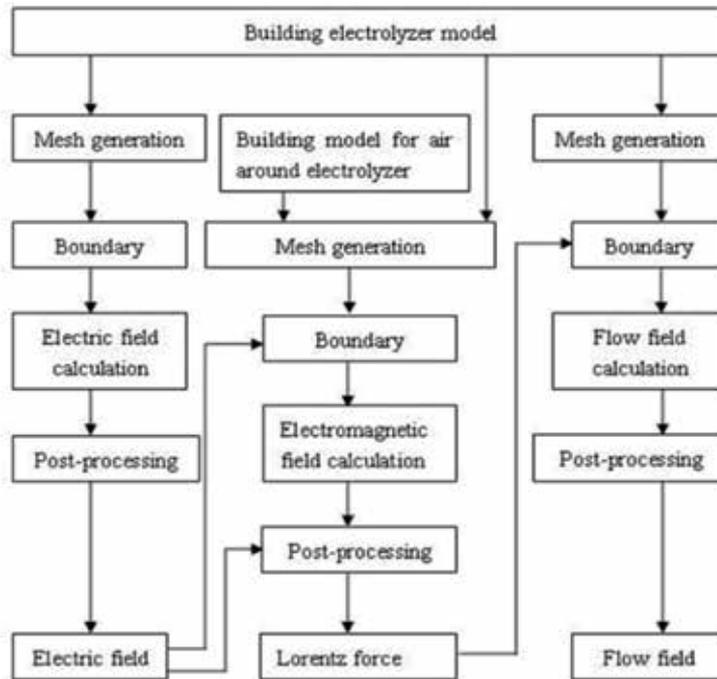


Fig. 11. Calculation procedures for the three-phase flow field under the electromagnetic field

4. Structure of the electrolytic cell and grids for CFD simulation

The advanced diaphragmless magnesium electrolyzer is with the dimension $2.91 \times 1.87 \times 1.40\text{m}$, eight anodes with dimension $0.95 \times 1.14 \times 0.15\text{m}$, nine cathodes with dimension $0.95 \times 1.14 \times 0.05\text{m}$, and the anode and cathode distance as 0.07m . The structure of the electrolytic cell and grids for CFD simulation are shown in Figure 12(a) and Figure 12(b), respectively.

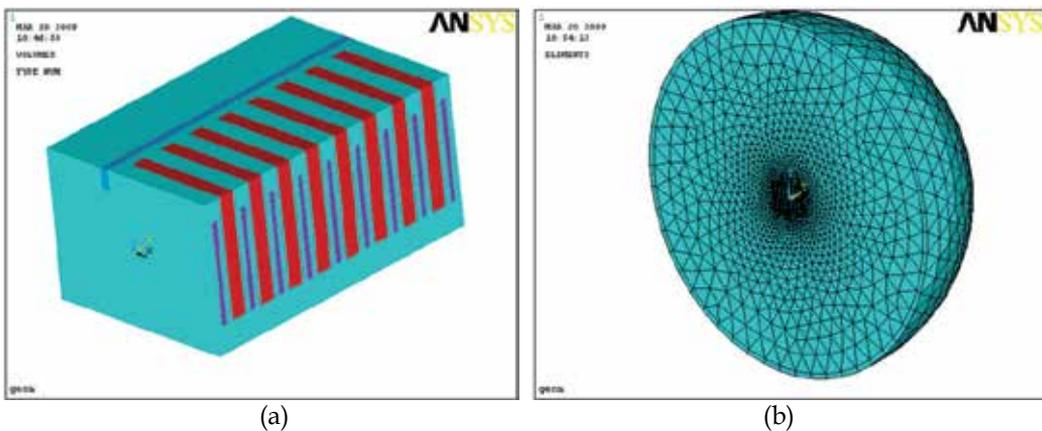


Fig. 12. Structure of the electrolytic cell (a) and grid division for the simulation of the electromagnetic field (b)

During the calculation of the electromagnetic field, the finite element method is employed. According to its symmetry, half of electrolyzer could be used to solve this issue. Therefore, the geometric model for calculation given by Figure 12(a) is the right part of the electrolyzer only, and its mesh partitioning is shown as Figure 12(b). During the calculation of the three-phase flow field, Fluent was employed. The GAMBIT software release 2.2 was used to generate structure and unstructured meshes. The grid in the electrolyzer is showed in Figure 13

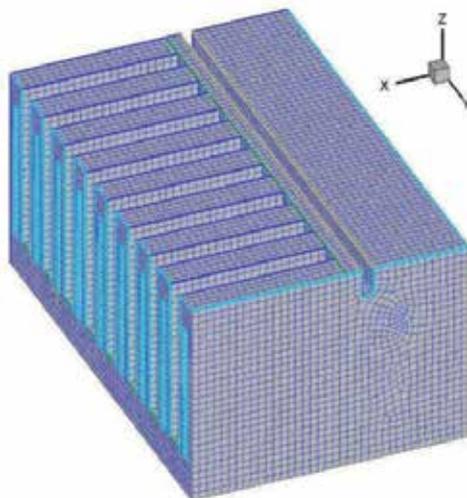


Fig. 13. Grids of the calculation domain for the simulation of the three-phase flow field in electrolyzer

The distribution of the electromagnetic field in the magnesium electrolyzer is calculated using ANSYS 11.0 software, and the properties of materials used for the calculation is shown in table 5. Figure 14 (a) shows the contour of the electric field in the electrolyzer. The voltage energy is concentrated in electrolysis compartment, especially in the space between anode and cathode. From anode to the top of cathode, the value of voltage decreases from maximum to zero. So, in the collection compartment of magnesium, magnesium is not affected by the electric field because of magnesium without charge. This favors the collection of magnesium in collection compartment. Figure 14 (b) shows the contour of voltage on the surface paralleling to work surfaces of electrodes. In the picture, maximum voltage is a constant on the surface of cathode. The value of voltage declines rapidly with distance from the cathode surface.

Material	Material				
	Electrolyte	Graphite	Cast steel	Refractory material	Air
Properties					
Magnetic Permeability	1.0	1.0	1.0	1.0	1.0
Electrical conductivity ($\Omega \cdot m$)	4.0×10^{-3}	8.5×10^{-6}	1.3×10^{-7}	∞	∞

Table 6. Properties of materials used for the calculation of the electromagnetic field

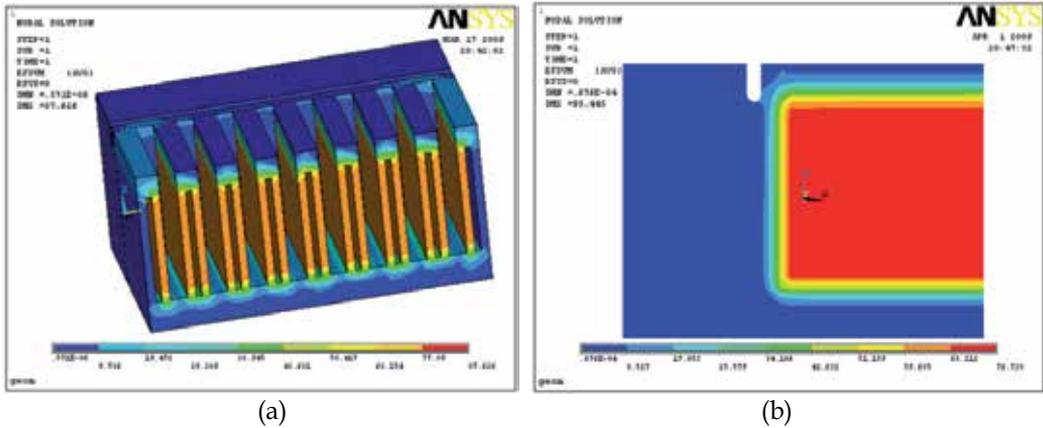


Fig. 14. Contour of the electric field in the magnesium electrolyzer

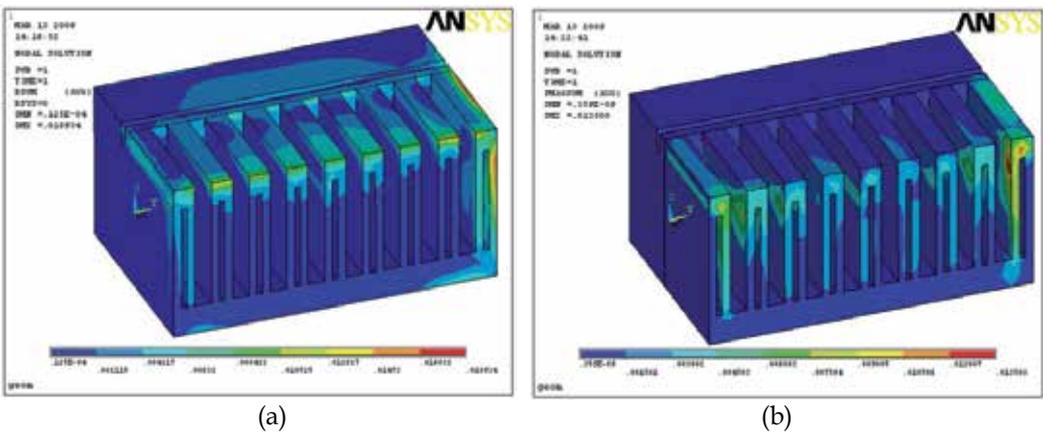


Fig. 15. Contours of the electromagnetic field (a) and contours of Lorentz forces(b) in the magnesium electrolyzer

The simulation results in Figure 15 show that the distributions of both the magnetic Field and Lorentz force are vertically symmetrical along the electrolysis compartment, decreasing gradually from the end toward the middle, and the maximum Lorentz force reaching 0.0135N at corner. The typical distribution of Lorentz forces in a plane vertical to work surfaces of electrodes is shown in figure 16 (b). Figure 16(a) shows the relative planes arrangement. As shown in Figure 16, vectors of Lorentz forces pointed to the collection compartment from electrolysis compartment. These directions of Lorentz forces will favor the circulation of electrolyte. Table 6 lists the typical values of Lorentz Forces calculated by ANSYS 11.0 software, these values will be added to the fluid momentum equations (Eq.20) as a source term, in order to couple the effect of the electromagnetic field on the three-phase flow behavior. This is done by using an in-house UDF code during simulations.

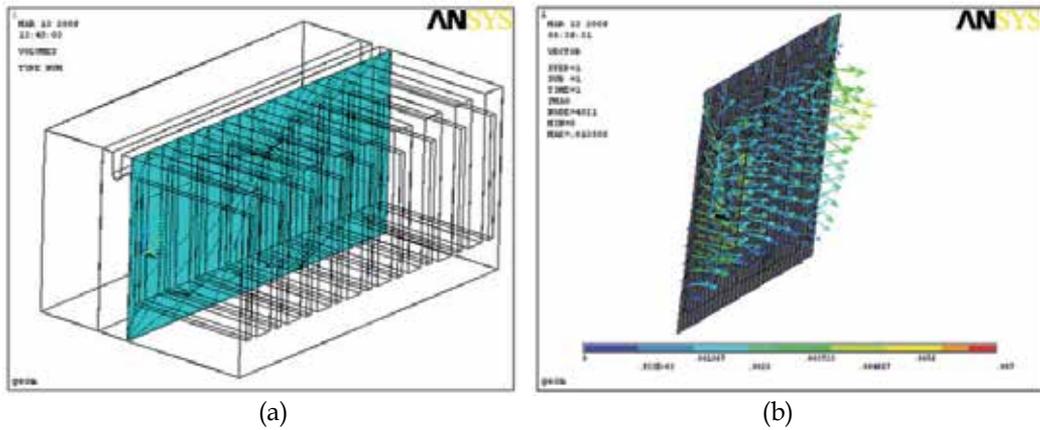


Fig. 16. Location of plane(b) in the electrolytic cell and Vectors of Lorentz Forces(a)

X /m	Y/m	Z/m	Fx/N	Fy/N	Fz/N
-0.900E-01	0.275E+01	0.934E-01	0.206E-02	0.857E-17	-0.124E-02
-0.400E-01	0.275E+01	0.935E-01	0.167E-02	0.693E-17	-0.111E-02
0.100E-01	0.275E+01	0.936E-01	0.117E-02	0.481E-17	-0.815E-03
...
-0.302E+00	0.284E+01	0.520E+00	0.262E-05	-0.234E-05	0.437E-06
-0.302E+00	0.289E+01	0.520E+00	0.263E-05	-0.373E-05	0.115E-05

Table 7. Distribution of Lorentz force in the electrolyzer

Table 7 lists the properties of fluid used in the simulation of the flow field in the electrolyzer. The typical simulation result is shown in Figure 17. Based on the simulation results, there exist three kinds of circulations in the electrolytic cell, circulation A and D as the same kind near walls, circulation B in the middle of electrolyzer and circulation C is the last one.

Fluid \ Properties	Electrolyte	chlorine	magnesium
	Density(kg/cm ³)	1620	0.9
Viscosity (kg/m·s)	1.68×10 ⁻³	2.93×10 ⁻⁵	1.04×10 ⁻³

Table 8. Properties of fluids used in the simulation of three-phase flow field

Though analyzing the flow field obtained, we can see that the whole circulation distributes symmetrically along the electrolyzer. Electrolyte in the pair of the cathode and anode does parallel movement between the surfaces of the electrodes such as circulation B. They are the best circulations in the whole electrolyzer, because the circulation had shortest distance from the electrolysis compartment to the collection compartment. Liquid magnesium can be

delivered effectively to the collection compartment by the circulation B of electrolyte. Other circulation between the cathode and anode deflect to the side wall, instead of parallel to the surface of cathode and anode such as circulation C. They are not parallel to the surfaces of electrodes in collection compartment. They are worse than circulation B in the process of production. The worst circulations are circulation A and D, because they are near the wall. The side wall had effects on circulations, so circulation A and D had to run longer distance in electrolyzer. That would decline the efficiency of current. In a word, the structure optimization of the electrolyzer should be designed to ensure electrolyte circulation parallel to the surfaces of electrodes.

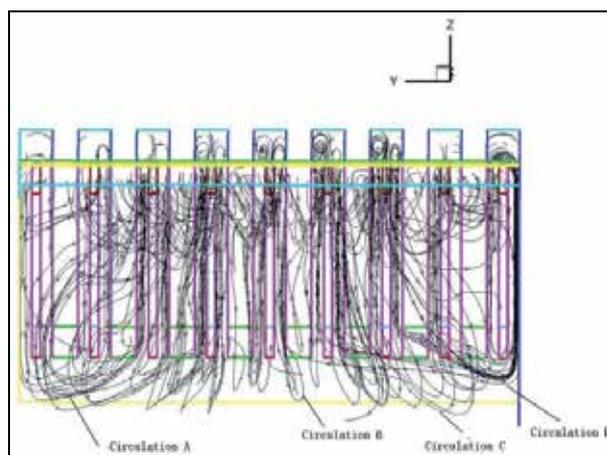


Fig. 17. Streamlines of flow field in the electrolyzer at X direction(in front)

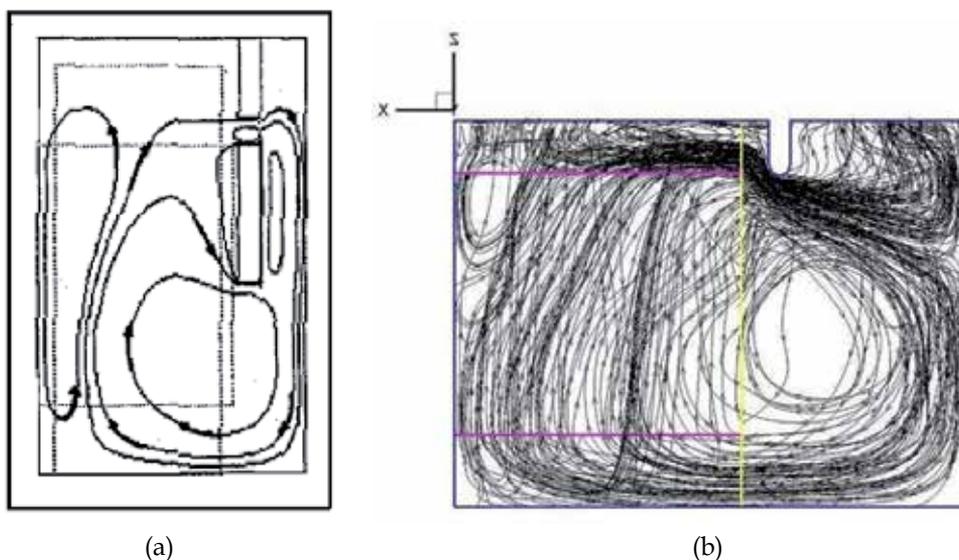


Fig. 18. Comparisons of the simulation results of the electrolyte circulation between Agalakov's results reported in the literature (a) and our work (b)

In Figure 18, from side view, our simulation result(b) has the main characteristics of flow field in electrolyzer and is similar with Agalakov's result(a) [14]. That is shown our math model is correct. Two main circulations appear as shown in Figure 18, one was big and the other was small. During production, the big one can transport magnesium from the electrolysis compartment to the collection compartment. The small one was near back wall, it will damage the back wall, the phenomena can be tested by an industrial electrolyzer, as shown in Figure 19.



Fig. 19. Damaged back wall of an industrial magnesium electrolyzer due to the frequently washing of the small electrolyte circulation

5. Conclusions

1. After variance analysis, the significance for the recovery rate of silicon oil is in the sequence of θ , ACD and D, and finally W.
2. The largest recycle rate of silicon oil is 91.34 wt-%. After the data processing, the best experimental conditions is 60 mm (ACD), 28 mm (D), 127 mm (W), 12° (θ).
3. In the single parameter experiment, when θ is larger than 12° , φ maintain at the largest level. The deeper of the electrode, the larger recovery rate was. When it reaches 28 mm, the recovery rate tend to be changeless. Near the best experiment conditions, φ tent to be unrelated to ACD.
4. The spray angle for gas from anodic bottom to top maintains at 9° , not change with anodic shape and flow rate.
5. Electric field concentrated in the electrolysis compartment; magnetic Field and Lorentz force distributes vertically symmetrical along the electrolysis compartment, decreasing gradually from the end toward the middle.
6. Electrolyte circulations were divided three different types: parallel to working surfaces of electrodes, not parallel to working surfaces of electrodes and near wall's. The circulations paralleling to working surfaces of electrodes can transport liquid magnesium efficiently from electrochemical compartment to service compartment.
7. The research indicates that the structure optimization of the electrolyzer should be designed to ensure electrolyte circulation parallel to the surfaces of electrodes and removing the adverse effects of the electrolyzer flow field.

6. Acknowledgment

We acknowledge the financial support provided by the National Natural Science Foundation of China (Grant 50874048), the National High-Tech R&D Program (Grant 2009AA06Z102), the Fundamental Research Funds for the Central Universities, and Special Foundation for Excellent Young Teacher in East China University of Science and Technology (YB0157115).

7. References

- [1] <http://www.roskill.com/reports/magnesium>
- [2] International Magnesium Association, www.intlmag.org
- [3] Friedrich H. E.; Mordike B. L. *Magnesium Technology*; Springer: Germany, 2006
- [4] Emley E.F. *Principles of Magnesium Technology*; Pergamon: London, 1966
- [5] Sivilotti O. G.; Kingston C. Procedures and apparatus for electrolytic production of metals. 1977, US Patent 4055474
- [6] Andreassen K. A.; Skien Y. B.; Johnsen H. K.; Ognedal L. B.; Solheim P. R. Method and electrolyzer for production of magnesium. 1981, US Patent 4308116
- [7] Holliday R. D.; McIntosh P. Laboratory cell and hydrodynamic model studies of magnesium chloride reduction in low-density electrolytes. *J. Electrochem. Soc.* 1973, 120, 858
- [8] Burnakin V. V.; Shestakov V. M.; Sorokous V. G.; Borutto G. M. Hydrodynamics of flow in the interelectrode space of magnesium electrolyzers. *Tsvetn. Metall.* 1989, (4), 62
- [9] Burnakin, V. V.; Polyakov, P. V.; Shestakov, V. M. Kolesnikov V. A. Analytical calculation of characteristics of two-phase flows in magnesium electrolytic cells. *Tsvetn. Met.* 1979, (8), 73
- [10] Korobov M. A. Mathematical model of a magnesium electrolytic cell. *Tsvetn. Met.* 1983, (5), 53
- [11] Scherbinin S. A.; Yakovleva G. A.; Fazylov A. F. Mathematical model of thermal and electric fields in a magnesium electrolytic cell. *Tsvetn. Met.* 1994, (4), 60
- [12] Shcherbinin S. A.; Yakovleva G. A.; Kazylov A. K. Numerical study of thermal and electric fields of a magnesium electrolytic cell. *Tsvetn. Met.* 1994, (6), 68
- [13] Shcherbinin S. A.; Fazylov A. R.; Yakovleva G. A. Mathematical simulation of three-dimensional thermal and electric fields of magnesium electrolyzer. *Tsvetn. Met.* 1997, (5), 79
- [14] Agalakov V. V.; Shcherbinin S. A.; Yakovleva G. A. Mathematical simulation of gas-hydrodynamic processes in a magnesium electrolyzer. *Tsvetn. Met.* 1997, (7), 74
- [15] Sun Z.; Zhang H. N.; Li P.; Li B.; Lu G. M.; Yu J. G. Modeling and simulation of the flow field in the electrolysis of magnesium". *JOM.* 2009, 61, 29
- [16] Gökhan D.; Karakaya İ. Electrolytic magnesium production and its hydrodynamics by using an Mg-Pb alloy cathode. *J. Alloy. Compd.* 2008, 465, 255
- [17] Dervede E.; Cambridge E.L. Gas induced circulation in an aluminum reduction cell. *light met.* 1975, 111.
- [18] Sun Z.; Zhang H. N.; Li B.; Lu G. M.; Yu J. G. Effect of Electromagnetic Field on Three-Phase Flow Behavior. *Ind. Eng. Chem. Res.* 2010, 49, 10798

- [19] Zhang Y. H.; Yang C.; Mao Z. S. Large eddy simulation of the gas-liquid flow in stirred tank. *AIChE J.* 2008, 54, 1963
- [20] Jain M.; Paranandi M.; Roush D.; Göklen K.; Kelly W. J. Using CFD to understand how flow patterns affect retention of cell-sized particles in a tubular bowl centrifuge. *Ind. Eng. Chem. Res.* 2005, 44, 7876
- [21] Panneerselvam R.; Savithri S.; Surender G. D. CFD simulation of hydrodynamics of gas-liquid-solid fluidised bed reactor. *Chem. Eng. Sci.* 2009, 64, 1119

Fluid-Structure Interaction Techniques for Parachute

Vinod Kumar¹ and Victor Udoewa²

¹*The University of Texas at El Paso*

²*American Association for the Advancement of Science Policy Fellow*
USA

1. Introduction

Fluid-Structure Interaction (FSI) is an inescapable feature of the complicated flow physics where strong coupling between fluid dynamics and structural dynamics occurs. One such example is to understand the behavior of parachutes during canopy inflation and decent. Such simulations can substantially reduce the design costs of parachutes by reducing the number of rather expensive experiments/airdrop tests required. Additionally, FSI simulations can augment experimental approaches by providing detailed fluid flow and structural deformation characteristics of the parachute systems under various scenarios. There are however many computational challenges that one faces in performing the high fidelity FSI simulations. Some of these are turbulence modeling, fluid-structure interaction coupling, the convergence of a nonlinear iteration loop, capturing the physical discontinuity in the flow field, efficiently and accurately solving a large set of linear equations on parallel computers, and improving the parallel performance. Here, we present a high fidelity FSI technique that addresses the coupling issues on High Performance Computing (HPC) environments.

2. Mathematical formulations

FSI modeling of parachutes requires simultaneously solving the Navier-Stokes equations (a set of highly nonlinear Partial Differential Equations (PDEs)) for fluid dynamics, structural mechanics, and mesh motions. Parachutes are made of membrane type structure (usually nylon clothes) that goes through large deformation as a result of aerodynamic forces. This often causes the FSI simulations to break down because of the convergence issues of nonlinear and linear iterative solvers and large mesh stretching/distortions. It is, therefore, difficult to operate and requires in-depth knowledge of numerical techniques, fundamentals of fluid and structure dynamics, coupling behavior, programming languages and environments on High-Performance Computing (HPC) systems. Preparations required before FSI simulations for the parachute-systems are equally challenging and poses tremendous difficulties for novice users.

2.1 Fluid dynamics (FD)

The physics of fluid dynamics is mathematically represented by the Navier-Stokes equations for compressible and incompressible flows. These equations represent conservation of mass

(continuity), momentum, and energy equations. These equations are a set of time dependent non-linear partial differential equations (PDEs).

Incompressible flow: The governing equations of incompressible aerodynamic flows are given by the conservation of mass and momentum equations. The energy equation become decoupled from mass and momentum, therefore we do not need to simultaneously solve the energy equations for incompressible flow system. Most parachute flow lies in the incompressible flow regime.

Let $\Omega_t \subset \mathcal{R}^{n_{sd}}$ be the spatial domain with boundary Γ_t at any instant of time $t \in (0, T)$ where n_{sd} is the number of spatial dimensions (= 4 for 3D incompressible fluid i.e. pressure and three components of velocity vectors) and T is the total time of computations. The spatial coordinates and time are denoted by \mathbf{x} and t , respectively. The Navier-Stokes equations of incompressible flows are:

$$\nabla \cdot \mathbf{u} = 0 \text{ on } \Omega_t \forall t \in (0, T) \quad (1)$$

$$\rho \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} - \mathbf{f} \right) - \nabla \cdot \boldsymbol{\sigma} = \mathbf{0} \text{ on } \Omega_t \forall t \in (0, T) \quad (2)$$

where ρ , \mathbf{u} and \mathbf{f} are the density (constant for incompressible flows), velocity, and the external force, respectively. With \mathbf{I} as the identity matrix, the stress tensor is related to the pressure (p) and shear stress tensor ($\boldsymbol{\tau}$) by

$$\boldsymbol{\sigma}(p, \mathbf{u}) = -p\mathbf{I} + \boldsymbol{\tau} \text{ on } \Omega_t \forall t \in (0, T) \quad (3)$$

The shear stress ($\boldsymbol{\tau}$) and strain rate tensor ($\dot{\boldsymbol{\epsilon}}$) are related by the following constitutive relations

$$\boldsymbol{\tau} = 2\mu\dot{\boldsymbol{\epsilon}} \quad (4)$$

where the viscosity of the air (μ) is constant (Newtonian flow assumptions). In this case, it can be shown that the strain rate tensor is related to velocity by the following relationship

$$\dot{\boldsymbol{\epsilon}}(\mathbf{u}) = \frac{1}{2} \left((\nabla \mathbf{u}) + (\nabla \mathbf{u})^T \right) \quad (5)$$

Compressible flow: The governing equations of compressible aerothermodynamics flows are given by the conservation of mass, momentum, and energy equations. These equations are written as follows:

Continuity or conservation of mass equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (6)$$

Conservation of momentum equations:

$$\frac{\partial(\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u} \mathbf{u}) + \nabla p - \nabla \cdot \mathbf{T} = 0 \quad (7)$$

Conservation of energy equation:

$$\frac{\partial(\rho e)}{\partial t} + \nabla \cdot (\rho e \mathbf{u}) + \nabla(p\mathbf{u}) - \nabla \cdot \mathbf{T} \mathbf{u} + \nabla \cdot \mathbf{q} = 0 \quad (8)$$

Above conservative equations (Eqs 6-8) can be simplified and written in advective-diffusive flux terms as follows:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}_i}{\partial x_i} - \frac{\partial \mathbf{E}_i}{\partial x_i} = \mathbf{0} \quad (9)$$

where \mathbf{U} is a vector containing the primitive variables and is given as

$$\mathbf{U} = (\rho, \rho u_1, \dots, \rho u_{n_{sd}}, \rho e)^T \quad (10)$$

where ρ is density, u_i is the velocities, and e is the total energy per unit mass. The subscript denotes space dimension (i.e., 1 for x-direction, 2 for y-direction and 3-for z-direction, e.g., u_1 is x-component of the velocity) and n_{sd} stands for number of space dimension (i.e., $n_{sd} = 1$ for 1-D problem, $n_{sd} = 2$ for 2-D problem and $n_{sd} = 3$ for 3-D problem. Here, the Euler flux \mathbf{F}_i is given by

$$\mathbf{F}_i = \left(u_i \rho, u_i \rho u_1 + \delta_{i1} p, \dots, u_i \rho u_{n_{sd}} + \delta_{i n_{sd}} p, u_i (\rho e + p) \right)^T \quad (11)$$

where \mathbf{E}_i is viscous flux and is given by

$$\mathbf{E}_i = (0, \tau_{i1} \rho, \dots, \tau_{i n_{sd}}, -q_i + \tau_{ik} u_k)^T \quad (12)$$

Here, p is the mechanical pressure and τ_{ij} is the viscous stress tensor. Eq (9) is further simplified and written in Euler-Jacobian advective and diffusive matrices form as

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{A}_i \frac{\partial \mathbf{U}}{\partial x_i} - \frac{\partial}{\partial x_i} \left(\mathbf{K}_{ij} \frac{\partial \mathbf{U}}{\partial x_j} \right) = \mathbf{0} \quad (13)$$

where Euler-Jacobian advective matrix, \mathbf{A}_i , is given by

$$\mathbf{A}_i = \frac{\partial \mathbf{F}_i}{\partial \mathbf{U}} \quad (14)$$

and the diffusive matrix, \mathbf{K}_{ij} , is given by

$$\mathbf{K}_{ij} \frac{\partial \mathbf{U}}{\partial x_j} = \mathbf{E}_i \quad (15)$$

The advective and diffusive coefficients are not constant and often strongly depend on the local Mach number of the flows, speed of sound, and viscous dissipation and hence are strongly dependent on the solution itself. The coefficients are derived in (Le Beau & Tezduyar, 1991; Kumar, 2005).

The advective and diffusive coefficients are not constant and often strongly depend on the local Mach number of the flows, speed of sound, and viscous dissipation and hence are strongly dependent on the solution itself. The coefficients are derived in (Le Beau & Tezduyar, 1991; Kumar, 2005).

2.2 Structural dynamics (SD)

The deformation of the flexible parachute structure is governed by the equations of large deformation in the structure mechanics under the influence of external driving force arising from dynamic fluid pressure and shear stresses loading from FD. The governing equations for structural dynamics are obtained from the conservation of linear momentum and are given by:

$$\rho^s \left(\frac{d^2 \mathbf{y}}{dt^2} + \eta \frac{d\mathbf{y}}{dt} - \mathbf{f}^s \right) - \nabla \cdot \boldsymbol{\sigma}^s = 0 \quad (16)$$

where ρ^s is the material density, \mathbf{y} is the displacement vector, \mathbf{f}^s is the external force body forces, $\boldsymbol{\sigma}^s$ is the Cauchy stress tensor, and η is the mass proportional damping coefficient. The mass-proportional damping provides additional stability, but can significantly affect the dynamics of the structure. Here, we assume large displacements and rotations, but small strains for nonlinear analysis. The second Piola-Kirchhoff stress (force per unit area in the original configuration) tensor, \mathbf{S} , and the Green-Lagrange strain (in the original configuration) tensor, \mathbf{E} , are used to write the constitutive equations using the total Lagrangian formulation. Thus, stresses are expressed in terms of the 2nd Piola-Kirchhoff stress tensor. The 2nd Piola-Kirchhoff stress tensor is related to Cauchy stress (force per unit area in deformed configuration) tensor, $\boldsymbol{\sigma}^s$, by the kinematic transformation, $\mathbf{s} = \frac{\rho_0}{\rho^s} \mathbf{F} \boldsymbol{\sigma}^s \mathbf{F}^T$ where ρ_0 is the density in the original configuration and \mathbf{F} ($=\mathbf{g}\mathbf{G}$; \mathbf{g} =covariant tensor in deformed configuration, \mathbf{G} =contravariant tensor in original configuration). Firstly, we will assume linear stress-strain relations (Hookean materials) and plane stress conditions. The constitutive equations are given by

$$S^{ij} = \{\bar{\lambda} G^{ij} G^{jk} + \mu (G^{il} G^{jk} + G^{ik} G^{jl})\} E_{kl} \quad (17)$$

where $\bar{\lambda} = \frac{2\lambda\mu}{\lambda+2\mu}$ (λ & μ are the Lamé constants). The Lamé constants are related to the Young's modulus Y and the Poisson's ν ratio by: $\bar{\lambda} = \frac{\nu Y}{(1+\nu)(1-2\nu)}$ and $\mu = \frac{Y}{2(1+\nu)}$. Dirichlet- and Neumann-type BCs are $\mathbf{y}=\mathbf{g}^s$ and $\mathbf{n} \cdot \boldsymbol{\sigma}^s=\mathbf{t}^s$ where \mathbf{g}^s is the specified displacements and \mathbf{t}^s is traction forces (shear stress from the fluid dynamics). The initial conditions are $\mathbf{y} = \mathbf{0}$ & $\frac{d\mathbf{y}}{dt} = \mathbf{0}$.

2.3 Mesh deformation (MD)

Fluid mesh is considered as elastic materials which deforms along with SD deformation. The governing equations mesh deformation is given by

$$\nabla \cdot \boldsymbol{\sigma}^M = 0 \quad (18)$$

3. Finite element discretization

The finite element method is a numerical tool for obtaining solutions to boundary value engineering problems governed by partial differential equations as described in the previous section. It is especially attractive for problems that involve complex geometries where other numerical methods, such as spectral or finite difference methods, are difficult to apply. The principle of the method is to replace an entire continuous domain by a number of sub-domains in which the unknown function is represented by simple interpolation functions with unknown coefficients. Thus, the original boundary-value problem with an infinite number of degrees of freedom is converted into a problem with a finite number of degrees of freedom or, in other words, the solution of the whole system is approximated by a finite number of unknown coefficients. Here, we construct a discretization of a weighted residual formulation in order to arrive at a linear matrix equation. The discretization can be applied in space and time. This formulation is called the space-time finite element method. Alternatively, one can carry out the discretization in space only. Such a method is called a semi-discrete formulation. Details of these methods can be found in a number of references (Tezduyar, Behr, & Liou, 1992; Shakib, 1988; Behr & Tezduyar, 1994).

Galerkin approximation methods are used for the finite element formulations. Galerkin methods are among the most commonly used weighted residual methods. In structural analysis, where often the minimization of energy is the underlying idea, the application of Galerkin methods leads to symmetric matrices and provides *optimal* results. By *optimal* we mean that the solution possesses the best approximation property. The difference between the approximate and the exact solutions is minimized with respect to a certain norm as shown by (Brooks & Hughes, 1982).

The situation, however, is very different in the presence of advective terms. The matrix associated with the advective term is non-symmetric and the best approximation property is lost (Brooks & Hughes, 1982). As a result Galerkin methods applied to these problems are far from optimal and show spurious node-to-node oscillations in the solutions, worsening with growing advection-domination. This not only leads to qualitatively incorrect results but also violates basic physical principles like the second law of thermodynamics (Hirsch, 1988). The pollution of the solution with oscillations is dependent on the domination of the advection terms over other terms of the differential equation. Domination of advection is determined by dimensionless numbers such as Reynolds (ratio of inertial to viscous terms) or Peclet (ratio of advection to diffusion terms). The larger these numbers are, the more dominant is the advection term and the stronger is the numerical node-to-node oscillations in the results. Brooks and Hughes (Brooks & Hughes, 1982) introduced the Streamline-Upwind/Petrov-Galerkin (SUPG) method and this method can be considered as the first successful stabilization technique to prevent oscillations in advection-dominated problems. The SUPG method introduces artificial diffusion in the streamline direction. The introduction of artificial diffusion is done in a consistent way. This can be interpreted as a modification of the test function in the advection direction. Therefore, the weak (variational) form still satisfies the exact solution of the problem.

Another source of potential instabilities in standard Galerkin methods arises when velocity and pressure interpolation functions are not chosen from compatible spaces. Babuska and Brezzi (Babuska, 1973; Brezzi, 1974) showed that compatible spaces must satisfy the *inf-sup* conditions. If they are not chosen from compatible spaces, then oscillatory behavior is observed, primarily in the pressure field. Unfortunately, much desirable function spaces are precluded due to Babuska-Brezzi conditions. Most computationally attractive combinations are the ones which employ equal order interpolations. Hughes et al. proposed a consistent way to circumvent the *inf-sup* condition for the Stokes problem (Hughes, Franca, & Mallet, 1986). As a generalization, Tezduyar, et al., proposed a Pressure-Stabilized/Petrov-Galerkin (PSPG) formulation for finite Reynolds number flows (Tezduyar, Mittal, Ray, & Shih, 1992). The PSPG formulations reduces to the one proposed by Hughes et al. formulation in the limiting case when the Reynolds number tends to zero.

3.1 Deforming spatial domain/stabilized space-time (DSD/SST) formulations for incompressible flows

In order to construct the finite element function spaces for the space-time method, the time interval $(0, T)$ is partitioned into sub-intervals $I_n = (t_n, t_{n+1})$, where t_n and t_{n+1} belong to an ordered sequence of time levels $0 = t_0 < t_1 < \dots < t_N = T$. Let $\Omega_n = \Omega_{t_n}$ and $\Gamma_n = \Gamma_{t_n}$. The space-time slab Q_n is defined as the domain enclosed by the surfaces Ω_n, Ω_{n+1} , and P_n , where P_n is the surface described by the boundary Γ_t as t traverses I_n . As is the

case with Γ_t , surface P_n can be decomposed into $(P_n)_g$ and $(P_n)_h$ with respect to the type of boundary condition (Dirichlet or Neumann) being applied for a degree of freedom of unknown vector \mathbf{d} .

A space-time slab for a 2D spatial domain is schematically shown in Figure 1. The volume of the slab is the volume traversed by the two dimensional plane in time and is given by Q_n . The boundary is denoted by P_n . For a three spatial dimension problem, one can visualize this as a four dimensional space. The spatial domain at $t = t_n$ denoted by Ω_n and at $t = t_{n+1}$ by Ω_{n+1} . In the beginning of the computations at $t = 0$, the spatial domain and its boundary is denoted by Ω_0 and Γ_0 respectively. Defining spatial domain in this way allows one to handle moving and deforming bodies.

For each space-time slab, we define the following finite element interpolation function spaces for the conservation variables

$$(\mathcal{S}_u^h)_n = \{u^h | u^h \in [H^{1h}(Q_n)^{n_{sd}}], u^h \cong g_u^h \text{ on } (P_n)_g, d = 1, \dots, n_{sd}\} \tag{19}$$

$$(\mathcal{V}_u^h)_n = \{w^h | w \in [H^{1h}(Q_n)^{n_{sd}}], w^h \cong 0 \text{ on } (P_n)_g, d = 1, \dots, n_{sd}\} \tag{20}$$

$$(\mathcal{S}_p^h)_n = (\mathcal{V}_p^h)_n = \{q^h | q^h \in [H^{1h}(Q_n)^{n_{sd}}], w^h \cong 0 \text{ on } (P_n)_g\} \tag{21}$$

where $H^{1h}(Q_n)$ is the n -dimensional function space over the space-time slab Q_n . Over the element domain, this space is formed by using first-order polynomials in both space and time. The interpolation functions are continuous in space but discontinuous in time. The stabilized space-time (SST) formulation of momentum balance and continuity equations for deforming spatial domains (DSD) can be written as follows.

Given $(u^h)_n^-$, find $u^h \in (\mathcal{S}_u^h)_n$ and $p^h \in (\mathcal{S}_p^h)_n$ such that $\forall w^h \in (\mathcal{V}_u^h)_n$ and $q^h \in (\mathcal{V}_p^h)_n$

$$\begin{aligned} & \int_{Q_n} \mathbf{w}^h \cdot \rho \left(\frac{\partial \mathbf{u}^h}{\partial t} + \mathbf{u}^h \cdot \nabla \mathbf{u}^h + \mathbf{f}^h \right) dQ + \int_{Q_n} \hat{\mathbf{e}}(\mathbf{w}^h) : \boldsymbol{\sigma}(p^h, u^h) dQ + \int_{\Omega_n} q^h \nabla \cdot \mathbf{u}^h dQ + \\ & + \sum_{e=1}^{(n_{el})_n} \int_{Q_n^e} \tau \left[\rho \left(\frac{\partial \mathbf{w}^h}{\partial t} + \mathbf{w}^h \cdot \nabla \mathbf{w}^h + \mathbf{f}^h \right) - \nabla \cdot \boldsymbol{\sigma}(q^h, w^h) \right] \cdot \left[\rho \left(\frac{\partial \mathbf{u}^h}{\partial t} + \mathbf{u}^h \cdot \nabla \mathbf{u}^h + \mathbf{f}^h \right) - \nabla \cdot \boldsymbol{\sigma}(p^h, u^h) \right] dQ + \tag{22} \\ & + \sum_{e=1}^{(n_{el})_n} \int_{Q_n^e} \delta \nabla \cdot \mathbf{w}^h \rho \nabla \cdot \mathbf{u}^h dQ + \int_{\Omega_n} (\mathbf{w}^h)_n^+ \cdot \rho \left((\mathbf{u}^h)_n^+ + (\mathbf{u}^h)_n^- \right) d\Omega = \int_{(P_n)_h} \mathbf{w}^h \cdot \mathbf{h}^h dP \end{aligned}$$

This process is applied sequentially to all the space-time slabs Q_1, Q_2, \dots, Q_{N-1} . At the start of the computations, we assume $(u)_0^+ = u_0$. In the equation (22), the first three terms, the sixth term, and the right-hand-side comprise the Galerkin formulation of the problem.

The first element-level integrals (terms under the summation $\sum_{e=1}^{(n_{el})_n}$) in equation are least-squares terms based on the momentum equation. The second element-level integrals are added to the formulation for numerical stability at high Reynolds numbers. These least-squares terms are based on the continuity equation. The stabilization coefficients, τ and δ , are defined at the element level. Both stabilization terms are weighted residuals, and therefore maintain the consistency of the formulation. Since the interpolation functions are discontinuous in time, the sixth term weakly enforces continuity of the velocity field across the space-time slabs.

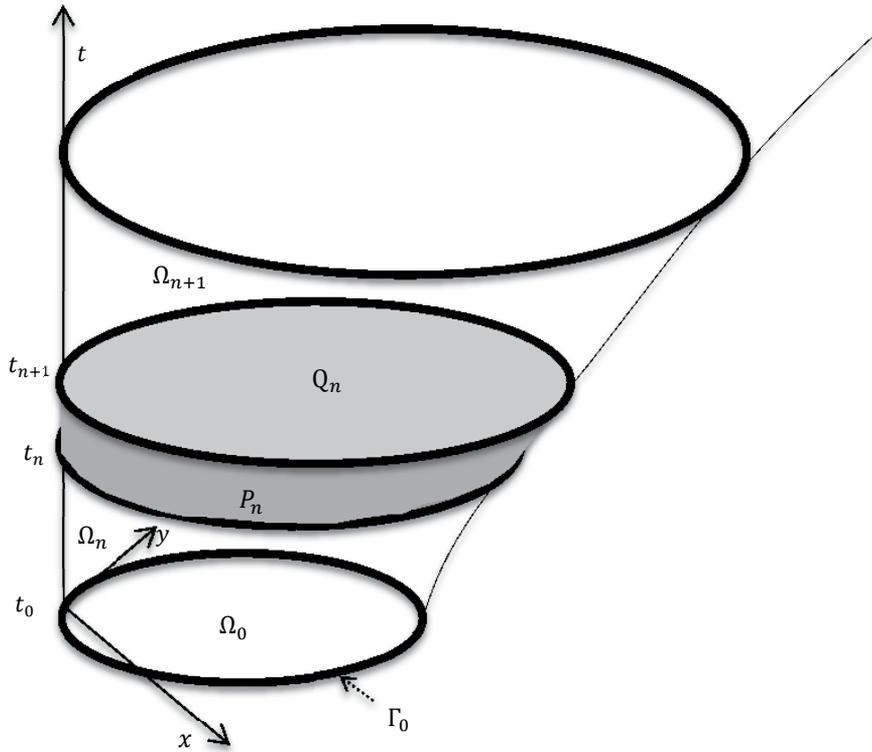


Fig. 1. Space time slab for deforming spatial domain

The stabilization coefficients, τ and δ , are given by

$$\tau = \left[\left(\frac{z}{\Delta t} \right)^2 + \left(\frac{2|u^h|}{h_e} \right)^2 + \left(\frac{4\nu}{h_e^2} \right)^2 \right]^{-\frac{1}{2}} \tag{23}$$

$$\delta = \frac{h_e}{2} |u^h| z \tag{24}$$

where ν is the kinematic viscosity, Δt is the time step, and h_e is a suitable measure of element length. One obvious choice for element length is the maximum edge length. This, however, gives rise to excessive numerical diffusion when the flow direction is not aligned along the maximum edge. A directional element length based on advection direction works better in this case. The parameter z is defined as

$$z = \begin{cases} \left(\frac{Re_u}{3} \right) & \text{if } Re_u \leq 3 \\ 1 & \text{if } Re_u > 3 \end{cases} \tag{25}$$

where $Re_u = \frac{|u^h| h_e}{2\nu}$ is the element Reynolds number.

The DSD/SST formulation compressible equations are written in the similar fashion. A detailed discussion of these stabilization terms and their origin can be found in (Mittal, 1992).

4. Coupling techniques

4.1 Multiphysics coupling techniques

Let us symbolically write the nonlinear system of equations arising from finite element discretization for the partial differential equations (PDEs) of a three physical systems, for example, *Physics 1*, *Physics 2*, and *Physics 3*, as a nonlinear set of equations denoted by scalar functions

$$\mathbf{N}_1(\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3) = 0 \quad (26)$$

$$\mathbf{N}_2(\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3) = 0 \quad (27)$$

$$\mathbf{N}_3(\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3) = 0 \quad (28)$$

where the subscript "1", "2", and "3" represent *Physics 1*, *Physics 2*, and *Physics 3* respectively (e.g., SD, MD, and FD) for our case. This representation can be extended to cases with more than three multiphysics processes. Here, \mathbf{d}_1 , \mathbf{d}_2 , and \mathbf{d}_3 are the vectors of unknown variables for *physics 1* (e.g., deformed coordinate systems and velocity), *physics 2* (e.g., deformed mesh coordinate systems), and *physics 3* (e.g., velocity, pressure, and density for FD). The coupling among different physics arises from boundary conditions at the interfaces of various systems.

Linearization of the nonlinear set of equations through first order Newton-Raphson iterative approximation results in

$$\left[\frac{\partial \mathbf{N}_1}{\partial \mathbf{d}_1}\right]^i \Delta \mathbf{d}_1^{i+1} + \left[\frac{\partial \mathbf{N}_1}{\partial \mathbf{d}_2}\right]^i \Delta \mathbf{d}_2^{i+1} + \left[\frac{\partial \mathbf{N}_1}{\partial \mathbf{d}_3}\right]^i \Delta \mathbf{d}_3^{i+1} = -[\mathbf{N}_1(\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3)]^i \quad (29)$$

$$\left[\frac{\partial \mathbf{N}_2}{\partial \mathbf{d}_1}\right]^i \Delta \mathbf{d}_1^{i+1} + \left[\frac{\partial \mathbf{N}_2}{\partial \mathbf{d}_2}\right]^i \Delta \mathbf{d}_2^{i+1} + \left[\frac{\partial \mathbf{N}_2}{\partial \mathbf{d}_3}\right]^i \Delta \mathbf{d}_3^{i+1} = -[\mathbf{N}_2(\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3)]^i \quad (30)$$

$$\left[\frac{\partial \mathbf{N}_3}{\partial \mathbf{d}_1}\right]^i \Delta \mathbf{d}_1^{i+1} + \left[\frac{\partial \mathbf{N}_3}{\partial \mathbf{d}_2}\right]^i \Delta \mathbf{d}_2^{i+1} + \left[\frac{\partial \mathbf{N}_3}{\partial \mathbf{d}_3}\right]^i \Delta \mathbf{d}_3^{i+1} = -[\mathbf{N}_3(\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3)]^i \quad (31)$$

where superscript i is the nonlinear iteration counter for the coupled systems. This linearized system of equations can be written in simplified matrix notation as

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33} \end{bmatrix}^i \begin{bmatrix} \Delta \mathbf{d}_1 \\ \Delta \mathbf{d}_2 \\ \Delta \mathbf{d}_3 \end{bmatrix}^{i+1} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \end{bmatrix}^i \quad (32)$$

where the Jacobians of the iterative solver are given by

$$\mathbf{A}_{jk} = \frac{\partial \mathbf{N}_j}{\partial \mathbf{d}_k} \quad (33)$$

and the right hand side vector is given by

$$\mathbf{b}_j = -[\mathbf{N}_j(\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3)] \quad (34)$$

These systems of equations represent a fully coupled multiphysics (three physics in this case) system. For most real life applications, the cross-Jacobians terms (e.g., \mathbf{A}_{12} & \mathbf{A}_{13}) that represents the coupling process (e.g., the instantaneous feedback that the structure gets from

the fluid flow for any small deformation in the structure) and are usually mathematically not well defined by partial differential equations.

Most commercial/opensource software treat fully coupled system through nonlinear iterations but assume one directional coupling inside the nonlinear iterations. This results in a diagonal Jacobian matrices (symbolizing one way coupling) and can be written by setting off diagonal terms equal to zeros as

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{0} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33} \end{bmatrix}^i \begin{bmatrix} \Delta \mathbf{d}_1 \\ \Delta \mathbf{d}_2 \\ \Delta \mathbf{d}_3 \end{bmatrix}^{i+1} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \end{bmatrix}^i \quad (35)$$

In case of FSI modeling, this implies that SD (*Physics 1*) is solved first using a forcing terms (shear and pressure coming out from FD i.e., *Physics 3*), then MD (*Physics 2*) is solved using deformed coordinate systems at the structure boundaries of the FD systems, and at the end FD (*Physics 3*) is solved using the deformed mesh from MD and velocity boundary conditions at the interfaces. In this, the coupling terms (\mathbf{A}_{21} , \mathbf{A}_{31} , and \mathbf{A}_{32}) are computed through Least-Square projections for incompatible meshes or mapping for compatible meshes.

One way coupling however fails if the coupling between structure and fluid is strong such parachute – aerodynamics systems. Other challenges are addressing the multiscale nature of such problems (e.g., temporal and spatial scales of structure dynamics usually occur at a lot smaller scales than the fluid dynamics). New advancement in computational technologies such multi-grid preconditioning for multi-physics/multi-scale advance linear solver on massively parallel computers come help address some of these challenges. Numerical convergence of the system can also be enhanced by advanced solvers. Discussions about these are beyond the scope of this chapter.

4.2 FSI coupling techniques for parachutes

As described in the previous sections, the FSI modeling of parachute involves solving the time dependent fluid dynamics (FD) equations together with structural dynamics (SD) and mesh deformation (MD) equations as described in the previous section. Coupling between structure and fluid is achieved by exchange of information (i.e., pressure p , velocity v , surface coordinates, and temperature) from the aerodynamics simulations to Interface (through mapping) and from the Interface to structural dynamics. We employ least-square projections if the meshes between two segments are incompatible. Here, we assume that the interface has the same mesh morphology as the fluid dynamics (FD) mesh on the object surface but it differs from the SD meshes (usually requires higher order finite element meshes to resolve the surface deformations). Fluid dynamics is governed by 3D gas dynamics equations and structural dynamics is modeled as 3D constitutive equations for solid mechanics. Coupling is achieved in block iterative fashion as shown in the figure 6. In block iterative method, first SD is solved using finite element methods for solid mechanics. Then, new mesh coordinates are determined by solving the elastodynamics equations for mesh deformation. Note that fluid meshes have no physical physics, so Young's modulus of the mesh elements is adjusted in such way those mesh-motion results in minimal element distortions. After mesh motion, FD is solved using the new coordinates in the fluid domain and velocity boundary conditions and the surface.

The whole process is repeated inside a non-linear Newton-Raphson iterations block (as shown in Figure 2) until a desired convergence or maximum number of iterations (set by user) is achieved.

For compressible flow, we compute pressure using the ideal gas conditions for air from the conservative (primitive) variables (i.e., $\rho, \rho \mathbf{u}, \rho e$) which are computed by solving fluid dynamics equations.

$$p = \rho R \theta = (\gamma - 1) \rho i = (\gamma - 1) \left[\rho e - \frac{|\rho \mathbf{u}|^2}{2\rho} \right] \quad (36)$$

where R is the ideal gas constant, i is the internal energy, θ is the temperature, and is the ratio of specific heat (usually = 1.4 for mono atomic gases). The boundary conditions for uid velocity on the structure interface boundaries are given by

$$[U_1, U_2, U_3] = [\rho u_1, \rho u_2, \rho u_3]_{interface} \quad (37)$$

4.3 Mesh generation and update

Proper selection of the appropriate mesh update approach depends on several factors. These include the complexity of the moving boundary (or interface) as well as the overall geometry, the unsteadiness of the moving boundary (or interface), and the qualities of the starting mesh. In general, the mesh update should have two components: moving the mesh for as long as it is possible, and full or partial remeshing (i.e. generating a new set of elements, and sometimes also a new set of nodes) when element distortion becomes too high.

In mesh moving strategies, our only requirement for the mesh motion is that at the moving boundary (or interface) the normal velocity of the mesh has to match the normal velocity of the uid. Beyond that, the mesh can be moved in any way desired, with the main objective being to reduce the frequency of remeshing. In 3D simulations that rely on an automatic mesh generator, the cost of automatic mesh generation becomes a major reason for trying to reduce the frequency of remeshing. Additionally, mesh generation is often done on a single CPU, losing the parallel performance. Furthermore, when we remesh we need to project the solution from the old mesh onto the new one. This step introduces projection errors along with an additional computational cost that is not trivial for 3D computations involving complex geometries. Projection errors can destroy the divergence free condition in the domain giving rise to pressure oscillations (Udoewa, 2005). All of these factors provide a strong motivation for utilizing mesh update strategies which minimize the frequency of remeshing.

Parachute simulations involve complex geometries and arbitrary motions for which it is difficult to design special purpose mesh moving techniques. For parachute clusters, the behavior of parachute interactions becomes even more erratic. For these problems I use an automatic mesh moving scheme (Johnson A. , 1995; Stein, Simulation and Modeling Techniques for Parachute Fluid-Structure Interactions, 1999; Udoewa, 2005) to move the nodal points, as governed by the equation of linear elasticity, and where the smaller elements enjoy more protection from mesh deformation. The motion of the internal nodes is determined by solving these additional equations. The boundary conditions for the mesh motion are specified in such a way that they ensure the matching of the normal velocity of the fluid at the interface.

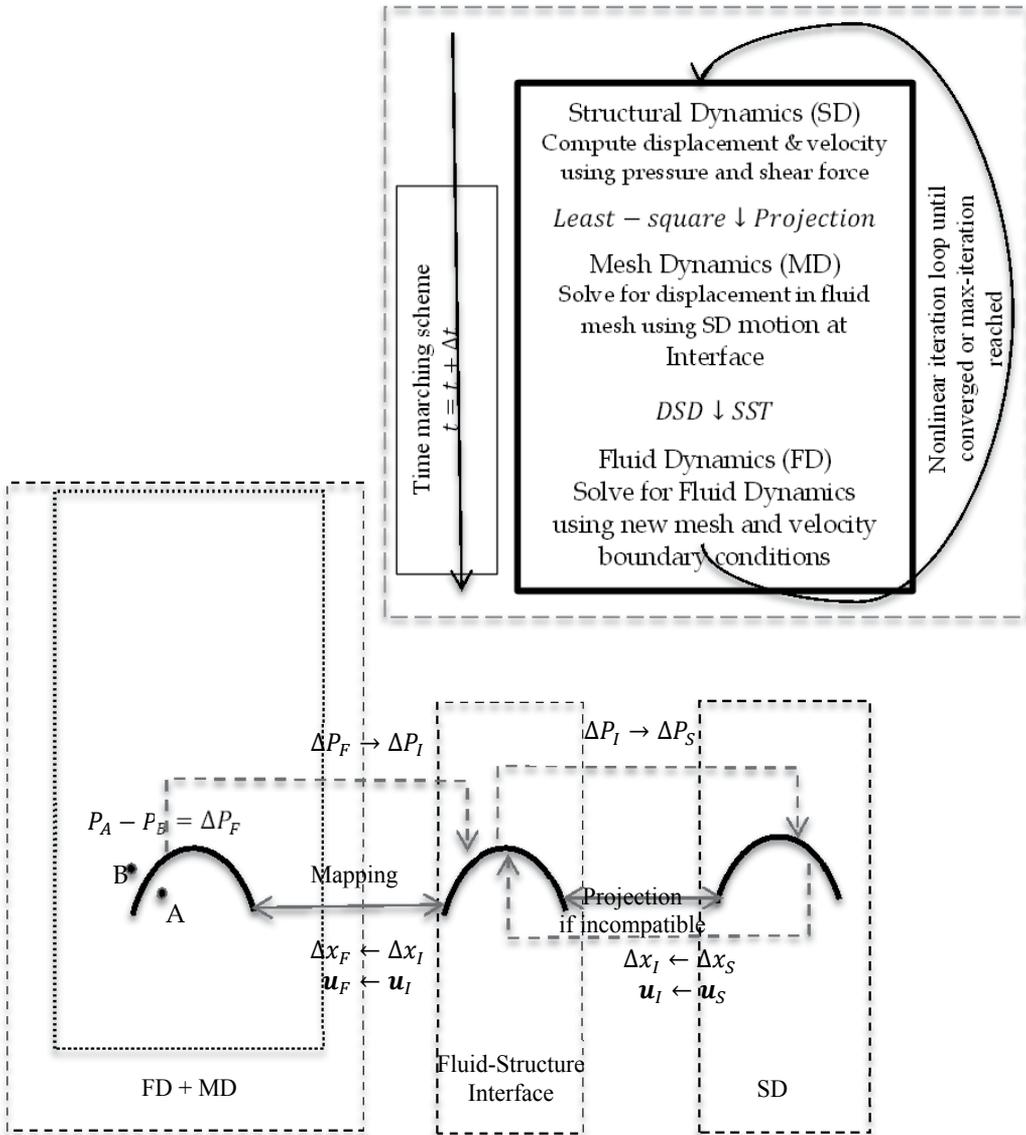


Fig. 2. Block iterative coupling technique (top-right) and information exchange processes to address the coupled multiphysics problems

4.4 Incompatible meshes

Incompatible meshes [40] for structure and interface surface are used. This allows one to choose different types (e.g. higher order bi-quadratic elements for parachute canopy and triangular mesh for fluid-structure interface) and different refinements. For parachute discretization, we choose the same type of triangular elements for both parachute canopy

and fluid-structure interface meshes but the canopy mesh for the structural dynamics was very refined to resolve the spatial scales of the gores of the canopy. The same level of mesh refinement for the fluid-structure interface mesh (in case compatible case) was not suitable due to resulting large number elements of the fluid meshes and computationally expensive Navier-Stokes equation solver. Additionally, the length scale associated with fluid dynamics process is smaller than the scale required for structural dynamics to achieve the same level of fidelity. Level of refinement of the meshes were decided based on extensive parachute modeling experiences and should be decided on case-by-case basis. This resulted in incompatible meshes and consequently a mechanism to transfer the deformed coordinate and velocity information from SD to FD and pressure and shear stresses from FD to SD.

Whether one uses different types of elements or different levels of refinements, a projection mechanism is needed to transfer the information from fluid to structure and vice-versa. Errors (usually called projection errors) come along with the projection. Some of the projection techniques are the linear projection and the least-square projection. We used the least-square projection technique to achieve this goal. The least-square projection technique minimizes the sum of residuals in the entire domain. The least-square projection is achieved by solving the following minimization problem

$$\min \int_{\Gamma^I} \|\mathbf{d}_S - \mathbf{d}_F\|^2 d\Gamma \quad (38)$$

where \mathbf{d}_F (the fluid variables, e.g., pressure drop across the canopy) is projected to \mathbf{d}_S (the structure variables, e.g., pressure applied on the canopies). Integration is carried over the interface domain Γ^I . Similar techniques are used to project the velocity and displacement from structure to the fluid-interface.

4.5 Pressure ramping

For some cases, the ramping of pressure was absolutely necessary to obtain converged solutions. It was desirable especially in those cases where the guess for the initial shape of the parachute canopies was not close to the real life parachute geometry. The convergence of nonlinear interactions at the start of the FSI simulations was very erratic in the absence of a ramping technique. Pressure ramping was successfully used in these cases. Here, the information exchange between structure and fluid is linearly ramped. The ramping is carried out such that the acting pressure on the canopy is equal to the guess pressure (used for inflating the parachute) at the start of the FSI simulations and equal to the fluid pressure at the end of the ramping. This is achieved by

$$\Delta P_S \leftarrow r(t)\Delta P_G + (1 - r(t))\Delta P_F \quad (39)$$

where ΔP_S is the effective pressure applied on the structure, ΔP_G is the guess pressure used to get the inflated shape of the parachute, and ΔP_F is the pressure drop across the canopies coming from fluid simulations. The ramping factor $r(t)$ is assumed to be linear in time and varies from 0 to 1 in time t_{ramp} where t_{ramp} is the ramping time. Ramping time were chosen case-by-case to achieve appropriate convergence of the Newton-Raphson iterations for the coupled non-linear FSI system.

5. FSI Simulations of parachutes

The parachute canopy is made of very flexible fabric materials. It responds quickly to any small changes in fluid behavior. Consequently, there is a strong interaction between the fluid and structure. Bigger the canopies, stronger will be the interactions. Stronger interactions make the problem very nonlinear in nature and hence increase the difficulty of numerical modeling.

5.1 Cluster of two G-12 parachutes

A single G-12 parachute is used to drop cargo weighing up to 2,200 lb. There is an interest in the Army for parachute systems that can drop cargos exceeding 2,200 lb. In order to achieve this objective, multiple parachutes in cluster configurations are often used.

The opening stage of a parachute to a fully inflated shape is a critical issue. Usually more than one stage is used to get the fully inflated shape of parachutes in a cluster configuration. One example of multi-stage opening is the use of drogue parachutes. First, a drogue parachute opens and then it pulls the other two G-12 parachutes. After all the parachutes are fully opened, the challenge is to safely drop the cargo into the hands of the right people. The strong parachute-to-parachute aerodynamic interaction can potentially destroy the efficiency of the system if it is not carefully designed. Various researchers have analyzed the cluster of parachute systems using both drop tests and computer simulations. Butler (Butler, 2001) presented the qualitative results from the airdrop tests using a drogue parachute to deploy the G-12 canopies. From the snapshots of a cluster of three G-12 canopy airdrop tests, one clearly observes strong interactions between fluid and canopies. Sahu, et al., (Sahu & Benney, 1997) carried out the numerical simulations using a quasi-static approach by imposing a symmetry condition for the three canopies in an attempt to predict the equilibrium configuration of a cluster of three half-scaled C-9 parachutes. Stein et al. presented results from the semi-discrete simulations for the aerodynamic interactions between the canopies of parachute clusters of varying numbers and arrangements. In these computations, the canopies were assumed to be rigid. In real life, however, the canopy of a parachute deforms in accordance with aerodynamic forces. The placement of parachutes in cluster configuration does not remain stationary. Therefore, FSI simulations are needed to replicate the real life scenario and to understand the dynamics of cluster of parachutes. Here, we present results from the FSI simulations of a cluster of two G-12 parachutes. The results from a cluster of three G-12 parachutes will be discussed in the next section. The DSD/SST formulations as discussed in the section 0, along with appropriate mesh-update strategies, allowed me to study the interaction of canopies in a cluster.

Problem setup:

The G-12 is a 64 ft diameter cargo parachute designed to deliver a payload of 2200 lb at a descent speed of approximately 28 ft/s. Clusters of G-12 parachutes are commonly used to deliver larger payloads. The G-12 is constructed with 64 suspension lines which extend to risers. For a single G-12 parachute, the confluence point of the risers is connected to a retraction cable which supports the payload. For a cluster of two G-12 parachutes, the retraction cable is connected to two cables, which connect to the confluence point of the risers for the two G-12 parachutes.

The structural model is composed of membranes, cables, and concentrated masses. The canopy is modeled with triangular membrane elements. Linear cable elements are used to model the suspension lines, radial reinforcements along the canopy, risers, and payload support cables. In each example, the payload is modeled with a single concentrated mass. Material properties are selected to be representative of the G-12. A model for the cluster of G-12 parachutes in constructed (unstressed) and inflated (prestressed) configurations is shown in Figure 3.

Several preparations are required for each fluid-structure interaction simulation. First, a stand-alone structural deformation simulation is carried out to determine the inflated (i.e. prestressed) shape of the G-12. The initial inflation pressure is assumed to be equal to stagnation pressure. From my experience, we observed that this pressure gives a better approximation for the initial shape of the parachute canopies. The prestressed configuration for the G-12 cluster is shown in Figure 3 (right). The unstressed configuration is shown in Figure 3 (left).

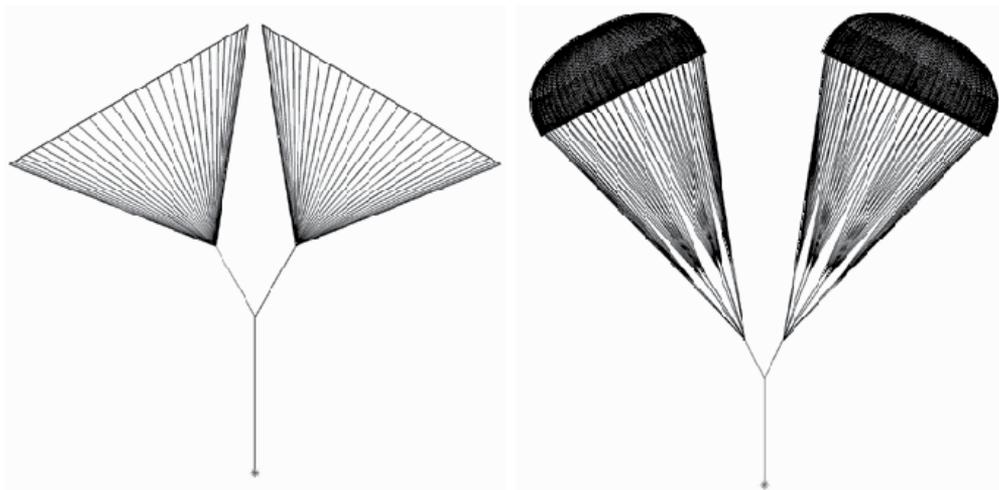


Fig. 3. Cluster of two G-12 parachutes: Constructed (left) and prestressed (right) structural model.

Using the parachute canopy from the prestressed configuration, a fluid mesh is generated and a stand-alone fluid simulation is carried out to obtain a developed flow. Fluid simulations are expensive and time consuming. Therefore, to arrive at developed flow quickly, at first semi-discrete (Johnson A. , 1995; Mittal, 1992) formulation is used to compute the flow field. Semi-discrete formulations are first order in time. Using this solution as the initial condition, space-time (Tezduyar T. E., Stabilized finite element formulations for incompressible flow computations, 1992) computations are carried out. We

start the simulations with a first order-time accurate integration scheme. This helps to clear the start-up vortices quickly. Startup vortices are generated due to large differences in the initial conditions and the exact solutions. After the start-up vortices are cleared from the domain, the second order accurate time-integration scheme is applied. Now, the stage is set to start fluid-structure interaction simulations. Using the results from space-time simulations, a fluid-structure interaction computation is carried out. To remove the mismatch in the initial guessed prestressed configuration, we use the pressure ramping to soften the exchange of information.

Pinned payload:

The pinned payload case corresponds to the wind tunnel testing where parachutes are fixed to a confluence point. The payload is pinned and is not allowed to move in any direction. The objective here is to understand the aerodynamics of two G-12 flexible canopies. Figure 4 presents the parachute cluster showing the deformed shapes and canopy pressure at different instances of time. Color coding ranging from blue to red represents low to high magnitudes of pressure, respectively. At the time $t = 00:27s$, two parachute canopies are close to each other. They move away from each other as time progresses. This may be because the initial configuration that I assumed to start the FSI simulations is not in equilibrium. The canopies become after ($t = 03:36s$) as a result of change in pressure distribution. One notices a strong interaction between the uid and the canopies. Pressure distribution keeps changing with time. Parachute canopies are made of very flexible fabric materials and modeled as a membrane structure in the FSI simulations. As a result, the canopies quickly respond to any change in pressure distribution by adjusting their shape as observed in Figure 4. As a result of dynamic behavior, two canopies come closer to each other and then move farther away. This is a time dependent phenomenon. At time $t = 03:36s$, these two canopies start going in conical motion in counterclockwise direction about their vertical axis. This motion can be clearly seen in Figs. 4.5 (left and right). They rotate by about 45° in $10:71s$. Interestingly, this conical motion has also been observed in real life scenarios. I am not sure how the direction of rotation (clockwise or counter-clockwise) is chosen. I believe that a slight asymmetry in the mesh generated by the automatic mesh generator can give rise to counter-clockwise as being the preferred direction. In Figure 3: Cluster of two G-12 parachutes: Constructed (left) and prestressed (right) structural model. Figure 5 shows the pressure distribution on a plane cutting through the volume mesh and passing through the parachute canopy surfaces. As expected, there is a high pressure region inside of the canopy and lower pressure outside. This pressure gradient keeps the canopy inflated. In Figure 5(middle), a pair of vortices (the blue color spot signifies lower pressure at the core of a vortex) that are shed by the canopy can be seen in the downstream. These pairs of vortices shed from the left and right canopies are of different strengths implying that the aerodynamic responses of parachutes are not symmetric. Figure 5(right) shows a close view of the mesh viewed from the top and colored with pressure at $t = 10:74s$. Mesh is very _ne close to the canopy surface. Whole simulations required 10 remeshes. Each remesh resulted in about 3 million elements and 0.5 million nodes. This implies that about 4 million unknown fluid variables were computed every nonlinear iteration in each time step. Total computational time required was about 200 hours using 32 processors on 16 nodes of a Linux cluster (2GB RAM/node, 1.7GHz P4 Xeon with 1.2GB/s Myrinet Switch).

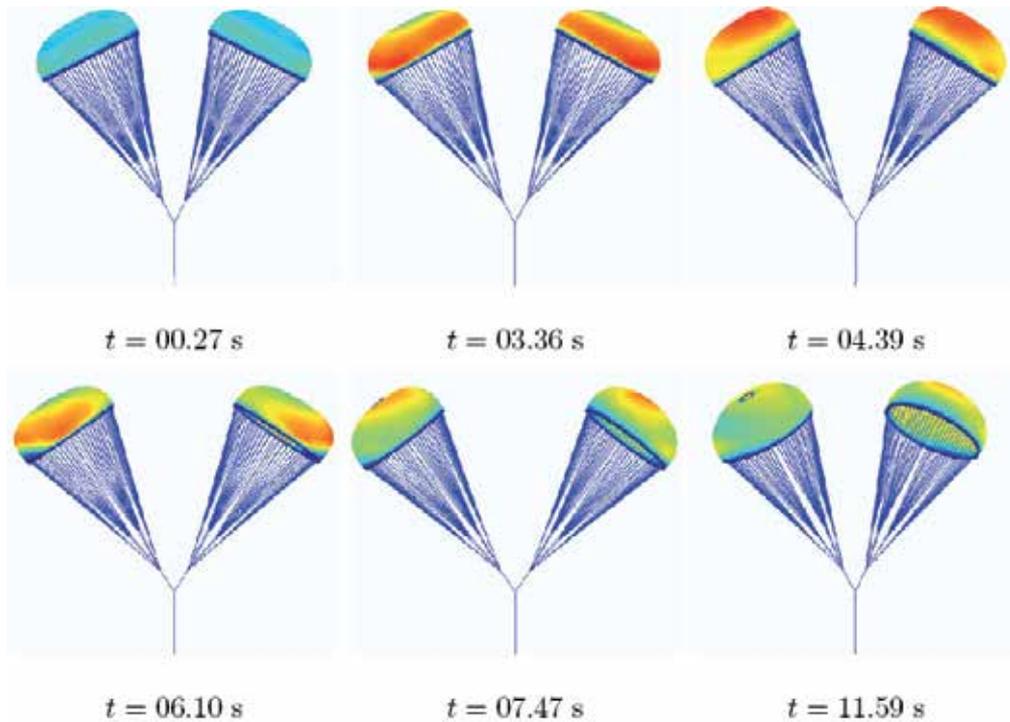


Fig. 4. G-12 parachute cluster with a fixed payload showing the deformed shapes and canopy pressure.

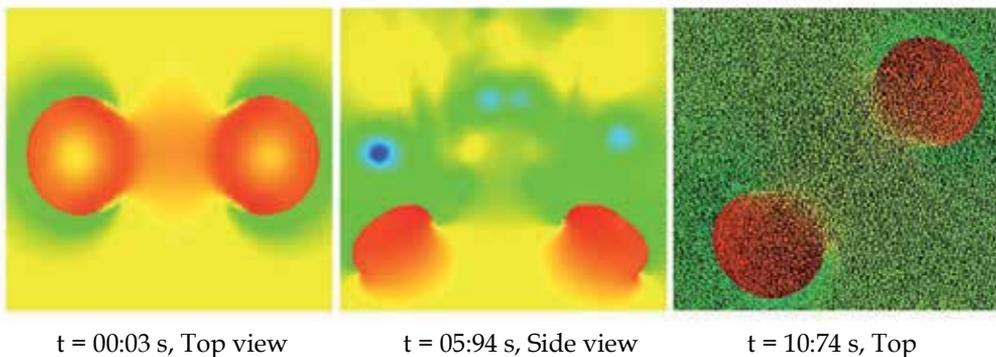


Fig. 5. G-12 parachute cluster with a fixed payload showing pressure (left and middle) and mesh colored with pressure (right) on a cross-section plane cutting through the parachutes.

Airdrop with free payload:

The pin was removed from the payload and the parachutes were released to fall with the payload weight. The total weight of the payload was 4,400 lb which is twice as heavy as that for a single parachute case. The flow results from a pinned case at $t = 06:10s$, when the FSI results were fully developed and free from startup conditions, were used as the initial conditions.

Figure 6 shows the shape of the parachutes and their fall in the vertical direction at three instants of time. In this figure, the fall position is determined with respect to an object falling with an assumed freestream velocity of 28 ft/s. The canopies are colored with pressure. With time they fall downward as expected implying that the terminal velocity in this case is higher than the assumed freestream velocity. Here, the cluster of parachutes with a payload of 4,400 lb is cruising at a terminal velocity of around 31 ft/s instead of 28 ft/s. Terminal velocity for a single parachute case is 28 ft/s. As expected, we lost the efficiency in cluster configuration. This implies that the impact velocity for the cluster at the time of landing would be higher than for a single parachute. This is not advisable for the safety of the payload. So, to find the terminal velocity of 28 ft/s in this case, we simulated a few more cases with lower payloads as shown in Figure 7. Following the trend of fall velocity for various payload weights from this figure, one notices that the lighter the payload, the lower the terminal velocity. One observes that the ideal weight for a payload would be about 3,600 lb. This weight will give a terminal velocity of 28 ft/s.

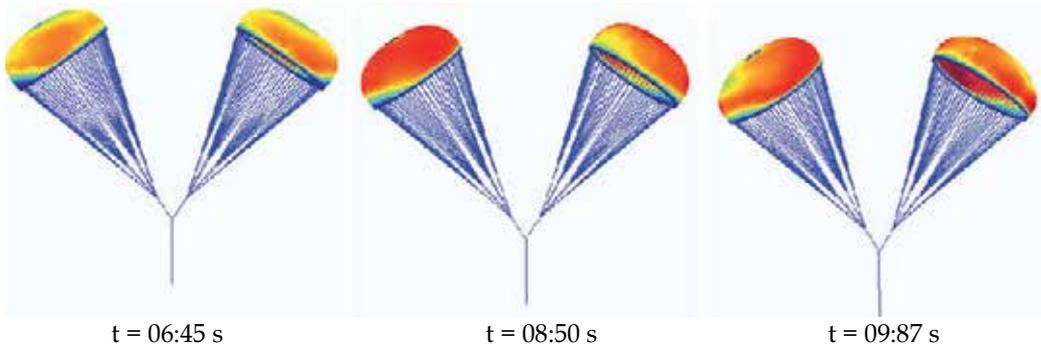


Fig. 6. Airdrop of payload using cluster of two G-12 parachutes: Parachute canopies colored with pressure.

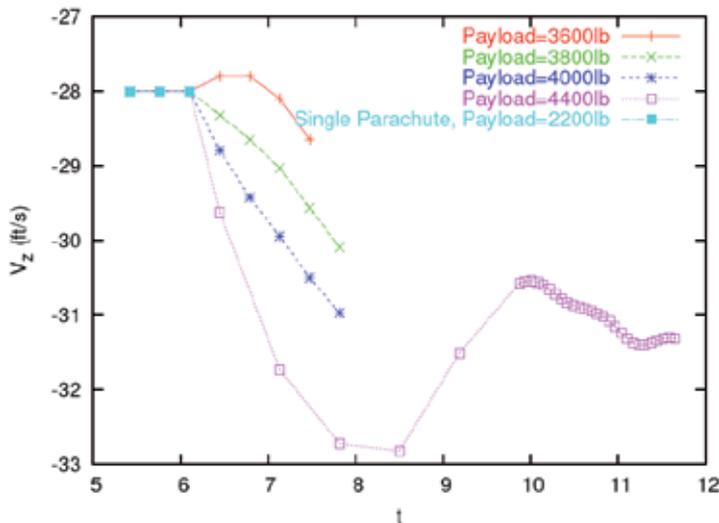


Fig. 7. Airdrop of a payload using a cluster of two G-12 parachutes: Fall velocity, V_z (ft/s) v/s time, t (s) for different weight of payload.

Soft landing:

We carried out simulations of two G-12 cargo parachutes to study the soft landing behavior of such systems. *Soft landing* implies making the landing of paratroopers or cargos softer. Various techniques are used to achieve this. Pneumatic Muscle Actuator (PMA) is one technology to achieve this objective. In PMA, the risers are made of pneumatic muscles. When the cargo is very close to the ground the muscle is pressurized. The length of PMA starts contracting and in the process the canopies get pulled downward and the payload gets pulled upward reducing the impact velocity. Previous soft landing simulations focused on the soft landing for single T-10 personnel parachute and on comparisons with drop test data (Stein K. , Tezduyar, Sathe, Benney, & Charles, 2005). These simulations provide a level of confidence for simulations of soft landings with our computational methods. A follow-on simulation was carried out for a single G-12 cargo parachute with a 2,200 lb payload and a G-12 parachute weight of 130 lb (Stein K. , et al., 2003). In this example, soft landing is modeled by reducing the natural length of the retraction cable from 12.80 ft to 2.88 ft in 1 second.

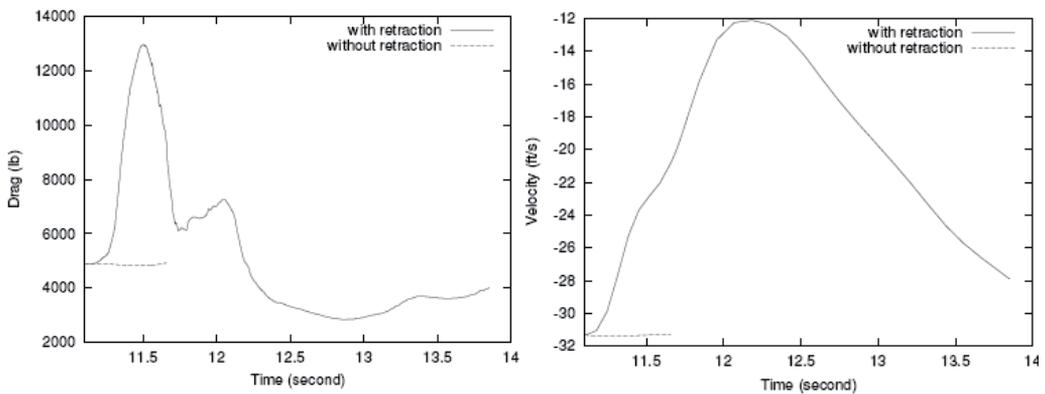


Fig. 8. Drag force (left) and payload velocity with a cluster of two G-12 parachutes during and after soft landing retraction.



Fig. 9. G-12 parachute cluster at the start of pneumatic muscle contraction (left), at the end of contraction (middle) and after soft landing retraction showing the deformed shapes and canopy pressures.

While soft landing systems have been shown to be effective for single parachutes, little is known about the retraction process for a cluster of parachutes. In this simulation, the soft landing of a 4,400 lb cargo with a cluster of two G-12 parachutes is modeled to study the behavior during and after the retraction process. The soft landing computation is carried out after the parachute has reached a state of terminal descent. In this computation, a 26 ft retraction device is modeled with a cable that connects the confluence point of the two parachutes to the payload. Parachute-payload retraction is modeled by reducing the natural length of the retraction cable by 7.68 ft in 0.27 seconds. Finally, a series of computations are carried out with the retraction cable held at its reduced length after the retraction is completed to study the post retraction behavior of the G-12 cluster. In addition to the soft landing simulation, a second computation is carried out without soft landing retraction. This no-retraction case is a shorter computation and serves as a baseline for comparison with the soft landing case in Figure 8. The aerodynamic drag force on the G-12 cluster during and after soft landing retraction is shown in Fig. 4.8. The effect of the soft landing is apparent from the sharp increase in the drag force during retraction. It is important to note the dramatic drop in drag shortly after retraction ends. This drop is accompanied by canopy collapse and suggests that *harder* landings can be experienced if ground impact is delayed too long after retraction. The payload velocity during descent of the G-12 cluster with and without soft landing retraction is shown in Figure 8 (right). Decreased descent speed resulting from soft landing retraction is very evident, with the payload descent rate decreasing from 31.0 ft/s to 12.0 ft/s. A sequence of snapshots of the two G-12 parachutes during and after the soft landing, colored with the corresponding differential pressures on the canopy is shown in Figure 9. The first two snapshots (left and middle) correspond to times at the start of retraction and at the end of retraction. The final snapshot (right) corresponds to a time well after retraction has finished, with the canopies showing more severe parachute deformations. At the end, before the simulations are stopped, few gores collapsed resulting in gore-to-gore contact. Further FSI simulations are not possible without a contact model.

5.2 Cluster of three G-12 parachutes

Varying numbers of parachutes have been used in cluster configurations by the Army to test the efficiency of each configuration. The most commonly used configurations are clusters of two and three parachutes. A cluster of three parachutes is one of the simplest clusters of parachute configurations. One benefit of this configuration is that this configuration does not have the natural tendency of going in conical motion as observed for a two-parachute cluster. Stein et al. (Stein, et al., 2001; Stein K. , et al., 2003)] presented the results from the stand-alone fluid dynamics simulations of a cluster of three rigid parachutes. The three parachute case presented more challenges in starting the FSI simulations than the two parachute case. Pressure ramping along with smaller time step was used to overcome the startup problems.

Problem setup:

A fluid mesh is generated by joining three pieces (domain box, refinement box and parachutes). Several preparation stages similar to the one discussed for a cluster of the two parachute case in the previous section are required. Geometry modeler software package called GAMBIT (Fluent, 2001), was used to model the computational domain and parachute

geometry and to generate the surface meshes. The volume mesh is generated by an unstructured automatic mesh generator (Johnson A. , 1995; Johnson & Tezduyar, Mesh update strategies in parallel finite element computations of flow problems with moving boundaries and interfaces, 1994; Johnson & Tezduyar, Advanced mesh generation and update methods for 3D flow simulations, 1999). A refinement boundary is used to get a satisfactory level of mesh refinements near the canopy's boundary to capture the flow physics accurately. The automatic mesh generator creates about 6 million tetrahedral elements and 1 million nodes of fluid mesh, resulting in approximately 8 million unknowns per nonlinear iteration at each time step with DSD/SST formulations. The size of the structure mesh is negligible compared to the fluid mesh. The Reynolds number, based on radius of a flat G-12 parachute, is 5.5 million. We used the parallel FSI solver on 128 Cray T3E processors to arrive at the results. The freestream velocity, which corresponds to terminal velocity, is assumed to be 20 ft/s.

Pinned payload:

The computed results from the FSI simulations are shown in Figure 10. We notice that the symmetric distribution is lost as a result of the dynamic behavior of parachutes in this cluster configuration. The changes in pressure distribution on the canopies surfaces result in stronger parachute-to-parachute aerodynamic interactions. In fact, two of the parachutes came very close to each other at the end of simulations. The closer they got, the higher the frequency of remeshes were and the more difficult it became to perform FSI simulations. At the onset of contact, the simulations have to be terminated in the absence of a contact model.

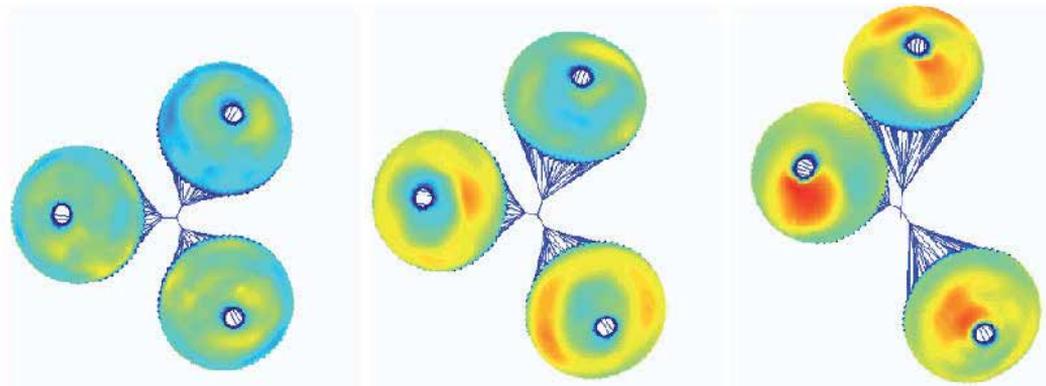


Fig. 10. Cluster of three G-12 parachutes at three different instances of time. Canopies are colored with pressure.

5.3 An example for compressible flow fluid-structure interaction simulation

Our targeted goal is to study the behavior of parachute dynamic in compressible flow regime (i.e., flow with Mach number > 0.3). Currently, the method has been tested for simple validation problems. Here, we present FSI results for a moving wedge problem in supersonic flow regime. The Mach number of the incoming fluid is 3.

At first, a solution for *fixed-wedge* case is arrived before switching moving algorithms. The *fixed-wedge* is used as the initial solution for the *moving-wedge*. The wedge is moving in the

downstream direction such that the equivalent freestream Mach number is 2.8. The computed values of Mach cone angles (θ) with the analytical values are compared. As observed in the table from Figure 11, there is a good agreement between these two results.



Fig. 11. Mesh with wedge (left), density at the beginning and end of the computations (center) and computed Mach cone angle (θ) compared for fixed and moving wedge with analytical solution.

6. Concluding remarks

In this chapter, we reviewed some computational challenges in fluid structure techniques. FSI simulations consisted of the following primary components: a solution method for the fluid dynamics, a solution method for the structural dynamics, an intelligent way to efficiently move and update a finite element mesh of fluid, and strategies for the coupling of fluid, and structural dynamics along the fluid-structure interface. The coupling is achieved in staggered fashion, with the fluid and structure coupled iteratively within a nonlinear iteration loop, and with multiple nonlinear iterations improving the convergence of the coupled system. The mesh domain was treated as a pseudo-elastic solid. To handle large structural-deformations, fluid dynamics solver required a methodology which can handle time-dependent spatial domain deformations. We presented Deforming-Spatial Domain / Stabilized Space-Time (DSD/SST) method which have a built-in capability to handle deforming structures. Semi-discrete formulations were used to carry out structural dynamics simulations which are based on the principle of virtual work. The SD model usually consists of membranes along with cables and payloads.

We presented results from the FSI simulations of a cluster of two G-12 parachutes under three for the parachute airdrop: fixed payload, airdrop, and soft-landing and cluster of three parachutes under fixed load conditions. Here, the parachute is represented as a structure composed of membranes, cables, and concentrated masses. The cables and membranes are assumed to have no flexural rigidity and experience large displacements and rotations. As a result, the interaction between a parachute system and the surrounding flow field is dominant in most of the parachute operations. Thus, the ability to predict parachute FSI is a challenge that must be faced in airdrop systems modeling. The G-12 parachute is a very large cargo parachute and its analysis presents convergence difficulties at the start of FSI simulations. A pressure ramping technique was proposed to deal with this start-up mismatch in the information exchange between the parachute's canopy and the fluid. It was found that a pressure ramping technique to smoothly transfer

the information between the fluid and structure solves the convergence problem encountered at the start of FSI simulations.

It was shown that one can ideally achieve an impact velocity as small as 12 ft/s using a soft-landing technique. However, for the 2-parachute cluster system, this minimum impact velocity is only achieved at a certain time after contraction. If one waits too long after contraction, the descent speed begins to increase, again, due to canopy collapse creating a harder landing. The amount of contraction is limited by the canopy strength because the loading on the canopy increases throughout the contraction process.

In a 3-parachute cluster system, we were able to predict that two parachutes collide as we observe in drop tests. We believe that this is caused by unsteady pressure distribution on the canopy surface as depicted in Figure 10. One can devise a control mechanism to prevent the collision of two canopies. One idea is to experiment with different control mechanisms such as heat-induced porosity-altering techniques for selected gores to introduce a sideways force.

The results from the FSI simulations of a cluster of three G-12 parachutes were also presented. The FSI simulations of both clusters of two and three parachute systems were found to suffer from contact issues. Two of the parachutes in this case came very close to contact. The FSI simulations faced a numerical barrier on the onset of contact. Similar contact issues were observed in the case of two parachute simulations when one or more gores collapsed. Further FSI simulations were not possible in the absence of a contact model.

In future research, we need to improve the structural modeling capabilities for the cables, risers, and membranes by instantaneously addressing the non-isotropic and non-linear deformation of the flexible structures caused by changes in the fluid dynamics forces to imitate the physical reality. Additionally, the addition of a contact model will enable the formulations to continue after contact is made to simulate what happens physically when contact is made while the parachutes are still falling through the air. All such additions to the code will greatly increase the cost, so increased accuracy in the model simulations depends on computational resources and continued innovation in clock speeds and floating point operations per second (FLOPS). Another route of research to address the contact is to experiment with different control mechanisms to prevent contact.

7. References

- Babuska, I. (1973). The finite element method with Lagrange multipliers. *Numerische Mathematik*, 20, 179-192.
- Behr, M., & Tezduyar, T. (1994). Finite element solution strategies for large-scale flow simulations. *Computer Methods in Applied Mechanics and Engineering*, 112, 3-24.
- Brezzi, F. (1974). On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *Modelisation Math. Anal. Number*, 8, 129-151.
- Brooks, A., & Hughes, T. (1982). Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 32, 199-259.
- Butler, M. C. (2001). Additional applications of bat sombrero slider technology. *Proceedings of AIAA 16th Aerodynamic Decelerator Systems Technology Conference*.

- Fluent, I. (2001). *GAMBIT 2, CFD Preprocessor*. Lebanon, NH: Fluent, Inc.
- Hirsch, C. (1988). *Numerical Computation of Internal and External Flows: Fundamentals of Numerical Discretization* (Vol. 1). Chichester: John Wiley and Sons.
- Hughes, T. J., Franca, L. P., & Mallet, M. (1986). A new finite element formulation for computational fluid dynamics: I. symmetric forms of compressible Euler and Navier-Stokes equations and the second law of thermodynamics. *Computer Methods in Applied Mechanics and Engineering*, 54, 223-234.
- Johnson, A. (1995). Mesh Generation and Update Strategies for Parallel Computation of Flow Problems with Moving Boundaries and Interfaces. *Aerospace Engineering, University of Minnesota, Ph.D. thesis*.
- Johnson, A. A., & Tezduyar, T. E. (1994). Mesh update strategies in parallel finite element computations of flow problems with moving boundaries and interfaces. *Computer Methods in Applied Mechanics and Engineering*, 119, 73-94.
- Johnson, A. A., & Tezduyar, T. E. (1999). Advanced mesh generation and update methods for 3D flow simulations. *Computational Mechanics*, 23, 130-143.
- Kumar, V. (2005). *Advanced Computational Techniques for Incompressible / Compressible Fluid-Structure Interactions*. Rice University: Rice University.
- Le Beau, G., & Tezduyar, T. (1991). Finite element computation of compressible flows with the SUPG formulation. In *Advances in Finite Element Analysis in Fluid Dynamics* (Vols. FED-Vol-123, pp. 21-27). New York: ASME.
- Mittal, S. (1992). *Stabilized Space-Time Finite Element Formulations for Unsteady Incompressible Flows Involving Fluid-Body Interactions*. University of Minnesota.
- Sahu, J., & Benney, R. (1997). *Prediction of terminal decent characteristics of parachute clusters using CFD*. Technical report, AIAA Paper No. 97-1453.
- Shakib. (1988). *Finite Element Analysis of the Compressible Euler and Navier-Stokes Equations*. Ph.D. thesis, Stanford University, Mechanical Engineering.
- Stein, K. (1999). Simulation and Modeling Techniques for Parachute Fluid-Structure Interactions. *Aerospace Engineering, University of Minnesota, Ph.D. thesis*.
- Stein, K., Benney, R., Tezduyar, T., Kumar, V., Thornburg, E., Kyle, C., et al. (2001). Aerodynamic interaction between multiple parachute canopies. *Proceedings of the First MIT Conference on Computational Fluid and Solid*. Boston.
- Stein, K., Tezduyar, S., Sathe, S., Senga, M., Ozcan, C., Soltys, T., et al. (2003). Simulation of parachute dynamics during control line input operations. *Proceedings of the 17th AIAA Aerodynamic Decelerator Systems Technology Conference*. Monterrey, California.
- Stein, K., Tezduyar, T. E., Sathe, S., Benney, R., & Charles, R. (2005). Fluid-structure interaction modelling of parachute soft-landing dynamics. *International Journal for Numerical Methods in Fluids*, 47, 619-631.
- Stein, K., Tezduyar, T., Kumar, V., Sathe, S., Benney, R., Thornburg, E., et al. (2003). Aerodynamic interactions between parachute canopies. *Journal of Applied Mechanics*, 70, 50-57.
- Tezduyar, T. E. (1992). Stabilized finite element formulations for incompressible flow computations. *Advances in Applied Mechanics*, 1-44.
- Tezduyar, T. E., Behr, M., & Liou, J. (1992). A new strategy for finite element computations involving moving boundaries and interfaces - the deforming-spatial-domain/space time procedure: I. the concept and the preliminary tests. *Computer Methods in Applied Mechanics and Engineering*, 339-351.

- Tezduyar, T., Mittal, S., Ray, S., & Shih, R. (1992). Incompressible flow computations with stabilized bilinear and linear equal-order-interpolation velocity pressure. *Computer Methods in Applied Mechanics and Engineering*, 95, 221-242.
- Udoewa, V. (2005). *Computational Techniques for Aerodynamic Interactions between*. Rice University: Mechanical Engineering and Material Science.

Part 3

Heat Transfer, Combustion, and Energy

Fluid Flow in Polymer Electrolyte Membrane Fuel Cells

Alfredo Iranzo, Antonio Salva and Felipe Rosa
*Thermal Engineering Group, Energy Engineering Department
School of Engineering, University of Seville
Spain*

1. Introduction

Polymer Electrolyte Membrane Fuel Cells (PEMFC) have attracted significant interest during the last few decades, as they are considered to be one of the most promising alternative clean power generation devices for portable, mobile and stationary applications. However, different technological barriers such as cost, durability, or heat and water management, are limiting the implementation of fuel cell systems into the global energy markets, and therefore significant research efforts and investments are being carried out.

Fuel cells are devices where electrochemical reactions transform chemical energy available in fuels into electrical energy. Fuel cells are not limited by the thermodynamic restrictions of conventional power generation systems, such as the Carnot efficiency, meaning that fuel cells can be operated with higher efficiency for energy conversion. Additionally, the environmental impact is low as no combustion processes occurs and no pollutants are generated (U.S. Department of Energy [DOE], 2004).

A typical fuel cell power system consists of different components:

- Single cells, where the electrochemical reactions occur.
- Stacks, consisting of the necessary number of cells electrically connected to provide the required power capacity.
- The balance of plant, or additional equipment to provide fuel and oxidants with the appropriate conditions, thermal management, electric power conditioning, and other functions.

Single or unit cells are the core of a fuel cell. They convert the chemical energy contained in a fuel into electrical energy, via electrochemical reactions. The basic configuration of a fuel cell consists of an electrolyte layer or membrane in contact with an anode and a cathode on either side.

In a PEM fuel cell, hydrogen is continuously supplied to the anode or negative electrode, and an oxidant, often oxygen or air, is also continuously supplied to the cathode or positive electrode. Electrochemical reactions occur at the electrodes, generating an electric current through the electrolyte thus driving the corresponding electric current that performs the electric work on the load.

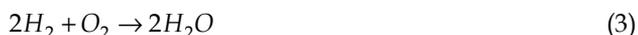
At the anode, hydrogen is fed to the cell and a reaction takes place at the catalyst layer:



The protons migrate through the polymeric membrane and react with the oxygen and the electrons at the cathode catalyst layer:



Therefore, the overall reaction taking place in the cell is:



Fuel cells are modular, so that unit cells can be combined into a stack to achieve the voltage and power required for the specific application. This involves connecting multiple unit cells in series via electrically conductive materials. The connection is achieved via the bipolar plate (BP), that also provides a gas barrier that separates the fuel and oxidant of adjacent cells (DOE, 2004), (Larminie & Dicks, 2003), (Barbir, 2005). The channels that distribute the gas to the electrode surface are usually imprinted in the plate.

The design of a fuel cell must therefore comprise a number of physical and chemical phenomena, such as electrochemical reactions, heat transfer and thermal management, material selection, fluid dynamics for laminar flow and flow in porous media, flow distribution of the reactants in the stack manifolds, and others. These phenomena are tightly coupled, making the design and optimization of a fuel cell a complicated task usually aided by numerical models and simulations, as well as by experimental techniques. A considerable effort and intense developments are being made by governments, public institutions, universities, research institutes and private companies in the field of fuel cell research.

In this chapter, the fluid flow within single cells and fuel cell stacks is analysed.

2. Fluid flow in PEM fuel cells

The fluid flow characteristics within the Fuel Cell are of major importance for the final performance of the system. This chapter deals with the main fluid flow mechanisms occurring within a Fuel Cell at two different levels: single cells and stacks.

The flow inside a single cell comprises a number of different transport mechanisms, namely laminar flow in channels, flow in porous media, multiphase flow with phase change, and diffusion mechanisms. This will be addressed in section 3. On the other hand, the flow in a Fuel Cell stack is mainly concern about the flow in the inlet/outlet stack manifolds, and the reactant distribution over the different cells. This will be addressed in section 4.

3. Fluid flow in a single cell

3.1 Description of a single cell: Main components and processes

The core of a single cell is a polymer membrane acting as an electrolyte. The membrane is a proton-conductive material, with porous electrodes on both sides of the membrane. The electrodes are porous materials to allow both reactant gases to reach the catalyst sites and products to leave the electrode. Electrochemical reactions occur at the catalyst surface on the electrodes. This multilayer assembly is called Membrane Electrode Assembly (MEA). The MEA is located between collector or separator plates called bipolar plates. Bipolar plates connect the cathode of one cell to the anode of the adjacent cell both physically and electrically. Reactant gases are fed to the Bipolar Plates, where imprinted channels distribute

the gases over the entire surface. In order to provide a better uniformity of the reactants over the electrode, a Gas Diffusion Layer (GDL) is placed between Bipolar Plate and electrode. The main components of a cell are depicted in Figure 1.

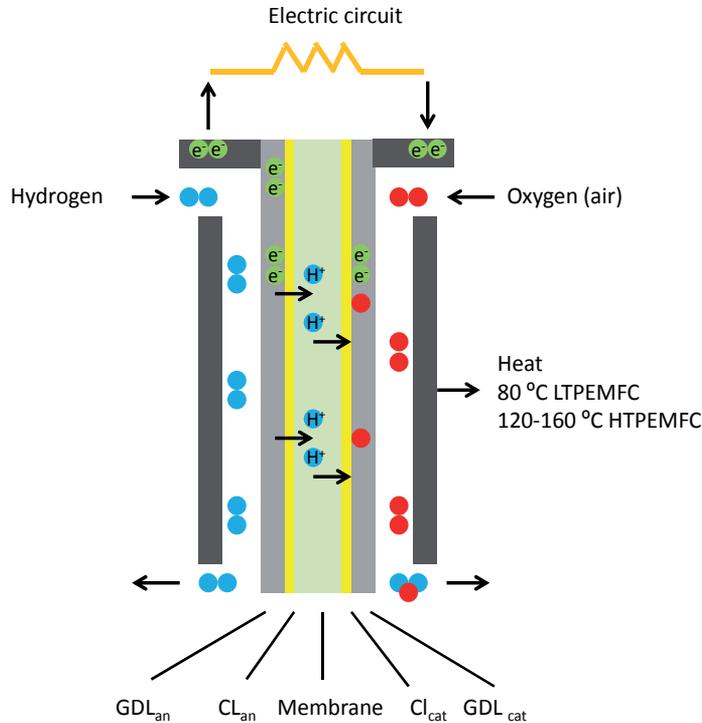


Fig. 1. Main components of a single fuel cell

The following processes take place inside the fuel cell:

- Gas flow through the channels.
- Gas diffusion through porous media.
- Electrochemical reactions at the electrodes or catalytic layers (CL).
- Proton transport through proton-conductive polymer membrane.
- Electron conduction through electrically conductive cell components.
- Water transport through polymer membrane including both electrochemical drag and back diffusion.
- Water transport (both vapour and liquid) through porous catalyst layer and gas diffusion layers.
- Two-phase flow of unused gas carrying water product.
- Heat transfer, including both conduction through solid components of the cell and convection to reactant gases and cooling medium.

A schematic diagram of the main transport mechanisms is depicted in Figure 2.

The design of the cell components must comprise a trade-off of these processes with minimum losses. A typical example is the porosity of the GDL: while high porosity would help the gas diffusion, this would hinder the electric conductivity and thus the cell performance. In general, the processes depicted above are tightly coupled in the cell. For

example, at the cathode side of the MEA a complex interrelation between various transport phenomena take place. The processes occurring at the cathode are (Djilali, 2007):

- Conduction of electrons between the current collectors and the reaction sites.
- Conduction of protons between the membrane and catalyst layer.
- Transport of the reactants to the catalyst layer.
- Transport of heat away from the reaction site primarily through the solid matrix.
- Condensation/evaporation and transport of liquid water and vapour.

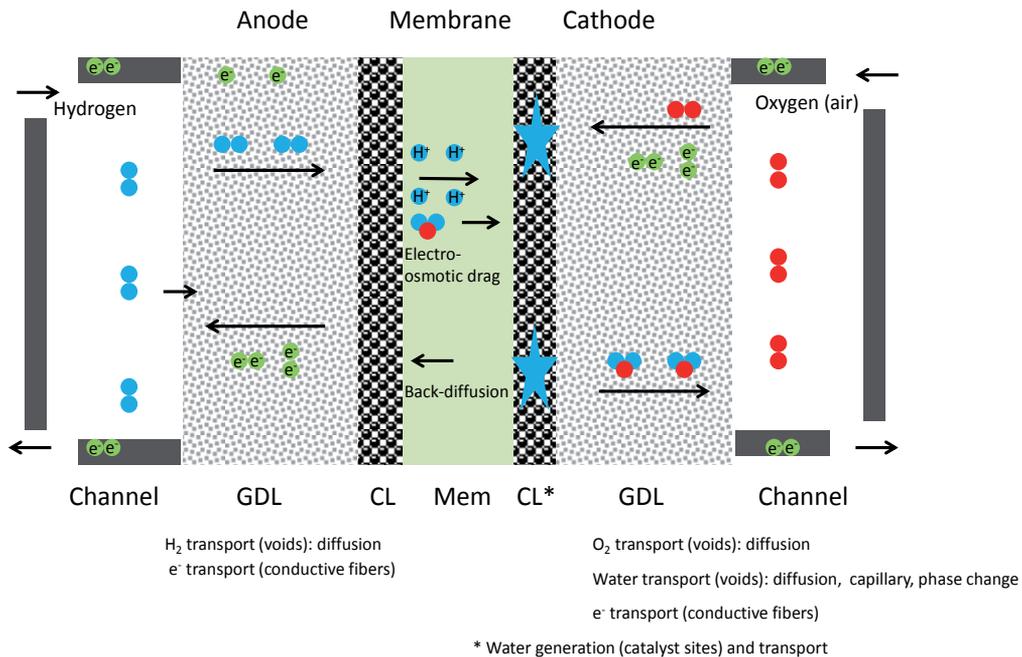


Fig. 2. Main processes within a single PEM fuel cell

On the cathode, the reaction is likely to suffer from mass transport limitations at high current densities, when depletion of oxygen in the air stream becomes significant. This limitation can be further extended by flooding of the electrode pores by the product water and blocking of pathways for oxygen transport to the reaction sites. Effective removal of excess water is thus essential, but must be balanced against the need to prevent drying of the adjacent proton conducting media (membrane), which not only causes deterioration in protonic conductivity (and thus higher ohmic losses and lower efficiency), but also impacts durability.

The performance of a fuel cell is usually represented by means of a polarization curve I-V as shown in Figure 3.

In the polarization curve, three different regions can be generally observed. The cell performance is dominated by kinetic or activation losses at low current densities, while at high current densities mass transport limitations have a major effect given the high amount of reactants required to maintain the intensity and the pore blockage produced by the water flooding the cathode. At medium current densities and voltages, ohmic losses become an important factor in the fuel cell performance.

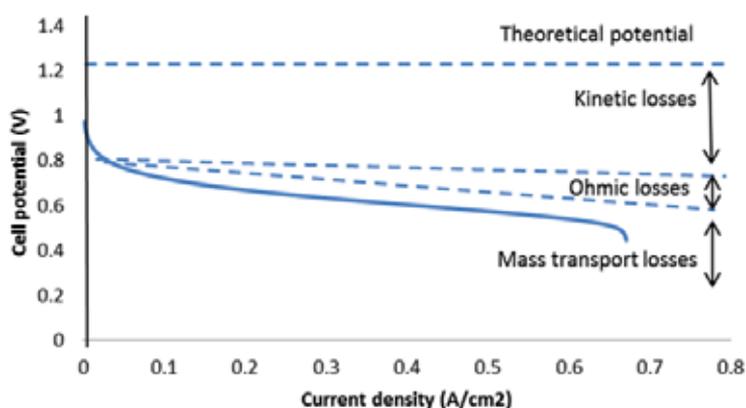


Fig. 3. Polarization curve of a single PEM fuel cell

The performance of a given fuel cell is affected by the parameters listed below (Barbir, 2005):

- Temperature
- Pressure
- Reactant flow rates (reactant utilization)
- Reactant composition (humidification, fuel and oxidant concentration)
- Current density
- Other parameters: reactant impurities, cell components degradation, etc.

From thermodynamics and electrochemical reaction kinetics, the Fuel Cell performance is higher when the operating pressure and temperature is increased. However, the complete system and balance of plant must be considered for the appropriate selection of the operating parameters (compression power, water management, etc.).

Regarding reactant flow rates, generally speaking higher flow rates result in improved performance of the fuel cell, as higher flow rate helps remove liquid water from the cell and maintain a high oxygen concentration within the catalyst layer. However, when hydrogen is considered, a too high flow rate will reduce the system efficiency as part of the hydrogen would be wasted.

In Nafion membranes, reactants must be supplied humidified in order to keep the membrane sufficiently hydrated and enhance its protonic conductivity. A complex water transport and management issues arises, with water being generated at the cathode, and transported by means of different mechanisms through the cell. The degree of humidification required depends on the cell design and membrane type. For example, thinner membranes such as Nafion-111 present a better cathode to anode water transport via diffusion. Water management is one of the most active research fields in PEM Fuel Cells, and detailed reviews can be found in the work of Dai et al. (Dai et al., 2009) and Jiao and Li (Jiao & Li, 2011).

The real potential obtained in a fuel cell is always less than the theoretical maximum, as certain irreversible losses exist. The main contributions to the losses are (Barbir, 2005; Mench 2008):

- Activation losses:
Activation losses area caused the activation energy needed for the electrochemical reactions at the electrodes. The kinetics of the reactions results in part of the voltage generated being used as driving force for the chemical reaction.

Therefore these losses depend on the reactions taking place, the properties of the electro-catalyst material and its microstructure, among others.

- Ohmic losses:

The voltage drop caused by ohmic losses represents the resistance to the flow of electrons through the materials and contact resistance, and the ionic resistance in the electrolyte. This is mainly proportional to the current density, and also depends on the material properties, cell geometry and working temperature.

- Mass transport or concentration losses:

Both fuel and oxidant must be transported to the electrodes surface as they are being consumed. There are transport limitations depending of the current density, reactant activity and electrode and GDL microstructure, and therefore the concentration of the reactants at the surface of the electrodes is reduced. At high current densities where water generation at the cathode is significant, liquid water being transported through the catalyst layer and GDL partially blocks the pores and the catalyst sites (flooding), thus leading to a limitation in the oxygen transport and reaction. This is considerably affecting the output voltage lowering the performance of the cell at high current densities (Li et al., 2008).

The high degree of coupling between all of the processes occurring within the cell makes necessary an overall approach for the analysis of the fuel cell. A well-established approach must account for all these processes, otherwise severe errors in the design can occur. For example, consider a design criteria based on the reactant transport from the channel to the catalytic layer. If no current transport is considered, no electric contact resistance between the GDL and the BP will be included in the study, and therefore the designer will probably reach the conclusion that a high channel-to land ratio optimises the performance. However, with a too high channel-to land ratio, the electric contact resistance between GDL and BP is increasing and the cell performance drops significantly.

The following subsections provide an insight into the main fluid flow phenomena within a single cell.

3.2 Fluid flow in the bipolar plate channels

The fluid flow in Bipolar Plate channels is always laminar in PEM Fuel Cells given the low Reynolds number found, generally in the range 50-500.

$$Re = \rho VD/\mu \quad (4)$$

where Re is the Reynolds number (dimensionless), ρ and μ are the fluid density and dynamic viscosity, V is the fluid velocity and D the channel hydraulic diameter. For anode and cathode, channel square section is around 1 mm x 1mm, with gas velocities in the range 0.1-5 m/s. Laminar flow is therefore found in the flow field of Bipolar Plates.

The equations governing the general fluid flow are the Navier-Stokes equations, also known as conservation or transport equations, which can be written in its general form as:

$$\frac{\partial}{\partial t} \int_V \rho \phi dV + \oint_A \rho \phi V \cdot dA = \oint_A \Gamma \nabla \phi \cdot dA + \int_V S_\phi dV \quad (5)$$

Where ϕ is the transported quantity, t is the time, A the superficial area, V the volume, Γ is transported quantity diffusivity, and S_ϕ is the source of ϕ . The first term in the equation corresponds to the transient transport of ϕ , the second term to the transport by convection

mechanism, the third term represents the transport of ϕ by diffusion, and the fourth term represents the source (or sink) of ϕ .

The different transport equations are assembled by using the appropriate variables, as shown in Table 1.

Equation	Variable ϕ
Continuity	1
x-momentum	u (velocity in x-direction)
y-momentum	v (velocity in y-direction)
z-momentum	w (velocity in z-direction)
Energy	h (enthalpy)
Chemical specie i	y_i (mass fraction of i)

Table 1. Main transport equations

There are different types of flow field designs for Bipolar Plates. Pin-type and parallel flow fields are found in many fuel cells, although the most widely used design in industry is the serpentine flow field (Li & Sabir, 2005). Some experimental techniques are available for flow visualization in Bipolar Plates (Barreras et al., 2005), but numerical modelling has been established as a powerful tool for flow analysis and cell components design (Djilali, 2007). For example, Fig. 4 presents CFD results for the flow distribution in the channels of a parallel flow field Bipolar Plate. The corresponding pressure fields are depicted in Fig. 5. Preferential paths that may lead to an uneven reactant distribution over the electrode are clearly observed, enabling the researcher to redesign the flow field for the cell improvement. The results correspond to the work done by Iranzo et al. (Iranzo et al, 2010, 2011a), where a commercial 50 cm² fuel cell was modelled by means of Computational Fluid Dynamics (CFD), and the numerical results were experimentally validated in a test station for different bipolar plate designs and operating conditions, using the real cells. The details of the CFD model, simulation, and results validation can be found in Iranzo et al. (Iranzo et al, 2010, 2011a).

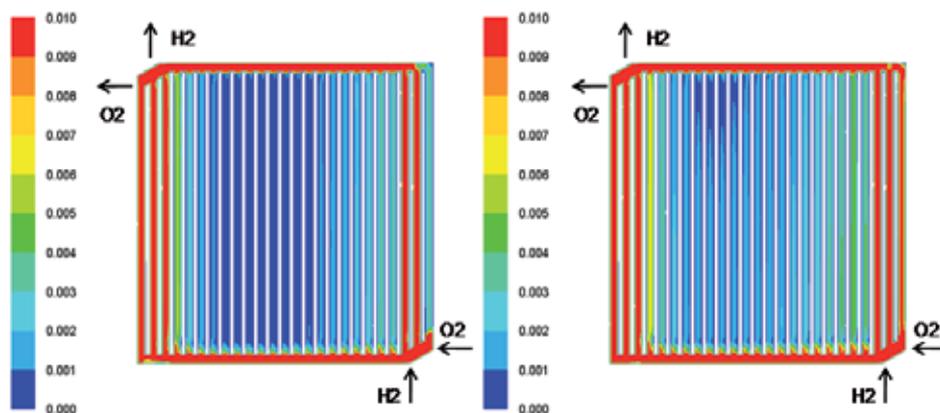


Fig. 4. Computational Fluid Dynamics simulation of the velocity distribution in a 50 cm² Bipolar Plate with parallel flow field design (cathode side). Left: 100 mA/cm². Right: 600 mA/cm².

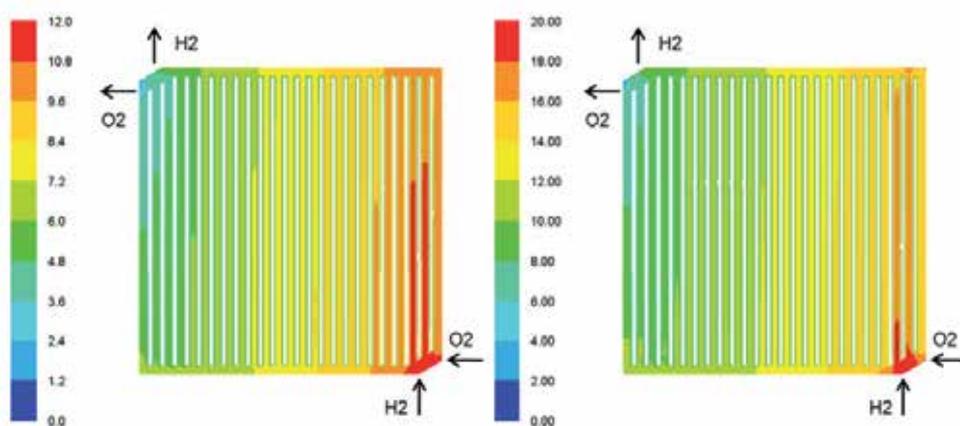


Fig. 5. Computational Fluid Dynamics simulation of pressure distribution in a 50 cm² Bipolar Plate with parallel flow field design (cathode side). Left: 100 mA/cm². Right: 600 mA/cm².

However, as the physical and electrochemical phenomena within a fuel cell is highly coupled, it is in general necessary to work with a more complete model considering the fundamental phenomena, in order to accurately represent the cell.

For that, the N-S equations listed above (Eq. 5) can be augmented to account for different physical and chemical phenomena relevant to fuel cells such as electrochemistry, flow in porous media and multiphase flow. An excellent review on computational heat and mass transfer modelling for PEMFCs is found in the work by Siegel (Siegel, 2008).

3.2.1 Multiphase flow in the bipolar plate channels

The flow in the Bipolar Plate channels is not always a single flow problem. Especially at high current densities the water generation (Eq. 2, Eq. 3) can be large enough for establishing a two-phase flow. This is particularly important in the catalyst layer and GDL, as the liquid water blocks the catalyst sites and GDL pores, inducing a high resistance for the oxygen transport. Regarding the channels, liquid water emerges from the GDL and enters the channel in the form of droplets. A complex droplet dynamics is established then in the channel, where the water is dragged by the reactant flow towards the exit. Depending on the gas flow rate and the wall contact angle, the flow regime in the channel can be a mist, a plug or slug flow, or a film flow. Liquid water must be efficiently evacuated from the cell, in order to ensure an adequate path for the oxygen to reach the catalyst sites. The two-phase flow can be visualized by means of advanced experimental techniques such as neutron imaging (Park et al., 2008; Wu et al., 2008b), but again Computational Fluid Dynamics is used by many researchers for the investigation of two-phase flow in fuel cells, although it is one of the most complex modelling tasks currently (Wang, 2004).

Fig.6 presents results for the liquid water concentration in a fuel cell with parallel flow field. This particular flow field design is known to be inadequate for handling liquid water, as droplets typically collapse one of the channels (red spots in the Fig. 6), the gas flow finds a high local resistance, and as other channels are still empty the flow is unable to flush the accumulated water (Li & Sabir, 2005).

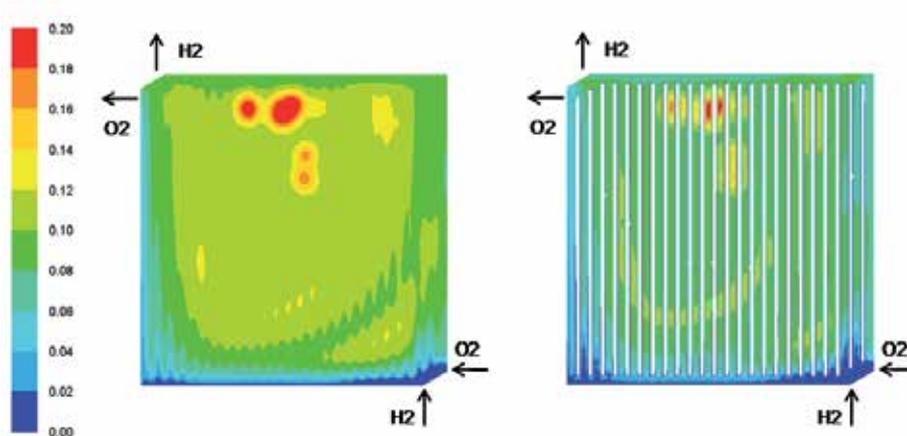


Fig. 6. Computational Fluid Dynamics simulation of the liquid water saturation (liquid water mass fraction) in a 50 cm² PEMFC with parallel flow field design (cathode side) operating at 600 mA/cm². Left: CL-GDL interface. Right: GDL-channel interface.

3.3 Fluid flow in the porous components: Gas diffusion layer and catalytic layer

The modelling of flow in porous media is a key issue in fuel cells, as both GDL and catalytic layers are porous materials. Typically, the porous media is represented by adding a negative source in the momentum equations, so that a pressure drop is generated accordingly. The consequent flow acceleration produced by the volume blockage or porosity can be accounted for. The source term is decomposed in two parts as in Eq. (6), the first one representing the viscous loss as in Darcy's law, and the second one representing inertial losses.

$$S_i = -\left(\frac{\mu}{\alpha} v_i + C_2 \frac{1}{2} \rho v_{mag} v_i\right) \quad (6)$$

where S_i is the source term in i direction, α is the permeability, μ the kinematic viscosity, ρ the fluid density, C_2 the inertial resistance factor, v_i the velocity in i direction and v_{mag} the velocity magnitude. For the velocity ranges typically found in fuel cell GDL and catalytic layers only the viscous loss is significant.

3.3.1 Diffusion mechanisms

One of the main transport mechanisms within the cell porous components is diffusion. Binary diffusion is generally considered, although multicomponent diffusion (Maxwell-Stefan diffusion) is also used for non-dilute mixtures with more than two components. The need for an effective diffusivity stems from the fact that porous media presents important geometric restrictions (pores, etc.). The Bruggemann correction is the most used correction factor for the diffusion coefficient (Sundén & Faghri, 2005), where:

$$D_{eff} = \epsilon^{1.5} D \quad (7)$$

The pores of the catalyst layer (CL) and the Microporous layer (MPL, a fine layer serving as an effective interface between CL and GDL) are significantly smaller than the pores in the GDL. In such porous media where the continuum regime can be not valid (when $Kn > 1$), it is

necessary to account for the Knudsen diffusion mechanism. The transition is determined by the Knudsen number:

$$\text{Kn} = \lambda / d_{\text{pore}} \quad (8)$$

The diffusion coefficient is then determined from the kinetic theory of gases (Sundén & Faghri, 2005).

3.3.2 Multiphase flow in porous components: Capillary forces

Multiphase flow on the cathode side represents a main concern for the practical operation and performance of a fuel cell, given the potential flooding of the porous CL and GDL, and represents currently one of the main areas of research in the field of PEM fuel cells.

PEM fuel cells operate under 100 °C, thus water vapour formed during the operation of the cell may condense to liquid water. This is especially true at higher current densities when the reaction rates are higher and therefore water production becomes important. An excess of reactant in the cathode side helps evacuating liquid water, but increases the parasitic power associated to gas compressors.

The production of water takes place in the cathode catalyst layer due to the oxygen reduction reaction (ORR). Another source of water is the humidification of the reactants. Different transport mechanisms through the membrane are available for water: electro-osmotic drag, proportional to the protonic flux, and diffusion, proportional to the water molar concentration gradient. Both effects operate in opposite directions and a correct water management implies a zero net water transport so that membrane and electrodes becomes neither flooded nor dried-out. Water management depends as well on the membrane thickness. Thin membranes such as Gore or Nafion-111 enhances the water back-diffusion from cathode to anode, but thick membranes such as Nafion-112 or 117 reduces the diffusion transport.

A correct amount of water is needed to keep the membrane hydrated, but excessive water blocks the porous media and reduces the mass transport rate by diffusion and also reduces the active reacting surface area, and therefore the cell performance. This is why correctly modelling water formation and transport is a key issue in predicting the cell behaviour with accuracy. Different types of multi-phase models have been used in the literature, such as the mixture model, the multifluid model, and the saturation model. A complete revision of the different modelling approaches can be found in the work by Wang (Wang, 2004). One of the modelling approaches is the saturation model, where a multi-component single phase system is resolved, with an additional equation accounting for the liquid saturation. The conservation equation for the liquid water volume fraction or water saturation, s , is solved:

$$\frac{\partial(\epsilon \rho_l s)}{\partial t} + \nabla \cdot (\rho_l \vec{V}_l s) = r_w \quad (9)$$

where the subscript l stands for liquid water, and r_w is the condensation rate, modelled as a function of the difference between vapour partial pressure and saturation pressure.

As a first approach the liquid velocity V_l can be assumed to be equivalent to the gas velocity inside the gas channel, so that the liquid droplets would be a fine mist. Inside the porous zones the use of the capillary diffusion term allows the replacement of the convective term in Equation (9).

The capillary pressure can be calculated as a function of s , depending on the wetting phase:

$$p_c = \begin{cases} \frac{\sigma \cos \theta_c}{\left(\frac{K}{\varepsilon}\right)^{0.5}} (1.42(1-s) - 2.12(1-s)^2 + 1.26(1-s)^3) \theta_c < 90^\circ \\ \frac{\sigma \cos \theta_c}{\left(\frac{K}{\varepsilon}\right)^{0.5}} (1.42s - 2.12s^2 + 1.26s^3) \theta_c > 90^\circ \end{cases} \quad (10)$$

Where σ is the surface tension, ε is the porosity, θ_c is the contact angle and K is the absolute permeability.

Such approach enables the modelling of various physical processes such as condensation, vaporization, capillary diffusion, and surface tension. Also the blocking of the porous media and the flooding of the reaction surface can be modelled by multiplying the porosity and the active surface area by a factor $(1-s)$, respectively.

The saturation model presented here is used by a number of groups, although it misses the main phenomena associated to liquid water transport and multiphase flow effects. The saturation model is in fact an enhanced multicomponent flow model, where different species coexist in a single-phase medium, rather than a true multi-phase flow model.

Other approaches for the modelling of water transport have been developed and applied by other research groups. Euler-Euler multiphase-flow is found in many fuel cell models reported in the literature, as well as Lattice-Boltzmann methods.

3.4 Membrane property models

In order to solve the governing equations for the fuel cell it is required to implement expressions for the various transport properties, especially in the membrane. The membrane phase is usually modelled as a porous media. There are many physical properties that must be correctly modelled in order to predict the fuel cell behaviour with accuracy, in particular: gas phase diffusivities, electrical conductivity, water diffusivity, and osmotic drag coefficient.

As seen in the modelling equations, the water content is a key variable. It represents a driving force and an integral part of PEMFC water management, and it is also a crucial parameter as almost all property expressions depends on it, such as conductivity. Because of the importance of water content, it has received more modelling attention than other membrane properties. Such models span from simple curve fits to more elaborate chemically based models.

There are different methods to determine the transport properties: empirically, semi-empirically, or by means of modelling. The semi-empirical method is used by most models. In particular, the polynomial fit of Springer et al. (Springer et al., 1991) is used in many fuel cell models:

$$\lambda = \begin{cases} 0.043 + 17.18a - 39.85a^2 + 36a^3 & (a < 1) \\ 14 + 1.4(a - 1) & (a > 1) \end{cases} \quad (11)$$

where λ is the membrane water content, and a is the water activity, defined as P_{wv}/P_{sat} .

The gas species diffusivities can be modelled as a function of the water saturation s in order to account for the pore blockage when liquid water is present. An expression is given in Um et al. (Um et al., 2000):

$$D_i = \varepsilon^{1.5} (1-s)^{r_s} D_i^0 \left(\frac{p_0}{p}\right)^{\gamma_p} \left(\frac{T}{T_0}\right)^{\gamma_t} \quad (12)$$

where D_i is the diffusivity of species i at reference temperature and pressure (P_0, T_0). The exponents and the pore blockage r_s are model parameters. Nam and Kaviany (Nam & Kaviany, 2003) provides a detailed discussion on the values of r_s .

The membrane electric conductivity can be modelled with the expression:

$$\sigma_{mem} = \varepsilon(0.514\lambda - 0.326)e^{1268(\frac{1}{303} - \frac{1}{T})} \quad (13)$$

where λ is the water content. The membrane osmotic drag coefficient follows the expression:

$$\alpha_d = 2.5 \frac{\lambda}{22} \quad (14)$$

3.5 Transient effects

Transient phenomena, such as the dynamic response of the cell with a load change or the cell start-up, adds additional challenges to the analysis of the flow and cell performance.

The different time scales associated with the main processes occurring within the cell must be considered. In fuel cells, the water transport phenomena are presenting larger timescales than electrochemical, mass transport, or gas flow timescales. Wang (Wang, 2004) compares the timescales for the different phenomena: the time constant for the electrochemical reaction is known to be typically in the range of micro- and milliseconds. Electrochemical reaction timescales can also be determined experimentally (Iranzo et al, 2011b). The time constant for the transport of the reactant gas through the GDL is estimated by its diffusion time:

$$\tau_{MT} = \delta_{GDL}^2 / D_g \quad (15)$$

Where δ_{GDL} is the GDL thickness and D_g is the diffusion coefficient. For typical cells, the time constant is in the range between 0.1 and 1 s as stated by Wang (Wang, 2004). Finally, the time constant for the membrane hydration process, associated to multiphase and water transport effects, can be estimated by (Wang, 2004):

$$\tau_H = \frac{\rho \delta_m \Delta \lambda}{EW} \left/ \left(\frac{I}{2F} \right) \right. \quad (16)$$

Where EW is the membrane equivalent weight, I the intensity, F the Faraday constant, δ_m the membrane thickness, and λ the water content. This results in a typical time constant of several seconds.

Therefore, the differences in the physical and chemical phenomena timescales in a fuel cell are at least within a 10^3 range.

3.6 Numerical and experimental analysis of fluid flow in PEM fuel cells

The main tools available for the analysis of the fluid flow inside fuel cells are the numerical modelling of the system and specific experimental techniques.

3.6.1 Numerical modelling of PEM fuel cells

Fuel cell modelling has received much attention over the past 15 years in an attempt to better understand the phenomena occurring within the cell.

Models can be used for a wide range of situations and problems within fuel cells, from very simple pre-dimensioning models to complex three-dimensional models covering different physical and chemical phenomena. Therefore, very different models have been developed.

In terms of the region of interest, numerical simulation can be divided into modelling at the electrode level or micro-modelling, or at cell or stack level (macro-modelling). The microscopic models provide the fundamental understanding of processes like diffusion and conduction in the membrane on a single-pore or even molecular level. They allow for the evaluation of small perturbations like pores heterogeneity as well as the incorporation of small-scale effects. Such models require a lot more knowledge of the microstructure. To allow modelling of entire PEFC behaviour, macroscopic models are more commonly used, where some microscopic details are included in the model. Computational Fluid Dynamics works at macro-modelling level, aiming at the optimization of different design alternatives and determining operational strategies, incorporating built-in models for the electrochemical performance of the given electrodes and electrolytes. CFD models work either for 2D or 3D problems, where during the last years a high increase in the number of 3D models have been reported in the literature.

3D models allows for a complete representation of the real fuel cell geometry, thus enabling for the evaluation of different cell geometries and for the understanding of the influence of operating conditions. Given the wide range of possible geometries and operating parameters that influence the fuel cell performance, the use of a model representing the key physical and chemical characteristics of fuel cells is in general necessary for the optimization of cell design under specific application requirements. In all 3D models, the geometry is discretised into finite elements or volumes, and different physical and chemical properties are assigned to the regions of the cell components.

As discussed, single cell models are preferred for CFD analysis given the computing power requirements. Models thus focus on the representation of unit cells without considering stack or manifold effects. However, their complexity is high as they consider in different degrees of detail the coupling between the different physical, chemical and electrochemical phenomena occurring in a fuel cell.

A considerable body of literature exists for the modelling and simulation PEM fuel cells, as summarized in excellent review articles from Yao et al. (Yao et al., 2004), Biyikoglu (Biyikoglu, 2005), Cheddie and Munroe (Cheddie & Munroe, 2005), Siegel (Siegel, 2008), Djilali (Djilali, 2007), Weber and Newman (Weber & Newman, 2004), and Wang (Wang, 2004).

Wang (Wang, 2004) states the necessary steps in the development of the so-called CFCD (Computational Fuel Cell Dynamics) or CFD Fuel Cell modelling. The steps are:

- Physicochemical model development
- Advanced numerical algorithms
- Materials characterization
- Model validation at detailed levels

All these fronts must be developed in order to reach CFD models with a high degree of accuracy, confidence, and direct applicability to the understanding and design of advanced fuel cells.

3.6.2 Experimental techniques for PEM fuel cells

Fuel cell modelling comprises a tightly coupled combination of different physical and chemical phenomena: fluid flow, flow in porous media, multiphase flow, capillary forces, electrochemical reactions, heat transfer, membrane properties, etc., mostly acting in very small space and time scales. Experimental techniques are sometimes limited due to the small scales to be resolved, but however very detailed and useful information can be obtained.

Apart from the determination of the cell polarization curve, more detailed and advanced techniques are available for fuel cell research and design:

- Electrochemical Impedance Spectroscopy

Electrochemical Impedance Spectroscopy (EIS) or AC Impedance consists of sending an AC signal through the cell and monitor the response signal for a wide range of frequencies. The signal is recorded by a Frequency Response Analyzer (FRA). The imaginary and real impedances of the signal can be plotted in a Nyquist diagram (Z_{Im} vs Z_{Re}), where the characteristics of the diagram can be interpreted in order to assess the different phenomena occurring within of the cell. At high frequencies, the value of Z_{Re} at $Z_{Im} = 0$ represents the total resistance of cell (electronic of the bulk materials and contact resistances, and protonic of the membrane). This value is known as High Frequency Resistance (HFR). The first semicircle at medium frequencies represents the charge transfer resistance of the electrochemical reaction, and the next semicircle at low frequencies represents the mass transport resistance of the reactant reaching the active sites of the electrode (Mench, 2008).

EIS can be used for a wide range of fuel cell research activities (Wu et al., 2008a; Fouquet et al., 2006), including the determination of diffusion and reactant mass transport within the porous media.

- Current Density Mapping

This technique consists of measuring the in-plane spatial distribution of the generated current. Different technologies are used, such as the printed circuit board approach using segmented current collectors, anode catalysts, and anode GDLs. Segmented cells are necessary as the uneven current distribution resulting at the reaction sites in the catalyst layers can quickly equalizes in the bipolar plates, especially if the conductivity in the in-plane direction is larger than in the through-plane direction. To map the current distribution the measurement must be taken using a segmented cell, where each segment is isolated from the others.

The current distribution measured in this way provides useful information about the local phenomena influencing the cell performance such as reactant starvation or electrode flooding (Higier & Liu, 2010; Hwnag et al., 2008).

- Neutron Imaging

Neutron imaging is a more sophisticated technique where a beam of neutrons are sent through the cell. A neutron detector provide the contrast necessary to image the hydrogen and water in the cell, allowing for real-time images that provides information about the water distribution within the cell (Park et al., 2008b; Wu et al., 2008).

4. Fluid flow in fuel cell stacks

4.1 Description of a fuel cell stack: Main components

A single PEM fuel cell can provide a voltage of 0,6 -0,8 V when drawing a useful current. In order to sufficiently increase the voltage for practical operation, many cells have to be connected in series. A collection of fuel cells in series is known as a "stack". For stack, the total current is proportional to the active electrode area of each cell in the stack and it is the same through all cells. The total stack voltage is the sum of the individual cell voltage.

$$V_{stack} = \sum_{j=1}^{N_{cell}} V_j \approx N_{cell} V_{cell} \quad (17)$$

One of the most important components in the stack are the manifolds. The stack has manifolds for feeding hydrogen over the anode, manifolds for feeding oxygen (or air) over

the cathode and manifolds for discharge of gases (hydrogen and air or oxygen). The design of these manifolds is one of most difficult engineering design in fuel cell stacks. The manifolds must have the following three characteristics (Mench, 2008):

- The manifold must be compact in order to enable high power density stacks. The drawback is that the pressure loss increases when the volume of the manifold decrease.
- The manifold must carry the same amount of flow to each cell. A flow maldistribution decreases the performance of the stack. The flow inside an exit manifold is difficult to calculate with analytical expressions because the uneven electrochemical reaction in each cell. The presence of water liquid droplets in the flow field increases the pressure drop for droplet removal, and produces transient pressure variations between the plates in the stack.
- The manifold must be properly sealed to prevent mixing of air, fuel and coolant.

Another design issue is the need to remove the heat produced by a PEM fuel cell, and prevent a too high increase of the temperature in the stack. Therefore it is necessary to use a refrigerant flow like air or water. For this reason, an additional plate is inserted to provide a coolant path between each cell, or the cooling path could be integrated in to the bipolar plate. The result is a block in which the electric current must pass efficiently through the cells.

An important approach to determine the flow and the pressure distribution within the stack is the flow network analysis. Various authors (Baschuk & Li, 2004; Ma et al., 2002; Park & Li, 2006; Karimi et al., 2005; Chang et al., 2006; Kohn et al., 2003) have developed this type of mathematic method to resolve this problem.

4.2 Flow network approach

Apart from ageing and durability issues, the performance degradation in a fuel cell stack is known to be caused by (Barbir, 2005):

- Non uniform distribution of reactant to each cell.
- Non uniform distribution of reactant inside each cell.
- Maintenance of required temperature in each cell.
- High resistive losses.
- Leak of reactant gases (internal between the cells or external).
- Mechanical sturdiness (internal pressure, including thermal expansive, external forces during handling and operation, including shocks and vibrations).

The fuel cell performance is sensitive to the flow rate of reactants therefore it is necessary that each cell in a stack receives approximately the same amount of reactant gases. It is important that the manifolds that feed the gases to the cells and the manifolds that collect the unused gases are properly sized. The cross-sectional area of the manifolds determines the velocity of gas flow and the pressure drop. As a rule of thumb, the pressure drop through the manifolds should be an order of magnitude lower than the pressure drop through each cell in order to ensure uniform flow distribution (Barbir, 2005).

The problem of flow and pressure distribution within the stack is coupled, and for this reason the mass and pressure distribution must be solved simultaneously (Baschuk & Li, 2004). Some mathematical techniques and computer software have been developed for solving the problem of flow and pressure distribution within the stack (Baschuk & Li, 2004; Ma et al., 2022; Park & Li, 2006; Karimi et al., 2005; Chen et al., 2007; Chang et al., 2006; Kohn et al., 2003).

In order to resolve the flow network it is usual to consider two types of stack manifold configuration, U and Z configuration. The U and Z stack manifold configurations are shown in Figure 7.

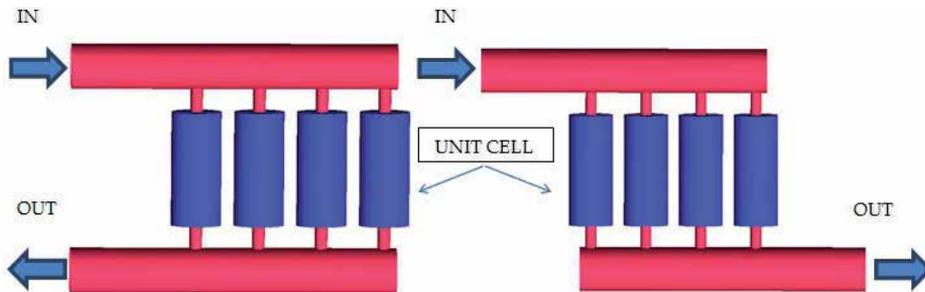


Fig. 7. U and Z stack manifold configurations.

The equation system to solve the flow network is composed in two blocks where the mass balance and the pressure loss are formulated respectively.

4.2.1 Equations of mass balance

First of all, it is important to define the different material flows that exist in the stack. Normally, there are six types of flow within the stack (see Fig. 8 for an explanation of the symbols):

- \dot{m}_{in}^{stack} : Total mass flow at the inlet of the stack (I).
- $\dot{m}_{cell,in}^j$: Mass flow that enters in the cell j (II).
- \dot{m}_{top}^j : Mass flow in the top section of the manifold in cell j (III).
- $\dot{m}_{cell,out}^j$: Mass flow that leaves the cell j (IV).
- \dot{m}_{bottom}^j : Mass flow in the bottom section of the stack manifold in cell j (V).
- $\Delta\dot{m}_{consumed}$: Mass flow consumed in the catalyst layer (VI).

In Figure 8, a schematic for Z stack manifold configuration is shown where all the fluid flows existing in a stack are represented.

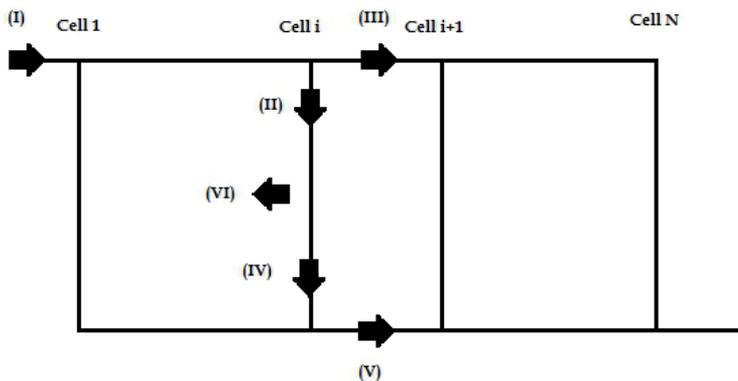


Fig. 8. Schematic for Z stack manifold configuration

To determine each mass flow is necessary to solve the following system of equations:

$$\dot{m}_{in}^{stack} = \sum_{j=1}^{N_{cell}} \dot{m}_{cell,in}^j \quad (18)$$

$$\dot{m}_{top}^i = \dot{m}_{in}^{stack} - \sum_{j=1}^i \dot{m}_{cell,in}^j \quad (19)$$

$$\dot{m}_{bottom}^i = \begin{cases} \sum_{j=1}^i \dot{m}_{cell,out}^j & \text{for Z configuration} \\ \sum_{j=1+1}^{N_{cell}} \dot{m}_{cell,out}^j & \text{for U configuration} \end{cases} \quad (20)$$

$$\dot{m}_{cell,in}^j = \dot{m}_{cell,out}^j - \Delta \dot{m}_{consumed} \quad (21)$$

$$\Delta \dot{m}_{consumed} = \begin{cases} \frac{IA_{cell}}{2F} M_{H_2} & \text{In the anode} \\ \frac{IA_{cell}}{4F} M_{O_2} - \frac{IA_{cell}}{2F} M_{H_2O} & \text{In the catode} \end{cases} \quad (22)$$

Where I (A) is the current density, F(C/mol), is the Faraday constant, A_{cell} (m^2) is the active area and M_i (kg/kmol) is the molecular weight of species 'i'.

4.2.2 Pressure loss

The pressure balance between two cells that are connected can be formulated as:

$$\Delta P_{cell}^j - \Delta P_{top\ manifold}^j - \Delta P_{cell}^{j+1} - \Delta P_{bottom\ manifold}^j = 0 \text{ for U configuration} \quad (23)$$

$$\Delta P_{cell}^j - \Delta P_{top\ manifold}^j - \Delta P_{cell}^{j+1} + \Delta P_{bottom\ manifold}^j = 0 \text{ for Z configuration} \quad (24)$$

Where ΔP_{cell}^j is the pressure loss within the single cell, $\Delta P_{top\ manifold}^j$ is the pressure loss within the top of the manifold and $\Delta P_{bottom\ manifold}^j$ is the pressure loss within the bottom of the manifold.

In the next sections it is shown how to calculate the manifold pressure loss and cell pressure loss respectively.

4.2.2.1 Manifold pressure loss

The pressure loss in the top or bottom can be calculated with the following expression:

$$\Delta P = \Delta P_m + \Delta P_f \quad (25)$$

Where ΔP_m and ΔP_f are the pressure loss due to the change in momentum of the fluid and to the wall friction respectively.

To calculate the pressure loss due to the change in momentum of the fluid, the following expression can be used:

$$\Delta P_m = \frac{\dot{m}}{A_{manifold}} (v_{out} - v_{in}) \quad (26)$$

Where \dot{m} (kg/s) is the flow in the section of the manifold, $A_{manifold}$ (m^2) is the manifold cross-sectional area, and v (m/s) is the velocity of the flow in the manifold.

On the other hand, to calculate the pressure loss due to the wall friction, the following expression can be used:

$$\Delta P_f = \frac{2C_f L_{manifold} \rho_{ave} (v_{ave})^2}{\theta_h^{manifold}} \quad (27)$$

Where C_f is the friction coefficient, L_{manifold} (m) is the length of the manifold section, ρ_{ave} (kg/m^3) is the average of the inlet and outlet density, v_{ave} (m/s) is the average of the inlet and outlet velocity, and $\theta_h^{\text{manifold}}$ (m) is the hydraulic diameter of the stack manifold. To calculate the friction coefficient, various authors (Baschuk & Li, 2004) use the following equation that depend of the Reynolds number.

$$C_f = \begin{cases} 16(Re_{\theta_h})^{-1} & \text{if } Re_{\theta_h} \leq 2000 \\ 0,079(Re_{\theta_h})^{-1/4} & \text{if } Re_{\theta_h} \geq 4000 \end{cases} \quad (28)$$

When the Reynolds number is between 2000 and 4000, the flow is neither laminar nor fully turbulent. In this case, it is necessary to interpolate in Equation 28 in order to obtain the friction coefficient.

4.2.2.2 Cell pressure loss

The method to calculate the cell pressure loss is similar to the manifold pressure loss. The only different is that the flow is modified due to the reaction that is produced within the cell. Hereby, the pressure loss due to momentum change can be calculated by the following expression:

$$\Delta P_m = \frac{1}{A_{\text{flow channel}}} (\dot{m}_{\text{cell out}} v_{\text{cell out}} - \dot{m}_{\text{cell in}} v_{\text{cell in}}) \quad (29)$$

Where $A_{\text{flow channel}}$ (m^2) is the cross-sectional area of a flow channel in the cathode or anode. The expression to calculate the loss pressure due to the wall friction is the same that have been show in section 1.2.2.1, it is only necessary to replace L_{manifold} with $L_{\text{flow channel}}$ (length of a flow channel in the cathode or anode) and $\theta_h^{\text{manifold}}$ with $\theta_h^{\text{flow channel}}$ (hydraulic diameter of the flow channel in the anode or cathode)

4.3 Stack heat transfer

There are many approaches to calculate the heat transfer in a stack. A simple general energy balance in a stack can be formulated as Energy of fuel reacted= Electricity generated + Heat generated (Barbir, 2005):

$$Q_{\text{reactant}} = W_{\text{electricity}} + Q_{\text{generated}} \quad (30)$$

The heat generated within a cell can be calculated by the following expression:

$$Q_{\text{generated}} = n_{\text{cell}} I (E_{\text{th}} - E_{\text{cell}}) = Q_{\text{conduction}} + Q_{\text{convection}} + Q_{\text{coolant}} \quad (31)$$

Where n_{cell} is the number of cells on the stack, I is the current (A), E_{th} (V) is the thermoneutral voltage, E_{cell} (V) is the real voltage in each cell, $Q_{\text{conduction}}$ (W) is the conduction heat transfer through the cell, $Q_{\text{convection}}$ (W) is the convection heat transfer between the reactants and bipolar plates with the environment, and Q_{coolant} (W) is the heat dissipation.

The heat conduction transfer can be calculated according to the Fourier Law:

$$Q_{\text{conduction}} = kA \frac{dT}{dx} \quad (32)$$

Where k ($\text{W}/(\text{mK})$) is the thermal conductivity, A (m^2) is the cross-sectional area, T (K) is the temperature, and x (m) is the direction through a finite cross section.

The convective heat transfer between each cell to the environment can be formulated by the following expression.

$$Q_{convection} = A_{environment} h (T_{cell} - T_{environment}) \quad (33)$$

Where h ($W/(m^2K)$) is the convection transfer coefficient, A (m^2) is the area for heat transfer and T (K) is the temperature.

The heat transfer from the bipolar plate to the coolant flow can be estimated by consideration of an energy balance on the fluid:

$$Q_{coolant} = \dot{m}_{coolant} C_{p_{coolant}} (T_{out_{coolant}} - T_{in_{coolant}}) \quad (34)$$

Using the equations described above, it is possible to calculate the heat transfer and temperature field in the stack. The main issue is that the temperature in each cell is not uniform, and it is complicated to determinate the precise temperature distribution of each component in each cell of the stack. For this reason, a lot of mathematical models usually assume that the temperature is constant in each cell (Park & Li. 2006). To calculate the temperature field in the stack, empirical methods or Computational Fluid Dynamics tools (CFD) are typically used (Liu et al., 2006; Shimpalee et al. 2009). The main problem to obtain the temperature field by CFD tools is the high computational expense that is required.

In summary, the final performance has a strong dependence on the flow and temperature distribution within the stack. For this reason, mathematical models are developed to estimate the flow and temperature fields inside the stack. In particular, the flow network approach is a well-established mathematical method able to estimate the final stack performance. Additionally, both reactant and coolant channels and manifolds must be appropriately designed for the enhanced operation of the stack.

5. Conclusion

In this chapter, the main characteristics of the fluid flow in Polymer Electrolyte Membrane (PEM) Fuel Cells have been addressed, both in single cells and in the complete stack. The fluid flow is highly coupled with other relevant phenomena in PEMFCs, such as electrochemical reactions and electric current transport, so that an analysis approach including all the phenomena is needed. Some techniques have been introduced, both at numerical (mainly analytical expressions and Computational Fluid Dynamics) and experimental levels.

6. Acknowledgment

Authors gratefully acknowledge InTech for their kind invitation to contribute with this chapter, and INTA (National Institute for Aerospace Technology) for the close collaboration in PEMFC CFD modelling and experimental validation.

7. References

- Barbir, F. (2005). *PEM Fuel Cells: Theory and Practice*. Elsevier Academic. ISBN 9780120781423, Amsterdam, London.
- Barreras, F.; Lozano, A.; Valiño, L.; Marín, C. & Pascau, A. (2005). Flow distribution in a bipolar plate of a proton exchange membrane fuel cell: experiments and numerical simulation studies. *Journal of Power Sources*. Vol.144, No.1, (2005), pp. 54-66, ISSN ISSN 0378-7753

- Baschuk, J & Li, X. (2004). Modeling of polymer electrolyte membrane fuel cells stacks based on a hydraulic network approach. *International Journal Energy Research*. Vol.28, No.8, (June 2004), pp.697-724, ISSN 1099-114X
- Biyikoglu, A. (2005). Review of proton exchange membrane fuel cell models, *International Journal of Hydrogen Energy*, Vol.30, No.11, (2005), pp. 1181-1212, ISSN 0360-3199
- Chang, P; St-Pierre, J; Stumper, J & Wetton, B. (2006). Flow distribution in a proton exchange membrane fuel cells stacks. *Journal of Power Sources*, Vol.162, No.1, (November 2006), pp.340-355, ISSN 0378-7753
- Cheddie, D. & Munroe N. (2005). Review and comparison of approaches to proton exchange membrane fuel cell modeling. *Journal of Power Sources*. Vol.147, No.1-2, (2005), pp. 72-84, ISSN 0378-7753
- Chen, C; Jung, S & Yen,S. (2007). Flow distribution in the manifold of a PEM fuel cell stack. *Journal of Power Sources*, Vol.173, No.1, (November 2007), pp.249-263, ISSN 0378-7753
- Dai, W.; Wang, H.; Yuan, X.-Z.; Martin, J.J.; Yang, D.; Qiao, J. & Ma, J. (2009). A review on water balance in the membrane electrode assembly of proton exchange membrane fuel cells. *International Journal of Hydrogen Energy*. Vol.34, No.23, (2009), pp. 9461-9478, ISSN 0360-3199
- Department of Energy. (2004). *Fuel Cell Handbook*. (7th edition), ISBN 26507-0880, Morgantown, West Virginia.
- Djilali, N. (2007). Computational modelling of polymer electrolyte membrane (PEM) fuel cells: Challenges and opportunities. *Energy*. Vol.32, No.4, (2007), 269-280, ISSN 0360-5442
- Fouquet, N.; Doulet, C.; Nouillant, C.; Dauphin-Tanguy, G. & Ould-Bouamama, B. (2006). Model based PEM fuel cell state-of-health monitoring via AC impedance measurements. *Journal of Power Sources*. Vol.159, No.2, (2006), pp. 905-913, ISSN 0378-7753
- Higier, A. & Liu, H. (2010). Optimization of PEM fuel cell flow field via local current density measurement. *International Journal of Hydrogen Energy*. Vol.35, No. 5, (2010), pp. 2144-2150, ISSN 0360-3199
- Hwnag, J.J.; Chang, W.R.; Peng, R.G.; Chen, P.Y. & Su, A. (2008). Experimental and numerical studies of local current mapping on a PEM fuel cell. *International Journal of Hydrogen Energy*. Vol.33, No.20, (2008), pp. 5718-5727, ISSN 0360-3199
- Iranzo, A.; Muñoz, M.; Rosa, F. & Pino, F. (2010). Numerical model for the performance prediction of a PEM fuel cell. Model results and experimental validation. *International Journal of Hydrogen Energy*. Vol.35, No.20, (October 2010), pp. 11533-11550, ISSN 0360-3199.
- Iranzo, A.; Muñoz, M.; Pino, F.; Rosa, F. (2011a). Update on Numerical model for the performance prediction of a PEM fuel cell. *International Journal of Hydrogen Energy*. Vol.36, No.15, (July 2011), pp. 9123-9127, ISSN 0360-3199.
- Iranzo, A.; Muñoz, M.; Pino, J.; Rosa, F. (2011b). Non-dimensional analysis of PEM fuel cell phenomena by means of AC impedance measurements. *Journal of Power Sources*. Vol.196, (2011), pp.4264-4269, ISSN 0378-7753
- Jiao, K. & Li, X. (2011). Water transport in polymer electrolyte membrane fuel cells. *Progress in Energy and Combustion Science*. Vol.37, No.3 (2011), pp. 221-291, ISSN 0360-1285

- Karimi, G; Baschuk, J & Li, X. (2005). Performance analysis and optimization of a PEM fuel cell stacks using flow network approach. *Journal of Power Sources*. Vol.147, No.1-2, (January 2005), pp.162-177, ISSN 0378-7753
- Kohn, J; Seo, H; Lee, C; Yoo, Y & Lim, H. (2002). Pressure and flow distribution in an internal gas manifolds of a fuel-cell stack. *Journal of Power Sources*. Vol.115, No.1, (March 2003), pp.54-65, ISSN0378-7753
- Larminie, J. & Dicks, A. (2003). Fuel cell systems explained (second edition), John Wiley & Sons, Ltd, ISBN 0-470-85857-X, Chichester, England
- Li X. & Sabir I. (2005). Review of bipolar plates in PEM fuel cells: Flow-field designs. *International Journal of Hydrogen Energy*. Vol.30, No.4 (2005), pp. 359-371. ISSN 0360-3199
- Li, H.; Tang, Y.; Wang, Z.; Shi, Z.; Wu, S.; Song, D.; Zhang, J.; Fatih, K.; Wang, H.; Liu, Z.; Abouatallah, R. & Mazza, A. (2008). A review of water flooding issues in the proton exchange membrane fuel cell. *Journal of Power Sources*. Vol.178, No.1, (2008), pp. 103-117, ISSN ISSN 0378-7753
- Litster, S. & Djilali, N. (2005). Two-phase transport in porous gas diffusion electrodes, In: *Transport Phenomena in Fuel Cells*, Sundén, B & Faghri, M., pp. 175-213, WIT Press, ISBN, 1853128400, Southampton, UK; Billerica, USA.
- Liu, Z; Mao, Z, Wang, C; Zhuge, W & Zhang, Y. (2006). Numerical simulation of a mini PEMFC stack. *Journal of Power Sources*, Vol.160, No.2 SPECISS, (October 2006), pp.1111-1121, ISSN 0378-7753
- Ma,Z; Jeter,S & Abdel-Khalik, S. (2002). Flow network analysis application in fuel cells (2002). *Journal of Power Sources*,Vol.108, No. 1-2, (December 2001), pp. 106-112, ISSN 0378-7753
- Mench, M. (2008). Fuel Cell Engines, John Wiley & Sons, Inc, ISBN 978-0-471-68958-4, Hoboken, New Jersey
- Nam, J.H. & Kaviany M. (2003). Effective diffusivity and water-saturation distribution in single- and two-layer PEMFC diffusion medium. *International Journal of Heat and Mass Transfer*. Vol.46, No.24, (2003), pp. 4595-4611, ISSN 0017-9310
- Park,J & Li,X. (2006). Effect of flow and temperature distribution on the performance of a PEM fuel cell stack. *Journal of Power Sources*. Vol.162, No.1, (July 2006), pp.444-459, ISSN 0378-7753
- Park, J.; Li, X.; Tran, D.; Abdel-Baset, T.; Hussey, D.S.; Jacobson, D.L. & Arif, M. (2008). Neutron imaging investigation of liquid water distribution in and the performance of a PEM fuel cell. *International Journal of Hydrogen Energy*. Vol.33, No. 13,(2008), pp. 3373-3384, ISSN 0360-3199
- Shimpalee, S; Ohashi, M; Van Zee, J; Stoeckmann, C; Sadeler, C & Hebling, C. (2009). Experimental and numerical studies of portable PEMFC stack. *Electrochimica Acta*. Vol.54, No. 10, (April 2009), pp.2899-2911, ISSN 0013-4686
- Siegel, C. (2008) Review of computational heat and mass transfer modeling in polymer electrolyte-membrane (PEM) fuel cells. *Energy*. Vol.33, No.9, (2008), pp. 1331-1352, ISSN 0360-5442
- Springer, T.E.; Zawodzinski, T.A. & Gottesfeld, S. (1991). Polymer electrolyte fuel cell model. *Journal of the Electrochemical Society*. Vol.138, No.8, (1991), pp. 2334-2342, ISSN 013-4651

- Um, S.; Wang, C.Y. & Chen, K.S. (2000). Computational fluid dynamics modeling of proton exchange membrane fuel cells. *Journal of the Electrochemical Society*. Vol.147, No.12 (2000), pp. 4485-4493, ISSN 0013-4651
- Wang, C.Y. (2004). Fundamental models for fuel cell engineering, *Chemical Reviews*, Vol. 104, No.10, (2004), pp. 4727-4765, ISSN 0009-2665
- Weber, A.Z. & Newman, J. (2004). Modeling transport in polymer-electrolyte fuel cells. *Chemical Reviews*, Vol.104, No.10, (2004), pp. 4679-4726, ISSN 0009-2665
- Wu, J.; Zi Yuan, X.; Wang, H.; Blanco, M.; Martin, J.J. & Zhang J. (2008). Diagnostic tools in PEM fuel cell research: Part I Electrochemical techniques. *International Journal of Hydrogen Energy*, Vol.33, No.6, (2008), pp. 1735-1746, ISSN 0360-3199
- Wu, J.; Zi Yuan, X.; Wang, H.; Blanco, M.; Martin, J.J. & Zhang J. (2008) Diagnostic tools in PEM fuel cell research: Part II: Physical/chemical methods. *International Journal of Hydrogen Energy*, Vol.33, No.6, (2008), pp. 1747-1757, ISSN 0360-3199
- Yao, K.Z.; Koran, K.; McAuley, K.B.; Oosthuizen, P.; Peppley, B. & Xie, T. (2004). A review of mathematical models for hydrogen and direct methanol polymer electrolyte membrane fuel cells, *Fuel Cells*, Vol.4, No.1-2, (2004), pp. 3-29, ISSN 1615-6846

Heat Transfer Enhancement in Microchannel Heat Sink Using Nanofluids

P. Gunnasegaran, N.H. Shuaib, H.A. Mohammed,
M.F. Abdul Jalal and E. Sandhita
*Department of Mechanical Engineering, College of Engineering
Universiti Tenaga Nasional, Km7
Jalan Kajang-Puchong, Selangor
Malaysia*

1. Introduction

1.1 Introduction to microchannel heat sink (MCHS)

The microchannel heat sink (MCHS) cooling concept is first proposed by Tuckerman and Pease [1] in 1981. MCHSs are the most common and cost-effective hardware employed for the thermal management of MEMS devices. Tuckerman and Pease [1] pointed out that decreasing liquid cooling channel dimensions to the micron scale will lead to increase the heat transfer rate. They demonstrated experimentally a forty-fold improvement in heat-sinking capability in Si-base microchannels anodically bonded to Pyrex cover plates. Since then, intense research on MCHS have been conducted to study the heat transfer and fluid flow characteristics of MCHS.

There are two main configurations for the application of microchannel cooling which are direct cooling and indirect cooling. Direct cooling requires a direct contact between the surface to be cooled and the coolant fluid as illustrated in Fig.1.1a. This scheme reduces the thermal resistance between the surface and the coolant and thus, enhances the cooling effectiveness. However, electrical and chemical compatibility between the coolant and device itself needs to be ensured for this system to work [2]. An alternative to the above configuration is the use of a metallic heat sink to conduct the heat away from the device to a coolant which is forced through circular or noncircular grooves in the heat sink. Such an indirect cooling configuration shown in Fig.1.1b allows for a greater flexibility in coolant selection at the cost of increased thermal resistance between the device and the heat sink due to the heat diffusion resistance in the heat sink itself [3].

Microchannels are very fine channels of the width of a normal human hair and are widely used for electronic cooling purposes. In a MCHS, multiple microchannels are stacked together as shown in Fig.1.1 (b) which can increase the total contact surface area for heat transfer enhancement and reduce the total pressure drop by dividing the flow among many channels. Liquid or gas is used as a coolant to flow through these microchannels. The large surface area of MCHS enables the coolant to take away large amounts of energy per unit time per unit area while maintaining a considerably low device temperature. Using these MCHS, heat fluxes can be dissipated at relatively low surface temperatures.

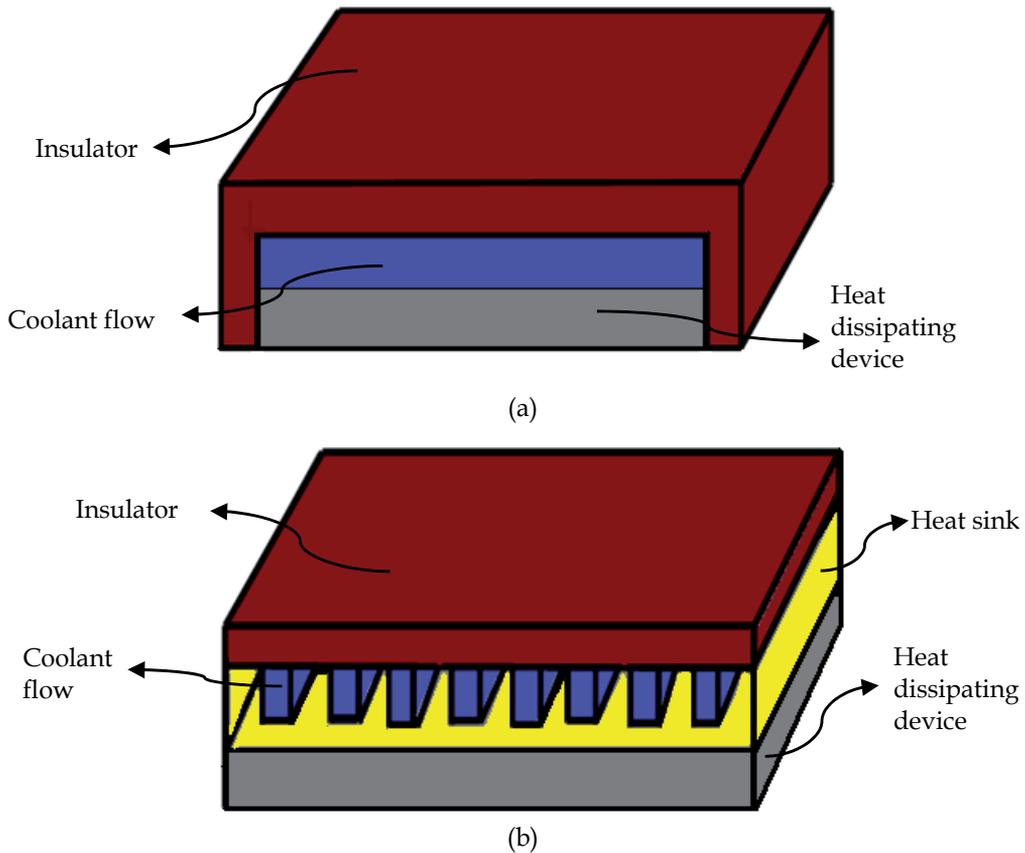


Fig. 1. Schematic diagram of the (a) direct cooling and (b) indirect cooling.

1.2 Cooling performance of MCHS

In order to drive the development of compact and efficient thermal management technology for advanced electronic devices, cooling devices have to be in light-weight, small in size and of high performance. Steinke and Kandlikar [4] presented a comprehensive review of friction factor data in microchannels with liquid flows. They indicated that entrance and exit losses need to be accounted for while presenting overall friction factors losses in microchannels. Most of the data that accounted for friction factor loss show good agreement with the conventional theory. They also provided a new procedure for correcting measured pressure drop to account for inlet and outlet losses.

Furthermore, three-dimensional fluid flow and heat transfer phenomena inside heated microchannels were investigated by Toh et al. [5]. They solved the steady laminar flow and heat transfer equations using a finite-volume method. It was found that the heat input lowers the frictional losses and viscosity leading to an increase in the temperature of the water, particularly at lower Reynolds numbers. Peng and Peterson [6,7] performed experimental investigations on the pressure drop and convective heat transfer of water flow in rectangular microchannels. They found that the cross sectional aspect ratio had a great influence on the flow friction and convective heat transfer both in laminar and

turbulent flows. Tiselj et al. [8] performed experimental and numerical analysis of the effect of axial conduction on the heat transfer in triangular microchannel heat sink. They pointed out that the bulk water and heated wall temperatures did not change linearly along the channel.

Inspired by the MCHS idea, new designs and modeling approaches of high performance cooling devices have been proposed, including using nanofluids as coolants in the MCHS study. Nanofluids are produced by dispersing nanometer-scale solid particles into base liquids such as water, ethylene glycol (EG), oils, etc. Lee et al. [9] used 38.4 nm of Al_2O_3 and 23.6 nm of CuO particles to enhance the thermal conductivity of water and EG. They showed that the enhancement percentage in thermal conductivity was not only a function of concentration and conductivities of the particles material and liquid, but it is also function of particle size and shape. Koo and Kleinstreuer [10] considered nanofluid flow in a representative microchannel, and conduction-convection heat transfer for different base fluids such as water and ethylene glycol with 20 nm CuO-nanoparticles. They come out with several suggestions which are a base fluid of high-Prandtl number such as ethylene glycol and oil should be used, using nanoparticles with high thermal conductivity are more advantageous, and a channel with high aspect ratio is desirable.

Chen and Huang [11] analyzed silicon microchannel heat sinks performance using nanofluids with a mixture of pure water and nanoscale Cu particles as coolants with various volume fractions. The MCHS with two specific geometries, one with $W_{\text{ch}}=100 \mu\text{m}$ and $L_{\text{ch}}=300 \mu\text{m}$, the other with $W_{\text{ch}}=57 \mu\text{m}$ and $L_{\text{ch}}=365 \mu\text{m}$, were examined. They found that the performances were greatly improved for these two specific geometries by using nanofluids as the coolants compared with pure water due to the increase in thermal conductivity of coolant and the nanoparticles thermal dispersion effect. The effectiveness of nanofluids for single-phase and two-phase heat transfer in microchannels is analyzed by Lee and Mudawar [12]. They indicated that the higher single-phase heat transfer coefficients are achieved in the entrance region of microchannels with increased nanoparticle concentration. However, the enhancement is weaker in the fully developed region. They suggested that nanoparticles should not be used in two-phase MCHS. This is because once boiling commences, particles begin to deposit into relatively large clusters near the channel exit due to localized evaporation. This clustering phenomenon quickly propagates upstream to fill the entire channel, preventing coolant from entering the heat sink and causing catastrophic failure of the cooling system. Jang and Choi [13] analyzed numerically the cooling performance of MCHS with nanofluids. They showed that the cooling performance of a MCHS with water-base nanofluids containing diamond (1%, 2 nm) at fixed pumping power of 2.25W was enhanced by about 10% compared with that of a MCHS with water.

There have been relatively few recent studies on nanofluid flow and heat transfer characteristics as comparing with those of pure fluid [14–16]. These studies showed that the heat transfer coefficient was greatly enhanced using nanofluid compared with pure fluid although there is a slight increase in pressure drop due to the presence of nanoparticles in MCHS operation. The enhancement depended on Reynolds number, particle volume fraction, and particle size and shape. It should be noted from the above literature review, however, that limited studies are available on nanofluid flow and heat transfer characteristics of rectangular shaped MCHS performance and this has motivated the present study.

The current study mainly focuses on 3D computational simulation of heat transfer and laminar liquid flow characteristics in MCHS. Following this introduction section, the governing equations and numerical model is explained, followed by studies on the geometrical parameters with various MCHS shapes (Section 3), effects of different nanoparticle volume fractions (Section 4), effects of different nanofluids types (Section 5), and effects of nanoparticle in different base fluids (Section 6).

2. Problem description and methodology

2.1 MCHS model

The physical configuration of the MCHS is schematically shown in Fig.2.1. The heat supplied to the MCHS substrate through a top plate is removed by flowing fluid through a number of 25 microchannels. This article focuses on heat transfer and liquid flow in three different cross sectional shapes of MCHS including rectangular, trapezoidal, and triangular. The dimensions of three different sets for each cross sectional shape of MCHS are given in Table 2.1-2.3.

	Case 1	Case 2	Case 3
H_{ch} (μm)	460	430	390
W_{ch} (μm)	180	280	380
L_c (μm)	10000	10000	10000
S (μm)	596	500	404
D_h (μm)	259	339	385
Number of channels	25	25	25

Table 2.1. Dimensions for three different sets of rectangular cross-section MCHS.

	Case 1	Case 2	Case 3
a (μm)	180	280	380
b (μm)	125	225	325
c (μm)	461	431	391
h (μm)	460	430	390
L_c (μm)	10000	10000	10000
S (μm)	596	500	404
D_h (μm)	229	318	370
Number of channels	25	25	25

Table 2.2. Dimensions for three different sets of trapezoidal cross-section MCHS.

	Case 1	Case 2	Case 3
k (μm)	180	280	380
l (μm)	469	452	434
h (μm)	460	430	390
L_c (μm)	10000	10000	10000
S (μm)	596	500	404
D_h (μm)	148	203	238
β	22.14°	36.07°	51.95°
Number of channels	25	25	25

Table 2.3. Dimensions for three different sets of triangular cross-section MCHS.

2.2 Governing equations

In the analysis of the entire domain of MCHS, it is necessary to set up the governing equations (continuity, momentum, and energy). For the specific case of heated flow through microchannels, the governing equations are solved with the following assumptions:

1. Both fluid flow and heat transfer are in steady-state and three-dimensional.
2. Fluid is in single phase, incompressible and the flow is laminar.
3. Properties of both fluid and heat sink material are temperature-independent.
4. All the surfaces of MCHS exposed to the surroundings are assumed to be insulated except the top plate where constant heat flux boundary condition, simulating the heat generation from electronic chip, is specified.

Thus, the governing equations in dimensionless form which are used in present study for heated MCHS can be written as [17]:

Continuity

$$\frac{\partial U}{\partial X} + \frac{\partial V}{\partial Y} + \frac{\partial W}{\partial Z} = 0 \quad (2.1)$$

Momentum

$$\text{X-Momentum: } \left(U \frac{\partial U}{\partial X} + V \frac{\partial U}{\partial Y} + W \frac{\partial U}{\partial Z} \right) = -\frac{d\hat{P}}{dX} + \frac{1}{\text{Re}} \left(\frac{\partial^2 U}{\partial X^2} + \frac{\partial^2 U}{\partial Y^2} + \frac{\partial^2 U}{\partial Z^2} \right) \quad (2.2a)$$

$$\text{Y-Momentum: } \left(U \frac{\partial V}{\partial X} + V \frac{\partial V}{\partial Y} + W \frac{\partial V}{\partial Z} \right) = -\frac{d\hat{P}}{dY} + \frac{1}{\text{Re}} \left(\frac{\partial^2 V}{\partial X^2} + \frac{\partial^2 V}{\partial Y^2} + \frac{\partial^2 V}{\partial Z^2} \right) \quad (2.2b)$$

$$\text{Z-Momentum: } \left(U \frac{\partial W}{\partial X} + V \frac{\partial W}{\partial Y} + W \frac{\partial W}{\partial Z} \right) = -\frac{d\hat{P}}{dZ} + \frac{1}{\text{Re}} \left(\frac{\partial^2 W}{\partial X^2} + \frac{\partial^2 W}{\partial Y^2} + \frac{\partial^2 W}{\partial Z^2} \right) \quad (2.2c)$$

Energy:

$$\left(U \frac{\partial \theta}{\partial X} + V \frac{\partial \theta}{\partial Y} + W \frac{\partial \theta}{\partial Z} \right) = \frac{1}{\text{Re.Pr}} \left(\frac{\partial^2 \theta}{\partial X^2} + \frac{\partial^2 \theta}{\partial Y^2} + \frac{\partial^2 \theta}{\partial Z^2} \right) \quad (2.3)$$

Where $X = \frac{x}{D_h}$, $Y = \frac{y}{D_h}$, $Z = \frac{z}{D_h}$, $U = \frac{u}{u_{in}}$, $V = \frac{v}{u_{in}}$, $W = \frac{w}{u_{in}}$, $\hat{P} = \frac{\Delta P}{\rho u_{in}^2}$ is the dimensionless pressure, and $\theta = \frac{T_f - T_i}{T_w - T_i}$ is the dimensionless temperature.

Given the complexity of these equations, computational methods of solving them are required. The methodology used in the present study to solve these governing equations is described in the following section.

2.3 Boundary conditions

Boundary conditions for all boundaries are specified for this simplified computational domain. Fig.2.1a shows the general MCHS computational model used in this study and Fig.2.1b shows the cross sectional shapes of channel considered in the study. Heat, supplied to the aluminum substrate through a top plate, is removed by flowing fluid through a number of microchannels.

At the entrance of the MCHS assembly ($z = 0$, from Fig.3.4a), two types of boundaries are encountered which are water flows through the microchannels and removes heat conducted to the surface of the MCHS. The remainder of the entrance is occupied by the aluminum substrate.

At the microchannels sections, the inlet fluid temperature was taken as 293K and the inlet fluid velocity was calculated using

$$u_{in} = \frac{Re \mu}{\rho D_h} \quad (2.4)$$

For rectangular cross-section MCHS,

$$D_h = \frac{4A}{P} = \frac{2H_{ch}W_{ch}}{H_{ch} + W_{ch}} \quad (2.5a)$$

where A is the channel flow area, P is the channel wet perimeter, H_{ch} is the channel height, and W_{ch} is the channel width.

For trapezoidal cross-section MCHS,

$$D_h = \frac{4A}{P} = \frac{2(a+b)h}{a+b+2c} \quad (2.5b)$$

where A is the channel flow area, P is the channel wet perimeter, a is the channel top width, b is the channel bottom width, c is the channel hypotenuse, and h is the channel height.

For triangular cross-section MCHS,

$$D_h = \frac{4A}{P} = \frac{2kh}{k+2l} \quad (2.5c)$$

where A is the channel flow area, P is the channel wet perimeter, k is the channel width, h is the channel height, and l is the channel hypotenuse.

The Reynolds number in this work was ranged from 100 to 1000. In calculating the velocity, the fluid is assumed to be evenly distributed into all microchannels. The transverse velocities at the inlet are assumed to be zero. On the aluminum substrate, the velocities are zero, and it is assumed to be an adiabatic surface.

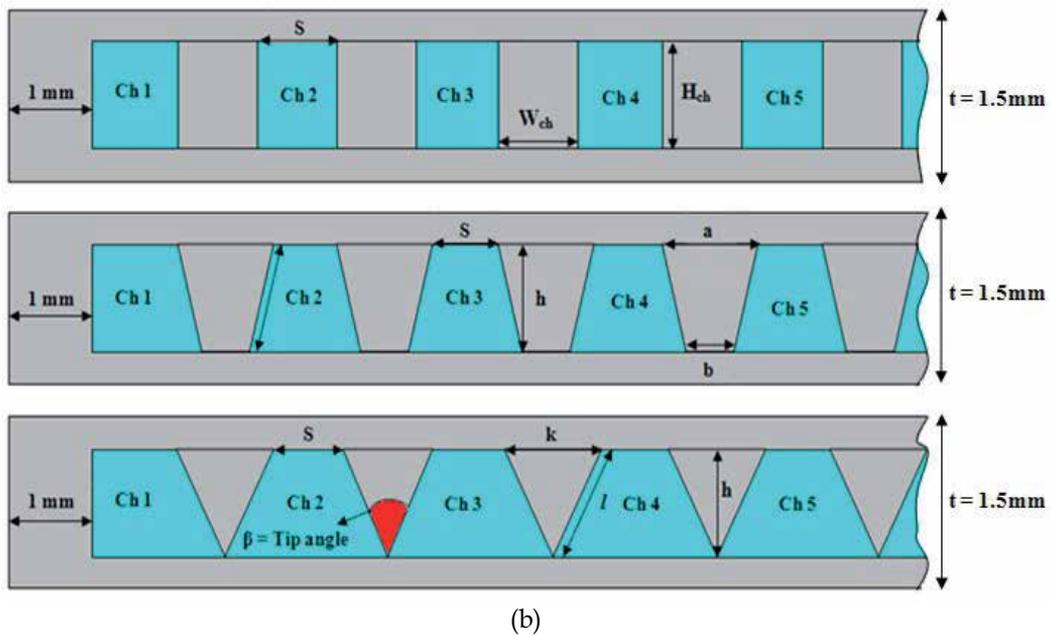
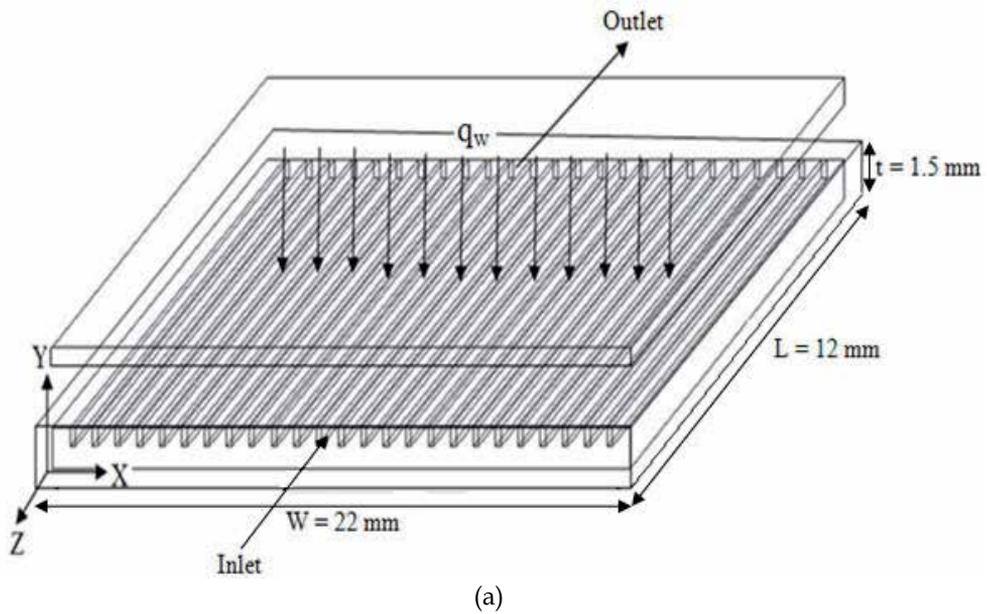


Fig. 2.1. (a) Schematic diagram of the MCHS (b) Section of the MCHS cross sectional shapes with its dimensions.

Boundary conditions at the inlet:

$$U = \frac{u}{u_{in}}, U = 1, \theta = 1 \quad (2.6a)$$

At the outlet:

$$P = P_{out}, \frac{\partial \theta}{\partial n} = 0 \quad (2.6b)$$

At the fluid-solid interface:

$$U = 0, \theta = \theta_s, -\kappa_s \frac{\partial \theta_s}{\partial n} = -\kappa \frac{\partial \theta}{\partial n} \quad (2.6c)$$

At the top plate:

$$q_w = -\kappa_s \frac{\partial \theta_s}{\partial n} \quad (2.6d)$$

In Eq. (2.6), U and θ are the dimensionless fluid inlet velocity and dimensionless temperature, respectively, P is the dimensionless pressure at the outlet, n is the direction normal to the wall or the outlet plane, and q_w is the heat flux applied at the top plate of the heat sink. The heat flux applied at the top plate was ranged from 100-1000 W/m².

2.4 Numerical solution using FVM

The governing conservation equations Eqs. (2) - (4) with the corresponding boundary conditions and equations for solid and fluid phases are simultaneously solved as a single domain conjugate problem using the standard finite volume method (FVM) with a hybrid differencing scheme [17]. The standard SIMPLE algorithm is used as the computational algorithm [18]. The iterations are continued until the sum of residuals for all computational cells became negligible (less than 10⁻⁷) and velocity magnitudes did not change from iteration to iteration. Because of the assumption of constant fluid properties and negligible buoyancy, the mass and momentum equations are not coupled to the energy equation. Therefore, the temperature field is calculated by solving the energy equation after a converged solution for the flow field is obtained by solving the momentum and continuity equations.

2.5 Numerical implementation, grid testing and code validation

The distribution of hexahedral cells in the computational domain is determined from a series of tests with different number of cells. For example, for the case of rectangular MCHS, computational cells with 1.9 × 10⁵, 2.8 × 10⁵, and 3.4 × 10⁵ grids are used to test the grid independence of the solution. The results are shown in Figs.2.2 and 2.3 which present the dimensionless temperature profiles and dimensionless pressure drop across the heat sink versus Reynolds number, respectively. The dimensionless temperature and dimensionless pressure are defined in Eq.3.1(Section 3) and Eq.5.1(Section 5), respectively. It can be seen that almost identical results obtained when 2.8 × 10⁵ and 3.4 × 10⁵ grids are used. Thus, based on the results shown in these figures, a computational cell with 2.8 × 10⁵ grids is employed for all the numerical computations in this study. Similar study was carried out for other MCHS shapes.

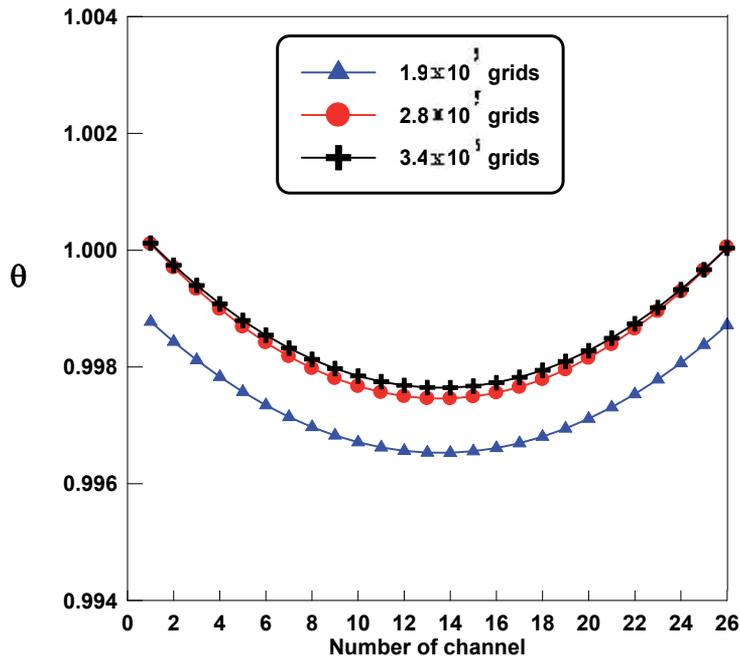


Fig. 2.2. Dimensionless temperature profiles for each channel in rectangular cross-section MCHS ($D_h = 0.259$ mm) using three different grids.

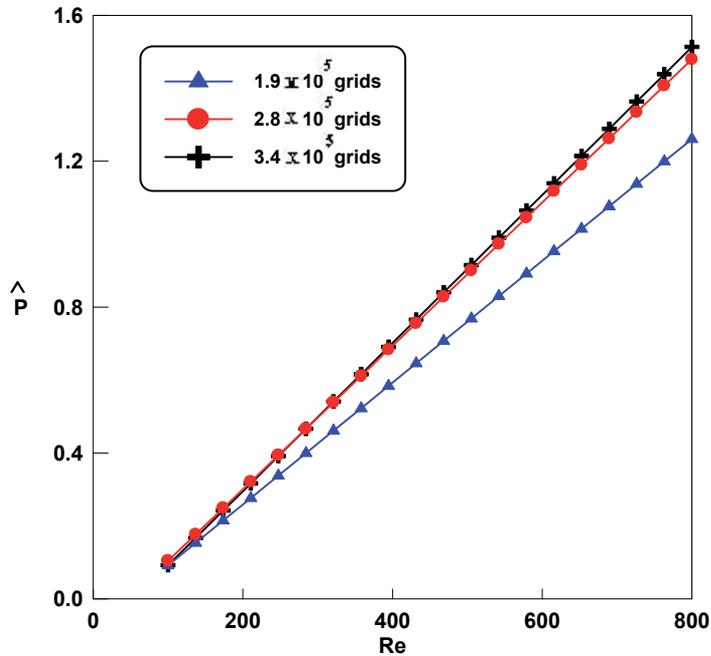


Fig. 2.3. Dimensionless pressure drop versus Reynolds number across the rectangular cross-section MCHS ($D_h = 0.259$ mm) using three different grids.

3. Effect of geometrical parameters of various MCHS shapes

This section investigates the effect of geometrical parameters on heat transfer and fluid flow characteristics for different cross sectional shapes of MCHS including rectangular, trapezoidal, and triangular with the dimensions as stated in Table 2.1-2.3, respectively. For this Section, water is used as the working fluid as the main aim is to investigate on the geometrical effects. For Section 4, 5, and 6, various nanofluids are used as the working fluid and only one particular cross sectional shape of MCHS with specified dimensions is considered for each Section. The dimensionless temperature profiles, heat transfer coefficient, friction factor, and thermal resistance which affected by geometrical parameters are discussed and presented.

3.1 Temperature profile

In the MCHS operation, the high temperature region should occur at the edge of the MCHS since there is no heat dissipation by fluid convection while the low-temperature region should occur in the region where microchannels are placed, especially at the middle regions of the MCHS due to the high heat transfer coefficient. To address this point, the dimensionless average temperature profiles of rectangular cross-section MCHS for different dimensions of the channel area at $Re = 500$ are presented in Fig.3.1. The dimensionless temperatures for each channel can be defined as:

$$\theta = \frac{T_f - T_i}{T_w - T_i} \quad (3.1)$$

Where T_f is the fluid temperature, T_i is the inlet temperature, and T_w is the wall temperature. It can clearly be seen from Fig.3.1 that the lower temperature profile occurs in the channels

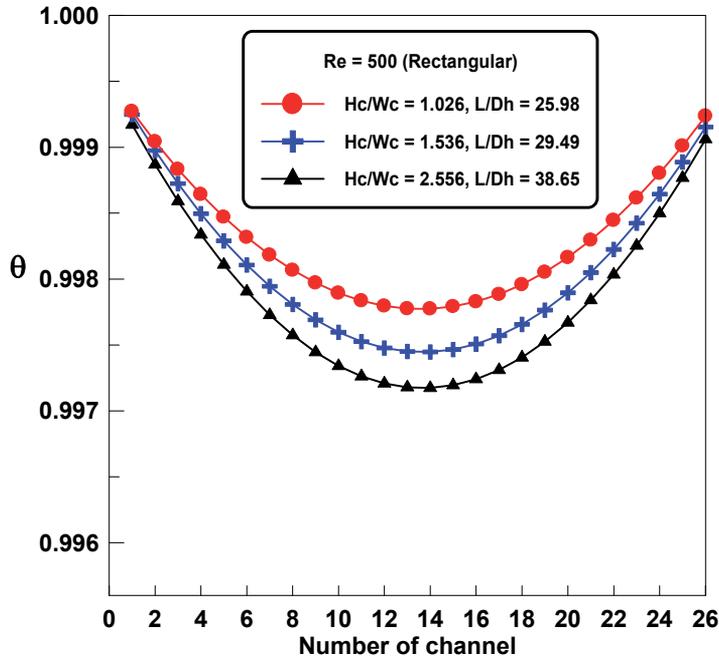


Fig. 3.1. Dimensionless average temperature profiles for each channel in rectangular cross-section MCHS.

located near the center of the MCHS. The validity of present result shows a similar trend with the result obtained by Chein and Chen [19]. For fluid in channels close to the edge of the MCHS, higher temperature profiles are observed due to the high heat transfer from the high-temperature edge of the MCHS.

3.2 Heat transfer coefficient

The computed averaged heat transfer coefficient in each channel of the rectangular cross-section MCHS for various hydraulic diameters is illustrated in Fig.3.2. The magnitude of heat transfer coefficient decreases with the increase in hydraulic diameter and the trends are the same for each cross sectional shape of the MCHS. This is due to the fact that lower pressure drop in larger hydraulic diameter corresponds to lower inlet velocity driven into the MCHS. Furthermore, due to the difference of the channel hydraulic diameter which is caused by the difference of channels' area and perimeter, the averaged heat transfer coefficient in each MCHS is also different under a given inlet velocity. For each type of the MCHS, the middle channel (channel number 14) has the highest averaged heat transfer coefficient value as expected. The averaged heat transfer coefficient value for other channels is seen to decrease depending on their distances from the wall. The averaged heat transfer coefficient distribution for all types of MCHS is almost symmetrical with respect to the centerline of the MCHS. It is shown that the heat transfer coefficient of the MCHS is greatly influenced by the hydraulic diameter of the channel, as the hydraulic diameter decreases, the heat transfer coefficient increases.

The computed average heat transfer coefficient versus Reynolds number for different cross sectional shapes of the MCHS is presented in Fig.3.3. It is apparent from this figure that the

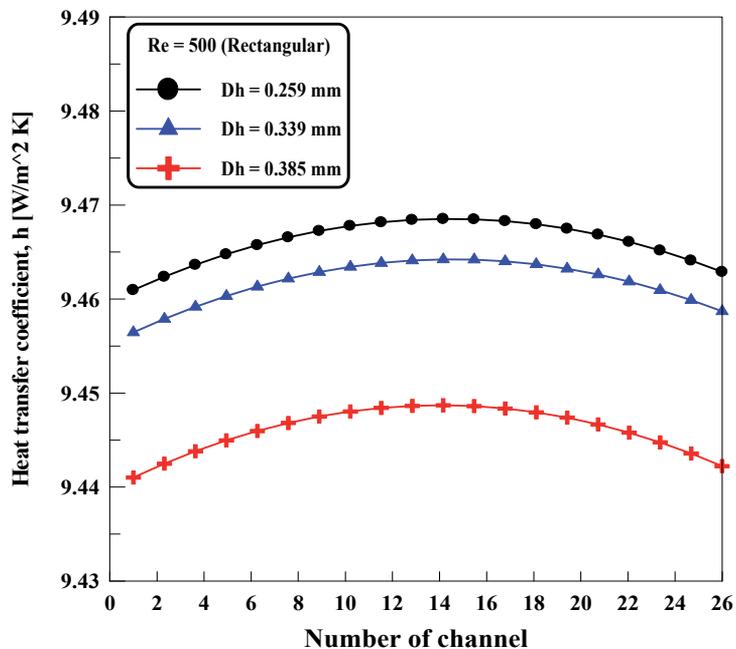


Fig. 3.2. Averaged heat transfer coefficient in each channel of the MCHS.

heat transfer coefficient for rectangular cross-section MCHS is the highest. The MCHS of triangular cross-section exhibits the lowest values of heat transfer coefficient. While, the heat transfer coefficient for trapezoidal cross-section MCHS is in between the rectangular and triangular.

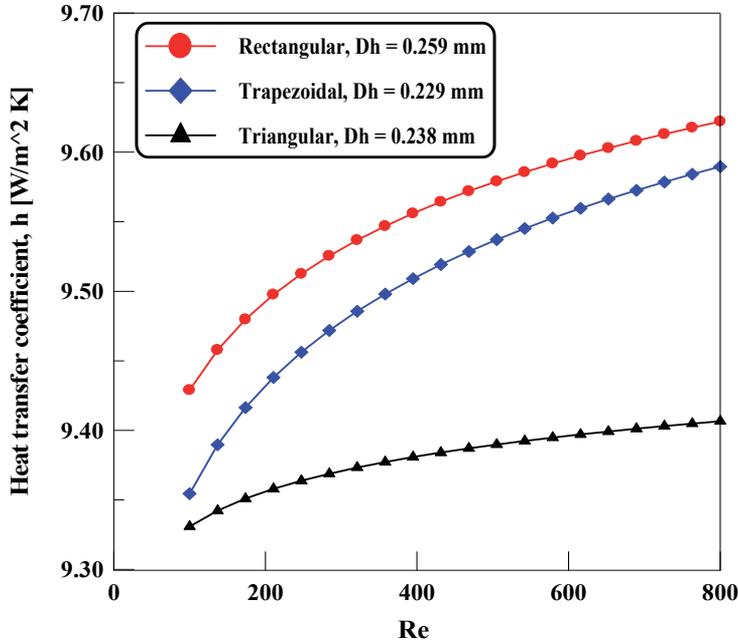


Fig. 3.3. Averaged heat transfer coefficient for different cross sectional shapes of the MCHS.

3.3 Friction factor

The effect of geometrical parameters on the friction factor, f for different microchannel shapes is discussed in this section. For the current study, the friction factor is calculated using Darcy equation [20]:

$$f = \frac{2D_h \Delta p}{\rho u_{in}^2 L_c} \quad (3.2)$$

Where D_h is the hydraulic diameter, Δp is the pressure drop, ρ is the density of water, u_{in} is the inlet velocity of water, and L_c is the length of channel.

Fig.3.4 shows the friction factor, f , at different W_o/H_c ratios for rectangular cross-section MCHS. It is clearly observed that f increases with the increase of W_o/H_c ratio of the channel. Therefore, the flow resistance will increase evidently when the ratio is increased. The reason behind this is that the increase in W_o/H_c ratio causes a decrease in the flow area and the pressure drop becomes significant. The validity of the present numerical results was also approved by Kandlikar et al. [21] results, where f increases with the increase of W_o/H_c ratio for rectangular cross-section MCHS.

Fig.3.5 shows f with three geometrical parameters including bottom-to-top width ratio (b/a), height-to-top width ratio (h/a), and length-to-hydraulic diameter ratio (L/D_h), which affect the friction and heat transfer in the trapezoidal cross-section MCHS. Fig.3.6 shows a comparison of predicted Poiseuille number which is the product of friction factor, f and

Reynolds number, fRe from the present numerical simulation for trapezoidal cross-section MCHS with the experimental work done by Wu and Cheng [22]. As noticed, a very good agreement is obtained between the two studies whereby it is shown that fRe rises with the increase of Reynolds number. As can be noticed from Fig.3.6, fRe increases linearly with the increase of Reynolds number due to the linear dependency of pressure drop on constant inlet velocity for a fixed hydraulic diameter in fully developed laminar region. Thus, fRe is linearly proportional to Reynolds number for fully developed channel flow.

The f with different tip angles of triangular cross-section MCHS is illustrated in Fig.3.7. It is clearly observed that f increases with the increases of tip angle of the channel. The validity of the present numerical simulation result was also approved by the results of Kandlikar et al. [21]. The results show that the flow behavior is very similar for all types of cross sectional shapes of MCHS where the friction factor decreases with the increase of Reynolds number.

In overall, From the Figs.3.4, 3.5, and 3.7, it is also apparent that f for rectangular shaped microchannel is most, for trapezoidal shaped microchannel is less, and for triangular shaped microchannel is least. This is due to the triangular shaped microchannel has a smallest area. For trapezoidal shaped microchannels, the flow area is changed suddenly from larger area to smaller area, thus the pressure drop would be large. Therefore, f value for trapezoidal shaped microchannels is also larger than the triangular shaped microchannels. For the rectangular shaped microchannels, the flow area is the largest which corresponds to the highest value of f . Therefore, f of water flowing in MCHS, having the same width and height but with different cross sectional shapes can be very much different due to the difference of the cross sectional shape and geometrical dimensions of the channels.

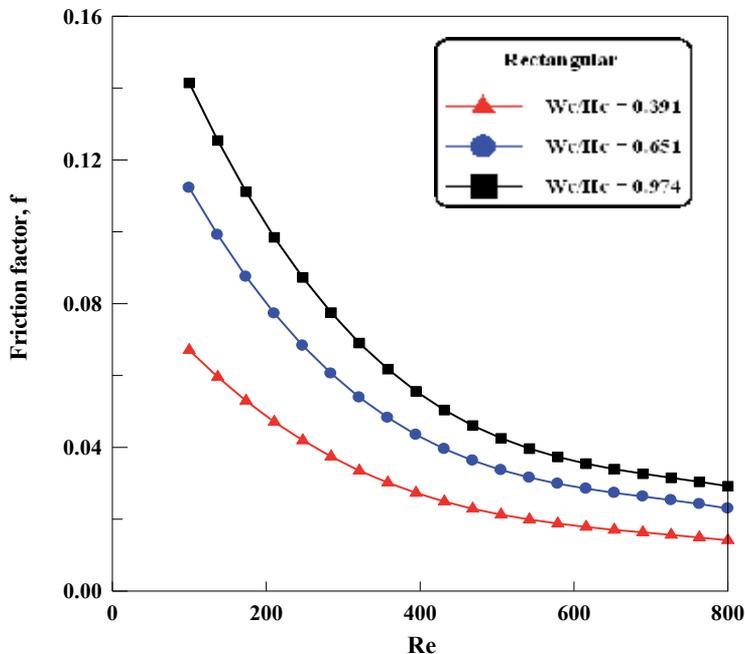


Fig. 3.4. Friction factor at different width-height (W_c/H_c) ratio for rectangular cross-section MCHS.

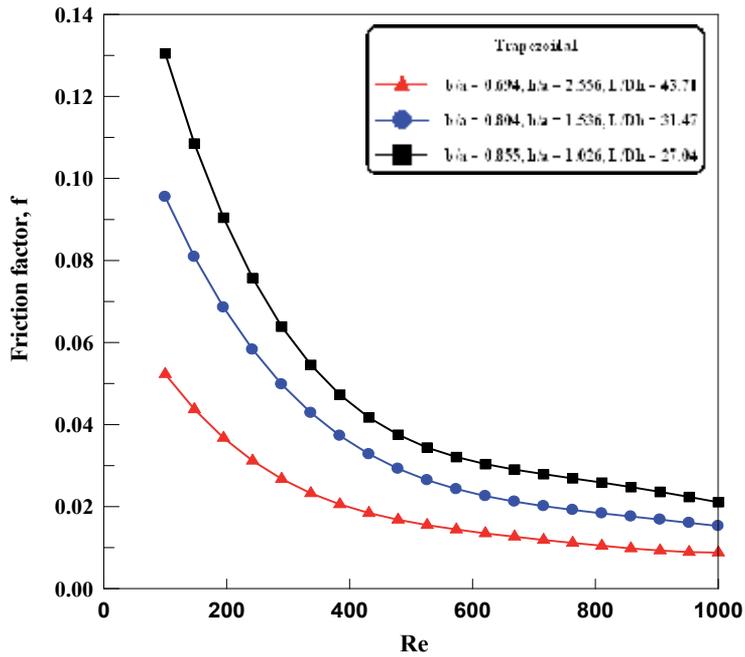


Fig. 3.5. Friction factor with three different geometrical parameters of trapezoidal cross-section MCHS.

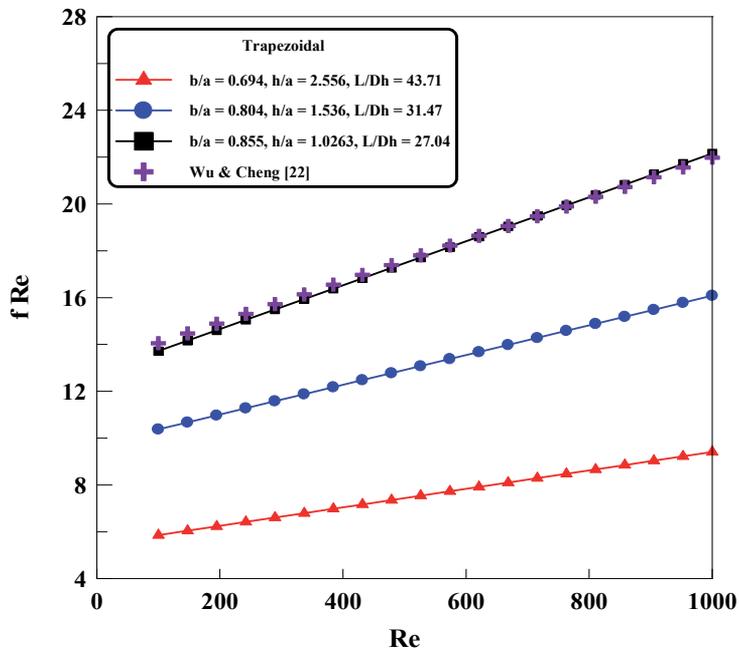


Fig. 3.6. Comparison of Poiseuille number of trapezoidal cross-section MCHS with [22].

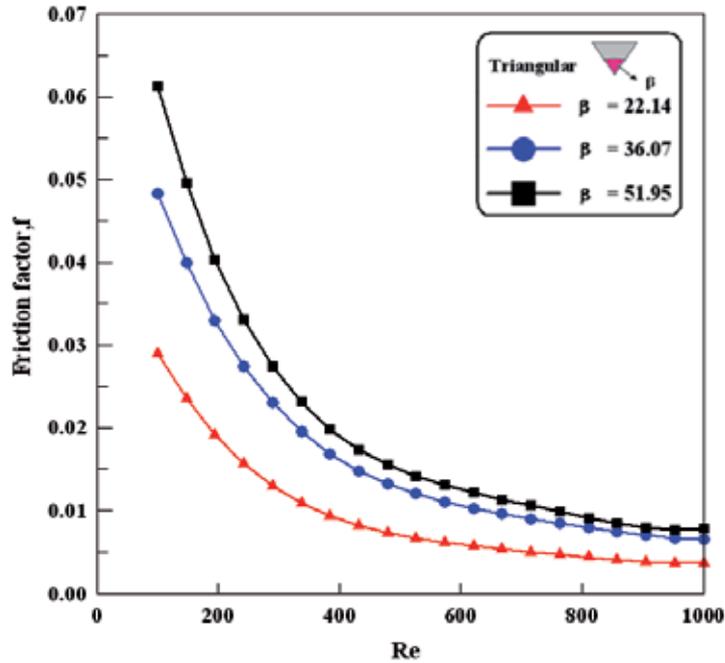


Fig. 3.7. Friction factor with different tip angles of triangular cross-section MCHS.

3.4 Thermal resistance

The performance of the MCHS is commonly presented by the thermal resistance (R_{th}) which is defined as:

$$R_{th} = \frac{T_{w,max} - T_{in}}{Q_{in}} \quad (3.3)$$

Another important parameter in MCHS operation is the pressure drop across the MCHS which relates to the coolant pumping power required. The pumping power is used to drive the coolant in MCHS operation. It is the product of the pressure drop across the heat sink, ΔP and volume flow rate, Q :

$$\text{Pumping power} = Q\Delta P \quad (3.4)$$

where Q is defined as:

For rectangular cross-section MCHS,

$$Q = N W_{ch} H_{ch} u_{in} \quad (3.5)$$

where N is the number of microchannels, W_{ch} is the width of channel, H_{ch} is the height of channel, and u_{in} is the inlet velocity of water.

For trapezoidal cross-section MCHS,

$$Q = 0.5 N (a + b) u_{in} \quad (3.6)$$

where N is the number of microchannels, k is the width of channel, h is the height of channel, and $u_{in,nf}$ is the inlet velocity of water.

For triangular cross-section MCHS,

$$Q = 0.5 N k h u_{in} \quad (3.7)$$

where N is the number of microchannels, k is the width of channel, h is the height of channel, and u_{in} is the inlet velocity of water.

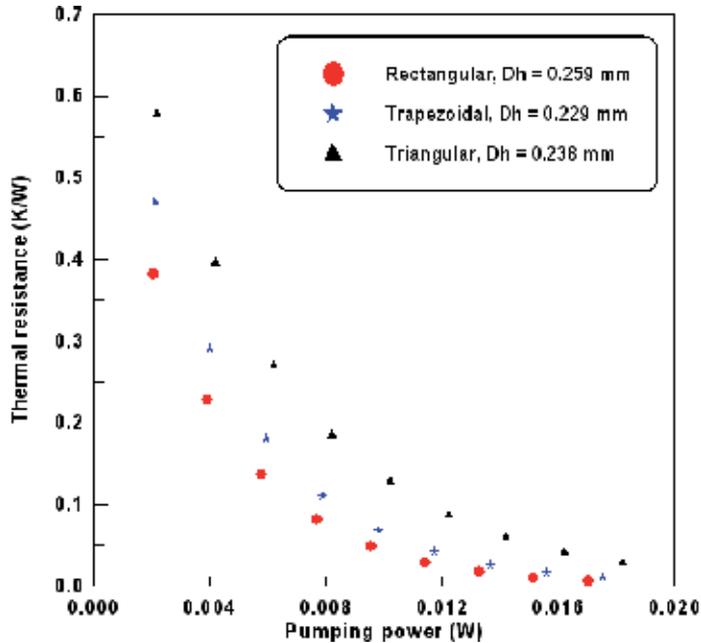


Fig. 3.8. Thermal resistance of MCHS with different cross-section shapes as function of pumping power.

The effect of using different cross sectional shapes of MCHS on thermal resistance, R_{th} versus pumping power is shown in Fig.3.8. The pumping power is determined from the calculated volume flow rate for rectangular, trapezoidal, and triangular cross-section shape using Eq.3.5-Eq.3.7, respectively, and computed pressure drop using Eq.3.4. It is evident from this figure that the R_{th} for rectangular cross-section MCHS is the lowest followed by trapezoidal and triangular cross-section MCHS. Thus, a MCHS with rectangular cross-section is capable for removing high heat flux due to its small difference between maximum wall temperature and inlet temperature at particular pumping power compared with that for MCHS with trapezoidal and triangular cross-sections.

4. Effect of nanoparticle volume fractions on MCHS performance

Nanofluids are produced from a suspension of nanoparticles with specified volume fraction in a conventional base fluid such as water, have been gaining interest recently due to its potential to greatly outperform traditional thermal transport liquids in MCHS [23]. However, the cooling effectiveness of using nanofluid as well as addressing other technical problems, such as agglomeration of the nanoparticles which yield to MCHS clogging due to the high particle volume fraction are vital to practical implementation of the nanofluid-

cooled MCHS. In the current study, the effect of using three different particle volume fractions on heat transfer and liquid flow characteristics in MCHS will be discussed. Since the cooling performance of MCHS with various cross sectional shapes is investigated in previous chapter, now a rectangular cross-section MCHS with specific dimensions as shown in Table 4.1 is considered. Results of interest such as temperature, heat transfer coefficient, friction factor, wall shear stress, and thermal resistance are reported in this section.

The forgoing thermophysical properties for the nanoparticle (alumina), base fluid (water), and the alumina-water nanofluid of three different volume fractions formulated are listed in Table 4.2. The density, thermal conductivity, and dynamic viscosity of the nanofluid appear significantly increased while the specific heat of nanofluid decreased with the increase of its particle volume fraction compared to pure water. The nanofluid of $\phi = 5\%$ alumina, for instance, has a relative increase of 14.88%, 15.17%, and 12.46% in the density, thermal conductivity, and dynamic viscosity and has a relative decrease of 14.14% in specific heat, respectively, compared to pure water.

D_h (μm)	H_{ch} (μm)	W_{ch} (μm)	L_{ch} (μm)	S (μm)
339.15	43	280	10000	500

Table 4.1. Dimensions of the rectangular cross-section MCHS.

Properties	Nanoparticle (alumina)	Base fluid (water)	Nanofluid (alumina-water)		
			$\phi = 1\%$	$\phi = 2.5\%$	$\phi = 5\%$
ρ (kg/m^3)	3970	998.2	1027.92	1072.5	114 .79
C_p ($\text{J}/\text{kg}\cdot\text{K}$)	765	4182	4050.03	3865.79	3590.55
k ($\text{W}/\text{m}\cdot\text{K}$)	40	0.613	0.631	0.658	0.706
μ (Ns/m^2)		0.001003	0.001028	0.001066	0.001128

Table 4.2. Thermophysical properties of alumina nanoparticles, water, and nanofluid at 293K.

4.1 Thermophysical properties of nanofluid

In this study, aluminum oxide with different particle volume fractions is used as the working fluid. The thermophysical properties required for flow as listed in Table 4.2 are calculated using the following equations [24, 25, 26, 27]:

$$\text{Density:} \quad \rho_{nf} = (1 - \phi) \rho_{bf} + \phi \rho_p \quad (4.1)$$

$$\text{Heat capacity:} \quad (\rho c_p)_{nf} = (1 - \phi) (\rho c_p)_{bf} + \phi (\rho c_p)_p \quad (4.2)$$

$$\text{Thermal conductivity: } k_{nf} = \frac{k_p + 2k_{bf} + 2(k_p - k_{bf})\varphi}{k_p + 2k_{bf} - (k_p - k_{bf})\varphi} k_{bf} \quad (4.3)$$

$$\text{Viscosity: } \mu_{nf} = \mu_{bf} (1 + 2.5\varphi) \quad (4.4)$$

Where φ is particle volume fraction, the subscript “*nf*” refers to nanofluid, “*bf*” refers to base fluid, and “*p*” refers to particle.

4.2 Temperature profile

The dimensionless temperature distribution at each channel of MCHS at $Re = 900$ for various particle volume fractions and different heat fluxes is shown in Figs.4.1–4.3. The dimensionless temperature is calculated using Eq.3.1 (Section 3). The presence of nanoparticles has an effect of reducing the temperature as the particle volume fraction of nanofluids increasing due to its higher dynamic viscosity and lower heat capacity compared to pure water. However, these conditions only can be greatly achieved for the case of heat flux = 1000 W/m^2 compared to 100 W/m^2 and 500 W/m^2 . For case 100 W/m^2 and 500 W/m^2 , there is no significant differences between nanofluids and pure water, and the results are inconsistent for temperature distributions. Therefore, the presence of nanoparticles could enhance cooling of MCHS under the extreme heat flux conditions, the same trend was reported by Tsai and Chein [15] and Murshed et al. [28].

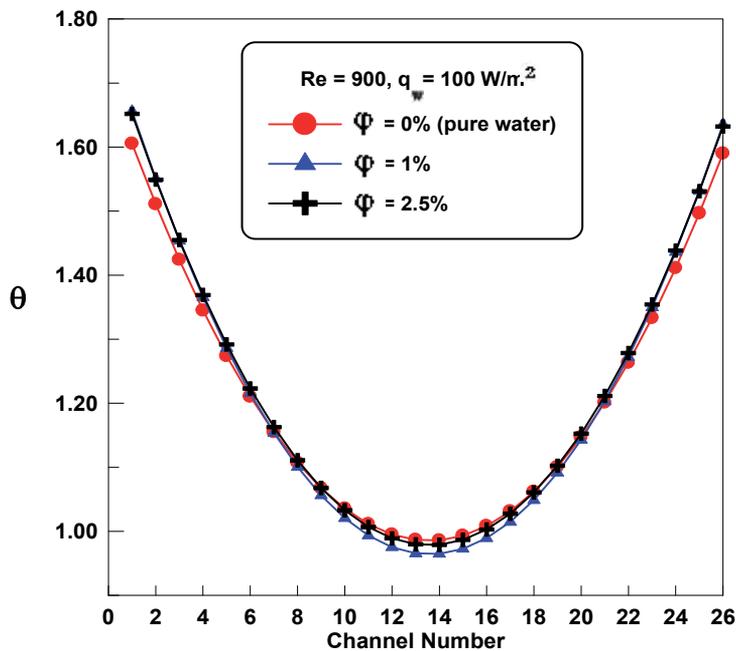


Fig. 4.1. Dimensionless temperature versus number of channels for 100 W/m^2 .

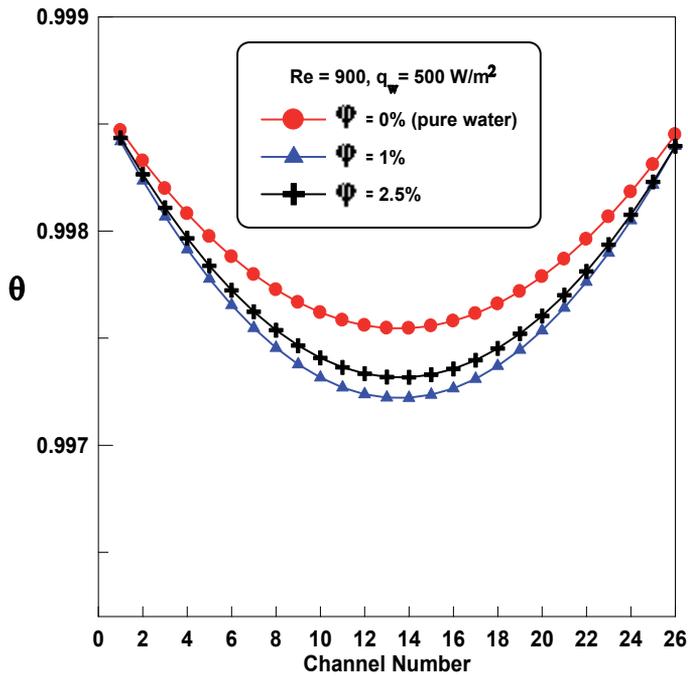


Fig. 4.2. Dimensionless temperature versus number of channels for 500 W/m^2 .

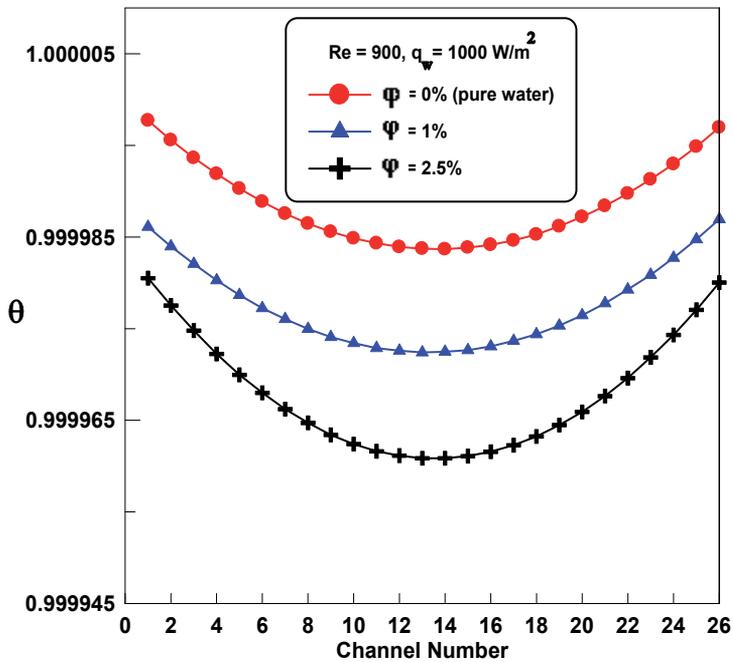


Fig. 4.3. Dimensionless temperature versus number of channels for 1000 W/m^2 .

4.3 Heat transfer coefficient

The effect of increasing nanoparticle percentage volume fraction on the dimensionless heat transfer coefficient of the nanofluid in each channel of MCHS is shown in Fig.4.4. The dimensionless heat transfer coefficient is calculated using:

$$\hat{h} = \frac{h_{nf}}{h_{avg,water}} \quad (4.5)$$

Where h_{nf} is the nanofluids heat transfer coefficient value and $h_{avg,water}$ is the average heat transfer coefficient of pure water. It is inferred that the nanofluids can enhance the heat transfer of MCHS as its volume fraction increases from 0% to 2.5%. However, nanofluid with 5% volume fraction is not able to enhance the heat transfer or performing almost the same result as pure water. This may due to agglomeration of the nanoparticles. The validity of the present results is confirmed with Murshed et al. [28] results. From Fig.4.4, it is also observed that the middle channel (channel 14) has the highest averaged heat transfer coefficient value as expected. The averaged heat transfer coefficient value for other channels is seen to decrease depending on their distance from the wall. The present results as illustrated in Fig.5.6 show similar trend with the results obtained by Liu and Garimella [29] and Chein and Chen [19]. Since higher temperatures are observed due to the heat transfer from the high-temperature edge of the heat sink, the averaged heat transfer coefficient is lower in channels close to the edge of the heat sink.

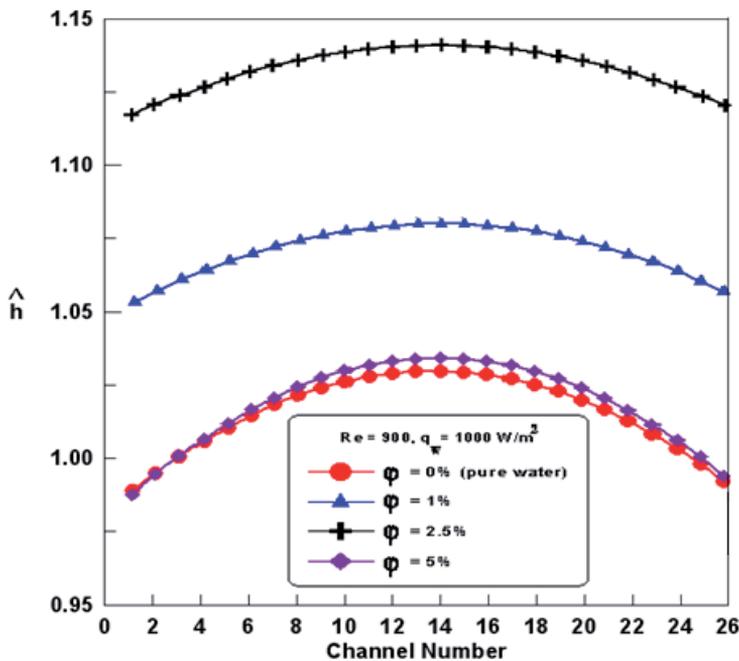


Fig. 4.4. Dimensionless averaged heat transfer coefficient versus number of channels for various particle volume fractions.

4.4 Friction factor

The variation of the friction factor for various particle volume fractions is shown in Figs.4.5 and 4.6 for Reynolds number ranged from 100-500 and 500-1000, respectively. The friction factor is calculated using Darcy equation [20]:

$$f = \frac{2D_h\Delta p}{\rho_{nf}u_{in,nf}^2L_c} \quad (4.6)$$

Where D_h is the hydraulic diameter, Δp is the pressure drop, ρ_{nf} is the density of nanofluid, $u_{in,nf}$ is the inlet velocity of nanofluid, and L_c is the length of channel.

The results show that the friction factor is similar for all particle volume fractions where the friction factor decreases with the increase of Reynolds number. Based on Figs.4.5 and 4.6, it can be stated that the increment of dynamic viscosity due to the presence of the nanoparticles in water, only appears to give a slight rise in friction factor especially for low Reynolds number range (Fig.4.5). For high Reynolds number range (Fig.4.6), the nanofluid flow yields significant effect on friction factor.

Fig.4.7 shows a comparison of the predicted friction factor from the present numerical simulation against previous experimental studies [12, 30]. A very good agreement was obtained for $\phi = 0\%$ (pure) and $\phi = 1\%$ alumina, which shows that the friction factor decreases with the increase of Reynolds number.

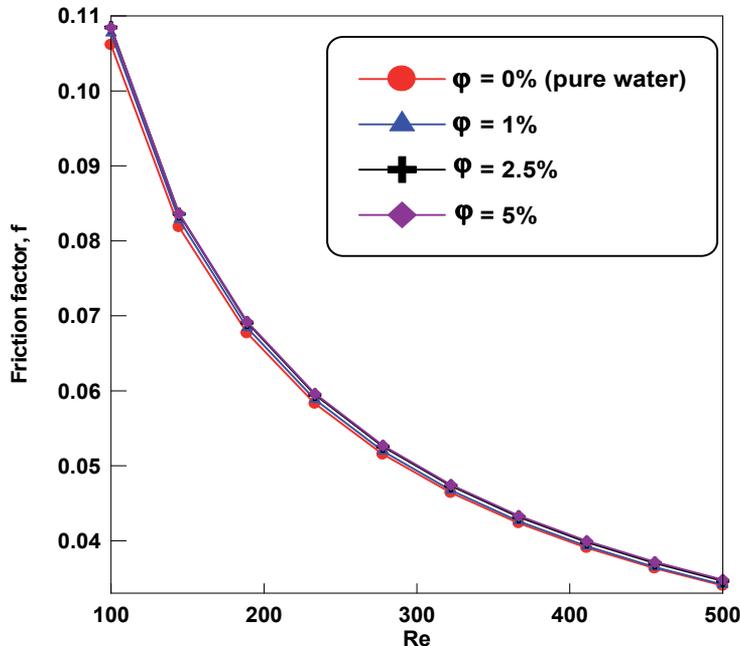


Fig. 4.5. Friction factor of MCHS for Reynolds number ranged from 100-500.

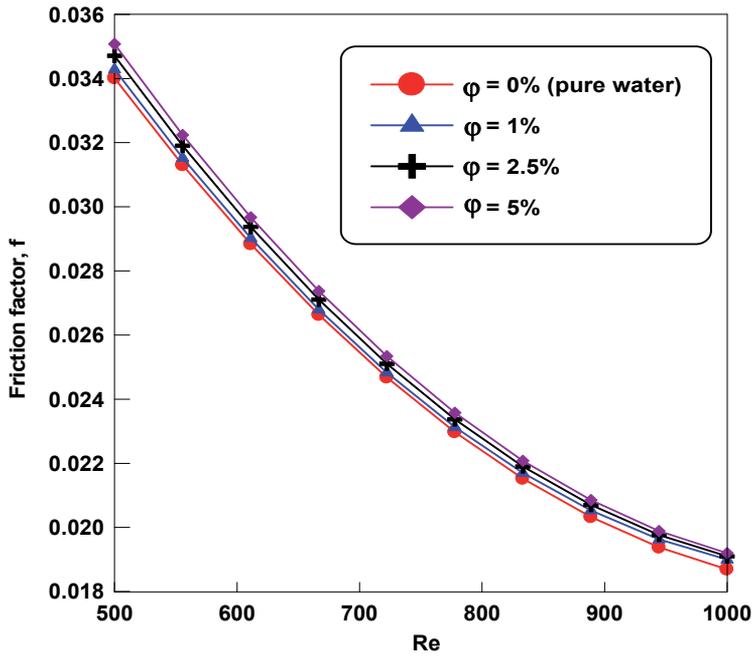


Fig. 4.6. Friction factor of MCHS for Reynolds number ranged from 500-1000.

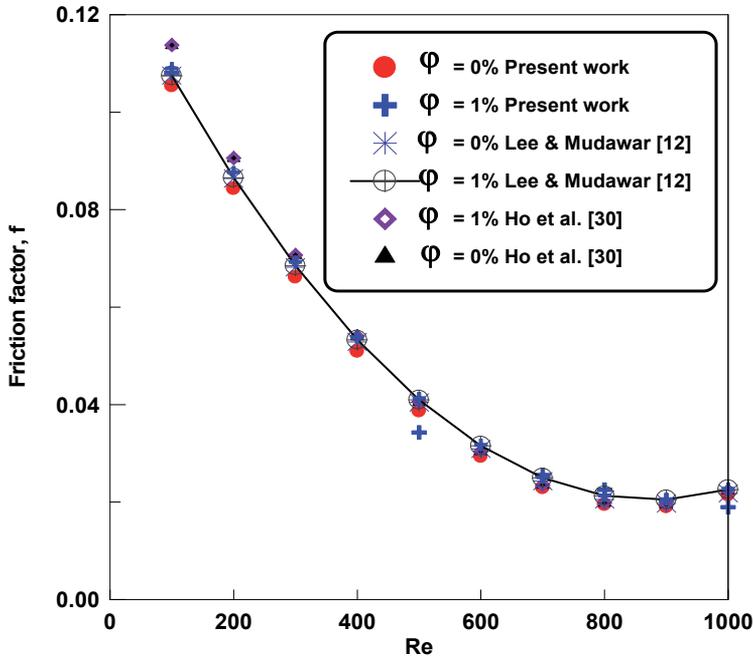


Fig. 4.7. A comparison of predicted friction factor with [12] and [30].

4.5 Wall shear stress

The effect of different volume fractions of nanoparticles on the local dimensionless wall shear stress is also investigated in the current study and the results are shown in Fig.4.8. The dimensionless wall shear stress is calculated using:

$$\hat{\tau} = \frac{\tau_{nf}}{\tau_{avg, water}} \quad (4.7)$$

Where τ_{nf} is the nanofluid wall shear stress and $\tau_{avg, water}$ is the average value of pure water wall shear stress.

As depicted in Fig.4.8, the wall shear stress increases with increasing Reynolds number for a specified percentage of volume fractions. It is clearly seen that for all cases investigated, the presence of nanoparticles substantially increases the wall shear-stress as compared with pure water. In general, for all cases examined, the least wall shear stress occurs at lower value of $\phi = 1\%$ while highest value occurs at $\phi = 5\%$.

4.6 Thermal resistance

Since the maximum wall temperature ($T_{w, max}$) of nanofluid-cooled MCHS is lower than pure water-cooled MCHS, reduction of thermal resistance in nanofluid-cooled MCHS is then expected. If the presence of nanoparticles in coolant did not produce extra pressure drop which relates to the coolant pumping power required, it can then be considered as another benefit of using nanofluid as the coolant in MCHS operation. The volume flow rate, Q for rectangular shaped microchannel is defined as:

$$Q = N W_{ch} H_{ch} u_{in, nf} \quad (4.8)$$

Where N is the number of microchannels, W_{ch} is the width of channel, H_{ch} is the height of channel, and $u_{in, nf}$ is the inlet velocity of nanofluid.

The cooling performance of the MCHS using nanofluid is assessed by exploring the results for the thermal resistance, R_{th} as defined in Eq.3.3. The variation of thermal resistance versus pumping power is presented in Fig.4.9. The pumping power is determined from the calculated volume flow rate as defined in Eq.4.8 and computed pressure drop using Eq.3.4 (Section 3). As depicted in Fig.4.9, the nanofluid is shown to have significantly lower value of R_{th} as its particle volume fraction increases than pure water. For the current geometry under study, the maximum reduction in R_{th} that can

be achieved is about 57.6% for nanofluid with particle volume fraction of 5% at 0.0145 W. This great improvement in R_{th} is mainly due to the increase in thermal conductivity and decrease in temperature difference between inlet and maximum wall temperature as the particle volume fraction increases. The additional reduction in R_{th} is clearly due to the thermal dispersion. Since water has the highest R_{th} value, it is not preferable for removing high heat flux. The percent enhancement of MCHS's cooling performance is actually calculated by using:

$$\frac{R_{th, nf} - R_{th, water}}{R_{th, water}} \times 100\% \quad (4.9)$$

Where $R_{th, nf}$ corresponds to thermal resistance of particular nanofluid and $R_{th, water}$ corresponds to thermal resistance of pure water.

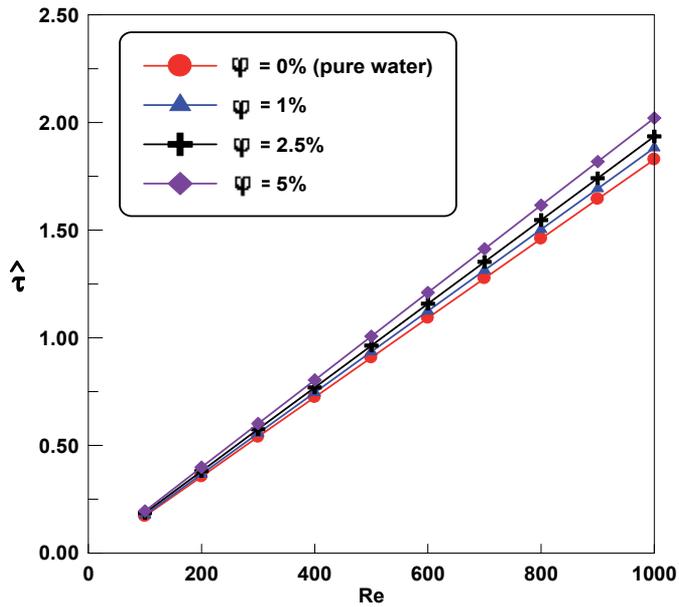


Fig. 4.8. Dimensionless wall shear stress versus Reynolds number for various particle volume fractions.

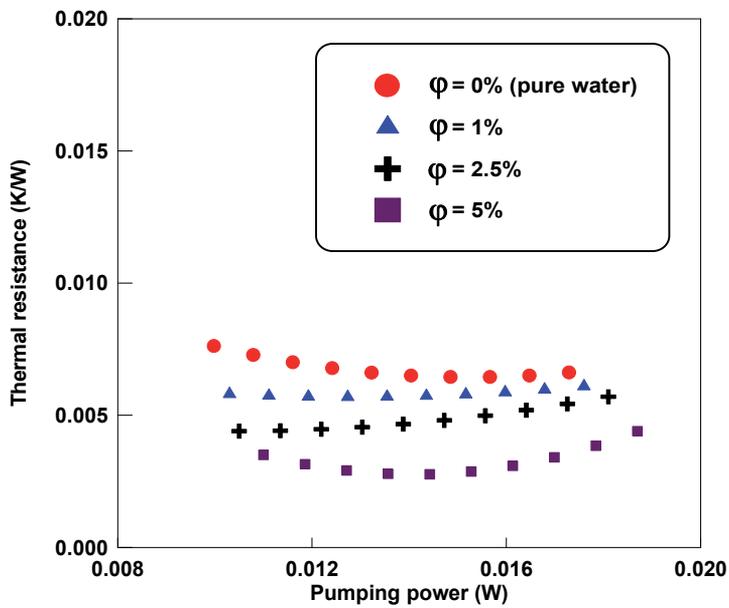


Fig. 4.9. Thermal resistance of MCHS as function of pumping power.

5. Effect of using different nanofluids on MCHS performance

Based on the results that obtained in Section 4 which focused on the effect of using different particle volume fractions of $\text{Al}_2\text{O}_3\text{-H}_2\text{O}$ on heat transfer and liquid flow characteristics in MCHS, it is known that nanofluids have the capability to enhance the heat transfer performance of MCHS with smaller pressure drop penalty. However, due to production of various nanofluids in market, further investigations on using different types of nanofluids in MCHS are necessary because the heat transfer enhancement of MCHS is strongly depend on thermophysical properties of dispersed nanoparticles in base fluid [31]. Thus, in this chapter, the effect of using six types of nanofluids on heat transfer and liquid flow characteristics in triangular cross-section MCHS with specified dimensions is comprehensively investigated.

The specific dimensions of the triangular MCHS used in this section are given in Table 5.1. A constant particle volume fraction value of 2% is considered for all types of nanofluids studied in this chapter. The thermophysical properties for all types of nanoparticles, base fluid (water), and the nanofluids with particle volume fraction of 2% are listed in Table 5.2. For all types of nanofluids studied, the density, thermal conductivity, and dynamic viscosity appear significantly higher while the specific heat of nanofluids is lower than pure water. For instance, diamond- H_2O nanofluid has a relative increase of 5.03%, 6.11%, and 5% in the density, thermal conductivity, and dynamic viscosity and has a relative decrease of 5.90% in specific heat, respectively, compared with pure water.

In this section, different types of nanofluids which are $\text{Al}_2\text{O}_3\text{-H}_2\text{O}$, $\text{Ag-H}_2\text{O}$, $\text{CuO-H}_2\text{O}$, diamond- H_2O , $\text{SiO}_2\text{-H}_2\text{O}$, and $\text{TiO}_2\text{-H}_2\text{O}$ are used as working fluids. The thermophysical properties as listed in Table 5.2 including density, heat capacity, thermal conductivity, and viscosity for all types of nanofluids which are involved in the governing equations are calculated using Eqs.4.1-4.4 (Section 4), respectively.

5.1 Heat transfer coefficient

The thermal performance of using different types of nanofluids in MCHS is examined by presenting the results of the dimensionless heat transfer coefficient along the length of channel number 1 in Fig.5.1. It is observed that all types of nanofluids-cooled MCHS could be able to enhance the heat transfer compared with pure water-cooled MCHS. The water-base nanofluid containing diamond has the highest heat transfer coefficient value followed by SiO_2 , CuO , TiO_2 , Ag and Al_2O_3 . However, there is no significant difference that can be detected between TiO_2 and CuO due to their small difference in thermal diffusivity. However, there is a very small increment in $\text{CuO-H}_2\text{O}$ -cooled MCHS especially at the outlet of the channel compared with that for $\text{TiO}_2\text{-H}_2\text{O}$ -cooled MCHS.

D_h (μm)	k (μm)	h (μm)	l (μm)	L_{ch} (μm)	S (μm)	β
230	280	430	452	10000	500	36.07o

Table 5.1. Dimensions of the triangular cross-section MCHS.

Properties	Nanoparticle (diamond)	Base fluid (water)	Nanofluid (diamond-H ₂ O) $\varphi = 2\%$
ρ (kg/m ³)	3510	998.2	1048.436
C_p (J/kg.K)	497.26	4182	3935.28
κ (W/m.K)	1000	0.613	0.65046
μ (Ns/m ²)		0.001003	0.00105315
Properties	Nanoparticle (silicon dioxide)	Base fluid (water)	Nanofluid (SiO ₂ -H ₂ O) $\varphi = 2\%$
ρ (kg/m ³)	2200	998.2	1022.236
C_p (J/kg.K)	703	4182	4032.25
κ (W/m.K)	1.2	0.613	0.62194
μ (Ns/m ²)		0.001003	0.00105315
Properties	Nanoparticle (copper oxide)	Base fluid (water)	Nanofluid (CuO-H ₂ O) $\varphi = 2\%$
ρ (kg/m ³)	6500	998.2	1108.236
C_p (J/kg.K)	535.6	4182	3754.26
κ (W/m.K)	20	0.613	0.64722
μ (Ns/m ²)		0.001003	0.00105315
Properties	Nanoparticle (titanium dioxide)	Base fluid (water)	Nanofluid (TiO ₂ -H ₂ O) $\varphi = 2\%$
ρ (kg/m ³)	4250	998.2	1063.236
C_p (J/kg.K)	686.2	4182	3902.53
κ (W/m.K))	8.9538	0.613	0.64364
μ (Ns/m ²)		0.001003	0.00105315
Properties	Nanoparticle (silver)	Base fluid (water)	Nanofluid (Ag-H ₂ O) $\varphi = 2\%$
ρ (kg/m ³)	10500	998.2	1188.236
C_p (J/kg.K)	235	4182	3484.44
κ (W/m.K)	429	0.613	0.65037
μ (Ns/m ²)		0.001003	0.00105315
Properties	Nanoparticle (aluminum oxide)	Base fluid (water)	Nanofluid (Al ₂ O ₃ -H ₂ O) $\varphi = 2\%$
ρ (kg/m ³)	3970	998.2	1057.636
C_p (J/kg.K)	765	4182	3925.47526
κ (W/m.K)	40	0.613	0.64882
μ (Ns/m ²)		0.001003	0.00105315

Table 5.2. Thermophysical properties of nanoparticles, water and nanofluids at 293K.

Since the velocity at the entrance of the channel was kept constant at $Re = 600$, it can be noticed that the heat transfer coefficient decreases linearly along the length of channel because of the development of thermal boundary layer for all cases. The highest value of heat transfer coefficient occurs near the entrance. The present results show a similar trend with the results obtained by Lee and Mudawar [12]. It is also depicted from Fig.5.1 that the difference in dimensionless heat transfer coefficient between various nanofluids is too small but it follows trend and not too far from reported values. From the results obtained, lowest

thermal resistance is expected with diamond- H_2O nanofluid cooled MCHS due to its lowest temperature distribution. Therefore, the presence of diamond nanoparticle in water could greatly enhance the cooling of MCHS compared with other types of nanofluids. This is due to the highest thermal transport capacity of diamond in nature and diamond particles are often used as filler in mixtures for upgrading the performance of a matrix (composition of mixtures) as reported by Xie et al. [32]. Thus, it is reasonable to expect that the addition of diamond nanoparticles in water would lead to heat transfer enhancement.

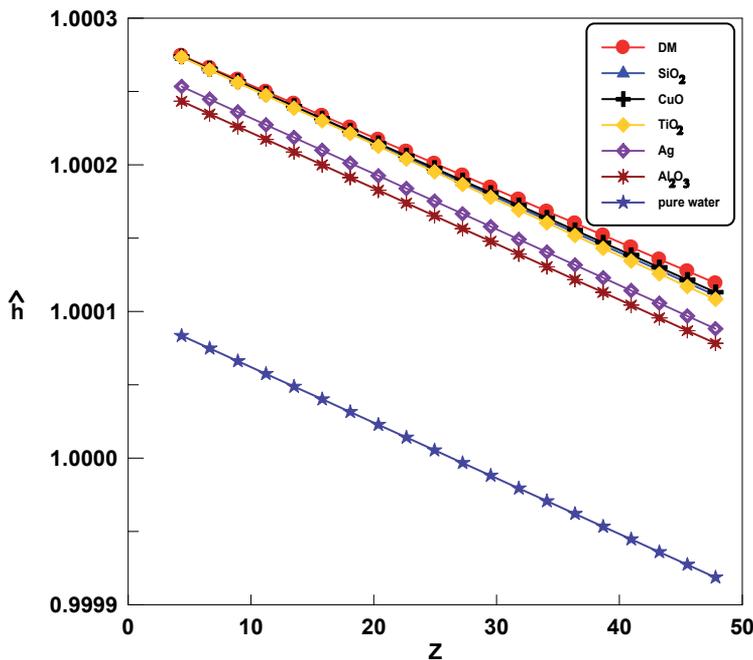


Fig. 5.1. Dimensionless heat transfer coefficient along the length of channel No.1.

5.2 Pressure drop

The dimensionless pressure distribution of the MCHS for various nanofluids and pure water along the length of channel number 1 is shown in Fig.5.2. The dimensionless pressure distribution is calculated using:

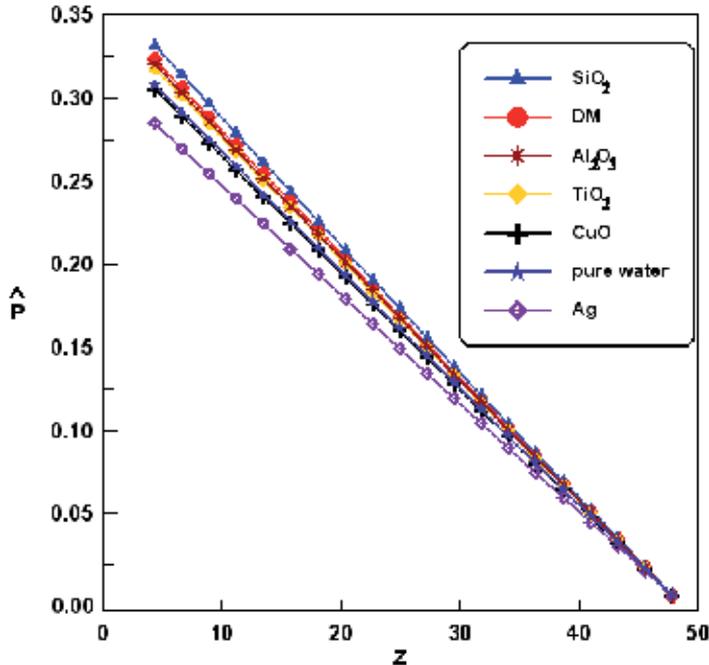


Fig. 5.2. Dimensionless pressure drop variation along the length of channel No.1.

$$\hat{p} = \frac{\Delta P}{\rho u_{in}^2} \quad (5.1)$$

where ΔP is the pressure drop of particular microchannel configuration, ρ is the water density, and u_{in} is the inlet water velocity.

It can obviously be seen that the high pressure region occurs at the entrance and the low pressure region occurs at the outlet of the heat sink. This is due to the needs of high pressure to push the fluid flow along the direction of the microchannels out from the outlet plenum of heat sink. Thus, the pressure decreases linearly along the length of channel for all types of nanofluids studied including pure water. Because of that, there is no apparent difference observed between pure water and nanofluids flows particularly at the outlet region of MCHS. However, compared with pure water cooled case, slightly larger pressure drops for nanofluid-cooled MCHS were found except for Ag-H₂O nanofluid which has the lowest value of pressure drop among nanofluids types. The water-base nanofluids containing SiO₂ has the highest value of pressure drop and followed by diamond-H₂O respectively. The pressure drop in water-base nanofluids-cooled MCHS containing Al₂O₃ and TiO₂ are slightly lower than diamond-H₂O-cooled MCHS.

Since the viscosity of nanofluids is larger than pure water as indicated in Table 5.2, larger pressure drop is expected in nanofluid-cooled MCHS. Another reason that contributes to increase the pressure drop is the deposition of nanoparticle in water is increasing the wall

roughness [28]. However, Ag-H₂O-cooled MCHS still has lower value of pressure drop than pure water-cooled MCHS and there is no extra pressure drop in CuO-H₂O-cooled MCHS as compared with pure water-cooled MCHS. The reason for this result is Ag-H₂O has the lowest Prandtl number ($Pr = 5.6$ and followed by CuO-H₂O ($Pr = 6.1$). The small difference in pressure drop between other types of nanofluids which can be seen in Fig.5.2 is due to small differences in their Prandtl numbers and SiO₂-H₂O yields the highest pressure drop across the MCHS due to its highest Prandtl number ($Pr = 6.8$). Therefore, the pressure drop penalty is not only affected by the viscosity of nanofluids but also the Prandtl number of nanofluids which need to be considered in nanofluids selection as a coolant for MCHS. Apart from that, slight increment in pressure drop particularly for Ag-H₂O and CuO-H₂O-cooled MCHS which can be considered as one of the benefits of using nanofluid in MCHS operations in order to achieve overall heat transfer enhancement with small pressure drop penalty.

5.3 Thermal resistance

The pumping power which is the product of the pressure drop across the heat sink, ΔP and volume flow rate, Q , is calculated using Eq.3.4. The volume flow rate, Q for triangular shaped microchannel is defined as:

$$Q = 0.5 N k h u_{in,nf} \quad (5.1)$$

Where N is the number of microchannels, k is the width of channel, h is the height of channel, and $u_{in,nf}$ is the inlet velocity of nanofluid.

The thermal resistance for all types of nanofluids as function of pumping power is shown in Fig.5.3. It is observed that all types of nanofluids reduce the thermal resistance of MCHS

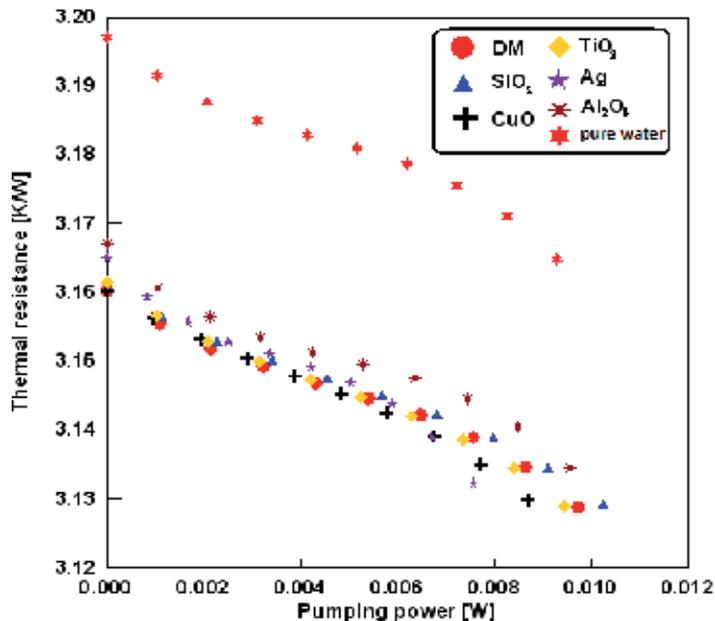


Fig. 5.3. Thermal resistance of MCHS as function of pumping power.

than pure water. The maximum reduction in R_{th} is achieved in diamond-H₂O MCHS is 3.1292 K/W when pumping power is 0.0097207 W. The results indicate that the cooling performance of a MCHS at pumping power = 0.005 W is enhanced by about 2 % for water-base nanofluids containing diamond, SiO₂, CuO, and TiO₂, 1.13% for water-base nanofluids containing Al₂O₃, and 1.04% for water-base nanofluids containing Ag respectively, compared with that of the MCHS with pure water. The percent enhancement of MCHS cooling performance is also calculated using Eq.4.9. The great improvement in R_{th} is mainly due to the increase in thermal conductivity and decrease in temperature difference between inlet and maximum wall temperature. The additional reduction in R_{th} is also due to the thermal dispersion of particle volume fractions.

6. Effect of using different base fluids on MCHS performance

Since base fluid for all types of nanofluids studied in Section 5 was water, it is important and necessary to know the heat transfer and liquid flow characteristics in MCHS using other types of conventional base fluids such as ethylene glycol (EG), glycerin, and oil. By using other types of conventional base fluids instead of using water, the MCHS cooling performance could be enhanced or may not. Thus, considering nanofluid flow in a microchannel, the effect of using different types of base fluids including water, EG, engine oil, and glycerin on heat transfer and liquid flow characteristics in trapezoidal shaped MCHS which was selected for this chapter with specified dimensions is extensively discussed in this chapter. The specified dimensions are given in Table 6.1.

Since the best uniformities in heat transfer coefficient and temperature can be achieved in MCHS using diamond nanoparticle as obtained in Section 5, the diamond nanoparticle is selected and considered for all types of base fluids studied in this chapter. The thermophysical properties for diamond nanoparticle, for all types of base fluids, and the nanofluids with particle volume fraction of 2% formulated are listed in Table 6.2. For all types of nanofluids, the density, thermal conductivity, and dynamic viscosity appear significantly higher while the specific heat of nanofluids is lower than the base fluids. For instance, the EG-diamond nanofluid has a relative increase of 4.28%, 6.12%, and 5% in the density, thermal conductivity, and dynamic viscosity and has a relative decrease of 4.77% in specific heat, respectively, compared with EG base fluid.

The thermophysical properties as listed in Table 6.2 including density, heat capacity, thermal conductivity, and viscosity for all types of base fluids for nanoparticles are calculated using Eqs.4.1-4.4 (Section 4), respectively.

Dh (μm)	a (μm)	b (μm)	c (μm)	Lch (μm)	h (μm)
318	280	225	431	10000	430

Table 6.1. Dimensions of the trapezoidal cross-section MCHS.

Properties	Nanoparticle (Diamond)	Base fluid (Water)	Nanofluid (H ₂ O-Diamond) $\phi = 2\%$
ρ (kg/m ³)	3510	998.2	1048.436
Cp (J/kg.K)	497.26	4182	3935.28
κ (W/m.K)	1000	0.613	0.65046
μ (Ns/m ²)		0.001003	0.00105315
Properties	Nanoparticle (Diamond)	Base fluid (Ethylene Glycol)	Nanofluid (EG-Diamond) $\phi = 2\%$
ρ (kg/m ³)	3510	1117.48	1165.33
Cp (J/kg.K)	497.26	2382.1	2268.56
κ (W/m.K)	1000	0.2492	0.26445
μ (Ns/m ²)		0.022	0.0231
Properties	Nanoparticle (Diamond)	Base fluid (Oil)	Nanofluid (Oil-Diamond) $\phi = 2\%$
ρ (kg/m ³)	3510	888.23	940.665
Cp (J/kg.K)	497.26	1880.3	1777.09
κ (W/m.K)	1000	0.145	0.15387
μ (Ns/m ²)		0.8451	0.88736
Properties	Nanoparticle (Diamond)	Base fluid (Glycerin)	Nanofluid (Glycerin-Diamond) $\phi = 2\%$
ρ (kg/m ³)	3510	1264.03	1308.95
Cp (J/kg.K)	497.26	2385	2283.76
κ (W/m.K))	1000	0.286	0.30349
μ (Ns/m ²)		1.5347	1.61144

Table 6.2. Thermophysical properties of diamond nanoparticels, base fluids and nanofluids at 293K.

6.1 Temperature profile

The effect of different types of base fluids on dimensionless temperature variation is presented in Fig.6.1. It can clearly be seen that the temperature of pure water is always highest among all types of base fluids. By comparing dimensionless temperature distributions between different types of base fluids, glycerin-diamond has the lowest value of temperature while water-diamond has the highest value of temperature and followed by EG-diamond and oil-diamond. Based on results shown in Fig.6.1, it can be stated that the lowest thermal resistance is expected in the glycerin-diamond fluid cooled MCHS due to its lowest temperature distribution. Therefore, the presence of diamond nanoparticle in glycerin could greatly enhance the cooling of MCHS compared with other types of base fluids studied. This is because of glycerin has the highest dynamic viscosity in nature compared to other base fluids and diamond particle mixed properly in glycerin which contribute to increase the thermal transport capacity of the mixture.

6.2 Heat transfer coefficient

The thermal performance of using different types of base fluids in MCHS is examined by plotting results of the dimensionless heat transfer coefficient along the length of channel number 20 as shown in Fig.6.2. Koo and Kleinstreuer [10] reported that higher Prandtl number base fluids-cooled MCHS could be able to enhance the heat transfer performance of MCHS as compared with water base-cooled MCHS. This was proven by the present results where the glycerin base-cooled MCHS has the highest \hat{h} value due to the highest Prandtl number value followed by oil, and EG.

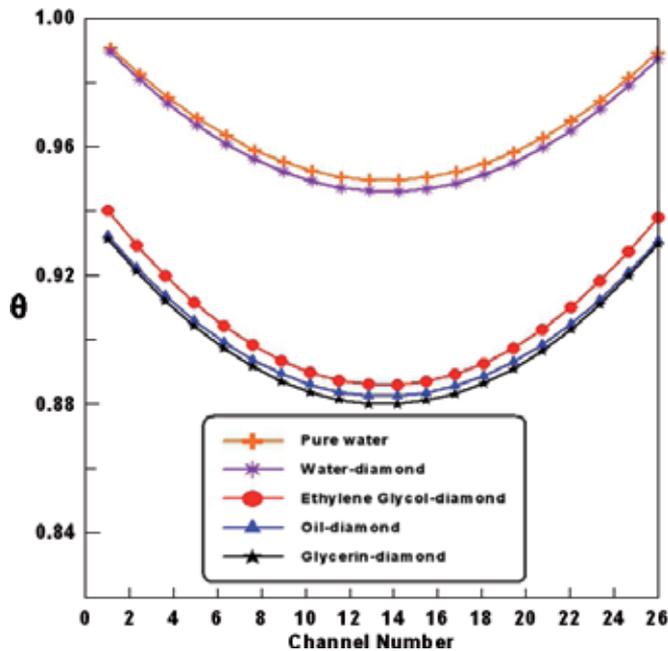


Fig. 6.1. Dimensionless temperature profile versus number of channels for pure water and various base fluids.

Fig.6.3 shows the dimensionless heat transfer coefficient in each channel of the heat sinks for various base fluids-cooled MCHS. The \hat{h} for high Prandtl number base fluids is always higher than that for water base fluid. This is due to the fact that high Prandtl number base fluids are experiencing stronger thermal flow development effects as reported by Koo and Kleinstreuer [10]. However, there is no significant differences in \hat{h} for EG, oil, and glycerine but when the plot is zoomed in, there are slight differences in \hat{h} where glycerin base-cooled MCHS has the highest \hat{h} value, oil is in between, and EG is the least. For each type of the heat sinks, the middle channel (channel number 14) has the highest \hat{h} value as expected. The averaged heat transfer coefficient value for other channels is seen to decrease depending on their distances from the wall. The averaged heat transfer coefficient distribution for all types of MCHS is almost symmetrical with respect to the centreline of the heat sink.

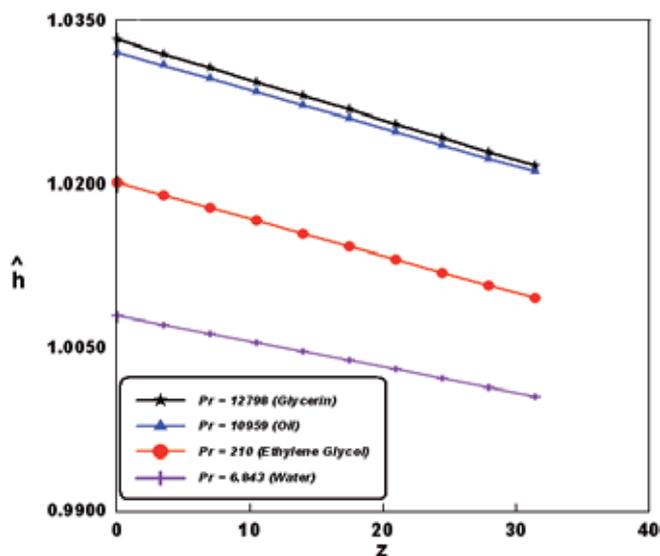


Fig. 6.2. Dimensionless heat transfer coefficient along the length of channel No.20 for various base fluids.

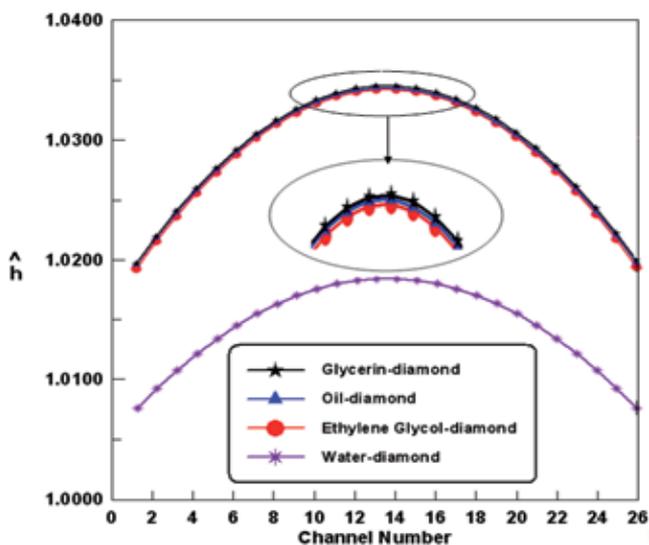


Fig. 6.3. Dimensionless averaged heat transfer coefficient versus number of channels for various base fluids.

Fig.6.4 shows the dimensionless heat transfer coefficient in each channel of various substrate materials for EG-base fluid. The highest \hat{h} is achieved in copper made microchannels, which has the lowest thermal diffusivity followed by aluminium, steel, and titanium. The increase in \hat{h} could be attributed to the decreased conduction resistance along the microchannel wall due to much lower thermal diffusivity in copper made microchannels. The averaged heat transfer coefficient of oil and glycerine base-cooled MCHS shows a similar trend as for EG base-cooled MCHS. However, the condition becomes different for low Prandtl number base

fluid (water) where the highest \hat{h} is present in steel made microchannels as depicted in Fig.6.5. This proves that the thermal diffusivity of substrate material on the heat transfer in micro domains has significant effect for low and high Prandtl number base fluids. Thus, based on the results presented in Fig.6.5, it can be stated that the heat transfer performance of a low Prandtl number base fluid such as water can be greatly enhanced by using it as a coolant in MCHS of a high thermal diffusivity made substrate material such as steel.

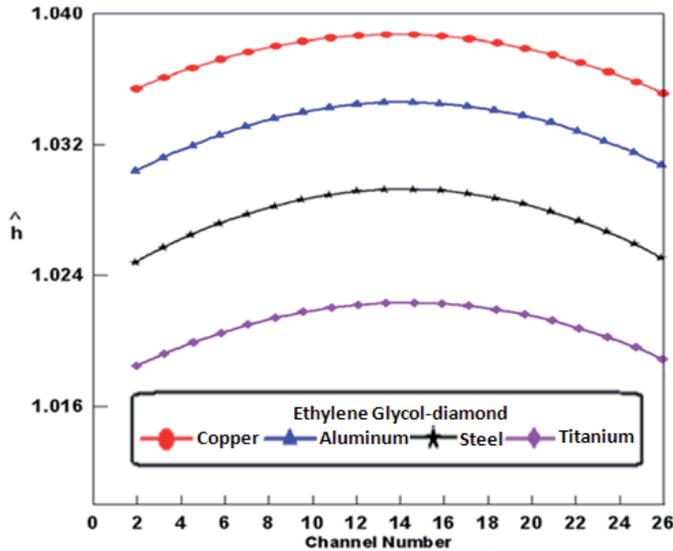


Fig. 6.4. Dimensionless averaged heat transfer coefficient versus number of channels for ethylene glycol-base nanofluid with different substrate materials.

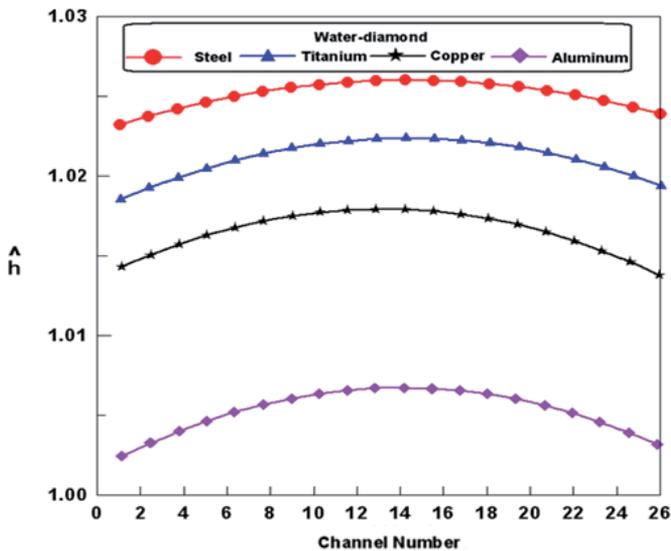


Fig. 6.5. Dimensionless averaged heat transfer coefficient versus number of channels for water-base nanofluid with different substrate materials.

6.3 Pressure drop

The variation of dimensionless pressure of the MCHS for various base fluids along the length of channel number 20 is presented in Fig.6.6. It can obviously be seen that the high pressure region occurs at the entrance and the low pressure region occurs at the outlet of MCHS. This is due to the needs of high pressure to push the fluid flow along the direction of the microchannels out from the outlet plenum of heat sink. Thus, the pressure decreases linearly along the length of channel for all types of base fluids studied.

The higher \hat{P} for nanofluid-cooled MCHS is found in high Prandtl number base fluids due to the highest dynamic viscosity as stated in Table 6.2 as reported by Koo and Kleinstreuer [10]. The water base-cooled MCHS has the lowest value of pressure drop due to its lowest dynamic viscosity among other base fluids. There is no apparent difference in \hat{P} between EG, oil and glycerin as can be seen in Fig.6.6 although glycerin has the highest value of \hat{P} followed by oil and EG, respectively. Since the dynamic viscosity of EG, oil and glycerin base fluids is larger than water, larger pressure drop is expected in the high Prandtl number base fluids than the low Prandtl number base fluids. Thus, selection of high Prandtl number base fluids as coolant in MCHS should be considered in balancing heat transfer enhancement and pressure drop penalty.

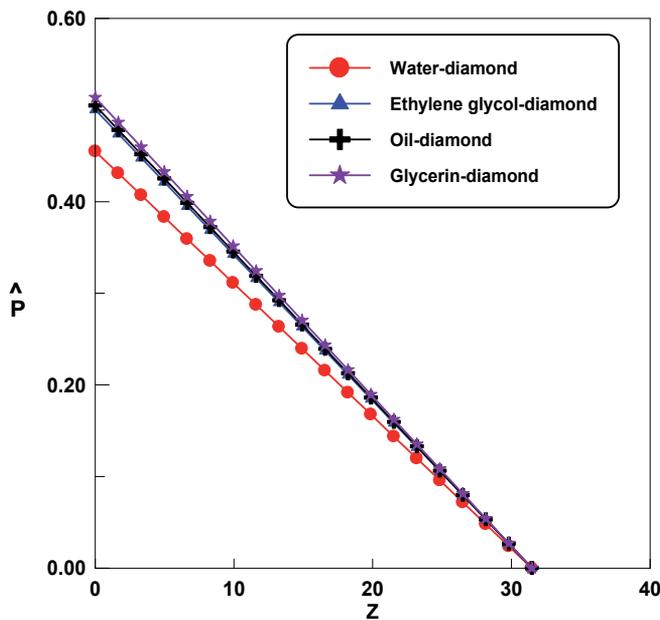


Fig. 6.6. Dimensionless pressure drop variation along the length of channel No.20.

7. Conclusions

The cooling efficiency of MCHS is numerically investigated using the FVM in this study. A three-dimensional model, fully developed laminar flow and heat transfer governing equations are solved using the finite volume method. Based on these effects, several conclusions can be drawn from all the simulation results.

For the effects of MCHS geometrical parameters:

- The width to height ratio (W_c/H_c) has a significant effect on the Poiseuille number for the rectangular cross-section microchannels. Poiseuille number increases linearly with the increase of W_c/H_c .
- For the trapezoidal cross-section MCHS, the height-to-top width ratio (H/a), the bottom-to-top width ratio (b/a), and length-to-hydraulic diameter ratio (L/D_h) are the important design parameters for trapezoidal microchannels. Poiseuille number increases linearly when H/a and L/D_h decreases while b/a increases.
- The tip angle (β) has a remarkable effect on the Poiseuille number for the triangular cross-section MCHS. Poiseuille number increases linearly with the increase of β .
- In order to achieve overall heat transfer enhancement in MCHS, it is envisioned that rectangular cross-section MCHS with a small hydraulic diameter was of greater benefit to the heat transfer coefficient with a lower penalty in pressure drop and friction factor compared with trapezoidal and triangular cross section MCHS.

For nanofluids in MCHS:

- The cooling performance of MCHS improved well when the particle volume fraction increases. However, the heat transfer becomes poor in larger amount of particle volume fraction of nanofluid flow as the heat transfer could not be enhanced compared with pure water.
- The presence of nanoparticles substantially increases the friction factor and wall shear stress. Both are rises with the increase of particle volume fractions.
- The nanofluid-cooled MCHS has lower thermal resistance than pure water-cooled MCHS. The thermal resistance decreases as the particle volume fraction increases.
- Diamond nanoparticles dispersed in water is preferable since diamond-H₂O-cooled MCHS has highest heat transfer coefficient value among the others.
- Ag nanoparticles dispersed in water is recommended as it has lowest pressure drop and wall shear stress among the other nanofluids studied.
- In order to achieve overall heat transfer enhancement, a base fluid of high Prandtl number (210-12800) such as glycerin, oil, and ethylene should be used to maximize the merits of adding nanoparticles for fluid flow in MCHS.
- Different types of base fluids do not have a significant influence on the friction factor.
- The heat transfer for low Prandtl number fluid flow such as water is greatly enhanced in high thermal diffusivity made material such as steel compared with low thermal diffusivity materials. These findings suggest that for low Prandtl number base fluids, the substrate material of high thermal diffusivity is desirable and vice versa in order to achieve a better heat transfer in MCHS.

8. Nomenclature

A	channel flow area, m ²
a	channel top width for trapezoidal cross-section MCHS, μm
Ag	silver
Al	aluminum
Al ₂ O ₃	aluminum oxide
b	channel bottom width for trapezoidal cross-section MCHS, μm
c	channel hypotenuse for trapezoidal cross-section MCHS, μm
c_p	specific heat, J/kg.K

Cu	copper
CuO	copper oxide
D_h	hydraulic diameter, μm
DM	diamond
EG	Ethylene Glycol
f	friction factor
f_{Re}	Poiseuille number
FV	Finite Volume
FVM	Finite Volume Method
h	heat transfer coefficient, $\text{W}/\text{m}^2\cdot\text{K}$
h	channel height for trapezoidal and triangular cross-section MCHS, μm
\hat{h}	dimensionless heat transfer coefficient
H_{ch}	channel height for rectangular cross-section MCHS, μm
H_2O	water
k	channel width for triangular cross-section MCHS, μm
κ	thermal conductivity, $\text{W}/\text{m}\cdot\text{K}$
κ_s	solid thermal conductivity, $\text{W}/\text{m}\cdot\text{K}$
l	channel hypotenuse for triangular cross-section MCHS, μm
L_{ch}	channel length, μm
MCHS	Microchannel Heat Sink
n	direction normal to the wall
\mathbf{n}	direction normal to the surface element
Nu	Nusselt number
P	channel wet perimeter, μm
\hat{P}	dimensionless pressure
P_{in}	inlet pressure, Pa
P_{out}	outlet pressure, Pa
Pr	Prandtl number
PW	pure water
P_{ow}	pumping power, W
Q	volume flow rate, m^3/s
q_w	heat flux at microchannel heat sink top plate, W/m^2
Re	Reynolds number
R_{th}	thermal resistance, K/W
S	distance between two microchannels, μm
SiO_2	silicon dioxide
t	substrate thickness, μm
T	fluid phase temperature, K
Ti	titanium
TiO_2	titanium dioxide
T_{in}	fluid inlet temperature, K
T_s	microchannel heat sink solid temperature, K
u	fluid velocity, m/s
u_{in}	inlet fluid velocity, m/s
U	dimensionless velocity in x-coordinate
V	dimensionless velocity in y-coordinate

W	dimensionless velocity in z-coordinate
W_{ch}	channel width for rectangular cross-section MCHS, μm
X, Y, Z	dimensionless cartesian coordinates

Greek Symbols

a	thermal diffusivity, m^2/s
β	tip angle for triangular cross-section MCHS
μ	viscosity, $\text{kg}/\text{m}\cdot\text{s}$
ρ	density, kg/m^3
τ	wall shear stress
$\hat{\tau}$	dimensionless wall shear stress
φ	particle volume fraction, %
θ	dimensionless temperature

Subscripts

avg	average
bf	base fluid
ch	channel
f	pure fluid
nf	nanofluid
h	hydraulic
i	inlet
o	outlet
s	solid
th	thermal
w	wall

9. Acknowledgements

The authors would like to sincerely thank the Ministry of Higher Education (MOHE) of Malaysia for the provision of a grant with code no FRGS01090010 to support this work. In addition, the supports and helps provided by all members of the Centre for Advanced Computational Engineering (CACE) at Department of Mechanical Engineering, UNITEN to carry out the research in this paper are deeply appreciated.

10. References

- [1] D.B. Tuckerman, R.F.W Pease, "High Performance Heat Sinking for VLSI", IEEE Elect. Dev. Lett., vol. 2, 1981, pp 126-129.
- [2] P.R. Parida, "Experimental Investigation of Heat Transfer Rate in Micro-Channels", Master Thesis, Indian Institute of Technology Guwahati, India, 2007.
- [3] I. Mudawar, "Assessment of High Heat Flux Thermal Management Schemes", IEEE Trans. Components and Packaging Tech., vol. 24, 2001, pp 122-128.
- [4] S.G. Kandlikar, M.E. Steinke, "Examples of Microchannel Mass Transfer Processes in Biological Systems", in Proc. 1st Int. Conf. Minichannels and Microchannels, ASME, vol. 11, April 24-25, 2003, pp 933-943.

- [5] K.C. Toh, X.Y. Chen, J.C. Chai, "Numerical Computation of Fluid Flow and Heat Transfer in Microchannels", *Int. J. Heat Mass Transfer*, vol. 45, 2002, pp 5133-5141.
- [6] X.F. Peng, G.P. Peterson, "Convective Heat Transfer and Flow Friction for Water Flow in Microchannels Structures", *Int. J. Heat Mass Transfer*, vol. 39, 1996, pp 2599-2608.
- [7] X.F. Peng, G.P. Peterson, "The Effect of Thermofluid and Geometrical Parameters on Convection of Liquids through Rectangular Microchannels", *Int. J. Heat Mass Transfer*, vol. 38, 1995, pp 755-758.
- [8] I. Tiselj, G. Hetsroni, B. Mavko, A. Mosyak, E. Pogrebnyak, Z. Segal, "Effect of Axial Conduction on the Heat Transfer in Microchannels", *Int. J. Heat Mass Transfer*, vol. 47, 2004, pp 2551-2565.
- [9] P.S. Lee, S.V. Garimella, D. Liu, "Investigation of Heat Transfer in Rectangular Microchannels", *Int. J. Heat Mass Transfer*, vol. 48, 2005, pp 1688-1704.
- [10] J. Koo, C. Kleinstreuer, "Laminar Nanofluid Flow in Microheat-Sinks", *Int. J. Heat Mass Transfer*, vol. 48, 2005, pp 2652-2661.
- [11] R. Chein, G. Huang, "Analysis of Microchannel Heat Sink Performance Using Nanofluids", *Appl. Therm. Eng.*, vol. 25, 2005, pp 3104-3114.
- [12] J. Lee, I. Mudawar, "Assessment of the Effectiveness of Nanofluids for Single-Phase and Two-Phase Heat Transfer in Micro-channels", *Int. J. Heat Mass Transfer*, vol. 50, 2007, pp 452-463.
- [13] S.P. Jang, S.U.S. Choi, "Cooling Performance of a Microchannel Heat Sink with Nanofluids", *Appl. Therm. Eng.*, vol. 26, 2006, pp 2457-2463.
- [14] R. Chein, J. Chuang, "Experimental Microchannel Heat Sink Performance Studies Using Nanofluids", *Int. J. Therm. Science*, vol. 46, 2007, pp 57-66.
- [15] T.H. Tsai, R. Chein, "Performance Analysis of Nanofluid-cooled Microchannel Heat Sinks", *Int. J. Heat Fluid Flow*, vol. 28, 2007, pp 1013-1026.
- [16] S.E.B. Maiga, S.J. Palm, C.T. Nguyen, G. Roy, N. Galanis, "Heat Transfer Enhancement by Using Nanofluids in Forced Convection Flows", *Int. J. Heat Fluid Flow*, vol. 26, 2005, pp 530-546.
- [17] S.V. Patankar, "Numerical heat transfer and fluid flow", Hemisphere Publishing Corporation, Taylor and Francis Group, New York, 1980.
- [18] H.K. Versteeg, W. Malalasekara, "An Introduction to Computational Fluid Dynamics", The Finite Volume Method, Prentice Hall, London, 1995.
- [19] R. Chein, J. Chen, "Numerical Study of the Inlet/Outlet Arrangement Effect on Microchannel Heat Sink Performance", *Int. J. Thermal Sciences*, vol. 48, 2009, pp 1627-1638.
- [20] L.F. Moody, "Friction Factors for Pipe Flow", *J. Heat Transfer ASME Trans*, vol. 66 1944, pp 671-684.
- [21] S. Kandlikar, S. Garimella, D. Li, S. Colin, M.R King, "Heat Transfer and Fluid Flow in Minichannels and Microchannels", Elsevier, USA, 2005.
- [22] H.Y. Wu, P. Cheng, "Friction Factors in Smooth Trapezoidal Silicon Microchannels with Different Aspect Ratios", *Int. J. Heat Mass Transfer*, vol. 46, 2003, pp 2519-2525.
- [23] H. Xie, J. Wang, T. Xi, Y. Liu, F. Ai, "Thermal Conductivity Enhancement of Suspensions Containing Nanosized Alumina Particles", *J. Appl. Phys.*, vol. 91, 2002, pp 4568-4572.
- [24] J.C. Maxwell, "A Treatise on Electricity and Magnetism", Clarendon Press, Oxford, UK, 1873.
- [25] Y. Xuan, W. Roetzel, "Conceptions for Heat Transfer Correlation of Nanofluids", *Int. J. Heat Mass Transfer*, vol. 43, 2000, pp 3701-3707.

- [26] A. Einstein, *"Investigations on the Theory of the Brownian Movement"*, Dover Publications, Inc., New York, 1956.
- [27] S.U.S. Choi, Z.G. Zhang, P. Keblinski, *"Nanofluids"*, Encyclopedia of Nanoscience and Nanotechnology, vol. 6, 2004, pp 757-773.
- [28] S.M.S. Murshed, K.C. Leong, C. Yang, *"Thermophysical and Electrokinetic Properties of Nanofluids-A Critical Review"*, Appl. Therm. Eng., vol. 28, 2008, pp 2109-2125.
- [29] D. Liu, S.V. Garimella, *"Investigation of Liquid Flow in Microchannels"*, J. Thermophys. Heat Transfer, vol. 18, 2004, pp 65-72.
- [30] C.J. Ho, L.C. Wei, Z.W. Li, *"An Experimental Investigation of Forced Convective Cooling Performance of a Microchannel Heat Sink with Al₂O₃/Water Nanofluid"*, Appl. Therm. Eng., vol. 30, 2010, pp 96-103.
- [31] D. Wen, G. Lin, S. Vafaei, K. Zhang, *"Review of Nanofluids for Heat Transfer Applications"*, Particuology, vol. 7, 2009, pp 141-150.
- [32] H. Xie, W. Yu, Y. Li, *"Thermal Performance Enhancement in Nanofluids Containing Diamond Nanoparticles"*, J. Appl. Phys, vol. 42, 2009, pp 9-13.

Modelling and Optimizing Operating Conditions of Heat Exchanger with Finned Elliptical Tubes

Stanisław Łopata and Paweł Ocloń
*Cracow University of Technology, Cracow
Poland*

1. Introduction

High performance heat exchangers are widely used in many industries such as petrochemical industry, automotive industry, energetic industry and many others. The name of this kind of heat exchangers is in consistency with their structure. Large heat transfer area leads to high thermal efficiency of the device. Its working principle is to cool rapidly large amount of gaseous or liquid medium. Because of its compact size, it is possible to use it for easy installation in various systems, like the heating, drying, air conditioning and the other systems.

High thermal performance of this type of heat exchangers (Nagatha, 2004; Ibrahim & Gomaa, 2009) is achieved due to large packed rectangular fins on tube surface, as well as the shape of the tubes. The elliptical shape ensures better distribution of velocity along the tube circumference during the flow, as well as lower pressure drop in comparison to the circular shape (Matos et al, 2004; Jang & Yang, 1998; Li et al, 2006; Nishiyama et al, 1988; Khan et al, 2004; Saboya & Saboya, 1981; Saboya & Sparrow, 1974; Bordallo & Saboya, 1999). Consequently, the coefficient of heat transfer in the elliptical shape is larger when compared to the circular shape (Idem et al, 1990; Rocha et al, 1997; Taler, 2009). It brings about that the heat exchange process is more efficient in elliptical than in circular tubes.

To use all the aforementioned advantages, it is necessary to ensure uniform distribution of velocity in all the tubes. Failing that, it leads to large differences in mean temperature in the tubes. Consequently, excessive thermal stress occurs, that may cause the heat exchanger to break down. Small volume of collectors of the heat exchangers implicates possibility of improper flow condition inside the tubes, causing unsuitable inner distribution of thermal and mechanical loads. Damages to a tube, Fig. 1, as well as a tube bundle prone to buckling, Fig. 2, prove, that such a failure is possible (Łopata & Ocloń, 2010; Ocloń & Łopata, 2011).

Compressible stresses are undesirable (Łopata & Ocloń, 2010; Łopata & Ocloń, 2011) in this type of construction, because the allowable compressible load (allowable buckling stress) for a slender tube is significantly lower than the safe tensile load.

The simplified method to assess failure of a high performance heat exchanger with finned elliptical tubes is presented in the paper. The diagram, shown on Fig. 3, presents the method. The geometrical representation of construction – a CAD model is prepared at first. Next, on the basis of prepared geometry, numerical representation of the model is created. Afterwards, the flow computation - CFD (Ferziger & Peric 1999; Anderson, 1995; Gresho &

Sani, 2000; Chung, 2010) is carried out using the commercial tool ANSYS CFX (ANSYS, 2008). It is desirable to obtain the uniform mass flow rates in all the tubes of the heat exchanger, because it ensures the appropriate heat transfer conditions for the device. When the mean velocities in tubes differ significantly (Łopata & Oclon, 2010), new modified collectors are applied to improve flow distribution inside the device (Oclon & Łopata, 2011). If the liquid distribution among all the tubes ensures safe heat transfer conditions, the CFD computation is stopped. The results obtained with the CFD analysis as the bulk temperatures T_{bulk} and the heat transfer coefficients h are transferred next to a code, that calculates temperature distribution inside the elliptical tube, using the Control Volume Method (Chung, 2010; Taler & Duda, 2006) as the mixed type boundary conditions (Incopera, 2001). The obtained temperature and pressure values are used as thermal loads for structural analysis, using the ANSYS Structural (ANSYS ADPL, 2009) software.

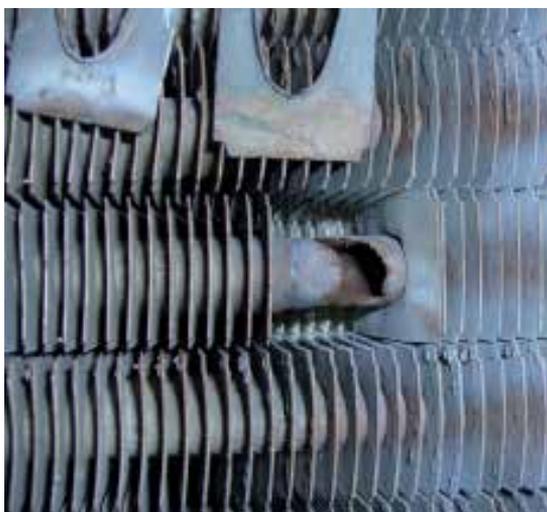


Fig. 1. View of damaged tube



Fig. 2. Buckled tubes of a high performance heat exchanger

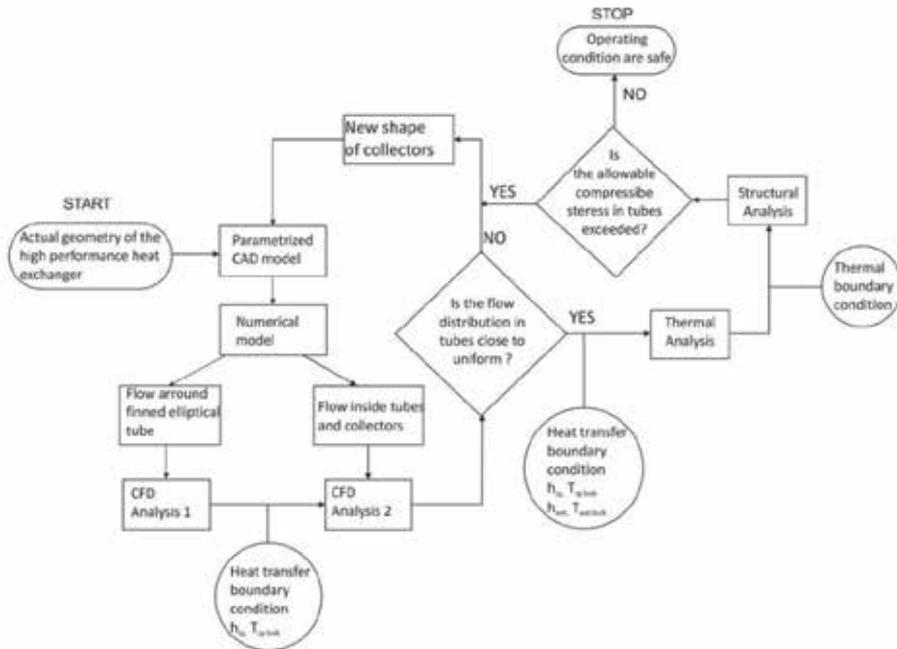


Fig. 3. Calculation methodology

This commercial tool computes displacement and stresses for the numerical model of the heat exchanger using the Finite Element Method (Reddy, 1993; Logan, 2011; Smith & Griffiths, 2004; Zienkiewicz & Taylor, 2000; Hutton, 2003). Structural analysis allows finding the zones and locations where allowable stresses are exceeded. When it happens, the computational flow loops back to the CFD analysis and the other geometrical parameters of the collector are set. The loop continues until allowable stresses in the construction are exceeded. If the maximum stress is lower than the critical stress, the construction is safe to use, and the computations stop.

The methodology presented in the paper may help to assess structural safety. The computations described in this chapter were successfully applied in analysis of two failures of a high performance heat exchanger that happened in an industrial company in Poland.

2. Hydraulic and thermal performance of heat exchanger

The proper operating conditions of the high performance heat exchanger improve its thermal efficiency and lead to significantly smaller pressure exerted on tube material.

The operating conditions depend strongly on flow distribution inside the tubes of the high performance heat exchanger. The main task of hydraulic computation is to confirm, that liquid flowing inside the tubes is distributed properly in all the tubes. It is ideally, when the mass flow rate \dot{m} in each tube equals to the mass flow at the inlet \dot{m}_{in} divided by the number of tubes n in a single pass of the heat exchanger. Nevertheless, it is impossible in fact to obtain the uniform flow distribution in all the tubes. The improper flow distribution has an influence on heat transfer process in each pass of the heat exchanger, causing high temperature differences in the adjacent tubes that can cause the stress limit of the tube material to be exceeded.

The heat exchanger, for which the computations are carried out, is presented schematically on Fig. 4.

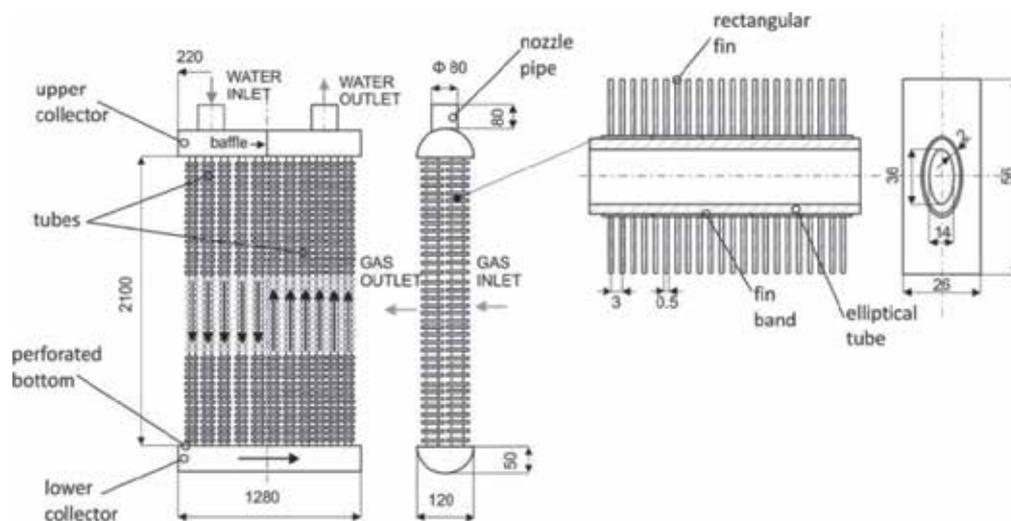


Fig. 4. Scheme of a two pass high performance heat exchanger

The operating conditions of the heat exchanger, as presented on Fig. 4, are listed in Tab. 1:

Name	Symbol	Value, Unit
Mass flow rate of water	\dot{m}_{wat}	19 kg /s, 14 kg /s, 9.5 kg/s
Velocity of combustion gas	w_{cg-in}	4 m/s
Number of tubes	n	42
Inlet and outlet temperature of water	$T_{wat-in}/T_{wat-out}$ (max value)	110/130 °C
Inlet and outlet temperature of combustion gas	$T_{cg-inlet}/T_{cg-outlet}$ (min value)	450/350 °C
Operating pressure for water	p_{op}	6 bar

Table 1. Operating conditions for the analysed high performance heat exchanger

Working principle of the analysed heat exchanger is to cool rapidly combustion gas. Heat is transferred from hot gas to water which flows into the device through the inlet nozzle pipe at the mean temperature $T_{wat-in} = 110$ °C. Water fills in the upper part of the collector. Next, it is distributed in the 1st pass to $n = 42$ tubes of the heat exchanger, where it flows down to the lower collector. Then, it turns back and flows through the other part of the tubes (2nd pass of the heat exchanger) into the upper collector and finally leaves the device through the outlet nozzle pipe with the mean temperature $T_{wat-out}$, that equals to no more than 130 °C. The tubes are arranged into two rows. Combustion gas, with the mean inlet temperature

$T_{cg-inlet} = 450$ °C flows perpendicularly to the tubes, exchanging heat with the first row of the tubes and subsequently with the second one. Finally, it leaves the device cooled to no less than $T_{cg-outlet} = 350$ °C. Later, the gas is cooled to 100 °C in other heat exchangers (not shown here) and released to the atmosphere. The presented heat exchanger is a part of the installation where the largest thermal stresses may occur, so its failure is most probable to happen.

2.1 Numerical model of flow around finned elliptical tube

Analysis of flow dynamic and heat transfer is divided into two parts. First, heat transfer from combustion gas to the tube is modelled, next the flow of water inside the heat exchanger is analysed, knowing the thermal boundary conditions - heat transfer coefficient and bulk temperature distribution on the outer surface of the tube. Discretization of flow domain (lower case) and boundary condition types (uppercase) for analysis of the flow around the finned elliptical tube are presented on Fig. 5.

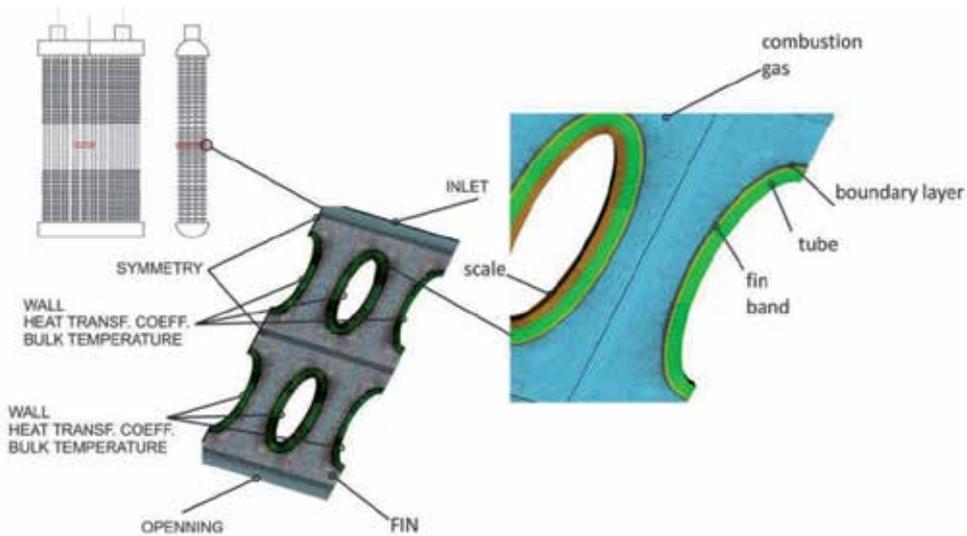


Fig. 5. Numerical model of flow around finned elliptical tube

The combustion gas is modelled as an ideal gas with density $\rho_{cg} = 3$ kg/m³ and dynamic viscosity $\eta_{cg} = 27 \cdot 10^{-6}$ Pa/s. Mass, momentum and energy equation (Bird, 2001) are used in computation. In order to model the turbulence, the SST (shear-stress transport) transport equations (Egorov & Menter, 2008) are used in the CFD analysis. The mixed boundary condition simulates the heat transfer from water to a tube. The bulk temperature of water T_{wat-b} is assumed to be equal $T_{wat-out}$ and the heat transfer coefficient h_{wat} is calculated, knowing the mean velocity in a tube, which is calculated according to the following formula:

$$\overline{w_{wat}} = \frac{\dot{m}_{wat}}{n \cdot A_{tube} \cdot \rho_{wat}}, \quad (1)$$

where A_{tube} is equal to the area of flow through an elliptical duct and ρ_{wat} is the density of water calculated for T_{wat-b} .

The Nusselt Number of water Nu_{wat} is calculated according to (Gnielinski, 1976):

$$\text{Nu}_{\text{wat}} = \frac{\frac{1}{8}f \cdot (\text{Re}_{\text{wat}} - 1000) \cdot \text{Pr}_{\text{wat}}}{1 + 12.7 \cdot \left(\frac{1}{8}f\right)^{\frac{1}{2}} \cdot (\text{Pr}_{\text{wat}}^{\frac{2}{3}} - 1)} \quad (2)$$

It was stated in the paper (Łopata & Oclon, 2010), that velocity of water in this type of a heat exchanger can be significantly different for tubes in the 1st and in the 2nd pass of the high performance heat exchanger. The zones and locations in the tube were indicated where mean velocity is over 20 times lower than in the other tubes. The thermal load is significantly larger in those zones than in the ones with proper velocity distribution. Moreover, low velocities of liquid favour deposition of boiler scale (Najibi et al., 2007; Quan et al., 1997; Nebot et al., 2007; Freeborn & Lewis, 1962), especially when water is not prepared properly that causes the flow resistance growth. Consequently, water velocity decreases even more. With regard to the above, it is not the average velocity of water, but the $w_{\text{wat}} = 1/2 \cdot \bar{w}_{\text{wat}}$ that is used to calculate the Reynold's number of water - Re_{wat} . It simulates the real operating conditions of a device, when scale starts to deposit inside the tube and thermal load increases significantly.

The Prandtl number of water Pr_{wat} is obtained for the specific heat capacity $c_{p-\text{wat}}$, dynamic viscosity η_{wat} and thermal conductivity k_{wat} of water that on the turn are measured at bulk temperature equal to $T_{\text{wat}-b}$:

$$\text{Pr}_{\text{wat}} = \frac{c_{p-\text{wat}} \cdot \eta_{\text{wat}}}{k_{\text{wat}}} \quad (3)$$

The Darcy friction factor f is calculated according to formula (White, 2003):

$$f = (0.79 \ln \cdot \text{Re}_{\text{wat}} - 1.64)^{-2} \quad (4)$$

Finally h_{wat} is computed as follows:

$$h_{\text{wat}} = \frac{\text{Nu}_{\text{wat}} \cdot k_{\text{wat}}}{d_h} \quad (5)$$

The equivalent diameter d_h of an elliptical tube is computed according to the following dependence:

$$d_h = \frac{4 \cdot A_{\text{tube}}}{P_{\text{tube}}}, \quad (6)$$

where P_{tube} is equal to the perimeter of an elliptical duct. The values of h_{wat} , obtained from eq. (5) and $T_{\text{wat}-b}$ for all the computational cases are listed in Table 2.

\dot{m}_{wat} [kg/s]	h_{wat} [$\frac{W}{m^2 \cdot K}$]	$T_{\text{wat}-b}$ [°C]
19	3630	130
14	2720	130
9.5	1815	130

Table 2. Values of $T_{\text{wat}-b}$ and h_{wat} used in analysis of flow around finned elliptical tube

The cross - section of the tube, which fractured during the device operation, is presented on Fig. 6. It is possible to observe deposition of scale on the inner surface of the tube. For computational cases, two thicknesses of scale $t_s = 1$ mm and 1.5 mm are analysed. Thermal conductivity of scale depends on quantity of calcium carbonate (Najibi et al. 1997). The thermal conductivity of boiler scale is assumed to be equal to 0.1 W/(m K) for the present

computation that models the location of the fracture. Scale deposition is modelled only for the tubes in the middle of computational domain, because the existence of scale in neighbouring tubes wasn't noticed during the site inspection.

It is possible to observe, on the outer wall of the elliptical tube, the layer of fouling that also influences heat transfer inside the device, worsening considerably thermal contact between the fin and the tube.

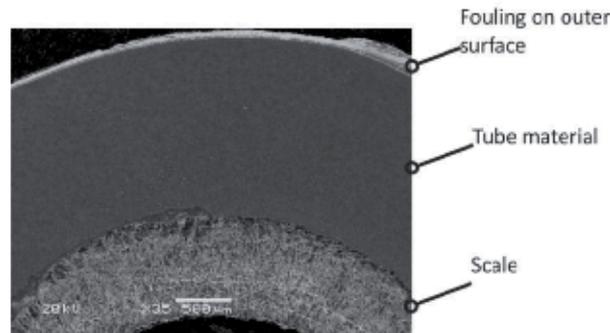


Fig. 6. Picture of damaged tube (cross - section)

This layer is modelled in ANSYS CFX as thermal resistance on the interface between the fin band and the tube. Thickness of material is assumed to be 0.1 mm and its thermal conductivity equals to $1 \text{ W}/(\text{m K})$.

Heat flux density \dot{q} is computed on the interface between the tube (Fig. 7) and the fin band so that the interface area is divided into $n_{int} = 40$ intervals along the y direction.

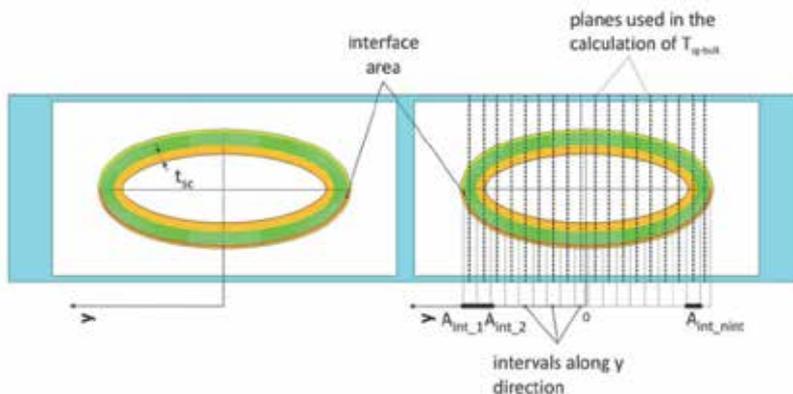


Fig. 7. Division of interface area into intervals

For the n_{el} number of elements that belong to each of the y intervals the area averaged heat flux is computed as follows:

$$\dot{q}_{int} = \frac{\sum_{j=1}^{n_{el}} q_j \cdot A_j}{\sum_{j=1}^{n_{el}} A_j}. \quad (7)$$

The area averaged wall temperature is calculated according to the formula:

$$T_{wall} = \frac{\sum_{j=1}^{n_{el}} T_j \cdot A_j}{\sum_{j=1}^{n_{el}} A_j}. \quad (8)$$

The mass flow averaged temperature is computed in the middle of each y interval:

$$T_{cg-bulk} = \frac{\sum_{j=1}^{n_{el}} T_j \cdot \dot{m}_j}{\sum_{j=1}^{n_{el}} \dot{m}_j}. \quad (9)$$

Knowing values of the area averaged heat flux density for each interval \dot{q}_{int} , the area averaged wall temperature for intervals T_{wall} and the bulk temperature computed in the middle of interval $T_{cg-bulk}$, the heat transfer coefficient from combustion gas to the tube is determined according to the formula:

$$h_{cg}(y) = \frac{\dot{q}_{int}}{(T_{cg-bulk} - T_{wall})} \quad (10)$$

The sample results presented below were obtained for $\dot{m}_{wat} = 19$ kg/s. The bulk temperature and heat transfer coefficient distribution for all the \dot{m}_{wat} values listed in Table 2 are used as the thermal boundary condition for analysis of flow distribution inside the high performance heat exchanger.

On Figure 8 temperature and velocity distribution is presented when combustion gas flows around the finned elliptical tube.

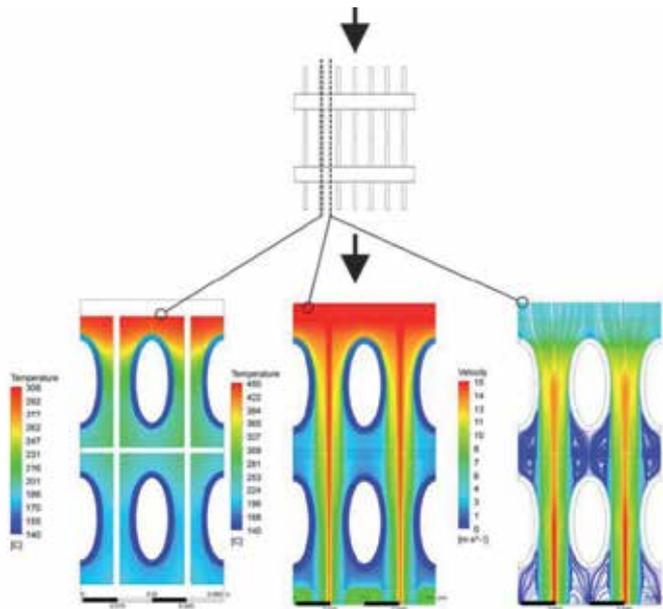


Fig. 8. Temperature and velocity distribution for $t_s = 0$ mm

It is possible to observe, that the stagnation zones occur under the 1st row as well as above and under the 2nd row of tubes. Therefore the fins don't operate properly, and the largest amount of heat is transferred only to the 1st row of tubes. It may lead to high temperature differences on the circumference of the tubes and consequently may cause significant bending stresses to the tube.

The temperature distribution in the fin and in gas for $t_s = 1$ mm is presented on Fig 9. The boiler scale deposition significantly increases temperatures of the tube and the fin. Consequently thermal loading of the tubes in this case is considerably greater than when $t_s = 0$ mm.

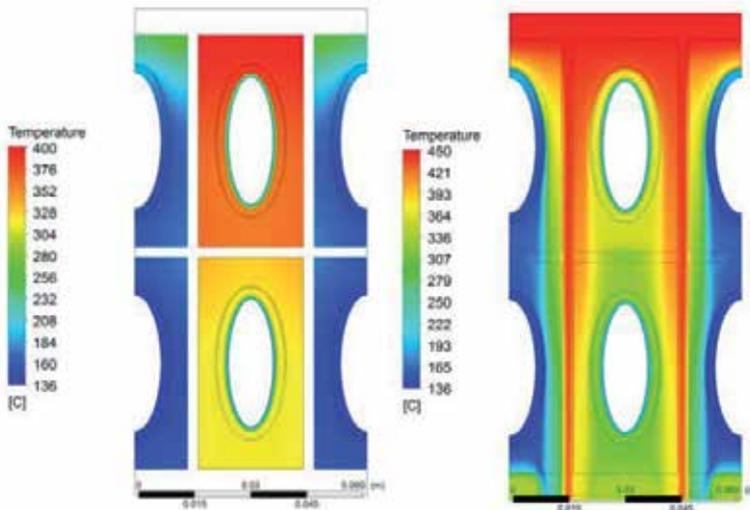


Fig. 9. Temperature distribution for $t_s = 1$ mm

Increase in scale thickness from 1 mm to 1.5 mm (see Fig. 10) causes the temperature of the tube, the fin and combustion gas to grow up. It is evident that boiler scale reduces heat flux from combustion gas to water, causing temperature of the fins and the tube to increase. This effect is more pronounceable with increase in boiler scale thickness.

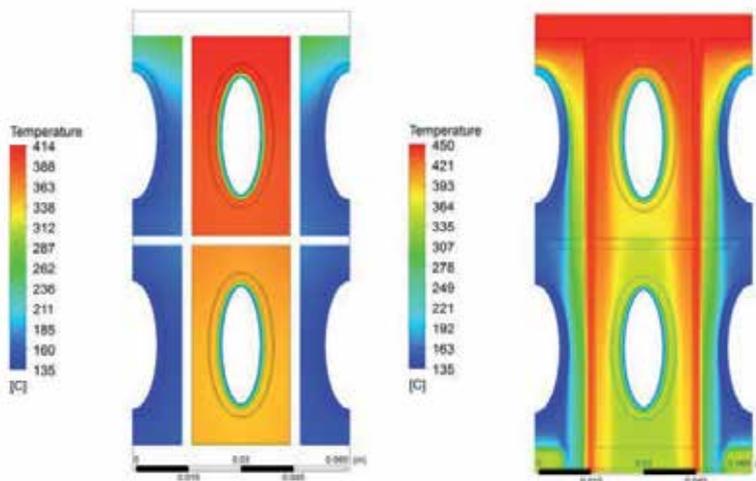


Fig. 10. Temperature distribution for $t_s = 1.5$ mm

On Fig. 11 the h_{cg} and $T_{cg-bulk}$ distribution along y direction for t_s is presented

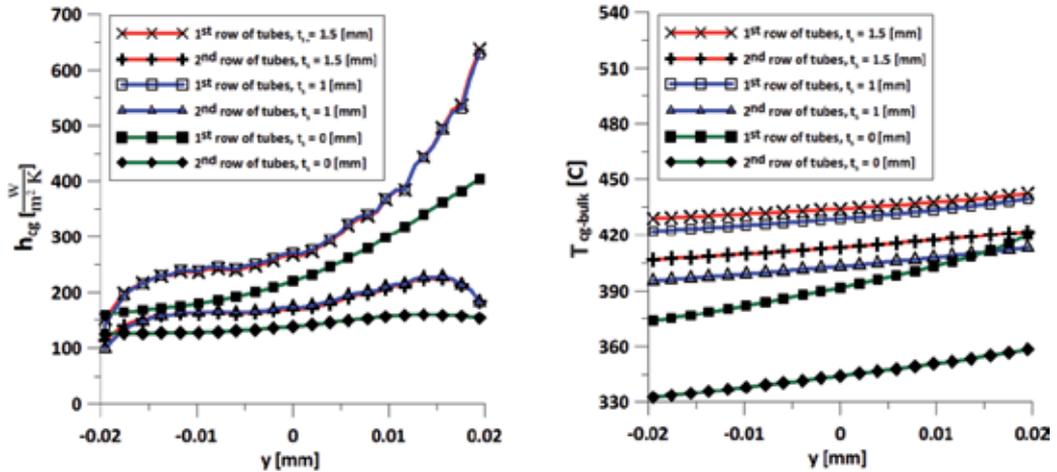


Fig. 11. Heat transfer coefficient and bulk temperature distribution for combustion gas - $t_s = 0$ mm, $t_s = 1$ mm and $t_s = 1.5$ mm

Heat transfer coefficient from combustion gas to the tube is larger when boiler scale thickness increases. Rapid growth of $T_{cg-bulk}$ combined with isolative behaviour of scale, causes high temperature increase of the tube. Therefore, operating condition of the device worsens significantly, because the normal stresses acting on the tube grow up with increase in temperature, too.

2.2 Flow distribution inside collectors and tubes of high performance heat exchanger

Analysis of flow distribution inside the high performance heat exchanger is carried out using the mass, momentum and energy balance equations (Bird, 2001; Incopera & De Witt, 2001). The turbulence model SST is applied to flow behaviour simulation. The numerical model of the flow is presented on Fig. 12.

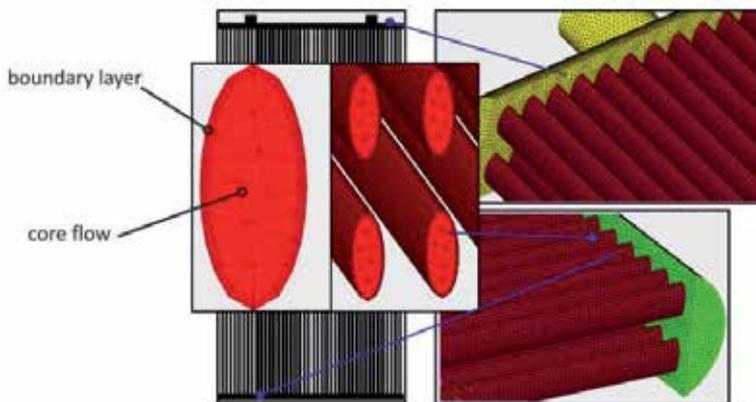


Fig. 12. Numerical model used for analysis of flow distribution inside heat exchanger

The boundary layer is modelled using the hexahedral elements and the flow core is discretized using the tetrahedral elements. The boundary conditions set for $\dot{m}_{wat} = 19 \text{ kg/s}$ is presented on Fig. 13.

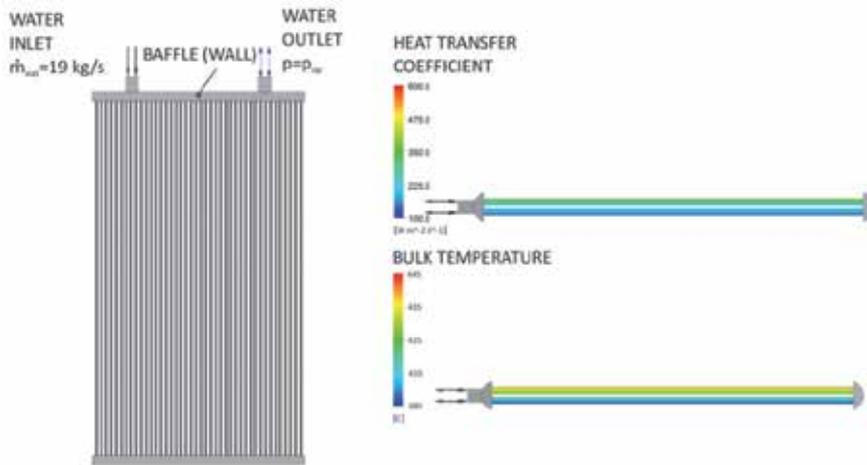


Fig. 13. Boundary condition for analysis of water flow - $\dot{m}_{wat} = 19 \text{ kg/s}$

As mentioned in section 1, it is difficult to ensure the uniform flow distribution to all the tubes, due to limited space in the collectors. The zones of the tubes where the velocity is significantly lower than the \overline{w}_{wat} are marked on red, see Fig. 14.

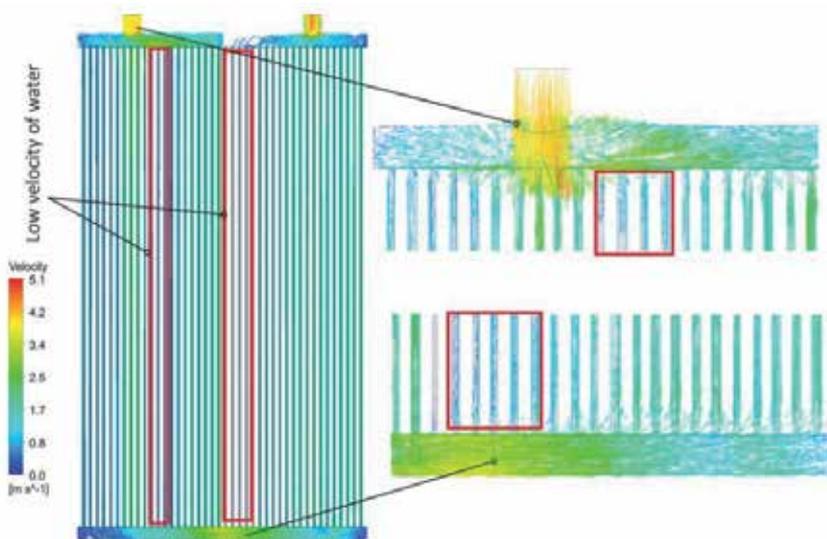


Fig. 14. Velocity distribution in high performance heat exchanger - $\dot{m}_{wat} = 19 \text{ kg/s}$

These zones are formed because pressure changes occur (see Fig. 15) in the direction parallel to the perforated bottom of the collector.

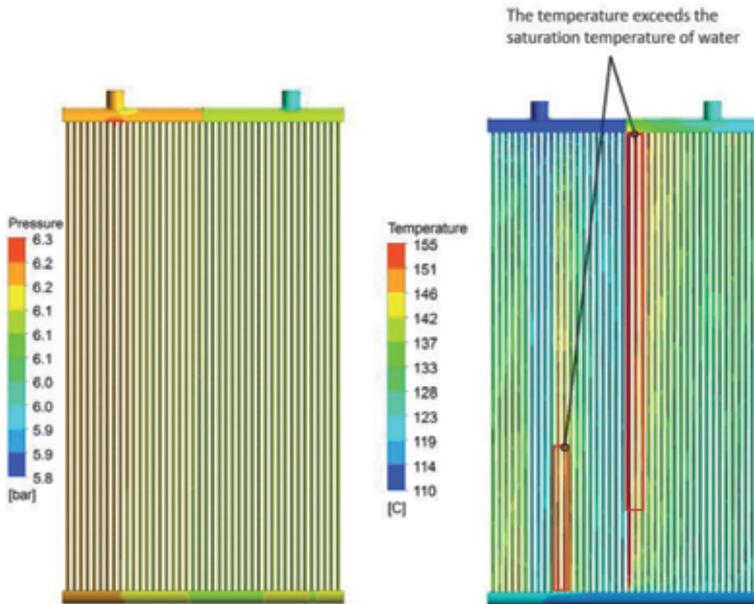


Fig. 15. Pressure and temperature distribution inside high performance heat exchanger - $\dot{m}_{wat} = 19 \text{ kg/s}$

The pressure is lower in the tubes with velocity below \overline{w}_{wat} than in the others. Therefore, the driving force of the flow falls down and liquid doesn't come into this segment of the tubes. E.g., in the first pass of the heat exchanger, liquid flows into the tubes of the heat exchanger, those are located directly under the inlet nozzle pipe. Velocity of liquid in these tubes is larger than \overline{w}_{wat} . Amount of liquid flowing into the other tubes in this pass is much lower so are velocities. For the second pass of the heat exchanger, the component of velocity vector, parallel to the perforated bottom of the collector, is significantly larger than the other components. Therefore, due to the large inertia forces, water encounters difficulties when changing the flow direction by 90 degrees especially when entering the first tubes of the 2nd pass of the heat exchanger. The mean velocity in tube w (calculated as area averaged velocity) is the smallest there. It is even 20 times lower than the mean value (see Fig. 16).

Low flow velocities inside the tubes cause temperature of water to increase significantly. The zones and the locations where saturation temperature is exceeded are marked in red on Fig 15., the vapour bubbles starts to form there. When saturated steam flows inside the tube, the value h_{wat} increases rapidly. As a result, heat transfer conditions in the device change considerably and differ from the design predictions.

The ratio of average velocity in tube w , computed from numerical simulation, to mean velocity of water \overline{w}_{wat} , presented on Fig. 16, shows that difficulties with improper flow distribution inside the heat exchanger exist in all the computational cases. For $\dot{m}_{wat} = 14$

kg/s in the 2nd pass of the heat exchanger, w can be even 50 times smaller than \overline{w}_{wat} , that can lead to fracture of the first tube near the axis of symmetry (2nd pass), because temperature increases and stress level exceeds the allowable limits. Therefore the construction of the collector should be modified to improve the ratios w/\overline{w}_{wat} , for the 1st and the 2nd pass of the high performance heat exchanger.

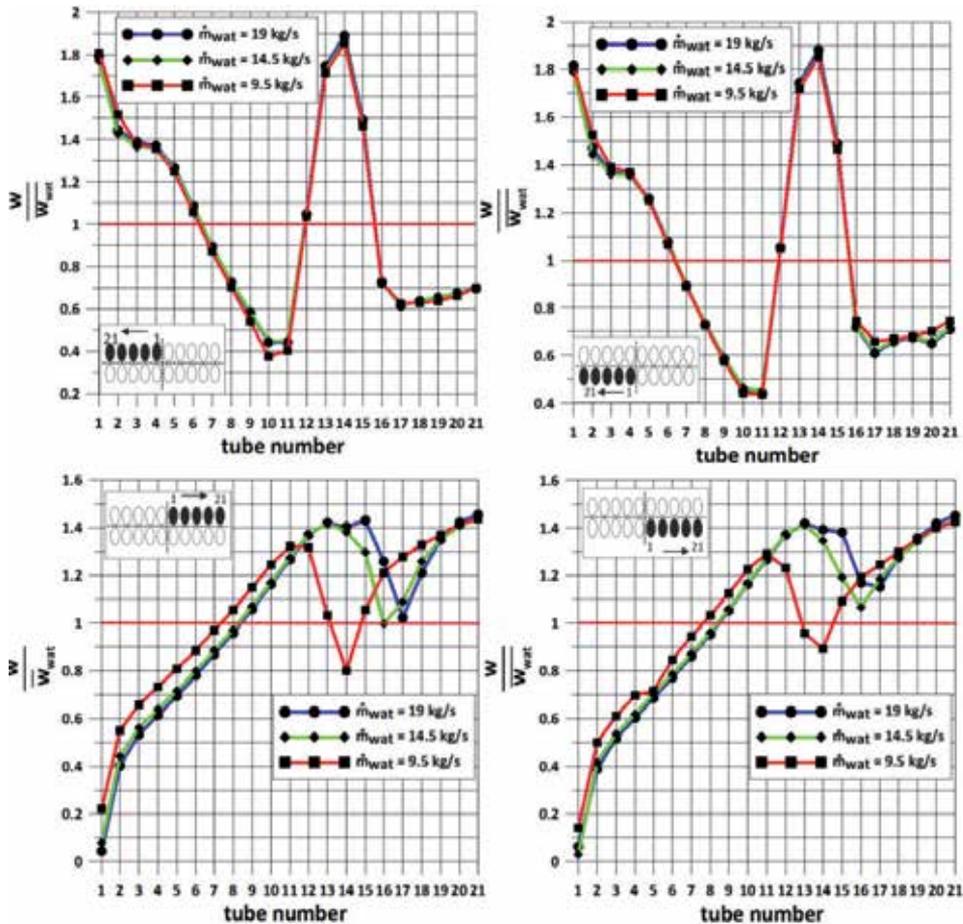


Fig. 16. Ratios of mean velocity in tube w to estimated velocity \overline{w}_{wat} for $\dot{m}_{wat} = 9.5$ kg/s, 14 kg/s and 19 kg/s

The modified version of the collector is presented in the next subchapter.

2.3 Analysis of flow distribution inside modified collector of high performance heat exchanger

As mentioned before, difficulties in ensuring uniform flow velocity inside all the tubes of the high performance heat exchangers have forced us to find a new shape of the collector that improves flow distribution inside the tubes of the 1st and the 2nd pass. The shape of the collector, presented on Fig. 17, is being proposed. The greatest advantages of using it are its simplicity and low costs of manufacturing.

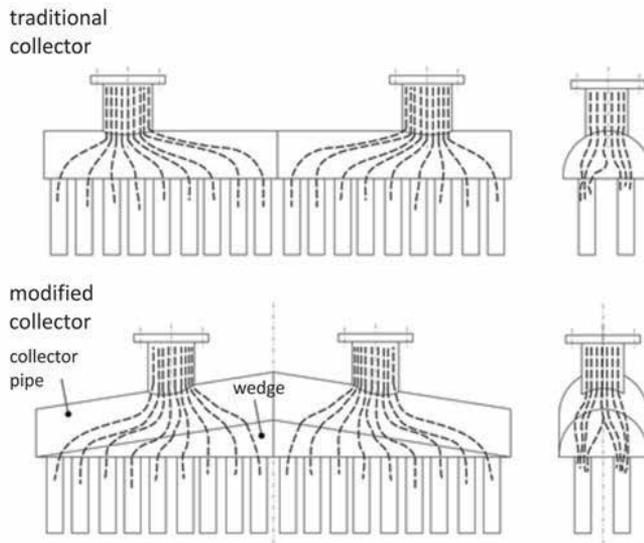


Fig. 17. Traditional and modified shape of collector

The wedge is being welded to the wall of the collector pipe. It increases flow space to enable the more uniform flow distribution inside all the tubes. It is the influence on flow distribution of three parameters (Fig. 18): the h_w , which is the height of wedge, the tube height h_t and x that is the distance between the side of collector and the axis of nozzle pipe, that are investigated.

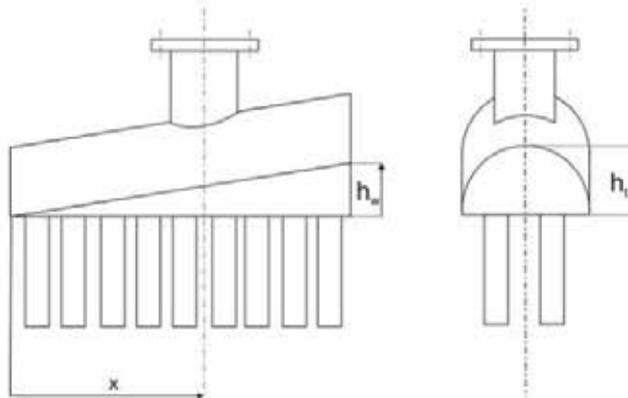


Fig. 18. Geometrical parameters of modified collector

The presented results from analysis are carried out for $\dot{m}_{wat} = 19 \text{ kg/s}$. The ratio of maximal velocity in the tube for the 1st and the 2nd pass of the heat exchanger to mean velocity of water in the tube $w_{max}/\overline{w_{wat}}$, as well as the ratio of minimal velocity for the 1st and the 2nd pass of the heat exchanger to mean velocity of water in the tube $w_{min}/\overline{w_{wat}}$ can be observed on Fig. 19 for $h_w = 40 \text{ mm}$.

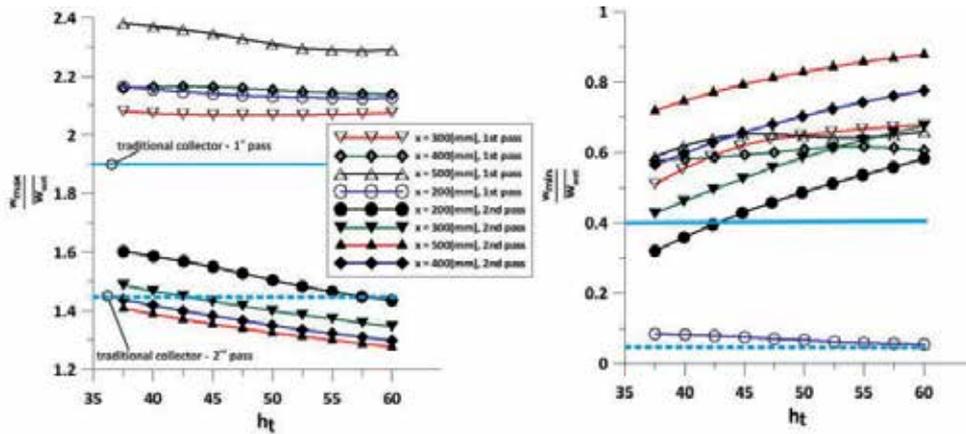


Fig. 19. Highest ratios of maximal and minimal velocity in tube w_{max} and w_{min} to $\overline{w_{wat}}$ for $\dot{m}_{wat} = 19$ kg/s and $h_w = 40$ mm.

The presented results show that it is difficult to obtain geometrical parameters of the collector for that the ratios $w_{max}/\overline{w_{wat}}$ and $w_{min}/\overline{w_{wat}}$ are optimal for the 1st and the 2nd pass of the heat exchanger. For instance, the parameters $x = 500$ mm, $h_t = 38$ mm and $h_w = 40$ mm ensure proper flow distribution to the 2nd pass of the heat exchanger, but in the 1st pass the ratio $w_{max}/\overline{w_{wat}}$ has the highest value. For the presented case it is not as dangerous as low values of $w_{min}/\overline{w_{wat}}$ for the construction; but when the function of the heat exchanger is to cool water, the large value of ratio $w_{max}/\overline{w_{wat}}$ is not desired. The worst results are for $x = 200$ mm, $h_t = 60$ mm and $h_w = 40$ mm, when the ratio $w_{min}/\overline{w_{wat}}$ equals 0.05 for the 1st pass of the heat exchanger. In the case of heating the medium that flows inside the heat exchanger it can bring about significant compressible stresses acting on the tube.

The results obtained for $h_w = 50$ mm doesn't differ significantly from the previously presented. However, it is possible to observe on Fig. 20 that for the 1st pass of the heat exchanger the ratio $w_{min}/\overline{w_{wat}}$ is lower for $x = 500$ mm than for the same values of x and h_t if $h_w = 40$ mm. This means, that the increase in h_w from 40 mm to 50 mm doesn't improve the flow condition inside the device considerably.

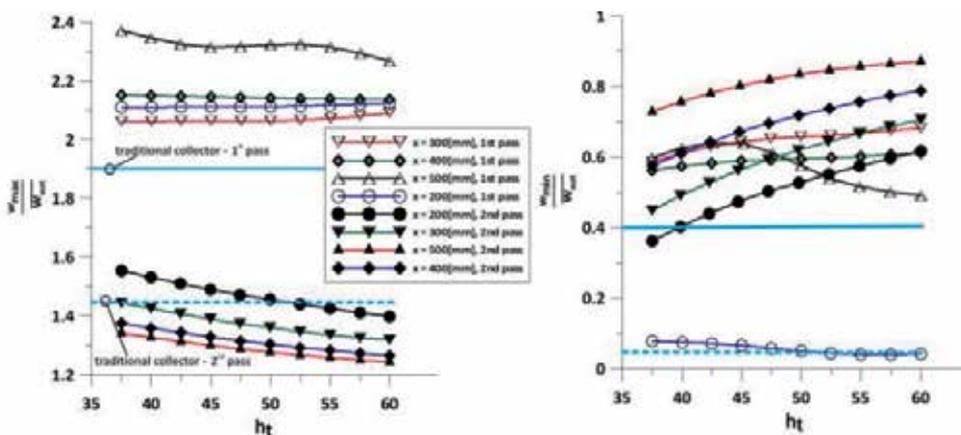


Fig. 20. Highest ratios of maximal and minimal velocity in tube w_{max} and w_{min} to estimated velocity $\overline{w_{wat}}$ for $\dot{m}_{wat} = 19$ kg/s and $h_w = 50$ mm.

The results for $h_w = 60$ mm are shown on the Fig. 21. Extreme positioning of the nozzle pipe: $x = 200$ mm and $x = 500$ mm causes improper flow distribution for the 1st pass of the heat exchanger.

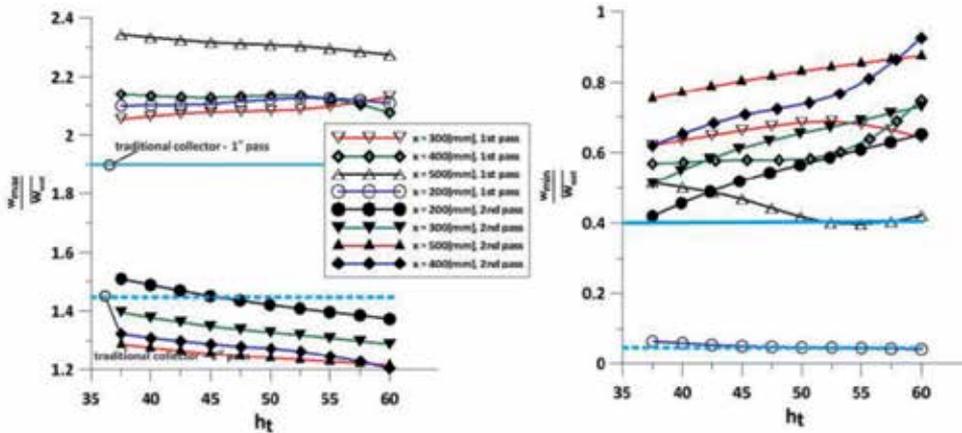


Fig. 21. Highest ratios of maximal and minimal velocity in tube w_{max} and w_{min} to estimated velocity \bar{w}_{wat} for $\dot{m}_{wat} = 19$ kg/s and $h_w = 60$ mm.

The last set of results, Fig. 22, is presented for $h_w = 70$ mm. In contrary to the results presented before, increasing h_w with increasing h_t up to 55 mm, improves w_{min}/\bar{w}_{wat} in the 1st pass of the heat exchanger for $x = 200$ mm. The largest improvement is in ratio w_{min}/\bar{w}_{wat} especially for the 1st pass, that is 0.4 for $x = 200$ mm.

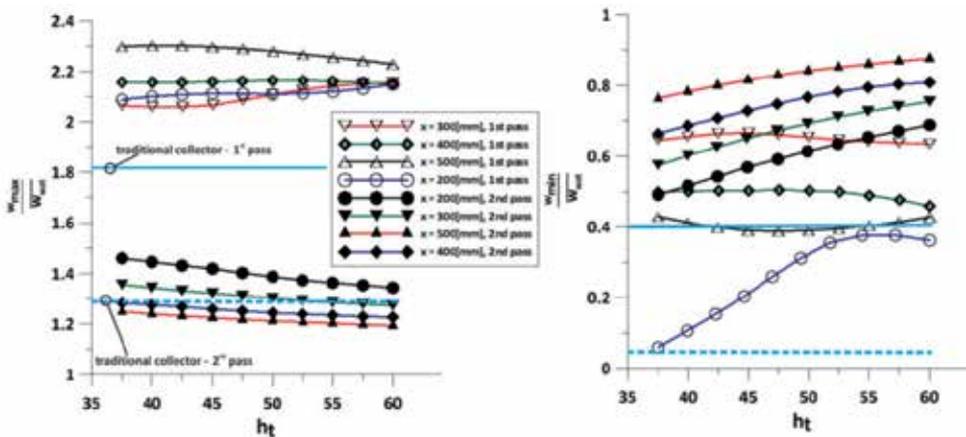


Fig. 22. Highest ratios of maximal and minimal velocity in tube w_{max} and w_{min} to estimated velocity \bar{w}_{wat} for $\dot{m}_{wat} = 19$ kg/s and $h_w = 70$ mm.

All of the presented results show that it is difficult to find optimal solution for geometry presented on Fig.18. However, it is possible to reach compromise between large values of w_{max}/\bar{w}_{wat} and low values of w_{min}/\bar{w}_{wat} , e.g. for the set of parameters $x = 300$ mm, $h_w = 40$ mm and $h_t = 55$ mm. In the 1st pass of the heat exchanger, the w_{max}/\bar{w}_{wat} is slightly larger than for the traditional collector. However, the biggest advantage of using the modified shape of the collector, with the mentioned parameters is that the w_{min}/\bar{w}_{wat} is similar for both the 1st

and the 2nd pass of the heat exchanger and doesn't exceed 0.6. Therefore the differences in the wall temperature will be not as large as for the traditional shape of the collector.

3. Temperature distribution inside elliptical tube

Numerical code in *c++* has been written for the thermal analysis, in order to obtain temperature distribution inside the elliptical tube. The boundary condition for this analysis is taken from the CFD computation (water flow distribution inside the tubes). On the outer surface of the elliptical tube, the h_{cg} and $T_{cg-bulk}$ values are applied. On the inner surface of the elliptical tube, the h_{wat} values obtained for the CFX calculated mass flow rate and mass flow averaged temperature in each tube are taken into computation. If it is necessary, the code can incorporate the boiler scale layer into calculation, as well as fouling resistance on the outer surface of the tube.

The results obtained from thermal analysis for all the tubes are then used as the thermal loading in structural analysis.

Temperature distribution inside the elliptical tube is obtained in the Finite Volume Method via the Finite Element Method (Taler & Duda, 2006; Chang, 2010). The two-dimension steady state heat transfer is assumed; therefore the isoparametric linear quadrilateral elements (Chang, 2010, Lewis, 2004) are applied. The isoparametric element creating the control volume is presented on Fig. 23.

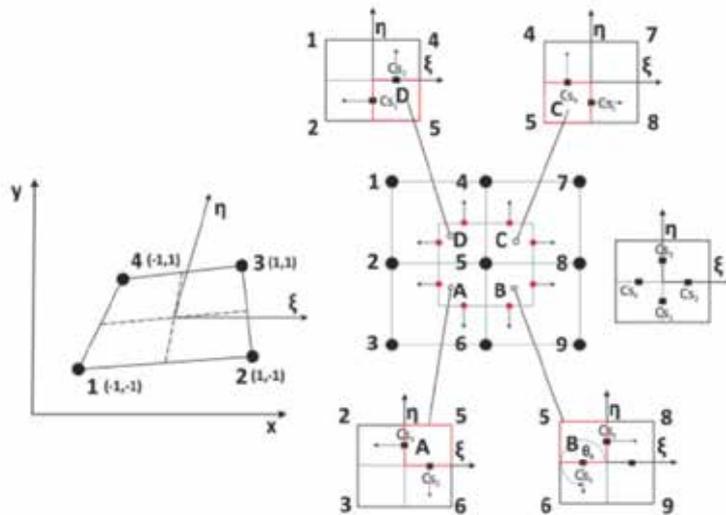


Fig. 23. Finite Volume Method via Finite Elements

The real coordinates (x,y) are mapped into the natural coordinates (ξ,η) . Therefore, the quadrilateral element can be treated as the rectangular element. The lines connecting the middles of the rectangle sides and the nodes of the quad element create the sub-control volumes, as presented on Fig. 23. Four sub-control volumes create one large control volume. Each sub-control volume is also treated as the quadrilateral isoperimetric element; therefore, the shape functions are applied to approximate temperature distribution and coordinates inside the quad element. For the isoperimetric elements, the temperature distribution inside the element domain is estimated, using the following shape function (Taler & Duda, 2006):

$$\begin{aligned}
\Phi_1^e &= \frac{1}{4} \cdot (1 - \xi) \cdot (1 - \eta) \\
\Phi_2^e &= \frac{1}{4} \cdot (1 + \xi) \cdot (1 - \eta) \\
\Phi_3^e &= \frac{1}{4} \cdot (1 + \xi) \cdot (1 + \eta) \\
\Phi_4^e &= \frac{1}{4} \cdot (1 - \xi) \cdot (1 + \eta)
\end{aligned} \tag{11}$$

Temperature inside the element is obtained by multiplying the nodal temperature T_i and the function Φ_i^e that corresponds to the node shape:

$$T(\xi, \eta) = \sum_{i=1}^4 T_i \cdot \Phi_i^e \tag{12}$$

The x and y coordinates inside the element domain are interpolated in a similar manner as temperature:

$$\begin{aligned}
x &= \sum_{i=1}^4 x_i \cdot \Phi_i^e \\
y &= \sum_{i=1}^4 y_i \cdot \Phi_i^e
\end{aligned} \tag{13}$$

As mentioned before, the control volume consists of four sub-control volumes A, B, C, D. The heat flux balance for the control volume can be written in the form:

$$\begin{aligned}
k_A \cdot \sum_{Cs2,3} \left(\frac{dT}{dx} \cdot n_x + \frac{dT}{dy} \cdot n_y \right) \cdot \overline{\Delta\Gamma} + k_B \cdot \sum_{Cs3,4} \left(\frac{dT}{dx} \cdot n_x + \frac{dT}{dy} \cdot n_y \right) \cdot \overline{\Delta\Gamma} + y = \sum_{i=1}^4 y_i \cdot \Phi_i^e \\
k_C \cdot \sum_{Cs4,1} \left(\frac{dT}{dx} \cdot n_x + \frac{dT}{dy} \cdot n_y \right) \cdot \overline{\Delta\Gamma} + k_D \cdot \sum_{Cs1,2} \left(\frac{dT}{dx} \cdot n_x + \frac{dT}{dy} \cdot n_y \right) \cdot \overline{\Delta\Gamma} = 0
\end{aligned} \tag{14}$$

Where n_x and n_y are the components of the normal vector \vec{n} and $\overline{\Delta\Gamma}$ is the boundary vector. Heat flows between control volumes through control surfaces Cs (two for each sub-control volume). In numerical code, the energy equation is applied in the following manner for sub-control volumes A, B, C and D:

- For sub-control volume A:

$$Q_A = k_A \cdot \left(\frac{dT}{dx} \cdot \cos \theta_2 + \frac{dT}{dy} \cdot \sin \theta_2 \right) \cdot \overline{\Delta\Gamma}_2 + k_A \cdot \left(\frac{dT}{dx} \cdot \cos \theta_3 + \frac{dT}{dy} \cdot \sin \theta_3 \right) \cdot \overline{\Delta\Gamma}_3 \tag{15}$$

- For sub-control volume B:

$$Q_B = k_B \cdot \left(\frac{dT}{dx} \cdot \cos \theta_3 + \frac{dT}{dy} \cdot \sin \theta_3 \right) \cdot \overline{\Delta\Gamma}_3 + k_B \cdot \left(\frac{dT}{dx} \cdot \cos \theta_4 + \frac{dT}{dy} \cdot \sin \theta_4 \right) \cdot \overline{\Delta\Gamma}_4 \tag{16}$$

- For sub-control volume C:

$$Q_C = k_C \cdot \left(\frac{dT}{dx} \cdot \cos \theta_4 + \frac{dT}{dy} \cdot \sin \theta_4 \right) \cdot \overline{\Delta\Gamma}_4 + k_C \cdot \left(\frac{dT}{dx} \cdot \cos \theta_1 + \frac{dT}{dy} \cdot \sin \theta_1 \right) \cdot \overline{\Delta\Gamma}_1 \tag{17}$$

- For sub-control volume D:

$$Q_D = k_D \cdot \left(\frac{dT}{dx} \cdot \cos \theta_1 + \frac{dT}{dy} \cdot \sin \theta_1 \right) \cdot \overline{\Delta\Gamma}_1 + k_D \cdot \left(\frac{dT}{dx} \cdot \cos \theta_2 + \frac{dT}{dy} \cdot \sin \theta_2 \right) \cdot \overline{\Delta\Gamma}_2 \tag{18}$$

Sum of fluxes to control volume equals to 0:

$$Q_A + Q_B + Q_C + Q_D = 0 \quad (19)$$

Derivatives of temperature are obtained from the following equation:

$$\begin{aligned} \frac{dT}{dx} &= \frac{1}{|J|} \cdot \left(\frac{dT}{d\xi} \cdot \frac{dy}{d\eta} - \frac{dT}{d\eta} \cdot \frac{dy}{d\xi} \right) \\ \frac{dT}{dy} &= \frac{1}{|J|} \cdot \left(\frac{dT}{d\eta} \cdot \frac{dx}{d\xi} - \frac{dT}{d\xi} \cdot \frac{dx}{d\eta} \right) \end{aligned} \quad (20)$$

And the Jacobian of isoparametric transformation is obtained according to the formula:

$$|J| = \frac{dx}{d\xi} \cdot \frac{dy}{d\eta} - \frac{dy}{d\xi} \cdot \frac{dx}{d\eta} \quad (21)$$

$T(\xi, \eta)$ is differentiated with respect to ξ and η , and the values of $\frac{dT}{d\xi}$ and $\frac{dT}{d\eta}$ are computed in the middle of each control surface (see Tab. 3):

Control Surface	ξ	η
CS ₁	0	-0.5
CS ₂	0.5	0
CS ₃	0	0.5
CS ₄	-0.5	0

Table 3. Values of isoparametric parameters ξ, η for control surfaces

The system of heat balance equations (see eq. 19) is being solved for each control volume. Assuming the mixed type boundary conditions on the inner and the outer wall of tube, the temperature distribution inside the elliptical tube is obtained. The sample results for the tube from the 1st row are presented on the Fig. 24 and Fig 25. Applied boundary conditions on the inner wall of the tube are: $h_{wat} = 5000 \text{ W}/(\text{m}^2 \text{ K})$, $T_{wat-bulk} = 120 \text{ }^\circ\text{C}$. On the outer wall, h_{cg} and $T_{cg-bulk}$ - obtained for $t_s = 0 \text{ mm}$ and adequately $t_s = 1 \text{ mm}$ is assumed - compare with Fig. 11.

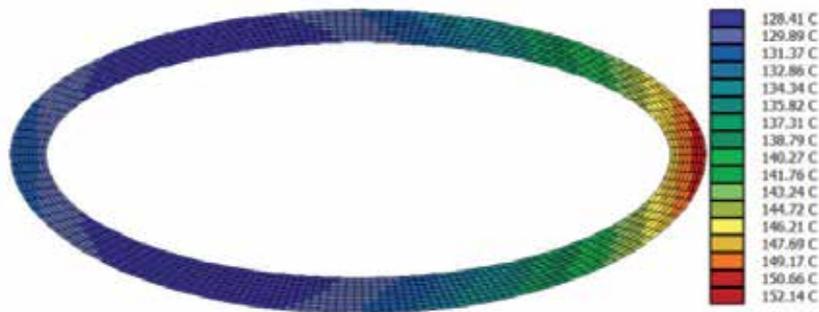


Fig. 24. Temperature distribution for $h_{wat} = 5000 \text{ W}/(\text{m}^2 \text{ K})$, $T_{wat-bulk} = 120 \text{ }^\circ\text{C}$, $t_s = 0 \text{ mm}$

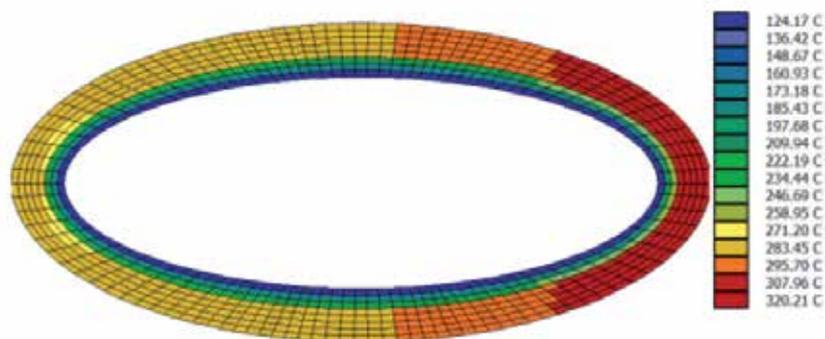


Fig. 25. Temperature distribution for $h_{wat}=5000 \text{ W}/(\text{m}^2 \text{ K})$, $T_{wat-bulk}=120 \text{ }^\circ\text{C}$, $t_s = 1 \text{ mm}$

It is possible to observe, on Fig. 25, that when boiler scale exists inside the tube, e.g. Fig. 26, temperature increases significantly.



Fig. 26. Boiler scale observed in tubes of heat exchanger

The large increase of temperature brings about considerable thermal stresses inside all the tubes that may lead to fracture of the tube; see Fig. 1 & Fig. 2.

4. Analysis of stress state inside construction of high performance heat exchanger

As mentioned before, structural analysis is carried out on basis of the Finite Element Method using the commercial code ANSYS. The manner in which the construction is supported is illustrated on Fig. 27. As it can be observed, the heat exchanger consists of two collectors with perforated bottoms to which all the tubes are welded. A shell connects the collectors and flat bars welded to the shell stiffen its structure. The frame increases the stiffness of the perforated bottom.

The other components of construction (a confuser and a diffuser), not shown on Fig. 27 are welded to the shell. These parts constrain structure movement in the y direction. The heat exchanger is also clamped at the distance $z = 525 \text{ mm}$ from the upper perforated bottom. The isoparametric linear shell element, SHELL 181 (ANSYS, 2009), used for the structural analysis is presented on Fig. 28. The CAD model of the heat exchanger is divided into a set

of finite elements. Each part of the construction consists of elements with different thickness, t_{shell} , (see Tab. 4). Thickness of heat exchanger components are listed on Table 4.

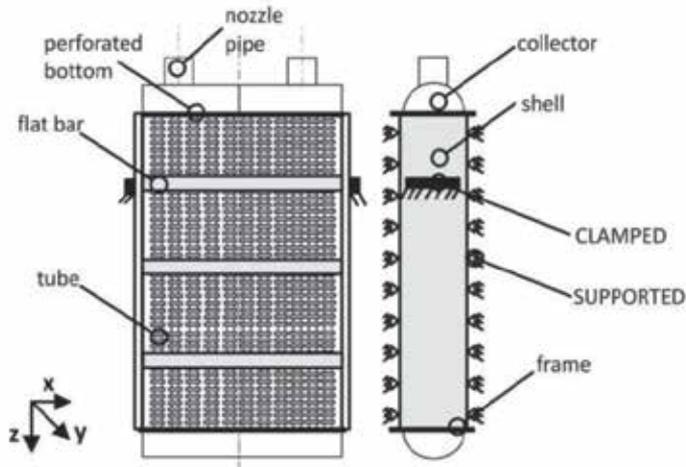


Fig. 27. Construction and support conditions for heat exchanger

Component	t_{shell} [mm]
tube	2
flat bar	15
perforated bottom	12
nozzle pipe	2.5
frame	20
collector	6
shell	8

Table 4. Thickness of heat exchanger components

The midsurface is geometrical representation of the element. The maximal stress value is computed at the outer fibbers of the element.

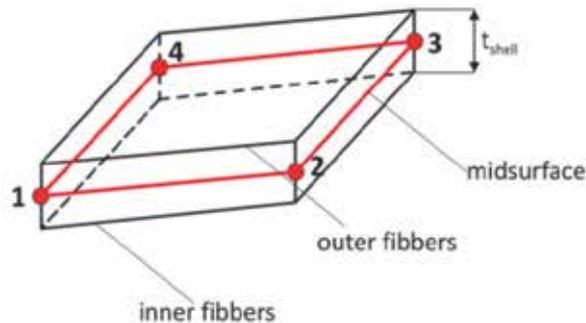


Fig. 28. Isoparametric shell element for structural calculations

The set of equation being solved in structural analysis (ANSYS, 2009) is presented below:

$$[K_e] \cdot \{u\} - \{F_e^{th}\} = [M_e] \cdot \{\ddot{u}\} + \{F_e^{pr}\}, \quad (22)$$

where the element stiffness matrix $[K_e]$ is obtained from the following formula:

$$[K_e] = \int_V [B]^T \cdot [D] \cdot [B] dV, \quad (23)$$

where $[B]$ is the strain displacement matrix based on the element shape functions, $\{u\}$ is the nodal displacement vector and $[D]$ is the elasticity matrix. The thermal load vector is calculated on the basis of the equation:

$$\{F_e^{th}\} = \int_V [B]^T \cdot [D] \cdot \{\varepsilon_{th}\} dV, \quad (24)$$

The thermal strain vector $\{\varepsilon_{th}\}$ is calculated on the basis of the following dependence:

$$\{\varepsilon_e^{th}\} = \alpha \cdot (T - T_{ref}), \quad (25)$$

where α is the thermal expansion coefficient. Nodal temperatures T are obtained from the thermal calculations, section 3. T_{ref} is the reference temperature assumed as 22 °C.

The element mass matrix $[M_e]$ is calculated on the basis of the following equation:

$$[M_e] = \rho \cdot \int_V [N]^T \cdot [N] \cdot [D] dV, \quad (26)$$

where $[N]$ is the element shape function matrix and ρ is material density that the element consists of. For the computation steel P355NH is used. Yield strength Y_s of this material equals 355 MPa and the density $\rho = 7850$ kg/m³. The element pressure vector $\{F_e^{pr}\}$ is calculated according to the formula:

$$\{F_e^{pr}\} = \int_A [N_n]^T \cdot \{P\} dA, \quad (27)$$

where $[N_n]$ is the matrix of shape functions for normal motions at the surface and $\{P\}$ is the nodal pressure vector assumed equal to p_{op} - Tab. 2.

The material properties used in the structural analysis are presented in Tab. 5.

Name	Symbol	Value
Young's modulus	E [MPa]	$2.1 \cdot 10^5$
Tangent modulus	T [MPa]	470
Thermal expansion coeff.	α [1/K]	$1 \cdot 10^{-5}$
Poisson's ratio	ν	0.3

Table 5. Material properties of P355NH steel

4.1 Analysis of stress state for construction of high performance heat exchanger with traditional collectors

Results of numerical computation for nodal displacements, compressible stresses and equivalent stress acc. to von Mises are presented in this subchapter. Computations are carried out using the model of bilinear isotropic hardening (ANSYS, 2009; Chakrabarty, 2006), because computations in elastic range show that the allowable stresses inside the construction are exceeded.

Thermal loadings for all the tubes are obtained using a program described in section 3 (computational case with $t_s = 0$ mm). Because combustion gas flow is on the outer surface of the elliptical tubes, where heat flux density is the highest - avg. 80000 W/m². Heat flux density on the outer wall of perforated bottoms, collectors and nozzle pipes is low - avg. 5000 W/m². Therefore, temperature of walls for these parts is assumed to be equal to near wall temperature of liquid obtained from the flow distribution analysis - subsection 2.2. Thermal loading of construction is presented on Fig. 29. The contour map of nodal displacements $\{u\}$ is presented on Fig. 30.

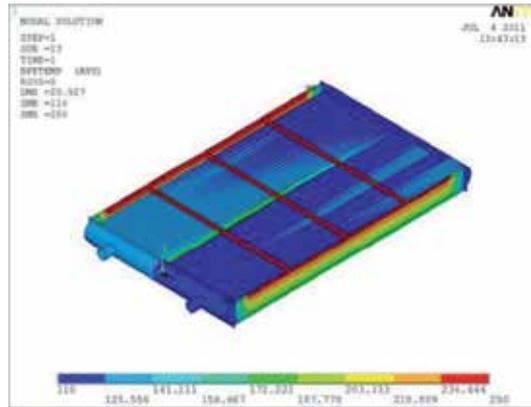


Fig. 29. Thermal loading for construction of high performance heat exchanger - traditional collector

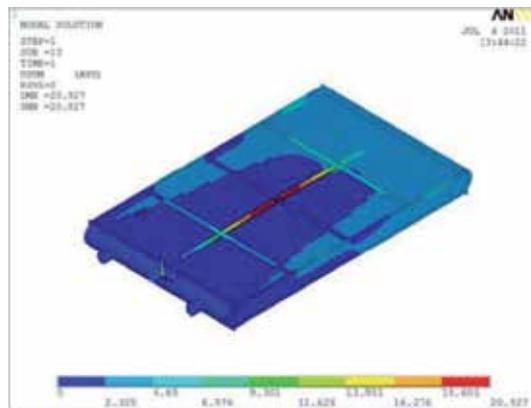


Fig. 30. Nodal displacement - high performance heat exchanger with traditional collectors

Displacements of the elliptical tube, that is the nearest to the axis of symmetry of the high performance heat exchanger are the largest at all. Comparing the results presented on Fig. 30 to Fig. 2, it is evident that reasons for the heat exchanger to break down are improper heat transfer conditions inside the construction. In the tube that broke down the temperature was significantly higher than in the others.

The next two figures - Fig. 31 & Fig. 32 present the contour of equivalent stress distribution σ_e acc. to von Mises (Bower, 2009), and the map of compressible stresses σ_3 inside the construction. The von Mises stress is calculated according to the formula:

$$\sigma_e = \frac{1}{2} \cdot \sqrt{(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2}, \quad (28)$$

where σ_1 , σ_2 and σ_3 are principal stresses. It is possible to observe that in the shell and in the tube that was broken the stress level is the largest. According to the adopted plasticity model, maximum compressible stress in the broken tube equals to 205 MPa.

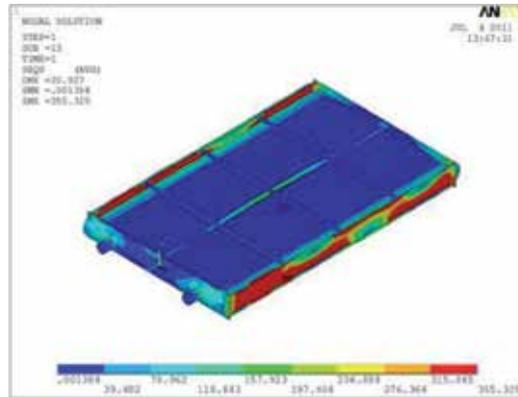


Fig. 31. Equivalent stress acc. to von Mises – heat exchanger with traditional collector

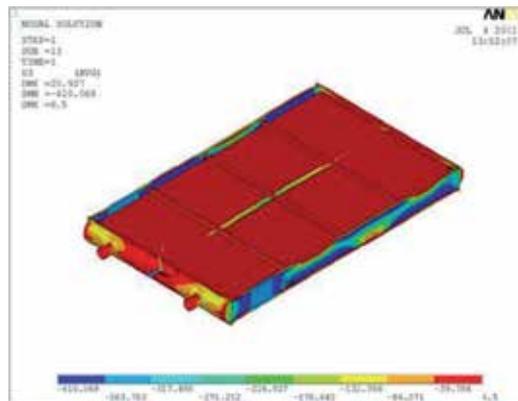


Fig. 32. Compressible stresses – heat exchanger with traditional collector

Because the tube is slender and thin, it is prone to buckling, that is caused by compressible loads. Buckling computations (Brush & Almroth, 1975) are carried out in order to understand the manner how the tube was destroyed. The assumed model of buckling is presented on Fig 33. The left end is clamped as opposed to the right one that is supported in vertical direction.



Fig. 33. Buckling model of elliptical tube

The effective length of the tube equals:

$$L_e = K \cdot L, \quad (29)$$

where K equals to the tube effective length factor and L is the unsupported length of the tube.

The slenderness ratio s for the tube presented on Fig. 33 equals:

$$s = \frac{L_e}{\sqrt{\frac{I_{min}}{A}}}, \quad (30)$$

where I_{min} equals to minimal area moment of inertia for the tube cross section and A equals to the cross - sectional area of the tube. The critical slenderness ratio s_{cr} is calculated according to the formula:

$$s_{cr} = \pi \cdot \sqrt{\frac{E}{Y_s}}. \quad (31)$$

The following values, Tab. 6, are used in calculations:

Name	Symbol	Value
Tube effective length factor	K	0.5
Length of unsupported tube	$L[m]$	2.1
Minimal area moment of inertia	$I_{min}[m^4]$	$4.85 \cdot 10^{-9}$
Cross - sectional area of tube	$A[m^2]$	$1.7 \cdot 10^{-4}$
Poisson's ratio	ν	0.3
Critical slenderness ratio	s_{cr}	73.1
Slenderness ratio	s	168

Table 6. Parameters taken into buckling calculations

Since s is larger than s_{cr} , see Tab.6, buckling occurs in the elastic range. The following formula allows calculating the critical compressible stress:

$$\sigma_{crit} = \frac{\pi^2 \cdot E}{s^2}. \quad (32)$$

The allowable stresses obtained on the basis of procedure (28 - 32) are less than $\sigma_{crit} = 66$ MPa. Comparing σ_{crit} to the maximal value from the contour map of compressible stresses in the construction, it is evident, that the critical stress is exceeded in the tube nearest to the axis of symmetry. This explains the manner of tube deformation, after the damage, compare with Fig 2.

4.2 Analysis of stress state for construction of high performance heat exchanger with modified collectors

For the construction of the high performance heat exchanger, computations are carried out in the elastic range, because the σ_e for this computations is lower than Y_s . On Fig. 34 & Fig. 35 thermal loading and the nodal displacements of construction are presented.

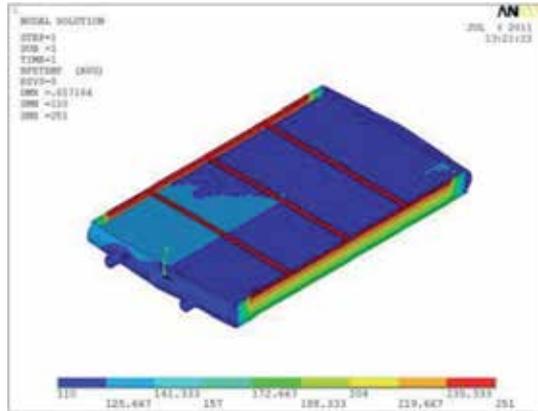


Fig. 34. Thermal loading for construction of high performance heat exchanger – modified collectors

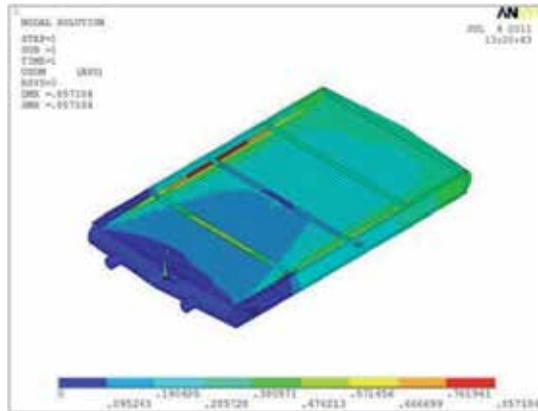


Fig. 35. Nodal displacement – high performance heat exchanger with modified collectors

It is possible to observe that, for the proposed construction of the collector, displacements are significantly lower than for the typical one. The maximum displacement is over 20 times lower – compare Fig. 35 and Fig. 30.

For the modified construction, distribution of equivalent stress acc. to von Mises is presented on Fig. 36. The σ_e value obtained from computations does not exceed Y_S , therefore the construction is safe to use. Moreover the compressible stresses σ_3 in the tubes, Fig. 37, do not exceed 22 MPa and are significantly lower than σ_{crit} .

The modified construction of the collectors improves flow distribution inside the tubes through the heat exchanger. Consequently, thermal load acting on the tubes is lower than in the case of the traditional construction of the collectors. Comparing the maps of nodal displacements $\{u\}$, effective stresses σ_e according to von Mises and compressible stresses σ_3 for the traditional and the modified construction of the collectors, it is evident that the proposed construction carries thermal loads better than the one actually available in manufacturing.

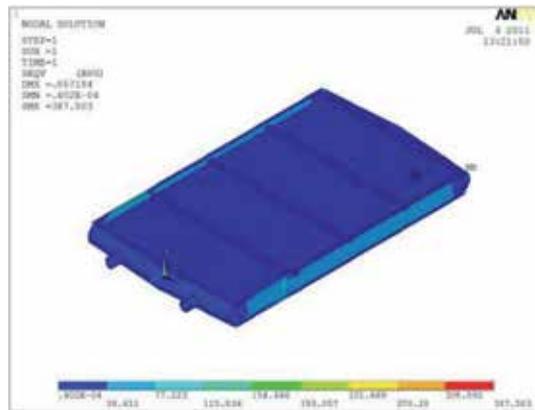


Fig. 36. Equivalent stress acc. to von Mises – heat exchanger with modified collector

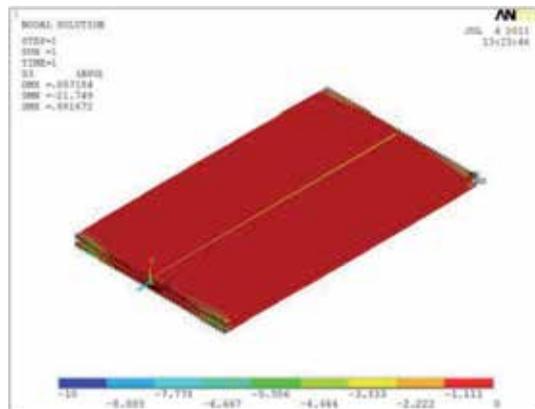


Fig. 37. Compressible stresses – heat exchanger with modified collector

5. Conclusions

Improper liquid distribution in the tubes of the presented heat exchanger causes the mass flow rates of medium flowing inside all the tubes being far from uniform. This implicates that heat transfer conditions inside the tubes differ significantly. Consequently, the temperatures as well as the thermal stresses inside all the tubes are different, often exceeding the allowable stress limits.

Reasons of high performance heat exchanger failure, on the basis of the real model applied in the industry, were investigated. The zones and the locations where the allowable stresses are exceeded were presented in the paper. Presented analysis may be applied not only in analysing the stress state in the high performance heat exchanger, but also in accessing safety of boilers, pressure vessels, shell and tube heat exchangers and many other devices applied in the thermal power engineering.

Furthermore, the new and easy in manufacturing model of collectors is presented. On the basis of numerical computation it was proven that application of the new collectors ensures

proper flow distribution inside all the tubes that consequently reduces significantly thermal stresses inside the construction.

With continuous increase in computational power it will be possible to investigate in more detail the flow behaviour and heat transfer conditions inside the high performance heat exchanger, including e.g. the phase change effect. If liquid evaporates, the heat transfer coefficient increases significantly and due to the local pressure drop, the cavitation may occur inside the tube that leads to erosion of the tube wall. Consequently, material of the tube is damaged and less resistible on acting stresses. Therefore, the real values of σ_e and σ_3 may be larger than the presented in the paper. This effect should be encountered into calculations in the future.

Wall temperature increases significantly in the tubes where the mass flow rates are low. If water is not prepared properly, boiler scale formation is more intensive than in the other tubes. Influence of boiler scale thickness and its thermal conductivity on the stress state in the high performance heat exchanger should be investigated in the future.

6. References

- Anderson, J. D. (1995). *Computational Fluid Dynamics: The Basics with Applications 6th Ed.*, McGraw Hill, ISBN: 0070016852
- ANSYS, (2008). *CFX – Solver Theory Guide*,
<http://www1.ansys.com/customer/content/documentation/120/cfx/xthry.pdf>, USA
- ANSYS, (2009). *Theory Reference for the Mechanical ADPL and Mechanical Applications*,
http://www1.ansys.com/customer/content/documentation/121/ans_thry.pdf
- Bansat, B.; Mtiller-Steinhagen, H. (1993). *Crystallization fouling in plate heat exchangers.*, ASME Journal of Heat Transfer, pp. 584-591, ISSN 0022 -1481, Vol. 115, No 3
- Bird, R. B. (2001). *Transport Phenomena 2nd Edition*, Wiley, ISBN-10: 0471410772, ISBN-13: 978-0471410775
- Bordallo, S. N. & Saboya, F. E. M. (1999). *Pressure Drop Coefficients for Elliptic and Circular Sections in One, Two and Three-Row Arrangements of Plate Fin and Tube Heat Exchangers*, Journal of the Brazilian Society of Mechanical Sciences, Vol. 21, No. 4, ISSN 0100-7386
- Bower, A. F. (2009). *Applied Mechanics of Solids*, CRC Press, ISBN-10: 9781439802472, ISBN-13: 978-1439802472
- Brush, D. O & Almroth, B. O. (1975). *Buckling of bars, plates, and shells*, McGraw Hill, ISBN-10: 0070085935, ISBN-13: 978-0070085930
- Chakrabarty, J. (2006). *Theory of Plasticity*, Butterworth Heinemann, ISBN-10: 0-7506-6638-2, ISBN-13: 978-0-7506-6638-1
- Chung, T.J. (2010): *Computational Fluid Dynamics 2nd Ed.*, Cambridge University Press, ISBN-10: 0521769698, ISBN-13: 978-0521769693
- Egorov, Y. & Menter F. (2008). *Development and Application of SST-SAS Turbulence Model in the DESIDER Project*, Notes on Numerical Fluid Mechanics and Multidisciplinary Design, pp. 261 – 270, Vol. 97
- Ferziger, J. H & Peric, M. (1999). *Computational Methods for Fluid Dynamics 2nd Ed.*, Springer, ISBN: 3540653732
- Freeborn, J. & Lewis, D. (1962). *Initiation of boiler scale formation*, Journal of Mechanical Engineering Sciences, pp. 46 – 52, Vol. 4

- Gnielinski, V. (1976). *New equations for heat and mass transfer in turbulent pipe and channel flow*, International Journal of Chemical Engineering, pp. 359–367, Vol. 16, No. 1
- Gresho, P. M. & Sani, R. L. (2000). *Incompressible Flow and the Finite Element Method*, Willey, ISBN: 0471492493
- Hutton, D. V. (2003). *Fundamentals of Finite Element Analysis*, McGraw Hill, ISBN - 10 0072922362, ISBN - 13 978-0072922363
- Ibrahim, T. A. & Gomaa, A. (2009). *Thermal performance criteria of elliptic tube bundle in crossflow*, International Journal of Thermal Sciences, pp. 2148-2158, Vol. 48, No. 11
- Idem, S.; Jacobi, A. M. & Goldschmidt, V. W. (1990). *Heat Transfer Characterization of a Finned-Tube Heat Exchanger (with and without Condensation)*. Transaction of the ASME, pp. 64-70, Vol. 112
- Incopera, F. P & De Witt, D. P. (2001). *Fundamentals of Heat and Mass Transfer 5th Ed.*, Willey, ISBN-10: 0471386502, ISBN-13: 978-0471386506
- Jang, J. Y.; Yang, J. Y. (1998). *Experimental and Numerical Analysis of the Thermal-Hydraulic Characteristics of Elliptic Finned-Tube Heat Exchangers*, Heat Transfer Engineering, pp. 55-67, Vol.19, No. 4
- Khan, M.G.; Fartaj, A. & Ting, D. S. K. (2004). *An experimental characterization of cross-flow cooling of air via an in-line elliptical tube array*, International Journal of Heat and Fluid Flow, pp. 636-648, Vol. 25, No. 4
- Lewis, R. W.; Nithiarasu, P. & Seetharamu, K. N. (2004). *Fundamentals of the finite element method for heat and fluid flow*, Willey, ISBN-10: 0470847891, ISBN-13: 978-0470847893341
- Li, B.; Feng, B.; He, Y.L. & Tao W.Q. (2006). *Experimental study on friction factor and numerical simulation on flow and heat transfer in an alternating elliptical axis tube*, Applied Thermal Engineering, pp. 2336-2344, Vol. 26, Issue 17-18
- Logan, D.L. (2011). *A First Course in the Finite Element Method 5th Ed.*, CL-Engineering, ISBN-10 0495668257, ISBN-13 978-0495668251
- Łopata, S. & Ocloń, P. (2010) *Coupled thermal – structural analysis of the high performance heat exchanger with finned elliptical* (in polish), Systems, Technologies and Energetic Devices, Collective Work edited by J. Taler, Cracow University of Technology press, pp. 826 – 840, Cracow
- Łopata, S. & Ocloń, P. (2010). *Investigation of the flow conditions in a high-performance heat exchanger*, Archives of Thermodynamics; pp. 37-50, Vol. 31, no. 3, Gdańsk
- Łopata, S. & Ocloń, P. (2011). *The Stress State analysis of the high performance heat exchanger*, Technical Transactions; Cracow University of Technology press, pp. 360-368, Issue 7, Year 108, Cracow
- Matos, R. S.; Vargas, J. V. C.; Laursen, T. A. & Bejan, A. (2004). *Optimally staggered finned circular and elliptic tubes in forced convection*, International Journal of Heat and Mass Transfer, pp. 1347-1359, Vol.47, No. 6-7
- Nagatha, H. (2004). *The Development of Heat Exchanger with Elliptic Tubes for High Viscous Liquid*: Mitsui Zosen Technical Review, pp. 29-33 ISSN: 0026-6825, No. 183, Japan
- Najibi, S. H; Miiller-Steinhagen, H. & Jamialahmadi, M. (1997). *Calcium sulphate scale formation during subcooled flow boiling*, Chemical Engineering Science, pp. 1265 1284, Vol 52, No. 8, Great Britan

- Nebot, E.; Casanueva, J. F.; Casanueva, T. & Sales, D. (2007). *Model for fouling deposition on power plant steam condensers cooled with seawater: Effect of water velocity and tube material*, International Journal of Heat and Mass Transfer, pp. 3351–3358, Vol. 50
- Nishiyama, H.; Ota, T. & Matsumo, T. (1988). *Heat Transfer and Flow around an Elliptic Cylinders in Tandem Arrangement*, JSME International Journal Series II, ISSN 0914-8817, pp. 410-419, vol. 31
- Oclon, P. & Łopata, S. (2011). *Modelling of the flow distribution inside the collectors of the high performance heat exchanger*, Technical Transactions Mechanics 4-M/2011/B, pp. 393 – 400, Cracow University of Technology press, , Cracow
- Quan, Z., Chen, Y. & Chongfang, M. A. (1997). *Experimental Study of Fouling on Heat Transfer Surface During Forced Convective Heat Transfer*, Chinese Journal of Chemical Engineering, pp. 535 – 540, Vol. 16, No. 4
- Reddy, J. N. (1993). *Introduction to Finite Element Method*, McGraw Hill, ISBN10: 0070513554, ISBN- 13: 978-0070513556
- Rocha, L. A. O.; Saboya, F. E. M. & Vargas, J.V.C. (1997). *A comparative study of elliptical and circular sections in one- and two-row tubes and plate fin heat exchangers*, International Journal of Heat and Fluid Flow, pp. 247–252, Vol. 18, Issue 2
- Saboya, F. E. M. & Sparrow, E. M. (1974). *Local and Average Transfer Coefficients for One-Row Plate Fin and Tube Heat Exchanger Configurations*, Journal of Heat Transfer, pp. 265-272, Vol. 96
- Saboya, S. M. & Saboya, F. E. M. (1981). *Transfer Coefficients for Plate Fin and Elliptical Tube Heat Exchangers*, Proceedings of the VI Brazilian Congress of Mechanical Engineering, pp. 153-162
- Smith, I. M & Griffiths, D. V. (2004). *Programming the Finite Element Method*, Willey, ISBN-10 047084969X, ISBN-13 978- 0470849699
- Taler, D. (2009). *Dynamics of Tube Heat Exchangers*, Monograph 193, AGH University of Science and Technology Press, ISSN 0867-6631, Cracow, Poland (in Polish)
- Taler, J. & Duda, P. (2006). *Solving Direct and Inverse Heat Conduction Problems*, Springer, ISBN 978-3-540-33470-5
- White F. (2003). *Fluid Mechanics*, McGraw Hill, ISBN - 10 0071215662, ISBN - 13 978-0071215664
- Zienkiewicz, O. C. & Taylor, R. L. (2000). *The Finite Element Method 5th Ed.*, Elsevier, Vol.1-3, ISBN 978-0-7506-5055-7

Simulation of H₂-Air Non-Premixed Flame Using Combustion Simulation Technique to Reduce Chemical Mechanisms

Kazui Fukumoto and Yoshifumi Ogami
*Ritsumeikan University, 1-1-1 Noji-Higashi, Kusatsu, Shiga
Japan*

1. Introduction

Combustion simulation that uses computational fluid dynamics (CFD) has been widely adopted as the design tool for combustion equipment. Because flow inside such equipment is generally turbulent, turbulence and combustion models are needed to simulate combustion; many combustion simulations have been performed to verify a system's internal state, such as velocity, pressure, mole fractions of chemical species, and temperature.

Combustion simulation of a confined impinging jet reactor has been performed by the large eddy simulation (LES) model serving as the turbulence model and the presumed probability density function (PDF) serving as the combustion model (Daniele, 2009). The analysis shows that a confined impinging jet reactor is indeed an interesting device because of its high mixing efficiency and absence of stagnant and recirculation zones.

Under the condition of moderate or intense low-oxygen dilution (MILD), the effect of H₂ on H₂-CH₄ turbulent non-premixed flames was investigated with the improved standard $k-\epsilon$ model as the turbulence model and with the eddy dissipation concept (EDC) model (Amir et al., 2010). Simulation results show that H₂ addition to CH₄ leads to improved mixing, increase in turbulent kinetic energy decay along the flame axis, increase in flame entrainment, higher reaction intensities, and increase in mixture ignitability and rate of heat release.

Although combustion simulation was considered to be an efficient designing tool, considerable computational time was needed to calculate the chemical reaction. Combustion models that detail chemical mechanisms require reaction calculations involving n -dimensional ordinary differential equations (ODEs) that are solved according to the number of chemical species involved. Therefore, reducing computation time for the combustion simulation is a significant problem. If computation time could be easily reduced according to the required prediction accuracy, we would be able to obtain the results more quickly. For example, O is a significant species whose mass fraction is necessary to compute the amount of NO present; therefore, the accuracy of the mass fraction of O cannot be neglected. To determine this mass fraction with sufficient accuracy, it is necessary to build a reduced mechanism including O. Generally, a quasi-steady state or partial equilibrium is assumed when the reduced mechanism is built (Warrants et al., 2006).

The chemical equilibrium method (Nagai et al., 2002) does not use reaction equations; instead, the equilibrium composition of a chemical system is determined by minimizing the

Gibbs free energy, which is subject to the conservation of the chemical elements involved in the combustion. The composition is obtained by solving a system of simultaneous nonlinear equations with five unknowns (one for each of the elemental species involved in that study, namely, C, H, O, and N, plus one). In another study, a partial non-premixed flame of CH₄ was modeled using the vortex method in combination with the chemical equilibrium method (Ogami et al., 2010). Although this method did not use any chemical equations, the calculated temperature and main products were in good agreement with experimental results. In some cases, the data obtained from the simulation concerning the minor products did not agree with the experimental results.

A combustion simulation technique based on a combination of the chemical equilibrium method and the EDC model was developed in a previous study and validated by simulation of a H₂-air turbulent non-premixed flame (Fukumoto et al., 2010) and a CO-H₂-air turbulent non-premixed flame (Fukumoto et al., 2009). An advantage of the technique is the ease with which a reduced chemical mechanism can be built according to the accuracy requirement for the chemical species. Thus, the technique can predict intermediate species with high accuracy when a reduced mechanism for the minor species is built. The key to this proposed technique is the incorporation of the concept of equilibrium into the chemical kinetics, which decreases the dimensionality of the chemical reaction calculation and reduces the computational time for a combustion simulation.

This chapter presents a technique for combustion simulation based on a chemical equilibrium method and an EDC model (Magnussen, 1989, 2005). The advantage of the proposed technique is that it makes it easy for us to construct a reduced mechanism according to the accuracy requirements of the chemical species. The proposed technique is validated through a simulation of a H₂-air non-premixed flame (Barrow et al., 2003a). The results computed using the proposed technique are then compared with experimental data and computational data obtained by the EDC model.

2. Method

Ansys FLUENT version 13.0 (Ansys, Inc., 2011) was used in this study. The temperature, enthalpy, and reaction rate of the chemical species were calculated by a user-defined function (UDF) in FLUENT. The UDF is a function that can be programmed to load dynamically with the FLUENT solver in order to enhance the standard features of FLUENT. Moreover, the momentum equation, turbulence model, and mass fraction equation were computed using Ansys FLUENT.

To solve the momentum equation, velocity and pressure were coupled by the SIMPLE method; the basic Reynolds stress model with a function of the enhanced wall treatment in Ansys FLUENT (Ansys, Inc., 2010) was chosen as the turbulence model. The temperature was obtained by computation of the following energy equation:

$$\frac{\partial(\bar{\rho}\tilde{u}_i\tilde{h})}{\partial x_i} = -\frac{\partial}{\partial x_i}(\bar{q}_J + \bar{\rho}\tilde{u}_i h') + \bar{q}_{\text{rad}}, \quad (1)$$

$$\bar{\rho}\tilde{u}_i h' = -\frac{\mu_t}{\sigma_h} \frac{\partial \tilde{h}}{\partial x_i}, \quad (2)$$

$$\bar{q}_J = -\lambda \frac{\partial \bar{T}}{\partial x_i}, \quad (3)$$

where $\bar{\rho}$ represents the density, \tilde{u}_i represents the gas velocity of i components, \tilde{h} is the enthalpy, \bar{q}_{rad} is the source term of the radiation, x_i is the coordinate in the i direction, μ_t is the turbulent viscosity, σ_h is the effective Prandtl number, λ is the heat conductivity, and \bar{T} is the temperature. The symbols “ \sim ” and “ $-$ ” placed over the variables indicate density weighted and time average, respectively. \bar{q}_{rad} was modeled as described on the TNF web site (Barrow, 2003b), and \bar{T} was computed as

$$\bar{T} = T_0 + \frac{\tilde{h} - \sum_J \tilde{Y}_J h_{J0}}{Cp_J}, \quad (4)$$

where $T_0 = 298.15$. h_{J0} represents the standard enthalpy of the formation of chemical species J , \tilde{Y}_J is the mass fraction of chemical species J , and Cp_J is the mean specific heat at constant pressure of chemical species J .

2.1 Combustion simulation technique

The proposed technique is explained using the example of a H₂-air turbulent non-premixed flame. Figure 1 illustrates the scheme of the proposed technique. In this technique, chemical species are divided, roughly, into two groups: group a—the species categorized by the chemical equilibrium, and group b—the species categorized by the chemical kinetics (including the overall reaction).

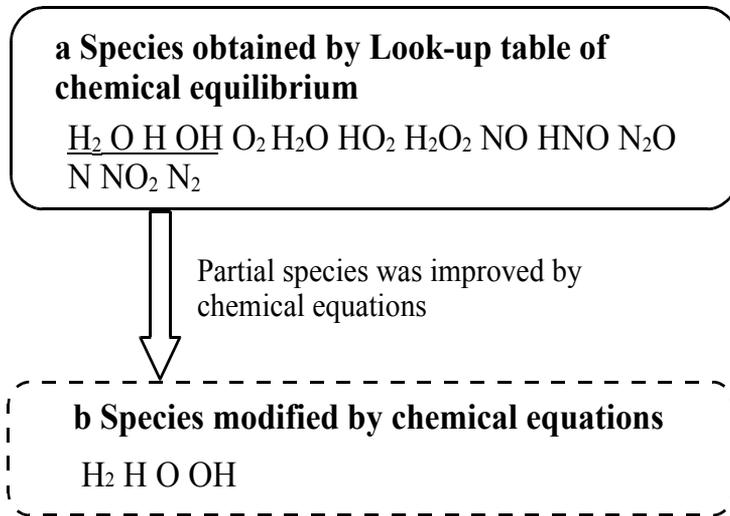


Fig. 1. Overview of proposed technique

In the proposed technique, the mass fraction equation of chemical species J was used as

$$\frac{\partial}{\partial x_j} (\bar{\rho} \tilde{u}_j \tilde{Y}_J) = -\frac{\partial}{\partial x_j} (\bar{j}_{jJ} + \bar{\rho} \tilde{u}_j \tilde{Y}_J') + \bar{w}_J, \quad (5)$$

$$\bar{j}_{jj} = -\bar{\rho} D_J \frac{\partial \bar{Y}_J}{\partial x_j}, \quad (6)$$

$$\bar{\rho} \bar{u}_j Y_J' = -\frac{\mu_t}{\sigma_m} \frac{\partial \bar{Y}_J}{\partial x_i}, \quad (7)$$

where \bar{w}_J represents the reaction rate of chemical species J ; D_J , the effective diffusion coefficient of chemical species J ; and σ_m , the effective Schmidt number. The eddy dissipation concept (Magnussen, 1989; Gran et al., 1996; Magnussen, 2005) was used to model \bar{w}_J , and D_J was estimated using Smooke's model. To obtain \bar{w}_J in Equation (5), calculations of the chemical reaction and chemical equilibrium were performed; the mass fraction equation can be computed by assignment of a value to \bar{w}_J .

To perform the chemical equilibrium calculation, we needed the mass fraction of elemental species I and enthalpy in group b. We therefore added the mass fraction equation of elemental species J in group a:

$$\frac{\partial(\bar{\rho} \bar{u}_j \tilde{Z}_I)}{\partial x_j} = \frac{\partial}{\partial x_j} \left(\frac{\mu_t}{\sigma_m} \frac{\partial \tilde{Z}_I}{\partial x_i} \right) + \bar{w}_I, \quad (8)$$

where \tilde{Z}_I represents the mass fraction of elemental species I . \bar{w}_I is the source term that exchanges the mass in group b for that in group a:

$$\bar{w}_I = -M_I \left[\sum_{J_b} \frac{a_{I J_b}}{M_{J_b}} \bar{w}_{J_b} \right], \quad (9)$$

where M_I represents the atomic weight of species I , J_b the chemical species J in group b, $a_{I J_b}$ the number of elemental species I included in chemical species J_b , \bar{w}_{J_b} the reaction rate of chemical species J_b , and M_{J_b} the mole weight of chemical species J_b .

The enthalpy in group a was approximated as

$$\tilde{h}_a = \tilde{h} \sum_I \tilde{Z}_I, \quad (10)$$

where \tilde{h}_a represents the enthalpy in group a.

The source term in Equation (5) for the equilibrium species J_a was modeled as

$$\bar{w}_{J_a} = \frac{\bar{\rho}(\xi^*)^2}{\tau^*} (Y_{J_{\text{aeq}}} - Y_{J_a}^0), \quad (11)$$

where J_a is the chemical species J in group a, \bar{w}_{J_a} is the reaction rate of chemical species J_a , ξ^* is the fine scale, τ^* is the time scale of the reaction, and $Y_{J_{\text{aeq}}}$ is the mass fraction of chemical species J_a obtained from the lookup table, which was prepared before the combustion simulation. The lookup table stored the results of the chemical equilibrium calculation when \tilde{Z}_I and \tilde{h}_a were varied, and these values were also used to extract the equilibrium values from the lookup table.

$Y_{J_a}^0$ in Equation (11) was estimated as

$$\tilde{Y}_J = \xi^3 Y_J^* + (1.0 - \xi^3) Y_J^0 \quad (12)$$

where J indicates J_a .

ξ^* was modeled as

$$\xi^* = C_\xi \left(\frac{\mu \epsilon}{\rho k^2} \right)^{1/4}, \quad (13)$$

where $C_\xi = 2.1377$. μ represents the viscosity, ϵ the turbulent dissipation rate, and k the turbulent kinetic energy. τ^* was given as

$$\tau^* = C_\tau \left(\frac{\mu}{\rho \epsilon} \right)^{1/2}, \quad (14)$$

where $C_\tau = 0.408$.

The proposed technique was needed to compute \bar{w}_J in Equation (5), and different expressions were used to compute \bar{w}_J , depending on the group that the species belonged to. The reaction rate of chemical species in group b can be obtained with the EDC model. The source term in Equation (5) for these modified chemical species was computed as

$$\bar{w}_{J_b} = \begin{cases} \frac{\bar{\rho}(\xi^*)^2}{\tau^*} \left[\max(Y_{J_b}^* - Y_{J_b}^0, Y_{J_{beq}}) \right], & (Y_{J_b}^* - Y_{J_b}^0 \geq 0.0) \\ \frac{\bar{\rho}(\xi^*)^2}{\tau^*} \left[Y_{J_b}^* + Y_{J_{beq}} - Y_{J_b}^0 \right], & (Y_{J_b}^* - Y_{J_b}^0 < 0.0), \end{cases} \quad (15)$$

where \bar{w}_{J_b} represents the reaction rate in group b; $Y_{J_b}^*$, the chemical species J_b in the fine structures; $Y_{J_b}^0$, the chemical species J_b in the surrounding gas; and $Y_{J_{beq}}$, the mass fraction of the chemical species J_b obtained from the lookup table.

Since the lookup table included the mass fraction of the chemical species in group b, the production resulting from using the values take from the table must be taken into account. When $Y_{J_{beq}} = 0$, these equations were similar to what was seen in the EDC model.

$Y_{J_b}^*$ was obtained by calculating the detailed chemical mechanism. The EDC model assumes that combustion occurs in a constant pressure reactor. The following equation was integrated under the initial conditions obtained from the current mass fractions, density, and temperature in each computational cell:

$$\frac{dY_{J_b}^*}{dt} = \frac{w_{J_b}^*}{\rho}, \quad (16)$$

where $w_{J_b}^*$ represents the reaction rate of chemical species J_b in the fine structures, and $w_{J_b}^*$ can be obtained by computing detailed chemical mechanisms based on the Arrhenius equation. Equation (16) is a stiff ODE, and therefore, the robust computational code CVODE (Hindmarsh et al., 2009) was used to acquire the steady solution.

Although the reaction rates of fuels can be computed by Equation (15), the flame sometimes blows out. That is why we used the following equations to estimate the reaction rates of the

fuels in this study. If we assumed that the reaction rates of fuels were determined with an overall reaction such as $1(\text{kg})\text{fuel} + st(\text{kg})\text{oxidant} \Rightarrow (1 + st)(\text{kg})\text{products}$, Y_{fu}^* can be modeled based on the expression of the overall reaction rate as

$$Y_{\text{fu}}^* = \tilde{Y}_{\text{fu}} - \Delta Y_{\text{fu}} \quad (17)$$

$$\Delta Y_{\text{fu}} = A_{\text{fu}} \min(\tilde{Y}_{\text{fu}}, \frac{\tilde{Y}_{\text{ox}}}{st}). \quad (18)$$

The subscripts fu and ox represent the fuels and oxidants, respectively. st represents the theoretical weight of the oxidants, and A_{fu} is the mixing parameter of fuels set to 1.0 (equal to the value in the original EDC model).

2.2 Chemical equilibrium method and reading lookup table

When preparing the lookup table, we developed a different program from the UDF based on the chemical equilibrium method, by minimization of the Gibbs free energy. This method assumes that the equilibrium composition of a chemical system can be determined by minimizing the Gibbs free energy. Under the minimum conditions of Gibbs free energy, the following simultaneous nonlinear equations are obtained:

$$X_0 \sum_J a_{IJ} \exp \left[-\sum_K \frac{\alpha_{KJ} \beta_I}{RT} - \left(\frac{G_J}{RT} + \ln p \right) \right] = d_I, \quad (19)$$

$$\sum_J \exp \left[-\sum_K \frac{a_{KJ} \beta_I}{RT} - \left(\frac{G_J}{RT} + \ln p \right) \right] = 1, \quad (20)$$

where X_0 represents the molar concentration, a_{IJ} the number of elemental species I included in chemical species J , a_{KJ} the number of elemental species K included in chemical species J , β_I the negative chemical potential only when the system is in chemical equilibrium, R the gas constant, G_J the Gibbs free energy of chemical species J , p the pressure, and d_I the concentration of elemental species I .

When d_I and T were assigned as constraints, β_I and X_0 were obtained by solving the simultaneous nonlinear Equations (19) and (20). In this study, the computational code KINSOL (Collier et al., 2009) was used to solve these equations, and the concentration of each species at given T was obtainable by

$$X_J = X_0 \exp \left[-\sum_I \frac{a_{IJ} \beta_I}{RT} - \left(\frac{G_J}{RT} + \ln p \right) \right], \quad (21)$$

where X_J represents the molar concentration of chemical species J , and thermodynamic properties such as G_J and h_J were computed as functions of the temperature (Mcbride et al., 2002). According to this procedure, the concentration of the chemical species J at given T was acquired. However, the additional constraint that the enthalpy was constant before and after combustion was required so that the chemical equilibrium composition in combustion could be found. Thus, T must be found under the constraint

$$\sum_J h_J X_J = h_r, \quad (22)$$

where h_J represents the enthalpy of chemical species J , and h_r represents the enthalpy of reactants (before combustion). When we found T using Equation (22), the obtained results were the concentrations after combustion.

The equilibrium value Y_{jeq} was obtained from reading the lookup table using \tilde{Z}_H , \tilde{Z}_O , \tilde{Z}_N , and \tilde{h}_a . Then, a simple multi-linear interpolation was used to extract information from the lookup table.

2.3 NO prediction using the extended Zeldovich mechanism

FLUENT version 13.0 has a function for computing the amount of the thermal NO_x using the extended Zeldovich mechanism as



Because we used the basic function in FLUENT, we describe only the outline of the method to estimate the amount of NO here (Warnatz et al., 2006; Ansys, Inc, 2010). The reaction rate constants used in FLUENT were selected based on the evaluation of Hanson and Salimian (Hanson et al., 1984). To calculate the NO concentration in turbulence combustion, the time-averaged formation rate of NO must be computed at each computational cell. Methods of modeling the mean turbulent reaction rate of NO can be based on the presumed PDF model (Peng et al., 2009). FLUENT adopted the PDF model to predict the mass fraction of NO, and the reaction rate of the NO was expressed as

$$\bar{w}_{\text{NO}} = \int w_{\text{NO}} P(\phi_1) d\phi_1, \quad (26)$$

alternately, the PDF was related to two variables:

$$\bar{w}_{\text{NO}} = \iint w_{\text{NO}} P(\phi_1, \phi_2) d\phi_1 d\phi_2, \quad (27)$$

where w_{NO} represents the instantaneous reaction rate of NO, and $P(\phi_1)$ and $P(\phi_1, \phi_2)$ represent the PDFs of the variables. In this study, the mass fraction of O₂ and T were chosen as ϕ_1 and ϕ_2 , respectively. As shown in Equations (23)–(25), the extended Zeldovich mechanism includes intermediate species such as O, N, H, and OH. The mass fractions of these species can be obtained by using the proposed technique. Thus, only the reaction rate of NO was computed by the presumed PDF model, while those of the intermediate species were taken into account using the proposed technique.

2.4 Numerical conditions

A H₂-air non-premixed flame was simulated to verify the accuracy of the proposed technique. The numerical conditions were based on Barrow's experimental data (Barrow et al., 2003a). Figure 2 shows the sketch of the computational domain used in this study. \tilde{Y}_{H_2}

in the fuel was 1.0; \tilde{Y}_{O_2} , \tilde{Y}_{N_2} , and \tilde{Y}_{H_2O} in air were 0.231, 0.761, and 0.00720, respectively. Since \tilde{Y}_{O_2} , \tilde{Y}_{N_2} , and \tilde{Y}_{H_2O} belong to group b, $\tilde{Z}_H = 0.008$, $\tilde{Z}_O = 0.23772$, and $\tilde{Z}_N = 0.76148$ were assigned in air. The fuel was injected from a nozzle with a radius of 3.75 mm. The velocity of the fuel and ambient air were 296 and 1.0 m/s, respectively. The temperature of the fuel and ambient air was set to 300 K. In all cases, a second-order upwind difference was applied to the governing equations of the momentum, mass fraction of chemical species J and elemental species I , and the energy equation. The two-dimensional axisymmetric solver in FLUENT was chosen using the axis indicated by the centerline in Fig. 2. The Reynolds stress model was chosen as the turbulence model; $C_{2\epsilon} = 1.80$, $\sigma_{h_i} = 0.7$, and $\sigma_m = 0.7$ of the model parameters were used to adjust the axial profile of the mixture fraction of the experimental data. In this study, the detailed chemical mechanism of H_2 suggested by Barrow (Barrow et al., 1999) was used to compute the reaction rates of the modified species in the proposed technique, and those of the chemical species taken into account by the EDC model.

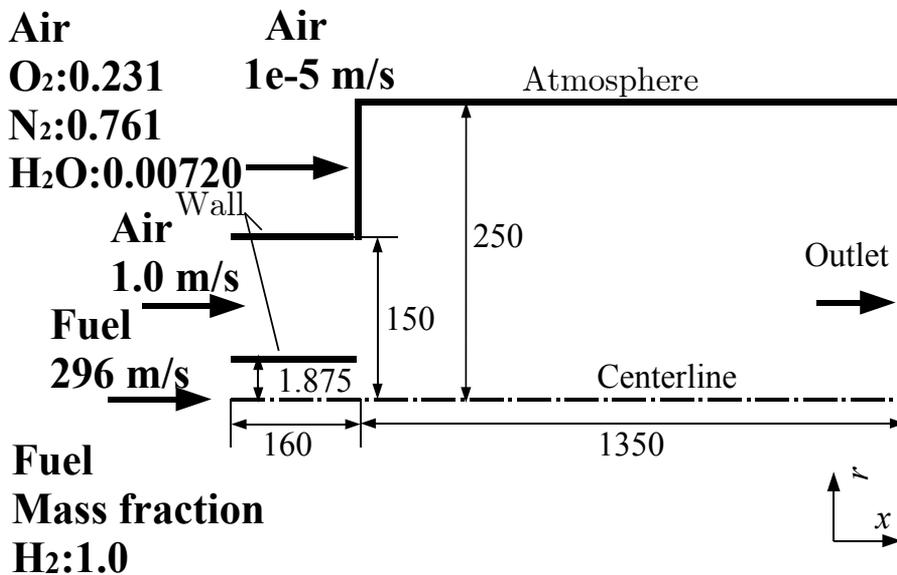


Fig. 2. Sketch of computational domain

3. Results

3.1 Prediction of mixture fraction, temperature, and mole fraction

First, to validate the prediction accuracy of the proposed technique, we compared the obtained results with those of the EDC model and the experiment. The lookup table was prepared before the combustion simulation, and it included the following species: H_2 , O_2 , H_2O , O , H , OH , HO_2 , H_2O_2 , NO , HNO , N_2O , N , NO_2 , and N_2 . In Figs. 3–18, x represents the axial distance from the exit of the nozzle; r , the radial distance from the centerline of the nozzle; L , the approximate visible flame length, which is equal to 675 mm in this flame; D , the diameter of the nozzle, which is equal to 3.75 mm in this flame; T , the temperature; EXP, the experimental data; EDC, the computational data predicted by the EDC model with

Barrow's detailed mechanism (Barrow et al., 1999); and $A_{H_2} = 100$, the computational data obtained by the proposed technique with $A_{H_2} = 100$. Although $A_{H_2} = 100$ was seemingly large, choosing this value means that H₂ is mixed with O₂ and is immediately consumed. Fig. 3 shows the axial profile of the mixture fraction. Because the parameters $C_{2\epsilon} = 1.80$, $\sigma_h = 0.7$, and $\sigma_\epsilon = 0.7$ of the Reynolds stress model were chosen to match the axial mixture fraction obtained by the experiment, the computational results were close to the experimental data at $0.2 < x/L$. However, as shown in Fig. 3, those at $x/L < 0.2$ were slightly smaller than the experimental data. The mixture fraction was related to mass diffusion, which in turn depended on turbulence models. It can be seen from this result that mass diffusion was not similar to the experimental data at $x/L < 0.2$ (near the nozzle), and the turbulence model would have to be modified to obtain the closer result. However, the objective in this section is to confirm the validity of the proposed technique. To that end, we performed the simulation using the EDC model for reference. The EDC model has been widely applied to the simulation of combustion equipment (Stefanidis et al., 2006; Minotti et al., 2010), and the results obtained using the EDC model have been considered to be reliable for the engineered approach. The prediction of the proposed technique must approach that of the EDC model, because the expression of the reaction rate in the proposed technique was based on the EDC model.

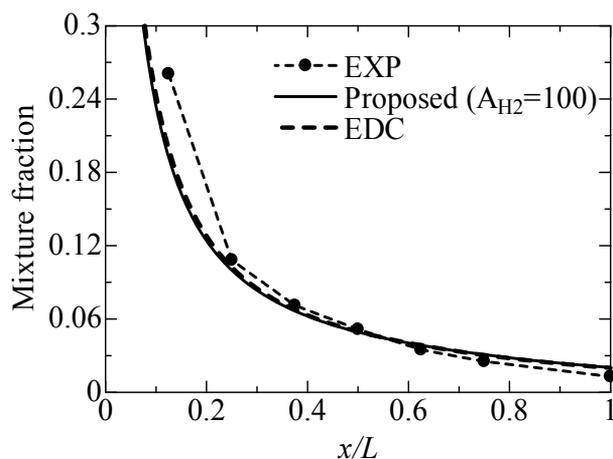


Fig. 3. Axial mixture fraction

The axial temperature T and mole fraction of H₂O are shown in Fig. 4. The mixture fractions of the computational results are in good agreement with each other, although those at $0.6 < x/L$ were slightly larger than the experimental data.

Fig. 5 shows the axial mole fraction of H₂ and O₂. Both the computational results of H₂ and O₂ were lower than those of the experimental data. It was found that the simulations slightly overestimated the consumption of H₂ and O₂ as compared with those of the experiment. The mole fraction of H₂ at $0.3 < x/L < 0.6$ obtained by the proposed technique was slightly larger than that of the EDC model. The reaction rate of H₂ was determined by the overall reaction, choosing $A_{H_2} = 100$ in the case of the proposed technique; the rate was computed

from the Barrow's detailed mechanism in the case of the EDC model. The overall reaction with choosing $A_{H_2} = 100$ in the proposed technique caused the difference of H_2 at $0.3 < x/L < 0.6$.

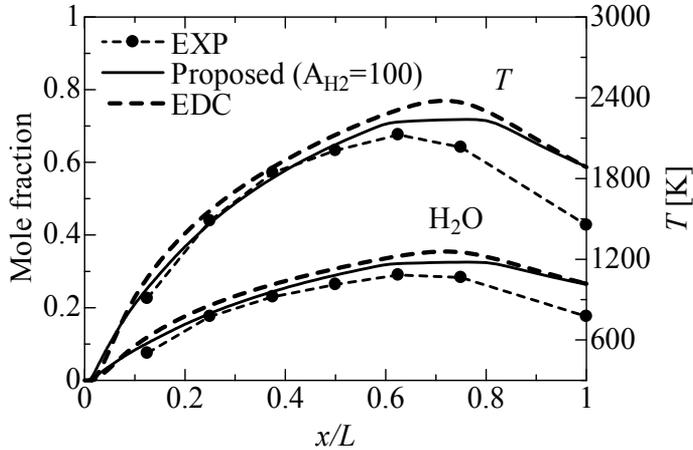


Fig. 4. Axial temperature T and mole fraction of H_2O

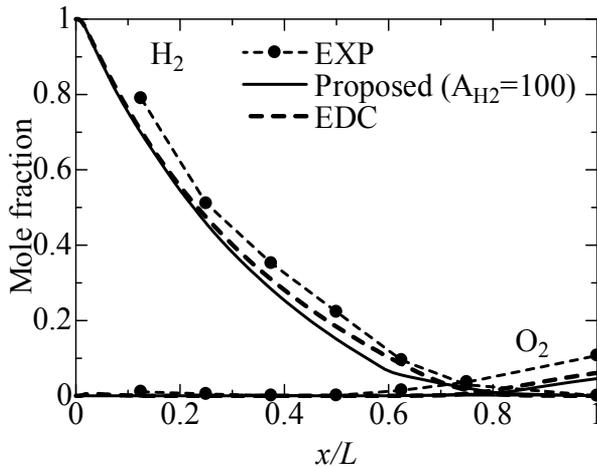


Fig. 5. Radial temperature and mole fraction of H_2 and O_2

The mole fraction of OH is plotted in Fig. 6. Generally, reaction equations including OH are needed to obtain the mole fraction of OH , but it was not required in the proposed technique. (However, the reaction equation of OH can be taken into account in the proposed technique using Equations (15)–(16) to improve the prediction accuracy.)

As shown in Fig. 6, the result of OH was comparatively small as compared with major species such as H_2 , O_2 , and H_2O ; the amount of the radical species, such as H , O , and OH , were generally very subtle. For this reason, it was important to estimate the tendency of yield and consumption of the mole fraction of OH . In Fig. 6, the coordinates of the peaks of the mole fraction of OH obtained by the simulations were similar to the experimental data,

and we found that the mole fraction of OH obtained using the proposed technique was qualitatively similar to that obtained by the EDC model and experimental data.

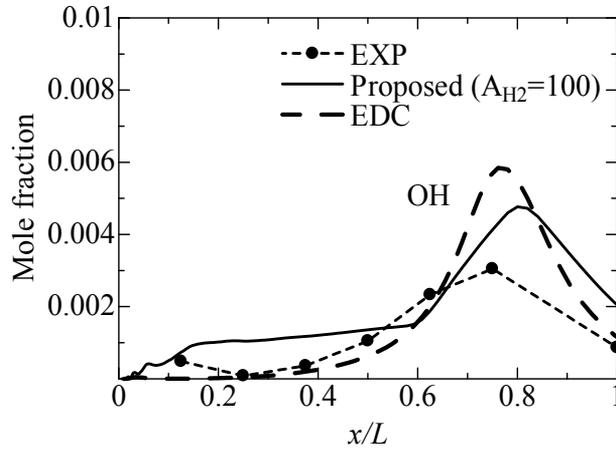


Fig. 6. Radial temperature and mole fraction of OH

Fig. 7 shows the radial mixture fraction at $x/L = 1/8$. Although the results obtained by the EDC model and the proposed technique were in good agreement, they were not close to the experimental data. As mentioned above, the difference of the mixture fraction between the simulations and experiment was related to the turbulence model, and not to either the EDC model or the proposed technique.

As we can observe from Fig. 8, the peaks of the temperature and the mole fraction of H₂O predicted with the use of the EDC and proposed models were at approximately $r/D = 4$, while that of the experimental data was at approximately $r/D = 2.9$. The difference of coordinates of the peaks between the simulation data and the experimental data derived mainly from turbulence model accuracy.

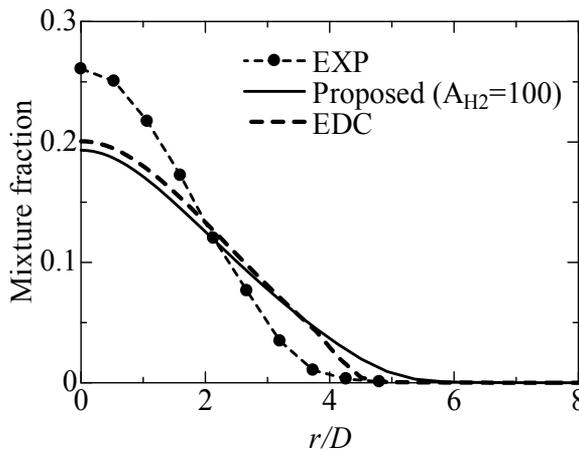


Fig. 7. Radial mixture fraction at $x/L = 1/8$

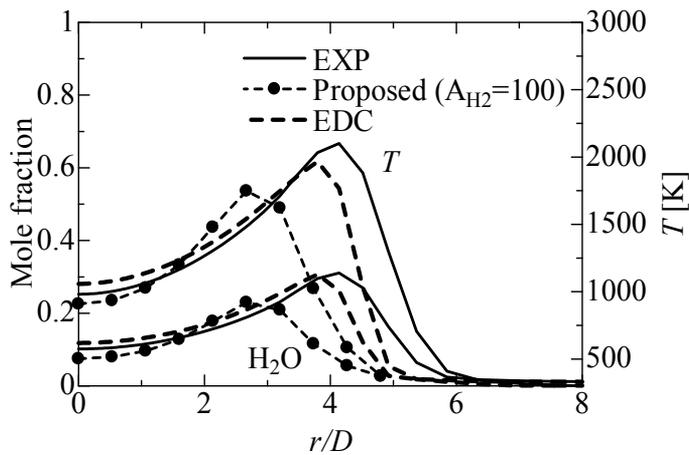


Fig. 8. Radial temperature and mole fraction of H_2O at $x/L = 1/8$

Fig. 9 shows the radial mole fraction of H_2 and O_2 . The mole fraction of O_2 obtained by the proposed technique was lower than that predicted by the EDC model. In the proposed technique, the mole fraction of O_2 was obtained from the lookup table, which was prepared using the chemical equilibrium method. For this reason, consumption of O_2 was sometimes overestimated at points where the reaction was in the state of strong nonequilibrium.

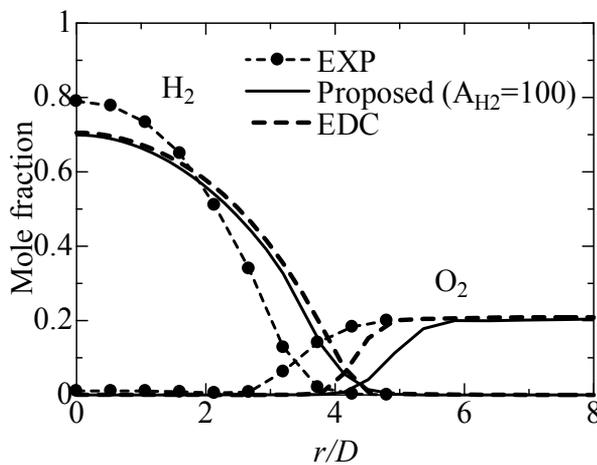


Fig. 9. Radial mole fraction of H_2 and O_2 at $x/L = 1/8$

As shown in Fig. 10, the coordinate values of the peak of the mole fraction of OH obtained by the simulations and experimental data are located at approximately $r/D = 4.1$ and $r/D = 2.7$, respectively. These values were slightly different from what was seen in the experiment. The coordinate of the peak mole fraction of OH was close to that of the mole fraction of H_2O and T . It can be seen from this tendency that OH was the indicative species of extinction and ignition phenomena.

The mixture fraction and temperature as well as mole fraction of H₂O are shown in Figs. 11 and 12, respectively. Both the computational data of the mixture fraction, temperature, and mole fraction of H₂O are in good agreement.

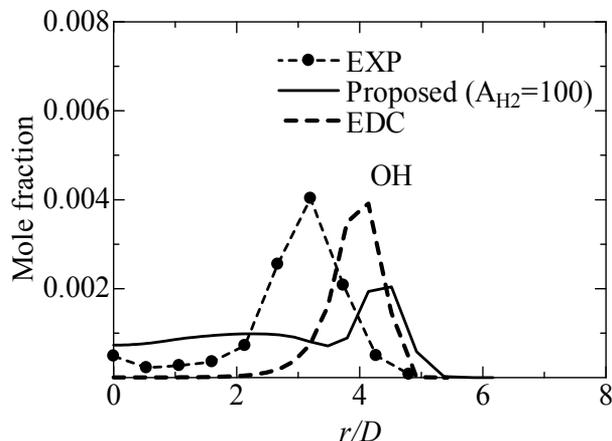


Fig. 10. Radial mole fraction of OH at $x/L = 1/8$

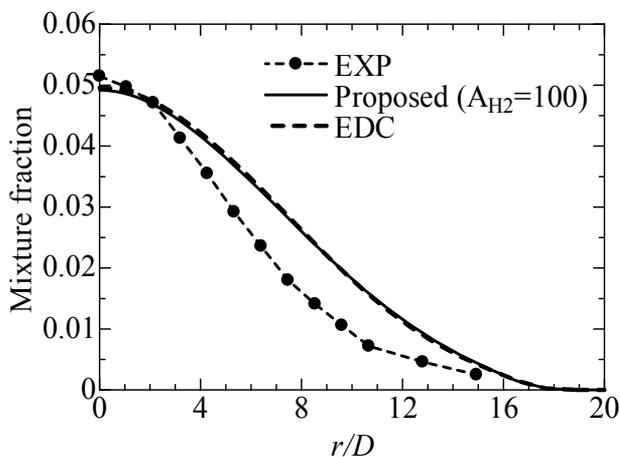


Fig. 11. Radial mixture fraction at $x/L = 1/2$

Fig. 13 shows the radial mole fraction of H₂ and O₂ at $x/L = 1/2$. Although the mole fractions of O₂ obtained by the proposed model and EDC model were almost similar, the result of H₂ computed using the proposed technique was not similar to that of the EDC model. This was because the proposed technique used $A_{H_2} = 100$ to estimate the reaction rate of H₂, while the reaction rate was computed using the detailed chemical mechanism in the EDC model. However, the difference of the results between the proposed technique and

EDC model were small as compared to the maximum mole fraction of H_2 . Basically, the maximum mole fraction of H_2 was equal to 1.0 taken at $x/L=0$; the difference at $r/D=0.0$ between both computational results was about 0.03.

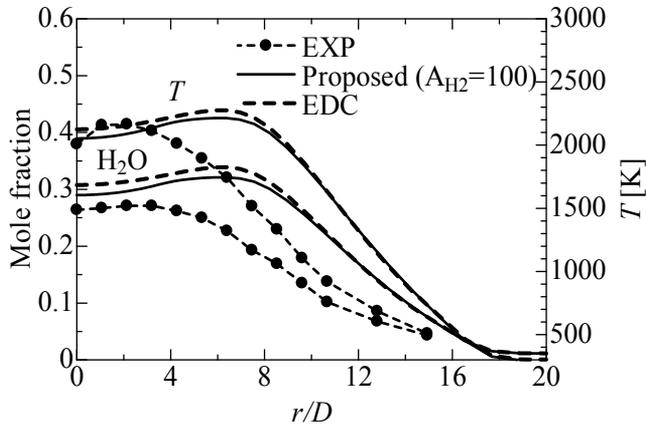


Fig. 12. Radial temperature and mole fraction of H_2O at $x/L = 1/2$

As shown in Fig. 14, the peaks of the mole fractions of OH obtained by the simulations and that of the experiment were located at approximately $r/D=7.6$ and $r/D=6.4$, respectively. The coordinates of the peaks obtained by the simulations were similar to each other, but they were slightly different from that of the experiment.

As shown in Figs. 15, 16, and 17, computational results for EDC and proposed models were in good agreement, but were either slightly greater or less than the experimental data, which showed a similar tendency to that of the other area, such as the radical direction at $x/L=1/8$ and $x/L=1/2$.

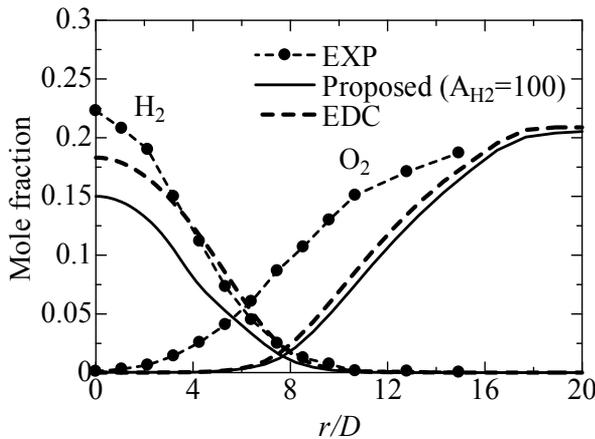


Fig. 13. Radial mole fraction of H_2 and O_2 at $x/L = 1/2$

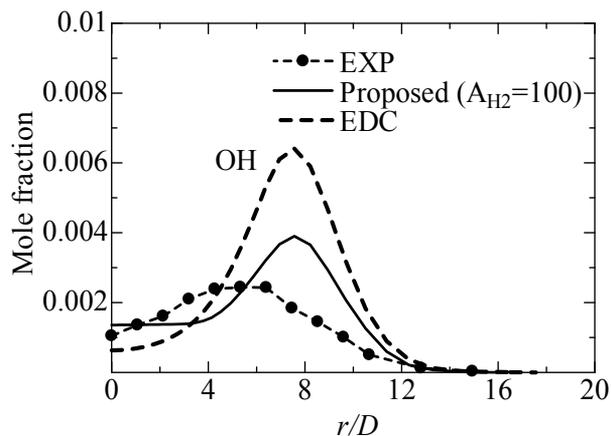


Fig. 14. Radial mole fraction of OH at $x/L = 1/2$

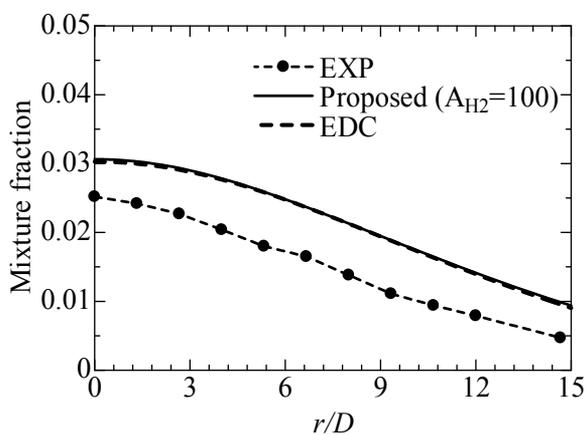


Fig. 15. Radial mixture fraction at $x/L = 3/4$

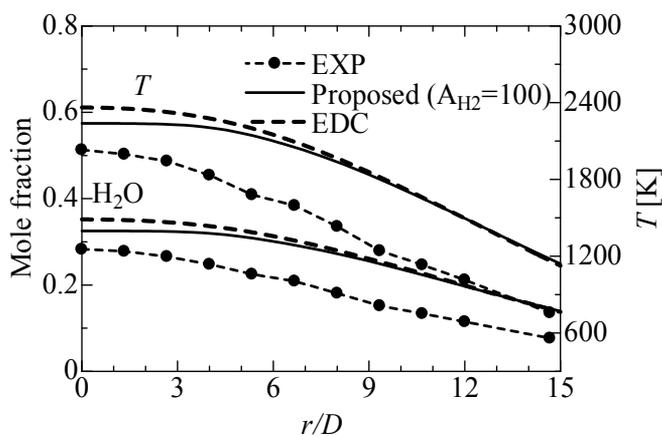


Fig. 16. Radial temperature and mole fraction of H₂O at $x/L = 3/4$

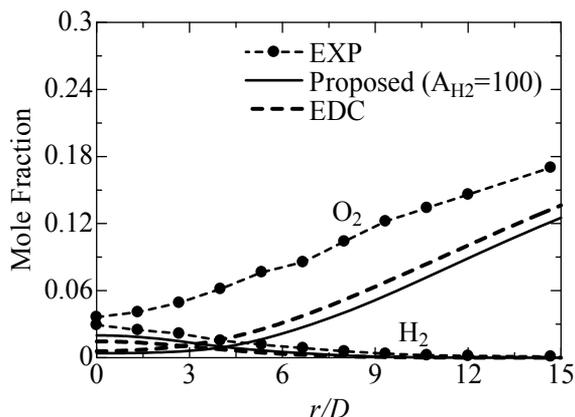


Fig. 17. Radial mole fraction of H_2 and O_2 at $x/L = 3/4$

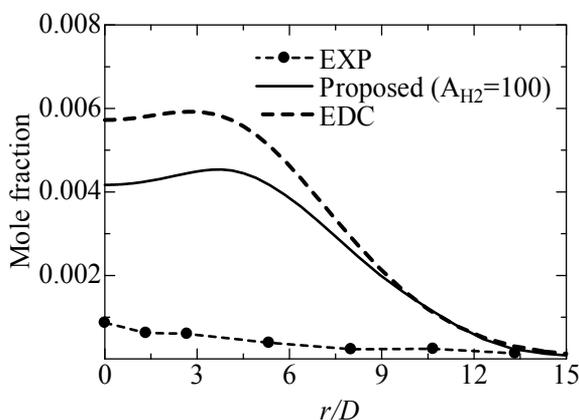


Fig. 18. Radial mole fraction of OH at $x/L = 3/4$

The radial mole fractions of OH at $x/L = 3/4$ are depicted in Fig. 18. The computational results were overestimated as compared with the experimental data, and the mole fraction of OH at $9 < r/D$ obtained using the proposed technique was similar to that obtained using the EDC model.

In this portion of the experiment, we found that the results of the mixture fraction, major species, and temperature obtained by the proposed technique were roughly similar to those predicted by the EDC model. The mole fraction of OH was not quantitatively similar to that of the EDC model; however, the coordinate values of the peaks of mole fraction of OH predicted by the proposed technique and EDC model were in good agreement with one another. Thus, the tendency of yield and consumption of OH was predicted with the proposed technique, which did not use the chemical equation of OH.

The computations show slight differences in the experimental data, even when the EDC model was applied to the combustion simulation. Moreover, we found that it is important to accurately predict the non-premixed flame using not only the combustion model but also the turbulence model. For this reason, further improvement of the combustion and turbulence models is a significant task for combustion simulation.

3.2 Prediction of NO

We show the application of the proposed technique to the simulation of NO in this section. In Figs. 19–30, “Proposed” indicates computational data obtained by the proposed technique with $A_{\text{H}_2} = 100$; “Proposed (H, O, OH)” indicates the computational data obtained by the proposed technique with $A_{\text{H}_2} = 100$ and reaction rates of H, O, and OH modified using Equations (15)–(16). Other symbols in Figs. 19–30 are as they were in the previous section. Two lookup tables were prepared before the combustion simulation. One table includes the species H₂, O₂, H₂O, O, H, OH, HO₂, H₂O₂, NO, HNO, N₂O, N, NO₂, and N₂; it was used for the case of the proposed technique with $A_{\text{H}_2} = 100$. The other table includes the species H₂, O₂, H₂O, HO₂, H₂O₂, NO, HNO, N₂O, N, NO₂, and N₂; it was prepared for the case of the proposed technique $A_{\text{H}_2} = 100$ and reaction rates of H, O, and OH modified using Equations (15)–(16).

As can be seen in Fig. 19, the peaks of the mole fraction of NO obtained by the simulation and that of the experiment were located at approximately $x/L = 0.82$. Those of the EDC model and proposed technique with modified reaction rates of H, O, and OH were bigger than the experimental data, while the peak obtained using the proposed technique without modified reaction rates was similar to that of the experiment. On the other hand, as shown in Fig. 20, the mole fraction of NO obtained by the proposed technique without the modified reaction rates was not similar to the experimental data. Thus, we did not determine which cases were most accurate in these simulations. As can be observed from Figs. 19 and 20, the amount of NO was very small as compared to the major species such as H₂, O₂, and H₂O. Therefore, the most important aim of the NO prediction was to simulate the tendency of yield and consumption of NO.

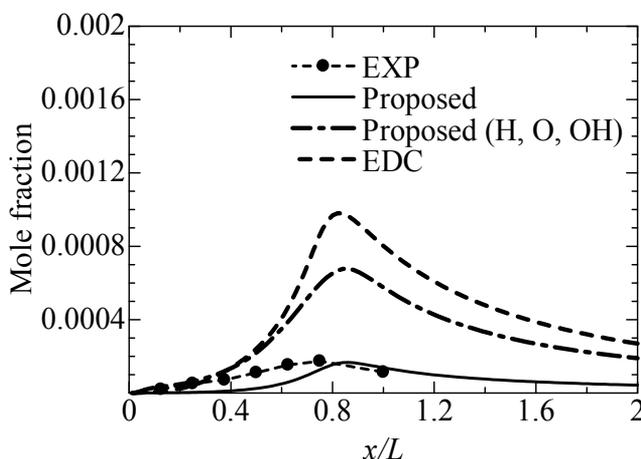


Fig. 19. Axial mole fraction of NO predicted with the extended Zeldovich mechanism

In Figs. 20 and 21, the coordinates of the peaks obtained by both the simulations were in good agreement. Moreover, as shown in Fig. 22, the peaks of the mole fractions were located at roughly $r/D = 6$, and the mole fraction of NO decreased at $6 < r/D < 15$.

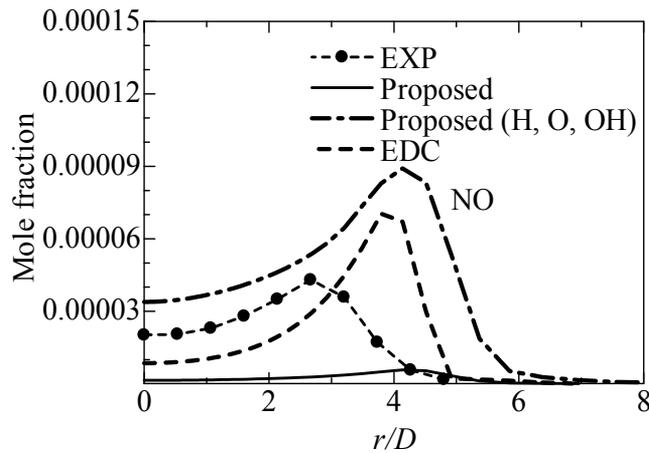


Fig. 20. Radial mole fraction of NO at $x/L=1/8$ predicted with the extended Zeldovich mechanism

It can be seen from Figs. 19–22 that the amounts of the NO prediction were not quantitatively correct; however, the peaks of the mole fractions obtained by the simulations were similar. Thus, we found that the tendency of yield and consumption of NO was simulated by all the cases in this part of the experiment.

The prediction accuracy of O is significant in the computation of the mass fraction of NO because the first reaction of the extended Zeldovich mechanism, expressed as $O + N_2 \rightarrow N + NO$, is the rate-limiting step (Warnatz et al., 2006). Thus, the amount of NO increases with an increase in O. As shown in Fig. 23, the axial mole fraction of O obtained using the proposed technique with modified reaction rates of H, O, and OH was larger than that of the EDC model.

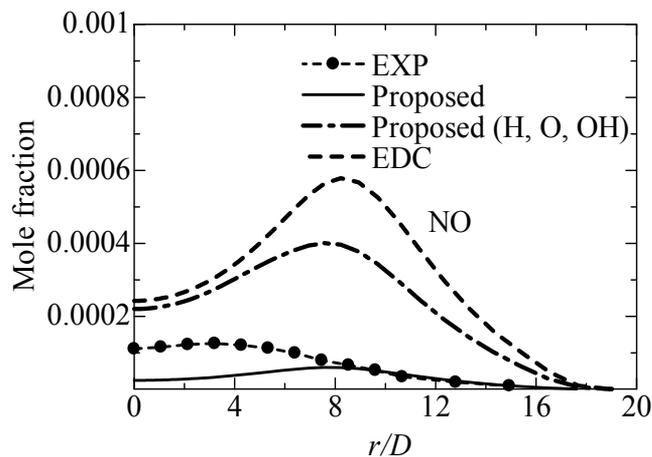


Fig. 21. Radial mole fraction of NO at $x/L = 1/2$

However, in Fig. 19, the axial mole fraction of NO obtained by the proposed technique with modified reaction rates of H, O, and OH was smaller than that of the EDC model. This was because axial T computed by the EDC model was overestimated as compared to that

computed by the proposed technique with modified reaction rates of H, O, and OH, as can be seen in Fig. 24. Thus, the first reaction of the extended Zeldovich mechanism became more reactive because of the increasing T of the EDC model.

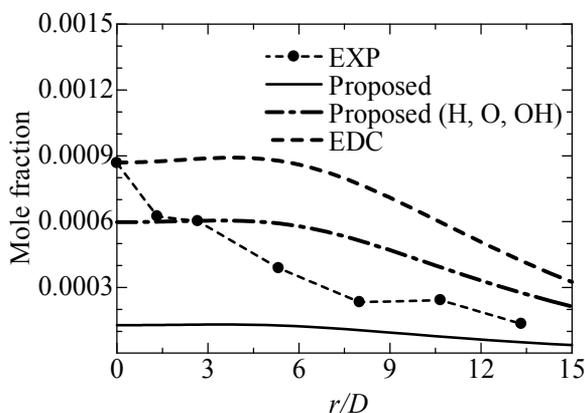


Fig. 22. Radial mole fraction of NO at $x/L = 3/4$

On the other hand, as can be seen in Figs. 25 and 26, either the mole fraction of O or T obtained using the proposed technique with modified reaction rates of H, O, and OH were larger than those of the EDC model. Therefore, the radial mole fraction of NO at $x/L = 1/8$ was also larger than that computed by the EDC model.

We found that the amount of NO was determined by the amount of O and T , and those of prediction accuracy were so important of obtaining the accurate mole fraction of NO. In the proposed technique, the reaction rates of the intermediate species such as H, O, and OH were computed, and the mole fractions can be obtained taking into account the chemical equations. Generally, when the reaction rates were modified using the chemical equations, the prediction accuracy of the chemical species was also increased (Fukumoto et al., 2010). However, even without the modified reaction rates, the proposed technique well simulated the mole fraction of O, and the improvement of the modified reaction rates was relatively small. Therefore, the computing the reaction rates of H, O, and OH can be omitted in this flame.

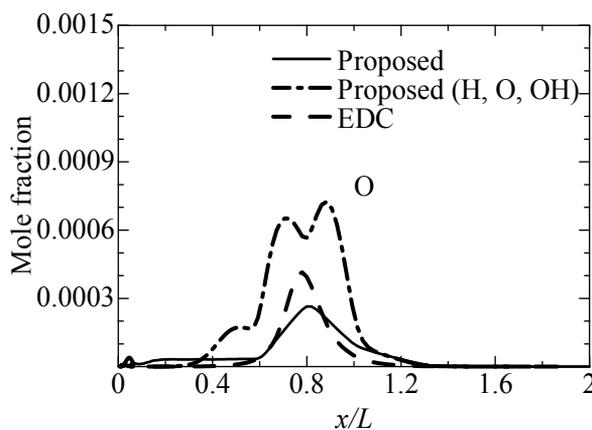


Fig. 23. Axial mole fraction of O

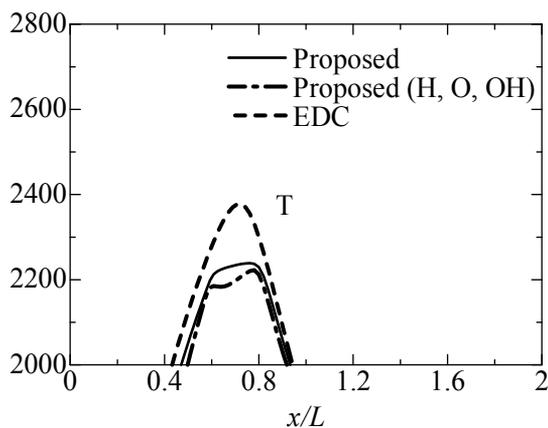


Fig. 24. Axial T ranging from 2000 K to 2800 K

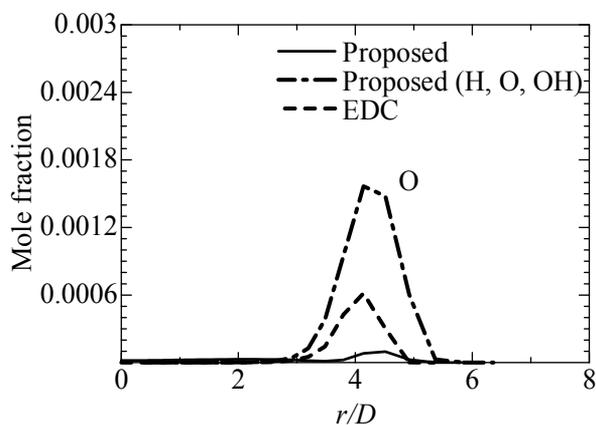


Fig. 25. Radial mole fraction of O at $x/L = 1/8$

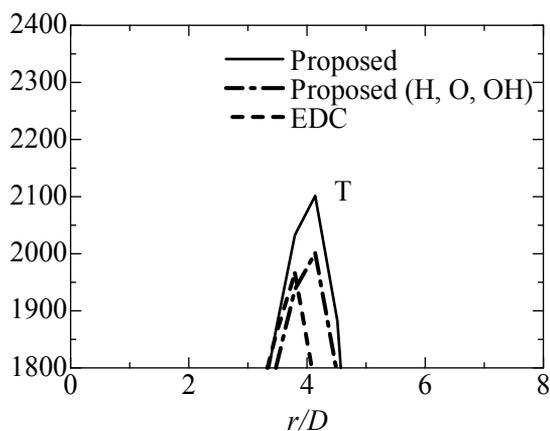


Fig. 26. Radial mole fraction of T at $x/L = 1/8$ ranging from 1800 K to 2400 K

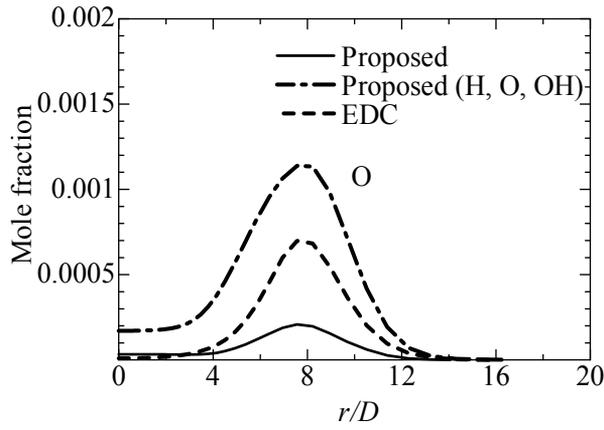


Fig. 27. Radial mole fraction of O at $x/L=1/2$

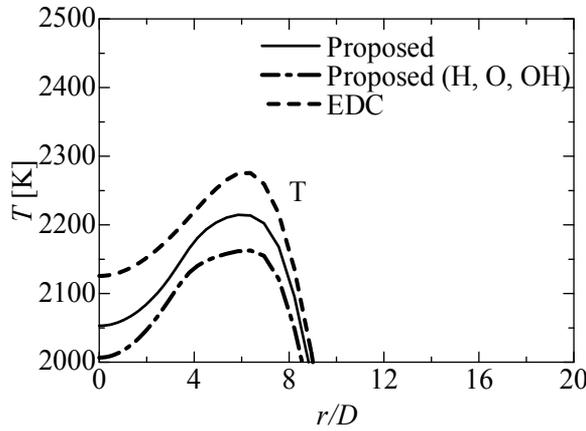


Fig. 28. Radial mole fraction of T at $x/L=1/2$ ranging from 2000 K to 2500 K

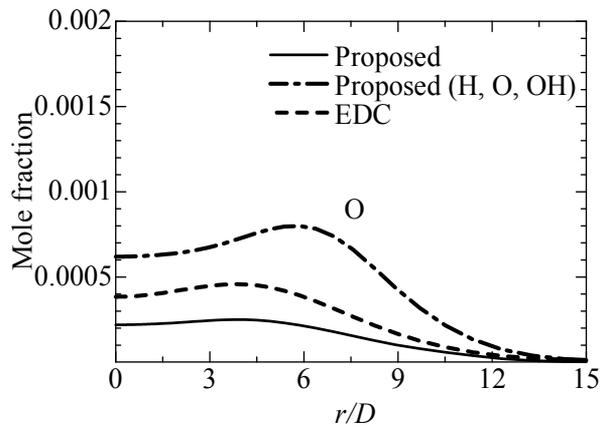


Fig. 29. Radial mole fraction of O at $x/L=3/4$

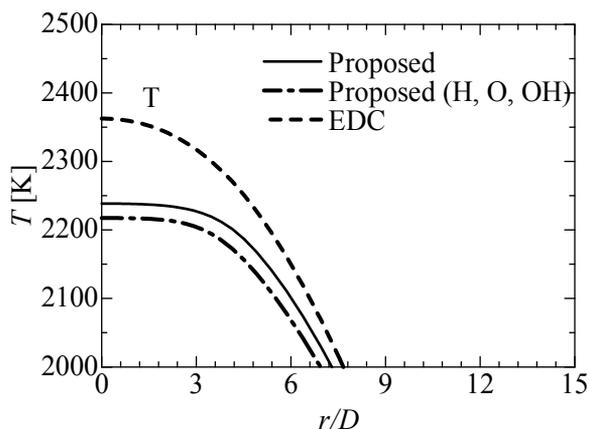


Fig. 30. Axial T at $x/L=3/4$ ranging from 2000 K to 2500 K

3.3 Influence of number of modified species and computational time

Due to the time required to compute reaction rates of these species, the computational time is dependent on the number of the chemical species in group a. Our proposed technique enables the shortening of these computational times. The purpose of this section is to discuss the prediction accuracy for the minor species and computational time required by the proposed technique. The numerical condition of the EDC model conformed to that presented in the previous sections, and the other two cases of the proposed technique were performed in this part of the experiment. In the first case of the proposed technique, the reaction rate of H_2 was computed by the estimated by Equations (17)–(18). In the second case, those of H, O, and OH were modified, and the reaction rate of H_2 was estimated.

Fig. 31 shows a comparison of the computational time in the case of the EDC model and two cases of the proposed technique. Each computational time per iteration was measured after the solution was converged and was measured during the 100 iterations of performing the reaction calculation and reading the lookup table. The CPU used in these simulations was the Intel Core i7-980X Processor Extreme Edition with 12M Cache, and 3.33 GHz. GFORTRAN 4.4.4 was chosen as a compiler, and CVODE 2.6 (Hindmarsh et al., 2009) was adopted to solve the ODEs of the reaction calculation. As can be seen in Fig. 31, the computational times for the EDC model—the two proposed techniques with modified H, O, and OH (second case) and with no modified species (first case)—were 8.99 s, 2.83 s, and 0.04 s, respectively. We found that the lower the number of included modified species, the fewer were the demands on lower computational time. On the other hand, the prediction accuracy also depended on the number of modified species included. Therefore, when the proposed technique is used, the number of modified species should be determined according to the required prediction accuracy. In the proposed technique, the computational time required increased with an increase in the number of modified species. However, the proposed technique has an advantage: We only chose the modified species and then the reduced mechanism was automatically built from the detailed mechanism.

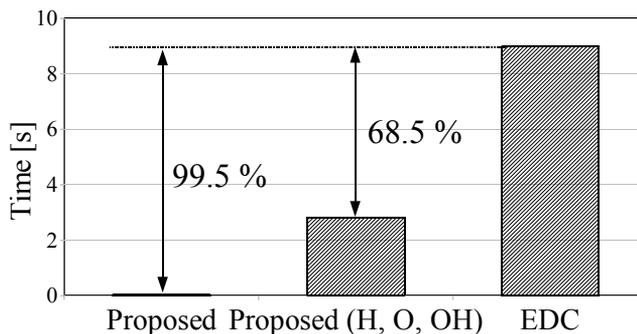


Fig. 31. Comparison of computational time

4. Conclusion

In this chapter, we presented a combustion simulation technique based on a chemical equilibrium method combined with the EDC model. The proposed technique makes it easy to construct a reduced mechanism according to the accuracy requirements of the chemical species, so that the computational time can be minimized.

We draw the following conclusions:

- The prediction accuracy of the proposed technique for the mixture fraction, major mole fraction, and T was similar to that of the EDC model.
- The coordinates of the peak OH mole fractions predicted by the proposed technique and EDC model were in good agreement. Thus, the tendency of yield and consumption of OH can be predicted using the proposed technique, which did not use the chemical equation of OH.
- The tendency of yield and consumption of NO was simulated by the proposed technique with the extended Zeldovich mechanism.
- A lower number modified species led to lower computational time. However, the prediction accuracy also relied on the number of modified species included. Modified species should be included according to the required prediction accuracy.

5. References

- Amir, M. & Sadegh, T. (2010). Effect of Hydrogen on Hydrogen-Methane Turbulent Non-premixed Flame under MILD Condition. *Hydrogen Energy*, Vol. 35, No. 20, (October 2010), pp. 11324–11331, ISSN: 0360-3199
- Ansys, Inc. (2011). Ansys FLUENT, (2010), Available from: <http://www.ansys.com/>
- Ansys, Inc. (2010). ANSYS FLUENT Theory Guide, Ansys, Inc. Retrieved from <http://www.ansys.com/>
- Barlow, R S. & Smith, N S A. (1999). Nitric Oxide Formation in Dilute Hydrogen Jet Flames: Isolation of the Effects of Radiation and Turbulence-Chemistry Submodels. *Combustion and Flame*, Vol. 117, No. 1, (April 1999), pp. 4–31, ISSN: 0010-2180
- Barlow, R S. (2003a). Sandia H₂/He Flame Data-Release 2.0, In: *International Workshop on Measurement and Computation of Turbulent Non-premixed Flames*, June 2011, Available from: <<http://www.sandia.gov/TNF/simplejet.html>>
- Barlow, R S. (2003b). Radiation Models, In: *International Workshop on Measurement and Computation of Turbulent Non-premixed Flames*, June 2011, Available from: <<http://www.sandia.gov/TNF/radiation.html>>

- Collier, A M. & Hindmarsh, A C. (2009). KINSOL version 2.6.0, In: *SUNDIALS web site*, 30 June 2011, Available from <https://computation.llnl.gov/casc/sundials/main.html/>
- Daniele, L M. (2009). Large Eddy Simulation of mixing and reaction in a Confined Impinging Jets Reactor. *Computers and Chemical Engineering*, Vol. 33, No. 2, (February 2009), pp. 393–534, ISSN: 0098-1354
- Fukumoto, K. & Ogami, Y. (2010). Turbulent Diffusion Combustion Model using Chemical Equilibrium Combined with the Eddy Dissipation Concept for Reducing Detailed Chemical Mechanisms: An Application of H₂-air Turbulent Diffusion Flame. *Heat Transfer-Asian Research*, Vol. 39, No. 5, (July 2010), pp. 292–313, ISSN: 1523-1496
- Fukumoto, K. & Ogami, Y. (2009). Simulation of CO-H₂-air Turbulent Diffusion Flame by the Combustion Model Combined Chemical Equilibrium Method with the Eddy Dissipation Concept Model. *Journal of High Temperature Society*, Vol. 39, No. 4, pp. 159-168, ISSN: 0387-1096
- Gran, I R. & Magnussen, B F. (1996). A Numerical Study of a Bluff-body Stabilized Diffusion Flame. Part 2. Influence of Combustion Modeling and Finite-Rate Chemistry. *Combustion Science Technology*, Vol. 119, (1996), pp. (191–217), ISSN: 0010-2202
- Hanson, R K. & Salimian S. (1984). Survey of Rate Constants in H/N/O Systems, In: *Combustion Chemistry*. Gardine, W C Jr, 36, Springer-Verlag, ISBN-10: 038790963X, ISBN-13: 978-0387909639, New York, USA
- Hindmarsh, A C. & Serban, R. (11 May 2009). CVODE version 2.6.0, In: *SUNDIALS web site*, 30 June 2011, Available from: <https://computation.llnl.gov/casc/sundials/main.html/>
- Magnussen, B F. (1989). Modeling of NO and Soot Formation by the Eddy Dissipation Concept, *Proceedings of 1st topic Oriented Technical Meeting*, Amsterdam, Netherlands, October 1989
- Magnussen, B F. (2005). The Eddy Dissipation Concept: A Bridge Between Science and Technology, *Proceedings of ECCOMAS Thematic Conference on Computational Combustion*, Lisbon Portugal, June 2005
- Mcbride, B J. & Zehe, M J. (2002). NASA Glenn Coefficients for Calculating Thermodynamic Properties of Individual Species. In: *NASA Technical Report TP-2002-211556*, 30 June 2011, Available from: <http://gltrs.grc.nasa.gov/>
- Minotti, A. & Sciubba, E. (2010). LES of a Meso Combustion Chamber with a Detailed Chemistry, *Energies*, Vol. 3, No. 12, pp. 1943–1959, ISSN 1996–1073
- Nagai, N. & Takagi, T. (15 December 2002). *JSME Combustion Handbook* (2nd edition), JSME, ISBN: 4-88898-074-8, Tokyo, Japan.
- Ogami, Y. & Fukumoto, K. (2010). Simulation of Combustion by Vortex Method *Computer & Fluids*, Vol. 39, No. 4, (April 2010), pp. 592–603, ISSN: 0045-7930
- Patankar, S. V. (1 January 1980). *Numerical Heat Transfer and Fluid Flow*. Taylor & Francis; ISBN-10: 0891165223, ISBN-13: 978-0891165224.
- Peng, L. & Zhang, J. (2009). Simulation of Turbulent Combustion and NO Formation in a Swirl Combustor, *Chemical Engineering Science*, Vol. 64, No. 12, (June 2009), pp. 2903-2914, ISSN: 0009-2509
- Stefanidis, G D. & Merci, B. (2006). CFD Simulations of Steam Cracking Furnaces using Detailed Combustion Mechanisms, *Computers & Chemical Engineering*, Vol. 30, No. 4, (15 February 2006), pp. 635–649, ISSN: 00981354
- Warrants, J. & Maas, U. (2006). *Combustion* (4th edition), Springer-Verlag, ISBN-10: 3-540-25992-9, ISBN-13: 978-3-540-25992-3, Berlin, Germany

Nuclear Propulsion

Claudio Bruno

*DIMA, University of Rome (La Sapienza), Roma
Italy*

1. Introduction

Nuclear propulsion (NP) concepts go back to the very end of WW II. Scientists informed about the effects of the US atomic bomb thought of exploiting its energy release for applications like commercial electric power generation, but also rockets and space flight [Shepherd and Cleaver, 1948, 1949; Bussard and DeLauer, 1958]. However, space flight was still considered science fiction, and the military had to deal with more concrete things, like the Cold War. Thus, besides power generation, second stages of ICBM, submarine propulsion, long range and long duration airplanes and missiles became the focus of nuclear energy applications.

It was the second-stage and airplane application that drove R&D in nuclear propulsion. With the advent of reliable ICBM (the Atlas missile) and lighter fission and thermonuclear warheads, a nuclear-powered second stage became no longer necessary. Airplane applications were found impractical: the Convair NB-36 required such a heavy lead shield for the crew that testing and operation were much restricted. Nuclear-powered missiles were easier to design, e.g., project PLUTO, but still far more complicated compared to conventional. The Soviets investigated airplanes and rockets powered by nuclear power as well, and discarded them too.

The history of NP can be found in [Czys and Bruno, 2009, Chapter 7; Lawrence, 2008; Lawrence et al, 1995; Gunn and Ehresman, 2003; Dewar, 2004] and will not be reported here. Basic technology is also discussed in the references above, in particular reactor design is in [Lawrence et al, 1995]. In the US investing in NP technology by NASA and LASL stopped in 1972, but was kept alive by the USAF, and to some extent also at Los Alamos Science Laboratories (LASL, now LANL) who focused on unconventional concepts. In the former Soviet Union, now Russia, research was slowed down, but never completely stopped, because it was thought indispensable for future space missions.

The present work focuses rather synthetically on NP based on fission. (Fusion has been demonstrated, but the net power gain budget is still negative, and power generation is still far in the future, so NP based on this technology will not be discussed.) Note that only reluctantly space agencies or international study groups like ISECG [ISECG] are beginning to include *fission-based NP* in their exploration scenarios. The cost of R&D and experimenting with NP technology is in fact much higher than with conventional rocket propulsion, and until now MIR first, and then ISS, have absorbed most of the human exploration R&D moneys. After completing the ISS, space agencies are now discussing in-house and among themselves what should be the next goal of *human* space exploration. This

question is critical to the survival of space industry in the US, EU and Japan; less so in Russia, and definitely not in emerging space-faring nations such as India, China and S.Korea. Note the *satellite* industry is in much better financial shape and has no such problem.

2. Human exploration: Constraints

In a long-term vision, the question becomes, what will be future manned interplanetary missions, and what do they need? NASA and Roskosmos are looking at 'far' destinations, i.e., not the Moon, although the Moon is still seen by most space planners an indispensable stepping stone to Mars and near Earth objects or asteroids (NEO or NEA). NASA favors, or favored, NEO missions as less expensive than going to the Moon, since no lander is needed (currently, no space agency is planning to build a man-rated Moon lander), but the life support system for a NEO mission does not exist yet, and, in fact, NASA is in the process of reassessing NEO vs. Moon goals. Such issues are more and more influenced by the growing consciousness of the health risks posed to crews by solar and cosmic-galactic radiation (SR and GCR). These risks are still difficult to quantify, but caution suggests to shorten interplanetary missions, e.g., to Mars as much as feasible. In turn, this means discarding conventional Hohmann trajectories (minimum energy trajectories). A Hohmann-type, chemical propulsion (CP)-powered mission to Mars lasting 9-mo or more each way would result in a total dose to the crew > 1 Sv, while the average dose per person on Earth, due to all sources (solar and background) is about 0.6 mSv per year. Although much research is necessary to quantify *the damage* due to space radiation, estimates of excess cancer probability due to 1 Sv range from a few percent to 30% [Durante and Bruno, 2010]. Even the lower bound is considered unacceptably high by most life scientists.

Note that the Earth magnetic field protect us from radiation provided we are within the van Allen belt. Belt-crossing itself exposes humans to much increased doses; outside the belt, in space, radiation dose depend on SR + GCR, including the SR contribution due to [still unpredictable] solar flares.

The SR + GCR dose to astronauts is directly proportional to the duration of exposure, i.e., dose = (radiation flux) \times (exposure time). To reduce dose there are two logical solutions: either to reduce flux to crew by some form of shielding, or to shorten exposure time (faster missions).

2.1 Shielding

Concerning the first solution, shielding depends mainly on the areal mass density (usually given in g/cm²) of the structure surrounding the astronauts. Typical ISS values, depending on location and thickness of scientific equipment (that contributes to areal mass), vary between 5 and 20 g/cm². This would be insufficient during a solar flare, where proton flux intensity may increase by a factor 10 to 50, or on the Moon. Dividing areal density by material density gives a good idea of the material thickness of the shield, so increasing areal density is effective in reducing dose, but translates also into increasing shield mass and thus orbiting cost; secondary radiation, due to interaction between high energy GCR and shield, remains a problem. *Active* shielding (electrostatic or electro-magnetic, EM) has been considered, e.g., by the Nobel Prize in physics Sam Ting, who worked on the original superconducting (SC) magnet of the Alpha Magnetic Spectrometer recently installed outside the ISS to detect dark matter. After discarding the superconducting magnet ($B \sim 1 - 2$ Tesla $\equiv T$; $1 T = 10,000$ gauss) in favor of simpler permanent magnets [Covault, 2011], Ting thought of testing SC magnets in space to

deviate high energy GCR particles and thus protect future astronauts. However, the B field required is of order 10-12 T or higher for a sufficiently large crew habitat. Besides feasibility, the high-B health effects on humans are still unknown; and even 10 T cannot deviate or stop the highest energy band of the GCR spectrum.

While research in this area continues, most experts have concluded that active shielding is for the foreseeable future still unfeasible [Parker, 2006], and that the cost of orbiting to LEO the massive shield required to reduce crew dose during a CP-powered ('long') Mars mission is much too high.

2.2 Interplanetary orbits

Thus, for the time being, the second solution (shortening travel time) should be sought. Any way to speed up interplanetary travel pays off in terms of astronauts safety. We have *some* understanding of the effects of microgravity on bone mass and metabolism, but we are just now beginning to explore the *psychological* consequences of crew confinement inside cramped quarters for many months (the recently completed Mars 500 ground experiment is the first example).

The question at this point becomes of if and how interplanetary trips may be shortened. Here this question is discussed by using basic physics, leaving details (but no show stoppers) to future engineering. In this context the basic physics of interplanetary travel is presented below, starting from what is current practice.

In fact, all present or planned mission are dominated by the key constraint to save mass to be orbited. Anything to be used for space missions must first be orbited, at a cost, for LEO, roughly between 10 and 20 k\$/kg, depending on launcher provider. With CP, most of the mass is propellants, see Eqn (1). Physics tells us that once in LEO, the minimum energy (or the minimum ΔV) required to change orbit, for instance to go to Mars, is that of Hohmann trajectories. These are *ellipses* osculating the starting orbit (for instance, that of the Earth around the Sun) and the final one (for instance that of Mars), having assumed the two are on the same ecliptic plane. Hohmann trajectories require the least expenditure of energy, thus the least expenditure of propellants, and are realized by applying thrust (F) for a time $\Delta t \ll$ of the total interplanetary travel time T . In the limit $\Delta t / T \rightarrow 0$ thrust becomes impulsive, and it is this feature that minimizes energy losses due to non-parallel spacecraft F and instantaneous velocity V . If F is impulsive, F and V are parallel, and all thrust is used to increase V by an assigned ΔV . It is always possible for a spacecraft to change its orbit by using slow acceleration ($\Delta t / T$ finite \rightarrow low F); however, in significant gravitational fields such as Earth's, the trajectory becomes initially a *spiral* where F and V form an angle and thrust work is partly wasted. Low thrust ('spiraling') losses may be relatively large, e.g., in a low thrust trajectory from LEO to L1 (theoretical ΔV about 4 km/s) the actual 'expenditure' of ΔV may become about 6 km/s; see [Martinez-Sanchez and Pollard, 1998] for a LEO to GEO quantitative example.

2.3 Mass consumption

Using CP consumes much propellants mass. Tsiolkovski's equation states

$$\Delta V = I_{sp} \ln(M_{in} / M_{fin}), \text{ or } M_{in} / M_{fin} = \exp(\Delta V / I_{sp}) \quad (1)$$

and, with CP the specific impulse I_{sp} , defined by $I_{sp} = F / (dm/dt)$, happens to be of the same order of the ΔV needed to change orbit in our solar system, see Figure 1 showing the ideal ΔV

with Hohmann orbits associated, for instance, to a Mars mission. The numbers tell the mass fraction will be $\gg 1$. In Eqn (1) the (dm/dt) is the rate of mass consumption of propellants ejected from the spacecraft.

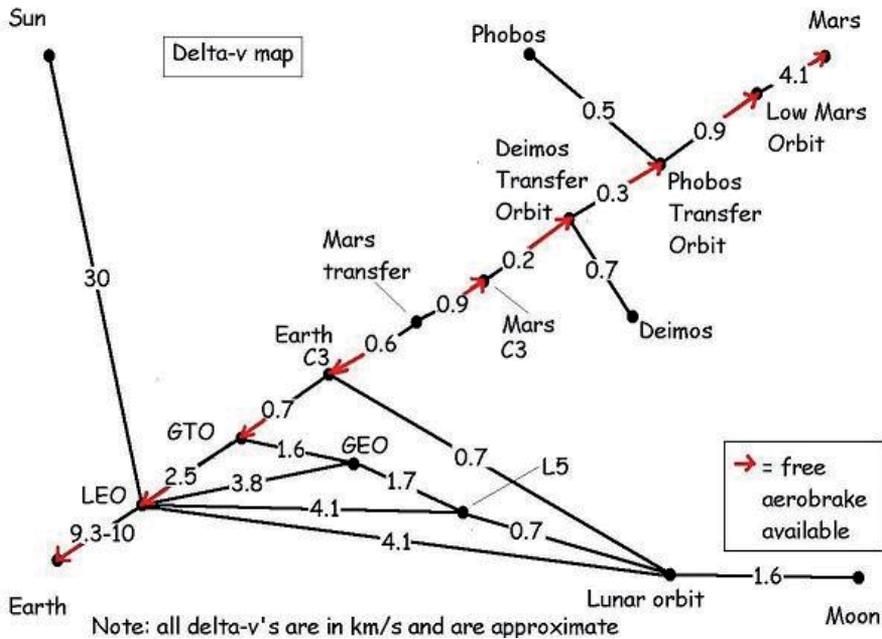


Fig. 1. Minimum ΔV using Hohmann's orbits associated to a Mars mission

Note that, according to Newton's second principle, F is the rate of change of momentum of the mass ejected. Assuming for simplicity that this mass has zero velocity prior to being accelerated, the instantaneous momentum change is V_e times (dm/dt) . Provided the acceleration process is ideal, i.e., 100% efficient, the Isp just defined coincides with V_e , the ejection velocity. In real chemical rockets V_e and Isp differ, because the gasdynamic expansion of combustion products is incomplete, that is, never reaches zero pressure and temperature.

Actual Isp values depend on fluid and on energy conversion process. In CP this means: maximum and minimum cycle temperatures; in practice: combustion temperature, average molecular weight of combustion products, and expansion ratio. The highest practical Isp is that of the LOX/LH₂ combination: in space, about 4500 m/s (or 4500/g seconds, g being the standard gravitational acceleration, about 9.81 m/s², when using old engineering units). Isp may be higher with 'exotic' combinations, e.g., LF₂ + LO₂ and others, but their logistics and safety implications are forbidding. In practice, CP limits Isp to about 4.5 km/s. The Tsiolkowski equation shows that to achieve a certain ΔV the mass to eject depends exponentially on $\Delta V/Isp$, this ratio being of order unity.

Thus ways to raise Isp are the key to economically feasible fast interplanetary travel.

3. Propulsion

The Tsiolkowski equation (1) is the form taken by the second Newton's principle when the mass of an object being accelerated varies with time. In fact, the physics we can bring to bear to solve the problem of faster travel is based solely on Newton's three principles. The third,

the most important in all kinds of propulsion, states that in space velocity of an object can be increased only *by ejecting mass*, that is, accelerating it from 0 to V_e (in the spacecraft reference system): so far, no massless ‘space drive’ has been invented.

Mass may be ejected either 1. via *internal* molecular collisions (i.e., through thermodynamic expansion), or 2. by applying an *external* force *directly* to atoms/molecules. These two different strategies are common to all space propulsion systems.

1. CP is based on the first strategy: in special relativity, the relativistic potential energy $PE = (\Delta m) c^2$ of a mass Δm becomes kinetic energy (KE) of the remaining mass m , that is $\frac{1}{2} m (V_e)^2$. The percentage $\Delta m / (m + \Delta m)$ of mass converted into KE depends on the *fundamental force* used. CP uses the potential of the *second force* (electro-weak), the potential energy released when chemical bonds are broken and rearranged.

In fact, the Standard Model of physics includes *only three* fundamental forces: gravity, electro-weak (the result of unifying electro-dynamic and weak forces in the ‘80s), and the strong, or nuclear, force. The electroweak force binds molecules and atoms together, and is responsible for the existence of matter as we know it. The strong force binds sub-atomic particles (nucleons) together, preventing them from disintegrating due to Coulomb repulsion.

The $\Delta m/m$ fraction varies from 10^{-27} of gravity to about 10^{-3} of the nuclear force. In classic mechanics, the simple 0-D relationship between PE change and KE gain is

$$\Delta(PE) = \frac{1}{2} m (V_e)^2 \tag{2}$$

predicting that the V_e (the Isp) scales with the *square root* of $\Delta(E)/m$, a weak dependence. (Note the equation above does not refer necessarily to a fluid.)

Table 1 shows conversion ratios and energy/unit mass, J, of the three forces. (metastable nuclear isomers are still a controversial subject, see [Czysz and Bruno, 2009].)

Type of force	Potential	alpha	Energy density, J (J/kg)
Gravity	gravitational	10^{-27}	$10^{-11} (*)$
Electro-weak	chemical (H_2/O_2 combustion)	1.5×10^{-10}	1.35×10^7
Strong Force	Nuclear: Fission (^{235}U)	9.1×10^{-4}	8.2×10^{13}
	Fusion (D-T)	3.75×10^{-3}	3.4×10^{14}
	Metastable (^{180m}Ta)	2×10^{-7}	1.8×10^{10}
	Annihilation ($p^+ - p^-$)	1.0	9×10^{16}

Table 1. Force, potential, conversion factor $\alpha = \Delta m/m$ and energy density $J = PE/m$ of the three fundamental forces. The J of gravity refers to two unit masses at 1 m distance [Czysz and Bruno, 2009].

Comparing these three forces shows that the J of CP, based on the second, is eighteen orders of magnitude larger than gravitation, but six smaller than the third. Dividing J of the second force by the specific heat of chemical reactions products yields the maximum temperature of a CP cycle, not much higher than 3500 K, and anyway limited by the melting point of materials

(ruled by the same force). When using realistic numbers, Eqn. (2) shows that the V_e and I_{sp} are correspondingly limited to 4-5 km/s. Hence, as said, I_{sp} is of order of the ΔV needed to change orbit, and Eqn (1) predicts large mass consumption: for instance, the ΔV from ground to LEO (in practice, 9-9.5 km/s) results in a $(M_{in} - M_{fin})/M_{in}$ ratio, the payload fraction, of order of a few percent. Thus, during a satellite launch, almost all propellants mass and energy release is used to lift and accelerate the propellants themselves, not the payload.

2. When adopting the *second* strategy, [ionized] molecules/atoms are still accelerated by electroweak forces, but the Coulomb or Lorentz force is external to the plasma and requires an *external electric power source*. The V_e possible are much larger than in a thermodynamic expansion, since they are neither limited by the upper and lower temperatures of the cycle, nor by the melting point of materials. Accelerating ionized matter by Coulomb or Lorentz is, in principle, limited only by relativity, although in practice voltages and magnetic fields cannot be made large at will by arcing, magnet mass, critical current density and Meissner effect.

Using *classical* (Newton's) mechanics, the conclusions for space propulsion are as follows:

- I_{sp} scales with V_e
- Thrust, F , scales with (flowrate) times (V_e) , thus with $(V_e)^2$
- Power of the ejected flowrate is the kinetic energy (of the mass ejected) per unit time, or $\frac{1}{2}$ (flowrate) times $(V_e)^2$, and thus scales with $(V_e)^3$.

(*Relativistic* relationship are formally different, but predictions differ only when V_e is $> 0.99c$, and will be ignored here).

Based on these simple scaling laws, to reduce mission time F must be increased. This implies V_e must be as large as possible; when doing so also mass consumption decreases. There is a price, however, and that is the power required, scaling with $(I_{sp})^3$. As said, CP is *fundamentally limited in I_{sp}* (in V_e) by $J = \Delta(E)/m$ that can be extracted from rearranging chemical bonds.

Thus the next logical step is to seek *other potentials* providing *larger energy density* J . Of the three fundamental forces we know, only the *nuclear force* remains, since gravity is so weak that only assist (flybys) manoeuvres, exploiting the mass of entire planets, can supply enough energy to a spacecraft, at the expense of lengthening, not shortening, mission time. The *nuclear force*, see Table 1, converts about 0.09% of the 'fuel' mass into KE in the case of ^{235}U fuel, and 0.3% in the case of D-T (deuterium-tritium) fusion, a factor $\sim 10^6$ larger than possible in chemistry. It is logical to look at nuclear reactions as the *only solution* to the need to increase J , thus I_{sp} , thus thrust, finally enabling faster missions with drastically lower mass consumption. Thrust, and travel time, scaling with $(V_e)^2$, will depend on the NP strategy outlined already. The first, nuclear thermal propulsion, or NTP, has produced in time the family of NP devices called nuclear thermal rockets (NTR), the second, nuclear electric rockets, or thrusters, or (generally) nuclear thermal propulsion (NTP). Strategy 1 began to be explored in the '50s. Strategy 2 was explored by Stuhlinger during WW II, but started being investigated only much later.

4. Nuclear thermal propulsion (NTP)

4.1 Principles and performance

In both US and SU NTP took the form of a compact nuclear reactor (NR) transferring heat to a coolant (e.g., LH2) brought to a temperature T limited only by materials. This temperature was about 2300 K in US reactors, close to 3000 K in SU. Heated hydrogen was then

expanded in a conventional nozzle, similarly to what done in CP rockets. The Isp of such NTR is higher than in CP, because the molecular weight of GH₂ (about 2) is a factor ~ 5 lower than that of a LOX/LH₂ rocket (about 10). See [Lawrence, 2008] for details.

The heat produced by a NR is due to nuclear reactions 'burning' (that is, fissioning) nuclear 'fuel', e.g., ²³⁵U-enriched uranium, for instance, encapsulated in bars, rods or pins. Since U has a melting point about 900°C, it is alloyed with O, N or C to form U-ceramics, with much higher melting point. Fissioning splits U nuclei into smaller fragments (FF), including neutrons fissioning other U nuclei. This chain reaction produces therefore FF consisting of unstable isotopes (most belonging to one or another of two families, with mass centered around 90 and 130, respectively); neutrons; and gamma photons. FF possess kinetic energy in the O(1) to O(100) MeV range. By colliding with nuclei of the rod material and of the fuel itself, they deposit their kinetic energy as heat ('thermalize'). In a solid-core NR this heat is transferred to the coolant/propellant flowing inside the many channels present in each rod, see Figure 2.

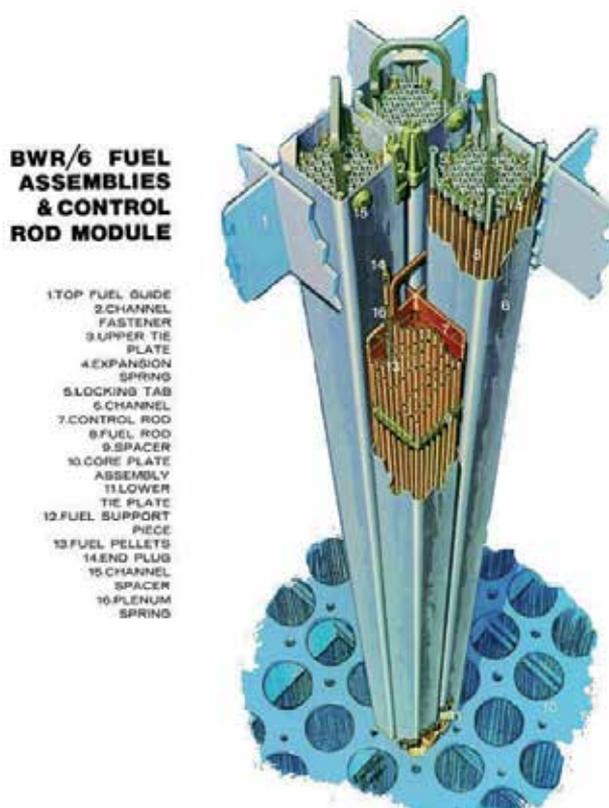


Fig. 2. Sketch of a single fuel bar in a solid-core, commercial Boiling Water NR

This strategy is the same of LRE, except the heat source is external to the propellant, not due to its chemistry. It started in the US, where space NR research (under the ROVER project)

became the responsibility of Los Alamos Scientific Labs (LASL, now LANL). Westinghouse was responsible for engineering the LASL reactors into *rocket engines*. After ROVER was terminated in 1972 by the Nixon administration, the USAF took over and built NR using ceramic fuel (CERMET) capable of withstanding hundreds of on-off thermal cycles without cracking, and more compact NR using fuel *pebbles* instead of bars (pebble bed reactors, PBR). In terms of sheer power the high point of ROVER was the Phoebus IIA reactor, tested at 4200 MW for 12 minutes. Less powerful (10-100 MW) but more versatile NR were designed and tested; see [Turner, 2005; Gunn and Ehresman, 2003] for details.

Not as well documented is NR work by the US Navy. The Navy wanted torpedoes faster than conventional (60-70 knots). The Brookhaven National Labs (BNL) designed NR for a standard-size torpedo (about 25 in dia.). Power was in the tens of MW range (note that life expectancy was only few minutes). Thus thermal power/unit volume was roughly of order MW/liter. In time, this work was applied conceptually to a family of space NTR concepts with the MITEE acronym by the US Plus Ultra Technologies company. Plus Ultra Technologies looked at even more compact NR using fuels with low critical mass, e.g., Am 242 and others [Powell et al, 1998, 1999]. Lower critical mass means lower fuel volume and thus lighter shielding, since the thickness of a neutron + gamma shield is a constant for a given material. (with graphite, a few decimetres), and thus shield mass scales with NR area. Am-based fuel was also proposed in Israel by Prof. Y. Ronen at Ben Gurion University, and by C. Rubbia in Italy for their independently developed FF NTR powering manned Mars missions (M3), see [Accettura et al, 2008; Ronen et al, 2000].

Still not translated is most SU/Russian work. Contrary to what happened in the US, SU work in NTP never stopped, and All Union conferences have been held until recently. In 2010 General A. Perminov (then head of Roskosmos) claimed to have the NP technology for a M3, and his statement was echoed by a statement by NASA Administrator Bolden, who went in print wishing to collaborate.

Russian capability is apparent from articles, e.g., [Koroteev et al, 2007] emphasizing the importance of bi-modal approach also proposed at NASA-Glenn by S. Borowski. In this approach part of the thermal energy of the NR is used for thrust, and part to generate electricity to power all spacecraft systems and also high Isp electric thrusters (ET). This strategy can produce large thermal thrust (many kN or even tens of kN, depending on NR power), at moderate Isp (880 s were demonstrated by Westinghouse NTR in the '60s; Russian engines were capable of more than 900 s; 1000 s are claimed feasible with MITEE technology). The thermal thrust would allow a spacecraft to quickly reach escape speed from LEO; then the NTR would be switched off and the much lower thrust ET, with Isp in the 4,000 to 15,000 s range, would be turned on, saving propellant mass during most of the interplanetary trajectory.

Dual mode operation avoids in principle the spiral trajectories inevitable with the low thrust typical of all ET, and thus avoids crossing the deadly van Allen belt tens or hundreds of times. Plus Ultra Technologies has also proposed a dual-mode concept based on MITEE [Powell et al, 2003].

In sum, NTP is a well explored and mature technology in Russia, and still viable in the US, although experts are retiring, retired or passed away. It is capable of thrust in the tens of kN, with Isp in the 800 to 1000 s range. Uranium nitrides and carbides (and more advanced) fuels are available. The large thrust of NTP enables interplanetary missions much faster than with CP, but Isp is only about a factor 2 better than with CP, so LH2 consumption implies substantial LEO orbiting cost. This said, in combination with ET could be a viable and very promising propulsion technology, indeed suggesting itself as a candidate system for M3 in

the framework of future global collaboration. Although much less mature, the Ronen/Rubbia engine is another strong NTP candidate, with F in the $O(1)$ kN, and a very interesting $I_{sp} \sim 2000$ s.

4.2 Unconventional NTR

4.2.1 Gas-core NR

Almost all NTR tested in the US and the former SU have a solid core (rods, bars or pins) containing the fissionable fuel. Liquid- and gas-core NR have also been investigated, but (to this author's knowledge) never built or tested. US and Russian work [Howe et al, 1998; Koroteev et al, 2002] has focused on gas reactors because their I_{sp} (and thrust) may be much higher than in solid-core NR, e.g., in the many 1000s of s. Their main difficulty is how to transfer energy from fissioning hot gas to propellant. The only heat transfer mechanism found to reduce the inevitable mixing losses is radiation, and that requires a heat exchanger made of a radiation-transparent material. This concept goes under the name of "nuclear light bulb", see [Czysz and Bruno, 2009, Section 7.12].

4.2.2 Rubbia and Ronen FF engine

Another unconventional NTR concept briefly mentioned was that originally proposed in [Shepherd and Cleaver, 1947] and rediscovered independently by C. Rubbia and Y. Ronen in the late '90s [Accettura et al, 2008; Ronen, 2000]. It is based on thermalizing the fission fragments generated by fissioning an unconventional fuel (e.g., ^{242m}Am) directly inside the propellant, without first heating solid rods. In this concept GH2 flows inside a chamber, or channel, where walls have been coated with a fissionable thin layer of fuel (Y. Ronen has theoretically proved the feasibility of fission inside a thin layer). The isotropically emitted FF thermalize partly inside the flow of GH2, that heats and then expands in a nozzle, and partly inside the coating, that must be cooled. The coolant may be the propellant itself, f.i., LH2, that is injected and gasifies in the channels. This strategy is similar to regenerative LRE cooling. The advantage of the Rubbia/Ronen Engine is its I_{sp} , predicted in the 2000 - 2500 s range, coupled to the significant thrust typical of all NTR. Assembling a cluster of such engines can supply many hundreds, perhaps thousands, of N in a compact volume. Extant problems are cooling of the reactor walls, of the nozzle, and materials.

4.2.3 Propellant-less NTR concepts

This form of NTR avoids all external propellants. That is, the NTR uses as propellant the FF themselves [Czysz and Bruno, 2009, Section 7.13]. Neutrons and gamma photons do not contribute to thrust, being too light, but isotopes created by fissioning U do. These have energies of order $O(100)$ MeV, with theoretical I_{sp} in the 10^6 to 10^7 s range, but their thrust is very small, and collimation hard to achieve. The concept was pushed to its limit by Eugen Saenger, the inventor of the photonic rocket [Saenger, 1958]. This rocket uses a NR to heat a paraboloid surface to high T . The surface emits and collimates photons, so the *ideal* I_{sp} available is the speed of light $c = 3 \times 10^8$ m/s). In fact, when accounting for consumption of fissile fuel mass the effective $I_{sp} = F/(dm/dt)$ is less than c , and Saenger calculated actual V_e of about $O(100)$ km/s. The same concept was proposed in the '90s by C. Rubbia, who was not aware of the Saenger work. In Rubbia's example, a 1-Gw NR perfectly converting all its energy to light would produce 3 N, with $I_{sp} = c$ (Rubbia did not include in I_{sp} the mass consumed by the Einstein relationship). Although fascinating, the S&T problems posed by the photonic rocket are immense: to achieve the performance predicted by Rubbia

the temperature and emissivity of the paraboloid surface would have to be that of a perfectly black body at about 4000 K. This said, the concept is interesting and (to 0-th order) physically sound, although hardly realizable with foreseeable technology.

4.2.4 The Orion project

This concept offers an alternative to the dichotomy “either high I_{sp} and low F , or viceversa”. The Orion is an idea (1947) of Stan Ulam (the father of the US thermonuclear bomb). Ulam thought of exploding atomic ‘bombs’ inside a rocket chamber to produce thrust, but quickly realized 1. that atomic explosions created pulses of energy and momentum many orders of magnitude larger than chemical, and therefore 2. that there was no way of confining an atomic explosion inside a chamber. Thus the solution was to have the atomic explosion(s) take place *outside* a vehicle, unfortunately with possibly destructive effects on its structure.

Thus his remained a pure idea until the 1954 US thermonuclear bomb test on Bikini: heavy steel spheres coated with graphite and suspended before the test, were found intact but at some distance away. The data analysis that followed the test showed that the ablating graphite protected the spheres, and that their motion was the effect of gas-dynamic (and radiation) momentum transfer from the expanding fireball gas to the spheres. These facts fueled interest in a proposal to the AEC, and the “Orion Project” took off in 1958. It lasted until the 1963 Partial Test Ban Treaty banned all atmospheric tests that killed it.

The Orion concept consisted of accelerating a spacecraft by periodically exploding astern atomic ‘bombs’ of yield from 0.03 kton to 0.35 kton each. The yield depended on spaceship mass, while the number of explosions was fixed at 800 to reach LEO. A massive thrust plate fitted with shock absorbers received the momentum pulse from each explosion and accelerated the spaceship at 2-4 g.

In fact, under Orion a family of spaceships was developed capable of orbiting to LEO payload from 300 to 6100 ton. The word ‘spaceship’ is appropriate, as all were single-stage (taking off from the ground) see [Ewig and Andrews, 2003] for details. The ‘bombs’ were in fact sophisticated atomic shaped charges. A sketch of an Orion spaceship is in Figure 3.

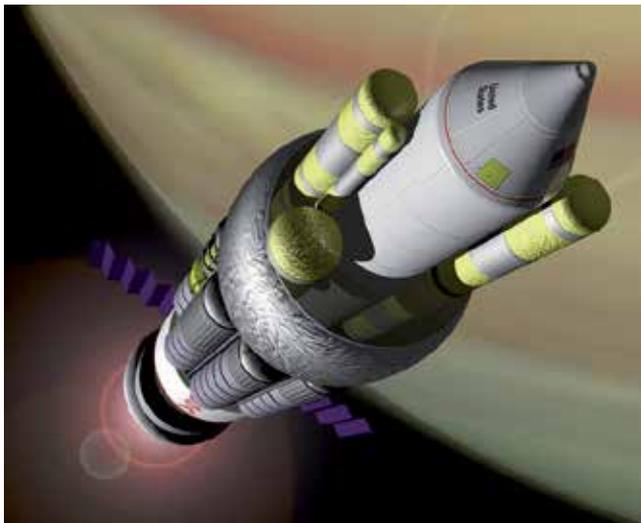


Fig. 3. Notional sketch of an Orion spaceship near Jupiter. Thrust plate and spacing between plate and shock absorbers assembly are shown.

All mechanical, thermal and control issues associated with this method of propulsion were analyzed, in particular the thrust plate, and all eventually were solved with engineering know-how of the time. A test with conventional HE confirmed performance and scaling expected. That requires discussion of Orion peculiar gas dynamics.

In Orion spaceships the energy release occurs *outside* the vehicle, not internally. The fundamental point, as Ulam realized, is the fact that the velocity of the shock wave created by each fireball is *at least* an order of magnitude larger than the V_e of CP-rockets, and actually depends on distance between fireball and thrust plate. This distance is a compromise between capturing as much as possible the momentum pulse and avoiding plate destruction by radiative heat transfer and by the expanding gas momentum. A typical calculated distance was 200 m, where the debris velocity, here still named V_e , was $(2 - 3) \times 10^4$ m/s.

In fact, fundamental thermodynamics, i.e., Eqn (2), shows that a 10 kton (4.18×10^7 MJ) explosion produces after a few ms a 100-m fireball at temperature ~ 50 to 70 k°C. The fireball expands at speed V_e scaling of course with (temperature)^{1/2}. For the 10 kton yield, initial V_e is ~ 100 km/s, dropping with the inverse of distance squared. With higher yield the V_e scales according to Eqn (2).

It is the hot gas at velocity V_e that transfers most of its momentum (some is also imparted by radiation pressure) to a surface; if the energy release is isotropic, the fraction transferred depends roughly on the solid angle of the cone with apex in the center of the fireball and base equal to the thruster plate surface. In fact, each 'bomb' was conceived as a *shaped charge*, that is, containing a 'reaction' mass designed on purpose to impinge on the plate, so the momentum calculated was higher than that predicted solely on the basis of the angle.

In practice, the *empirical* (but physically consistent) relationship for the 'equivalent I_{sp} ' of each explosion pulse is $I_{sp} = (C_0) \times (V_e)$, where C_0 is a collimation factor of order 0.5 when fireball diameter and thrust plate diameter are well matched.

The calculations indicated this NP method ensures equivalent I_{sp} *about 2-3 times larger* than any NTR, with a thrust F scaling with $(I_{sp}) \times (\text{reaction mass of each bomb}) \times (\text{explosion frequency})$. In most of the Orion family spaceships the pulse frequency was 1 Hz. F can also be estimated from ship mass and its impulsive acceleration data. The F reported for the smallest Orion spaceship (880 ton at take-off) was about 10^6 N.

If restricted to space, the Orion concept is still valid. In his "Cosmos" TV series Carl Sagan proposed, in fact, to use the charge of nuclear warheads dismantled because of nuclear ban treaties, as fuel for future Orion-type spacecraft. It was estimated that a single Orion spaceship could lift a complete human outpost to the Moon.

Based on physics, the Orion concept offers the largest possible I_{sp} and F combination for a given energy pulse (yield). Engineering it into a spacecraft has its problems, but the work done indicated they are solvable. Rather, the issues are political (exploding bombs in space), and the availability and consumption of nuclear fuel to build the shaped charges. Given the political will (a rare commodity these years) Orion could be made into a true, *reusable* spaceship. In fact, Orion showed how to exploit the third force to create the ultimate (high I_{sp} , high F) NP system, the only capable of exploring our planetary system in reasonable times.

The Orion concept was resurrected in recent years and molded into an ET called MagOrion, then scaled down to the MiniMagOrion (MMO) described in Section 5.1.

5. Nuclear Electric Propulsion (NEP)

The second NP strategy consists in using a NR to generate electric power fed to an ET. ET may use either the Coulomb force, e.g., Gridded Ion Thrusters, GIE, or the Lorentz force, as in magneto-plasma-dynamic (MPD) thrusters. Hall thrusters, VASIMR and others are of this type. Their description is documented, for instance, in [Goebel and Katz, 2008; Rossetti et al, 2008]. A review of NEP has been recently published in [Cassady et al, 2008]. Here the focus will be on near term to long term ET powered by NR rather than on current ET technology.

GIE are mature and are commercialized, e.g., as the Boeing XIPS family, with Isp 4000 to 4500 s. Xe is the propellant of choice for GIE. Ionizing Xe absorbs 5-6% of the total electric power. More importantly, in all types of ET ionization means low pressure, because recombining ions and e^- need a third body, and thus recombination rate grows with (density)³. Because of low pressure operation, the actual thrust of GIE, and ET in general, is very small. The thrust of a state of the art 'high thrust' satellite GIE (powered by solar power) may be a fraction of N. F is basically limited by solar power available, in turn limited by solar panel area and efficiency. If significant thrust is desired, GIE must be assembled in clusters. These would be orders of magnitude bulkier than CP thrusters in the same thrust class. However, their mass consumption, ruled by Isp, is orders of magnitude less.

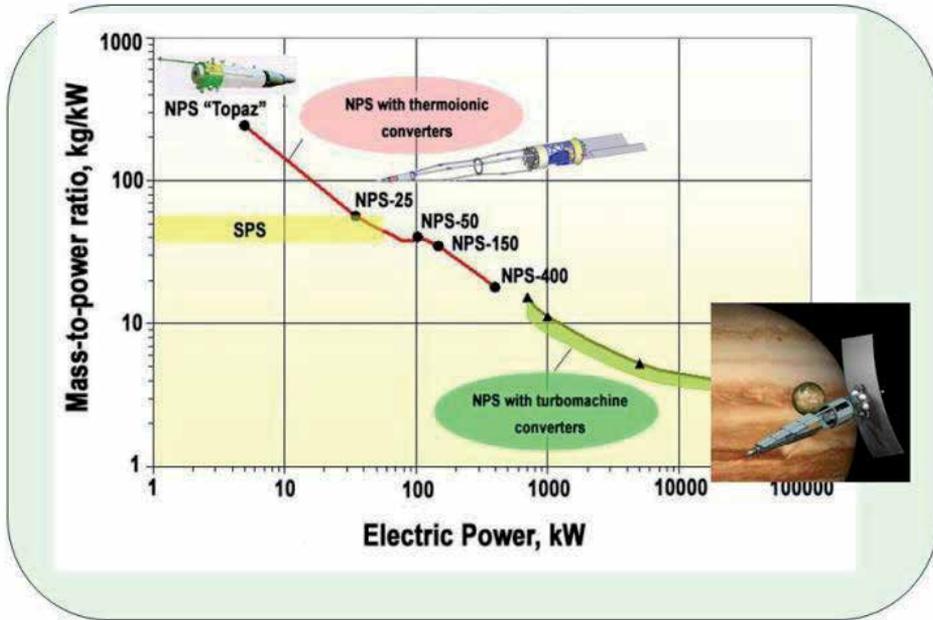
The low F of ET means that significant spacecraft ΔV requires either engine-on time comparable to the entire mission duration, $\Delta t/T \sim 1$, or that an ET consisting of tens or hundreds of individual modules, with a corresponding increase in electric power. A key issue with high power (high voltage) GIE is grid life, since some of the ion beam current impacts the grid system. With most MPD thrusters the key issue is anode/cathode erosion. In both cases these issues are due to ion impingement, and determine ET life. Because ET thrust is 'low' (at least using solar power), mission time with ET propulsion may be years, as in ESA's SMART-1 or JAXA's Hayabusa spacecraft.

The VASIMR engine being developed by the Ad Astra Company [AdAstra, 2011] by F. Chang-Diaz, and to be tested on the ISS in the near future, is as promising a MPD thruster as it is ambitious. It is electrode-less, with ionization and energy feed realized by plasma radio-frequency. Its technology is borrowed from tokamak fusion technology. As reportedly planned, it should deliver *any* combination of F and Isp while keeping their product (\sim power) fixed. The Isp of VASIMR tested so far is still about 5000 s rather than the 10,000 s planned, seemingly due to power losses ~ 30 to 40 %. These are under investigation but are probably due to radiative heat transfer and incomplete conversion inside the magnetic nozzle of the ionic circumferential velocity component to axial. A 200 kWe VASIMR will be tested on the ISS in the near future, but the power needed by tests will be from batteries, not a NR. VASIMR systems in the 10s and 100s of MW have been proposed; if built, they will need nuclear power: solar power (and thermo-ionic power) is inadequate above a few 100s of kWe.

The disadvantages of NEP noted, its advantages are remarkable in terms of Isp and ease of coupling between power generator and ET. Note that EU and the US have extensive GIE know-how, while Russia leads in Hall (MPD) thrusters. In addition, Russians have designed many conceptual dual-mode engines, and advocate NP as the best solution for a M3 [Karash, 2011]. VASIMR is the most promising concept in ET, but is still not mature. In any case, low F remains the Achille's heel of all ET.



SPECIFIC MASS OF SOLAR AND NUCLEAR POWER SYSTEMS OF SPACECRAFTS DEPENDING ON GENERATED POWER



7

Fig. 4. The alpha parameter (kg/kW of propulsion system) of thermo-ionic and NR-powered electric propulsion systems. Courtesy of A. Koroteev, Keldysh Research Center (KERC), Russia.

Clustering nuclear ET may solve these problems, but if the NR must function reliably and autonomously for most of the mission, the issue becomes that of NR refuelling. The NP group led by A. Koroteev at KERC has presented NR concepts where refuelling is done mechanically. To date, this concept has not been tested.

NEP has other drawbacks besides low F . The thermal to electric conversion step is one. Standard conversion is thermodynamic (e.g., a Brayton closed cycle with heat exchanger). The mechanical power drives alternators/dynamos. In the MWth range this technology is considered mature by many. Unofficially, Soviet NR using liquid Na or NaK eutectic are known to have powered some of their reconnaissance satellites requiring $O(100)$ kWe. Conversion efficiency (thermal to electric) $\eta \sim 25\%$ is feasible with total mass is ~ 1 to 2 ton, including shielding. Heat extraction may recommend a gas cycle, a liquid metal cycle, and may use heat pipes, see [Lenard, 2008].

Because of thermodynamics, the third drawback of NEP is the 'bottom' T at which the $(1-\eta)$ fraction of the thermal power not converted to mechanical work must be dumped. Dumping must be done by a space radiator, and it is a fact that such radiator may be more massive than the NR itself, and certainly bulkier. Note that the bottom temperature of a standard NEP cycle is critical: too low, it cannot radiate away enough waste power; too high, the η drops.

Thus NEP not only may waste up to $\frac{3}{4}$ of the NR power, but adds much weight and bulk to the spacecraft. This explains why the parameter $\alpha = \text{mass/power of electric propulsion systems}$ (measured in kg/kWe) is at the moment about 10 or a little less, see Figure 4. For comparison, NTR systems have α about 100 times lower. α is very important in determining the performance of EP-, or NEP-powered missions, see [Stuhlinger, 1964] for a must-read reference.

Work in NEP is at the moment being carried on in Russia, the US and France, with Russia favoring dual mode NTR/NEP operation. Within the next 1-2 years a decision about an international Mars mission is likely and will indicate also the preferred propulsion system for future M3.

5.1 Unconventional NEP

Like NTR, all NEP thrusters mentioned (GIE, Hall, VASIMR) use steady operation. The Orion history suggested to Andrews Space & Technology (AS&T) Corporation, in the US, to investigate NEP whereby pulsed nuclear explosions become a Lorentz force periodically pushing the spacecraft. Thus energy pulses are converted to momentum pulses electro-magnetically rather than through gas dynamics. Andrews Space's concept (called MagOrion) is therefore a pulsed MPD thruster. The frequency of atomic detonations was assumed 1 Hz. This concept was called MagOrion. Each explosion produces ionized debris, that interacts electro-magnetically with a 2-km superconducting ring. This accelerates the debris and produces thrust, see [Ewig and Andrews, 2003; Lenard and Andrews, 2006] for details. The thrust/weight ratio was predicted to be in the 0.2 to 10 range (0.2 to 10 m/s^2 acceleration).

Because of size (especially that of the superconducting ring, impossible to achieve in practice) MagOrion was eventually scaled down, taking the name MiniMagOrion (MMO). In MMO the superconducting ring was replaced by a magnetic nozzle analyzed in collaboration with the University of Washington. MMO was tested piecewise, meaning that some ersatz components were tested using the unique facilities and energy sources at the Sandia National Laboratories.

This partial testing confirmed part of MMO physics. The future of this concept envisaged in 2006 was a manned spacecraft, for instance shaped as in Fig. 5. Calculations indicated I_{sp} of order 10,000 s is potentially feasible, with thrust depending on vehicle size and fuel yield, but, typically, in the 10^2 ton. An interesting feature of MMO is that each nuclear fuel pellet is [periodically] compressed to the critical target density (the fuel density times pellet radius, in g/cm^2) by the plasma effect called Z-pinch, driven by fast rising current density. (This method of compression is similar to that proposed by the late R. Bussard to adiabatically compress and fuse D-T fuel in his "QED" inertial confinement fusion concept.)

MMO research at AS&T investigated and compared several fuels, ^{235}U , ^{233}U , and ^{239}Pu among others, and settled on ^{245}Cm , a fuel with the theoretical lowest critical target density. ^{245}Cm is a component of commercial NR waste, with convenient neutronics properties and good 'burn' fraction. In the MMO is in the form of hollow pellets coated with a Be neutron reflector layer.

Estimated performance of the baseline MMO vehicle in [Ewig and Andrews, 2003; Andrews and Lenard, 2006] shows $I_{sp} = 10,000$ s, thrust 1100 kN, and that a $\Delta V = 100$ km/s is feasible with a 100 ton payload. The parameter α is of order 2.5×10^{-4} kg/kWe , to be compared with $\alpha = 10$ or so for solar-powered ET. Although appealing,

the MMO physics has not been investigated in the same depth as Orion's. The efficiency of converting energy pulses to thrust still awaits confirmation; but to this author's knowledge, all research in MMO has stopped.

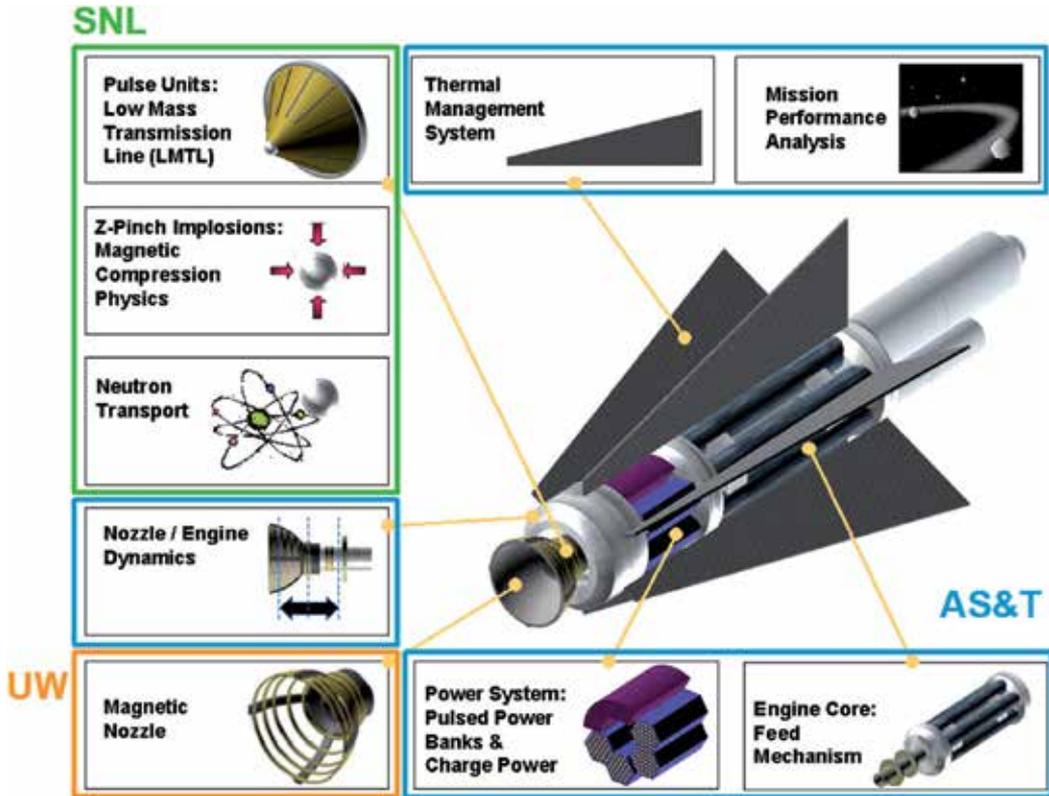


Fig. 5. Sketch of a notional MiniMagOrion spacecraft and of its main components (Ewig and Andrews, 2003).

5.2 Pulsed operation

Orion concepts beg the question of whether releasing energy in pulses is more convenient, or performing, than steady-state operation. Time-averaging constant frequency pulses shows thrust is lower than steady state operation of the same amplitude. In fact, there might be some advantages. It is known, see Section 2.2, that impulsive F results in Hohmann minimum energy transfer trajectories. Besides, ideal propulsion efficiency depends on the ratio between spacecraft speed and exhaust jet speed, and peaks when this ratio is unity. Thus impulsive thrust may be more efficient, as the ratio between [time-averaged] ΔV and exhaust speed may be made close to one 'instantaneously', rather than slowly. In rocket launches the lift-off, when V_e is of order 3 km/s and V is a few m/s, is notoriously the most inefficient fraction of the launcher trajectory.

A third consideration is that fluid dynamics of a suddenly accelerated fluid (f.i., expanding in vacuo, as is the case of Orion), predicts its "limit velocity" higher by the factor $[2/(\gamma - 1)]^{1/2}$ than the steady state limit velocity for the same reservoir conditions [Thompson, 1972].

This said, the case for impulsive Orion and MMO concepts is made simply because the effects of detonating a nuclear charge (V_e and thrust) are per se very large, and because a steady, or quasi-steady nuclear explosion is unlikely (possible exception: X-ray laser technology). In this light the three effects mentioned are probably insignificant.

The limit to this strategy is posed by maximum acceleration tolerable by structure, payload and especially crew. Humans can tolerate steady acceleration ~ 3 g, much higher for much shorter times even if periodically repeated, e.g., 10g can be tolerated for times < 0.1 s. A spacecraft shock absorbing system is a must in this context. In the case of MMO, the EMP may damage payload and humans: these have evolved in a $B = 0.6$ gauss (6×10^{-5} T) environment, and do not take easily to EMP of order 1 T or higher. For instance, anecdotal reports of the effects of $B \sim 1$ T for periods ~ 1 s include 'seeing' colored lights.

5.3 The future of NEP

NEP is being considered by many space agencies and experts as potentially more suited than NTP to manned interplanetary missions. The reasons are partly scientific and partly based on engineering.

With NEP, electric thruster (ET) and NR are effectively uncoupled, and lend themselves to preliminary separate optimization (as the aeroengines for commercial airliners, for instance). Mathematical criteria to pick the best I_{sp} for given mission speed and payload have been developed since the '60s [Stuhlinger, 1964], and reliable ET have been space-qualified and flown for many tens of thousands of hours. The coupling between power source and ET is more flexible and less sensitive to feedback than in NTR. In principle, at constant NR power, that is at constant $(\text{Thrust}) \times (I_{sp})$, thrust and I_{sp} may be chosen to optimize a mission, for instance by using high F to reach quickly escape speed from LEO, and then switching to lower F and high I_{sp} , thus consuming less propellant, during the interplanetary leg.

The specific type of ET may influence the propulsion architecture considerably, and may ultimately determine its future. The thrust of GIE depends on the *square root* of voltage, and 70 kV is cited as the future limit in [Fearn, 2008]. With this voltage the future thrust of a single GIE may eventually reach 10 N. Thus 100 N, a thrust significant for interplanetary missions, will need a (3x3)-module GIE. With high current 60 cm dia. modules, the total assembly bulk is not significant. An advantage of clustering is that the \mathbf{F} vector can be modulated step-wise. At the same time, heat dissipation and arcing may be problematic.

On the MPD side, high power thrusters (e.g., Hall, VASIMR) absorb many 1000s A. Ohmic losses may become limiting, suggesting superconductor wiring and magnets. Using LH2 propellant would enable the switch to this technology. Ground testing of MPD ET becomes another problem at the power level enabled by NEP, dictating individual or of downscaled module testing, when feasible. One positive side of MPD thrusters is their higher thrust per unit power, even though that means inevitably lower I_{sp} at fixed NR power.

In the long term, the jury is still out on which of the two ET families will become the standard, and probably the choice will depend on mission optimization; however, it seems that future nuclear-powered MPD lend themselves to better integration of propellant, thrust and I_{sp} management by using superconductivity. In both families ionization of the propellant will eventually do away with electrodes, that either because of massive current or of very high voltage tend to erode, degrade and break down. Electrode-less ionization by RF, as in fusion tokamak, has been implemented successfully

in the VASIMR thruster, and there is no reason why it should not work also in future ET absorbing tens or even hundreds MW.

As for performance, future nuclear-powered GIE built with tested tokamak refueling technology may accelerate hydrogen to 100 km/s ($I_{sp} = 10^4$ s) producing $F = 10$ N. Ideal power is then 1 MWe [Fearn, 2008]. This number, times a factor 3-4 gives the thermal power of the NR.

Future GIE and MPD with even higher I_{sp} ($\sim 10^5$ s) are potentially feasible with the same technology and will need 10 MWe, and hundreds of MWe for thrust of order 100 N required by fast missions (missions lasting only a few months).

These figures give pause to any ET expert, used to solar-powered ET in the 10 kW range. But power of order 100 MWe is nothing to commercial NR engineers. Thus, IF space agencies worldwide will collaborate, for instance in a M3, such figures may well become commonplace.

5.4 When is NP useful?

The crucial advantage of NP is the power the NR can supply, thus thrust, and I_{sp} . The energy density of NR is orders of magnitude larger than any other power source, in particular of solar panels (SP), see Table 1. A rule of thumb is that above 250 - 300 kWe a NR is lighter and less bulky than SP, but that depends also on mission and solar panel technology. Certainly it is true for missions toward the outer planets, since solar power decreases with $1/r^2$.

The issue of power becomes the key issue when looking at the main objectives of human exploration under current consideration by space agencies. According to the ISECG international organization coordinating future human exploration, they are the Moon, NEO/NEA, and Mars, see [ISECG].

Going to the Moon does not require NP, but building and maintaining an outpost there will definitely need a NR during the long lunar nights and especially to carry on power intensive surface activities, e.g., ISRU, or to dig regoliths to build radiation shelters. In fact, NR in the 1-MWth range are included in most lunar scenarios. Mining will of course need far more power.

For human mission to [still unspecified] NEO/NEA the question of power becomes acute. Distance from the sun will likely rule out *present* technology SP. Restricting this discussion to propulsion, NP power will depend on planned mission time. Due to SR+ GCR dose, crew exposure time must be less than 5-6 months to limit cancer risks, although these numbers must be taken with a grain of salt, waiting for more research on the effects of radiation.

Roughly speaking, these times dictate thrust in the many 100s of N at the very least, to avoid crossing the van Allen belt too many times while escaping. This thrust would be too large for SP-powered clusters of GIE or [future] MPD ET. Note that NTP systems in this thrust range were never built, so NEP is the obvious choice for the trans-NEA portion of a trajectory. To avoid spiraling around the Earth and crossing several times the van Allen belt while accelerating toward escape speed, pure NEP must be complemented by either a chemical rocket, or be replaced by a hybrid NEP-NTR combination. This conclusion is actually tentative: VASIMR progress may indeed result in a ET capable of varying F and I_{sp} inversely at fixed NR power. One may roughly estimate that for F of order $O(100)$ N and a reasonably high I_{sp} , say, $O(10,000)$ s [Coletti, 2011], electric power needs be $O(10)$ MWe. While significant, this power is modest in terms of NR capability. Thrust in this range

enables a roundtrip mission to some asteroids lasting only a few months, with a significant reduction in NR and ET life requirements.

This said, the future of SP includes [initial] efficiencies up to 42% (current high performance cells are at 28%). When these are mature, SEP may, or might, become competitive for certain NEO missions, even those requiring beyond 300 kWe. This tentative conclusion is important in view of the 2025-2030 timeline envisaged by ISECG for future deep space missions.

A M3 is a good deal more challenging. A conceptual pure-NEP-driven M3 mission has been already studied [Ferraro et al, 2009], admittedly using GIE with I_{sp} equal to the highest tested in the lab, $\sim 37,500$ s (this I_{sp} is nine times higher than commercially available now, but the tokamak fusion technology already mentioned will in due time result in I_{sp} in the many 100s of ks [Fearn, 2008]. With these assumptions a NEP-powered Mars transfer was predicted to take between 200 and 100 d with GIE power between 100 and 300 MWe. Total spacecraft mass was between 150 and 100 ton, respectively. A significant fraction of mission time was due to spiraling to reach Earth escape speed.

The crew dose for such fast trip would be quite tolerable with structural areal density (Al, or Al + equipment), 5 to 20 g/cm² as on the ISS [Durante and Bruno, 2010]. The latter showed that NEP is a key enabling technology to lower radiation risk. Presumably, in combination with NTR, or based on VASIMR-type ET, NEP may be even more convenient.

A second example of the capability of NP is associated to the future of ISS. The ISS life has been extended beyond 2020 and perhaps even beyond, to 2028. In this context a concept has been proposed [Lorenzoni and Bruno, 2011] fitting the ISTAR vision illustrated at the recent ISS Mars Conference in Washington by NASA's C. Bolden and supported by US industry. This vision has also been endorsed by Russia [Savinyk, 2011]. In the ISTAR scenario the now-complete ISS is used not only for scientific purposes, but becomes incrementally a testbed of technologies indispensable to deep space exploration. ISTAR includes eventually installing propulsion on the ISS, that (whole or in part) would be moved from LEO to some location making it a deep space habitat (DSH). A possible location is L1. The reasoning behind is that a DSH is needed as a stepping stone toward a NEA and/or future Mars mission, but, in the current financial scenario, there is no money to build one. The space community will make do with whatever space assets already exist, and that is the ISS. The testing of the 200 kWe VX-200 VASIMR thruster [AdAstra, 2011], under discussion between Ad Astra Company and NASA may be coincidental, or may be consistent with the ISTAR vision.

At the ISS Mars Conference mentioned [Lorenzoni and Bruno, 2011] compared the ISS transfer from LEO to L1 using SEP to that using NEP. The ISS mass was assumed 190 ton, and a $\Delta V = 6$ km/s accounted for the spiral transfer using ET. SP power was assumed limited to 110 KWe. A GIE and a Hall thruster, with $I_{sp} = 10,000$ s and 3200 s, respectively, were the ET chosen for this comparison. (The I_{sp} of the GIE is that predicted feasible in the '20s by the U. of Southampton, UK.) Note that the ISS inertia moments are huge, limiting maximum transfer acceleration and thrust. Thus expected transfer times from LEO to L1 with ET are significant. A simplified 0-th order solution indicated ~ 10.6 yr the transfer and 9.9 ton of Xe propellant with GIE, and ~ 6.4 yr and 33.7 ton with Hall thruster. These numbers may improve with future (3-junction) solar cells, that at Spectralab have already shown 42% initial efficiency.

For comparison, with a modest 1 MWe NEP system using the same GIE, but powered by a 1 to 2 ton NR, the transfer time becomes 620 d, and the propellant consumed is about 9.8 ton,

a factor 4 better in terms of transfer time, and 3.3 in propellant consumption in the worst case. Although the fraction of time spent within the van Allen belt is non-negligible, the comparison is noteworthy, even not accounting for the difference in orbiting cost due to the much bulkier solar panels when compared to a compact NR system.

Finally, the future may see a new class of interplanetary missions combining robotics with human activity. This is asteroid capturing. Far from being science fiction, this concept has been proposed originally in [Tantardini, 2010] and is now being preliminarily explored at JPL by John Brophy and others. The purpose of such mission is manifold, i.e., scientific, but also to train astronauts near/on very low gravity bodies, for instance to dock, and to fasten objects by drilling or by other means. (this is a crucial task to deviate asteroids from their orbit, as in planetary defence).

There is a dearth of small asteroids passing close to Earth, , e.g., 2011 MD went by on June 27, 2011 ~ 7600 mi. above the Atlantic Ocean [Chang, 2011]. Moving asteroids in the 20 to 30 m size (meteoroids) from their orbit to L1, or even close to the ISS, is technically feasible in the near future. For instance, using current SEP technology (Isp = 3000 s, 150 kWe panels feeding thrusters capable of 50 mN/kWe), the thrust available, $F = 7.5 \text{ N}$, could be applied to the 2000 LG6 meteoroid. Its size (5m) and mass (200 ton) can be handled relatively well: its acceleration would be $3.75 \times 10^{-5} \text{ m/s}^2$, its ΔV , after a year of tugging, about 1.2 km/s. With solar power this mission requires 8 ton/yr of propellant and about 500 m^2 of SP. With the same [modest] NEP system of the previous example the ΔV time could be halved.

Note that for this class of missions a NTR would be much more convenient (thrust and acceleration would be many orders of magnitude larger), and given the modest ΔV required, the larger propellant consumption would probably not matter much. However, no known NTR was ever designed with power less than many *tens* of MW.

6. Conclusions

NP is an outstanding, if not *the* outstanding propulsion candidate for deep space human missions to Mars and possibly also to NEA, although future high efficiency solar panels might become a good competitor. Because of its inherent power density NP is capable of reducing drastically travel time and thus crew radiation dose, with its known and newly emerging health risks. While there are still engineering issues to be solved to apply NEP, NTP, or their combination to future human missions, its capability and potential outclass any CP performance, enabling faster travel with drastically reduced mass consumption. This last is a very important factor in planning financially affordable human exploration.

7. References

- Ad Astra (2011), see www.adastrarocket.com/aarc/AboutUs for the VASIMR MPD ET description.
- Accettura, A., Bruno, C., and Del Rossi, A., (2008, "The Case for Nuclear Propulsion: the Rubbia's Engine", in: *Advanced Propulsion Systems and Technologies, Today to 2020*, ed. by C. Bruno and A. Accettura, AIAA, Reston, VA, Chapter 13.
- Andrews, D., and Lenard, R.X., (2006), "Proposed Follow-on Mini-MagOrion Pulsed Propulsion Concept", AIAA Paper 2006-5223, presented at the 42nd AIAA/ASME/SAE/ASEE Joint Propulsion Conference, 9-12 July 2006

- Bussard, R.W., and DeLauer, R.D., (1958), "Nuclear Rocket Propulsion", McGraw-Hill, New York.
- Cassady, R.J., et al (9 co-Authors), (2008), "Recent advances in nuclear powered electric propulsion for space exploration", Energy Conversion Management, Vol. 49, pp. 412-435.
- Chang, A., (2011), "Small asteroid swings harmlessly past Earth", AP press release, see http://old.news.yahoo.com/s/ap/20110627/ap_on_sc/us_sci_asteroid_pass
- Coletti, M., and Gabriel, S.B., (2010), "Design of a dual stage ion engine for the Hiper project" paper presented at the AIAA/ASME/SAE/ASEE 46th Joint Propulsion Conference, Nashville, 25-28 July, 2010.
- Covault, C., (2011), "AMS: Shedding light on the dark", Aerospace America, Vol. 49, No. 6, June 2011, pp. 40-45.
- Czysz, P.A., and Bruno, C. (2009), "Future Spacecraft Propulsion Systems", Springer-Praxis, London.
- Dewar, J.A., (2004), "To the End of the Solar System: The Story of the Nuclear Rocket", The University Press of Kentucky, Lexington, KY.
- Durante, M., and Bruno, C., (2010), "Impact of rocket propulsion technology on the radiation risk in missions to Mars", Eur. Phys. J. D, published online, 16 February 2010, DOI: 10.1140/epjd/e2010-00035-6.
- Dyson, G., (2002), "Project Orion -The True Story of the Atomic Spaceship", H. Holt and Co., New York.
- ESRIN (2011), Workshop on "Future Human Spaceflight and Exploration Scenarios", Frascati, Italy, May 2-3, 2011. See www.msmftp.estec.esa.int (need to ask ESA for id and password).
- Ewig, R, and Andrews, D., (2003), "Mini-Mag Orion: A Pulsed Nuclear Rocket for Crewed Solar System Exploration", AIAA Paper 2003-4525, presented at the 39th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Huntsville, AL.
- Fearn, D.G., (2008), "Application of Ion Thrusters to High-Thrust, High-Specific-Impulse Nuclear Electric Missions", in: Nuclear Space Power and Propulsion Systems, ed. by C. Bruno, AIAA, Reston, VA, Chapter 3.
- Ferraro, F., D'Elia, R., Paternostro, S., Simonetti, A., and Bruno, C., (2009), "Analysis of a Manned Mars Mission with Nuclear Electric Propulsion (NEP) System", IAC Paper IAC-09.C4.7.-C3.5.2, presented at the 60th IAC, Daejeon, Korea, 12-16 October 2009.
- Goebel, D.M. and Katz, I., (2008), "Fundamentals of Electric Propulsion: Ion and Hall Thrusters" J. Wiley and Sons, New York, 2008.
- J. Wiley, Hoboken, NJ. Gunn, S.V., and Ehresman, C.M., (2003), "The Space Propulsion Technology Base Established Four Decades Ago for the Thermal Nuclear Rocket is ready for Current Applications", AIAA Paper 2003-4590, presented at the 39th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, 20-23 July 2003, Huntsville, AL.
- Howe, S.D., DeVolder,, B., Thode, L., and Zerkle, D., (1998), "Reducing the Risk to Mars: the Gas Core Nuclear Rocket", in: STAIF-1998, Ed. by Mohamed El-Genk, The Am. Inst. Of Physics, Publication CP-420, p. 1138.

- ISECG: see www.globalspaceexploration.org; www.nasa.gov/pdf/178109mainges_framework.pdf]
- Karash, Y., (2011), see in www.bernama.commy/bernama/v5/newsworld.php?id=575134 his statement on the MARPOST mission to Mars.
- Koroteev, A.S., Prishletsov, A.B., Martishin, V.M., Pavelyev, A.A., Shcherbinin, V.P., Reshmin, A.I., and Iosilevskii, I.L., (2007), "Rocket Engines and Powerplants Based on Gas-core Nuclear Reactor, ed. by A-S. Koroteev, Machinostroenie Publisher, Moscow (in Russian).
- Koroteev, A.S., Akimov, V.N., and Gafarov, A.A., (2007), "Creation and Perspectives of Application of Space Nuclear Energy in Russia"(in Russian), *Polyot (Flight)*, Vol. 7, pp. 3-15.
- Lawrence, T.J., (2008), "Nuclear thermal rocket propulsion systems", in: C.Bruno, ed., *Nuclear Space Power and Propulsion Systems*, AIAA, Reston, VA, Chapter 2.
- Lawrence, T.J., Witter, J.K., and Humble, R.W., (1995), "Nuclear Rocket Propulsion Systems", in: *Space Propulsion Analysis and Design*, ed., by R.W. Humble, G.N. Henry and W.J. Larson, McGraw-Hill, New York, Chapter 8.
- Lorenzoni, A., and Bruno, C., (2011), "Upgraded ISS to L1", presentation given at the ISS and Mars Conference, see <http://exploremars.org/iss/>, George Washington University, Washington, DC, April 6-7, 2011.
- Martinez-Sanchez, M. and Pollard, J.E.,(1998), "Spacecraft Electric Propulsion –An Overview", *J. Prop. and Power*, Vol. 14, No. 5, September –October 1998, pp. 688-699.
- Parker, E.N., (2006), "Shielding Space Travelers", *Sci. Am.*, Vol. 294, No. 9, March 2006, pp. 22-29.
- Powell, J., Paniagua, J., Ludewig, H., and Maise, G., (1998), "MITEE: A New Nuclear Engine Concept for Ultra Fast, Lightweight Solar System Exploration Missions", in: STAIF-1998, The Am. Inst. of Physics, Publication CP-420, New York, p. 1131.
- Powell, J., Paniagua, J., Ludewig, H., and Todosow, M., (1999), "The MITEE Family of Compact, Ultra Lightweight Nuclear Thermal Propulsion Engines for Planetary Exploration Missions", Paper IAF-99-S.6.03, presented at the 50th Int. Astron. Congress, 4-8 Oct. 1999, Amsterdam, The Netherlands.
- Powell, J., Maise, G., and Paniagua, J., (2003), "HIP: a Hybrid NTP/NEP Propulsion System for Ultra Fast Robotic/Lander Missions to the Outer Solar System", IAC Paper presented at the 54th IAC, Bremen, Sept. 29 to Oct. 3, 2003.
- Ronen, Y., Aboudy, M., and Regev, D., (2000), "A Nuclear Engine with ^{242m}Am as a Nuclear Fuel", *Ann. Nucl. Energy*, Vol. 27, p. 85.
- Rossetti, P., Saverdi, M., and Biagioni, L., (2008), *Electric-Propulsion Systems*", in: *Advanced Propulsion Systems and Technologies, Today to 2020*, ed. by C.Bruno and A.Accettura, AIAA, Reston, VA, Chapter 10.
- Saenger, E., (1958), "Die Erreichbarkeit der Fixsterne", in: *Mitteilungen der Landesgruppe Nordbayern der DGRR vom 13.05.1958*.
- Savinyk, V., (2011), see www.bernama.commy/bernama/v5/newsworld.php?id=575134 for the statement by Russian astronaut Viktor Savinykh.
- Shepherd, L.R., and Cleaver, A.V., (1948, 1949), "The Atomic Rocket-1,2,3 and 4", *J. British Interplanet. Soc.*, Vol. 7, no.5 and No. 6, 1948; Vol. 8, No. 1 and No. 2, 1949.

Stuhlinger, E., (1964), "Ion propulsion for space flight", MacGraw-Hill, New York, Chapter 4.

Tantardini, M., (2011), personal communication and presentation made at JPL.

Thompson, P.A., (1972), "Compressible Fluid-Dynamics", McGraw-Hill, New York, Section 8.12

Turner, M.J.L., (2005), "Rocket and Spacecraft Propulsion", Springer-Praxis London Chapter 7.

Fluid Dynamics in Microchannels

Jyh-tong Teng et al.*

*Department of Mechanical Engineering
Chung Yuan Christian University, Chung-Li
Taiwan*

1. Introduction

1.1 Need for microchannels research

In contrast to external flow, the internal flow is one for which the fluid is confined by a surface. Hence the boundary layer develops and eventually fills the channel. The internal flow configuration represents a convenient geometry for heating and cooling fluids used in chemical processing, environmental control, and energy conversion technologies [1].

In the last few decades, owing to the rapid developments in micro-electronics and biotechnologies, the applied research in micro-coolers, micro-biochips, micro-reactors, and micro-fuel cells have been expanding at a tremendous pace. Among these micro-fluidic systems, microchannels have been identified to be one of the essential elements to transport fluid within a miniature area. In addition to connecting different chemical chambers, microchannels are also used for reactant delivery, physical particle separation, fluidic control, chemical mixing, and computer chips cooling.

Generally speaking, the designs and the process controls of Micro-Electro-Mechanical-Systems (MEMS) and micro-fluidic systems involved the impact of geometrical configurations on the temperature, pressure, and velocity distributions of the fluid on the micrometer (10^{-6} m) scale (Table 1.1). Therefore, in order to fabricate such micro devices effectively, it is extremely important to understand the fundamental mechanisms involved in fluid flow and heat transfer characteristics in microchannels since their behavior affects the transport phenomena for the bulk of MEMS and micro-fluidic applications.

Overall, the published studies based on an extensive literature reviews include a variety of fluid types, microchannel cross-section configurations, flow rates, analytical techniques, and channel materials. The issues and related areas associated with the microchannels are summarized in the following table (Table 1.2).

*Jiann-Cherng Chu¹, Chao Liu², Tingting Xu², Yih-Fu Lien¹, Jin-Hung Cheng¹, Suyi Huang², Shiping Jin², Thanhtrung Dang³, Chunping Zhang², Xiangfei Yu², Ming-Tsang Lee⁴, and Ralph Greif⁵

¹*Department of Mechanical Engineering, Chung Yuan Christian University, Chung-Li, Taiwan*

²*School of Energy and Power Engineering, Huazhong University of Science and Technology, Wuhan, China*

³*Department of Heat and Refrigeration Technology, Hochiminh City University of Technical Education, Hochiminh City, Vietnam*

⁴*Department of Mechanical Engineering, National Chung Hsing University, Taichung, Taiwan*

⁵*Department of Mechanical Engineering, University of California at Berkeley, CA, USA*

Definition	The range of channel dimension
Conventional channels	$D_c > 3 \text{ mm}$
Minichannels	$3 \text{ mm} \geq D_c > 200 \text{ }\mu\text{m}$
Microchannels	$200 \text{ }\mu\text{m} \geq D_c > 10 \text{ }\mu\text{m}$
Transitional Microchannels	$10 \text{ }\mu\text{m} \geq D_c > 1 \text{ }\mu\text{m}$
Transitional Nanochannels	$1 \text{ }\mu\text{m} \geq D_c > 0.1 \text{ }\mu\text{m}$
Nanochannels	$D_c \leq 0.1 \text{ }\mu\text{m}$

Table 1.1. Channel classification schemes [2]

General research areas	Related studies
1. Cross-sections of microchannels <i>-Triangle, Rectangle, Trapezoid, Circle, Square and Non-uniform.</i>	1. Size effect (Hydraulic diameter) 2. Aspect ratio 3. Entrance effect
2. Materials of microchannels <i>- Silicon, Nickel, Polycarbonate, Polyamide, Fused silica, Stainless steel, Copper, Aluminum, Brass, Glass, Oxidized silicon, SiO₂, Polyvinylchloride, Poly-dimethylsiloxane(PDMS), poly-methyl methacrylate (PMMA).</i>	1. Surface roughness 2. Contact angle 3. Hydrophilic and hydrophobic property 4. Electrical double layer (EDL) 5. Thermo-physical properties
3. Types of flows <i>- N₂, H₂, Ar, Water, R-134a, Methanol, Iso-propanol, Aqueous KCl, Fluorinert fluid FC-84, Vertrel XF, Air, Helium, Silicon oil.</i>	1. Polarity 2. Rarefaction effect 3. Compressibility 4. Temperature jump 5. Non-slip and Slip 6. Joule heating
4. Flow rates <i>- Reynolds number, Mach number</i>	1. Critical Reynolds number of the transition to turbulence 2. Viscous heating or Viscous dissipation 3. Mal-distribution
5. Analytical techniques for microchannel flows <i>- Numerical simulation</i> <i>- Experimental analysis</i>	1. Flow pattern and visualization 2. Velocity field 3. Temperature distribution 4. Friction factor 5. Nusselt number 6. Poiseuille number

Table 1.2. Summary of research areas and discussed issues related to fluid flow in microchannel [3]

1.2 Research methodology

As the field of micro-fluidic systems continues to grow, it is becoming increasingly important to understand the mechanisms and fundamental differences involved in micro-scale fluid flow.

To study the thermal and hydrodynamic characteristics of fluid flow in microchannels, this work used experimental measure and numerical simulation to investigate the behavior of flow and temperature fields in microchannels. In this chapter, each part of the study will be described separately as follows.

1.2.1 Experimental work

To carry out the experiments of the flow in microchannels, first and foremost, a fluid flowing and measurement system, together with microchannel structures should be properly designed and built up. In this study, the details of experimental procedure involving the manufacturing of test chip, construction of experimental system, and analysis of experimental uncertainties will be described in the following sections.

In this chapter, an experimental flow chart related to the experimental procedure is shown in Fig. 1.1.

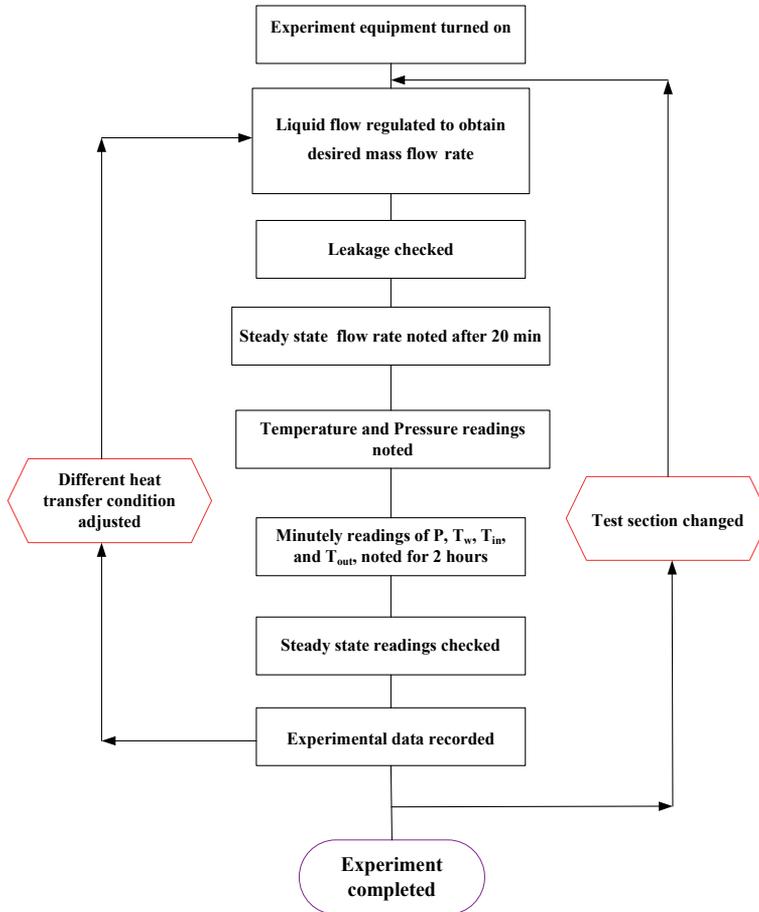


Fig. 1.1. Experimental flow chart [3].

1.2.2 Experimental system

The experimental system was divided into three parts – the test section, the water driving system, and the dynamic data acquisition section. In addition, DI water (deionized water) was used as the working fluid. The experimentally-measured data were composed of pressure drop through the microchannel, temperatures of DI water and substrate, and the mass flow rate. Each part of the experimental system will be described separately as follows:

1. Test section

Since measuring the local pressure and temperature along the flow path was difficult inside the microchannels, two sumps were machined in the PMMA block and were connected by microchannels with holes made by laser processing. A diaphragm type differential pressure transducer with ranges of 0-35 bars was connected to the sumps to measure the corresponding pressure drop across the inlet and outlet of the microchannel. Concerns had been raised that this kind of fitting could cause some dead volume resulting in the detection of false signals. However, to minimize the dead volume that might appear in the flow channel, extreme caution had been taken to ensure that no visible dead volume was observed there.

2. Water-driven system

Precision-controlled fluidic HPLC pump or injection pump was used to transport the DI water through the test section, and the flow rate was controlled in a range from 0.1 to 40 ml/min for the DI water flow. A filter (with a screening size of 0.2 μm) was installed midway to remove any possible particulates and contaminants that might be present in water under testing. For the case with heat generation, the stainless steel plate was electrically heated by directly connecting the bottom of test sections to a DC power supply that provided low voltage and high electric current. Once the working fluid was flowing into the test section, measurements of the pressure drop and temperature were done, followed by the weight measurement of the collected fluid using a precision electronic balance with an accuracy of 0.0001 g to obtain the mass flow rate of the system. In the meantime, the temperature of working fluid in the microchannel was measured by calibrated T-type Cu-Ni thermocouples to determine accurately the values of the water density and viscosity.

3. Dynamic data acquisition section

This section used an experimental platform with the test section laying horizontally on the platform, above which a digital microscope hooked with video capture card and signal cable to send the digital image to the PC for data-processing, ensuring that the working fluid passed through microchannel and that no left-over bubbles or impurities existed to block the flow channel that could lead to erroneous signals being obtained for the tests. In the meantime, a notebook PC and a network system were used to transmit data at a speed of one per 500 milliseconds. The data acquisition system for recording the electronic signals was implemented to obtain data from the differential pressure transducer and T-type thermocouples; the system was integrated through the instant monitoring software to record and analyze the data received. For each data point being measured, the flow was considered to be at a steady-state condition when the measured pressure drop and temperature remained unchanged for at least 10 minutes. Each case was repeated for at least three times to make sure that the arrangement could always produce reliable and reproducible results.

1.2.3 Numerical simulation

Numerical simulations were done by the CFD-ACE+ software [4] which provided an integrated numerical analysis of the continuity, momentum, and energy equations for the fluid flow and heat transfer. After specifying the boundary condition, CFD-ACE+ uses an iterative, segregated solution method where in the equation sets for variables such as pressure, velocity, and temperature are solved sequentially and repeatedly until a converged solution is obtained. In general, iterative equation solvers are preferred for this task because they are more economical in memory requirements than direct solvers. In CFD-ACE+ program, conjugate gradient solvers and algebraic multi-grid solver are provided to obtain the converged solution. In Chu et al. [3], the latter method was adopted. The basic idea of a multi-grid solution is to use a hierarchy of grids, from fine to coarse, to solve a set of equations, with each grid being particularly effective for removing errors of the wavelength characteristic of the mesh spacing on that grid.

In CFD-ACE+, the finite-volume approach is adopted due to its capability of conserving solution quantities. The solution domain is divided into a number of cells known as control volumes. In the finite volume approach of CFD-ACE+, the governing equations are made discrete and finite, and then numerically integrated over each of these computational cells or control volumes. An example of such a control volume is shown in Fig.1.2 [5].

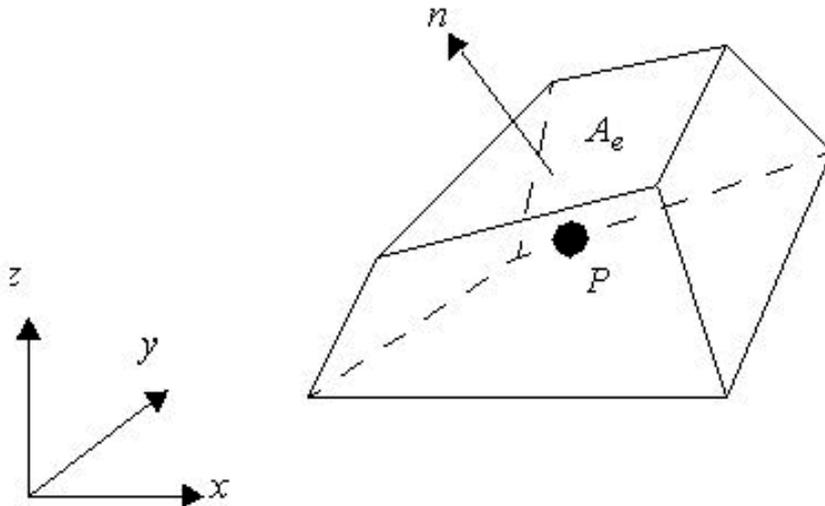


Fig. 1.2. 3D Computational Cell (Control Volume).

The geometric center of the control volume, which is denoted by P , is also often referred to as the cell center. CFD-ACE+ employs a co-located cell-centered variable arrangement, i. e., all dependent variables and material properties are stored at the cell center P . In other words, the average value of any quantity within a control volume is given by its value at the cell center. Most of the governing equations can be expressed in the form of a generalized transport equation as shown in Eq. (1-1), which is also known as the generic conservation equation for a quantity Φ [5].

$$\underbrace{\frac{\partial(\rho\phi)}{\partial t}}_{\text{transient}} + \underbrace{\nabla \cdot (\rho\bar{V}\phi)}_{\text{convection}} = \underbrace{\nabla \cdot (\Gamma\nabla\phi)}_{\text{diffusion}} + \underbrace{S_\phi}_{\text{source}} \quad (1-1)$$

where S is the source term. The overall solution procedure in flowchart form for the solution algorithm is shown in Fig. 1.3. The number of iterations (NITER) can be defined to dictate how many times a procedure is repeated [6].

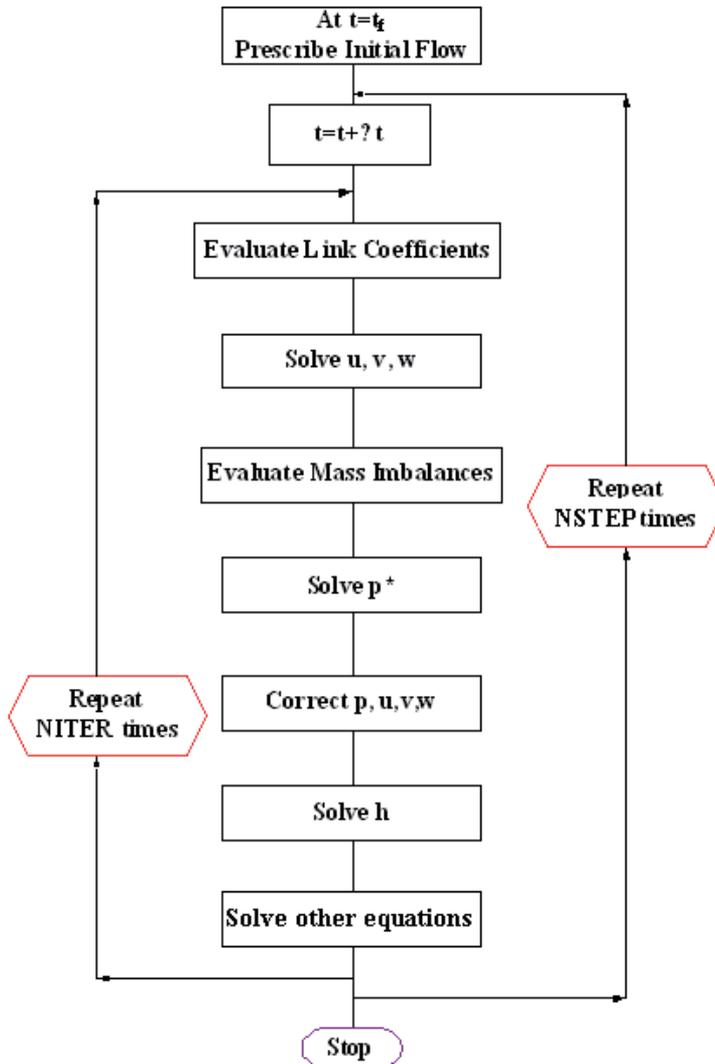


Fig. 1.3. Solution Flowchart.

2. Fundamental theory about flow motion

2.1 Pressure drop in single liquid-phase flow

2.1.1 Basic pressure drop correlations

For a non-circular cross section of the flow channels, the calculated hydraulic diameter D_h of a rectangular channel is computed by the following equation:

$$D_h = 4A_{ch}/P_w = 2WH/(W+H) \quad (2-1)$$

where A_{ch} , P_w , D_h , W and H represent as the areas of microchannel, wetted perimeter, hydraulic diameter, microchannel width and height, respectively.

The Reynolds number Re is defined as:

$$Re = \rho_m u_m D_h / \mu_m \quad (2-2)$$

where μ_m and ρ_m are mean dynamic viscosity and mean density of fluid at an arithmetic mean temperature ($T_m = (T_o + T_i)/2$), respectively. It should be noted that the fluid properties are functions of the temperature and values are obtained from correlations for dynamic viscosity (μ) correlations for dynamic viscosity (μ), thermal conductivity (k), specific heat (c_p), and density (ρ) of DI water.

Under actual conditions, the measured pressure drop includes the effect of the losses (1) in the bends and (2) at the entrance and exit, together with the frictional pressure drop in the microchannel. Phillips [7] suggested that the measured pressure drop was the sum of these components.

$$\Delta P_t = \frac{\rho u_m^2}{2} \left(K_c + K_e + 2(A_{ch}/A_p)^2 K_{90} + \frac{4f_{app}L}{D_h} \right) \quad (2-3)$$

where ΔP_t , A_p , K_c , K_e , K_{90} , f_{app} and L are measured pressure drop, plenum area, contraction loss coefficient, expansion loss coefficient, bends loss coefficient and apparent friction factor, respectively. The loss coefficient K_{90} was recommended by Phillips [7]. K_c and K_e can be obtained from Kays and London [8]. According to the published investigations with regard to microchannels [9-12], these values of loss coefficient are usually obtained from the traditional relationships in macro-scale flow.

In addition, the method described above for determining minor losses was supported by the data obtained by Abdelall et al. [13], which showed that the experimentally measured loss coefficients associated with single-phase flow in abrupt area changes in microchannels were comparable to those obtained for large channels with the same area ratios.

2.1.2 Pressure drop in fully-developed laminar flow

For hydrodynamically fully developed flow, the velocity gradient at the channel wall can be readily calculated from the well-known Hagen-Poiseuille parabolic velocity profile for the fully developed laminar flow in a pipe. The Fanning friction factor f_c is expressed in the following form:

$$f_c = P_o / Re \quad (2-4)$$

where the Poiseuille number P_o is defined as $P_o = fRe$, the product of the friction factor and the Reynolds number.

For incompressible flow through horizontal channels of constant cross-sectional area, f_c can be calculated by Eq. (2-5), based on the mass flow rate and the pressure drop ΔP where the latter is due to the friction occurred inside the rectangular microchannel.

$$f_c = \frac{2\tau_w}{\rho_m u_m^2} = \frac{D_h \Delta P}{2\rho_m u_m^2 L} \quad (2-5)$$

where τ_w is the wall shear stress, L is the channel length, and u_m is the mean flow velocity.

A simple equation proposed by Shah and London [14] for fully developed, incompressible and laminar flow in a rectangular channel was used to predict the friction factor of straight rectangular microchannels. This equation, which has been used and proven to be adequate for predicting liquid flows in rectangular microchannels by several researchers [15-17], is expressed as follows:

$$f_c = 24(1 - 1.3553\alpha_c + 1.9467\alpha_c^2 - 1.7012\alpha_c^3 + 0.9564\alpha_c^4 - 0.2537\alpha_c^5) / Re \quad (2-6)$$

Here f_c is the Fanning friction factor for a straight channel and α_c is the aspect ratio, which is the ratio of the dimension for the short side to that of the long side.

2.1.3 Pressure drop in developing laminar flow

The hydrodynamic entry length L_h for rectangular microchannels is given as follows [2].

$$L_h / D_h = 0.05Re \quad (2-7)$$

The apparent friction factor (f_{app}) includes the combined effects of frictional losses (pressure losses in developed region) and the additional losses in the developing region. The difference between the apparent friction factor (f_{app}) over a length x , measured from the entry location, and the fully developed Fanning friction factor (f_c) is expressed in terms of an incremental pressure defect $K(x)$ as follows [3]:

$$K(x) = (f_{app} - f_c)(4x) / D_h \quad (2-8)$$

where $K(x)$ is the Hagenbach factor in the above equation.

2.1.4 Fully developed and developing turbulent flow

Regarding fully developed turbulent flow in smooth microchannels, a number of correlations with comparable accuracies are available in the literatures. Blasius put forward the following equation which is used extensively nowadays.

$$f = 0.0791Re^{-0.25} \quad (2-9)$$

Covering both the developing and developed flow regions, Phillips [18] used a more accurate equation for Fanning friction factor in a circular tube.

$$f = ARe^B \quad (2-10)$$

where $A = 0.09290 + \frac{1.01612}{x / D_h}$ and $B = -0.26800 - \frac{0.32930}{x / D_h}$

For rectangular microchannels, Re is replaced with the laminar-equivalent Reynolds number given by

$$Re^* = \frac{\rho u_m D_{le}}{\mu} = \frac{\rho u_m [(2/3) + (11/24)(1/\alpha_c)(2 - 1/\alpha_c)] D_{le}}{\mu} \quad (2-11)$$

where D_{le} is the laminar-equivalent diameter calculated by the term in the brackets.

2.1.5 Laminar-to-turbulent transition

The laminar-to-turbulent flow transition is another topic investigated by lots of researchers. In the entrance region of rectangular tubes with abrupt area change, the laminar-to-turbulent transition was reported to take place at a transition Reynolds number of $Re_t = 2200$ for $\alpha_c = 1$ and at $Re_t = 2500$ for parallel plates with $\alpha_c = 0$. For the other aspect ratios, a linear interpolation is recommended.

Some of the initial studies presented an early transition to turbulent flow in microchannels. However, several recent studies stated that the laminar-to-turbulent transition remained unchanged. For circular microtubes with diameter 171-520 μm , Bucci et al. [19] pointed that the transition occurred around $Re_t = 2000$. The result of Baviere et al. [20] also indicated that the dimensions of smooth microchannels didn't affect the laminar-to-turbulent transition and the critical Reynolds number was still around 2300. It can be supported by a number of investigators, such as Schmitt and Kandlikar [21] for minichannels with $D_h < 1$ mm, and Li et al. [22] for $80 \mu\text{m} \leq D_h \leq 166.3 \mu\text{m}$.

2.1.6 Pressure drop related to the change liquid properties

Eq. (2-6) provides the theoretical value of Fanning friction factor in rectangular microchannels at constant thermal properties of liquid suggested by Shah and London [23]. In the present experiments, the effect of liquid property variations cannot be neglected for a large temperature difference between inlet and outlet. Kays and London [8] suggested a corrected correlation for temperature dependent properties.

$$f / f_m = (\mu_w / \mu_m)^M \quad (2-12)$$

where $M = 0.58$ for liquid heating, and $M = 0.5$ for liquid cooling; subscripts m and w are for the condition at the arithmetic mean fluid temperature and the condition at the wall temperature, respectively.

Then, the corrected form for the Shah and London correlation [23] according to the present study is

$$f'_c = f_c (\mu_w / \mu_m)^M \quad (2-13)$$

2.2 Basic heat transfer correlations in single liquid-phase flow

Results for the total heat transfer rate and the axial distribution of the mean temperature are derived as follows for the constant surface temperature condition (taking three heated walls in a channel for example, as shown in Fig. 2.1). Defining ΔT as $T_s - T_m$, the equation may be expressed as

$$\frac{dT_m}{dx} = -\frac{d(\Delta T)}{dx} = \frac{P_w}{\dot{m}c_p} h \Delta T \quad (2-14)$$

Separating variable and integrating from the tube inlet to the outlet yield

$$\int_{\Delta T_i}^{\Delta T_o} \frac{d(\Delta T)}{\Delta T} = -\int_0^L \frac{P_w}{\dot{m}c_p} h dx \quad (2-15)$$

or,

$$\ln\left(\frac{\Delta T_o}{\Delta T_i}\right) = -\frac{P_w L}{\dot{m} c_p} \left(\int_0^L \frac{1}{L} dx \right) \quad (2-16)$$

From the definition of the average convection heat transfer coefficient, it follows that

$$\ln\left(\frac{\Delta T_o}{\Delta T_i}\right) = -\frac{P_w L}{\dot{m} c_p} \bar{h}_L = -\frac{A_s}{\dot{m} c_p} \bar{h}_L, \quad T_s = \text{constant} \quad (2-17)$$

where \bar{h}_L , or simply \bar{h} , is the average value of h for the entire channel, A_s is the heat exchange area between the working fluid and wall surface inside the channel. Rearranging,

$$\frac{\Delta T_o}{\Delta T_i} = \frac{T_s - T_{m,o}}{T_s - T_{m,i}} = \exp\left(-\frac{A_s}{\dot{m} c_p} \bar{h}\right), \quad T_s = \text{constant} \quad (2-18)$$

For a general form of Eq. (2-18), one can obtain

$$\frac{T_s - T_m(x)}{T_s - T_{m,i}} = \exp\left(-\frac{P_w x}{\dot{m} c_p} \bar{h}\right), \quad T_s = \text{constant} \quad (2-19)$$

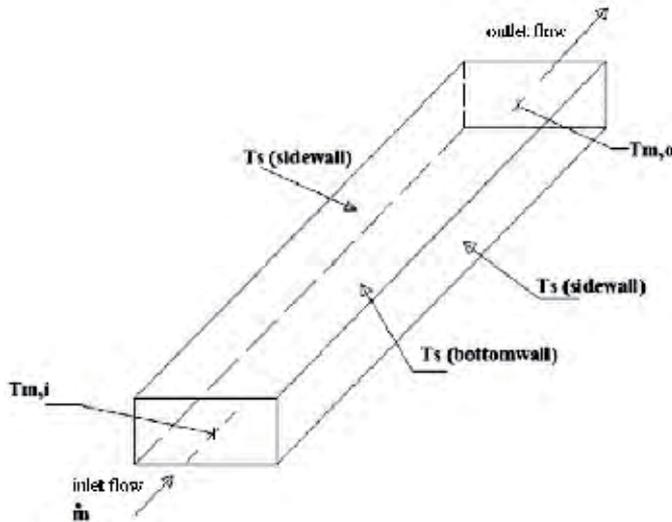


Fig. 2.1. Flow through a rectangular microchannel.

Since, by definition, Nusselt number equals to $h \cdot D_h / \lambda$, the average value of Nu for the entire channel can be expressed as

$$Nu = \frac{-D_h}{\lambda_m} \ln\left(\frac{\Delta T_o}{\Delta T_i}\right) \frac{\dot{m} c_p}{A_s}, \quad T_s = \text{constant} \quad (2-20)$$

where λ_m and c_p are the mean thermal conductivity and heat capacity at constant pressure of DI water at an arithmetic mean temperature, respectively; A_s is the heat exchange areas between the walls and fluid.

For the heat exchange at the constant heat flux, the deduction of correlations can be found in [1]. Moreover, Lee et al. [24] examined the validity of conventional correlations and numerical analysis approaches in predicting the heat transfer behavior in microchannels for correctly matched inlet and boundary conditions.

3. Flow and heat transfer in microchannels of various configurations

3.1 Flow and heat transfer in V-shaped microchannels

A number of experimental and numerical investigations of single phase flow in microchannel have been extensively performed. However, most studies have experimental results obtained for microchannels with rectangular and circular cross-sections. The studies of thermo-fluidic characteristics in microchannels with a V-shaped cross-section are quite limited in this field. Thus, it is necessary to provide the present knowledge of V-shaped microchannel.

Flockhart and Dhariwal [25] measured the flow characteristics of distilled water inside trapezoidal microchannels with hydraulic diameters ranging from 50 to 120 μm and the Reynolds numbers below 600. They found that the flow characteristics of water in trapezoidal microchannels could be predicted by the numerical analysis based on the conventional theory.

Qu et al. [26, 27] used experimental and theoretical methods for the study of the thermo-fluidic characteristics of the trapezoidal microchannels with hydraulic diameters ranging from 62 to 169 μm . Results of their study indicated that the wall roughness of the microchannels might lead to lower Nusselt numbers as comparing with those obtained from the theoretical predictions. In addition, they also found that the experimentally determined friction factors in the trapezoidal microchannels were higher than those obtained from the conventional theory. They used a roughness-viscosity model developed by Mala and Li [16] to interpret the difference of friction factors obtained from the experimental data and those obtained from the conventional theory.

Hetsroni et al. [28] performed an experimental investigation of a microchannel heat sink for cooling of electronic devices. The heat sink had parallel triangular microchannels with a base of 250 μm . The results indicated that the temperature distribution at the exit of the triangular microchannel had a nonlinear distribution, and the instabilities caused fluctuations in the pressure drop and decrease in the heat transfer coefficient.

Wu and Cheng [29] conducted a series of experiments to measure the friction factor and convective heat transfer coefficient in trapezoidal silicon microchannels with different surface conditions. The results indicated that the geometric parameters had significant effect on the Nusselt number and Poiseuille number of trapezoidal microchannels, and the hydrophilic property at the surface of the microchannel enhanced the heat transfer capability of the trapezoidal silicon microchannels.

Wu and Cheng [30] experimentally studied the laminar flow of de-ionized water in smooth silicon microchannels of trapezoidal cross-sections with hydraulic diameters ranging from 25.9 to 291 μm . The measured results indicated that the Navier-Stokes equations were still valid for the laminar flow in the trapezoidal microchannel having a hydraulic diameter as small as 25.9 μm .

Tiselj et al. [31] performed the experimental and numerical analyses to evaluate the effect of axial heat flux on heat transfer in triangular microchannels with hydraulic diameter of 160 μm in the range of Reynolds numbers from 3.2 to 64. The experimental results revealed that

the temperature distribution of the triangular microchannels on the heated wall was in agreement with the numerical predictions obtained from conventional Navier-Stokes and energy equations.

Tiselj et al. [32] used three-dimensional numerical simulations for the study of the heat transfer characteristics of the fluid flowing through triangular microchannels. Their results indicated that a singular point existed near the exit of the channel. In addition, for the flow with higher Reynolds number, the singular point was closer to the exit of the channel.

Based on the conventional theory, Morin [33] developed a model to predict the viscous dissipation effects in a microchannel with an axially unchanging cross section. The microchannel geometries having rectangular, trapezoidal, and double trapezoidal were discussed in his work. The water and isopropanol were used as working fluids. The analytical results demonstrated that for different fluids the effect of viscous dissipation could play different roles and that the effect of viscous dissipation could become very significant for liquid flows when the hydraulic diameter was less than 100 μm . In addition, the rising temperature in an adiabatic microchannel could be expressed as a function of the Eckert number (defined as $Ec = u^2/c_v\Delta T$), Reynolds number, and Poiseuille number.

3.1.1 Model description

For the manufacture of V-shaped microchannels, photolithographic processes are particularly utilized for silicon wafers, and these processes initiated in the electronic field are well developed. When a photolithography-based process is employed, the microchannels having a cross-section fixed by the orientation of the silicon crystal planes can be fabricated; for example, the microchannels etched in $\langle 100 \rangle$ or in $\langle 110 \rangle$ silicon by using a potassium hydroxide (KOH) solution can form a V-shaped cross-section (as shown in Fig.3.1).

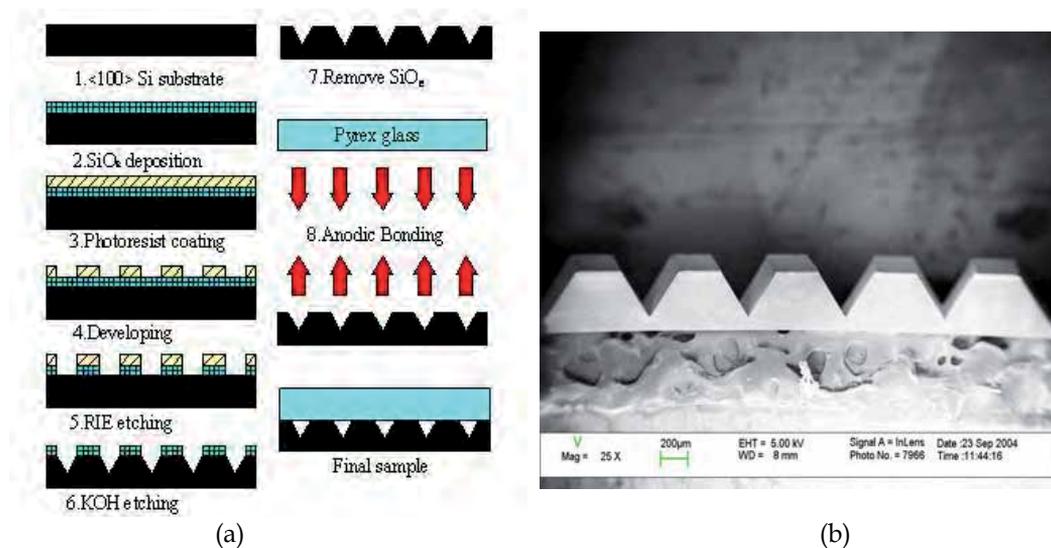


Fig. 3.1. (a) Flowchart of micro-manufacture processing; (b) V-shaped microchannels.

3.1.2 Results and discussion

For incompressible, fully-developed laminar flow, the friction factor can be expressed in terms of the two experimentally obtained parameters - pressure drop and mass flow rate.

$$f_{\text{exp}} = \frac{D_h}{L} \left(\frac{2\Delta P_{\text{exp}}}{\rho V_{\text{av}}^2} - \sum K_L \right) \tag{3-1}$$

where K_L is the friction factor for the minor loss. For the comparison of values of f vs. Re as shown in Fig. 3.2, the differences between the results obtained from numerical simulation and those from traditional correlation are within 2.5% of each other, within 6% between the numerical simulations and the experimental data, and with the f vs. Re values obtained from the experimental data and those obtained from the numerical simulations approaching a fixed value which is slightly lower than the value of 53.3 predicted by the traditional theory.

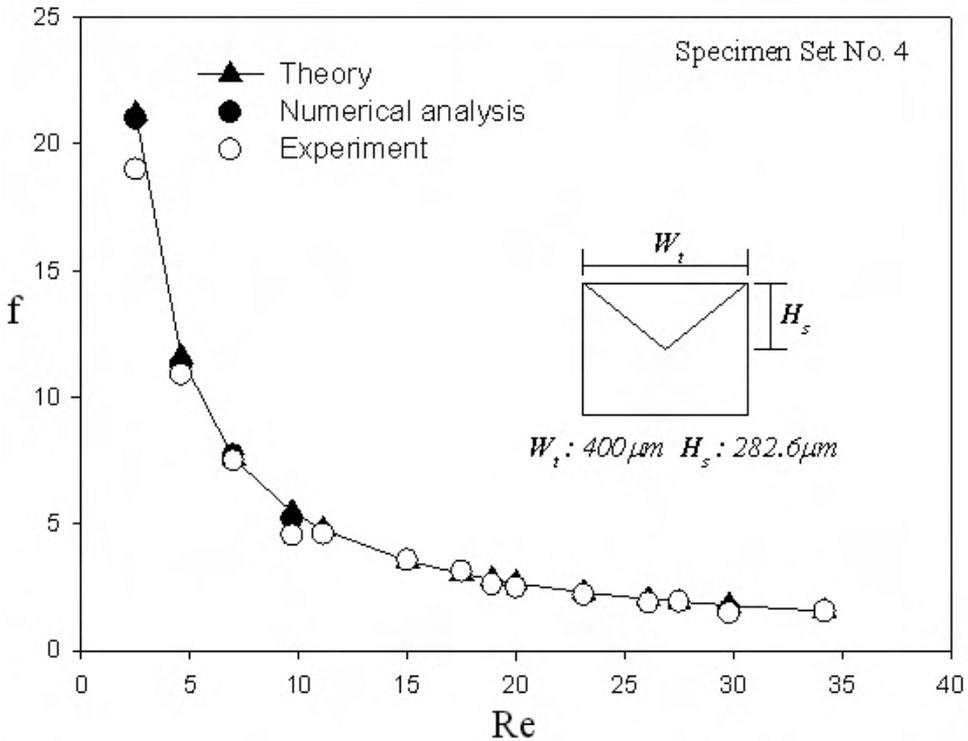


Fig. 3.2. Comparison of f vs. Re for theoretical values, predicted values, and experimental data for Specimen Set No. 4 [3].

Wu and Cheng [29] proposed a correlation for the V-shaped microchannels ($W_b/W_t = 0$) for fluid at low Reynolds numbers as follows.

$$Nu = 6.7 Re^{0.946} Pr^{0.488} \left(1 - \frac{W_b}{W_t} \right)^{3.547} \left(\frac{W_t}{H} \right)^{3.577} \left(\frac{\epsilon}{D_h} \right)^{0.041} \left(\frac{D_h}{L} \right)^{1.369}, \quad Re < 100 \tag{3-2}$$

where W_b and W_t are the bottom and the top width of microchannel, respectively. And ε is the surface roughness.

Referring Wu and Cheng [29], Chu [3] proposed an empirical correlation, based on experimentally obtained data from four sets of triangular microchannel test specimens (with different channel widths) under low Reynolds number conditions ($Re < 50$).

$$Nu = \frac{1}{1.2 + (-23.1 + 25.4W_t^{0.5})^2 Re^{-2}} \quad (3-3)$$

Generally speaking, the trends of the predicted results obtained from the correlation specified by Eq. (3-3) are in agreement, as shown in Fig. 3.3, while the widths of the microchannels have an obvious impact on the behavior of the development of the Nusselt numbers for the microchannels under study. It is noted that the magnitude of the Nusselt number increases at a slower rate as the Reynolds number becomes larger than 20.

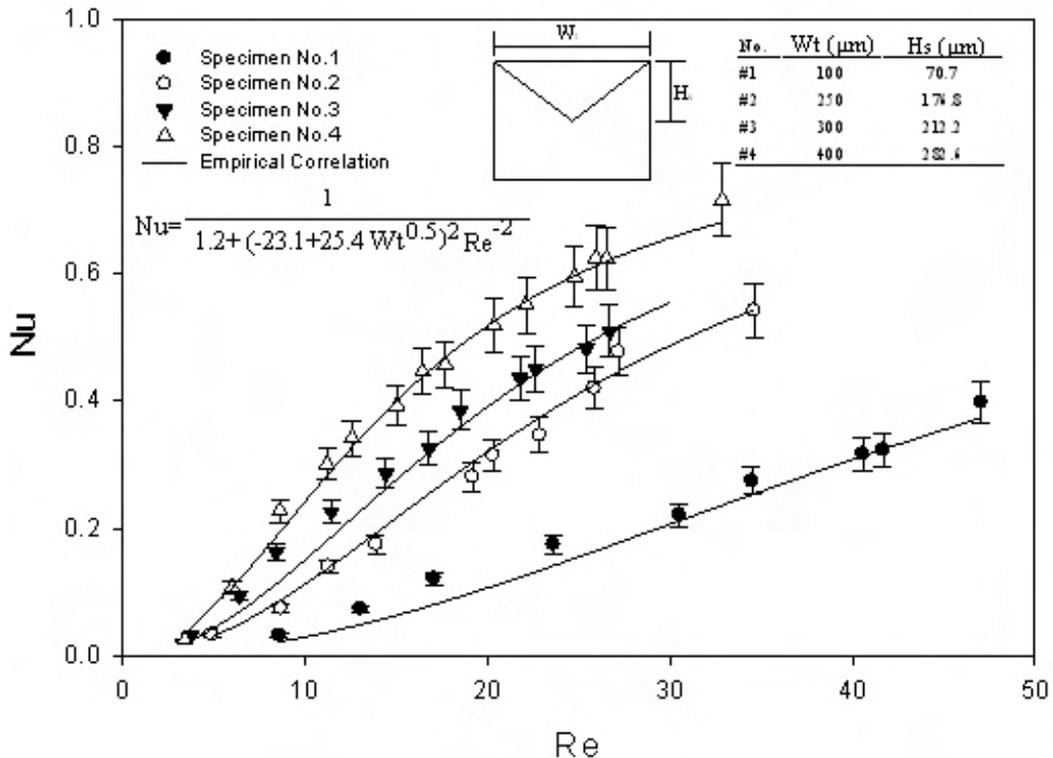


Fig. 3.3. Comparison of Nu vs. Re among empirical correlation and experimental data.

It is also noted that high temperature gradients at the inlet and exit were observed from the temperature distributions of microchannels for all sets of the test specimens. In addition, the Nusselt numbers increase as the Reynolds number increases, as shown in Fig. 3.3.

For the range of the Reynolds number being tested ($Re < 50$), the average discrepancy of the values calculated from the correlation of Nu obtained in [3] and those obtained from the experimental data is within 15%; the difference is judged to be in fair agreement.

3.2 Flow in circular curved microchannels

Among various micro-fluidic systems such as micro-coolers, micro-biochips, micro-reactors, and micro-fuel cells [3], the curved or bended microchannel (as shown in Fig. 3.5) has been identified as being one of the essential elements for shifting the direction of fluid flow, increasing the length of the path of the fluid flow, enhancing mixing efficiency, and improving heat transfer performance within a confined and compact space [24-26]. Therefore, it is extremely important to acquire a fundamental understanding of the flow behavior of fluid in curved microchannels, since its behavior affects the transport phenomena for the design and process control of micro-fluidic systems.

Up to the present time, there have been numerous investigations in the characteristics of the flow inside straight microchannels. However, a review of the literatures relative to researches conducted in straight microchannels during the last decade [37, 38] has revealed that only a handful of experimental or computational evaluations were done on the study of flow characteristics in curved microchannels [39-41].

For the manufacture of curved microchannels, Chu [3] demonstrated that the curved microchannel could be constructed by standard etching processes; the curved microchannel was etched on a silicon wafer with a 4-inch diameter and a 550 μm thickness. The processes included SiO_2 deposition, photoresist coating, developing, baking, etc. Subsequently, an inductively coupled plasma (ICP) process accounting for the crystal directional characteristics was used to finish the fabrication of the curved microchannel structure (as shown in Figs. 3.4 and 3.5).

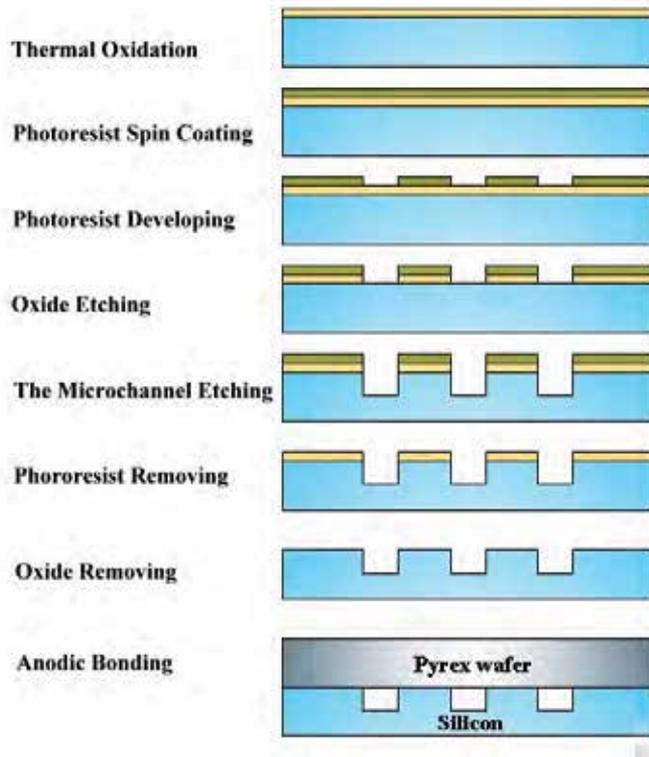


Fig. 3.4. Process flow of fabrication for curved rectangular microchannels.

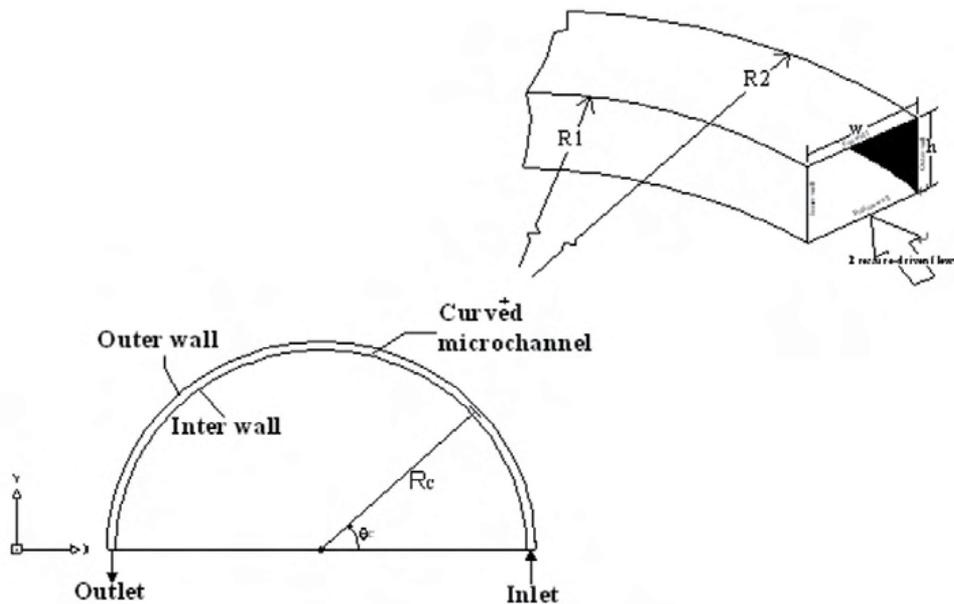


Fig. 3.5. Schematic diagram of the geometry for the curved rectangular microchannel.

For curved microchannels, the geometrical configurations used for testing are given in Table 3.1 [3].

Channel type	Channel width, w (μm)	Channel height, h (μm)	Radius of curvature, R_c (μm)	Aspect ratio, α_c	Curvature ratio
C1	200	200	5,000	1	0.04
C2	200	200	10,000	1	0.02
C3	300	200	5,000	0.667	0.048
C4	300	200	7,500	0.667	0.032
C5	300	200	10,000	0.667	0.024
C6	400	200	5,000	0.5	0.0533
C7	400	200	7,500	0.5	0.0355
C8	400	200	10,000	0.5	0.0266
C9	200	40	5,000	0.2	0.0133
C10	200	40	7,500	0.2	0.0088
C11	200	40	10,000	0.2	0.0066
C12	300	40	7,500	0.133	0.0094
C13	300	40	10,000	0.133	0.007
C14	400	40	5,000	0.1	0.0146
C15	400	40	10,000	0.1	0.0072

Table 3.1. Geometrical parameters of the curved microchannels used for testing

For incompressible flow through horizontal channels of constant cross-sectional area, the Fanning friction factor f_c is based on the mass flow rate and the pressure drop ΔP_f , where the latter is due to the friction occurred inside the curved microchannel.

$$f_c = \frac{2\tau_w}{\rho u_m^2} = \frac{D_h \Delta P_f}{2\rho u_m^2 L} = \frac{180}{\pi R_c \theta_c} \times \frac{w^3 h^3 \Delta P_f}{\rho Q^2 (h+w)} \quad (3-4)$$

where τ_w is the wall shear stress, L is the channel length, u_m is the mean flow velocity, Q is the volumetric flow rate of the working fluid, R_c is the radius of curvature, and θ_c is the angle of the microchannel.

According to the recommendations and method described by Holman [42], Chu [3] proposed the expression of uncertainties associated with Re , f and the product fRe for curved microchannels.

$$\frac{U_{Re}}{Re} = \left[\left(\frac{\delta\rho}{\rho} \right)^2 + \left(\frac{\delta\mu}{\mu} \right)^2 + \left(\frac{\delta Q}{Q} \right)^2 + \left(\frac{\delta h + \delta w}{h+w} \right)^2 \right]^{1/2} \quad (3-5)$$

$$\frac{U_f}{f} = \left[\left(\frac{\delta\rho}{\rho} \right)^2 + \left(\frac{\delta\Delta p}{\Delta p} \right)^2 + 4 \left(\frac{\delta Q}{Q} \right)^2 + \left(\frac{\delta L}{L} \right)^2 + 9 \left(\frac{\delta h}{h} \right)^2 + 9 \left(\frac{\delta w}{w} \right)^2 + \left(\frac{\delta h + \delta w}{h+w} \right)^2 \right]^{1/2} \quad (3-6)$$

$$\frac{U_{fRe}}{fRe} = \left[\left(\frac{\delta\mu}{\mu} \right)^2 + \left(\frac{\delta\Delta p}{\Delta p} \right)^2 + \left(\frac{\delta Q}{Q} \right)^2 + \left(\frac{\delta L}{L} \right)^2 + 9 \left(\frac{\delta h}{h} \right)^2 + 9 \left(\frac{\delta w}{w} \right)^2 + 4 \left(\frac{\delta h + \delta w}{h+w} \right)^2 \right]^{1/2} \quad (3-7)$$

In Eqs. (3-5) - (3-7), it is observed that the measurement errors in the height h and width w of the microchannel dimensions have a significant influence on the overall uncertainty.

In order to compare the flow characteristics of the curved microchannel with those of the conventional dimensions, the relationship between the friction factor and the Reynolds number estimated from the empirical correlation proposed by Yang et al. [43] are plotted in Figs. 3.6 and 3.7 (it should be noted that the correlation was originally proposed by Hua and Yang [44]). The correlation is expressed by the following equation.

$$f = \frac{5}{Re^{0.65}} \left(\frac{w}{2R_c} \right)^{0.175} \quad (3-8)$$

From Figs. 3.6 and 3.7, it is observed that when $Re < 600$, all experimentally-determined friction factors decrease nonlinearly with an increase of the Reynolds number. For the curved microchannels with aspect ratios varying from 0.5 to 1, it was observed that the friction factor of the curved microchannels was mainly affected by the curvature ratio and the Reynolds number. However, for curved microchannels with aspect ratios varying from 0.1 to 0.2, the aspect ratio had a significant effect on the relationship between the friction factor and the Reynolds number.

Another important parameter used to describe the laminar flow fluid behavior in channels is the Poiseuille number Po , which is a product of the friction factor and the Reynolds number. According to the numerical investigation presented by Wang and Liu [45] for a curved microchannel with an aspect ratio of 1 and a curvature ratio of 5×10^{-6} , the predicted relationship between the friction factor ratio and De number, defined by $De = (\rho u_m D_h / \mu) (D_h / 2R_c)^{1/2}$, can be expressed by the following equation.

$$fRe_c / fRe_s = 0.96194 + 0.01035De^{0.78715}, \quad \text{where } 0 < De \leq 450 \quad (3-9)$$

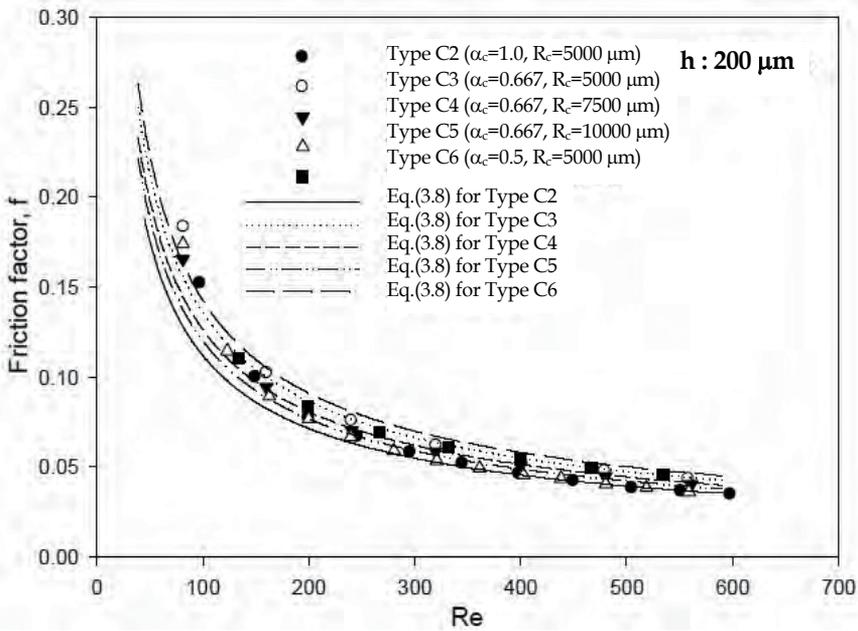


Fig. 3.6. Comparison of friction factor versus Re number for Type C2 to C6 microchannels.

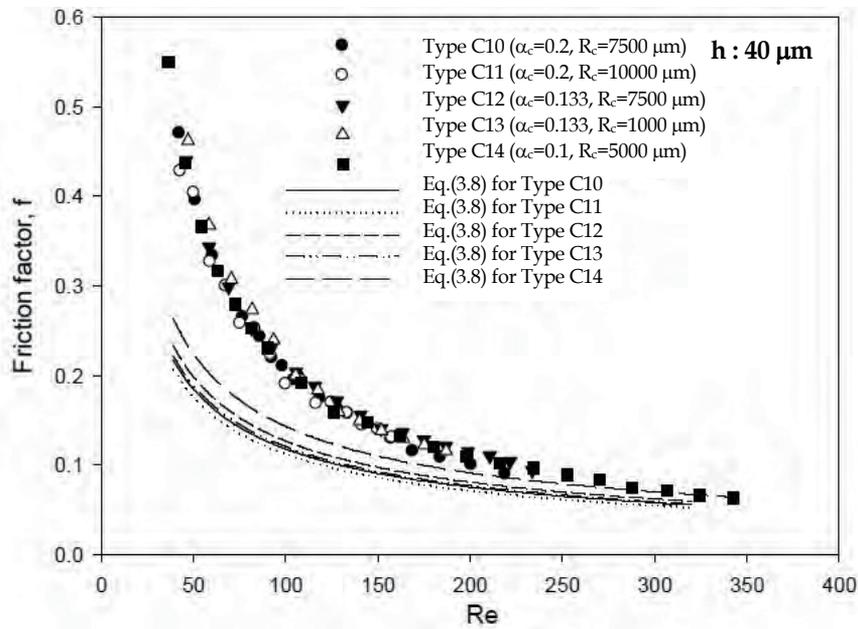


Fig. 3.7. Comparison of friction factor versus Re number for Type C10 to C14 microchannels.

3.3 Flow in fractal-like microchannels

Since the paper published by Tuckerman and Pease [46], manifold microchannel heat sinks using single-phase fluid as coolant have emerged as one of the effective and promising cooling techniques for microelectronic cooling for the last two decades. Due to the high surface-to-volume ratio and compactness for the microchannels, application of microchannels possesses the potential to be an attractive method for cooling micro-system devices. For this reason, numerous investigations have been devoted to fluid flow and heat transfer characteristics of forced convection of water in parallel microchannels [47, 48, 49, 50, 51, 52], and the majority of investigations mainly focused on single phase flow and heat transfer in circular, trapezoidal, rectangular, and parallel plate microchannels by analyzing the variations in the physical behaviors associated with the friction factor, region of transition, and the Nusselt numbers [53]. Moreover, among the extensive studies of micro-thermo-fluidics in straight microchannels, discrepancies were found among the results obtained from the experimental data and those obtained from the classical theories.

Generally speaking, the conventional design of flow architectures such as parallel plate microchannels is based on a unique length scale that is distributed uniformly throughout the available space. However, a network of straight microchannels could cause non-uniform temperature distribution [54, 55], high pressure gradient [56, 57, 58], and flow maldistribution [59, 60, 61, 62] in the micro-fluidic systems. Thus, it is necessary to develop new type of microchannel geometries to improve the hydrodynamic characteristics of straight microchannel networks.

Consequently, many researchers [63-71] have tried to obtain useful guidelines from the efficient transport properties of nature systems for the design of flow architectures in micro-system.

3.3.1 Model description

Flow architectures are ubiquitous in nature systems such as mammalian circulatory and respiratory networks, arteries and veins in animals, stems and leaves in plants, and river basins [72]. The structure of mammalian lungs is a typical example of a distribution system with a nearly tree-shaped structure.

Bejan and Errera [73] first investigated the architecture of the volume-to-point path such that the flow resistance is minimal. They found that fractal-like networks configuration could provide a minimal flow resistance. Later, Lorente et al. [74] proposed a simpler and direct route to determine the construction of effective fractal-like flow structure in thermal and fluidic systems and discussed the importance of the simplified design method.

To generate a fractal-like microchannel network, the analytical configuration of fractal-like microchannel is characterized by the following constant ratios of the length and hydraulic diameter of the channel at the (k+1)th branching level to the length and hydraulic diameter of the channel at the kth branching level, respectively. The ratios γ and β are defined as:

$$\gamma = \frac{L_{k+1}}{L_k} = N^{-1/D_l} \quad (3-10)$$

$$\beta = \frac{D_{k+1}}{D_k} = N^{-1/D_d} \quad (3-11)$$

where N represents the number of branches into which a single channel is bifurcated, D_L and D_d are fractal dimensions associated with the length and diameter of the channels, respectively, and L_k and D_k represent the length and hydraulic diameter of the fractal-like channel section at branching k , respectively, with k originated from zero.

An initial channel of length L_0 and diameter D_0 bifurcates at one end, and the new channels of length L_1 and diameter D_1 bifurcate at each end to produce the first branching level of the fractal networks, as shown in Fig. 3.8. The bifurcations at the ends of the newly formed channels may be reproduced until the required branching level of the fractal-like microchannel networks is obtained, as shown in Fig. 3.9.

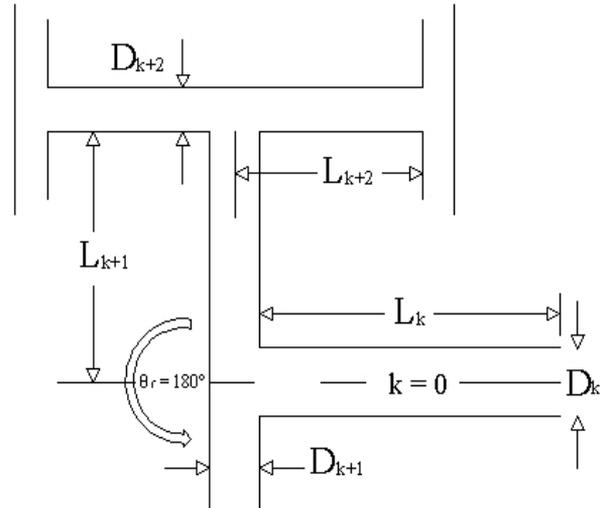


Fig. 3.8. Generation of a fractal-like microchannel networks.

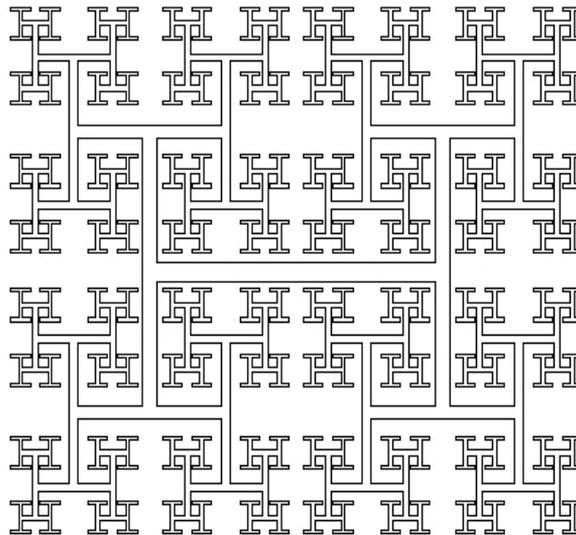


Fig. 3.9. Schematic of the geometric structure of fractal-like microchannel networks with 180° branch angle.

According to Murray's study [75] on cardiovascular system, it has been found that there is an optimal dimension associated with the hydraulic diameter D_d for the fractal-like networks, such that the global flow resistance is minimized. In the study done by Chu [3], the choice of $D_d = 3$ was taken in the numerical calculations by following the Murray's study, and the values of D_L was taken as 1.4 in setting up the computation domain of fractal-like microchannel networks. In addition, the branching angle θ_f was chosen to be 180° , N was set to be equal to two for the present analysis, and a rectangular cross-section with fixed channel depth was assumed for all of channel branches. The detail dimensions of the fractal-like microchannel networks based on Eqs. (3-10) and (3-11) are given in Table 3.2.

k	0	1	2	3	4	5
H_k (mm)	0.5	0.5	0.5	0.5	0.5	0.5
W_k (mm)	1.5	0.74	0.45	0.30	0.21	0.15
D_k (mm)	0.75	0.60	0.47	0.37	0.30	0.24
L_k (mm)	15	9.14	5.57	3.40	2.07	1.26

Table 3.2. Channel dimensions of the fractal-like microchannel networks

Furthermore, the pumping power in the fractal and parallel channel networks is compared with the theoretical correlation based on the recommendation proposed by Chen and Cheng [53] as follows.

$$P_f/P_p = \frac{[1-(\gamma/(N\beta^4))^{m+1}](1-N\gamma)}{[1-\gamma/(N\beta^4)][1-(N\gamma)^{m+1}]} \quad (3-12)$$

where P_f is the pumping power in fractal-like channel networks and P_p is the pumping power in parallel channel networks.

3.3.2 Results and discussion

The flow in straight microchannels with low Reynolds number is mainly regarded as laminar flow, with a parabolic profile under fully-developed flow condition. However, due to the disturbance effect of channel pattern for the fractal-like microchannel networks, the water flow in straight channel deviates from the laminar situation when the fluid flows through the T-shaped bifurcations. Fig. 3.11 displays the pressure drop variation along Path A_{a-b-c} and Path C_{d-e-f} and in the middle of the channel (Path B). The measured path of pressure distribution of the fractal-like microchannel networks is shown in Fig. 3.10.

As the water flows through, the sharp increase of pressure is developed in the center and outer sides of the channel, and the inner pressure is rapidly decreasing behind the sharp corner. Then, behind the branches of these T-shaped bifurcations, the distribution of outer and center pressures are dropping immediately, and the velocity boundary layer redevelopment is observed before the next T-shaped bifurcation, as shown in Fig. 3.10. Due to the curvature of the bifurcation, the water flow is directed into a new direction. At this time, the centrifugal forces push the water flow from the center of the inlet channel to the outer wall, and the pressure on the outer side of the channel is increased.

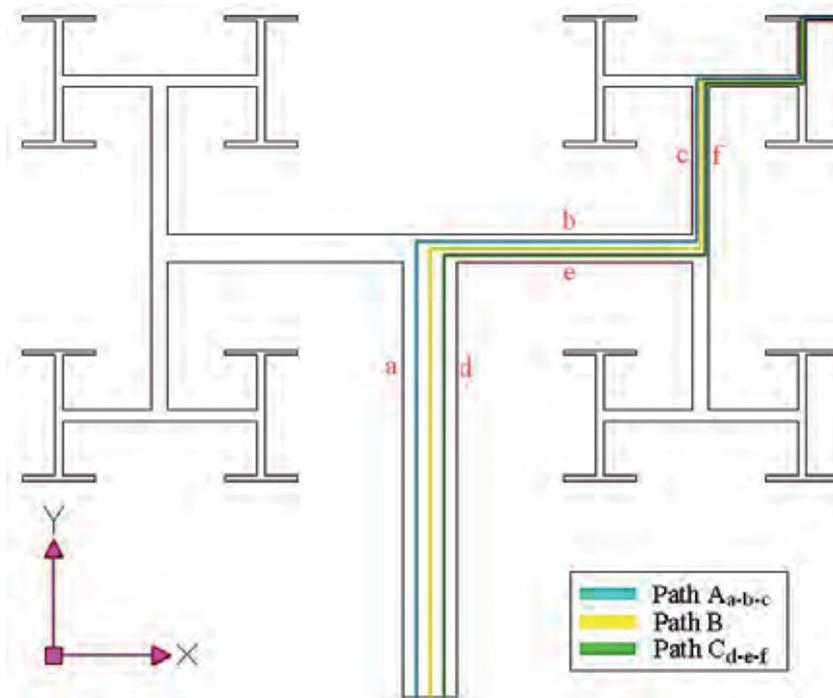


Fig. 3.10. Paths A, B, C (includes branches $k=0$, $k=1$, $k=2$, $k=3$, $k=4$, and $k=5$) and positions of outer and inner walls.

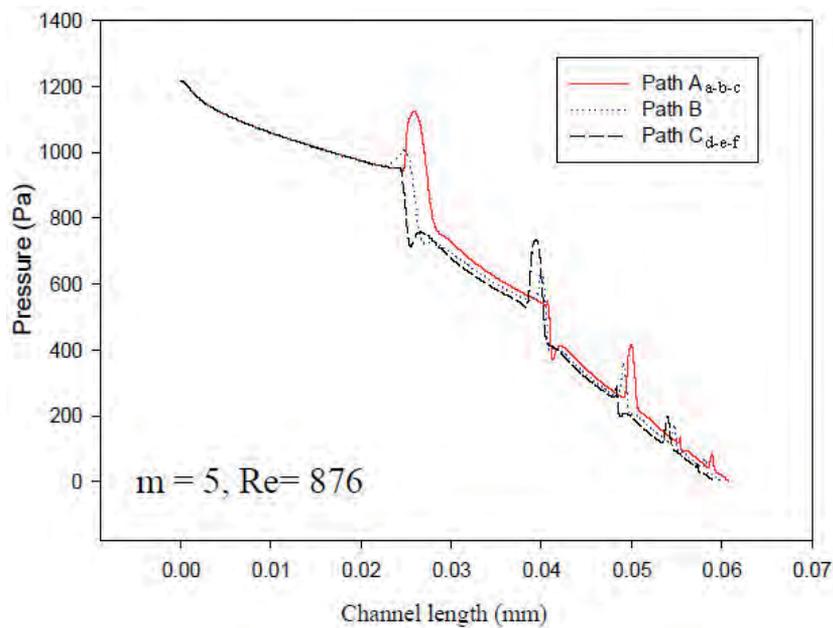


Fig. 3.11. Pressure distribution in fractal-like channel for $m=5$ case.

In the fractal-like microchannel networks, as the water flow passes through the T-shaped bifurcation, a symmetric double vortex is generated near the top and bottom walls with respect to the centerline of the cross-section as shown in Fig. 3.12. Note that the outside wall is to the right of each flow pattern shown in the Fig. 3.12.

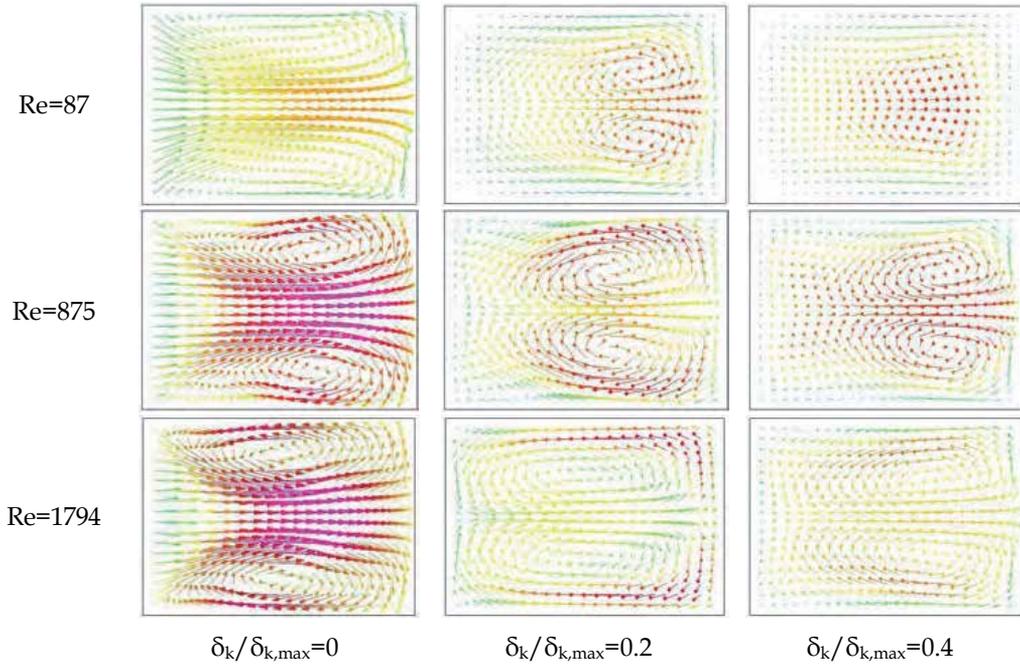


Fig. 3.12. Secondary flow pattern at cross-sections of channels at branch $k = 1$ with three different Re numbers (which δ_k is the local coordinate indicating the distance from the inlet of the k th bifurcation).

As observed in Fig. 3.12, the secondary flow is composed of two-vortex flow and the type of two-vortex flow rotates in the opposite direction. The flow direction on the centerline is toward the outside wall. As the inlet Reynolds number increases from the $k = 1$ segment region, the spanwise component of velocity is stronger and the larger secondary velocities are concentrated near the outside wall. However, it can be seen that the secondary flow initiated at the inlet of the branch ($k=1$) gradually diminishes with increasing distance from the inlet of the branch ($k=1$).

The variation of pumping power ratios versus the Reynolds numbers at different branching levels is plotted in Fig. 3.13. It is seen that the ratio of pumping powers for the cases with $m = 3$ to 5 varies linearly as Reynolds number increases, and the variation of pumping powers ratio versus Reynolds number diminishes gradually with increasing branching levels. The region under the $P_f/P_p = 1$ line indicates that the fractal-like microchannel network exhibits better hydrodynamic performance relative to that of the parallel channel network.

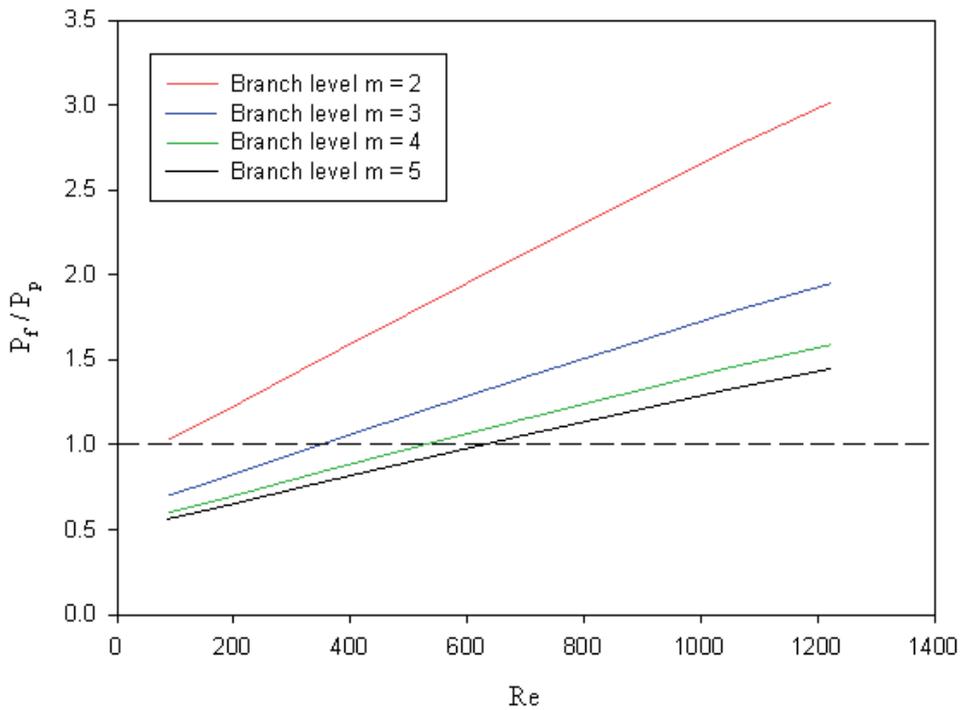


Fig. 3.13. Variation of P_f/P_p versus Re with different branch levels.

4. Flow motion and heat transfer in microchannels with turbulence generators

An experimental study [76] was conducted on the rectangular microchannels with longitudinal vortex generators (LVGs) in the Reynolds number up to 1,200 by using DI water as working fluid. The results can be summarized by three concluding remarks: 1) heat transfer was enhanced with the help of longitudinal vortices in rectangular microchannel while encountering larger pressure drop; 2) it was also found that laminar-to-turbulent transition occurred earlier in rectangular microchannel with LVGs than that for the smooth rectangular microchannel; 3) different configurations of LVGs in rectangular microchannel resulted in different overall heat transfer performance (the ratio of heat transfer enhancement to pressure drop increase), which increased with an increase of the Reynolds number. Discussed below were results of the research studies done by Chu [3] and Liu et al. [76].

4.1 Model description

Dimensions of rectangular microchannels with longitudinal vortex generators are shown in Fig. 4.1. H , W and L are the height, width, and length of microchannels, respectively. The geometrical configuration of LVGs is also shown in Fig. 4.1, which illustrates the length, width, and angle of attack for the LVGs. More details about the locations of LVGs (channel types G1~G7) in test chips are shown in Table 4.1.

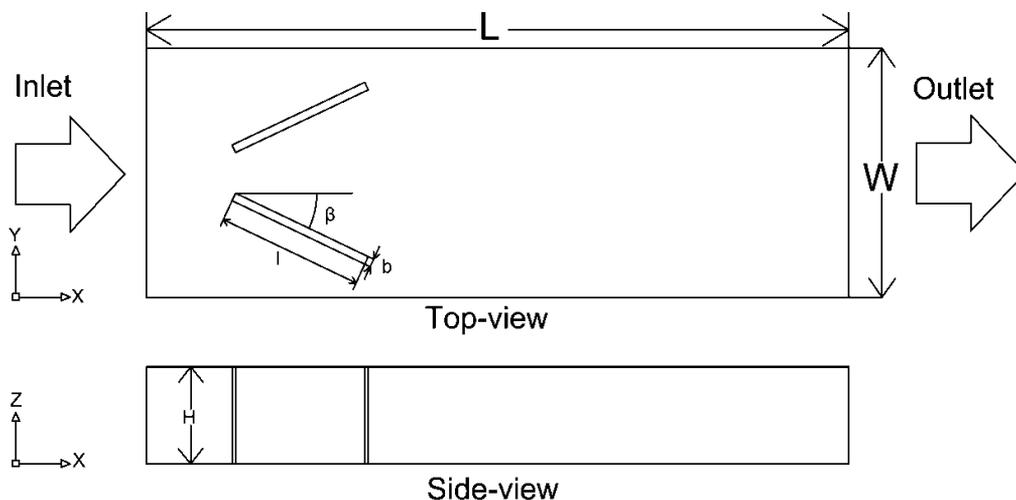


Fig. 4.1. Schematic diagram of rectangular microchannel with longitudinal vortex generators [3, 76].

Channel	Configurations	Schematic Diagram
G1	(5 × 30°) or (5 × 150°)	
G2	(5 × 45°) or (5 × 135°)	
G3	Smooth channel	
G4	(3 × 30°) or (3 × 150°)	
G6	(30°, 150°, 30°) or (150°, 30°, 150°)	
G7	(45°, 135°, 45°) or (135°, 45°, 135°)	

Table 4.1. List of rectangular microchannel configurations

4.2 Results and discussion

From Fig. 4.2, it can be observed that the rectangular microchannels with LVGs clearly have better heat transfer enhancement than the smooth rectangular microchannel (channel type G4). It can also be seen from Fig. 4.2 that for the rectangular microchannels with LVGs, the slopes of the Nusselt number curves change abruptly when the Reynolds number reaches the range of 600-730. From the conclusion of reference [54], the laminar-to-turbulent transition occurs in a similar range of the Reynolds numbers.

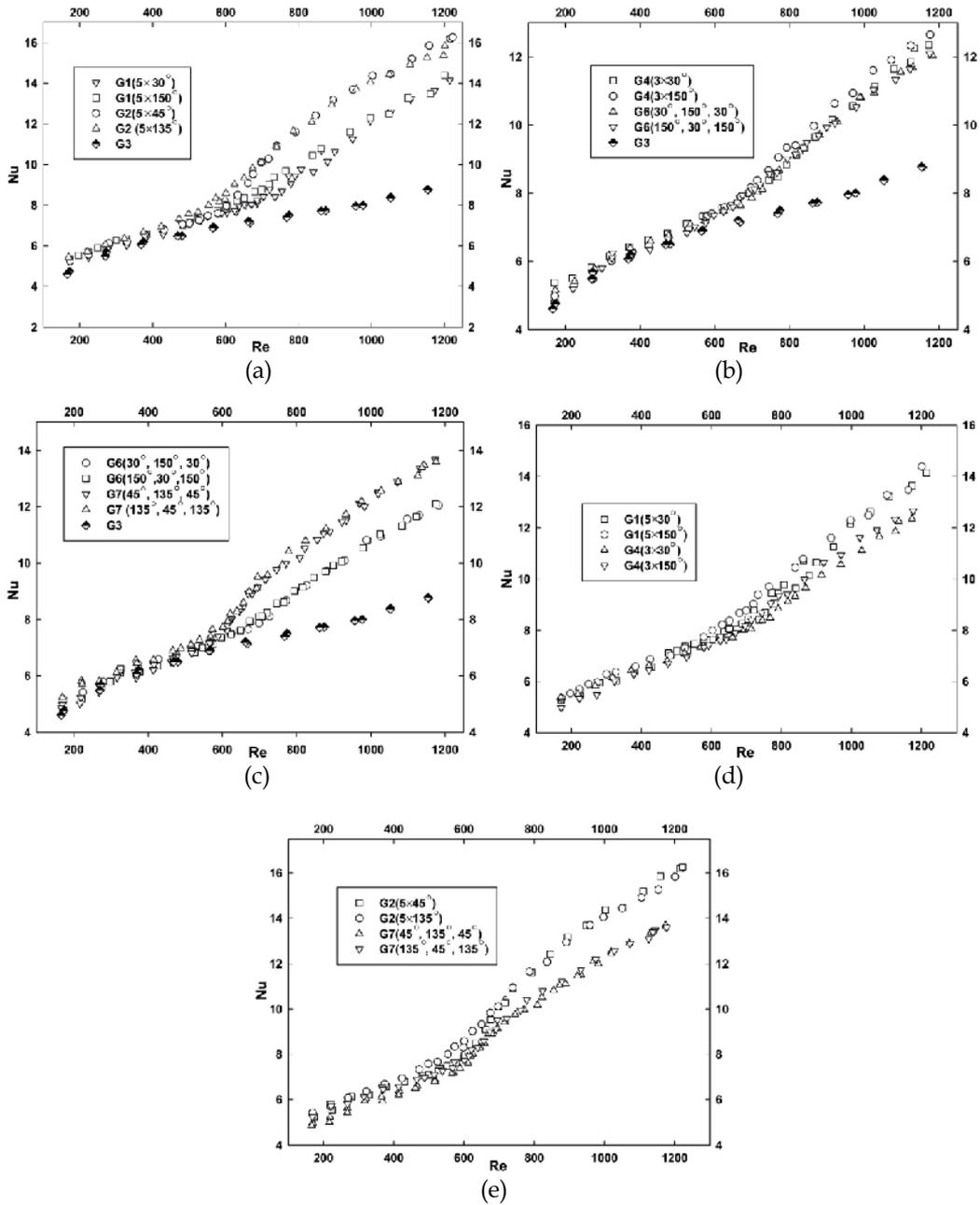


Fig. 4.2. Nusselt number as a function of Reynolds number for microchannels with longitudinal vortex generators [3, 76].

From Fig. 4.3, one can observe that the rectangular microchannels with LVGs result in much larger pressure drop than that for the smooth rectangular microchannel. In addition, different configurations of LVGs in rectangular microchannel demonstrate different fluid flow characteristics.

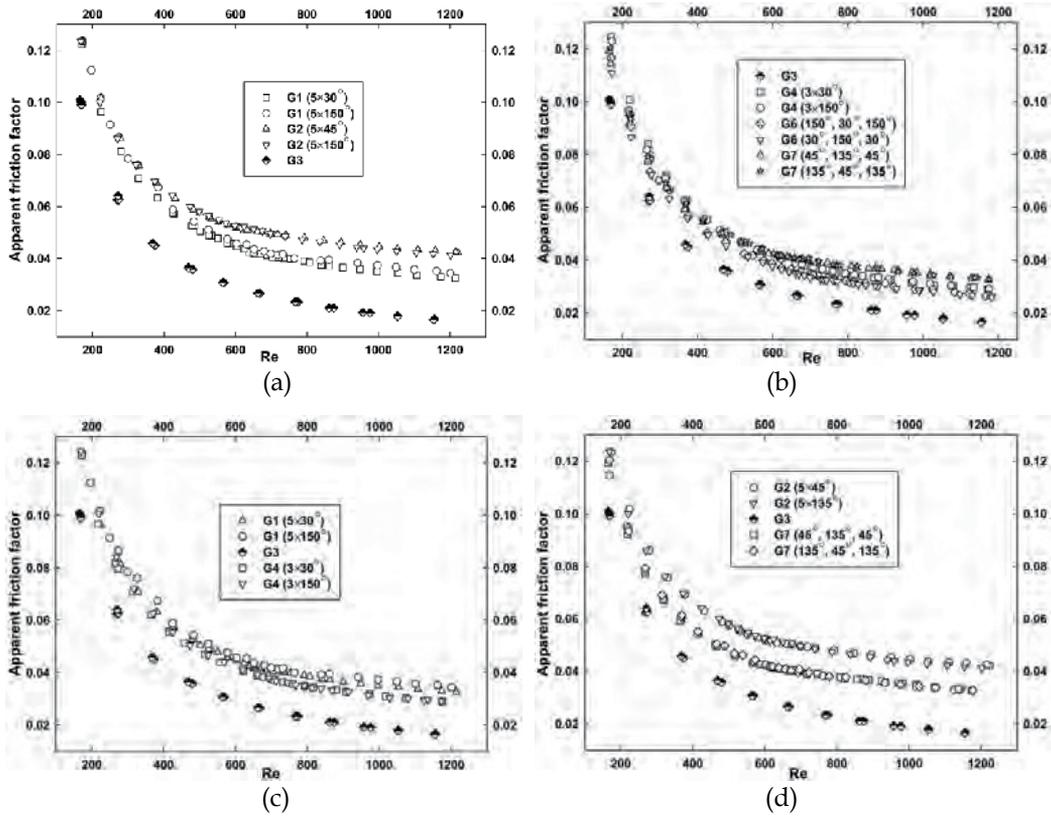


Fig. 4.3. Apparent friction factor as a function of Reynolds number for (a) G1, G2 and G3, (b) G3, G4, G6 and G7, (c) G1, G3 and G4 and (d) G2, G3 and G7 channels [3, 76].

4.3 Empirical correlations

4.3.1 Apparent friction factor correlations

The empirical correlations of experimental data for apparent friction factors are listed in Table 4.2.

	Laminar flow regime	Ranges of applicability	Turbulent flow regime	Ranges of applicability
G1	$f = 7.1 / Re^{0.792}$	$7.8 < RePrD_h/L < 36$	$f = 0.657 / Re^{0.424}$	$36 < RePrD_h/L < 61$
G2	$f = 4.657 / Re^{0.707}$	$7.8 < RePrD_h/L < 31$	$f = 0.324 / Re^{0.286}$	$31 < RePrD_h/L < 61$
G4	$f = 10.016 / Re^{0.859}$	$7.8 < RePrD_h/L < 36$	$f = 0.875 / Re^{0.482}$	$36 < RePrD_h/L < 61$
G6	$f = 9.088 / Re^{0.858}$	$7.8 < RePrD_h/L < 37$	$f = 0.734 / Re^{0.471}$	$37 < RePrD_h/L < 61$
G7	$f = 7.443 / Re^{0.815}$	$7.8 < RePrD_h/L < 32$	$f = 0.431 / Re^{0.364}$	$32 < RePrD_h/L < 61$

Table 4.2. Empirical correlations for apparent friction factors

4.3.2 Heat transfer correlations

For laminar and turbulent flows in rectangular microchannels with LVGs, the empirical correlations (obtained by curve-fitting) of the experimental data are shown in Table 4.3.

Channel	Laminar flow regime	Ranges of applicability
G1	$Nu = 4.76 + \frac{0.366(Re Pr D_h / L)^{1.253}}{1 + 5.56(Re Pr D_h / L)^{0.095}}$	$7.8 < Re Pr D_h / L < 36$
G2	$Nu = 4.67 + \frac{0.329(Re Pr D_h / L)^{1.322}}{1 + 4.041(Re Pr D_h / L)^{0.189}}$	$7.8 < Re Pr D_h / L < 31$
G4	$Nu = 4.26 + \frac{0.432(Re Pr D_h / L)^{1.358}}{1 + 2.595(Re Pr D_h / L)^{0.471}}$	$7.8 < Re Pr D_h / L < 36$
G6	$Nu = 4.07 + \frac{0.418(Re Pr D_h / L)^{1.357}}{1 + 2.321(Re Pr D_h / L)^{0.479}}$	$7.8 < Re Pr D_h / L < 37$
G7	$Nu = 4.5 + \frac{0.364(Re Pr D_h / L)^{1.168}}{1 + 5.269(Re Pr D_h / L)^{0.015}}$	$7.8 < Re Pr D_h / L < 31$
Channel	Turbulent flow regime	Ranges of applicability
G1	$Nu = 0.011 Re^{0.934} Pr^{1/3}$	$36 < Re Pr D_h / L < 61$
G2	$Nu = (19.85 - 372.1 Re^{-0.5}) Pr^{1/3}$	$31 < Re Pr D_h / L < 61$
G4	$Nu = 0.0182 Re^{0.845} Pr^{1/3}$	$36 < Re Pr D_h / L < 61$
G6	$Nu = 0.0311 Re^{0.763} Pr^{1/3}$	$37 < Re Pr D_h / L < 61$
G7	$Nu = (39.03 - 221.5 / \ln Re) Pr^{1/3}$	$31 < Re Pr D_h / L < 61$

Table 4.3. Empirical correlations for Nusselt numbers [3]

5. Concluding remarks

The above topic was chosen to be included in this chapter on Fluid Dynamics in Microchannels due to the fact that one of the key research categories done in Thermo-Fluids Analysis Group (TFAG) Lab at Department of Mechanical Engineering, Chung Yuan Christian University in Chung-Li, Taiwan, is associated with the study on the behaviors of fluid flow and heat transfer for water flowing through microchannels. In addition, fluid flow of micro-scale channels is of interest to many researchers, academicians, and practitioners; thus, the topic was deemed to be an appropriate one to be included in the book on Fluid Dynamics.

This chapter summarized the work performed and the results obtained in both fluid flow and heat transfer done by TFAG over the last several years. The authors would like to express their deep appreciation for the financial supports obtained from National Science Council, Taiwan (Grant Nos. NSC93-2212-E-033-012, NSC94-2212-E-033-017, NSC95-2212-E-033-066, NSC96-2212-E-033-039, NSC97-2212-E-033-050, and NSC99-2212-E-033-025) and Chung Yuan Christian University (Grant No. CYCU-98-CR-ME).

6. Nomenclature

A	Microchannel Area, m^2
A_{ch}	Microchannel Area, m^2
c_p	Specific heat at constant pressure, $kJ\ kg^{-1}\ K^{-1}$
c_v	Specific heat at constant volume, $kJ\ kg^{-1}\ K^{-1}$

D_c	Diameter of microchannel, m
D_d	Fractal dimensions associated with the diameter
De	Dean number
D_h	Hydraulic diameter, m
D_L	Fractal dimensions associated with the length
D_{le}	Laminar equivalent diameter, Eq. (2-11)
Ec	Eckert number
f	Friction factor
G	Vortex generator
H	Microchannel height, m
k	Branch serial number
K	Friction factor for minor loss (Eq. (3-8)) Hagenbach factor (Eq. (2-8))
L	Microchannel length, m
\dot{m}	Mass flow rate, kg s ⁻¹
N	Number of branches in fractal-like microchannels
Nu	Nusselt number
P_o	Poiseuille number
Pr	Prandtl number
P_w	Wetted perimeter, m
Q	Volumetric flow rate, m ³ s ⁻¹
R_c	Radius of curvature, m
Re	Reynolds number
S	Source term (Eq. (1-1))
T	Temperature, K
u	Velocity, m s ⁻¹
W	Microchannel width, m
W_b	Bottom width of trapezoidal microchannels
W_t	Top width of trapezoidal microchannels
ε	Surface roughness, m

Greek symbols

a_c	Aspect ratio
β	Diameter ratio
γ	Length ratio
θ	Angle of vortex generator
λ	thermal conductivity, W m ⁻¹ K ⁻¹
μ	Dynamic viscosity, kg m ⁻¹ s ⁻¹
ρ	Density, kg m ⁻³
τ	Shear stress, N m ⁻²

Subscripts

app	Apparent
c	Contraction
ch	Channel
e	Expansion
h	Hydraulic

<i>i</i>	Inlet
<i>m</i>	Arithmetic mean
<i>o</i>	Outlet
<i>s</i>	Bottom side of microchannel test chip
<i>t</i>	Turbulent

7. References

- [1] Frank P. Incropera, David P. Dewitt, *Fundamentals of Heat and Mass Transfer*, Third edition, Wiley, 1990.
- [2] S. G. Kandlikar, S. Garimella, D. Li, S. Colin and M. R. King, *Heat Transfer and Fluid Flow in Minichannels and Microchannels*, Elsevier Pte Ltd, Singapore, 2006.
- [3] J. C. Chu, *A Study on the Thermo-Fluidic Behaviors of Water Flowing through Microchannels*, Ph.D. Dissertation, Department of Mechanical Engineering, Chung Yuan Christian University, July 2010.
- [4] J. P. Van Doormaal and G. D. Raithby, "Enhancements of the SIMPLE method for predicting incompressible fluid flows," *Numerical Heat Transfer*, Vol.7, No. 2, pp. 147-163, 1984
- [5] M. Peric, R. Kessler, and G. Scheuerer, "Comparison of finite-volume numerical methods with staggered and collocated grids," *Computers and Fluids*, Vol. 16, No. 4, pp. 389-403, 1988.
- [6] J. P. Van doormaal and G. D. Raithby, "Enhancements of the SIMPLE method for predicting incompressible fluid flows," *Numerical Heat Transfer*, Vol. 7, pp. 147-163, 1984.
- [7] R. J. Phillips, *Forced convection, liquid cooled, microchannel heat sinks*, MS Thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, 1987.
- [8] W. M. Kays and A. L. London, *Compact Heat Exchangers*, New York, NY, McGraw-Hill, pp. 105-112, 1984.
- [9] W. L. Qu, G. M. Mala, and D. Q. Li, "Pressure-driven water flows in trapezoidal silicon microchannels," *International Journal of Heat and Mass Transfer*, Vol. 43, pp. 353-364, 2000.
- [10] J. Judy, D. Maynes, and B. W. Webb, "Characterization of frictional pressure drop for liquid flows through microchannels," *International Journal of Heat and Mass Transfer*, Vol. 45, pp. 3477-3489, 2002.
- [11] D. Lelea, S. Nishio, and K. Takano, "The experimental research on microtube heat transfer and fluid flow of distilled water," *International Journal of Heat and Mass Transfer*, Vol. 47, pp. 2817-2830, 2004.
- [12] M. J. Kohl, S. I. Abdel-Khalik, S. M. Jeter, and D. L. Sadowski, "An experimental investigation of microchannel flow with internal pressure measurements," *International Journal of Heat and Mass Transfer*, Vol. 48, pp. 1518-1533, 2005.
- [13] F. F. Abdelall, G. Hahn, S. M. Ghiaasiaan, S. I. Abdel-Khalik, S. S. Jeter, M. Yoda, and D. L. Sadowski, "Pressure drop caused by abrupt flow area changes in small channels," *Experimental Thermal and Fluid Science*, Vol. 29, pp. 425-434, 2005.
- [14] R. K. Shah and A. L. London, *Laminar flow forced convection in ducts*, *Advances in heat transfer supplement*, Academic Press, New York, 1978.

- [15] F. F. Abdelall, G. Hahn, S. M. Ghiaasiaan, S. I. Abdel-Khalik, S. S. Jeter, M. Yoda, and D. L. Sadowski, "Pressure drop caused by abrupt flow area changes in small channels," *Experimental Thermal and Fluid Science*, Vol. 29, pp. 425-434, 2005.
- [16] P. Hrnjak and X. Tu, "Single phase pressure drop in microchannels," *International Journal of Heat and Fluid Flow*, Vol. 28, pp. 2-14, 2007.
- [17] G. H. Tang, Z. Li, Y. L. He, and W. Q. Tao, "Experimental study of compressibility, roughness and rarefaction influences on microchannel flow," *International Journal of Heat and Mass Transfer*, Vol. 50, pp. 2282-2295, 2007.
- [18] R. J. Phillips, *Microchannel heat sinks, Advances in Thermal Modeling of Electronic Components and Systems*, New York, NY: Hemisphere Publishing Corporation, 1990, Chapter 3
- [19] A. Bucci, G. P. Celata, M. Cumo, E. Serra, and G. Zummo, "Water single-phase fluid flow and heat transfer in capillary tubes," Paper No. ICMM2004-2406, Second International Conference on Microchannels and Minichannels, Rochester, NY USA, pp. 221-228, June 17-19, 2004.
- [20] R. Baviere, F. Ayela, S. Le Person, and M. Farve-Marient, "An experiment study of water flow in smooth and rough rectangular microchannels," Paper No. ICMM2004-2338, Second International Conference on Microchannels and Minichannels, Rochester, NY USA, pp. 221-228, June 17-19, 2004.
- [21] D. J. Schmitt, and S. G. Kandlikar, "Effects of repeating microstructures on pressure drop in rectangular minichannels," Paper No. ICMM2005-75111, ASME, Third International Conference on Microchannels and Minichannels, Toronto, Canada, June 13-15, 2005.
- [22] Z. X. Li, D. X. Du, and Z. Y. Guo, "Experiment study on flow characteristics of liquid in circular microtubes," *Proceeding of the International Conference on Heat Transfer and Transport Phenomena in Microscale*, Banff, Canada, pp. 162-167, October 15-20, 2000.
- [23] R. K. Shah and A. L. London, *Laminar flow forced convection in ducts*, *Advances in heat transfer supplement*, Academic Press, New York, 1978.
- [24] Pho-Seng Lee, Suresh V. Garimella, and Dong Liu, "Investigation of heat transfer in rectangular microchannels," *International Journal of Heat and Mass Transfer*, Vol. 48, pp. 1688-1704, 2005.
- [25] S. M. Flockhart and R. S. Dhariwal, "Experimental and numerical investigation into the flow characteristics of channels etched in <100> silicon," *Journal of Fluids Engineering*, Vol. 120, pp. 291-295, 1998.
- [26] W. Qu, G. M. Mala, and D. Li, "Heat transfer for water flow in trapezoidal silicon microchannels," *International Journal of Heat and Mass Transfer*, Vol. 43, pp. 3925-3936, 2000.
- [27] W. L. Qu, G. M. Mala, and D. Q. Li, "Pressure-driven water flows in trapezoidal silicon microchannels," *International Journal of Heat and Mass Transfer*, Vol. 43, pp. 353-364, 2000.
- [28] G. Hetsroni, A. Mosyak, Z. Segal, and G. Ziskind, "A uniform temperature heat sink for cooling of electronic devices," *International Journal of Heat and Mass Transfer*, Vol. 45, pp. 3275-3286, 2002.

- [29] H. Y. Wu and P. Cheng, "An experimental study of convective heat transfer in silicon microchannels with different surface conditions," *International Journal of Heat and Mass Transfer*, Vol. 46, pp. 2547-2556, 2003.
- [30] H. Y. Wu and P. Cheng, "Friction factors in smooth trapezoidal silicon microchannels with different aspect ratios," *International Journal of Heat and Mass Transfer*, Vol. 46, pp. 2519-2525, 2003.
- [31] I. Tiselj, G. Hetsroni, B. Mavko, A. Mosyak, E. Pogrebnyak, and Z. Segal, "Effect of axial conduction on the heat transfer," *International Journal of Heat and Mass Transfer*, Vol. 47, pp. 2551-2565, 2004.
- [32] I. Tiselj, G. Hetsroni, B. Mavko, A. Mosyak, E. Pogrebnyak and Z. Segal, "Effect of axial conduction on the heat transfer in microchannels," *International Journal of Heat and Mass Transfer*, Vol. 47, pp. 2551-2565, 2004.
- [33] G. L. Morini, "Viscous heating in liquid flows in microchannels," *International Journal of Heat and Mass Transfer*, Vol. 48, pp. 3637-3647, 2005.
- [34] M. Yi and H. H. Bau, "The kinematics of bend-induced mixing in micro-conduit," *International Journal of Heat Fluid Flow*, Vol. 24, pp. 645-656, 2003.
- [35] N. Kockmann, M. Engler, D. Haller, and P. Wolas, "Fluid Dynamics and Transfer Processes in Bended Microchannels," *Heat Transfer Engineering*, Vol. 26, pp. 71-78, 2005.
- [36] F. Schonfeld and S. Hardt, "Simulation of helical flows in microchannels," *AIChE Journal*, Vol. 50, pp. 771-778, 2004.
- [37] G. L. Morini, "Single-phase convective heat transfer in microchannels: a review of experimental results," *International Journal of Thermal Sciences*, Vol. 43, pp. 631-651, 2004.
- [38] T. Bayraktar and S. B. Pidugu, "Characterization of liquid flows in microfluidic systems," *International Journal of Heat and Mass Transfer*, Vol. 49, pp. 815-824, 2006.
- [39] Y. Yamaguchi, F. Takagi, K. Yamashita, H. Nakanura, H. Maeda, K. Sotowa, K. Kusakabe, Y. Yamasaki, and S. Morooka, "3-D Simulation and visualization of laminar flow in a microchannel with hair-pin curves," *AIChE Journal*, Vol. 50, pp. 1530-1535, 2004.
- [40] W. H. Yang, J. Z. Zhang, and H. E. Cheng, "The study of flow characteristics of curved microchannel," *Applied Thermal Engineering*, Vol. 25, pp. 1894-1907, 2005.
- [41] L. Wang and F. Liu, "Forced convection in slightly curved microchannels," *International Journal of Heat and Mass Transfer*, Vol. 50, pp. 881- 897, 2007.
- [42] J. P. Holman, *Experimental Methods for Engineers*, McGraw-Hill, New York, 1984.
- [43] W. H. Yang, J. Z. Zhang and H. E. Cheng, "The study of flow characteristics of curved microchannel," *Applied Thermal Engineering*, Vol. 25, pp. 1894-1907, 2005.
- [44] S. Z. Hua and X. N. Yang, *Actual Fluid Friction Manual*, National Defense Industry Press, Beijing, pp. 269, 1985.
- [45] L. Wang and F. Liu, "Forced convection in slightly curved microchannels," *International Journal of Heat and Mass Transfer*, Vol. 50, pp. 881- 897, 2007.
- [46] D. B. Tuckermann and R. F. W. Pease, "High performance heat sinking for VLSI," *IEEE Electron Device Letters*, EDL-2, pp. 126-129, 1981.
- [47] W. Qu, G. M. Mala, and D. Li, "Heat transfer for water flow in trapezoidal silicon microchannels," *International Journal of Heat and Mass Transfer*, Vol. 43, pp. 3925-3936, 2000.

- [48] G. Hetsroni, A. Mosyak, Z. Segal, and G. Ziskind, "A uniform temperature heat sink for cooling of electronic devices," *International Journal of Heat and Mass Transfer*, Vol. 45, pp. 3275-3286, 2002.
- [49] H. Y. Wu and P. Cheng, "An experimental study of convective heat transfer in silicon microchannels with different surface conditions," *International Journal of Heat and Mass Transfer*, Vol. 46, pp. 2547-2556, 2003.
- [50] I. Tiselj, G. Hetsroni, B. Mavko, A. Mosyak, E. Pogrebnyak, and Z. Segal, "Effect of axial conduction on the heat transfer," *International Journal of Heat and Mass Transfer*, Vol. 47, pp. 2551-2565, 2004.
- [51] D. Lelea, S. Nishio, and K. Takano, "The experimental research on microtube heat transfer and fluid flow of distilled water," *International Journal of Heat and Mass Transfer*, Vol. 47, pp. 2817-2830, 2004.
- [52] W. Owhaib and B. Palm, "Experimental investigation of single-phase convective heat transfer in circular microchannels," *Experimental Thermal and Fluid Science*, Vol. 28, pp. 105-110, 2004.
- [53] G. L. Morini, "Single-phase convective heat transfer in microchannels: a review of experimental results," *International Journal of Thermal Sciences*, Vol. 43, pp. 631-651, 2004.
- [54] G. Hetsroni, A. Mosyak, Z. Segal, and G. Ziskind, "A uniform temperature heat sink for cooling of electronic devices," *International Journal of Heat and Mass Transfer*, Vol. 45, pp. 3275-3286, 2002.
- [55] I. Tiselj, G. Hetsroni, B. Mavko, A. Mosyak, E. Pogrebnyak and Z. Segal, "Effect of axial conduction on the heat transfer," *International Journal of Heat and Mass Transfer*, Vol. 47, pp. 2551-2565, 2004.
- [56] W. L. Qu, G. M. Mala and D. Q. Li, "Pressure-driven water flows in trapezoidal silicon microchannels," *International Journal of Heat and Mass Transfer*, Vol. 43, pp. 353-364, 2000.
- [57] M. J. Kohl, S. I. Abdel-Khalik, S. M. Jeter, and D. L. Sadowski, "A microfluidic experimental platform with internal pressure measurements," *Sensors and Actuators A*, Vol. 118, pp. 212-221, 2005.
- [58] M. J. Kohl, S. I. Abdel-Khalik, S. M. Jeter, and D. L. Sadowski, "An experimental investigation of microchannel flow with internal pressure measurements," *International Journal of Heat and Mass Transfer*, Vol. 48, pp. 1518-1533, 2005.
- [59] J. M. Commenge, L. Falk, J. P. Corriou, and M. Matlosz, "Optimal design for flow uniformity in microchannel reactors," *AIChE Journal*, Vol. 48, pp. 345-358, 2002.
- [60] E. R. Delsman, A. Pierik, M. H., J. H. De Croon, G. J. Kramer, and J. C. Schouten, "Microchannel plate geometry optimization for even flow distribution at high flow rates," *Chemical Engineering Research and Design*, Vol. 82, pp. 267-273, 2004.
- [61] G. Griffini and A. Gavriilidis, "Effect of microchannel plate design on fluid flow uniformity at low flow rates," *Chemical Engineering Technology*, Vol. 30, pp. 395-406, 2007.
- [62] J. T. Teng, J. C. Chu, M. S. Liu, C. C. Wang and R. Greif, "Investigation of the flow maldistribution in microchannels," *ASME International Mechanical Engineering Congress (IMECE03)*, Washington, D.C., November 15-21, 2003.

- [63] D. V. Pence, "Reduced pumping power and wall temperature in microchannel heat sinks with fractal-like branching channel networks," *Microscale Thermophysical Engineering*, Vol. 6, pp. 319-330, 2002.
- [64] Y. Chen and P. Cheng, "Heat transfer and pressure drop in fractal tree-like microchannel nets," *International Journal of Heat and Mass Transfer*, Vol. 45, pp. 2643-2648, 2002.
- [65] S. M. Senn and D. Poulikakos, "Laminar mixing, heat transfer and pressure drop in tree-like microchannel nets and their application for thermal management in polymer electrolyte fuel cells," *Journal of Power Sources*, Vol. 130, pp.178-191, 2004.
- [66] A. Y. Alharbi, D. V. Pence, and R. N. Cullion, "Thermal characteristics of microscale fractal-like branching channels," *Transactions of the ASME*, Vol. 126, pp. 744-752, 2004.
- [67] Y. Chen and P. Cheng, "An experimental investigation on the thermal efficiency of fractal tree-like microchannels," *International Communications in Heat and Mass Transfer*, Vol. 32, pp. 931-938, 2005.
- [68] N. Kockmann, M. Engler, D. Haller, and P. Wolas, "Fluid Dynamics and Transfer Processes in Bended Microchannels," *Heat Transfer Engineering*, Vol. 26, pp. 71-78, 2005.
- [69] L. Ghodoossi, "Thermal and hydrodynamic analysis of a fractal microchannel network," *Energy Conversion and Management*, Vol. 46, pp. 771-788, 2005.
- [70] X. Q. Wang, A. S. Mujumdar, and C. Yap, "Thermal characteristics of tree-shaped microchannel nets for cooling of a rectangular heat sink," *International Journal of Thermal Sciences*, Vol. 45, pp. 1103-1112, 2006.
- [71] F. J. Hong, P. Cheng, H. Ge, and G. T. Joo, "Conjugate heat transfer in fractal-shaped microchannel network heat sink for integrated microelectronic cooling application," *International Journal of Heat Mass Transfer*, Vol. 50, pp. 4986-4998, 2007.
- [72] B. Mandelbrot and B. Benoit, *The Fractal Geometry of Nature*, W. H. Freeman and Company, New York, 1982.
- [73] A. Bejan and M. R. Errera, "Deterministic tree networks for fluid flow: geometry for minimal flow resistance between a volume and one point," *Fractals*, Vol. 5, pp. 685-695, 1997.
- [74] S. Lorente, W. Wechsato, and A. Bejan, "Tree-shaped flow structure designed by minimizing path lengths", *International Journal of Heat and Mass Transfer*, Vol. 45, pp. 3299-3312, 2002.
- [75] C. D. Murray, "The Physiological principle of minimal work in vascular system, and the cost of blood-volume," *Proceedings of the National Academy of Sciences*, Vol. 12, pp. 207-214, 1926.
- [76] C. Liu, J. T. Teng, J. C. Chu, Y. L. Chiu, S. Y. Huang, S. P. Jin, T. T. Dang, R. Greif, and H. H. Pan, "Experimental investigations on liquid flow and heat transfer in rectangular microchannel with longitudinal vortex generators," *International Journal of Heat and Mass Transfer*, Vol. 54, pp. 3069-3080, 2011.

Part 4

Medical and Biomechanical Applications

Modelling Propelling Force in Swimming Using Numerical Simulations

Daniel A. Marinho^{1,2}, Tiago M. Barbosa^{2,3}, Vishveshwar R. Mantha^{2,4},
Abel I. Rouboa^{2,5} and António J. Silva^{2,4}

¹*University of Beira Interior, Department of Sport Sciences, Covilhã*

²*Research Centre in Sports, Health and Human Development, Vila Real*

³*Polytechnic Institute of Bragança, Department of Sport Sciences, Bragança*

⁴*University of Trás-os-Montes and Alto Douro, Department of Sport Sciences,
Exercise and Health, Vila Real*

⁵*University of Trás-os-Montes and Alto Douro,
Department of Engineering, Vila Real
Portugal*

1. Introduction

In the sports field, numerical simulation techniques have been shown to provide useful information about performance and to play an important role as a complementary tool to physical experiments. Indeed, this methodology has produced significant improvements in equipment design and technique prescription in different sports (Kellar et al., 1999; Pallis et al., 2000; Dabnichki & Avital, 2006). In swimming, this methodology has been applied in order to better understand swimming performance. Thus, the numerical techniques have been addressed to study the propulsive forces generated by the propelling segments (Rouboa et al., 2006; Marinho et al., 2009a) and the hydrodynamic drag forces resisting forward motion (Silva et al., 2008; Marinho et al., 2009b).

Although the swimmer's performance is dependent on both drag and propulsive forces, within this chapter the focus is only on the analysis of the propulsive forces. Hence, this chapter covers topics in swimming propelling force analysis from a numerical simulation technique perspective. This perspective means emphasis on the fluid mechanics and computational fluid dynamics methodology applied in swimming investigations. One of the main aims for performance (velocity) enhancement of swimming is to maximize propelling forces whilst not increasing drag forces resisting forward motion, for a given trust. This chapter will concentrate on numerical simulation results, considering the scientific simulation point-of-view, for this practical application in swimming.

Basically, numerical simulations consist of a mathematical model that replaces the Navier-Stokes equations with discretized algebraic expressions that can be solved by iterative computerized calculations. The Navier-Stokes equations describe the motion of viscous non-compressible fluid substances. These equations arise from applying Newton's second law to fluid motion, together with the assumption that the fluid stress is the sum of a diffusing viscous term (proportional to the gradient of velocity), plus a pressure term. A

solution of the Navier–Stokes equations is called a velocity field or flow field, which is a description of the velocity of the fluid at a given point in space and time. Numerical simulations are based on the finite volume approach, where the equations are integrated over each control volume. It is required to discretize the spatial domain into small cells to form a volume mesh or grid, and then apply a suitable algorithm to solve the equations of motion (Marinho et al., 2010a). Additionally, several studies reported the importance of numerical simulations on testing and experimentation, reducing the total effort required in the experimental design and data acquisition (Lyttle & Keys, 2006; Bixler et al., 2007). For instance, Lyttle and Keys (2006) referred that these numerical simulations can provide answers into many complex problems that have been unobtainable using physical testing techniques. One of its major benefits is the possibility to test many variations until one arrives at an optimal result, without physical/experimental testing.

Although some doubts on the accuracy of numerical simulations results, this numerical tool has been validated as being feasible in modelling complicated biological fluid dynamics, through a series of stepwise baseline benchmark tests and applications for realistic modelling of different scopes for hydro and aerodynamics of locomotion (Liu, 2002). Bixler et al. (2007) studied the accuracy of numerical analysis of the passive drag of a male swimmer. Comparisons of total drag force were performed between a real swimmer, a digital model of this same swimmer and a real mannequin based on the digital model. Bixler et al. (2007) found drag forces determined from the digital model using numerical simulations to be within 4% of the values assessed experimentally for the mannequin, although the mannequin drag was found to be 18% less than the real swimmer drag. In fact, this study has reinforced the idea of the validity and accuracy of numerical simulations in swimming research. Some differences were due to little body movements during the gliding position and to differences between the swimmer and the model since the swimmer's skin is flexible while the mannequin's skin is rigid. So, it is usually assumed that numerical simulations have ecological validity even for swimming research.

In the first part of the chapter, we introduce the issue, the main aims of the chapter and a brief explanation of the computational fluid dynamics methodology. Then, the contribution of different studies for swimming using numerical studies and some practical applications of this methodology are presented. During the chapter the authors will attempt to present the computational fluid dynamics data and to address some practical concerns to swimmers and coaches, comparing as well the numerical data with other experimental data available in the literature.

2. Numerical simulation methodology

2.1 Mathematical formulation

The flow around the swimmer seems to be turbulent (Sanders, 2001; Toussaint & Truijens, 2005). Due to that reason Reynolds averaged Navier–Stokes equations with the Boussinesq hypothesis to model the Reynolds stresses are usually used (Hinze, 1975). The closure problem of the turbulent modeling is solved using k-epsilon model with appropriate wall functions. The system of equations for solving three-dimensional, incompressible fluid flow in steady-state regime is as follows:

$$\text{div}V = 0 \tag{1}$$

$$\frac{\partial V}{\partial t} \pm V \cdot \nabla V + \nabla p \pm \nabla \left(v + c_\mu \frac{k^2}{\varepsilon} \right) (\nabla V + \nabla V^t) = 0 \quad (2)$$

$$\frac{\partial(\rho k)}{\partial t} + \frac{\partial(\rho V_x k)}{\partial x} + \frac{\partial(\rho V_y k)}{\partial y} + \frac{\partial(\rho V_z k)}{\partial z} = \frac{\partial \left(\frac{\mu_t}{\sigma_k} \frac{\partial k}{\partial x} \right)}{\partial x} + \frac{\partial \left(\frac{\mu_t}{\sigma_k} \frac{\partial k}{\partial y} \right)}{\partial y} + \frac{\partial \left(\frac{\mu_t}{\sigma_k} \frac{\partial k}{\partial z} \right)}{\partial z} + \mu_t \Phi - \rho \varepsilon \quad (3)$$

$$\frac{\partial(\rho k)}{\partial t} + \frac{\partial(\rho V_x \varepsilon)}{\partial x} + \frac{\partial(\rho V_y \varepsilon)}{\partial y} + \frac{\partial(\rho V_z \varepsilon)}{\partial z} = \frac{\partial \left(\frac{\mu_t}{\sigma_\varepsilon} \frac{\partial \varepsilon}{\partial x} \right)}{\partial x} + \frac{\partial \left(\frac{\mu_t}{\sigma_\varepsilon} \frac{\partial \varepsilon}{\partial y} \right)}{\partial y} + \frac{\partial \left(\frac{\mu_t}{\sigma_\varepsilon} \frac{\partial \varepsilon}{\partial z} \right)}{\partial z} + \mu_t \frac{\varepsilon}{k} \Phi - C_2 \frac{\rho \varepsilon^2}{k} \quad (4)$$

Where k is the turbulent kinetic energy and ε is the turbulent kinetic energy dissipation ratio. V_x , V_y and V_z represent the x, y and z components of the velocity V . μ_t is the turbulent viscosity and ρ represents the fluid density. ν is the kinematic viscosity, Φ is the pressure strain, C_2 , C_μ , σ_ε and σ_k are model constants, 1.92, 0.09, 1.30 and 1.00, respectively.

The detailed terms of the k-epsilon model transport equations used during simulations are provided in user manual of Fluent documentation (Fluent, 2006).

2.2 Digital model

One of the main tasks to carryout numerical simulations is to define the digital model of the swimmer. The majority of the studies performed on this field used segments of the human body, representative of propelling segments, i.e., the hand and the forearm (Bixler & Schloder, 1996; Rouboa et al., 2006).

In order to create the three-dimensional digital model computer tomography scans of a hand and forearm segments of an Olympic swimmer were applied. With these data we converted the values into a format that can be read in *Gambit*, *Fluent*[®] pre-processor. *Fluent*[®] software is used to simulate the fluid flow around structures, allowing the analysis of values of pressure and speed around (i.e. the hand and forearm of a swimmer). With these values we can calculate force components through integration of pressures on the hand/forearm surfaces, using a realistic model of these human segments.

Eighteen cross-sectional scans of the right arm (hand and forearm) were obtained using a Toshiba[®] Aquilion 4 computer tomography scanner. Computer tomography scans were obtained with configuration of V2.04 ER001. A 2 mm slice thickness with a space of 1 mm was used. The subject was an Olympic level swimmer, who participated in the 2004 Olympic Games in Athens. The subject was lying with his right arm extended upwards and fully pronated. The thumb was adducted and the wrist was in a neutral position (Marinho et al., 2010b). The appropriate ethical committee of the institution in which it was performed has approved this protocol, and the subject consented to participate in this work.

The transformation of values from the computer tomography scans into nodal coordinates in an appropriate coordinate system demands the use of image processing techniques. The image-processing program used in this study was the *Anatomics Pro*[®]. This program allowed obtaining the boundaries of the human segments, creating a three-dimensional reconstruction of the swimmer hand and forearm.

At first, before processing and converting procedures the data was prepared, namely by observing the computer tomography data and erasing the non-relevant parts of the anatomical model. For example, surfaces supporting the subject were also scanned, reason

why it had to be defined the relevant points and deleted the irrelevant ones. This step was also conducted using the software FreeForm Sensable®. Finally, the data was converted into an IGES format (*.igs), that could be read by Gambit/Fluent® to define the finite volume approach through the three-dimensional surfaces (Marinho et al., 2010b).

2.3 Simulations

The dynamic fluid forces produced by the hand, lift (L) and drag (D), were measured in this study. These forces are functions of the fluid velocity and they were measured by the application of the equations 5 and 6, respectively:

$$D = C_D \frac{1}{2} \rho A v^2 \quad (5)$$

$$L = C_L \frac{1}{2} \rho A v^2 \quad (6)$$

In equations 5 and 6, v is the fluid velocity, C_D and C_L are the drag and lift coefficients, respectively, ρ is the fluid density and A is the projection area of the model for different angles of attack used in this study.

The whole domain was meshed with a hybrid mesh composed of prisms and pyramids. Significant efforts were conducted to ensure that the model would provide accurate results by decreasing the grid node separation in areas of high velocity and pressure gradients.

Angles of attack of hand models of 0° , 15° , 30° , 45° , 60° , 75° and 90° , with a sweep back angle of 0° (thumb as the leading edge) were used for the calculations (Schleihauf, 1979).

Steady-state analyses were performed using the Fluent® code and the drag and lift coefficients were calculated for a flow velocity of 2.0 m/s (Lauder et al., 2001; Rouboa et al., 2006).

We used the segregated solver with the standard k-epsilon turbulence model because this turbulence model was shown to be accurate with measured values in a previous research (Moreira et al., 2006).

All numerical computational schemes were second-order, which provides a more accurate solution than first-order schemes. We used a turbulence intensity of 1.0% and a turbulence scale of 0.10 m. The water temperature was 28° C with a density of $998.2 \text{ kg}\cdot\text{m}^{-3}$ and a viscosity of $0.001 \text{ kg}/\text{m}\cdot\text{s}$. Incompressible flow was assumed. The measured forces on the hand models were decomposed into drag (C_D) and lift (C_L) coefficients, using equations 5 and 6.

3. Propelling force

3.1 Definition

Swimming is characterized by the intermittent application of a propulsive force (thrust) to overcome a velocity-dependent water resistance (hydrodynamic drag). The thrust is generated by a combination of arm, leg and body movements and lead to variations of thrust and velocity. Different fluctuations in thrust, drag and velocity among different techniques and different level of skills contribute to the highly variable performance in swimming. Swimming performance can be studied by analysing the interaction of propelling and resistive forces. In this sense, a swimmer will only enhance performance by minimizing resistive forces that act on the swimming body at a given velocity and/or by increasing the propulsive forces produced by the propelling segments. Furthermore, a third performance-enhancing factor would be to do this with a minimal enhancement of physiological or energetic costs (Barbosa et al., 2010).

Effective propelling force can be defined as the component of the total propulsive force acting in the direction of moving. This force is produced due to the interaction of the swimmer with the water allowing overcoming hydrodynamic drag forces resisting forward motion. Thus, it is a hydrodynamic force with the same direction of the movement but opposite to drag force.

There are several mechanisms responsible to produce propelling forces, although some of them seemed to be more efficient than others. This fact seems to distinguish swimmers of different level, since some mechanisms allow producing the same work with less energy consumption (Barbosa et al., 2010). Knowing the most effective way to produce propulsive force should play a major role in swimming training.

3.2 Relative contribution of drag and lift forces

The relative contribution of drag and lift forces to overall propulsion is one of the most discussed issues in swimming hydrodynamics research.

Bixler and Riewald (2002) evaluated the steady flow around a swimmer's hand and forearm at various angles of attack and sweep back angles. The digital model was created based upon an adult male's right forearm and hand with the forearm fully pronated using similar procedures abovementioned. Force coefficients measured as a function of angle of attack showed that forearm drag was essentially constant and forearm lift was almost null. Additionally, hand drag presented the minimum value near angles of attack of 0° and 180° and the maximum value was obtained near 90° , when the model is nearly perpendicular to the flow. Hand lift was almost zero at 95° and presented the highest values near 60° and 150° . Regarding the water channel analysis, Schleihau (1979) reported that lift coefficient values increased up to an attack angle around 40° and then decreased, although some differences with respect to the sweepback angle were observed. Drag coefficient values increased with increasing the attack angle and were less sensitive to sweepback angle changes.

Silva et al. (2008), using a real digital model of a swimmer hand and forearm, confirmed the supremacy of the drag component. They also revealed an important contribution of lift force to the overall propulsive force production by the hand/forearm in swimming phases, when the angle of attack is close to 45° . The drag coefficient presented higher values than the lift coefficient for all angles of attack. In fact, the drag coefficient increased with the angle of attack showing the maximum values with an angle of attack of 90° and the minimum values with an angle of attack of 0° . The lift coefficient of the model presented the maximum values with an angle of 45° . Silva et al. (2008) obtained values of lift coefficient very similar for the angles of attack of 0° and 90° , although the minimum values were obtained with an angle of attack of 90° . Sato and Hino (2002) showed values of drag coefficient higher than lift coefficient at all angles of attack. From the results of the simulations the authors suggested that the resultant force was maximal with an angle of attack of 105° and the direction of the resultant force in that situation was -13° . Based on this analysis, the authors suggested stroke backward and with a little-finger-ward, out sweep motion, as the best stroke motion to produce the maximum thrust during underwater path.

Another interesting numerical report, when the sweep back angle is considered, was that more lift force is generated when the little finger leads the motion than when the thumb leads (Bixler & Riewald, 2002; Silva et al. 2008).

3.3 Contribution of arms and legs to propulsion

Another important issue is related to the contribution of arms and legs to propulsion. It is almost consensual that most propulsion is generated by the arms' actions. In front crawl

swimming, it was found (Hollander et al., 1988; Deschodt, 1999) that about 85 to 90% of propulsion is produced by the arms' movements. Accordingly, the majority of the research under this scope is performed on arm's movements. Nevertheless, leg's propulsion should not be disregarded and future studies under this field should be addressed, helping swimmers to enhance performance. In this section it seems pertinent to present some interesting and pioneer studies of Lyttle and Keys (2006). These authors performed a three-dimensional numerical analysis, modelling the swimmer performing two kinds of underwater dolphin kick used after starts and turns, after the swimmer push-off from the wall: (i) high amplitude and low frequency dolphin kick and; (ii) low amplitude and high frequency dolphin kick. Main data demonstrated an advantage of using the large slow kick, over the small fast kick, concerning the velocity range that underwater dolphin kicks are used. In addition, changes were also made into the input kinematics (ankle plantar flexion angle) to demonstrate the practical applicability of the model. While the swimmer was gliding at 2.18 m/s, a 10° increase in ankle plantar flexion created greater propulsive force during the kick cycle. These results demonstrated that increasing angle flexibility would increase the stroke efficiency for the subject that was modelled, although some caution should be made when transferring this data into other swimmers with different anthropometrical profile.

Regarding arms' propulsion, Lecrivain et al. (2008) reported that the arm (and not only the hand and forearm) provided effective propulsion through most of the stroke, and this must be considered when studying the arm propulsion. In fact, Gardano and Dabnichki (2006) underlined the importance of the analysis of the entire arm rather than different parts of it. Thus, the authors concluded that drag profiles differed substantially with the elbow flexion angle, as the maximum value could vary by as much as 40%. In addition, Gardano and Dabnichki (2006) stated that maximum drag force was achieved by 160° of elbow angle. A prolonged plateau between 50° and 140° indicated greater momentum generated at 160° in comparison with the other configurations. This fact suggests a strong possibility for the existence of an optimal elbow angle for the generation of a maximum propulsive force. However, these findings are only possible to confirm if an entire model of the swimmer's arm, its movement relative to the body and the body's movement relative to the water are computed (Marinho et al., 2009c). This concern seems also an interesting topic to address in further studies.

3.4 Fingers relative position

Regarding arms' actions, a large inter-subject range of fingers relative position can be observed during training and competition, regarding thumb position and finger spreading. Due to the inherent inefficiency of human swimming, the question is: do any of these strategies enhance performance or is it just a more comfortable hand posture that swimmers assumed?

Regarding thumb position, although some differences in the results of different studies (Schleihauf, 1979; Takagi et al., 2001; Marinho et al., 2009a), main data seemed to indicate that when the thumb leads the motion (sweep back angle of 0°) a hand position with the thumb abducted would be preferable to an adducted thumb position. Additionally, Marinho et al. (2009a) found, for a sweep back angle of 0°, that the position with the thumb abducted presented higher values than the positions with the thumb partially abducted and adducted at angles of attack of 0° and 45°. At an angle of attack of 90°, the position with the thumb adducted presented the highest value of resultant force. Schleihauf (1979), using experimental procedures, found that the position with the thumb fully abducted showed a maximum lift coefficient at an angle of attack of 15°, whereas the models with partial thumb abduction

showed a maximum value of lift coefficient at higher angles of attack (45° - 60°). In these orientations, the position with the thumb partially abducted presented higher values than with the thumb fully abducted. Moreover, Takagi et al. (2001) also applying experimental measurements revealed that the thumb position influenced the lift force. For a sweep back angle of 0° (as used in the study of Marinho et al., 2009a) the model with abducted thumb presented higher values of lift force, whereas for a sweep back angle of 180° (the little finger as the leading edge), the adducted thumb model presented higher values of lift force. In addition, the drag coefficient presented similar values in the two thumb positions for a sweep back angle of 0° and higher values in the thumb adducted position for a sweep back angle of 180° . Regarding different finger spreading, Marinho et al. (2010c), using a numerical analysis, studied the hand with: (i) fingers close together, (ii) fingers with little distance spread (a mean intra finger distance of 0.32 cm, tip to tip), and (iii) fingers with large distance spread (0.64 cm, tip to tip), following the same procedure when Schleihauf (1979) conducted his experimental research. Marinho et al. (2010c) found that for attack angles higher than 30° , the model with little distance between fingers presented higher values of drag coefficient when compared with the models with fingers closed and with large finger spread. For attack angles of 0° , 15° and 30° , the values of drag coefficient were very similar in the three models of the swimmer's hand. Moreover, the lift coefficient seemed to be independent of the finger spreading, presenting little differences between the three models. Nevertheless, Marinho et al. (2010c) were able to note slightly lower values of lift coefficient for the position with larger distance between fingers. In the same line of research, Minetti et al. (2009) showed, through numerical simulation of a three-dimensional model of the hand, that an optimal finger spacing (12° , roughly corresponding to the resting hand posture) increases the drag coefficient (+8.8%), which is 'functionally equivalent' to a greater hand palm area, thus a lower stroke frequency can produce the same thrust, with benefits to muscle, hydraulic and propulsive efficiencies. These results suggested that the hand seems to create more propulsive force when fingers are slightly spread. Flow visualization, through numerical simulations, provides an explanation for the increased force associated with the optimum finger spacing.

3.5 Steady vs. unsteady flow conditions

The majority of the abovementioned studies were conducted only under steady state flow conditions. However, one knows (Schleihauf, 1979) that swimmers do not move their arms/hands under constant velocity and direction motions. Therefore, some authors (Sanders, 1999; Bixler & Riewald, 2002; Sato & Hino, 2002; Rouboa et al., 2006) referred that it is important to consider unsteady effects when swimming propulsion is analysed. For instance, Bixler and Schloder (1996) analysed the flow around a disc with a similar area of a swimmer hand. Different simulations with different initial velocity and acceleration were conducted to model identical real swimming conditions, especially during insweep and upsweep phases of the front crawl stroke. According to the obtained results the authors reported that the hand acceleration could increase the propulsive force by around 24% compared with the steady flow condition. Sato and Hino (2002) using also numerical and experimental data showed that the hydrodynamic forces acting on the accelerating hand was much higher than with a steady flow situation and these forces amplifies as acceleration increases. Rouboa et al. (2006) analysed the effect of swimmer's hand/forearm acceleration on propulsive forces generation using numerical simulation techniques. The main data underlined that under the hand/forearm acceleration condition, the measured values for propulsive forces were approximately 22.5% higher than the forces produced under the steady flow condition. Thus, these data suggests that drag and lift forces produced by the

swimmers' hand in a determined time are dependent not only on the surface area, the shape and the velocity of the segment but also on the acceleration of the propulsive segment.

3.6 Equipment

Research in this scope is very scarce and when applied the main focus is related to analyse how can different equipment improve swimming performance due to a decrease in hydrodynamic drag (Neiva et al., 2011). The study of the effects of using different swimsuits is a good example (Roberts et al., 2003; Pendergast et al., 2006), emphasising the importance of compression effects due to swimsuits on drag reduction (Neiva et al., 2011). However, to the best of our knowledge, the effects on improving propelling force had lower attention by swimming scientific community. Some equipment used by swimmers, especially during training, can be tested using numerical simulation techniques, allowing understand the specific effects on propelling force. For instance, the effects of wearing fins, and different types of fins (Tamura et al., 2002), paddles, and other devices using during training should be tested attempting to elucidate coaches to improve training efficiency.

4. Future research in swimming using numerical simulations

Throughout this chapter, several future ideas have been presented to improve the application of numerical simulations in swimming research. One of our major aims is to be able to evaluate biomechanical situations that can be used by coaches and swimmers to swim faster and, thus to enhance performance. Therefore, the effective evaluation of true swimming conditions should be a main focus. Under this scope, main concerns should be addressed to analyse unsteady flow conditions, studying arms and legs propulsion during actual swimming, adding body roll, movement of the body, rotations and accelerations of the propelling segments, on different swimming techniques.

As mentioned above, the analysis of the effects of different equipment and facilities on propelling forces seems to be an interesting and an important issue to be dealt in future studies.

5. Conclusion

During this chapter, the authors attempted to present some important studies that have been conducted in swimming research using numerical simulations. Although there are some limitations of these studies, it seems that this numerical tool should not be disregarded. Numerical simulations can be used to evaluate several hydrodynamic issues, hence helping swimmers moving faster. In the current work some issues regarding the effect of propelling force on swimming performance were discussed. The authors are aware of some limitations, although they believe that they were able to show the practical applications of numerical simulations to swimmers and their coaches.

Moreover, it was an attempt to address some concerns to be improved in future investigations.

6. Acknowledgement

This work was supported by the Portuguese Government by Grants of the Science and Technology Foundation (PTDC/DES/098532/2008).

7. References

- Barbosa, T.M.; Bragada, J.A.; Reis, V.M.; Marinho, D.A.; Carvalho, C. & Silva, J.A. (2010). Energetics and biomechanics as determining factors of swimming performance: updating the state of the art. *Journal of Science and Medicine in Sports*, 13, 262-269
- Bixler, B. & Schloder, M. (1996). Computational fluid dynamics: an analytical tool for the 21st century swimming scientist. *Journal of Swimming Research*, 11, 4-22.
- Bixler, B.S. & Riewald, S. (2002). Analysis of swimmer's hand and arm in steady flow conditions using computational fluid dynamics. *Journal of Biomechanics*, 35, 713-717.
- Bixler, B.; Pease, D. & Fairhurst, F. (2007). The accuracy of computational fluid dynamics analysis of the passive drag of a male swimmer. *Sports Biomechanics*, 6, 81-98.
- Dabnichki, P. & Avital, E. (2006). Influence of the position of crew members on aerodynamics performance of two-man bobsleigh. *Journal of Biomechanics*, 39, 2733-2742.
- Deschodt, V. (1999). Relative contribution of arms and legs in human to propulsion in 25 m sprint front crawl swimming. *European Journal of Applied Physiology*, 80, 192-199.
- Fluent (2006). *Fluent 6.3 Documentation*. Fluent Inc., Hanover.
- Gardano, P. & Dabnichki, P. (2006). On hydrodynamics of drag and lift of the human arm. *Journal of Biomechanics*, 39, 2767-2773.
- Hinze, J.O. (1975). *Turbulence*. McGraw-Hill Publishing Co., New York.
- Hollander, A.P.; de Groot, G.; van Ingen Schenau, G.; Kahman, R. & Toussaint, H. (1988). Contribution of the legs to propulsion in Front Crawl swimming. In: *Swimming Science V*, B. Ungerechts, K. Wilke & K. Reischle, (Eds.), 39-43, Human Kinetics Books, Illinois.
- Kellar, W.P.; Pearse, S.R.G. & Savill, A.M. (1999). Formula 1 car wheel aerodynamics. *Sports Engineering*, 2, 203-212.
- Lauder, M.; Dabnichki, P. & Bartlett, R. (2001). Improved accuracy and reliability of sweepback angle, pitch angle and hand velocity calculations in swimming. *Journal of Biomechanics*, 34, 31-39.
- Lecrivain, G.; Slaouti, A.; Payton, C. & Kennedy, I. (2008). Using reverse engineering and computational fluid dynamics to investigate a lower arm amputee swimmer's performance. *Journal of Biomechanics*, 41, 2855-2859.
- Liu, H. (2002). Computational biological fluid dynamics: digitizing and visualizing animal swimming and flying. *Integrative and Comparative Biology*, 42, 1050-1059.
- Lyttle, A. & Keys, M. (2006). The application of computational fluid dynamics for technique prescription in underwater kicking. *Portuguese Journal of Sport Sciences*, 6, Suppl. 2, 233-235.
- Marinho, D.A.; Rouboa, A.I.; Alves, F.B.; Vilas-Boas, J.P.; Machado, L.; Reis, V.M. & Silva, A.J. (2009a). Hydrodynamic analysis of different thumb positions in swimming. *Journal of Sports Science and Medicine*, 8, 1, 58-66.
- Marinho, D.A.; Reis, V.M.; Alves, F.B.; Vilas-Boas, J.P.; Machado, L.; Silva, A.J. & Rouboa, A.I. (2009b). The hydrodynamic drag during gliding in swimming. *Journal of Applied Biomechanics*, 25, 3, 253-257.
- Marinho, D.A.; Barbosa, T.M.; Kjendlie, P.L.; Vilas-Boas, J.P.; Alves, F.B.; Rouboa, A.I. & Silva, A.J. (2009c). Swimming simulation: a new tool for swimming research and practical applications. In: *Lecture Notes in Computational Science and Engineering - CFD for Sport Simulation*, M. Peters (Ed.), 33-62. Springer, Berlin.
- Marinho, D.A.; Barbosa, T.M.; Kjendlie, P.L.; Mantripragada, N.; Vilas-Boas, J.P.; Machado, L.; Alves, F.B.; Rouboa, A.I. & Silva, A.J. (2010a). Modeling hydrodynamic drag in

- swimming using computational fluid dynamics, In: *Computational Fluid Dynamics*, H.W. Oh (Ed.), 391-404. INTECH Education and Publishing, Vienna.
- Marinho, D.A.; Reis, V.M.; Vilas-Boas, J.P.; Alves, F.B.; Machado, L.; Rouboa, A.I. & Silva, A.J. (2010b). Design of a three-dimensional hand/forearm model to apply Computational Fluid Dynamics. *Brazilian Archives of Biology and Technology*, 5(2), 437-442.
- Marinho, D.A.; Barbosa, T.M.; Reis, V.M.; Kjendlie, P.L.; Alves, F.B.; Vilas-Boas, J.P.; Machado, L.; Silva, A.J. & Rouboa, A.I. (2010c). Swimming propulsion forces are enhanced by a small finger spread. *Journal of Applied Biomechanics*, 26, 87-92.
- Minetti, A.E.; Machtsiras, G. & Masters, J.C. (2009). The optimum finger spacing in human swimming. *Journal of Biomechanics*, 42, 13, 2188-2190
- Moreira, A.; Rouboa, A.; Silva, A.; Sousa, L.; Marinho, D.; Alves, F.; Reis, V.; Vilas-Boas, J.P.; Carneiro, A. & Machado, L. (2006). Computational analysis of the turbulent flow around a cylinder. *Portuguese Journal of Sport Sciences*, 6(Suppl. 1), 105.
- Neiva, H.P.; Vilas-Boas, J.P.; Barbosa, T.M.; Silva, A.J. & Marinho, D.A. (2011). 13th FINA World Championships: Analysis of Swimsuits Used By Elite Male Swimmers. *Journal of Human Sport and Exercise*, 6, 1, 87-93.
- Pallis, J.M.; Banks, D.W. & Okamoto, K.K. (2000). 3D computational fluid dynamics in competitive sail, yacht and windsurfer design, In: *The Engineering of Sport: Research, Development and Innovation*, F. Subic & M. Haake (Eds.), 75-79. Blackwell Science, Oxford.
- Pendergast, D.R.; Capelli, C.; Craig Jr, A.B.; di Prampero, P.E.; Minetti, A.E.; Mollendorf1, J.; Termin, A. & Zamparo, P. (2006). Biophysics in swimming. *Portuguese Journal of Sport Sciences*, 6, Suppl. 2, 185-189.
- Roberts, B.S.; Kamel, K.S.; Hedrick, C.E.; MLean, S.P. & Sharpe, R.L. (2003). Effect of Fastskin suit on submaximal freestyle swimming. *Medicine and Science in Sports and Exercise*, 35, 519-524.
- Rouboa, A.; Silva, A.; Leal, L.; Rocha, J. & Alves, F. (2006). The effect of swimmer's hand/forearm acceleration on propulsive forces generation using computational fluid dynamics. *Journal of Biomechanics*, 39, 1239-1248.
- Sanders, R.H. (1999). Hydrodynamic characteristics of a swimmer's hand. *Journal of Applied Biomechanics*, 15, 3-26.
- Sanders, R.B.; Rushall, H., Toussaint, H., Steager, J. & Takagi, H. (2001). Bodysuit yourself but first think about it. *American Swimming Magazine*, 5, 23-32.
- Sato, Y & Hino, T. (2002). Estimation of thrust of swimmer's hand using CFD. In: *Proceedings of 8th symposium on nonlinear and free-surface flows*, 71-75. Hiroshima.
- Schleihauf, R.E. (1979). A hydrodynamic analysis of swimming propulsion. In: *Swimming III*, J. Terauds & E.W. Bedingfield (Eds.), 70-109. University Park Press, Baltimore.
- Silva, A.J.; Rouboa, A.; Moreira, A.; Reis, V.; Alves, F.; Vilas-Boas, J.P. & Marinho, D. (2008). Analysis of drafting effects in swimming using computational fluid dynamics. *Journal of Sports Science and Medicine*, 7, 1, 60-66.
- Takagi, H.; Shimizu, Y.; Kurashima, A. & Sanders, R. (2001). Effect of thumb abduction and adduction on hydrodynamic characteristics of a model of the human hand. In: *Proceedings of Swim Sessions of the XIX International Symposium on Biomechanics in Sport*, J. Blackwell & R. Sanders (Eds.), 122-126. University of San Francisco, San Francisco.
- Tamura, H.; Nakazawa, Y.; Sugiyama, Y.; Nomura, T. & Torii, N. (2002). Motion analysis and shape evaluation of a swimming monofin. *The Engineering Sports*, 4, 716-724.
- Toussaint, H. & Truijens, M. (2005). Biomechanical aspects of peak performance in human swimming. *Animal Biology*, 55, 1, 17-40.

Surfactant Analysis of Thin Liquid Film in the Human Trachea via Application of Volume of Fluid (VOF)

Sujudran Balachandran
Bumi Armada Berhad
Malaysia/Singapore

1. Introduction

The human tracheobronchial tree is a complex branched distribution system in charge of renewing the air inside the acini which are the gas exchange units. The surfactant factor existing in the acini of the human tracheobronchial tree is exposed to thin liquid film flow where the application of volume of fluid (VOF) can significantly determine the membrane effects on the branching asymmetry. The thin film application in this chapter will focus on the breathing airway which is commonly known as the windpipe (trachea) of an adult human. The upper human airway is the primary conduit for inspiration in the breathing process. Air entering the mouth passes through pharynx and flows into the trachea via the glottal region. The air which enters the windpipe applies a surfactant pressure at the lipo-protein complex, where this complex is developed as a thin layer on the trachea (Caro et al. 2002). The surfactant is a lipo-protein complex, which is a highly surface-active material found in the fluid lining of the air-liquid interface in the trachea surface. This surfactant plays a dual function of preventing sudden collapse during the breathing cycle and protection from injuries and infections caused by foreign bodies and pathogens. The varying degrees of structure-function abnormalities of surfactant have been associated with obstructive trachea diseases, respiratory infections, respiratory distress syndromes, interstitial lung diseases, pulmonary alveolar proteinosis, cardiopulmonary bypass surgery and smoking habits. For some of the pulmonary conditions, especially respiratory distress syndrome, surfactant therapy is on the horizon. In order to understand the behaviour and relevant condition of the surfactant in the human trachea, it is important to apply the volume of fluid method on these surfaces. The phenomena that occur on the trachea will ensure that the surfactant responsibility in resolving the potential obstruction of breathing. The surfactant factor may occur with non-lateral conditions in space as well as during the inspiration of breathing in the human body. The fluid interaction in the thin surface results in serious impairment by obstructive trachea diseases as mentioned earlier. The pulmonary surfactant is essential for normal breathing, alveolar stability and as a host defence system in the lungs. The interface of surfactant films reduces the surface tension to extremely low values when it is compressed during expiration. This protects our lungs from collapse during breathing out. Thus application of the Volume of Fluid (VOF) method is introduced in this paper to study the behaviour of the pulmonary surfactant in the human trachea.

2. Model selection

Figure 1 shows the three-dimensional model of the lung. Detailed geometries of this model were extracted from the anatomical model by Schmidt et al. (2004). This explicit human lung was derived from High-Resolution Computed Tomography (HRCT) imaging of an in vitro preparation. This model was a selection of data extracted from the anatomy of a healthy human lung of an adult male. This lung model is free from pathological alteration (Gemci et al. 2008). In Figure 2, an isometric view of the human lung is presented. The notation of G0, G1, G2 and G3 are the cross-sectional gate to facilitate the model drawn for indication segments. This model was selected based on the actual structure of the human lung. No additional tissue configuration was created due to investigation which only focused on the airway blockage area. At each generation, the branching is essentially dichotomous; each airway being divided into two smaller daughter airways (Farkas et al. 2007). The tree starts at the trachea (G0-G0) whose average diameter and length are respectively $D_0 = 1.8$ cm and $L_0 = 12$ cm in the healthy human adult (Allen et al. 2004), and ends in the terminal bronchioles. From the trachea to terminal bronchioles, this element is located on average in generation G1-G1, these airways are purely conducting pipes. Two important features have

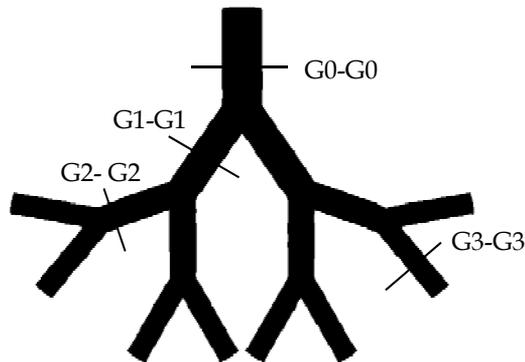


Fig. 1. General three-dimensional lung airway model



Fig. 2. Isometric view of lung airway

been included. First, the airway dimensions at the first generations (G0-G0 and G1-G1) are specific to the human anatomy and they are essentially independent of physiological variability (Martonen et al. 2003). Second, for higher generations (G2-G2 and G3-G3), a systematic branching asymmetry has been modelled in the different tree bifurcations.

Figure 3a and 3b illustrate the overall mesh of the model selected. The meshes are generated using a tetrahedral unstructured mesh format. The total number of grids generated is approximately 0.62 million. Concentrated mesh is applied near the branches of the bifurcation to increase the result interpretation. The flow details are critical as the effect of the boundary layer becomes significant in this region.

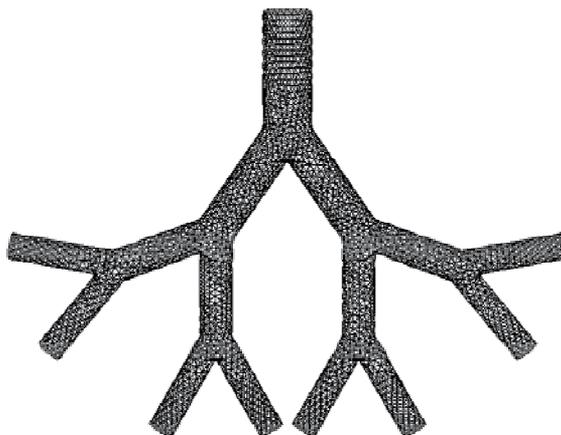


Fig. 3a. Mesh concentration on lung model

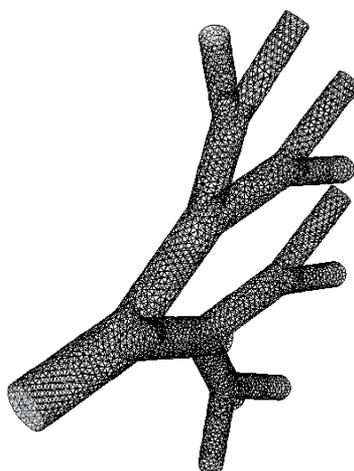


Fig. 3b. Mesh concentration on lung model (isometric view)

3. CFD simulation

The categorization of the anatomy of tracheobronchial tree airways constitutes the first step in the examination of the respiratory flow field and surfactant impact with volume of fluid

(VOF). The intention of this study is to understand the behaviour of the membrane which acts as a thin liquid film and the role of the surfactant on the model. In this study, the homogenous airflow is a comprehensive digital reference model with a maximum of four sections of pulmonary airway tree computed using the commercial CFD code FLUENT® (version 6.2). The computational mesh was generated using the FLUENT mesh generation code GAMBIT®. FLUENT employs a finite-volume method to solve the Navier-Stokes and continuity equations on an arbitrarily shaped flow domain with appropriate boundary conditions. The steady-state solution of the flow fields assumed converged when the residuals reduced to less than 10^{-4} . A typical run time was approximately 38 hours on a normal processor. In this CFD study of surfactant in the human pulmonary tree model, computations were performed at 28.3 L/min for a quasi steady-state volumetric adult inhalation flow rate for pulmonary surfactant.

3.1 Volume of Fluid (VOF) method in Fluent 6.2

The Volume of Fluid (VOF) model has been widely used in many research fields and commercial software. Known for its wide range of usage, implementation of this model in commercial code has been popular and FLUENT 6.2 is one of the software where their codes have been implemented. Isothermal VOF formulation was considered in this project with heat transfer and is assumed negligible due to analysis being based on surface interface. Formulation of laminar flow is presented here. For the turbulence effect, it can be done by averaging the Navier Stokes equation and detailed formulation can be found in Versteeg and Malalsekera (1995).

3.1.1 The basic of VOF for interface development

The VOF model is based on a volume fraction denoted by α where this volume of fraction is using a two fluids mixture in a fixed computational grid. The transport of the two fluids results in solving the multiphase fluid equation using the α values and interface location is reconstructed from the volume fraction field. The following procedures are used in solving the VOF model:

- i. Initial interface shape is modelled in computation domain.
- ii. From the existing interface in the model, volume of fraction is obtained. (Truncated volumes calculation was done at each interfacial cell, this discrete volume fraction is used instead of the exact interface information.)
- iii. The volume fraction field is advected. (Transport of fluid.)
- iv. Finally, the interface geometry and location are obtained using the new volume fraction.

All these steps are repeated until a convergence criterion is achieved. In VOF, especially in FLUENT 6.2, one can conduct a simulation for two different fluids without mixing the respective fluids. There will be an interface region between the both liquids and for each region α_q of 1 and 0 values will be given.

$\alpha_q = 0$: the cell is empty

$\alpha_q = 1$: the cell is full

$0 < \alpha_q < 1$: the cell contains the interface between the q^{th} fluid and one or more other fluids.

Using the same momentum equation, both fluids shared the same velocity field and it is solved throughout their domain. The tracking interface between both phases is accomplished by the solution of a continuity equation for the volume fraction of one of the phases. The equation is:

$$\frac{\partial \alpha_q}{\partial t} + \bar{v} \cdot \nabla \alpha_q = \frac{S_{\alpha_q}}{\rho_q} \quad (1)$$

The primary phase of the fluid will not be solved by the volume fraction equation, but will be computed through the following constrains:

$$\sum_{q=1}^n \alpha_q = 1 \quad (2)$$

For each of the phases, the properties will be presented in control volume of the respective component phase. For the two phase system, for example, if the phases are represented by the subscripts 1 and 2, and if the volume fraction of these is being tracked, the density in each cell is given by:

$$\rho = \alpha_2 \rho_2 + (1 - \alpha_2) \rho_1, \quad (3)$$

and for every n - phase system , the volume fraction averaged density takes on the following form:

$$\rho = \sum \alpha_q \rho_q \quad (4)$$

3.1.2 The momentum equation

As mentioned earlier, both fluids will share the same momentum equation throughout the domain. The momentum equation is:

$$\frac{\partial}{\partial t} \rho u_j + \rho u_i \frac{\partial u_j}{\partial x_i} = - \frac{\partial P}{\partial x_j} + \mu \frac{\partial}{\partial x_i} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) + \rho g_j + B_j, \quad (5)$$

Velocity component and direction denoted by u_j and x_j in j direction and t is the time. The summation of convention is applied when an index is repeated and $i, j = 1, 2$ for this problem. Pressure and gravitational acceleration are denoted by P and g respectively and B is a source term due to any additional force applied (per unit volume). This momentum equation dependant on the volume fraction α_q where in this case $q = 1, 2$ of the phases through the density ρ and viscosity μ . The computational domain is calculated through the following equation:

$$\frac{D\alpha_2}{Dt} = \frac{\partial \alpha_2}{\partial t} + u_i \frac{\partial \alpha_2}{\partial x_i} = 0 \quad (6)$$

The value of α_1 for the first phase is not solved directly at any stage, but it is solved based on the volume condition $\alpha_2 + \alpha_1 = 1$. By applying the properties of density ρ and viscosity μ , the following equations are obtained since ρ and μ depend on the volume fraction values:

$$\rho = \alpha_2 \rho_2 + \alpha_1 \rho_1, \quad (7)$$

$$\mu = \alpha_2 \mu_2 + \alpha_1 \mu_1 \quad (8)$$

The volume transport equation then, yields an updated volume fraction field with discrete values in computational cell. The values will later be used in reconstructing the new interface location.

3.1.3 Interface reconstruction

Reconstruction of new interface needs formulation from the discrete values of volume fraction inside the domain. In FLUENT 6.2, there are four interface reconstruction schemes, which are geometric reconstruction, donor-acceptor, Euler explicit and Euler implicit. Both Euler explicit and Euler implicit schemes treat cells that lie near the interface using the same interpolation as other cells. Meanwhile, the geometric reconstruction and the donor acceptor scheme treat the interfaces in an advanced interpolation. Full description of these schemes can be found in the FLUENT 6.2 User Manual (2002). A brief explanation on geometric reconstruction is explained here as it is recommended in FLUENT 6.2.

In the FLUENT 6.2 User Manual, the geometric reconstruction scheme approach standard interpolation to obtain face fluxes as the cell fills with one phase or another. Only at the cell near the interface between two phases is this method widely applied. This method is generalized for unstructured meshes and obtained from the work of Young (Piecewise linear 1982). It is assumed that for the calculation of the advection of fluid through the cell faces, the interface between two fluids has a linear slope within each cell. The following are the procedures of reconstruction on interface using geometric reconstruction scheme:

- i. First, calculation of position of the linear interface relative to the centre of each partial cell is done based on values of volume fraction.
- ii. Second, the advection amount of fluid through each face is calculated using computed interface representation. Normal and tangential velocity distribution on the faces is also used as information in calculating the advection of fluid.
- iii. Finally, each cell of volume fraction is calculated by computing the balance fluxes from the previous step.

Aside from the steady state problem, the geometric reconstruction scheme also available for transient cases. User defined functions can be used in order to obtain the desired solution of interest.

3.1.4 Effect of surface tension in volume of fraction

In VOF, attraction between the molecules in a fluid causes the rises of surface tension effect. This surface tension is a force only acting on the surface. The surface tension model in FLUENT is the continuum surface force (CSF) model proposed by Brackbill et al. 1992. In this model, the addition of surface tension to the VOF calculation results in a source term in the momentum equation. The pressure drop across the surface depends upon the surface tension coefficient and the surface curvature as measured by two radii in an orthogonal direction, R1 and R2:

$$\Delta p = p_2 - p_1 = \sigma \left(\frac{1}{R_1} + \frac{1}{R_2} \right) \quad (9)$$

where,

$$\kappa = \left(\frac{1}{R_1} + \frac{1}{R_2} \right), \quad (10)$$

hence,
$$\Delta p = \sigma \kappa \tag{11}$$

In equation $\Delta = p$ is the pressure drop across the surface, σ is the surface tension and κ is the mean curvature.

The formulation of CSF model is used in surface curvature computation and it is located from the local gradient in the surface normal at the interface. For example, let n be the surface normal and the gradient is α_q . Then, the volume fraction will be:

$$n = \nabla \alpha_q \tag{12}$$

And the curvature represented by divergence of the unit normal, $\hat{n} \kappa = \nabla \cdot \hat{n}$

where
$$\hat{n} = \frac{n}{|n|} \tag{13}$$

With this equation, the pressure variation across the interface is assumed linear and given by:

$$p_2 = p_1 + \sigma \kappa (\Delta \alpha_q) \tag{14}$$

The force at the surface can be expressed as a volume force using divergence theorem:

$$F_{vol} = \sum_{pairs_{ij}, i < j} \sigma_{ij} \frac{\alpha_i \rho_i \kappa_j \nabla \alpha_j + \alpha_j \rho_j \kappa_i \nabla \alpha_i}{\frac{1}{2}(\rho_i + \rho_j)} \tag{15}$$

And if only two phases are present in a cell, then $\kappa_i = -\kappa_j$ and $\nabla \alpha_i = -\nabla \alpha_j$, this equation can be simplified to:

$$F_{vol} = \sigma_{ij} \frac{\rho \kappa_i \nabla \alpha_i}{\frac{1}{2}(\rho_i + \rho_j)} \tag{16}$$

ρ is the volume averaged density computed using Equation 4. The simplified volume force shows that the surface tension source term for a cell is proportional to the average density in the cell.

3.2 Boundary conditions

The liquid was considered Newtonian with viscosity μ , density ρ and surface tension α flows in the region. The film thicknesses were obtained from the lipo lipid thickness in the human airway and it was set to be constant from the entrance of the tracheobronchial tree and uniform to its outlet. The membranes were patched with a similar thickness of the film, in order to reduce the computation period. A very fine mesh was created at the edge of the lung walls. The number of cells used was 198200 and a finer grid was placed near the thin film where the location of the interface was used. Two velocities inlets and a pressure outlet were used as boundary condition inputs. Inlet velocities were separated into two parts; small inlet and big inlet. The small inlet was selected for the thin film boundary condition and the flows maintained at 0.0001 m/s. Meanwhile the big inlet was set to 0.001 m/s to develop consistent laminar flows throughout the geometry.



Fig. 4. Patch region in the semi symmetrical airway model

These velocities were applied as the Reynolds number for thin films considered to be ≈ 1 . The density of water and viscosity of the lipo liquid was fixed to be 1023 kg/m^3 and 0.00093 kg/ms and for air the density was 1.225 kg/m^3 and viscosity $1.7894 \times 10^{-5} \text{ kg/ms}$. The calculations were performed transiently using FLUENT 6.2 until the steady state system was reached. Initially the study was done to anticipate that the solution could be run in a steady state so as to avoid time consuming transient calculations, but for better results for the interface it was run under transient mode. A grid dependency study was done until it was found that the location of the interface was not changing with the grid refinement. The time step was set to be 0.001s and simulation was done until 6750 iteration for the solution to converge. Later the iteration was reduced to 4600 iteration as the convergence residual was achieved sooner than expected. After total time of 38 hours, all the simulations reached convergence state. In order to achieve the state of convergence, severe computational time step needed to be selected. Initially time step was set to be 0.00001s and it not only prolonged the state of convergence, but also required more computational time. An assumption was made to set the time step to be 0.001s after 2500 iterations with results not showing prominent converging state of residuals. It was important to select the correct values of velocity based on thin film liquid flow. As per the initial guesses done with different velocity magnitude, the results show the region of multiphase fluids washed by the greater velocity of air. Based on several recalculations, attempts were made to maintain the thin film velocity to be as low as 1 (Reynolds number) in order to achieve the desired profile.

4. Results and discussion

The simulations on lipo lipid with interface were done with two different selections of velocity intake based on nominal inhalation in the human body. To obtain more detailed studies on each of these cases, the effect of surface tension were also included later to the simulation to provide better understandings of the membrane function of the pulmonary surfactant (lipo lipid in the human airway). A contour plot of lipid air volume fraction is

shown in Figure 5. The contours are from the calculation in which the effect of surface tension was neglected.

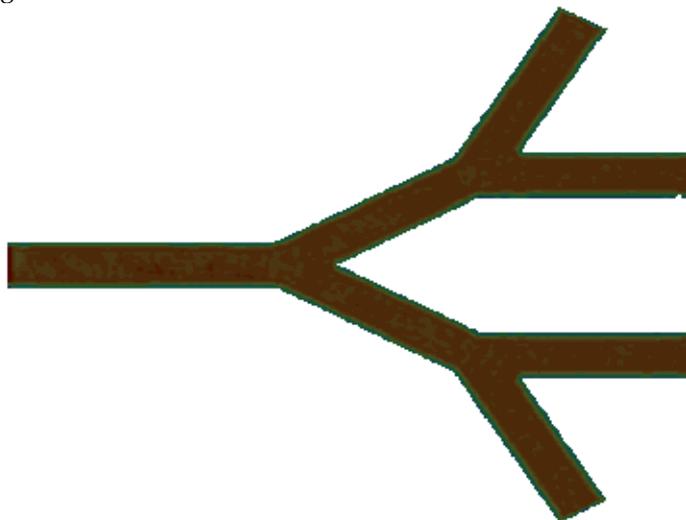


Fig. 5. Contour of volume of fraction with surface tension neglected of airway mode

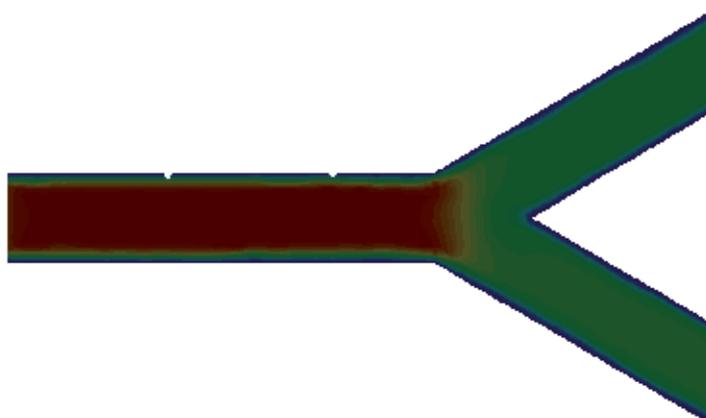


Fig. 6. Velocity profile of contour Figure 6

There are only small effects of interface changes shown for the simulation above. Near the edge of both membranes bordering adhesive to the human lung wall, the region bordering both fluids shows a minimum deformation over each fluid. Conversely there is some numerical diffusion involved where smeared effects can be seen at the interface resolution. In order to improve this factor, finer grids are able to solve this problem, but it occupies more computation time and it was not attempted here. This simulation was repeated by adding surface tension factor. The surface tension coefficient was taken to be 2 Nm as per lipo lipid configuration in the human body to obtain the effect of surface tension at the interface. The contours of volume of fraction are shown in Figure 7 based on the interface region taken approximately $F=0.5$ from the VOF calculation.

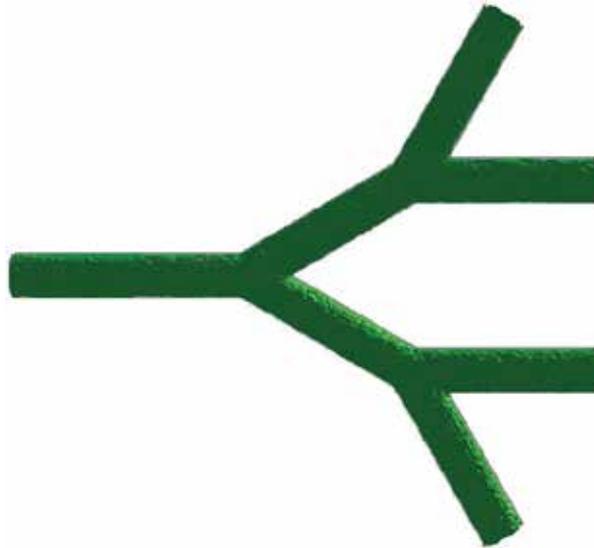


Fig. 7. Interface section $F=0.5$ without surface tension, 0 Nm

In order to identify the effect of the pulmonary surfactant, the edge perimeter of each simulation was selected to obtain the interface section for both, with and without surface tension. Figure 8 and Figure 9 will provide details of the edge perimeter membrane interface of pulmonary surfactant. The figure was studied from the entrance of the airway up until the first bifurcation section and the pulmonary effects were uniform throughout the model. As can be seen from the figure, the effect of surface tension plays a vital role in surface reconstruction. It is very prominent to see that as the value of surface tension increases from 0 to 2 ($0 = \text{no surface tension}$), the effect on the interface is to deliver a smoother curve on the edge of bifurcation and wall. It can be concluded that with surface tension effect, the volume fraction creates an even structure when surface tension increases. This proves the understanding of surface tension effect where the results show smoother effects as surface tension is applied. This also agrees with human lipid membrane function where the existence of the membrane as a thin liquid film provides a better breathing cycle and protects from injuries by preventing the collapse or absence of surfactant.

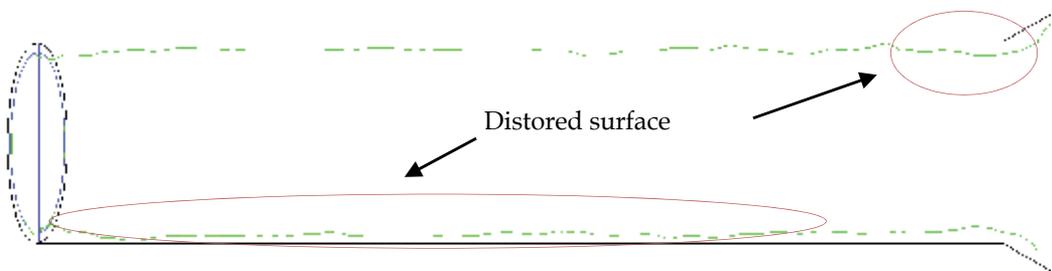


Fig. 8. Model without surface tension 0 Nm

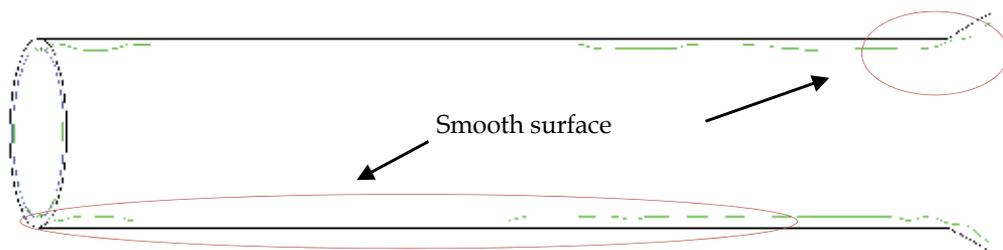


Fig. 9. Model with surface tension 2 Nm

The iso-surface obtained above was computed in a polynomial condition to obtain a better understanding between the two simulations. It is understood that the two phases are separated by an interface and the normal velocity of the interface must denote the value of zero. From the results obtained, the interface region of the volume fraction varies from 1 (lipo lipid) to zero (air). It can be concluded that the interface can be located along this region. Starting from the volume fraction contour plot, such as the one shown in Figure 10, the coordinates and velocities for each point (defined by the grid) on the assumed interface can be obtained.

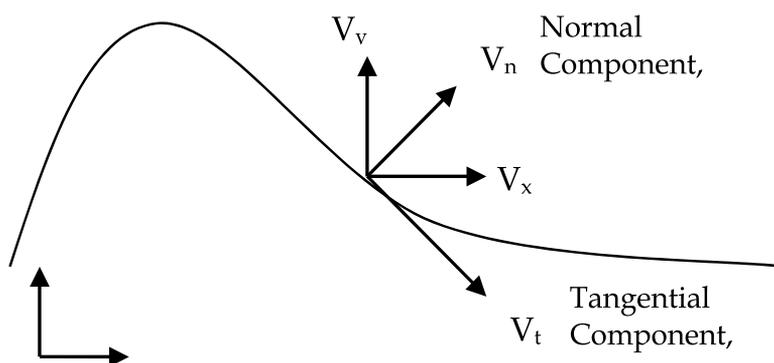


Fig. 10. Schematic of the interface

Considering a point on the interface, see Figure 10, the CFD calculated velocity components (V_x , V_y) correspond to the x , y values of the coordinate system. Normal and tangential velocities need to be calculated to evaluate the interface boundary conditions. An interface curve was fitted through sets of computed points corresponding to each of the volume fraction values investigated. The height of the interface can be expressed through the corresponding value of y along coordinate distance x as:

$$y = f(x) \tag{17}$$

This interface function of $y(x)$ defined, which is a continuous function, can then be differentiated to obtain the interface slope:

$$\frac{dy}{dx} = f'(x) \tag{18}$$

From the geometry, values of the unit normal vector can be obtained as follows:

$$\hat{n}_x = -\frac{dy/dx}{\sqrt{1+\left(\frac{dy}{dx}\right)^2}}, \quad \hat{n}_y = \frac{1}{\sqrt{1+\left(\frac{dy}{dx}\right)^2}} \quad (19)$$

Similarly the unit tangential vectors are given by:

$$\hat{n}_y = -\frac{dy/dx}{\sqrt{1+\left(\frac{dy}{dx}\right)^2}}, \quad \hat{n}_x = \frac{1}{\sqrt{1+\left(\frac{dy}{dx}\right)^2}} \quad (20)$$

Having obtained these unit vectors, the normal and tangential velocity components of the interface are respectively given by:

$$V_n = V_x \hat{n}_x + V_y \hat{n}_y \quad (21)$$

$$V_t = V_x \hat{t}_x + V_y \hat{t}_y \quad (22)$$

The calculations of first derivative dy/dx ensure the accuracy of the result. Here dy/dx represent the slope at any point x on the curve. Consequently, it is important to determine the precise interface curve representation $f(x)$ that passes through all data points.

Computation of normal velocity at the interface was done based on equation 21 and was repeated for both cases of volume of fraction on the conducted simulation. The objective of the investigation was to verify the accurateness of the obtained interface and its sensitivity to selection of significant volume fraction. The value of 0.5 of volume of fraction used here was based on previous work that shows at this particular value the interface can be presented more significantly. Using equation 21, the normal velocity along the x direction of the interface was calculated. The plot of normal velocity with respect to x direction is shown in Figure 11.

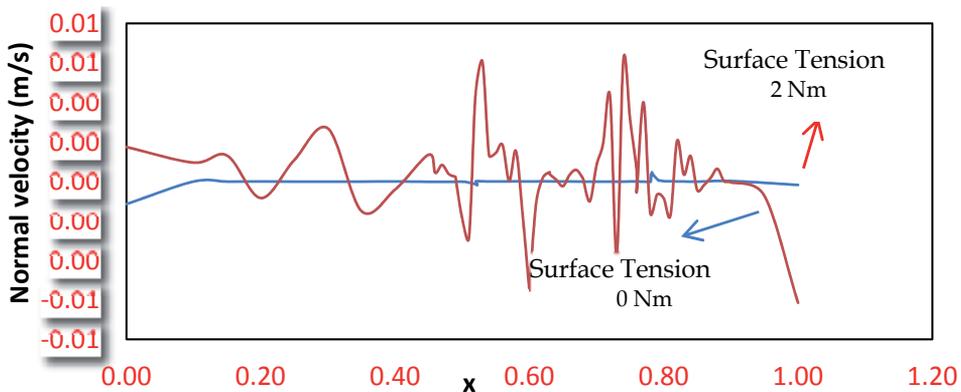


Fig. 11. Graph of normal velocity of iso -0.5 of both simulations

The comparison was done between the plot of normal velocity without surface tension and with surface tension. From Figure 11 the surface tension creates a smoother surface. The values of normal velocities of each case provide sufficient value of zero. It is understood that the interface region is zero normal velocity and with the application of the surface tension in the simulation it denotes as the numerical value. The results obtained clearly indicate that using VOF in FLUENT facilitate a good approach in determining the interface of new surface reconstruction. In the cases of using VOF method in the thin liquid film flow, this provides significant results on pulmonary surfactant condition in thin film flow.

5. Conclusion

In this chapter, a 3D model of the human tracheobronchial tree was investigated by using commercial software FLUENT 6.2. The major study was done to understand the behaviour of pulmonary surfactant in the human airway whereby the surface tension presences in airways are significant and without the existence of surfactant, it will severely affect the condition of inhalation. Apart from this, presenting the effects of surface tension in VOF provide significant results with application in thin liquid film studies. Finally, the use of the VOF model to calculate thin film was considered. It is a known fact that in a thin film flow, surface tension plays a vital role and is prominent in controlling the film surface. The results of VOF show that application of this method in the human airway with thin film flow accommodates a high level of accuracy. The conclusion was made that use of VOF in thin film flow is appropriate in identifying the real interface using volume fraction of VOF and the pulmonary surfactant in the human airway plays a major role for inhalation.

6. Acknowledgment

I have taken great effort in writing this book chapter, however, it would not have been possible without the kind support and help of many individuals and my organization. I would like to extend my sincere thanks to all of them. I am highly indebted to Bumi Armada Berhad and my family members for their guidance and constant motivation as well as for providing necessary support in completing this chapter. I would like to express my special gratitude and thanks to InTech and the University of Rijeka, Croatia, for giving me such attention and time.

7. References

- Allen G.M, Shortall B.P, Gemci T, Corcoran T.E, and Chigier N.A, (2004). Computational simulations of airflow in an in vitro model of the pediatric upper airways, *ASME Journal of Biomechanical Engineering* 126, pp 604–613
- Brackbill, J.U., Kothe, D.B. and Zemach, C. (1992). A continuum method for modelling surface tension, *Journal of Computational Physics* 100, pp 335–354
- Caro C.G, Schroter R.C, Watkins N, Sherwin S.J, and Sauret V, (2002). Steady inspiratory flow in planar and non planar models of human bronchial airways, *Proceedings of the Royal Society*, pp 458:791 - 809.
- Cerne, G., Petelin, S. and Tiselj, I. (2000). Upgrade of the VOF method for the simulation of the dispersed flow, *Proceedings of the ASME 2000, Fluid Engineering Division Summer Meeting Boston, Massachusetts*

- Farkas A, and Balashazy I, (2007). Simulation of the effect of local obstructions and blockage on airflow and aerosol deposition in central human airways, *Journal of Aerosol Science* 38: pp 865–884
- FLUENT 6.2 User Manual, (2002)
- Gemci T, Ponyavin V, Chen Y, Chen H, Collins R, (2008). Computational model of airflow in upper 17 generations of human respiratory tract, *Journal of Biomechanics*, Volume 41, Issue 9, pp 2047-2054
- Martonen, T. B., Fleming, J., Schroeter, J., Conway, J., & Hwang, D. (2003). In silico modelling of asthma, *Advanced Drug Delivery Reviews*, 55, pp 829–849
- Schmidt A, Zidowitz S, Kritete A, Denhard T, Krass S, and Peitgen H.O, (2004). A digital reference model of the human bronchial tree, *Computerized Medical Imaging and Graphics*, pp 28:203-211
- Versteeg H. K., and Malalasekera W. Book - An introduction to computational fluid dynamics, Pearson: (1995)
- Youngs, D.L. (1982). Time-dependent multi-material flow with large fluid distortion, in *Numerical Methods for Fluid Dynamics*, K.W.Morton and M.J.Baines (Eds.), pp 273-285

3D Particle Simulations of Deformation of Red Blood Cells in Micro-Capillary Vessel

Katsuya Nagayama¹ and Keisuke Honda²
Kyushu Institute of Technology
Hitachi Cooperation
Japan

1. Introduction

With the increase in arteriosclerosis, thrombosis, etc., in order to find out the cause, research of the flow characteristic of blood attracts attention. As for the analysis of the flow phenomenon of the RBC (Red Blood Cell or Erythrocyte), the numerical simulation (Wada et al., 2000, Tanaka et al., 2004) as well as experiment observation (Gaehtgens et al., 1980) is becoming a strong tool. Particle methods, such as SPH method (Monaghan J., 1992) and the MPS method (Koshizuka, 1997), treats both solid and liquid as particles, and can be applied to complicated flow analysis. When applying a particle method to the flow analysis of RBC, RBC is divided into the elastic film and internal liquid, and its deformation was analyzed in detail (Tanaka et al., 2004, Tsubota et al., 2006).

The RBC which is actually flowing in our body occupies 40-60% by volume ratio of blood (hematocrit), and is numerous. The objective of our research is clarifying the flow characteristic of the blood flow containing many RBCs. We reported preliminarily simulation of 2D blood flow (Nagayama et al., 2004), where many RBCs were simply treated as a lump of an elastic particle, the flow was analyzed qualitatively. Moreover, three dimensional RBC was modelled with double structure, and the RBC shape in flow was more realistic (Nagayama et al., 2005). The relation of the blood vessel diameter and the blood-flow with many RBCs was studied by 2D model (Nagayama, 2006) and by 3D model (Nagayama et al., 2008a). The model was also applied to 3D blood flow in capillary bend tube (Nagayama et al., 2008b).

The objective is to understand the fundamental flow phenomenon in a blood vessel. In this paper, 3 dimensional blood flows with RBCs in capillary tube were simulated.

In Section 2, simulation model was described. And the shape of single red blood cell in static fluid was shown.

In Section 3, blood flows with RBCs in capillary straight tube were simulated. And the relations of the blood vessel diameter and the hematocrit were investigated. Furthermore, transient phenomena of interacting red blood cells and their shape were investigated.

In Section 4, the model is applied to the capillary vessel flow at finger tip edge. The capillary vessel is modelled as two cases. One case is bent tube and another is bent and twisted tube, and RBC deformation were investigated.

2. Model descriptions

The particle model used for simulation is described. Then calculation conditions will be explained.

2.1 Mathematical descriptions

Particle method considers the interaction between particles and pursues motions of particles in Lagrangian way. Instead of NS equation, a momentum equation in particle model (1) consists of inertial force, inter-particle force, viscous diffusion and external force. Inter-particle force is attracting force or repulsing force between particles using particle pressure as shown in equation (2), so that to keep the density uniform in the domain.

Spring model is also considered for the elastic RBC surface. For the viscous diffusion term, MPS method (Koshizuka, 1997) is used. As for the external force, pressure difference between both ends of blood vessel was taken into consideration. The symbols are, \mathbf{u} : velocity vector, t : time, ω : weighting function, n : number density, \mathbf{r} : position, d and λ : constants. In addition i : particle number, j : surrounding particle number, n_0 : basic condition. RBC film particle is tied by surrounding particles using springs. In addition, resistance against bending is modelled as the force to the center of mass of surrounding particles. In addition, damping force is treated as viscous force in Eq. 1. The size of RBC is about $8 \mu\text{m}$ in diameter and $3 \mu\text{m}$ in thickness.

$$\frac{\partial \mathbf{u}_i}{\partial t} = -\frac{1}{\rho} \frac{d}{n^0} \sum_{j \neq i} \left[(P_j - P_i) \omega_j \frac{\mathbf{r}_{ij}}{|\mathbf{r}_{ij}|^2} \right] + \nu \frac{2d}{n^0 \lambda} \sum_{j \neq i} \left[\omega_j (\mathbf{u}_j - \mathbf{u}_i) \right] + \mathbf{F} \quad (1)$$

$$P_i = \frac{1}{\kappa} \left(1 - \frac{n_0}{n_i} \right) \quad (2)$$

2.2 RBC shape in static fluid

RBC is double structure: surface film and plasma liquid inside. RBC film particle is tied by surrounding film particles by springs with coefficient of $7.52 \times 10^{-2} \text{N/m}$ as shown in Fig. 1. In addition, resistance against bending with coefficient of $3.76 \times 10^{-4} \text{N/m}$, is modelled as the force to the center of mass of surrounding particles. Surface area is $140 \mu\text{m}^2$ and volume is $90 \mu\text{m}^3$. In the simulation, starting from sphere shape and removing 42% of plasma particle inside, the shape of RBC is formed. Fig.2 is the simulated RBC shape in static fluid. The size of RBC is about $8 \mu\text{m}$ in diameter and $3 \mu\text{m}$ in thickness. RBC shape will change with flow in blood vessel.

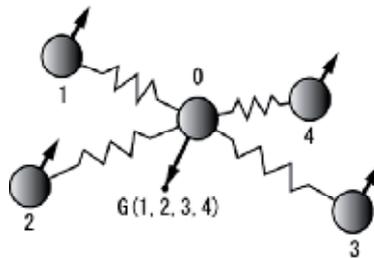


Fig. 1. Elastic film model

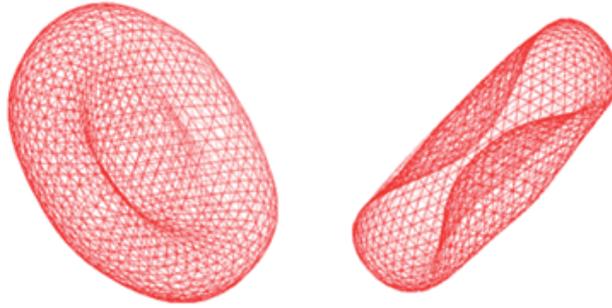


Fig. 2. Simulated RBC shape in static fluid

3. Deformation of RBCs in various inner diameter capillaries and hematocrit

Deformation of RBCs in various inner diameter capillaries and hematocrit was studied. Next, transition from single-file to multi-file flow as a function of hematocrit in capillaries of various diameters was discussed.

3.1 Simulation conditions

Simulation conditions are shown in Table 1. Simulations were carried out using normalized value. The velocity is normalized by 1 mm/s and the length is normalized using 10 μm . Physical properties are also shown in Table 1. Total simulation time is 0.3 s (300000 iterations with time step 1 μs), which is enough to reach stable flow.

Cases for simulation with various inner diameter (ID) capillaries and hematocrit are shown in Table 2. To study transition phenomena from single-file to multi-file flow, hematocrit 0.24-0.54 and capillaries of various inner diameters 5.5-8.7 μm were chosen for simulation.

Velocity of normalization	1	[mm/s]
Length of normalization	10	[μm]
Viscosity	0.001	[Pa s]
Density	1000	[kg m ³]
Simulation time	0.3	[s]
Elasticity of stretching	7.52 E-04	[N/m]
Elasticity for bending	2.63 E-05	[N/m]

Table 1. Simulation conditions

Case	ID [μm]	Ht
(a)	5.5	0.31
(b)	7.37	0.24
(c)	7.37	0.49
(d)	8.5	0.21
(e)	8.7	0.54

Table 2. Cases for simulation

3.2 Results for cases in various inner diameter capillaries and hematocrit

In this section, first of all, results of deformation of RBCs in various ID and hematocrit will be shown. Next, transition from single-file to multi-file flow will be discussed. The deviation of RBC distribution in a capillary blood vessel will also be shown.

3.2.1 Deformation of RBCs in various inner diameter capillaries and hematocrit

Simulation conditions are shown in Table1. Simulations were carried out for 5 cases with various inner diameter capillaries and hematocrit. Results are shown in Fig. 3 and the RBC shape was studied for each cases.

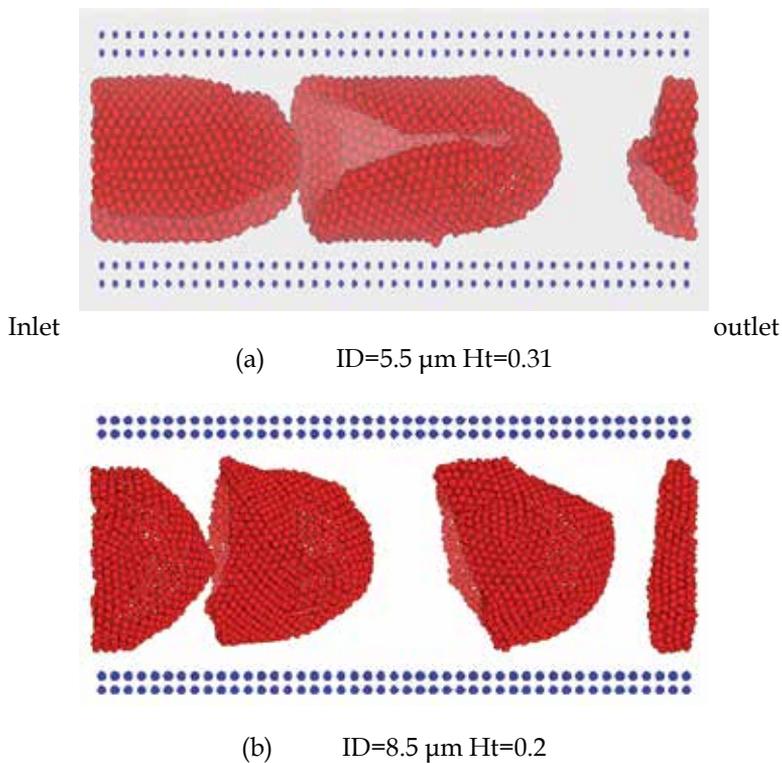
In case of (a) ID=5.5 μm Ht=0.31 (narrowest capillary), RBC flows in lines (single-file flow). RBC contacts with wall and deforms to consistently non-axisymmetric rocket shape 'torpedo' exhibiting a membrane-fold which extends from the open rear-end along one side toward the leading end (Gaehtgens et al., 1980).

In case of (b) ID=8.5 μm Ht =0.2, RBC flows in lines (single-file flow), rarely contact with the wall. RBCs flow at center of the blood vessel, parachute type deformation appeared.

In case of (c) ID=7.37 μm Ht =0.24, RBC flows basically in lines (single-file flow). RBCs do not flow at center of blood vessel. RBC interact each other, and sometimes contact with another RBC.

In case of (d) ID=7.37 μm Ht=0.49, RBC interacts (multi-file flow) with each other and contact with the wall, forming zipper shape.

In case of (e) ID=8.7 μm Ht=0.54, RBC interacts (multi-file flow) strongly with each other and contact with the wall, forming strong and complex deformation.



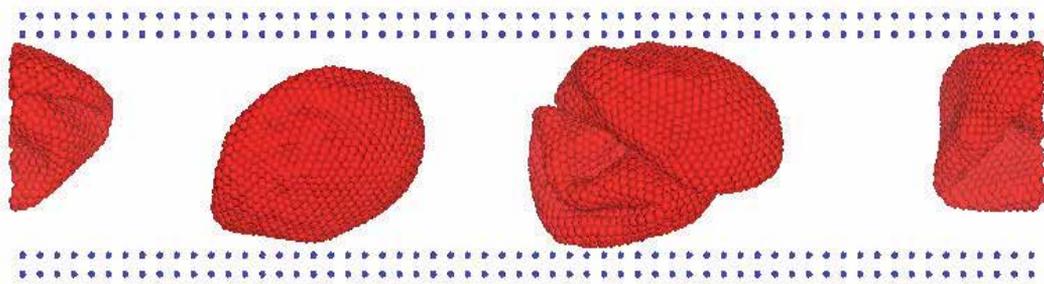
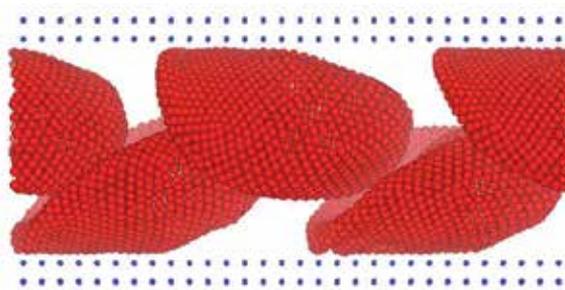
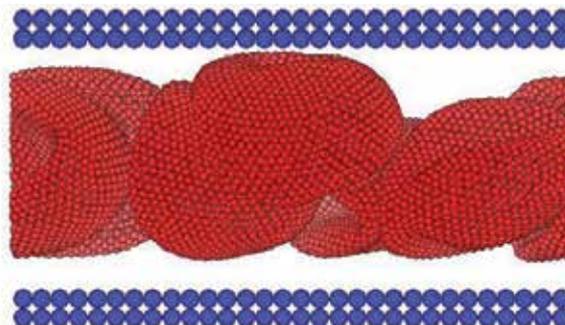
(c) ID=7.37 μm Ht=0.24(d) ID=7.37 μm Ht=0.49(e) ID= 8.7 μm Ht=0.54

Fig. 3. Deformations of RBCs in various inner diameter capillaries and hematocrit

3.2.2 Transition from single-file to multi-file flow

Transition from single-file to multi-file flow as a function of hematocrit in capillaries of various diameters is shown in Fig.4. Overall tendency in the experiment (Gaehtgens et al., 1980) and simulation are similar qualitatively. RBCs are single-file in narrow tube and at low hematocrit, while they are multi-file as the tube diameter increases or hematocrit increases. A line in Fig.2 is $Ht = 2.8/ID$. By a rough classification, RBCs are single-file when $Ht < 2.8/ID$, while they are multi-file when $Ht > 2.8/ID$.

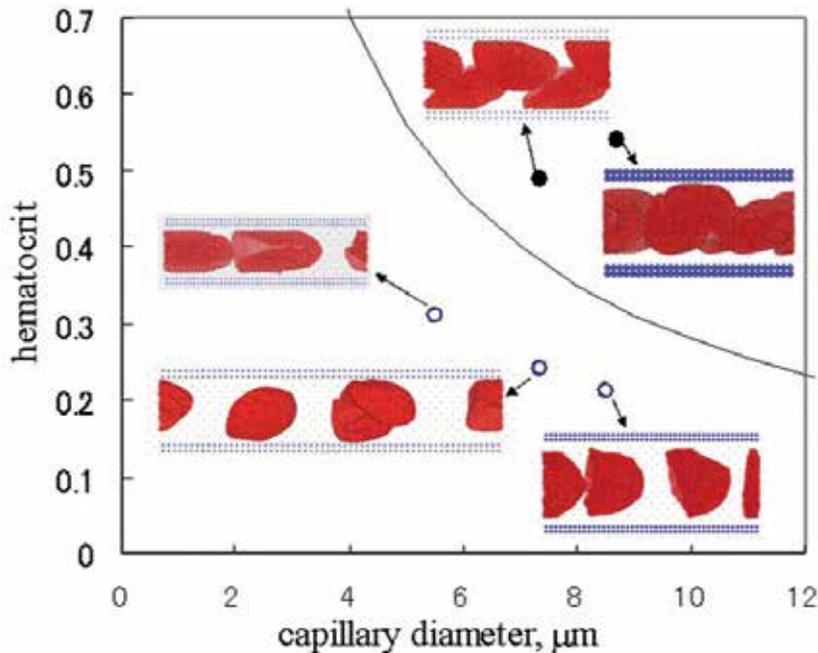


Fig. 4. Transition from single-file to multi-file flow as a function of hematocrit in capillaries of various diameters

3.2.3 Particle simulation about the deviation of RBC distribution in a capillary blood vessel

In Case (d), although RBCs were placed in line initially, they interacted and flows like zipper shape finally. This transition will be described in detail here. In Fig.5, RBC shape are shown for Case (d) ID=7.37 μm Ht =0.49, at 0 ms, 40 ms, 80 ms, 140 ms, 200 ms and 260 ms.

At 40 ms, RBCs are at the center of the blood vessel, parachute type shape, and flows in line. The back of the erythrocyte is dented, and it can also be said the bowl type.

In addition, plasma flow without the RBCs was calculated, the flow was Hagen-Poiseuille flow and velocity distribution at the cross section in pipe was parabola-shaped. When there were RBCs, RBC particles flows together due to the elastic film, and the velocity distribution was near in a trapezoid.

In 80 ms, the inclination occurs to arrangement of RBCs, and intervention happened. Uniformity in the axis center of the parachute-shaped collapsed, and intervention with a face of wall and the surrounding erythrocyte happened.

In time 140 ms, more mutual intervention of an erythrocyte was seen, and the shape is being changed complicatedly while having contact and rallying. A back RBC enters into the indent of the previous RBC, and the transfer state to the zipper type was seen.

Moreover its tendency was strengthened in time 200 ms. Back erythrocyte was entering into the indent of the previous erythrocyte, and 4 erythrocytes have ranged. At 260 ms, RBCs flow in the zipper shape, and it became stable.

As shown, RBCs flow with intervention of a tube wall and between the erythrocytes mutually. Initially RBCs flows with the parachute shape at early stage, but they begun to fluctuate and became unstable state. Mutual intervention of an erythrocyte was seen, and

the shape is being changed complicatedly while having contact and rallying. Finally RBCs flow in the zipper shape, and it became stable.

It is thought that the placement of the red blood cell changes when the condition changes. It is expected that the stable state changes by pipe diameter, a red count, the properties of matter of the red blood cell.

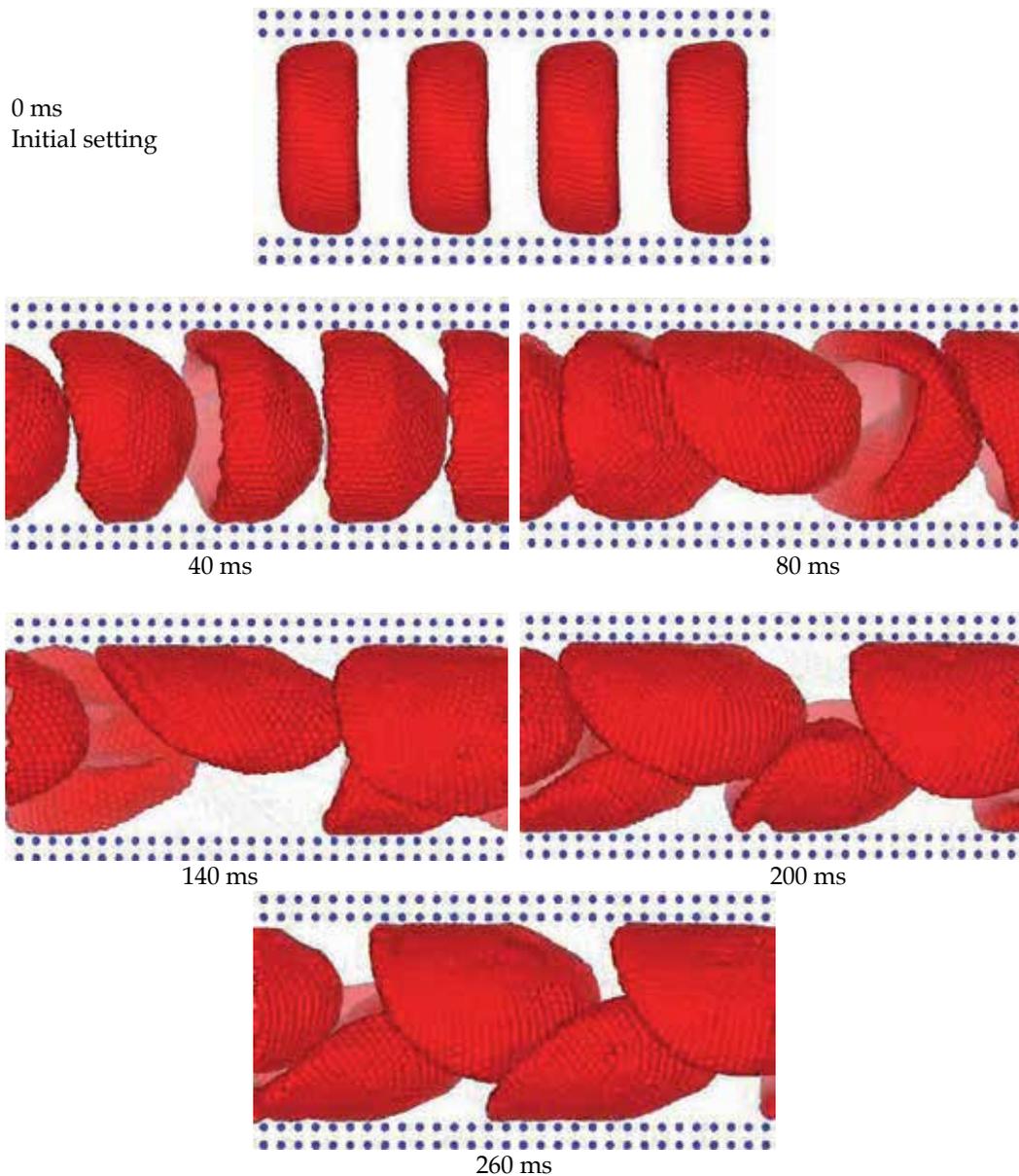


Fig. 5. RBC shape change for ID=7.37 μm Ht =0.49

4. Particle simulations of blood flow in bent and twisted capillary vessel with red blood cells

In this section, particle model is applied to simulate the capillary vessel flow in the turning point of blood vessel at finger tip. The capillary vessel is modelled as two cases: one is bent tube and another is bent and twisted tube.

4.1 Simulation conditions

Using microscope, capillary blood vessels can be observed at the finger tip as shown in Fig. 6. They change their shape depending on the health. To supply nutrition and to remove wastes, bent shape is usual and in good health. Winding and twisted shape tend to be observed in unhealthy condition such as high viscosity, although the reason is not clear. Blood flow analysis is important to study blood circulation. Here, particle model is applied to the capillary vessel flow to clarify the flow characteristics and mechanism of shape change.

Physical properties of plasma are density 1030 kg/m^3 , and viscosity $1.2 \text{ mPa}\cdot\text{s}$. Pressure gradient of 100 kPa/m is applied which cause blood velocity about 1 mm/s . (Note pressure gradient is fixed and the velocity depends on the tube shape and RBC number). Reynolds number is 0.86×10^{-2} and flow field is laminar. Periodic boundary condition is used at inlet and outlet, exit particles are supplied from inlet again.

Two type of basic shape of the capillary vessel is modelled as shown in Fig.7. One is bent tube and another is bent and twisted tube. All the RBC, blood vessel and plasma fluid are modelled as particles and Fig.2 is the initial setting of particles (plasma particles are not shown). Bent tube is inner diameter $6.82 \mu\text{m}$, length of straight portion $37.2 \mu\text{m}$. Twisted tube is inner diameter $7.37 \mu\text{m}$, height $90 \mu\text{m}$.

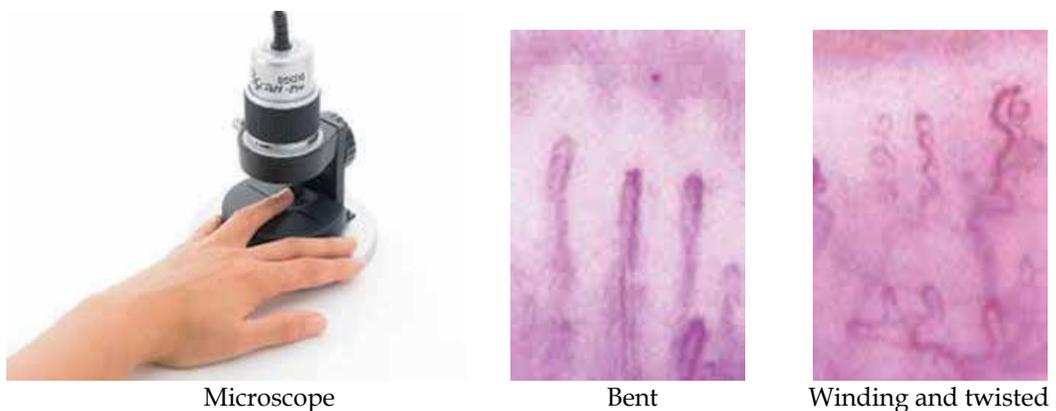


Fig. 6. Observed blood vessel at the finger tip

Particles are set every $0.62 \mu\text{m}$ in bent tube and $0.67 \mu\text{m}$ in twisted tube. At the surface of RBC, particle density is high to increase the accuracy of RBC shape. Cases for simulation with various RBC number for bent tube and twisted tube are shown in Table 3. To study the effect of RBC on flow field, cases with different number of RBC are simulated. Volume ratio (Hematocrit) of one RBC is 2.23% in bent tube and 1.57% in twisted tube. Hematocrit is 31% in bent tube and 38% in twist tube, which is close to the typical range of 40-50%. Total particle number including plasma, wall and RBC is minimum 29490 in plasma flow in bent

tube, maximum 85304 in 24 RBC flow in twisted tube. Total simulation time is 100 ms (100000 iterations with time step 1 s).

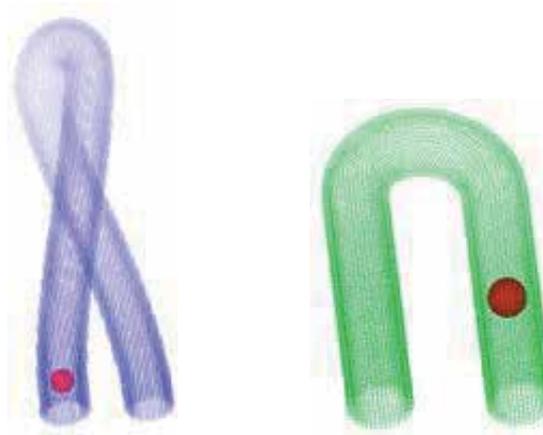


Fig. 7. Model for bent tube and twisted tube (with 1RBC)

Tube type	Number of RBC	Particle number for simulation
Bent tube	0	29490
	1	30742
	14	46995
Twisted tube	0	54446
	1	55728
	24	85304

Table 3. Cases for simulation

4.2 Results for blood flow in bent and twisted capillary vessel

In this section, results of RBC deformation for two cases: one is bent tube and another is bent and twisted tube will be shown for one RBC and many RBCs.

4.2.1 RBC deformation in bent tube

Simulated RBC shape in bent tube is shown in Fig.8. Fig.8 (a) is the result of one RBC case, and RBC shape is shown every 5 ms up to 40 ms. At 5 ms in the straight portion, RBC is parachute shape due to the fast flow at the tube center and slow flow close to the wall, which is typical in capillary tube. Velocity field around RBC is trapezoid, while parabolic Poiseuille flow in plasma flow far from RBC. At 10 ms to 20 ms, RBC is passing through the bent portion, starts to deform to asymmetric parachute shape. RBC particle tends to pass quickly inside the bent and slowly outside the bent, due to the flow length difference. After 25 ms, RBS shape tends to recover to symmetric parachute shape.

Fig.8 (b) is the snapshot of 14 RBC case at $t=80$ ms. RBCs tend to flow at the tube center, and the shape is between rocket and parachute. RBCs interact each other and tend to go in the back end of another RBC. When passing through the bent, RBC tends to keep the dent inside the bent.

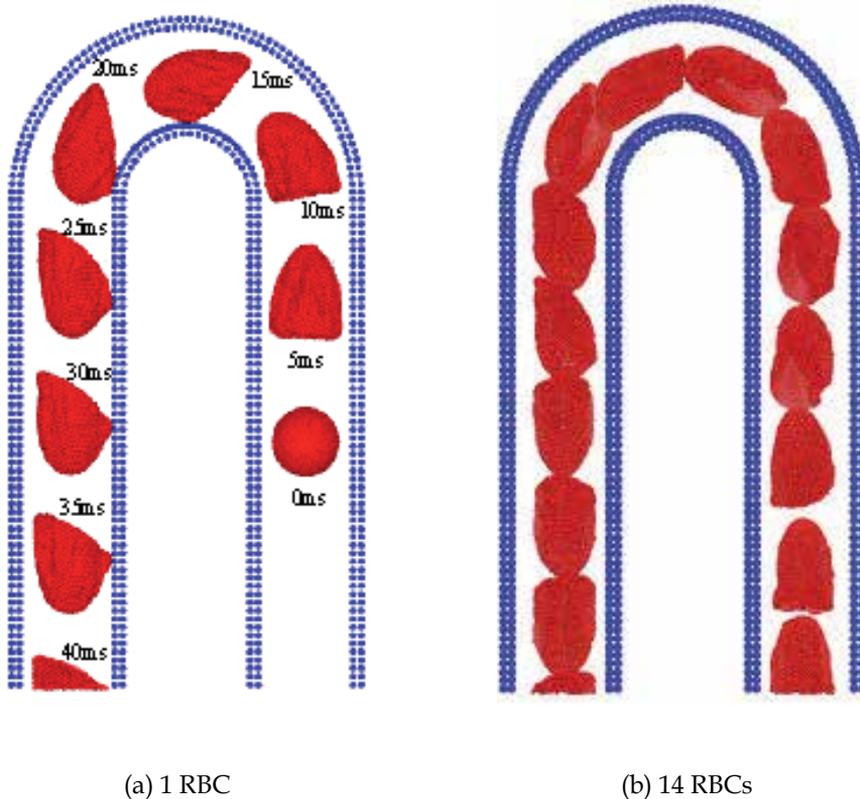


Fig. 8. Simulated RBC shape in bent tube

4.2.2 RBC deformation in twisted tube

Simulated RBC shape in twisted tube is shown in Fig.9 Fig.9 (a) is the result of one RBC case, and RBC shape is shown at 20, 40, 50, 60, 80 ms. RBC is parachute shape slightly asymmetric before the bent at 20 ms. At 40 ms to 60 ms, RBC is passing through the bent portion and deforms to asymmetric shape (between flat and parachute shape) due to the twist and bend flow. RBC particle tends to pass quickly inside the bend and slowly outside the bend, due to the flow length difference. After 60 ms, RBC shape tends to recover to symmetric parachute shape.

Fig.9 (b) is the snapshot of 24 RBC case at $t=90$ ms. RBC shapes are quite uneven by the strong interaction due to the twist and bent.

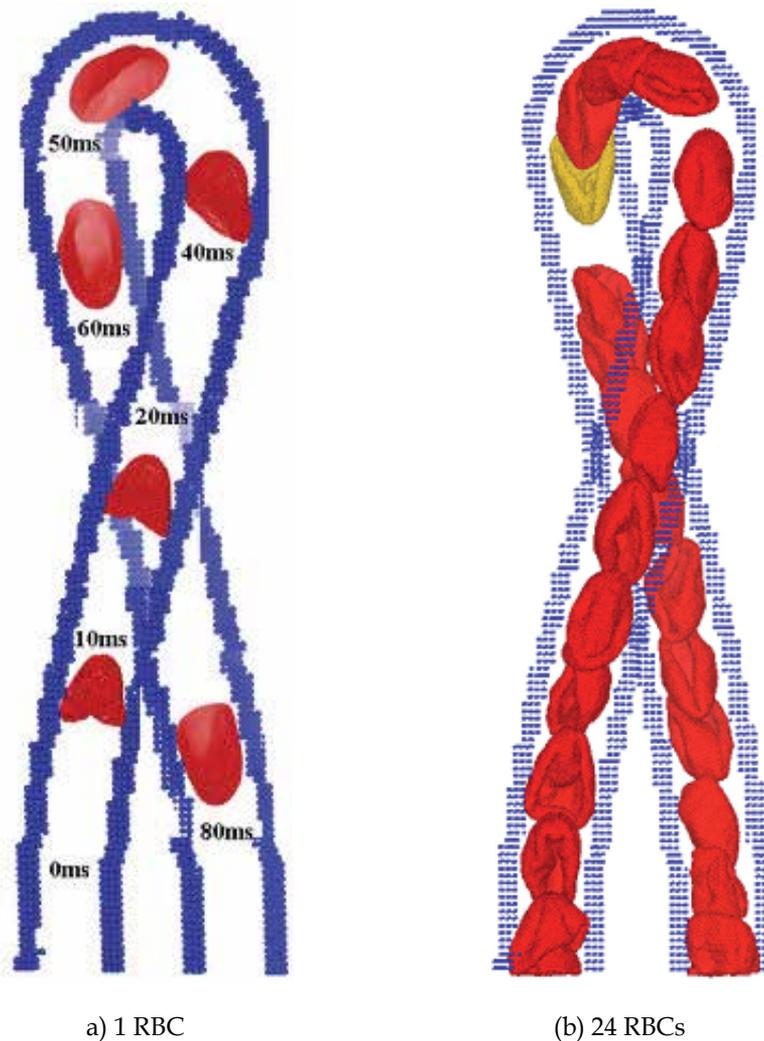


Fig. 9. Simulated RBC shape in twisted tube

5. Conclusion

3 dimensional particle model is applied to the capillary straight tube flow.

1. Deformations of RBCs in various ID (inner diameter) capillaries and hematocrit were studied. In case of $ID=5.5\mu m$ $Ht=0.31$, RBC flows in lines contacting with the wall and deforms to consistently non-axisymmetric rocket shape. In case of $ID=8.5\mu m$ $Ht=0.2$, RBCs flow at center of the blood vessel, parachute type deformation appeared. In case of $ID=7.37\mu m$ $Ht=0.49$, RBC interacts (multi-file flow) with each other and contact with the wall, forming zipper shape. In case of $ID=8.7\mu m$ $Ht=0.54$, RBC interacts (multi-file flow) strongly with each other and contact with the wall, forming strong and complex deformation.

2. Transition from single-file to multi-file flow as a function of hematocrit in capillaries of various diameters is studied. RBCs are single-file in narrow tube and at low hematocrit, while they are multi-file as the tube diameter increases or hematocrit increases.
3. RBC shape change in time was studied for a case of $ID = 7.37\mu\text{m}$ $Ht = 0.49$ in details. At first, RBCs were flowing in line like a parachute. But after that the shape fluctuated gradually, and it became stable in this case in the zipper state finally. A stable state is expected that it also changes by the pipe diameter, the number of erythrocytes and the physical properties of the RBCs.
4. Particle simulations were applied to the capillary vessel flow at finger tip. In case of bent tube, the RBCs is initially parachute shape at straight tube and then deforms to asymmetric parachute shape. In case of bent and twisted tube, initially RBC is parachute shape and then deforms to asymmetric shape (between flat and parachute shape) due to the twist and bend flow.

6. References

- Wada, Kobayashi, Takahashi and Karino (2000). A numerical simulation of the deformation of an erythrocyte, *Japanese Mechanical Engineering Congress*, MECJ-05, No.1226 pp.287-288, Japan, 2000
- Tanaka N., Takano T. and Masuzawa T. (2004). 3-dimensional micro-simulation of blood flow with SPH method, *Japanese Fluid Engineering Conference*, JSME, No. 712, Japan, 2004
- Gaehtgens P., Duhressen C. and Albrecht KH. (1980). Motions, Deformation and Interaction of Blood Cells and Plasma During Flow Through Narrow Capillary Tubes, *Blood cells* Vol.6, 799-812, 1980
- Monaghan J. (1992). *Annu. Rev. Astrophys.*, No.30, pp.543-574, 1992
- Koshizuka S., *Computational Fluid Dynamics* (1997). in Japanese, Baihuukan, Japan, 1997
- Tsubota K., Wada S. and Yamaguchi T. (2006). Particle method for computer simulation of red blood cell motion in blood flow, *Computer Methods and Programs in Biomedicine*, Vol. 83, pp. 139-146, 2006.
- Nagayama K. and Tanaka K. (2004). Particle Simulations of Two Phase Blood Flow with Red Blood Cell, *Japanese Fluid Engineering Conference*, JSME, No.G808, Japan, 2004, ISSN 1348-2882
- Nagayama K. and Tanaka K. (2005). Particle Simulations of three dimensional blood flow with a blood cell, *Proceedings of 2005 Annual Meeting*, Japan Society of Fluid Mechanics AM05-17-007, Japan, 2005
- Nagayama K. (2006), Particle Simulations of Blood flow in Vein with Many RBCs, *International Proceedings by Medimond from World Congress of Biomechanics*, pp. 557-562 Munich, Germany, 2006, Volume ISBN 88-7587-270-8, CD ISBN 88-7587-271-6
- Nagayama K. and Honda K. (2008a), Particle Simulations of the deformation of red blood cells in a capillary vessel. *Proceedings of the 12th Asian Congress of Fluid Mechanics*, Daejeon, Korea, 18-21 August 2008
- Nagayama K. and Honda K. (2008b), Particle Simulations of Blood Flow in Bent and Twisted Capillary Vessel with Red Blood Cells, *Proceedings of the TFEC*, Sapporo, Japan, 14-16 September 2008

Numerical Modeling and Simulations of Pulsatile Human Blood Flow in Different 3D-Geometries

Renat A. Sultanov and Dennis Guster
*Department of Information Systems and BCRL,
St. Cloud State University, St. Cloud, MN
USA*

1. Introduction

Cardiovascular diseases, such as ischemic heart disease, myocardial infarction, and stroke are leading causes of death in Western countries. All of these vascular diseases share a common element: atherosclerosis. They also share a common final event: the failure or destruction of the vascular wall structure, Dhein et al. (2005); Waite (2005).

Atherosclerosis reduces arterial lumen size through plaque formation and arterial wall thickening. It occurs at specific arterial sites. This phenomenon is related to hemodynamics and to wall shear stress (WSS) distribution, Fung (1993). From the physical point of view WSS is the tangential drag force produced by moving blood, i.e. it is a mathematical function of the velocity gradient of blood near the endothelial surface. A general description of WSS is presented in Landau & Lifshitz (1959). Arterial wall remodeling is regulated by WSS, Grotberg & Jensen (2004), for example, in response to high shear stress arteries enlarge. From the bio-mechanical point of view one can conclude, that the atherosclerotic plaques localize preferentially in the regions of low shear stresses, but not in regions of higher shear stresses. Furthermore, decreased shear stress induces intimal thickening in vessels which have adapted to high flow.

Final vascular events that induce fatal outcomes, such as acute coronary syndrome, are triggered by the sudden mechanical disruption of an arterial wall. Thus, we can conclude, that the final consequences of tragic fatal vascular diseases are strongly connected to mechanical events that occur within the vascular wall, and these in turn are likely to be heavily influenced by alterations in blood flow and the characteristics of the blood itself.

Currently researchers in the field of biomechanics and biomedicine conduct laboratory investigations of human blood flow in different shape and size tubes, which are designed to be approximate models of human vessels and arteries, see for example Huo & Kassab (2006); Taylor & Draney (2004). Some researchers also carry out intensive computer simulations of these bio-mechanical systems, see for example Chen & Lu (2004; 2006); Cho & Kensey (1991); Duraiswamy et al. (2007); Johnston et al. (2004); Morris et al. (2004; 2005); Mukundakrishnan et al. (2008); Peskin (1977); Sultanov et al., 2008 (a;b); Sultanov & Guster (2009).

Also, there have been laboratory experiments in which specific stents are incorporated in such artificial vessels (tubes). Stent implantation has been used to open diseased coronary blood vessels, allowing improved perfusion of the cardiac muscle. Used in combination with drug therapy, vascular repair and dilation techniques (angioplasty) the implantation of metallic

stents has created a multibillion dollar industry. Stents are commonly used in many different blood vessels, but the primary site of deployment is in diseased coronary arteries.

Stents represent a very special case in the modeling research problems mentioned above, Frank et al. (2002). Taking into account that stents have a very small size and rather complicated structure and shape, this situation makes it difficult to obtain precise measurements. Therefore high quality and precise computer simulations of blood flow through vessels with implanted stents would be most useful, Frank et al. (2002). Work of this type is already underway and we would like to mention several pertinent studies, Benard et al. (2006); Banerjee et al. (2007); Faik et al. (2007); Seo et al. (2005).

Nevertheless, there are still many problems in obtaining precise realistic geometries for the required vessels. Human arteries, especially the aorta, have complicated spatial-geometric and characteristic configurations. For example, the aortic arch centerline does not lie on a plane and there are major branches at the top of the arch feeding the carotid arterial circulation. One of the main problems in the field of bio-medical blood flow simulation is to obtain precise geometrical-mathematical representations of different vessels. This information in turn needs to be included in the simulation programs.

However, it seems logical that a first step in these investigations would be to apply simpler 3D-geometry forms and models, but at the same time to take into account the precise physical effects of blood movement such as the non-Newtonian characteristics of human blood, realistic pulsatile flow, and possible turbulent effects. Because of the applied pulsatile flow in our simulations turbulence may be significant to the final results of this study.

Therefore, in the current work we carried out real-time full-dimensional computer simulations of a realistic pulsatile human blood flow in actual size vessels, vessels with a bifurcation, and in a model of the aortic arch. We take into account different physical effects, such as turbulence and the non-Newtonian nature of human blood. The next section presents the mathematical methodology and the physical model used in this work. The general purpose commercial computational fluid dynamics program FLOW3D is used for its basic functionality, but we supplemented its capability by adding our routines to obtain the results presented in this work.

Sec. 3 presents results for three vessels of different geometries. The CGS unit system is used in all simulations, as well as for presentation of the results. Conclusions and discussion comparing our results to well respected previous studies are included in Sec. 4.

2. Physical models and mathematical methods

As we mentioned above, we undertook pulsatile human blood flow simulation experiments using different size and shape human vessel arteries. For each spatial configuration one needs to provide a specific approach for the numerical solution to the complicated second order partial differential equations of the fluid dynamics. These equations are also named as the Navier-Stokes (NS) equations.

For simple cylindrical vessels we used the cylindrical coordinate system: $\vec{r} = (r, \theta, Z)$. However, for the aortic arch or bifurcated vessels, where there is no cylindrical symmetry, we applied the Cartesian coordinate system: $\vec{r} = (x, y, z)$. In the cases of the aortic arch and bifurcated vessels we used up to five blocks of matched Cartesian coordinate subsystems. Below we represent the NS equations in a general form, because, for each of the special cases, considered in this work and the chosen coordinate system, the partial differential equations of the fluid dynamics may look different. At the same time we understand that the general differential operator form of these equations is unique.

2.1 Equations

The general form of the dynamics equation for viscous fluid can be written in the following way, Landau & Lifshitz (1959):

$$\rho \left(\frac{\partial v_i}{\partial t} + v_k \frac{\partial v_i}{\partial x_k} \right) = - \frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_k} \left\{ \eta \left(\frac{\partial v_i}{\partial x_k} + \frac{\partial v_k}{\partial x_i} - \frac{2}{3} \delta_{ik} \frac{\partial v_l}{\partial x_l} \right) \right\} + \frac{\partial}{\partial x_i} \left(\zeta \frac{\partial v_l}{\partial x_l} \right), \quad (1)$$

here v_i and x_k are velocity and coordinates of the fluid, ρ is the density of fluid, η and ζ are dynamical characteristics of the fluid, i.e. coefficients of viscosity. Because in a general case the pressure p , the temperature T and therefore the viscosity coefficients η and ζ are not constants in a flowing fluid, one cannot take them out of the partial differentials in the Eq. (1). However, in a very wide range of applications it is a good approximation to consider the variation of these coefficients to be negligible in the fluid, that is $\eta = const$ and $\zeta = const$. In these cases the Eq. (1) becomes the well known Navier-Stokes equation:

$$\rho \left[\frac{\partial \vec{v}}{\partial t} + (\vec{v} \nabla) \vec{v} \right] = - \text{grad} p + \eta \Delta \vec{v} + (\zeta + \eta/3) \text{grad} \text{div} \vec{v}. \quad (2)$$

Further, if the fluid is considered as incompressible: $\text{div} \vec{v} = 0$, then the NS equation becomes simpler in form:

$$\left[\frac{\partial \vec{v}}{\partial t} + (\vec{v} \nabla) \vec{v} \right] = - \frac{1}{\rho} \text{grad} p + \frac{\eta}{\rho} \Delta \vec{v}. \quad (3)$$

This equation is, probably, one of the most applicable mathematical results related to modeling and real time simulation of various physical systems, such as: air flow in aerodynamics, blood flow in medical applications and even cash flow in financial problems. The fundamental NS equation is nonlinear, diverse, rich, and, as we mentioned above, has strong practical applications in science and in various fields of modern technologies, for example, micro- and nano-fluidics. However, because of the considered atomistic level in the field of novel nano-fluidics problems a direct application of the NS equation to these systems might be problematic, especially at low temperatures of these systems. Further, from the mathematical point of view the NS equation is a very complicated nonlinear partial differential equation, which still remains as a "stumbling block" for mathematicians. The Clay Mathematics Institute (Boston, Massachusetts, USA) announced *Seven Millennium Problems* with a prize of US\$ 1,000,000.00 for each (<http://www.claymath.org/index.php>). One of these problems is related to the existence and smoothness of the NS equation. It is hard to believe that despite many successful practical applications of the NS equation its fundamental mathematical property is still open to question.

Nevertheless, because this work deals with different 3D geometries it would be useful to represent the NS equation in different forms. For example, in the case of cylindrical symmetry one can apply cylindrical coordinates (r, θ, z) and the NS equation could be written:

$$\frac{\partial v_r}{\partial t} + v_r \frac{\partial v_r}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_r}{\partial \theta} + v_z \frac{\partial v_r}{\partial z} - \frac{v_\theta^2}{r} = - \frac{1}{\rho} \frac{\partial p}{\partial r} + F_r + \nu \left(\nabla^2 v_r - \frac{v_r}{r^2} - \frac{2}{r^2} \frac{\partial v_\theta}{\partial \theta} \right), \quad (4)$$

$$\frac{\partial v_\theta}{\partial t} + v_r \frac{\partial v_\theta}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_\theta}{\partial \theta} + v_z \frac{\partial v_\theta}{\partial z} + \frac{v_r v_\theta}{r} = - \frac{1}{\rho r} \frac{\partial p}{\partial \theta} + F_\theta + \nu \left(\nabla^2 v_\theta + \frac{2}{r^2} \frac{\partial v_r}{\partial \theta} - \frac{v_\theta}{r^2} \right), \quad (5)$$

$$\frac{\partial v_z}{\partial t} + v_r \frac{\partial v_z}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_z}{\partial \theta} + v_z \frac{\partial v_z}{\partial z} = - \frac{1}{\rho} \frac{\partial p}{\partial z} + F_z + \nu \nabla^2 v_z, \quad (6)$$

together with the continuity equation:

$$\frac{1}{r} \frac{\partial}{\partial r}(rv_r) + \frac{1}{r} \frac{\partial v_\theta}{\partial \theta} + \frac{\partial v_z}{\partial z} = 0, \quad (7)$$

where $\nu = \eta/\rho$ is the kinematic viscosity Landau & Lifshitz (1959).

However, in the general case when there is no symmetry it is useful to apply the well known Cartesian coordinates x , y , and z : accordingly the equations of motion for the fluid velocity components (u, v, w) are:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} = -\frac{1}{\rho} \frac{\partial p}{\partial x} + X + \nu \nabla^2 u, \quad (8)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} = -\frac{1}{\rho} \frac{\partial p}{\partial y} + Y + \nu \nabla^2 v, \quad (9)$$

$$\frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} = -\frac{1}{\rho} \frac{\partial p}{\partial z} + Z + \nu \nabla^2 w. \quad (10)$$

In this case the continuity equation has the form:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0. \quad (11)$$

In the specific case of the FLOW3D program the equations of motion for the fluid velocity components (u, v, w) with special additional terms included in the program are written:

$$\frac{\partial u}{\partial t} + \frac{1}{V_F} (uA_x \frac{\partial u}{\partial x} + vA_y R \frac{\partial u}{\partial y} + wA_z \frac{\partial u}{\partial z}) - \zeta \frac{A_y v^2}{xV_f} = -\frac{1}{\rho} \frac{\partial p}{\partial x} + G_x + f_x - b_x - \frac{R_{sor}}{\rho V_f} (u - u_w - \delta \cdot u_s) \quad (12)$$

$$\frac{\partial v}{\partial t} + \frac{1}{V_F} (uA_x \frac{\partial v}{\partial x} + vA_y R \frac{\partial v}{\partial y} + wA_z \frac{\partial v}{\partial z}) + \zeta \frac{A_y uv}{xV_f} = -\frac{R}{\rho} \frac{\partial p}{\partial y} + G_y + f_y - b_y - \frac{R_{sor}}{\rho V_f} (v - v_w - \delta \cdot v_s) \quad (13)$$

$$\frac{\partial w}{\partial t} + \frac{1}{V_F} (uA_x \frac{\partial w}{\partial x} + vA_y R \frac{\partial w}{\partial y} + wA_z \frac{\partial w}{\partial z}) = -\frac{1}{\rho} \frac{\partial p}{\partial z} + G_z + f_z - b_z - \frac{R_{sor}}{\rho V_f} (w - w_w - \delta \cdot w_s). \quad (14)$$

Here, (u, v, w) are the velocity components in coordinate directions (x, y, z) respectively. For example, when Cartesian coordinates are used, $R = 1$ and $\zeta = 0$, see FLOW3D manual FLOW3D (2007). A_x is the fractional area open to flow in the x direction, analogously for A_y and A_z . Next, V_F is the fractional volume open to flow, R and ζ are coefficients which depend on the coordinate system: (x, y, z) or (r, θ, z) , ρ is the fluid density, R_{sor} is a mass source term. Finally, (G_x, G_y, G_z) are so called body accelerations FLOW3D (2007), (f_x, f_y, f_z) are viscous

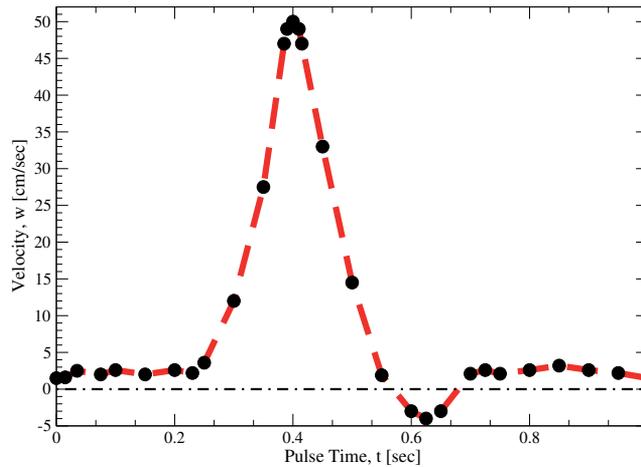


Fig. 1. The inflow time-dependent waveform used in these simulations. The function was taken from Fig. 3 of the work Papaharilaou et al. (2007).

accelerations, (b_x, b_y, b_z) are the flow losses in porous media or across porous baffle plates, and the final term accounts for the injection of mass at a source represented by a geometric component. Mathematical expressions for the viscous accelerations (f_x, f_y, f_z) are presented in the Appendix.

The term $U_w = (u_w, v_w, w_w)$ in equations (12-14) is the velocity of the source component, which will generally be non-zero for a mass source of a General Moving Object (GMO) FLOW3D (2007). The term $U_s = (u_s, v_s, w_s)$ is the velocity of the fluid at the surface of the source relative to the source itself. It is computed in each control volume as

$$\vec{U}_s = \frac{1}{\rho_s} \frac{d(Q\vec{n})}{dA} \quad (15)$$

where dQ is the mass flow rate, ρ_s fluid source density, dA the area of the source surface in the cell and \vec{n} the outward normal to the surface. The source is of the stagnation pressure type when in equations (12-14) $\delta = 0.0$. Next, $\delta = 1.0$ corresponds to the source of the static pressure type.

As we mentioned earlier, in all simulations we considered the blood flow as a pulsatile flow. The final result for the inflow waveform has been taken from Fig. 3 of the work Papaharilaou et al. (2007). The pulse was applied for 5.5 cycle times in our work. The waveform is shown in Fig. 1. These velocity values are used as time-dependent inflow initial boundary conditions. These numbers are included directly in the FLOW3D program.

Next, the general mass continuity equation, which is solved within the FLOW3D program has the following *general form*:

$$V_f \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u A_x) + R \frac{\partial}{\partial y}(\rho v A_y) + \frac{\partial}{\partial z}(\rho w A_z) + \zeta \frac{\rho u A_x}{x} = R_{dif} + R_{sor}, \quad (16)$$

R_{dif} is a turbulent diffusion term, and R_{sor} is a mass source. The turbulent diffusion term is

$$R_{dif} = \frac{\partial}{\partial x}(v_p A_x \frac{\partial \rho}{\partial x}) + R \frac{\partial}{\partial y}(v_p A_y R \frac{\partial \rho}{\partial y}) + \frac{\partial}{\partial z}(v_p A_z \frac{\partial \rho}{\partial z}) + \zeta \frac{\rho v_p A_x}{x}, \quad (17)$$

where the coefficient $v_p = C_p \mu / \rho$, μ is dynamic viscosity and C_p is a constant. The R_{sor} term is a density source term that can be used to model mass injections through porous obstacle surfaces.

Compressible flow problems require the solution of the full density transport equation. In this work we treat blood as an incompressible fluid. For incompressible fluids $\rho = \text{constant}$, and the equation (15) becomes the following:

$$\frac{\partial}{\partial x}(uA_x) + \frac{\partial}{\partial y}(vA_y) + \frac{\partial}{\partial z}(wA_z) + \zeta \frac{uA_x}{x} = \frac{R_{sor}}{\rho}. \quad (18)$$

It is assumed, that at a stagnation pressure source fluid enters the domain at zero velocity. As a result, pressure should be considered at the source to move the fluid away from the source. For example, such sources are designed to model fluid emerging at the end of a rocket or the simple deflating process of a balloon. In general, stagnation pressure sources apply to cases when the momentum of the emerging fluid is created inside the source component, like in a rocket engine. At a static pressure source the fluid velocity is computed from the mass flow rate and the surface area of the source. In this case, no extra pressure is required to propel the fluid away from the source. An example of such a source is fluid emerging from a long straight pipe. Note that in this case the fluid momentum is created far from where the source is located.

Turbulence models can be taken into account in FLOW3D. It allows us to estimate the influence of turbulent fluctuations on mean flow quantities. This influence is usually expressed by additional diffusion terms in the equations for mean mass, momentum, and energy. The turbulence kinetic energy per unit mass, q , is the following:

$$\frac{\partial q}{\partial t} + \frac{1}{V_F} \left(uA_x \frac{\partial q}{\partial x} + vA_y R \frac{\partial q}{\partial y} + wA_z \frac{\partial q}{\partial z} \right) = P + G + Diff - D, \quad (19)$$

where P is shear production, G is buoyancy production, $Diff$ is diffusion, and D is a coefficient FLOW3D (2007).

When the turbulence option is used, the viscosity is a sum of the molecular and turbulent values. For non-Newtonian fluids the viscosity can be a function of the strain rate and/or temperature. A general expression based on the Carreau model is used in FLOW-3D for the strain rate dependent viscosity:

$$\mu = \mu_\infty + \frac{\mu_0 E_T - \mu_\infty}{\lambda_{00} + [\lambda_0 + (\lambda_1 E_T)^2 e_{ij} e_{ij}]^{(1-n)/2}} + \frac{\lambda_2}{\sqrt{(e_{ij} e_{ij})}}, \quad (20)$$

where $e_{ij} = 1/2(\partial u_i / \partial x_j + \partial u_j / \partial x_i)$ is the fluid strain rate in Cartesian tensor notations, $\mu_\infty, \mu_0, \lambda_0, \lambda_1, \lambda_2$ and n are constants. Also, $E_T = \exp[a(T^* / (T - b) - C)]$, where T^*, a, b , and C are also parameters of the temperature dependence, and T is fluid temperature. This basic formula is used in our simulations for blood flow in vessels and in the aortic arch. For a variable dynamic viscosity μ , the viscous accelerations have a special form. That form is shown in the Appendix.

The equations of fluid dynamics should be solved together with specific boundary conditions. The numerical model starts with a computational mesh, or grid. It consists of a number of interconnected elements, or 3D-cells. These 3D-cells subdivide the physical space into small volumes with several nodes associated with each such volume. The nodes are used to store values of the unknown parameters, such as pressure, strain rate, temperature, velocity

components and etcetera. This procedure provides values for defining the flow parameters at discrete locations and allows specific boundary conditions to be set up. As the culminating step, one can start developing effective numerical approximations for the solution of the fluid dynamics equations, i.e. NS equation.

New pressure-velocity solutions have been implemented in FLOW-3D. We used the GMRES method. GMRES stands for the generalized minimum residual method. In addition to the GMRES solution, a new optional algorithm, the generalized conjugate gradient (GCG) algorithm, has also been implemented for solving viscous terms in the new GMRES routine. This new solver is a highly accurate and efficient method for a wide range of problems. It possesses good convergence, symmetry and speed properties; however, it does use more memory than the SOR or SADI methods.

3. Numerical results

Results of our simulations are presented below. One of the most important preliminary testing tasks is to check for numerical convergence. This test has been successfully accomplished in this work. A portion of the test calculation results are shown below in this paper. Next, in this work particular attention has been given to the calculations of the wall shear stress distribution (WSS). As we mentioned above WSS is the tangential drag force produced by moving blood, i.e. it is a mathematical function of the velocity gradient of blood near the endothelial surface:

$$\tau_w = \mu \left[\frac{\partial U(t, y, R_v)}{\partial y} \right]_{y \approx 0}. \quad (21)$$

Here μ is the dynamic viscosity, t is current time, $U(t, y, R_v)$ is the flow velocity parallel to the wall, y is the distance to the wall of the vessel, and R_v is its radius. It was shown, that the magnitude of WSS is directly proportional to blood flow and blood viscosity and inversely proportional to the cube of the radius of the vessel, in other words a small change of the radius of a vessel will have a large effect on WSS.

First, we present results for a simple geometry vessel in the shape of a tube. However, the human blood is treated as real and a non-Newtonian liquid. The necessary data for viscosity of the blood was found in previous laboratory and clinical measurements.

As we mentioned above, we take into account the real pulsatile flow, which is shown in Fig. 1. The data for Fig. 1 have also been obtained in clinical measurements, Papaharilaou et al. (2007). After such preliminary simulations we switched to a more complicated spatial configurations. This work begins with coronary bifurcation and the aortic arch. It is axiomatic that real people may have different size aortic arches with slightly different shapes. However, we carried out simulations for an average size and shape aortic arch.

The main goal of this work is to treat the above mentioned systems realistically, reveal the physics of the blood flow dynamics, and to obtain reliable results for pressure, dynamic viscosity, velocity profiles and strain rate distributions. Also, we tested the widely cited Newtonian and non-Newtonian models of the human blood.

3.1 Straight vessel: cylinder

First, we selected a simple vessel geometry, that is we considered the shape of a straight vessel to be a tube. In our simulations involving a straight cylinder type vessel we applied a cylindrical coordinate system: (r, θ, Z) with the axis OZ directed over the tube axis. Different quantities of cells have been used to discretize the empty space inside the tube. In the open space (inner part of the tube) the fluid dynamics equations have been solved using appropriate

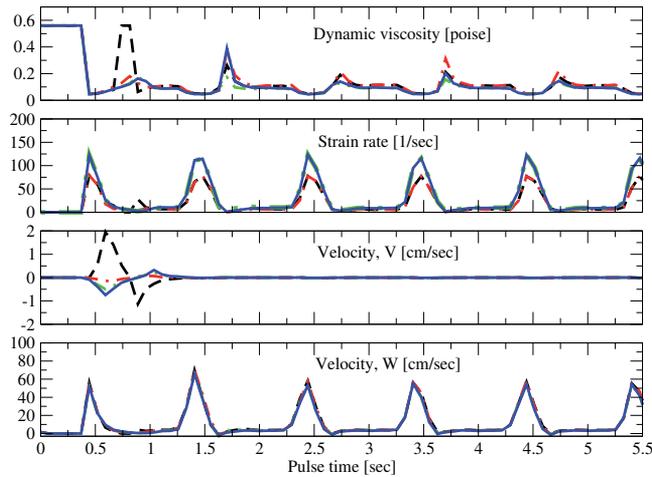


Fig. 2. Test of numerical convergence. Time-dependent dynamic viscosity, strain rate and velocity components V and W . Results for a vessel of simple geometry - cylinder type, for a specific spatial point inside the cylinder - the middle point. No turbulence effects are involved in these simulations with the realistic non-Newtonian viscosity of human blood. Black dashed line: calculations with 0.08 size for all cells FLOW3D (2007), red dot-dashed line with 0.07, green double dot - dashed line with 0.065, and blue bold line calculations with 0.062 size for all cells.

mathematical boundary conditions. The size of the tube is: $L = 8$ cm (in length) and $R = 0.34$ cm (length of inner radius). The thickness of the vessel wall is $s = 0.03$ cm. We have applied 5.5 cycles of blood pulse.

Let us now evaluate the expression (20). In these calculations we followed the work Cho & Kensey (1991), where the Carreau model of the human blood has also been used. To be consistent with Cho & Kensey (1991) we choose the following coefficients: $\lambda_2 = \lambda_{00} = 0$, $a = 0$ and $E_T = 1$, that is we don't take into account the temperature dependence of the viscosity. This investigation is to be addressed in our subsequent work. Next: $\lambda_0 = 1$, $\lambda_1 = 3.313$ sec, $\mu_\infty = 0.0345$ P, $\mu_0 = 0.56$ P, and $n = 0.3568$. The convergence was achieved when we used 52,800 cells, that is we used 100 points over OZ , 22 points over the radius of the inside space $R = 0.34$ cm, and 24 points over azimuthal angle Φ from 0 to 2π .

Time-dependent results for dynamic viscosity, strain rate and velocity components V and W are presented in Fig. 2. The turbulent effects are not taken into account. We decided to present only one precise geometrical point for comparison purposes: the middle point: $r = \theta = 0$, and $Z = 4.0$ cm. The data for Fig. 2 were obtained with the non-Newtonian model of human blood. We refer the reader to the comments provided for the figure. We were able to closely replicate the values for all previous cell sizes FLOW3D (2007) and obtain almost identical values, for example for pressure, wall shear stress and other parameters, for 0.065 mm and 0.062 mm cell sizes FLOW3D (2007). This means, that the convergence has been achieved.

Next, it would be very interesting to compare the results calculated with and without the turbulent effect. To support this endeavor we used the realistic non-Newtonian model of blood viscosity, the pulsatile flow, and the size of computation cells at which convergence has been achieved, that is the 0.062 mm size for all computational cells FLOW3D (2007). The results are presented in Fig. 3. As we see from Fig. 3 the effect of the turbulence is

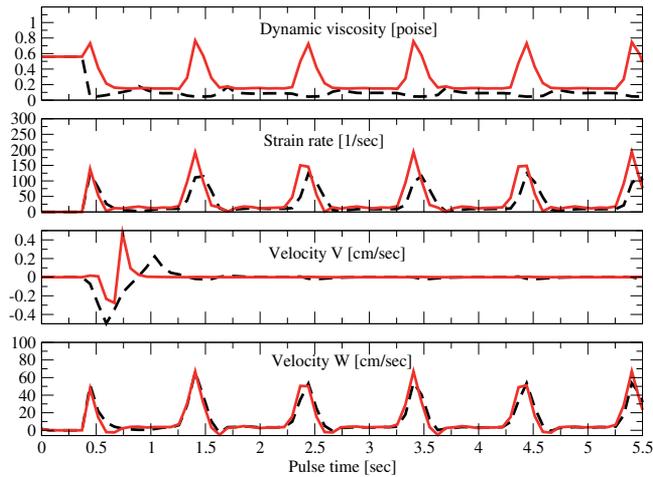


Fig. 3. Time-dependent results for a specific geometrical point inside the cylinder: the middle point. Black dashed line: simulations without taking into account the turbulence; red bold line results with the turbulence. The non-Newtonian viscosity is taken into account.

significant, particularly in regard to dynamic viscosity and strain rate. This result means that in the case of pulsatile flows and non-Newtonian viscosity the turbulent term should be taken into account. In Fig. 4 we separately show the results for the pulsatile pressure distribution and the turbulent energy, again using the middle point of the cylinder.

Finally, it would also be very interesting to make a comparison between the results calculated using both a Newtonian and non-Newtonian viscosity. However, as in previous simulations, we will apply the pulsatile flow with the turbulence included, since it has proved to be important. The results are shown in Fig. 5. As one can see, we obtain significant differences between these two calculations. We specifically observed that for the pressure distribution, dynamic viscosity and turbulent energy, we obtained significant variation.

Thus, we arrive at the important conclusion: within a time-dependent (pulsatile) flow of human blood it is necessary to take into account turbulence and non-Newtonian viscosity. The bold lines are results with the non-Newtonian viscosity (20) and the dashed lines are results with the Newtonian model when the viscosity μ has a constant value and is equal to 0.0345 P. As one can see the results are different for strain rate distributions and very different for pressure distributions. These results clearly indicate that in most cases when computer simulations are used in regard to human blood flow only the non-Newtonian model should be used.

3.2 Hemodynamics in the coronary bifurcation

Below we show the result of a subsequent simulation involving a 90° bifurcated coronary artery in Figs. 6 and 7. The geometrical model of the bifurcation consisted of a 90° intersection of two cylinders. This model represents the bifurcation between the left anterior descending coronary artery and the circumflex coronary artery. In our opinion, in the case of pulsatile flow it is more interesting to present results in a time-dependent way. This method can provide a wider picture of highly non-stationary flow systems. In this paper, because of space limitations, we just included time-dependent results for pressure, dynamic viscosity,

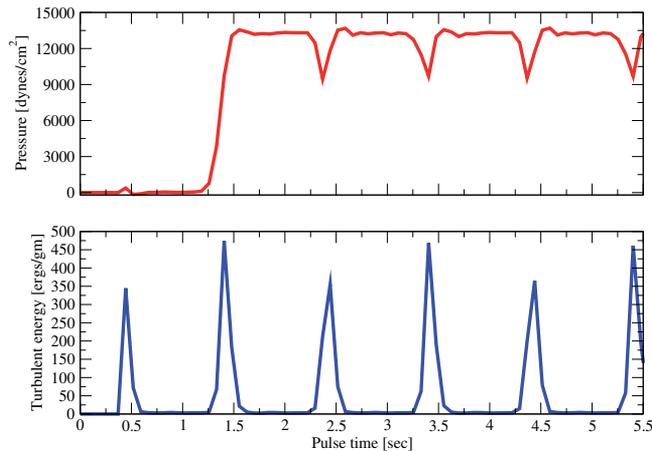


Fig. 4. Time-dependent results for pressure and the turbulent energy in the middle point of the cylinder. The non-Newtonian viscosity.

turbulent energy, and strain rate. However, we understand, that results which depend on spatial coordinates (r, θ, Z) for a few fixed moments of time are also highly useful.

In the case of the bifurcation shown in Fig. 6, we report the results for only two spatial points, which are the two outflow sides: the far right side and the farthest upper side of the bifurcation. The length of the lower horizontal vessel is 4 cm and its diameter is 0.54 cm. The length of the upper vertical vessel is 1.2 cm and its diameter is 0.4 cm. These sizes are consistent with average size human vessels.

Further, Fig. 6 represents our time-dependent results for the two outflow sides mentioned above. These results are for pressure, dynamic viscosity, turbulent energy and strain rate. The bold black lines are the results for the right outflow side, and the red dashed lines are the results for the farthest upper side (see comments to Fig. 6). In conclusion, the main goal of these calculations is to adopt them to investigate a case in which a stent is implanted in the bifurcation area Frank et al. (2002).

In Fig. 7 blood flows in from the left to the right with the imposed initial velocity profile taken from Fig. 1. The pressure, strain rate and turbulent energy distributions are shown for only one specific time moment $t=4.329$ s. The velocity vectors are also shown on these plots.

3.3 Blood flow in aortic arch

The geometry of the blood simulations inside the human aortic arch is shown in Figs. 8 and 9. On the top of the aortic arch three arteries are included. These arteries deliver the blood to the carotid artery and then to the brain. This configuration only models and approximately represents the real aortic arch. One of the goals of our simulations is to reveal the physics of the blood flow dynamics in this important portion of the human cardiovascular system.

The aortic arch is represented as a curved tube. The outer radius of the tube is 2.6 cm. A straight vessel (tube) is also merged to the arch. The length of the straight tube is about 4 cm. Again, the thickness of the wall is 0.03 cm, and the inner radius of the tube is $r = 0.34$ cm. The thickness is not important in these simulations, but it will be useful when, in future works, we will need to introduce elasticity of the walls of the tubes. The FLOW3D program allows to carry out fluid dynamic simulations with elastic (not only hard body) walls.

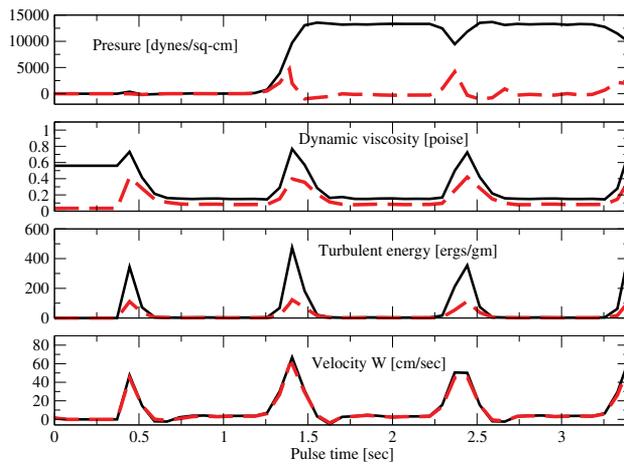


Fig. 5. Results for pressure, dynamic viscosity, turbulent energy and velocity W . Time-dependent results for the middle point of the cylinder. Bold black line calculations with non-Newtonian viscosity of the human blood; red dashed line with its Newtonian approximation.

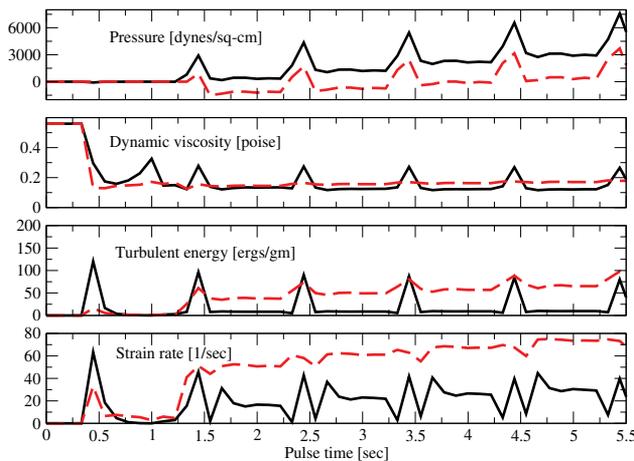


Fig. 6. Time-dependent results for a vessel with bifurcation. Pulsatile blood flow, non-Newtonian viscosity, and the turbulence effect is included. Bold black line: results for the far right outflow side $z = 0.0$; red dashed line results for the farthest up outflow side $y = 0.0$. Convergence test results for velocity components: U , V , W .

Once again we are using the Cartesian coordinate system. We also carried out a convergence test. To better represent the shape of the arch we applied five Cartesian sub-coordinate systems in our FLOW3D simulations. After the discretization the total number of all cubic cells reached about 900,000. It is important that once again we obtained full numerical convergence. Again, the geometry is shown in Figs. 8 and 9. In this work we computed pressure, velocity and strain rate distributions in the arch, while the human blood is treated as a non-Newtonian liquid and while the realistic pulsatile blood flow is used.

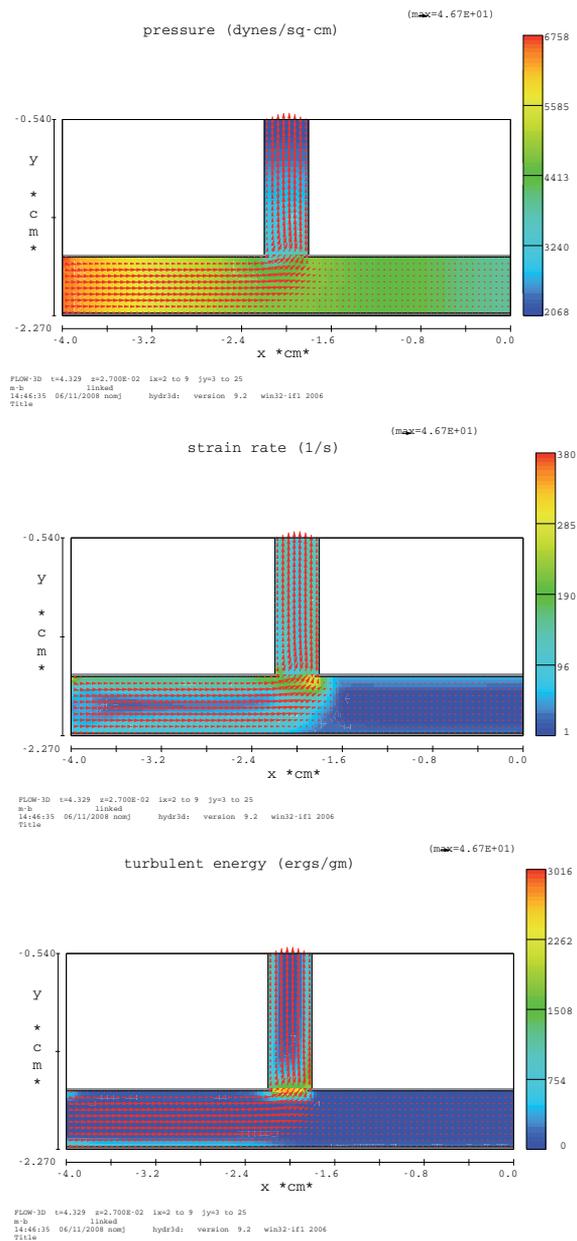


Fig. 7. The figures are 2D-plots showing the blood flow in the bifurcated vessels for only one precise moment of the discretized time $t_i = 4.329$ sec, the corresponding index is $i = 40$. Upper plot represents the result for the pressure distribution in the bifurcation, and the pressure ranges from 2068 dynes/sq-cm to 6758 dynes/sq-cm. The middle plot represents the results for the strain rate distribution and the lower plot shows results for the turbulent energy in the bifurcation. The range of the values is also shown.

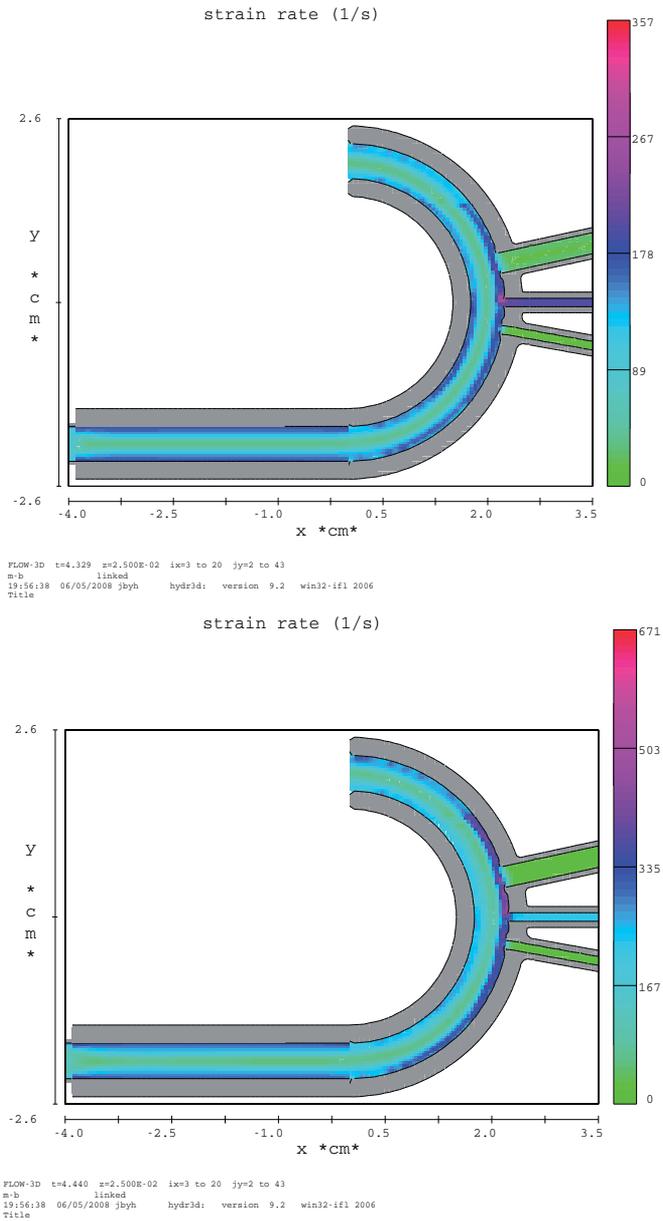


Fig. 8. Blood flow in the aortic arch. These two plots represent the full 2D-picture of the geometry used in these simulations. Shaded results for the strain rate are also shown, the bars on the right show the values. Results are for two specific moments of the time $t_{40} = 4.329$ sec and $t_{41} = 4.440$ sec. The values of the strain rate distribution range from 0.0 1/sec to 357.0 1/sec (upper plot) and from 0.0 to 671 1/sec (lower plot). The maximum values of the strain rate are localized in the region inside the arch. Blood flows from right to left in both pictures.

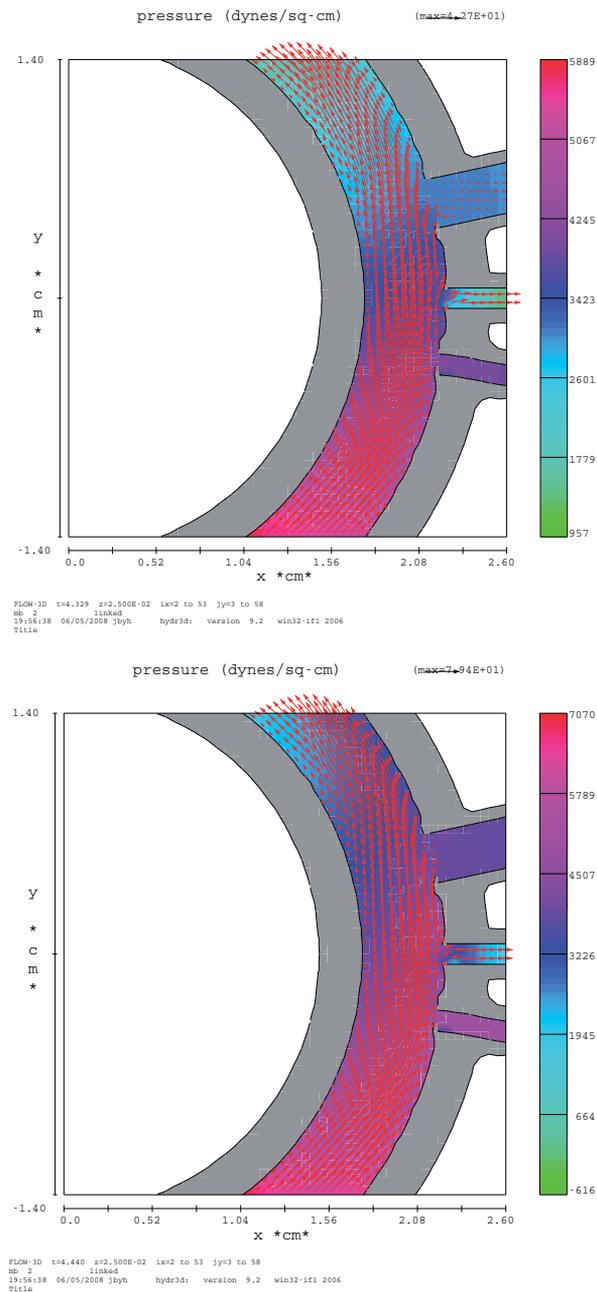


Fig. 9. Blood flow in the aortic arch. These two plots represent in more detail the region of the arch together with shaded results for the pressure distribution. The bars on the right show the values. These results are for two specific moments of the time $t_{40} = 4.329$ sec and $t_{41} = 4.440$ sec, where the pressure ranges from 957 dynes/sp-cm to 5889 dynes/sq-cm (upper plot), and from -616 dynes/sq-cm to 7070 dynes/sq-cm (lower plot).

In Fig. 8 we present the results of strain rate distributions inside the arch for two specific time moments. At the most left point, which is the inlet, we specify the pulsatile velocity source as the initial condition, that is the data from Fig. 1 are used. From the general theory of fluid mechanics Landau & Lifshitz (1959) it is possible to determine together with the blood density, viscosity, and spatial geometries, the dynamics of the blood according to the Navier-Stokes equation and its boundary conditions. Small vectors indicate the blood velocity. As can be seen from Fig. 8 blood flows from left to right in direction. However, because of pulsatility blood flows in the opposite direction too.

The values of the strain rate are also shown. These values are strongly oscillating. From the plots one can conclude that in the region of the arch the strain rate values become much larger than in the region of the straight vessel. This result represents clear evidence that in this part of the human vascular system atherosclerotic plaques should localize less than in the straight vessels. However, the higher wall shear stress values in the aortic arch could be the reason for sudden mechanical disruption of the arterial wall in this part of the human vascular system. These results are consistent with laboratory and clinical observations. In Fig. 9 we depict the pressure distribution in the arch.

4. Conclusion

In this work we applied computational fluid dynamics techniques to support pulsatile human blood flow simulations through different shape/size vessels and the aortic arch. The realistic blood pulse has been adopted and applied from the work Papaharilaou et al. (2007). The geometrical size of the vessels and the aortic arch have been selected to match the average real values. Human blood was treated in two different ways: (a) as a Newtonian liquid when the viscosity of the blood has a constant value, and (b) as a non-Newtonian liquid with the viscosity value represented by the equation (20). The numerical coefficients in (20) have been taken from work, Cho & Kensey (1991).

It is always difficult to obtain a steady-state cycle profile and stable computational results at the very beginning of time-dependent simulations. However, after a short stabilization period a steady-state cycle profile can be obtained. In our simulations we used up to 5.5 pulse cycles to reach complete steady state profiles. We obtained valid results for pressure, wall shear stress distribution and other physical parameters, such as the three velocity components of blood flow. All of these were shown in Figs. 2-6.

Our simulations showed that the FLOW3D program is capable of providing stable numerical results for all geometries included in this work. The time-dependent mathematical convergence test has been successfully carried out. Particular attention has been paid to this aspect of the calculations. It is a well known fact that fluid dynamics equations can have unstable solutions, Landau & Lifshitz (1959). Therefore, numerical convergence has been tested and confirmed in this work.

The result of computer simulations of blood flow in vessels for three different geometries have been presented. For pressure, strain rate and velocity component distributions we found significant disagreements between our results obtained with the realistic non-Newtonian treatment of human blood and the widely used method in literature: a simple Newtonian approximation.

Our results are in good agreement with the conclusions of the works, Chen & Lu (2004; 2006), where the authors also obtained significant differences between their results calculated with and without the non-Newtonian effect of blood viscosity. However, the recent work, Boyd & Buick (2007) should be mentioned, in which the authors performed 2-dimensional simulations of human blood flow through the carotid artery with and without the non-Newtonian effect of

the viscosity. They did not find any substantial differences in their results. Finally, we would like to mention the paper Agarwal et. al. (2008), where the authors also performed simulations for the carotid artery, but only the non-Newtonian viscosity was used.

Next, the influence of a possible turbulent effect has also been investigated in this work. It was found that the effect is important. We believe, that the physical reason of this phenomena lies in the strong pulsatility of the flow and in the non-Newtonian viscosity of the blood. The contribution of the turbulence is most significant in the area of bifurcated vessels.

Finally, a significant increase of the strain rate and, the wall shear stress distribution, is found in the region of the aortic arch. This computational result provides additional evidence to support recent clinical and laboratory observations that this part of the human cardiovascular system is under higher risk of disruption Carter et al. (2001); Pochettino & Bavaria (2006). In future work it would be interesting to include the elasticity of the walls of the aortic arch Fang et al. (1998) and other vessels.

In conclusion, we would like to specifically point out, that the developments in this work can be directly applied to even more interesting and very important situations such as when a stent is implanted inside a vessel Frank et al. (2002). In this case, for example, it would be very useful to determine blood flow disturbance, the pressure distribution, strain rate and values of other physical parameters. The results of this work should allow us to determine the optimal size and shape of effective stents. As we mentioned in the Introduction some research groups are carrying out laboratory and computer simulations of blood flow through vessels with implanted stents Frank et al. (2002). It is very difficult to underestimate the value of these works.

5. Acknowledgments

This work was partially supported by Office of Sponsored Programs (OSP), by Internal Grant Program of St. Cloud State University, St. Cloud, MN-56301-4498, USA, and by a private Minnesota based company: *RIE Coatings, Eden Valley, MN-55329-1646, USA* (www.riecoatings.com).

6. Appendix

For a variable dynamic viscosity μ , the viscous accelerations are

$$\rho V_F f_x = w_x^s - \left[\frac{\partial}{\partial x} (A_x \tau_{xx}) + R \frac{\partial}{\partial y} (A_y \tau_{xy}) + \frac{\partial}{\partial z} (A_z \tau_{xz}) + \frac{\xi}{x} (A_x \tau_{xx} - A_y \tau_{yy}) \right] \quad (22)$$

$$\rho V_F f_y = w_y^s - \left[\frac{\partial}{\partial x} (A_x \tau_{xy}) + R \frac{\partial}{\partial y} (A_y \tau_{yy}) + \frac{\partial}{\partial z} (A_z \tau_{yz}) + \frac{\xi}{x} (A_x + A_y \tau_{xy}) \right] \quad (23)$$

$$\rho V_F f_z = w_z^s - \left[\frac{\partial}{\partial x} (A_x \tau_{xz}) + R \frac{\partial}{\partial y} (A_y \tau_{yz}) + \frac{\partial}{\partial z} (A_z \tau_{zz}) + \frac{\xi}{x} (A_x \tau_{xz}) \right], \quad (24)$$

where

$$\tau_{xx} = -2\mu \left(\frac{\partial u}{\partial x} - \frac{1}{3} \left(\frac{\partial u}{\partial x} + R \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} + \frac{\xi u}{x} \right) \right) \quad (25)$$

$$\tau_{yy} = -2\mu \left[R \frac{\partial v}{\partial x} + \xi \frac{u}{x} - \frac{1}{3} \left(\frac{\partial u}{\partial x} + R \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} + \frac{\xi u}{x} \right) \right] \quad (26)$$

$$\tau_{zz} = -2\mu \left(\frac{\partial w}{\partial z} - \frac{1}{3} \left(\frac{\partial u}{\partial x} + R \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} + \frac{\xi u}{x} \right) \right) \quad (27)$$

$$\tau_{xy} = -\mu \left(\frac{\partial v}{\partial x} + R \frac{\partial u}{\partial y} - \frac{\zeta v}{x} \right) \quad (28)$$

$$\tau_{xz} = -\mu \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right) \quad (29)$$

$$\tau_{yz} = -\mu \left(\frac{\partial v}{\partial z} + R \frac{\partial w}{\partial y} \right). \quad (30)$$

In the above equations (22)-(24) the terms w_x^s , w_y^s and w_z^s are wall shear stresses. If these terms are equal to zero, there is no wall shear stress. This is because the remaining terms contain the fractional flow areas (A_x, A_y, A_z) which vanish at the walls FLOW3D (2007).

The wall stresses are modeled by assuming a zero tangential velocity on the portion of any area closed to flow. Mesh and moving obstacle boundaries are an exception because they can be assigned non-zero tangential velocities. In this case the allowed boundary motion corresponds to a rigid body translation of the boundary parallel to its surface. For turbulent flows, a law-of-the-wall velocity profile is assumed near the wall, which modifies the wall shear stress magnitude.

7. References

- Agarwal, R., Katiyar, V.K. & Pradhan, P. (2008). A mathematical modeling of pulsatile flow in carotid artery bifurcation, *Intern. J. of Engineering Science* 46 (11), pp. 1147-1156.
- Benard, N., Perrault, R. & Coisne, D. (2006). Computational Approach to Estimating the Effects of Blood Properties on Changes in Intra-Stent Flow, *Annals of Biomed. Engineering* 34 (8), pp. 1259-1271.
- Banerjee, R.K., Devarakonda, S.B., Rajamohan, D. & Back, L.H. (2007). Developed Pulsatile Flow in a Deployed Coronary Stent, *Biorheology* 44 (2), pp. 91-102.
- Boyd, J. & Buick, J.M. (2007). Comparison of Newtonian and non-Newtonian Flows in a Two-Dimensional Carotid Artery Model Using the Lattice Boltzmann Method, *Physics in Medicine and Biology* 52(20), pp. 6215-6228.
- Carter, Y.M, Karmy-Jones, R.C., Oxorn, D.C. & Aldea, G.S. (2001). Traumatic Disruption of the Aortic Arch, *European J. Cardiothoracic Surgery*, 20, pp. 1231.
- Chen, J. & Lu, X.-Y. (2005). Numerical investigation of the non-Newtonian blood flow in a bifurcation model with a non-planar branch *Journal of Biomechanics*, 37 (12), pp. 1899-1911.
- Chen, J. & Lu, X.-Y. (2006). Numerical investigation of the non-Newtonian pulsatile blood flow in a bifurcation model with a non-planar branch, *Journal of Biomechanics* 39 (5), pp. 818-832.
- Cho, Y.I. & Kensey, K.R. (1991). Effects of the non-Newtonian viscosity of blood on flows in a diseased arterial vessel. Part 1: Steady Flows, *Biorheology* 28(3-4), pp. 241-262.
- Dhein, S., Delmar, M., & Mohr, F.W. (2005). *Practical Methods in Cardiovascular Research*, Springer-Verlag New-York, LLC.
- Duraiswamy, N., Schoepfoerster, R.T., Moreno, M.R. & Moore Jr., J.E. (2007). Stented Artery Flow Patterns and Their Effects on the Artery Wall, *Ann. Rev. of Fluid Mechanics* 39, pp. 357-382.
- Faik, I., Mongrain, R., Leask, R.L., Rodes-Cabau, J., Larose, E. & Bertrand, O. (2007). Time-Dependent 3D Simulations of the Hemodynamics in a Stented Coronary Artery, *Biomedical Materials* 2 (1), art. no. S05, S28-S37.
- Fang, H., Lin, Z. & Wang, Z. (1998). Lattice Boltzmann Simulation of Viscous Fluid Systems with Elastic Boundaries, *Phys. Rev. E* 57(1), R25-R28.
- FLOW-3D Users Manual. (2007). Version 9.2, Flow Science, Santa Fe, New Mexico, USA.

- Frank, A.O., Walsh, P.W., & Moore Jr., J.E. (2002). Computational Fluid Dynamics and Stent Design, *Artificial Organs* 26(7), pp. 614-621.
- Fung, Y.C. (1993). *Biomechanics*, Springer.
- Grotberg, J.B. & Jensen, O.E. (2004). Biofluid Mechanics in Flexible Tubes, *Ann. Rev. Fluid Mech.* 36, pp. 121-147.
- Huo, Y. & Kassab, G.S. (2006). Pulsatile Blood Flow in the Entire Coronary Arterial Tree: Theory and Experiment, *Am. J. Physiol. Heart Circ. Physiol* 291(3) pp. H1074-H1087.
- Johnston, B.M., Johnston, P.R., Corney, S. & Kilpatrick, D. (2004). Non-Newtonian Blood Flow in Human Right Coronary Arteries: Steady State Simulations, *J. of Biomech.* 37(5), pp. 709-720.
- Landau, L.D. & Lifshitz, E.M. (1959). *Fluid Mechanics*, Volume 6, Pergamon Press Ltd.
- Morris, L., Delassus, P., Walsh, M. & McGloughlin, T. (2004). A mathematical model to predict the in vivo pulsatile drag forces acting on bifurcated stent grafts used in endovascular treatment of abdominal aortic aneurysms (AAA): *J. of Biomechanics* 37(7), pp. 1087-1095.
- Morris, L., Delassus, P., Callanan, A., Walsh, M., Wallis, F., Grace, P. & McGoughlin, T. (2005). 3-D Numerical Simulation of Blood Flow Throught Models of the Human Aorta, *J. of Biomech. Engineering* 127(5), pp. 767-775.
- Mukundakrishnan, K., Ayyaswamy, P.S. & Eckmann, D.M. (2008). Finite-sized gas bubble motion in a blood vessel: Non-Newtonian effects, *Phys. Rev. E* 78(3), art. no. 036303.
- Papaharilaou, Y., Ekaterinaris, J.A., Manousaki, E. & Katsamouris, A.N. (2007). A Decoupled Fluid Structure Approach for Estimating Wall Stress in Abdominal Aortic Aneurysms, *J. Biomechanics* 40(2), pp. 367-377.
- Peskin, C.S. (1977). Numerical Analysis of Blood Flow in the Heart, *Journal of Computational Physics*, 25 (3) pp. 220-252.
- Pochettino, A. & Bavaria, J.E. (2006). Aortic Dissection, in Book: *Mastery of Cardiothoracic Surgery*, Eds. L.R. Kaiser, I.L. Kron, T.L. Spray, Publisher: Lippincott Williams and Wilkins, pp. 534-544.
- Seo, T., Schachter, L.G. & Barakat, A.I. (2005). Computational Study of Fluid Mechanical Disturbance Induced by Endovascular Stents, *Annals of Biomed. Engineering* 33 (4), pp. 444-456.
- Sultanov, R.A., Guster, D., Engelbrekt, B. & Blankenbecler, R. (2008). A Full Dimensional Numerical Study of Pulsatile Human Blood Flow in Aortic Arch, *Proceedings of the 2008 International Conference on Bioinformatics and Computational Biology*, Vol. 2, pp. 437-443, Eds. H.M. Arabnia, M.Q. Yang, J.Y. Yang, CSREA Press-WORLDCOMP;
- Sultanov, R.A., Guster, D., Engelbrekt, B., & Blankenbecler, R. (2008). 3D Computer simulations of pulsatile human blood flows in vessels and in aortic arch: investigation of non-Newtonian characteristics of human blood, *Proceedings of the 2008 11-th IEEE International Conference on Computational Science and Engineering*, art. no. 4578268, pp. 479-485, IEEE Comp. Soc.
- Sultanov, R. A. & Guster, D. (2009). Full dimensional computer simulations to study pulsatile blood flow in vessels, aortic arch and bifurcated veins: Investigation of blood viscosity and turbulent effects, *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009*, art. no. 5334202, pp. 4704-4710.
- Taylor, C.A. & Draney, M.T. (2004). *Experimental and Computational Methods in Cardiovascular Fluid Mechanics*, Annual Review of Fluid Mechanics 36, pp. 197-231.
- Waite, L. (2005). *Biofluid Mechanics in Cardiovascular Systems*, Mc-Graw-Hill Professional Publishing.

Biomechanical Factors Analysis in Aneurysm

Kleiber Bessa¹, Daniel Legendre² and Akash Prakasan²

*¹Department of Environmental Sciences and Technological
Rural Federal University of Semi-Arid*

*²Institute Dante Pazzanese of Cardiology
Brazil*

1. Introduction

All the cells in the body need to receive food (nutrients, metabolic products) and to dispose of waste products. The responsible system for that is cardiovascular system. It is responsible to supply food through the arteries and to return waste products through the veins for all living cells in the human body. This task is reached by a circulating fluid, the blood. The central location which all lines of supply originate from and return to is a small, very small, pump, the heart. The heart keeps the fluid in circulation. In the heart, there are two pumps, propelling blood into the pulmonary and systemic circulation and are combined into a single muscular organ to synchronously beat. Any disruption in the blood flow causes a disruption in food supply. Life is not possible without blood, but in the truth life is not possible without the circulation of blood. It must pump at all times, which it does by contracting and relaxing in a rhythmic pattern, approximately once every second, more than 86 thousand times every day, and about 2 billion times in a lifetime of 75 years, nonstop (Zamir, 2005). The blood ejected by the heart follows in the direction the arterial tree. Along the arterial tree, the arteries successively decrease in size, increase in number, undergo structural changes, and finish in arterioles that are as little as 10 μm in diameter. The structure of the artery is quite complex. The main components of the vessel wall are endothelium, smooth muscle cells, elastic tissue, collagen, and connective tissue. The arteries are targets for diseases such as atherosclerosis or aneurysms that each year claims the lives of scores of people worldwide. The cardiovascular disease may be triggered or aggravated by mechanical stimuli, such as wall stress or stretch resulting from the blood pressure, or shear stress resulting from the blood flow (Wernig and Xu, 2002). Arteries can also adapt to long-term physiological conditions by thinning or thickening the muscular layer, and altering the relative composition and organization of the various assemblies of structural proteins in process generally know as remodelling. Bessa et al. (2011) showed that occurs remodelling in tail arterial bed from normotensive and hypertensive rats. As shown in Figure 1, the internal diameter of the proximal portion of the tail artery did not differ between Wistar rats and spontaneously hypertensive rats (SHR), whereas the diameter of the intermediate and distal portions of SHR tails arteries were significant smaller than those of normotensive rats.

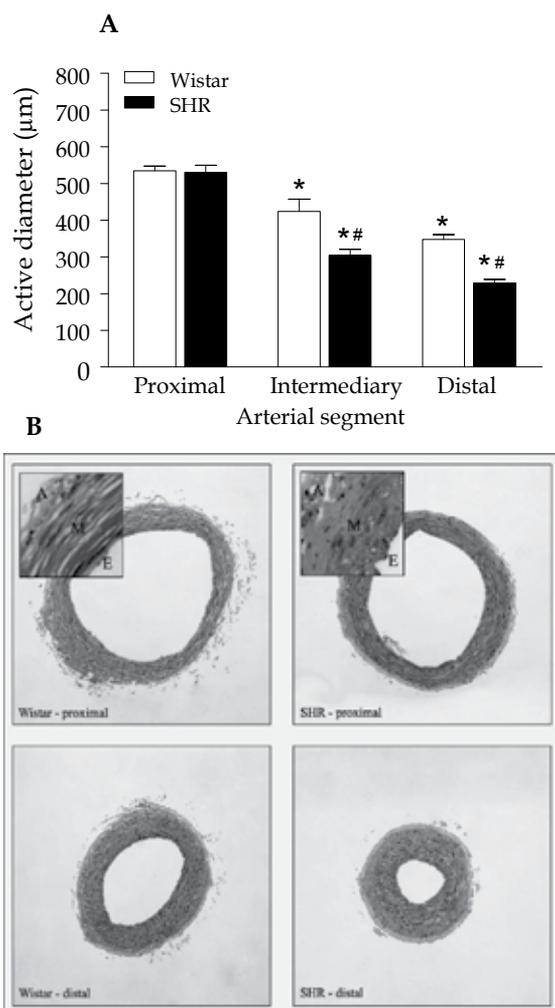


Fig. 1. A, Active internal diameters of tail arteries from Wistar rats (N=4animals) and spontaneously hypertensive rats (SHR, N=5 animals) fixed at a volumetric flow rate of 2.5 mL/min. *P < 0.05 vs proximal portion. #P < 0.05 vs wistar rats (two-way ANOVA followed by the Bonferroni *post hoc* test). B, Photomicrograph of transverse section of proximal and distal tail artery from Wistar rats and SHR. Magnification: 10X. Inset panels in the proximal tail artery photomicrographs of Wistar and SHR are amplified images showing the endothelial (E), medium (M) and adventitial (A) layers of the artery. Extracted from Bessa et al. 2011.

An aneurysm is defined as a focal dilation of a blood vessel when compared with the original artery. Aneurysms are widening of the lumen in any artery, most commonly in the aorta for fusiform types, and in the head for saccular types. The arterial fusiform aneurysms and intrinsic stenosis are possible complications of atheroma. Saccular aneurysms can occur as complications of arterial wall trauma or infection. Intracranial aneurysms (IA), rare in childhood and adolescence, are observed in 3% to 5% of the population, with a gender ratio

of 3 women for 2 men. Abdominal aortic aneurysms (AAA) most of the affect men between 40 and 70 years old (5-7% of people older than 60) (Thiriet, 2008). Genetics and risk factors play important roles in the development of the aneurysms, but is universally accepted that biomechanical factors (including increased pressures in hypertension) also play fundamental roles. Vascular endothelial and smooth muscle cell are constantly exposed to the biomechanical factors caused by the blood flow. Cellular responses to these biomechanical factors influence vessel wall homeostasis (Hisai, 2008). Once the aneurysm forms, the biomechanical factors caused by the pulsatile flow in the aneurysm can cause gradual expansion of it. When the wall of the distended artery fails to support the stress occurs rupture of the aneurysm. This rupture often leads to death or severe disability.

This chapter introduces analysis about pathogenesis, hemodynamic forces acting on the wall vessel that could be important factor to the origin and progression of the disease, and computational fluid dynamic (CFD) associated to the medical imaging. Thus, the interaction between pathogenesis of aneurysm, medical imaging and CFD are important to understand the development of aneurysm.

2. Pathogenesis of arterial aneurysms

2.1 General considerations

An arterial aneurysm is one of the most common vascular diseases causing disability and death. True aneurysm represents a degeneration of the artery wall with loss of structural integrity leading to gradual dilatation of all artery wall layers. Aneurysms have been reported in almost all segments of the arterial tree, but are more frequently localized at the aorta below the renal arteries. Aneurysms are most commonly diagnosed in the sixth and seventh decades of life, with a rising incidence for unknown reasons. The prevalence of aneurysms in a given population depends on the presence of risk factors associated, including older age, male gender, white race, positive family history, smoking, hypertension, hypercholesterolemia, peripheral vascular occlusive disease, and coronary artery disease. Although these risk factors are associated with increased abdominal aortic aneurysms (AAA), they may not be independent predictors and may be markers rather than causes of AAA prevalence.

Although a precise cause of AAA remains unknown, much has been learned about the pathophysiology of the aneurysmal aorta. Research has linked the development of AAA with chronic aortic wall inflammation, increased local expression of endogenous proteinases, and the degradation of structural connective tissue proteins (Shah, 1997).

Most arterial aneurysms arise at the bifurcation of major arteries, and this is also true for the intracranial location. Around 85% of all intracranial aneurysms originate from the anterior circulation. The most common location (30%–35%) is the anterior communicating artery. The prevalence of intracranial aneurysms among first-degree relatives of patients with cerebral aneurysms is higher than in the general population. The risk for a first-degree relative harbouring an aneurysm is about three to four times higher than for someone from the general population (Raaymakers 1999; Ronkainen et al. 1997).

Although the pathogenesis and etiology of cerebral aneurysms has been studied extensively, both are still poorly understood. Endogenous factors like elevated blood pressure, the special anatomy of the Circle of Willis or the effect of hemodynamic factors, particularly originating at vessel bifurcations, are all known to be involved in the growth and rupture of an aneurysm. Arteriosclerosis and inflammatory reactions, however, might also have an

impact. Exogenous factors like cigarette smoking, heavy alcohol consumption or certain medications are thought to be risk factors in the pathogenesis of an aneurysm or at least increase the risk of rupture.

2.2 Basic mechanisms

The normal aortic wall is composed of a thin endothelial lined intima, a thick elastic media dominated by vascular smooth muscle cells, and a fibrocollagenous adventitia. Medial elastin fibers and interstitial collagens are normally responsible for the tensile strength, resilience, and structural integrity of the aortic wall. Aneurysmal dilatation and rupture are due to mechanical failure of these fibrillar extracellular matrix proteins.

Basic research on the pathophysiologic processes of aortic aneurysm formation and growth has identified several putative mediators of the disease. These mediators include both bacteriologic and enzymatic agents, especially matrix metalloproteinases. Aneurysm formation and aneurysm growth represent two distinct phases of aortic aneurysm disease. Aneurysm formation encompasses the metabolic processes that cause degradation of the structural components of the aortic wall and loss of biomechanical integrity. This phase of aneurysm disease appears to be directed at the destruction of elastin within the aortic wall. Although all layers of the aortic wall may be involved, the most fundamental structural alteration that results in a loss of biomechanical integrity and aneurysm formation is degradation of adventitial elastin. The metabolic processes that cause adventitial elastin degradation are nonselective. They most likely begin in the intima or media, associated with an inflammatory cell infiltrate. Destruction of the elastin within these layers does not cause loss of the structural integrity of the aortic wall. These processes, however, ultimately invade the adventitia, cause degradation of the elastin within this layer, destroy the structural integrity of the aortic wall, and permit its pathologic expansion. Because humans are unable to synthesize and deposit elastin in the aortic wall in any detectable quantity beyond the first year of life, the degraded elastin is replaced by collagen I and III during the process of aneurysm formation. (Mesh et al., 1992) Once formed, even the smallest aneurysm represents the end-stage of aortic wall destruction, a significant alteration in the biomechanical properties of the aortic wall and a reduction in structural integrity. Aneurysm growth appears to be a self-sustaining process. Data from both human and animal aneurysm cells demonstrate a sustained increase in collagen types I and III mRNA and new collagen deposition within the aneurysm wall. These processes continue until the rate of collagen degradation exceeds the rate of deposition or the pressure per unit area of aneurysm wall exceeds the ability of the collagen fibers to withstand the load. In either of these situations, aneurysm rupture occurs.

3. Evaluation of aneurysms

3.1 Physical examination and imaging

Although most clinically significant AAA are potentially palpable during routine physical examination, the sensitivity of this technique depends on size, obesity of the patient, skill of the examiner, and focus of the examination. With physical examination alone, the diagnosis is made in 29% of AAAs 3 to 3.9 cm, 50% AAAs 4 to 4.9 cm, and 75% of AAAs 5 cm or larger (Lederle & Simel, 1999).

Several imaging modalities are available to confirm the diagnosis of AAA. Abdominal B-mode ultrasonography is the least expensive, least invasive, and most frequently used examination, particularly for initial confirmation of suspected AAA and follow up of small AAAs. Computed tomography is more expensive than ultrasound and involves exposure to radiation and intravenous contrast material, but it provides more accurate measurement of diameter, with 91% of studies showing less than 5 mm interobserver variability (Chervu et al., 1995).

3.2 Factors affecting clinical decision making

The choice between observation and prophylactic surgical repair of an AAA for an individual patient at any given time should take into account the risk for rupture under observation, the operative risk associated with repair, the patient's life expectancy, and the personal preferences of the patient. For the scope of this chapter we will only analyse the risk of rupture.

Estimates of the risk for AAA rupture are imprecise because large numbers of patients with AAAs have not been observed without intervention. Data are insufficient to develop an accurate prediction rule for AAA rupture in individual patients, which makes surgical decision making difficult. Some aspects may be considered:

- Aneurysm diameter

For the past 5 decades maximal aneurysm diameter has been the primary determinant of rupture risk. Several studies firmly established the effect of size on AAA rupture and provided a sound basis for recommending elective repair for large AAAs because of marked improvement in survival after repair. Despite differences in precise estimates, most studies show that rupture risk increase substantially with AAA diameter between 5 – 6 cm.

- Aneurysm wall stress

From a biomechanical perspective, AAA rupture occurs when the forces within a AAA exceed the wall's "bursting strength. The application of engineering principles to the analysis of actual aneurysms has only recently been possible. Multiple studies have demonstrated that finite element analysis of AAA wall stress with three dimensional CT reconstructions is better than diameter for estimating rupture risk.

- Finite element analysis

At this point the strength of the data and the size of the patient cohorts already rival or exceed that of the data initially used to determine the clinical use of aneurysm diameter to estimate rupture risk in the 1960's. The technique of using aneurysmal wall stress to predict rupture risk remains to be validated in a large multicenter cohort using a standardized, broadly applicable technique, although one such study is currently under way (Fillinger et al., 2004).

- Aneurysm shape

Clinical opinion holds that shape is important and eccentric or saccular aneurysms present a greater risk for rupture than do more diffuse fusiform aneurysms. Vorp et al. (1998) associates using computer modelling showed that wall stress is substantially increased by an asymmetric bulge in AAA. The presence of calcification in the wall may increase wall stress focally, but may not be useful as a clinical tool. The effect of intraluminal thrombus on rupture risk is also debated, with studies suggesting that thicker thrombus may increase the risk for rupture, decrease the risk for rupture, or have no effect. The practical impact of these variables on AAA rupture risk requires further study.

4. Biomechanical analysis

This section provide an overview of the fundamental basis of biomechanical factors acting on the vessel wall and their influences under endothelial cells and the development of the aneurysm, once that there are several studies showing disturbed flow conditions and unsteady turbulent stresses damage endothelial cells and may provide a first step to the degradation of the wall (Davies et al., 1984, 1995). There will be analysed biomechanical factors in intracranial aneurysm and an abdominal aortic aneurysm.

4.1 Intracranial aneurysm

At present, there is not completely satisfied the theory about the origin, growth and rupture of intracranial aneurysm. The walls of intracranial arteries exhibit the same general organization and composition of all arteries. However, in most cerebral arteries, there is no external elastic lamina. Moreover, in these arteries, the arterial walls are thin, there is less elastin, and there are "medial defects" (gaps in the muscle layer) present frequently, thus intracranial arteries are more susceptible to aneurysm formation than extracranial arteries. Other arteries show similar medial defects, particularly in renal, mesenteric, splenic, and coronary arteries. However, these defects are smaller and less frequent and the aneurysm formation is rarely (Crompton, 1966). In the literature, there are some hypothesis about the location and structural changes in the wall of the artery that contribute to development of the aneurysm. Eppinger (1887)(apud Thubrikar, 2007) showed that certain aneurysms occurred at the site of the medial defects, however, other researchers, such as, Forbus (1930), Toth et al. (1998) and Zhang et al. (2003) showed that the aneurysm could be developed through combination of degeneration of the elastic and defects in the muscular layer. Meng et al. (2007) recently demonstrated in animal models that the aneurysm initiation and development occurred at the apices of arterial bifurcations where there was high wall shear stress and aneurysm-type wall remodelling (disrupted internal elastic lamina and endothelium, thinned media and smooth muscle cells loss) at histology. These factors appear to render these arteries susceptible to a local weakening under the persistent action of hemodynamic loads, particularly in hypertension (Inci and Spetzler, 2000). The hemodynamic loads resulting from the direct impingement of the central streams at the apex of bifurcations are probably the most important factor contributing to the focal degeneration of the internal elastic membrane and the early origin the aneurysms. These forces could enlarge medial defects already present. The impingement of central axial streams results in a much greater velocity gradient and shear stress at the apex than is experienced in the main stem or branches of bifurcations. As flow is pulsatile, the peak force will be great, because there is a brief impact time and in the moment of impact the kinetic energy of the moving blood is changed to pressure energy (stagnation pressure) (Thubrikar, 2007). This extra pressure is the force responsible for focal degeneration of the internal elastic membrane and thus the cause of initiation of aneurysms at the apex. Shojima et al. (2005) reported that local rises in pressure due to flow impingement are less than 2 mmHg, which is small compared with nominal pressure levels in cerebral arteries, and concluded that dynamic pressures acting at bifurcations and on the walls of intracranial arteries may be less significant to enlargement and rupture than previously assumed. Acevedo-Bolton et al (2006) concluded that regions that continued to enlarge experienced low wall shear stress and speculated that this might be due to increased residence time of particles that degrade the aneurysm wall. There is a great interest to study the wall shear stress in aneurysms because there are many works show that it plays a role in the evolution of aneurysmal

disease. It is known that the wall shear stress is a major physiological stimulus for the vessel endothelium. The endothelial cells submitted the physiological levels of wall shear stress in the arteries [1 (10) – 7 (70) Pa (dyn/cm²)] and in the veins (0.1 (1) – 0.7 (7) Pa (dyn/cm²)] provide a physiological stimulus for the vessel endothelium contributing to selective barrier for macromolecular permeability, can influence vascular remodelling via the production of growth-promoting and -inhibiting substances, modulate hemostasis/thrombosis through the secretions of procoagulant, anticoagulant, and fibrinolytic agents, mediate inflammatory responses via the surface expression of chemotactic and adhesion molecules and release of chemokines and cytokines (Malek et al., 1999; Paszkowiak & Dardik, 2003; Li et al., 2005). On the other hand, when there is excess or lack of the stimulus can lead to pathological phenomena that cause changes in the arterial wall biomechanical properties. The high values of wall shear stress can cause damage in the endothelial cells (Fry, 1968). A prolonged high wall shear stress fragments the internal elastic lamina of vessels (Masuda et al., 1999) and gives rise to the initial change involved in the formation of cerebral aneurysm. Low wall shear stress has been reported to be related to aneurysmal growth (Jou et al., 2003) and rupture (Shojima et al., 2004) as it promotes various mechanisms that cause arterial wall remodelling. When the wall shear stress is lower than 0.4 Pa, it generates endothelial proliferation (Malek et al., 1999) and apoptosis (Kaiser et al., 1997). It is also responsible for abnormal vascular reactivity and vasospasm that can cause ischemia, angina, and myocardial infarction, increased permeability to macromolecules such as lipoproteins, increased expression of chemotactic molecules and adhesion molecules, enhanced recruitment and accumulation of monocytes/macrophages in the intima as foam cells, altered regulation in growth and survival of vascular cells (Lerman & Burnett, 1992; Bonetti et al., 2003; Gimbrone et al., 2000). This excessive low wall shear stress may be one of the main factors underlying the degeneration, indicating the structural fragility of the aneurysmal wall (Shojima et al., 2004). It appears that, after an initial injury that might result from excessive wall shear stress on the endothelial cells (Meng et al., 2007), progressive changes in aneurysm shape occur with a trend for the cross section to become more elliptical (Boussel et al., 2008). This generates a progressive decrease of wall shear stress leading to endothelial dysfunction, wall remodelling, and aneurysm growth (Ahn et al., 2007; Utter & Rossman, 2007).

4.2 Abdominal aortic aneurysm

Abdominal aortic aneurysms (AAAs) most of the affect men between 40 and 70 years old (5-7% of people older than 60) (Thiriet, 2008). AAAs rarely appear in individuals under 50 years old, but their incidence increases drastically at age 55 and peaks in the early 80s. Norway in 1994-1995 showed that AAAs are present in 8.9% of men and in 2.2% of women over 60 years old (Singh et al., 2001). Reed et al. (1992) showed that AAAs increased steadily with age after 60 years old and aortic dissections (aortic wall dissects and blood enters the wall causing the enlargement) showed peak between ages 70 and 75 years old and decreased after that (Figure 2).

Although aneurysms may develop throughout ht length of the aorta, AAAs are at least 5 times more prevalent than thoracic or thoracoabdominal aneurysms (Dua & Dalman, 2007). This fact can be explained from two points: physiologic and anatomic features unique to the distal aorta. The infrarenal aorta is the most common site of extracranial aortic aneurysm formation. Hemodynamic through the aorta combined with regional factors can explain this preferential distribution. In the infrarenal aorta the number of elastic lamellae (and therefore

elastin) is markedly decreased in comparison with the thoracic aorta which become fragmented and unorganized (Lakatta et al., 1987; MacSweeney et al., 1994) contributing to reducing elasticity and wall motion (Ailawadi et al., 2003). The degeneration of elastic fibers is accompanied by an increase in the collagenous substance (the stiffer structural component). As the ratio of elastin to collagen decreases, the vessel progressively loses its elasticity. The stiffening of the wall causes an increase in speed of the pulse wave. For example, in the aorta, the wave speed increases from 6.5 m/s in a 10-year-old child to upwards of 11 m/s in a 60-year-old adult (Nichols & O'Rourke, 1990). Reduced distal aortic elasticity, in combination with augmented pressure due to pulse wave reflections from the aortic bifurcation and other downstream arteries, may increase wall strain and aneurysm susceptibility (Humphrey and Taylor, 2008). The collagen-to-elastin ratio is the principal determinant of wall mechanics in the aorta. Changes in composition and structure of the arterial wall will alter the wall mechanics. An increase in collagen-to-elastin ratio results in a higher wall stiffness and lower tensile strength. Clinical observations show that most AAA walls become progressively stiffer as the diameter increases. This is because of biomechanical restructuring of the wall (Kleinstreuer et al., 2007). In the normal abdominal aorta, the collagen-to-elastin ratio is approximately 1.58 (Nichols and O'Rourke, 1990), however, the collagen-to-elastin ratio is much higher in AAAs (table 1).

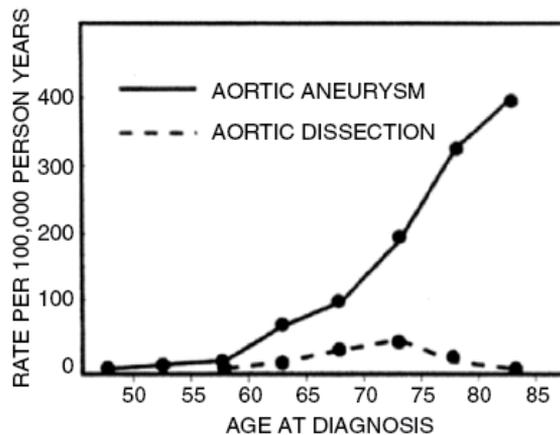


Fig. 2. Plots of incidence rates of aortic aneurysm and dissection by age at diagnosis. Extracted from Reed et al., (1992).

As commented above, hemodynamic of the infrarenal is one the main factors to cause the development of AAAs. Hemodynamic forces relevant to AAA pathogenesis can be obtained into three components: 1) wall shear stress (explained above), 2) hydrostatic pressure, the perpendicular force acting on the vascular wall; and 3) relative wall strain (RWS), the circumferential stretch of the vessel wall exerted by cyclic luminal pressure changes and resulting tensile stress. Several works in the literature showed that cultured vascular endothelial cells studies when submitted the disturbed flow conditions and unsteady turbulent stresses damage the endothelium, and the loss or malfunctioning of their regulatory processes may provide a first step to the degradation of the wall (Davies et al. 1984, 1995, 2009; Chiu and Chien, 2011). Hemodynamic conditions vary markedly along the aorta, from high Reynolds numbers (Re)

$$\text{Re} = \frac{\rho V d}{\mu} \quad (1)$$

where ρ is the specific mass, V is media velocity, d is the diameter of the vessel and μ is the viscosity dynamic, at the aortic root to low and oscillatory shear conditions at the aortic bifurcation (Greve et al., 2006). Most relevant to AAA disease pathophysiology, and its predilection for the distal-most aortic segment, is the marked difference between resting aortic wall shear stress in the thoracic and abdominal aorta. In suprarenal aortic segments, flow is antegrade throughout the cardiac cycle, providing continuous antegrade laminar wall shear stress. In infrarenal aorta, wall shear stress values are lower, and reverse flow is present in late systole and diastole. In response to reduced distal arterial resistance and increase flow, such as is demonstrated in the response to even modest lower extremity exercise, wall shear stress becomes antegrade and laminar throughout the cardiac cycle, mimicking those characteristic of more proximal aortic segments. These distinct regional differences in hemodynamic influences may account for some component of the differential aneurysm risk noted between the thoracic and abdominal aortic segments (Dua and Dalman, 2010).

	Normal Aorta	Aneurysm
Elastin		
Average	22.7	2.4
Maximum	32.5	6.7
Minimum	16.1	0.2
Muscle		
Average	22.6	2.2
Maximum	33.6	6.4
Minimum	15.5	0.4
Collagen and ground substances		
Average	54.8	95.5
Maximum	63	98
Minimum	48	91.4

Table 1. Composition of normal aorta and aneurysm. Extracted from Nichols & O'Rourke, 1990.

As commented above, there are several work in the literature that show a correlation between very low shear stresses and the loss of permeability of the endothelial cell membrane (Helmlinger et al., 1991; Chiu et al., 2003). Studies have been the basis of an alternative, or even complementary, mechanism responsible for the origin of these aneurysms. During the normal course of aging, the abdominal aortic artery gradually undergoes conformal changes in its geometry (increasing its length and diameter, thickening its wall, etc.). Over time, the relative unconstructed nature of this artery inside the abdominal cavity may lead to the formation of bends, kinks, and other morphological changes that, in turn, create "disturbed flow" conditions inside the vessel (i.e., unsteady flow separation and weak turbulence). It is argued that the anomalous response of the VEC to the high shear stresses, very low shear stresses, low, but oscillating shear stresses, and the anomalous temporal and spatial gradients of wall shear stress associated with these disturbed flow conditions could contribute to an unstable progressive degradation of the arterial wall and to the formation of the aneurysm (Lasheras, 2007). To understand more about the risk for and progression of aneurysm disease

is necessary introduce DFC to analyse hemodynamic in site specific, for example, in the aorta infrarenal. Compared with the suprarenal aorta, the infrarenal environment in resting subjects is characterized by increased peripheral resistance, increased oscillatory wall shear stress and stagnant flow (Dua and Dalman, 2010).

5. Computational fluid dynamics (CFD)

Since the beginning of the computer age, the computational study of fluid dynamic problems has been of interest to researchers studying both fundamental problems and engineering applications. Vast numbers of real-world problems require accurate viscous flow solutions to meet requirements for supporting engineering and science tasks – such as achieving ideal fluid dynamic performances and satisfying cost effectiveness. For example, computational analysis is indispensable, as well as economical, for developing advanced rocket-engine turbopumps and biomedical devices handling blood flow in humans. The computational fluid dynamics (CFD) for viscous, incompressible flow has been of interest for many decades to investigate fundamental fluid dynamic problems as well as engineering applications. The pioneering work by Harlow and Welch (1965) opened a new possibility of applying a computational approach to solving realistic incompressible fluid engineering problems, especially for three-dimensional problems (Kwak and Kiris, 2011).

Development of image-based modelling technologies for simulating blood flow began in the late 1990s. Since that time, many groups have developed and utilized these techniques to investigate the pathogenesis of occlusive and aneurysmal disease in the carotid artery (Long et al., 2000), the coronary arteries (Gijssen et al., 2007), the aorta (Tang et al., 2006) and the cerebral circulation (Cebal et al., 2005). Patient-specific modelling techniques have also been applied in solid mechanics analyses to predict rupture risk of aneurysms (Vorp, 2007).

There are many methods for quantifying vascular anatomy for patient-specific modelling of cardiovascular mechanics include noninvasive imaging techniques such as computed tomography (CT), magnetic resonance imaging (MRI), 3D ultrasound (3DUS) and an invasive method combining angiography and intravascular ultrasound (IVUS). Contrast-enhanced CT and MRI are particularly suited for generating high-resolution volumetric images of many parts of the vascular tree. Generally, iodinated contrast is used in CT angiography, and a gadolinium-based contrast agent is used in magnetic resonance (MR) angiography. MRI has the additional advantage of being able to quantify physiologic parameters, including blood flow, wall motion and blood oxygenation (Taylor & Figueroa, 2009).

A recent serial MRI-based case study demonstrated an association between the eventual site of plaque ulceration and elevated wall shear stress (WSS) (Groen et al., 2007) , a finding corroborated by an IVUS-based CFD study of 31 plaques that showed a clear association between elevated WSS and elevated strains within the plaque (Gijssen et al., 2008). Tang et al. (2006) quantified hemodynamic conditions in subject-specific models of the human abdominal aorta constructed from magnetic resonance angiograms (MRA) of five young, healthy subjects. Image-based modelling techniques are also being used to provide hemodynamic data in a clinical study testing the hypothesis that exercise can be employed to slow the progression of small (3-5 cm diameter) abdominal aortic aneurysms (Dalman et al., 2006). Rayz et al. (2008) performed a study on the effect of vertebral artery flow in four patients with basilar artery aneurysms. Computed flow fields were found to agree well with measurements made using PC-MRI. Subsequently, these authors reported good correspondence between regions of slow flow predicted by CFD at baseline and the deposition of thrombus observed at follow-up of three basilar aneurysms cases.

5.1 Real model of artery

One way to analyze and estimate the hemodynamic behaviour in an artery is the use of computational analysis of a real model of blood vessel (Legendre, 2010). The artery model is built based on images obtained from a patient. For example, in order to evaluate an abdominal aortic aneurysm in a patient, this model can be obtained by using multi-slice CT scan of the patient chest (Figure 3).

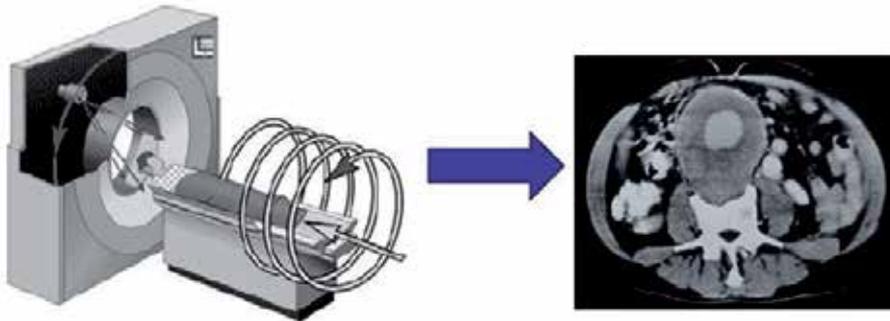


Fig. 3. CT images acquisition of a patient. Extracted from Invesalius (2008).

These DICOM (Digital Image Communication in Medicine) images should be treated by dedicated software, for example InVesalius, Mimics, etc., to select only the region of interest. Then, it is generated a compatible extension file that can be imported by a software to generate a three-dimensional mesh model of the arterial segment under study.

Overlapping the two dimensions (2D) images obtained through computed tomography or MRI, it can be created the computational model in three dimensions (3D) corresponding to the patient anatomical structure (Figure 4) and (Figure 5).

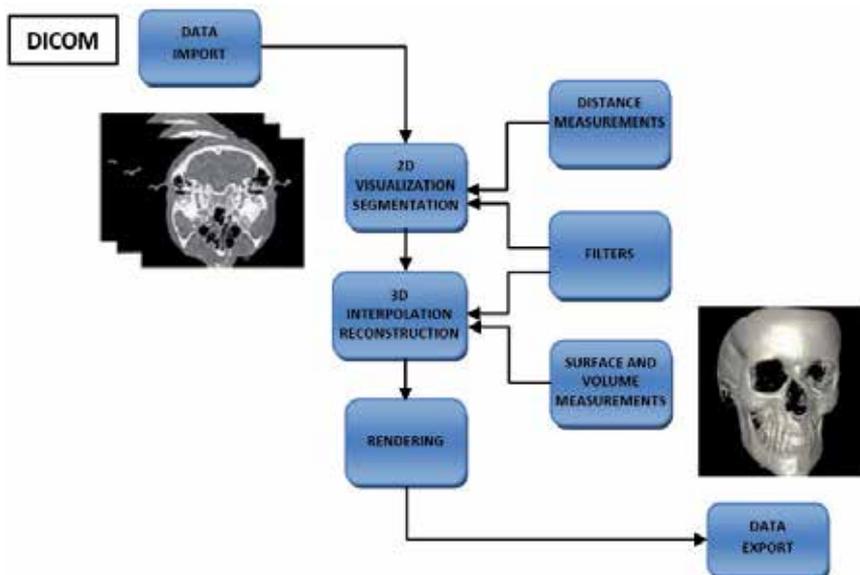


Fig. 4. Flowchart of the reconstruction of medical images. Extracted from Invesalius (2008).

The program makes, from tomographic medical imaging, a three-dimensional computational model in genuine size of several anatomical structures. This model can be manipulated and observed from different angles and it is also possible to separate specific parts of the image obtained from CT for a detailed analysis. The different densities of bone and tissue are easily identified by using tools such as color map. This feature is an important tool for creating computational models. Different methods of volume segmentation can be used for the treatment of the obtained images. Those methods are useful to adequately separate the regions of tissues, organs, anatomical structures etc.

The computational model generated by the software enables computer simulations in order to evaluate hemodynamic patterns developed by each patient according to their individual characteristics (Figure 6).

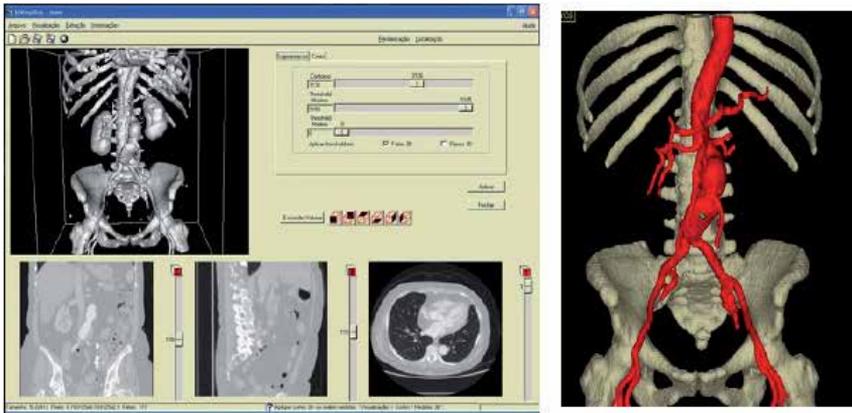


Fig. 5. Three-dimensional reconstruction of the abdominal aorta segment obtained from CT Scan. Extracted from Invesalius (2008).

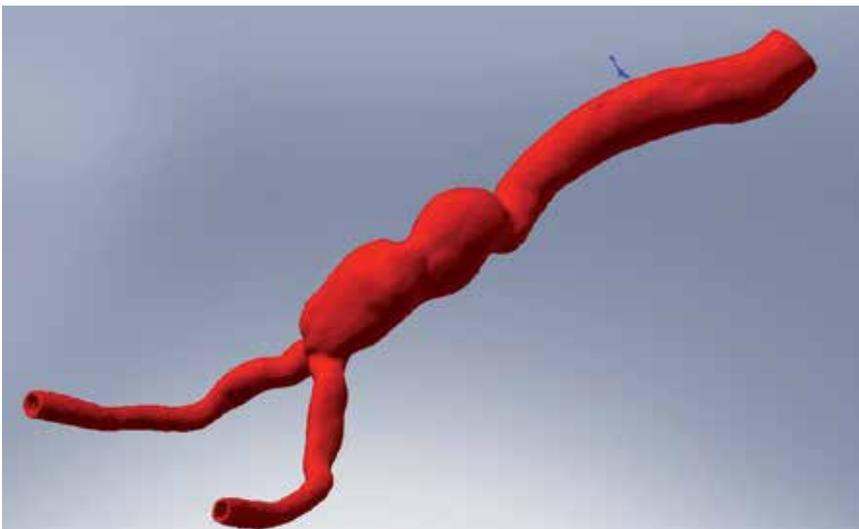


Fig. 6. Abdominal aortic aneurysm model obtained after image processing. Extracted from Legendre (2010).

5.2 Computer analysis

5.2.1 Finite volume method

The finite volume method (FVM) is a numerical technique for the solution of partial differential equations. The domain under study is subdivided into control volumes or computational cells. This method consists in the integration of the equations that governing the fluid flow over each control volume. The numerical solutions obtained by techniques of FVM (Patankar, 1975) have problems or errors known as false numerical diffusion. The numerical diffusion occurs when the interpolation function used in the discretization of equations differs from the exact solution. It can be understood as any effect that tends to moderate gradients or discontinuities present in the exact solution of a problem. There are some functions that were developed in order to minimize these effects of false diffusion, for example, the "upwind" method of second order.

For instance, the Fluent software solve the governing integral equations for the conservation of mass and momentum and also for scalars such as turbulence. The technique based on the volume control works as the following:

- The domain is divided in discrete control volumes using a computational mesh;
- Integration of the governing equations of the individual control volume to construct algebraic equations for discrete dependent variables (unknown), such as speed, pressure, etc.;
- Linearization of the discretized equations and solution of the resulting linear equation system to update the field values of the dependent variables.

5.2.2 Segregation solution method

In the segregated solution method, the governing equations are solved sequentially. Thus, to solve the equations that governing the phenomenon, many interactions are carried out until the solution convergence. Each interaction consists of the following steps outlined below:

- The fluid properties are updated based on the current solution. If the calculation has been started, the fluid properties are updated based on the initialized solution;
- The equations of momentum in x , y and z are solved one by one using the current values of pressure and mass flow to update the velocity field;
- Since the velocities obtained in step 2 cannot satisfy the conditions of continuity, pressure corrections are made so that continuity is satisfied;
- Other equations for scalar quantities such as turbulence are solved using the previously updated value of the other variables;
- It is made a convergence confirmation of the equations set.

These steps are performed until the convergence criteria are achieved.

5.2.3 Discretization

The technique based on control volume converts the governing equations into algebraic equations so that they can be solved numerically. This control volume technique consists of integrating the governing equations about each control volume, generating discrete equations that conserve each quantity on the volume control.

5.3 Overview

In general, the computational analysis tool can be viewed as a data interpolation in which the uncertainties of the results will depend on the model, the boundary conditions and mesh refinement adopted. To perform a study for hemodynamic performance evaluation, the

computational model adopted from the arterial segment should be able to represent the real model in a reliable way, as much as possible. The boundary conditions should be properly employed, since they are directly related to the consistency of the numerical results. The use of the finite volume method for solving the problem requires a proper discretization of the model in order to enable the convergence of analysis and also minimize the uncertainties in the results. On the other hand, it is also needed a good sense to perform the mesh refinement and computational cost in order to obtain a good quality of results in the shortest time. The quality of the mesh has an important role in the accuracy and stability of numerical computation. The convergence of results is intrinsically related to the size of mesh elements of the model and thus its quality is measured by its "asymmetry".

The choice of the optimal mesh for the simulation of the model should consider the performance and accuracy. The performance will depend on the total number of elements to be analyzed and thus a computational mesh made by larger elements will require lower computational cost due to lower number of elements required to cover the whole region of interest. However, this may affect the consistency of results. On the other hand, smaller mesh elements means better results, but this will require a greater computational cost (Figure 7).

One way to obtain more information regarding the arterial hemodynamic behavior is through the use of computer simulation. By using this kind of tool, it is possible to evaluate wall stress, recirculation areas, pressure, pathlines, velocity field and speeds distribution in the arterial segment.

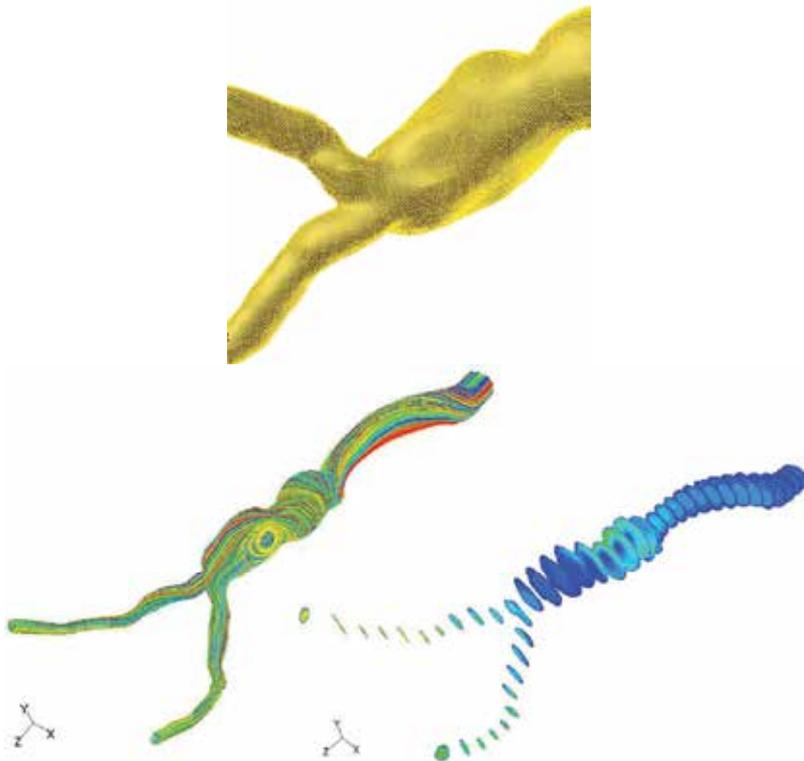


Fig. 7. Computational mesh, pathlines and velocity field of an arterial aneurysm. Extracted from Legendre (2010).

There are many works numerical and experimental fluid mechanics in the literature aimed at determining the characteristics of the flow shear stresses on the walls of AAAs at different stages of their development. These studies occur in ideal symmetric and nonsymmetric shapes of fusiform aneurysms and in realistic geometries obtained from high-resolution CT and angiographies (Egelhoff et al. 1999, Finol & Amon 2003, Salsac et al. 2006, Bessa & Ortiz 2009, Legendre 2010). Although, these studies show limitations about the appropriate initial and boundary conditions, as well as account for the precise elastic properties of the wall, they clearly show that once a fusiform aneurysm forms, the flow is dominated by the onset of an unsteady, massive separation from the walls that occurs immediately after the peak systole. When the flow separates from the walls during the deceleration portion of the cardiac cycle, a relatively coherent array of large vortices forms and the blood flow slowly recirculates (Lasheras, 2007). Bessa & Ortiz (2009) showed that occur massive separation from the walls and the velocity profile becomes inverted inside of the aneurysm, forming a vortex in the bottom and in the top of the aneurysm. After that, the velocity profile has been completely inverted and the vortex pair travels towards to the center of the aneurysm (Figure 8).

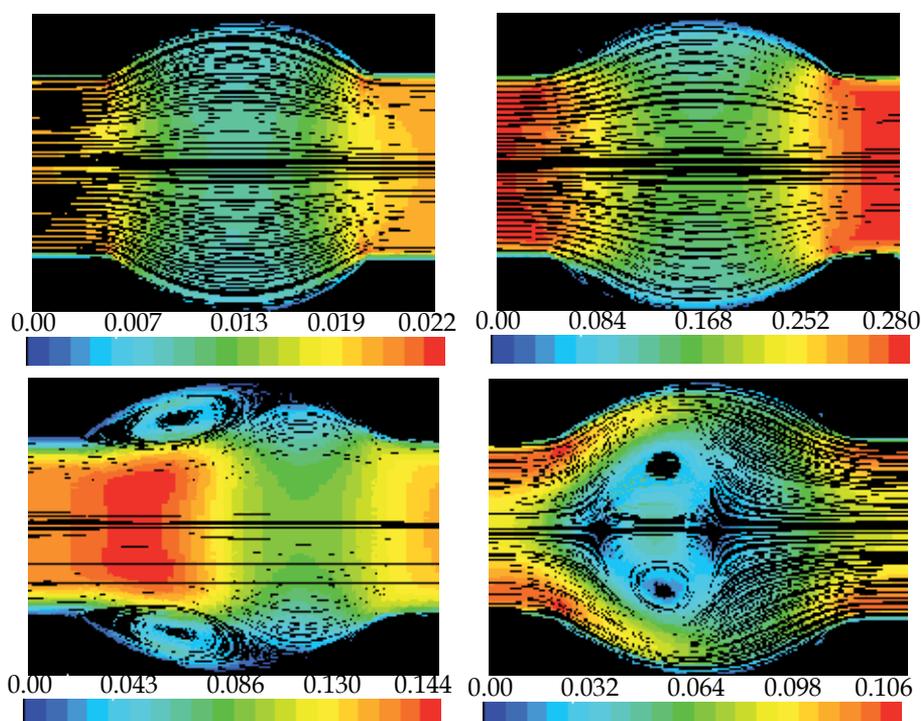


Fig. 8. Streamlines at various instant time of the cardiac cycle for aneurysm model. Extracted from Bessa & Ortiz (2009).

6. Conclusion

In the last few decades, it had a great effort of the scientific community to identify the beginning and the progression of the aneurysms. For this, several visualization techniques and 3D reconstruction of the vessels were developed associates with the computational fluid

dynamics with intention to analyze hemodynamic parameters that contributed to the development of the aneurysms. However, many questions exist on aneurysms that continue without answers: What cause aneurysms? When occur will the rupture of the aneurysms? Nowadays, advances in computational methods and computing hardware are also making it possible to solve increasingly more challenging problems, notably those involving flow instabilities and turbulence associated with a variety of vascular pathologies. However, to advance the research is the necessary to improve the related uncertainties the inserted boundary conditions in computational model. Exactly, even if these issues are fully resolved, still the mechanotransduction knowledge is essential to understand the endothelial cells behaviour. Therefore, still it has great challenges for the understanding of the origin and development of the aneurysm.

7. Acknowledgment

This chapter was supported, in part, by grants from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP).

8. References

- Acevedo-Bolton, G., Jou, L.D., Dispensa, B.P., Lawton, M.T., Higashida, R.T., Martin, A.J., Young, W.L. & Saloner, D. (2006). Estimating the Hemodynamic Impact of Intervantional Treatments of Aneurysms: Numerical Simulation with Experimental Validation: Technical Case Report. *Neurosurgery*, Vol. 59, pp. E429-E430.
- Ahn, S., Shin, D., Tateshima, S., Tanishita, K., Vinuela, F. & Sinha, S. (2007). Fluid-Induced Wall Shear Stress in Anthropomorphic Brain Aneurysm Models: Mr Phase-Contrast Study at 3 T. *Journal of Magnetic Resonance Imaging*, Vol. 25, pp. 1120-1130.
- Ailawadi, G., Eliason, J. & Upchurch Jr., G. (2003). Current Concepts in the Pathogenesis of Abdominal Aortic Aneurysm. *Journal of Vascular Surgery*, Vol. 38, 584-588.
- Bessa, K.L. & Ortiz, J.P. (2009). Flow Visualization in Arteriovenous Fistula and Aneurysm Using Computational Fluid Dynamics. *Journal of Visualization*, Vol. 12, pp. 95-107.
- Bessa, K.L., Belletati, J.F., dos Santos, L., Rossoni, L.V. & Ortiz, J.P. (2011). Drag Reduction by Polyethylene Glycol in the Tail Arterial Bed of Normotensive and Hypertensive Rats. *Brazilian Journal of Medical and Biological Research*, Vol. 44, pp. 767-777.
- Bonetti, P.O., Lerman, L.O. & Lerman, A. (2003). Endothelial Dysfunction: A Marker of Atherosclerotic Risk. *Arteriosclerosis, Thrombosis and Vascular Biology*, Vol. 23, pp. 168-175.
- Boussel, L, Rays, V., McCulloch, C., Martin, A., Acevedo-Bolton, G., Lawton, M., Higashida, R., Smith, W.S., Young, W.L. & Saloner, D. (2008). Aneurysm Growth Occurs at Region of Low Wall Shear Stress: Patient-Specific Correlation of Hemodynamics and Growth in a Longitudinal Study. *Stroke*, Vol. 39, pp. 2997-3002.
- Cebral, J.R., Castro, M.A., Burgess, J.E., Pergolizzi, R.S., Sheridan, M.J. & Putman, C.M. (2005). Characterization of Cerebral Aneurysms for Assessing Risk of Rupture by Using Patient-Specific Computational Hemodynamics Models. *American Journal of Neuroradiology*, Vol. 26, pp. 2550-2559.

- Chervu, A., Clagett, G.P., Valentine, R.J., Myers, S.I. & Rossi, P.J. (1995). Role of Physical Examination in Detection of Abdominal Aortic Aneurysms. *Surgery*, Vol. 117, pp. 454-457.
- Chiu, J.J. & Chien, S. (2011). Effects of Disturbed Flow on Vascular Endothelium: Pathophysiological Basis and Clinical Perspectives. *Physiological Reviews*, Vol. 91, pp. 327-387.
- Chiu, J.J., Chen, L.J., Lee, P.L., Lee, C.I., Lo, L.W., et al. (2003). Shear Stress Inhibits Adhesion Molecule Expression in Vascular Endothelial Cells Induced by Coculture with Smooth Muscle Cells. *Blood*, Vol. 1, pp. 2667-2674.
- Crompton, M.R. (1966). The Pathogenesis of Cerebral Aneurysms. *Brain: a journal of neurology*, Vol. 4, (December), pp. 797-814.
- Dalman, R.L., Tedesco, M.M., Myers, J. & Taylor, C.A. (2006). AAA Disease: Mechanisms, Stratification, and Treatment. *Annals of the New York Academy of Sciences*, Vol. 1085, pp. 92-109.
- Davies, P.F. (2009). Hemodynamic Shear Stress and the Endothelium in Cardiovascular Pathophysiology. Natural Clinical Practice. *Cardiovascular Medicine*, Vol. 6, pp.16-26.
- Davies, P.F., Dewey, C.F., Bussolari, S., Gordon, E. & Gibrone, M.A. (1984). Influence of Hemodynamics Forces on Vascular Endothelial Function. *The Journal of Clinical Investigation*, Vol. 4 (April), pp. 1121-1129.
- Davies, P.F., Mundel, T. & Barbee, K.A. (1995). A Mechanism for Heterogeneous Endothelial Responses to Flow in vivo and in vitro. *Journal of Biomechanics*, Vol. 28, (December), pp. 1553-1560.
- Dua, M.M. & Dalman, R.L. (2010). Hemodynamic Influences on Abdominal Aortic Aneurysm Disease: Application of Biomechanics to Aneurysm Pathophysiology. *Vascular Pharmacology*, Vol. 53, pp. 11-21.
- Egelhoff, C.J., Budwig, R.S., Elringer, D.F., Khraishi, T.A. & Johansen, K.H. (1999). Model Studies of the Flow in Abdominal Aortic Aneurysms During Resting and Exercise Conditions. *Journal of Biomechanics*, Vol. 32, pp. 1319-1329.
- Eppinger, H. (1887). Pathogenesis (Histogenesis und Aetiologie) der Aneurysmen einschliesslich des Aneurysma equi verminosum. Pathologisch-anatomische studien. *Arch Klin Chir*, Vol 35, pp. 1-563.
- Fillinger, M.F., Racusin, J., Baker, R.K., Cronenwett, J.L., Teutelink, A., Schermerhorn, M.L., Zwolar, R.M., Powell, R.J., Walsh, D.B., Rzucidlo, E.M. (2004). Anatomic Characteristics of Ruptured Abdominal Aortic Aneurysm on Conventional CT Scans: Implications for Rupture Risk. *Journal of Vascular Surgery*, Vol. 39, pp. 1243-1252.
- Finol, E.A. & Amon, C.H. (2003). Flow Dynamics in Anatomical Models of Abdominal Aortic Aneurysms: Computational Analysis of Pulsatile Flow. *Acta Científica Venezolana*, Vol. 54, pp. 43-49.
- Forbus, W.D. (1930). On the Origin of Miliary Aneurysm of the Superficial Cerebral Arteries. *Bull Johns Hopkins Hospital*, vol 47, pp 239-284.
- Fry, D.L. (1968). Acute Vascular Endothelial Changes Associated with Increased Blood Velocity Gradients. *Circulation Research*, Vol. 22, pp. 165-197.
- Gijssen, F.J., Wentzel, J.J., Thury, A., Lamers, B. & Schuurbijs, J.C. (2007). A New Imaging Technique to Study 3-D Plaque and Shear Stress Distribution in Human Coronary Artery Bifurcations in ViVo. *Journal of Biomechanics*, Vol. 40, pp. 2349-2357.
- Gijssen, F.J., Wentzel, J.J., Thury, A., Mastik, F., Shaar, J.A., Schuurbijs, J.C., Slager, C.J., van der Giessen, W.J. & de Feyter, P.J. (2008). Strain Distribution Over Plaques in

- Human Coronary Arteries Relates to Shear Stress. *American Journal of Physiology. Heart and Circulatory Physiology*, Vol. 295, pp. H1608-H1614.
- Gimbrone, M.A. Jr., Topper, J.N., Nagel, T., Anderson, K.R. & Garcia-Cardena, G. (2000). Endothelial Dysfunction, Hemodynamic Forces, and Atherogenesis. *Annals of the New York Academy of Sciences*, Vol. 902, pp. 230-240.
- Greve, J.M., Les, A.S., Tang, B.T., Draney Blomme, M.T., Wilson, N.M., Dalman, R.L., Pelc, N.J. & Taylor, C.A. (2006). Allometric Scaling of Wall Shear Stress from Mice to Humans: Quantification using Cine Phase-Contrast MRI and Computational Fluid Dynamics. *American Journal of Physiology. Heart and Circulatory Physiology*, Vol. 291, pp. H1700-H1708.
- Groen, H. C., Gijzen, F.J., van der Lugt, A., Ferguson, M.S., Hatsukami, T.S., van der Steen, A.F., Yuan, C. & Wentzel, J.J. (2007). Plaque Rupture in the Carotid Artery is Localized at the High Shear Stress Region: a Case Report. *Stroke*, Vol. 38, pp. 2379-2381.
- Harlow, F.H. & Welch, J.E. (1965). Numerical Calculation of Time-Dependent Viscous Incompressible Flow with Free Surface. *The Physics of Fluids*, Vol. 8, pp. 2182-2189.
- Helmlinger, G., Geiger, R.V., Schrech, S. & Nerem, R.M. (1991). Effects of Pulsatile Flow on Cultured Vascular Endothelial Cell Morphology. *Journal of Biomechanical England*, Vol. 113, pp. 123-131.
- Hisai, T.K. (2008). Mechanosignal Transduction Coupling between Endothelial and Smooth Muscle Cells: Role of Hemodynamics Forces. *American Journal of Physiology Cell Physiology*, 294, C659-661.
- Humphrey, J.D. & Taylor, C.A. (2008). Intracranial and Abdominal Aortic Aneurysms: Similarities, Differences, and Need for a New Class of Computational Models. *Annual Review of Biomedical Engineering*, Vol. 10, pp. 221-246.
- Inci, S. & Spetzler, R.F. (2000). Intracranial Aneurysms and Arterial Hypertension: A Review and Hypothesis. *Surgical Neurology*, Vol. 53, pp. 540-542.
- Invesalius (2008). Invesalius Software: Basic Course. Cenpra, Campinas, São Paulo, Brazil.
- Jou, L.D., Quick, C.M., Young, W.L., Lawton, M.T., Higashida, R., Martin, A. & Saloner, D. (2003). Computational Approach to Quantifying Hemodynamic Forces in Giant Cerebral Aneurysms. *AJNR American Journal of Neuroradiology*, Vol. 24, pp. 1804-1810.
- Kaiser, D., Freyberg, M.A. & Friedl, P. (1997). Lack of Hemodynamic Forces Triggers Apoptosis in Vascular Endothelial Cells. *Biochemical and Biophysical Research Communications*, Vol. 24, pp. 586-590.
- Kleinstreuer, C., Li, Z. & Farber, M.A. (2007). Fluid-Structure Interaction Analyses of Stented Abdominal Aortic Aneurysms. *Annual Review of Biomedical Engineering*, Vol. 9, pp. 169-204.
- Kwak, D. & Kiris, C.C. (2011). *Computation of Viscous Incompressible Flows*, Springer, ISBN 978-94-007-0192-2, New York, USA.
- Lakatta, E.G., Mitchell, J.H., Pomerance, A. & Rowe, G. (1987). Human Aging: Changes in Structure and Functions. *Journal of the American College of Cardiology*, Vol. 10, pp. 42A-47A.
- Lasheras, J.C. (2007). The Biomechanics of Arterial Aneurysms. *Annual Review of Fluid Mechanics*, Vol. 39, pp. 293-319.
- Lederle, F.A. & Simel, D.L. (1999). The Rational Clinical Examination. Does this Patient have Abdominal Aortic Aneurysm? *The Journal of the American Medical Association*, Vol. 281, pp. 77-82.

- Legendre, D. (2010). Study of Flow Behavior Through Physical and Computational Infrarenal Aortic Aneurysm Model Obtained by CT Scan. Academic Thesis, University of São Paulo, São Paulo, Brazil.
- Lerman, A. & Burnett, J.C. Jr. (1992). Intact and Altered Endothelium in Regulation of Vasomotion. *Circulation*, Vol. 86, pp. III12 - III9.
- Li, Y.S., Haga, J.H. & Chien, S. (2005). Molecular Basis of the Effects of Shear Stress on Vascular Endothelial cells. *Journal of Biomechanics*, Vol. 38, pp. 1949-1971.
- Long, Q., Xu, X.Y., Ariff, G., Thom, S.A., Hughes, A.D. & Stanton, A.V. (2000). Reconstruction of Blood Flow Patterns in a Human Carotid Bifurcation a Combined CFD and MRI Study. *Journal of Magnetic and Resonance Imaging*, Vol. 11, pp. 299-311.
- MacSweeney, S.T.R., Powell, J.T. & Greenhalgh, R.M. (1994). Pathogenesis of Abdominal Aortic Aneurysm. *The British Journal of Surgery*, Vol. 81, pp. 935-941.
- Malek, A.M., Alper, S.L. & Izumo, S. (1999). Hemodynamic shear stress and its role in atherosclerosis. *The Journal of the American Medical Association*, Vol. 282, pp. 2035-2042.
- Masuda, H., Zhuang, Y.J., Singh, T.M., Kawamura, K., Murakami, M., Zarins, C.K. & Glagov, S. (1999). Adaptive Remodeling of Internal Elastic Lamina and Endothelial Lining During Flow-Induced Arterial Enlargement. *Arteriosclerosis, Thrombosis and Vascular Biology*, Vol. 19, pp. 2298-2307.
- Meng, H., Wang, Z., Hoi, Y., Gao, L., Metaxa, E., Swartz, D.D. & Kolega, J. (2007). Complex Hemodynamics at the Apex of an Arterial Bifurcation Induces Vascular Remodeling Resembling Cerebral Aneurysm Initiation. *Stroke*, vol. 38, pp. 1924-1931.
- Mesh, C.L., Baxter, B.T., Pearce, W.H., Chisholm, R.L., McGee, G.S. & Yao, J.S. (1992). Collagen and Elastin Gene Expression in Aortic Aneurysms. *Surgery*, Vol. 112, pp. 256-261.
- Nichols, W.W. & O'Rourke, M.F. (1990). *McDonald's Blood Flow in Arteries: Theoretic, Experimental and Clinical Principles*. London: Edward Arnold.
- Nichols, W.W., O'Rourke, M.F. (1990). *Properties of the Arterial Wall*. In McDonald's Blood Flow in Arteries. London: Lea & Febiger, pp. 81-85.
- Paszowski, J.J. & Dardik, A. (2003). Arterial wall shear stress: Observations from the bench to the bedside. *Vascular and Endovascular Surgery*, Vol. 37, pp. 47-57.
- Patankar, S.V. (1975). *Numerical Prediction of Three-Dimensional Flows, in Studies in Convection: Theory Measurement and Applications*. Academic Press, London, Vol. 1.
- Raaymakers, T. (1999). Aneurysms in Relatives of Patients with Subarachnoid Hemorrhage: Frequency and Risk Factors. MARS Study Group. Magnetic Resonance Angiography in Relatives of Patients with Subarachnoid Hemorrhage. *Neurology*, Vol. 53, pp. 982-988.
- Rayz, V.L., Bousset, L., Acevedo-Bolton, G., Martin, A.J., Young, W.L., Lawton, M.T., Higashida, R. & Saloner, D. (2008). Numerical Simulations of Flow in Cerebral Aneurysms: Comparison of CFD Results and in vivo MRI Measurements. *Journal of Biomechanical Engineering*, Vol. 130, pp. 051011.
- Reed, D., Reed, C., Stemmermann, G. & Hayashi, T. (1992). Are Aortic Aneurysms Caused by Atherosclerosis? *Circulation*, Vol. 85, pp. 205-21.
- Ronkainen, A., Hernesniemi, J., Puranen, M., Niemitukia, L., Vanninen, R., Ryynanen, M., Kuivaniemi, H. & Tromp, G. (1997). Familial Intracranial Aneurysms. *Lancet*, Vol. 349, pp. 380-384.
- Salsac, A.V., Sparks, S.R., Chomaz, J.M. & Lasheras, J.C. (2006). Evolution of the Wall Shear Stresses During the Progressive Enlargement of Symmetric Abdominal Aortic Aneurysms. *Journal of Fluid Mechanics*, Vol. 560, pp. 19-51.

- Shah, P.K. (1997). Inflammation, Metalloproteinases, and Increase Proteolysis: an Emerging Paradigm in Aortic Aneurysm. *Circulation*, Vol. 96, pp. 2115-2117.
- Shojima, M., Oshima, M., Takagi, K., Torii, R., Hayakawa, M., Katada, K., Morita, A. & Kirino, T. (2004). Magnitude and Role of Wall Shear Stress on Cerebral Aneurysm: Computational Fluid Dynamic Study of 20 Middle Cerebral Artery Aneurysms. *Stroke*, Vol. 35, 2500-2505.
- Shojima, M., Oshima, M., Takagi, K., Torii, R., Nagata, K., Shirouzu, I., Morita, A. & Kirino, T. (2005). Role of the Bloodstream Impacting Force and the Local Pressure Elevation in the Rupture of Cerebral Aneurysms. *Stroke*, Vol. 36, pp. 1933-1938.
- Singh, K., Bonaa, K.H., Jacobsen, B.K., Bjork, L. & Solberg, S. (2001). Prevalence of and Risk Factors for Abdominal Aortic Aneurysms in a Population-Based Study. *American Journal of Epidemiology*, Vol. 154, pp. 236-244.
- Tang, B.T., Cheng, C.P., Draney, M.T., Wilson, N.M., Tsao, P.S., Herfkens, R.J. & Taylor, C.A. (2006). Abdominal Aortic Hemodynamics in Young Healthy Adults at Rest and During Lower Limb Exercise: Quantification Using Image-Based Computer Modeling. *American Journal of Physiology. Heart and Circulatory Physiology*, Vol. 291, pp. H668-H676.
- Taylor, C.A. & Figueroa, C.A. (2009). Patient-Specific Modeling of Cardiovascular Mechanics. *The Annual Review of Biomedical Engineering*, Vol. 11, pp. 109-134.
- Thiriet, M. (2008). *Biology and Mechanics of Blood Flows: Part II – Mechanics and Medical Aspects*, Springer, ISBN 978-0-387-74848-1, New York, USA.
- Thubrikar, M. J. (2007). *Vascular Mechanics and Pathology*, Springer, ISBN-10: 0-387-33816-0, New York, USA.
- Toth, M., Nadasy, G.L., Nyary, I., Kerenyi, T., Orosz, M., Molnárka, G. & Monos, E. (1998). Sterically Inhomogeneous Viscoelastic Behavior of Human Saccular Cerebral Aneurysms. *Journal of Vascular Research*, vol. 35, pp. 345-355.
- Utter, G. & Rossman, J.S. (2007). Numerical Simulation of Saccular Aneurysm Hemodynamics: Influence of Morphology on Rupture Risk. *Journal of Biomechanics*, Vol. 40, 2716-2722.
- Vorp DA, Raghavan ML, Webster MW.(1998) Mechanical wall stress in abdominal aortic aneurysm: influence of diameter and asymmetry. *J Vasc Surg*;27:632-639.
- Vorp, D.A. (2007). Biomechanics of Abdominal Aortic Aneurysm. *Journal of Biomechanics*, Vol. 40, pp. 1887-1902.
- Wernig, F. & Xu, Q. (2002). Mechanical Stress-Induced Apoptosis in the Cardiovascular System. *Progress in Biophysics & Molecular Biology*, Vol. 78, Issues 2-3, (February-April), pp. 105-137.
- Zamir, M. (2005). *The Physics of Coronary Blood Flow*, Springer, ISBN 978-0387-25297-1, New York, USA.
- Zhang, B., Fugleholm, K., Day, L.B., Ye, S., Weller, R.O. & Day, I.N. (2003). Molecular Pathogenesis of Subarachnoid Haemorrhage. *The International Journal of Biochemistry & Cell Biology*, Vol. 35, pp. 1341-1360.

Assessment of Carotid Flow Using Magnetic Resonance Imaging and Computational Fluid Dynamics

Vinicius C. Rispoli¹, Joao L. A. Carvalho¹, Jon F. Nielsen²
and Krishna S. Nayak³

¹*Universidade de Brasilia*

²*University of Michigan*

³*University of Southern California*

¹*Brazil*

^{2,3}*USA*

1. Introduction

Knowledge of blood flow patterns in the human body is a critical component in cardiovascular disease research and diagnosis. Carotid atherosclerosis, for example, refers to the narrowing of the carotid arteries. One symptom of atherosclerosis is abnormal flow. The carotid arteries supply blood to the brain, so early detection of carotid stenosis may prevent thrombotic stroke. The current clinical gold standard for cardiovascular flow measurement is Doppler ultrasound. However, evaluation by ultrasound is inadequate when there is fat, air, bone, or surgical scar in the acoustic path, and flow measurement is inaccurate when the ultrasound beam cannot be properly aligned with the axis of flow (Hoskins, 1996; Winkler & Wu, 1995). Two alternative approaches to 3D flow assessment are currently available to the researcher and clinician: (i) direct, model-independent velocity mapping using flow-encoded magnetic resonance imaging (MRI), and (ii) model-based computational fluid dynamics (CFD) calculations.

MRI is potentially the most appropriate technique for addressing all aspects of cardiovascular disease examination. MRI overcomes the acoustical window limitations of ultrasound, potentially allowing flow measurements to be obtained along any direction, and for any vessel in the cardiovascular system. MRI measurements are also less operator-dependent than those of Doppler ultrasound, and the true direction of flow can generally be precisely measured. The current gold standard for MRI flow quantification is phase contrast (PC) (O'Donnell, 1985). However, PC suffers from partial-volume effects when a wide distribution of velocities is contained within a single voxel (Tang et al., 1993). This is particularly problematic when flow is turbulent and/or complex (e.g., flow jets due to stenosis) or at the interface between blood and vessel wall (viscous sublayer). This issue is typically addressed by increasing the spatial resolution, which dramatically affects the signal-to-noise ratio (SNR) and increases the scan time. Therefore, PC may be inadequate for estimating the peak velocity of stenotic flow jets and for assessing wall shear rate.

An alternative approach for cardiovascular flow assessment is to reconstruct a complex flow field via CFD simulation, using vascular geometries and input/output functions derived from non-invasive imaging data (Kim et al., 2006; Papathanasopoulou et al., 2003; Steinman et al., 2002). The accuracy of conventional CFD routines hinges on many modeling assumptions that are not strictly true for *in vivo* vascular flow, including rigid vessel walls and uniform blood viscosity. Furthermore, conventional CFD routines typically employ a structured non-Cartesian finite-element mesh that conforms to the vessel geometry, which leads to relatively complex algorithms that incur significant computational costs.

This chapter proposes two new approaches for MRI assessment of carotid flow.

First, we introduce a rapid MRI method for fully-localized measurement of cardiovascular blood flow. The proposed method consists of combining slice-selective spiral imaging with Fourier velocity-encoded (FVE) imaging. The “spiral FVE” method provides intrinsically higher SNR than PC. Furthermore, it is not affected by partial-volume effects, as it measures the velocity distribution within each voxel. This makes this method particularly useful for assessment of flow in stenotic vessels. We show that spiral FVE measurements of carotid flow show good agreement with Doppler ultrasound measurements. We also introduce a mathematical model for deriving spiral FVE velocity distributions from PC velocity maps, and we use this model to show that spiral FVE provides good quantitative agreement with PC measurements, in an experiment with a pulsatile carotid flow phantom. Then, we show that it is possible to accurately estimate *in vivo* wall shear rate (WSR) and oscillatory shear index (OSI) at the carotid bifurcation using spiral FVE.

Finally, we propose a hybrid MRI/CFD approach that integrates low-resolution PC flow measurements directly into the CFD solver. The feasibility of this MRI-driven CFD approach is demonstrated in the carotid bifurcation of a healthy volunteer. We show that MRI-driven CFD has a regularizing effect on the flow fields obtained with MRI alone, and produces flow patterns that are in better agreement with direct MRI measurements than CFD alone. This approach may provide improved estimates of clinically significant blood flow patterns and derived hemodynamic parameters, such as wall shear stress (WSS).

2. Magnetic resonance flow imaging

2.1 Basic principles of MRI

MRI is a modality uniquely capable of imaging all aspects of cardiovascular disease, and is a potential “one-stop shop” for cardiovascular health assessment. MRI can generate cross-sectional images in any plane (including oblique planes), and can also measure blood flow. The image acquisition is based on using strong magnetic fields and non-ionizing radiation in the radio frequency range, which are harmless to the patient.

The main component of a MRI scanner is a strong magnetic field, called the B_0 field. This magnetic field is always on, even when the scanner is not being used. Typically, MR is used to image hydrogen nuclei, because of its abundance in the human body. Spinning charged particles (or “spins”), such as hydrogen nuclei, act like a tiny bar magnet, presenting a very small magnetic field, emanating from the south pole to the north pole. In normal conditions, each nucleus points to a random direction, resulting in a null net magnetization. However, in the presence of an external magnetic field (such as the B_0 field), they will line up with that field. However, they will not all line up in the same direction. Approximately half will point north, and half will point south. Slightly more than half of these spins (about one in

a million) will point north, creating a small net magnetization M_0 , which is strong enough to be detected. The net magnetization is proportional to the strength of the B_0 field, so MRI scanners with stronger magnetic fields (e.g., 3 Tesla) provide higher SNR.

Other important components of the scanner are the gradient coils. There are typically three gradient coils (G_x , G_y , and G_z), that produce an intentional perturbation in the B_0 field when turned on (“played”). This perturbation varies linearly along each spatial direction (x , y and z), such that no perturbation exists at the iso-center of the magnet when these gradients are used. In the presence of an external magnetic field, the spins rotate about the axis of that field. B_0 is (approximately) spatially uniform, so all spins initially rotate at the same frequency (the Larmor frequency), $\omega = \gamma B_0$, where γ is the gyromagnetic ratio ($\gamma = 42.6$ MHz/Tesla for hydrogen protons). However, when any of the gradients is played, the magnetic field becomes spatially varying, and so does the rotation frequency of the spins. Therefore, G_x , G_y , and G_z are used to frequency-encode (or phase-encode) spatial position along the x , y and z directions, respectively.

The final major component of the MR scanner is the radio-frequency (RF) coil. This is used to transmit an RF “excitation” pulse to the body, and also to receive the frequency-encoded signal from the “excited” portion of the body. In practice, independent coils may be used for transmission and reception. The RF pulse is typically modulated to the Larmor frequency. While B_0 is aligned with the z -axis (by definition), B_1 , which is a very weak magnetic field associated with the RF pulse, is aligned with the x -axis (also by definition). When the RF pulse is played, some of the spins which are in resonance with the RF pulse (i.e., rotating at the RF pulse’s frequency) will now begin to rotate around the x -axis (thus the name magnetic resonance). This tilts the net magnetization towards the x - y plane, and the net magnetization will now have a component in the x - y plane, M_{xy} .

The RF pulse is typically designed to have a somewhat rectangular profile in Fourier domain, centered at the modulation frequency (e.g., a modulated windowed sinc). This implies that the RF pulse in fact contains a certain range of frequencies, thus all spins rotating within that range become “excited”, or tilted towards the x -axis. So, by playing gradient(s) of appropriate amplitude, and designing the RF pulse accordingly, one can excite only a thin slice of the body, which correspond to the region containing all spins that are in resonance with the RF pulse’s range of frequencies. Excitation profiles other than “slices” may also be obtained (e.g., a pencil beam, or cylindrical excitation (Hu et al., 1993)), by designing an appropriate gradient/pulse combination.

When the RF pulse is turned off, M_{xy} begins to rotate (at the Larmor frequency) around the z -axis, as the net magnetization begins to realign with B_0 . This rotating magnetization generates an oscillating signal, which can be detected by the receive coil. The frequency content of the received signal can be used to obtain spatial information about the excited portion of the body. In order to frequency-encode spatial information, gradients are also played during signal acquisition. These are called readout gradients. For imaging a slice perpendicular to the z -axis (an axial image), G_z is played during excitation (for slice-selection), and G_x and G_y are played during acquisition. These can be switched, for acquiring sagittal or coronal images, or all three gradients may be used during both excitation and acquisition to image oblique planes.

When the readout gradients are played, the acquired signal at a particular time instant corresponds to the sum of sinusoidal signals generated by spins located at different regions of the body, each rotating at different frequencies corresponding to their spatial locations. If

an axial slice is being acquired, for example, the demodulated signal value is equivalent to a sample of the Fourier transform $M(k_x, k_y)$ of the cross-sectional image $m(x, y)$. In this case, by changing the amplitudes of G_x and G_y during acquisition, one may acquire different samples of $M(k_x, k_y)$. In fact, by playing G_x and/or G_y , one can move along the k_x - k_y plane (which is known in MRI as “k-space”), collecting samples of $M(k_x, k_y)$. When enough samples of $M(k_x, k_y)$ have been collected, an inverse Fourier transform produces $m(x, y)$.

The required coverage of k-space, and the number of samples, depend on the specified spatial resolution and field-of-view. For low spatial resolution imaging, only the central portion of k_x - k_y needs to be sampled. For higher spatial resolution, the periphery of k-space must also be covered. The field of view is associated with the spacing between samples. For a larger field-of-view, k-space needs to be more densely sampled, requiring an increased number of samples. If k-space is not sufficiently sampled, and the resulting field-of-view is not large enough to cover the entire object, overlap in spatial domain (aliasing) is observed.

Because signal amplitude is lost as the net magnetization realigns with B_0 (this is called relaxation), multiple acquisitions (excitation + readout) may be needed in order to cover the entire k-space. Some trajectories are more efficient in covering k-space than others. For example, spiral imaging, which uses oscillating gradients to achieve spiral k-space trajectories (Figure 1b), are generally faster than 2DFT imaging, i.e., require fewer acquisitions. In 2DFT imaging, each acquisition readout acquires a single line of k-space, sampling k_x - k_y in a Cartesian fashion (Figure 1a). This is generally slower, but may be advantageous in some applications with respect to the nature of associated image artifacts. The full sequence of RF pulses and gradients is called a “pulse sequence”. The time between acquisitions is called the pulse repetition time, or TR.

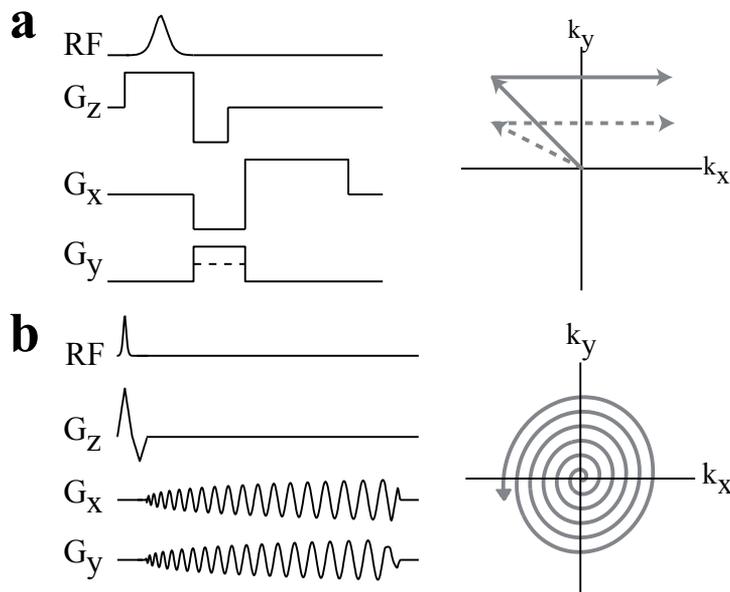


Fig. 1. Timing diagram (left) and corresponding k-space trajectories (right) for (a) 2DFT, and (b) spiral acquisitions.

2.2 Mathematical formalism

As discussed on section 2.1, the acquired MR signal $s(t)$ at a particular time instant corresponds to a sample of the Fourier transform $M(k_x, k_y)$ of the excited magnetization $m(x, y)$:

$$M(k_x, k_y) = \int_x \int_y m(x, y) e^{-j2\pi(k_x x + k_y y)} dx dy. \quad (1)$$

The Fourier coordinates k_x and k_y vary with time, according to the zeroth moment of the readout gradients G_x and G_y :

$$k_x(t) = \frac{\gamma}{2\pi} \int_0^t G_x(\tau) d\tau \quad (2)$$

$$k_y(t) = \frac{\gamma}{2\pi} \int_0^t G_y(\tau) d\tau. \quad (3)$$

These equations explain how the gradients can be used to “move” along k-space, as discussed in section 2.1. This formalism can be generalized for any combination of G_x , G_y , and G_z gradients:

$$M(\vec{k}_r) = \int_{\vec{r}} m(\vec{r}) \cdot e^{-j2\pi\vec{k}_r \cdot \vec{r}} d\vec{r} \quad (4)$$

$$\vec{k}_r(t) = \frac{\gamma}{2\pi} \int_0^t \vec{G}_r(\tau) d\tau, \quad (5)$$

where \vec{G}_r is the oblique gradient resulting from the combination of the G_x , G_y and G_z gradients, and \vec{r} is its corresponding axis along which the linear variation in magnetic field intensity is realized.

Given a spatial position function $\vec{r}(t)$ and a magnetic field gradient $\vec{G}_r(t)$, the magnetization phase is:

$$\phi(\vec{r}, t) = \gamma \int_0^t \vec{G}_r(\tau) \cdot \vec{r}(\tau) d\tau, \quad (6)$$

For static spins, $\vec{r}(t)$ is constant (\vec{r}), and this becomes:

$$\phi = \gamma \vec{r} \cdot \int_0^t \vec{G}_r(\tau) d\tau \quad (7)$$

$$= 2\pi \vec{k}_r \cdot \vec{r}, \quad (8)$$

as in the exponential in equation 4.

2.3 Principles of MR flow imaging

The basic principles of quantitative flow measurement using magnetic resonance were first proposed by Singer (1959) and Hahn (1960) in the late 1950s. However, clinical applications of MR flow quantification were not reported until the early 1980s (Moran et al., 1985; Nayler et al., 1986; Singer & Crooks, 1983; van Dijk, 1984). Current MR flow imaging methods are based on the fact that spins moving at a constant velocity accrue a phase proportional to the

velocity times the first moment of the gradient waveform along the direction in which they are moving.

For spins moving along the \vec{r} -axis with a constant velocity \vec{v} , and initial position \vec{r}_0 , we can write $\vec{r}(t) = \vec{r}_0 + \vec{v}t$. Rewriting equation 6, for $t = t_0$:

$$\phi = \gamma \int_0^{t_0} \vec{G}_r(t) \cdot (\vec{r}_0 + \vec{v}t) dt \quad (9)$$

$$= \gamma \vec{r}_0 \cdot \int_0^{t_0} \vec{G}_r(t) dt + \gamma \vec{v} \cdot \int_0^{t_0} \vec{G}_r(t) t dt \quad (10)$$

$$= \gamma \vec{r}_0 \cdot \vec{M}_0 + \gamma \vec{v} \cdot \vec{M}_1, \quad (11)$$

where \vec{M}_0 and \vec{M}_1 are the zeroth and first moments of the \vec{r} -gradient waveform at the time of signal acquisitions (“echo time”, or “time to echo” (TE)), respectively. Thus, if a gradient with null zeroth moment is used (e.g., a bipolar gradient, aligned with \vec{v}), the phase accrued for a constant velocity spin is $\phi = \gamma \vec{v} \cdot \vec{M}_1$.

Therefore, if a bipolar gradient waveform is played between the excitation and the readout, the phase measured in a pixel of the acquired image is directly proportional to the velocity of the spins contained within its corresponding voxel. However, factors other than flow (such as inhomogeneities of the magnetic field) may cause additional phase shifts that would cause erroneous interpretation of the local velocity (Rebergen et al., 1993).

2.4 Phase contrast

The phase contrast method addresses the problem mentioned above by using two gradient-echo data acquisitions in which the first moment of the bipolar gradient waveform is varied between measurements (O'Donnell, 1985). The velocity in each voxel is measured as:

$$v(x, y) = \frac{\phi_a(x, y) - \phi_b(x, y)}{\gamma(M_1^a - M_1^b)}, \quad (12)$$

where $\phi_a(x, y)$ and $\phi_b(x, y)$ are the phase images acquired in each acquisition, and M_1^a and M_1^b are the first moment of the bipolar gradients used in each acquisition.

2.5 Fourier velocity encoding

While phase contrast provides a single velocity measurement associated with each voxel, Fourier velocity encoding (Moran, 1982) provides a velocity histogram for each spatial location, which is a measurement of the velocity distribution within each voxel.

FVE involves phase-encoding along a velocity dimension. Instead of only two acquisitions, as in phase contrast, multiple acquisitions are performed, and a bipolar gradient with a different amplitude (and first moment) is used in each acquisition. Equation 10 can be rewritten as:

$$\phi(\vec{r}, \vec{v}, t) = 2\pi(\vec{k}_r \cdot \vec{r} + \vec{k}_v \cdot \vec{v}), \quad (13)$$

where \vec{k}_v is the velocity frequency variable associated with \vec{v} , and is proportional to the first moment of $\vec{G}_r(t)$:

$$\vec{k}_v = \frac{\gamma}{2\pi} \vec{M}_1. \quad (14)$$

Each voxel of the two-dimensional image is associated with a distribution of velocities. This three-dimensional function $m(x, y, v)$ is associated with a three-dimensional Fourier space $M(k_x, k_y, k_v)$. Thus, an extra dimension is added to k-space, and multiple acquisitions are required to cover the entire k_x - k_y - k_v space. In order to move along k_v , a bipolar gradient with the appropriate amplitude (and first moment) is played before the k_x - k_y readout gradients, in each acquisition. Placing the bipolar gradient along the z-axis will encode through-plane velocities. Placing the bipolar gradient along x or y will encode in-plane velocities. Oblique flow can be encoded using a combination of bipolar gradients along the x , y and z axes.

Each acquisition along k_v is called a velocity encode. The number of required velocity encodes depends on the desired velocity resolution and velocity field-of-view (the maximum range of velocities measured without aliasing). For example, to obtain a 25 cm/s resolution over a 600 cm/s field-of-view, 24 velocity encodes are needed. The spatial-velocity distribution, $m(x, y, v)$, is obtained by inverse Fourier transforming the acquired data, $M(k_x, k_y, k_v)$. If cine imaging (Glover & Pelc, 1988) is used, measurements are also time resolved, resulting in a four-dimensional dataset: $m(x, y, v, t)$.

2.6 FVE signal model

2DFT phase contrast provides two two-dimensional functions, $m(x, y)$ and $v_o(x, y)$, the magnitude and velocity maps, respectively. If these maps are measured with sufficiently high spatial resolution, and flow is laminar, one can assume that each voxel contains only one velocity, and therefore the spatial-velocity distribution associated with the object is approximately:

$$s(x, y, v) = m(x, y) \times \delta(v - v_o(x, y)), \quad (15)$$

where $\delta(v)$ is the Dirac delta function.

In 2DFT FVE, k-space data is truncated to a rectangular cuboid in k_x - k_y - k_v space. The associated object domain spatial-velocity blurring can be modeled as a convolution of the true object distribution, $s(x, y, v)$, with $\text{sinc}(x/\Delta x)$, $\text{sinc}(y/\Delta y)$, and $\text{sinc}(v/\Delta v)$, where Δx and Δy are the spatial resolutions along the x and y axes, respectively, and Δv is the velocity resolution, as follows:

$$\hat{s}(x, y, v) = [m(x, y) \times \delta(v - v_o(x, y))] * \text{sinc}\left(\frac{x}{\Delta x}\right) * \text{sinc}\left(\frac{y}{\Delta y}\right) * \text{sinc}\left(\frac{v}{\Delta v}\right), \quad (16)$$

where $\hat{s}(x, y, v)$ is the measured object distribution, and $*$ denotes convolution. This is equivalent to:

$$\hat{s}(x, y, v) = \left[m(x, y) \times \text{sinc}\left(\frac{v - v_o(x, y)}{\Delta v}\right) \right] * \left[\text{sinc}\left(\frac{x}{\Delta x}\right) \times \text{sinc}\left(\frac{y}{\Delta y}\right) \right]. \quad (17)$$

This approach for deriving FVE data from high-resolution velocity maps can be used for many simulation purposes (Carvalho et al., 2010).

3. Slice-selective FVE with spiral readouts (spiral FVE)

Phase contrast imaging is fast, but has limitations associated with partial-volume effects. On the other hand, 2DFT FVE addresses those limitations, but requires long scan times. Thus, we propose the use of slice-selective FVE MRI with spiral acquisitions. The proposed spiral FVE method is capable of acquiring fully localized, time-resolved velocity distributions in a short breath-hold. Scan-plane prescription is performed using classic protocols.

We present practical implementations for measuring blood flow through the carotid arteries, and comparisons with Doppler ultrasound and high-resolution 2DFT phase contrast MRI. The proposed method is demonstrated in healthy volunteers. Subjects provided informed consent, and were imaged using a protocol approved by the institutional review board of the University of Southern California.

3.1 Pulse sequence

The spiral FVE imaging pulse sequence (Figure 2) consists of a slice-selective excitation, a velocity-encoding bipolar gradient, a spiral readout, and refocusing and spoiling gradients. The dataset corresponding to each temporal frame is a stack-of-spirals in k_x - k_y - k_v space (Figure 3). The bipolar gradient effectively phase-encodes in k_v , while each spiral readout acquires one “disc” in k_x - k_y .

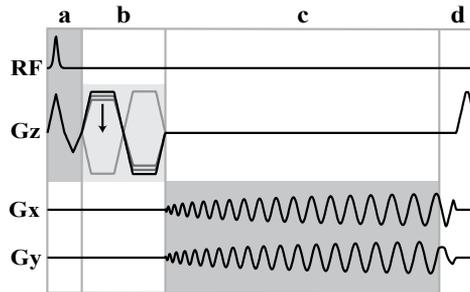


Fig. 2. Spiral FVE pulse sequence. It consists of (a) slice selective excitation, (b) velocity encoding bipolar gradient, (c) spiral readout, and (d) refocusing and spoiling gradients.

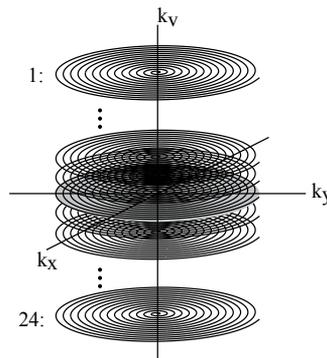


Fig. 3. Spiral FVE k -space sampling scheme. The dataset corresponding to each temporal frame is a stack-of-spirals in k_x - k_y - k_v space. Each spiral acquisition corresponds to a different k_v encode level.

3.2 Spiral FVE signal model

As spiral FVE acquisitions follow a stack-of-spirals pattern in k_x - k_y - k_v space (Figure 3), k -space data is truncated to a cylinder, i.e., a circle along k_x - k_y (with diameter $1/\Delta r$), and a rect function along k_v (with width $1/\Delta v$), where Δr and Δv are the prescribed spatial and velocity resolutions, respectively. Using the same approach we used in section 2.6 for 2DFT FVE, the associated object domain spatial-velocity blurring in spiral FVE can be modeled as a convolution of the true object distribution, $s(x, y, v)$, with $\text{jinc}(\sqrt{x^2 + y^2}/\Delta r)$ and $\text{sinc}(v/\Delta v)$, resulting in:

$$\begin{aligned} \hat{s}(x, y, v) &= [m(x, y) \times \delta(v - v_o(x, y))] * \text{sinc}\left(\frac{v}{\Delta v}\right) * \text{jinc}\left(\frac{\sqrt{x^2 + y^2}}{\Delta r}\right) \\ &= \left[m(x, y) \times \text{sinc}\left(\frac{v - v_o(x, y)}{\Delta v}\right) \right] * \text{jinc}\left(\frac{\sqrt{x^2 + y^2}}{\Delta r}\right). \end{aligned} \quad (18)$$

3.3 In vitro validation

An *in vitro* comparison of velocity distributions measured with spiral FVE with those derived from high-resolution 2DFT phase contrast — the current MR gold standard — was performed. The signal model presented in section 3.2 was used to generate simulated FVE data based on high-resolution 2DFT phase contrast data.

The validation experiments were performed using a pulsatile carotid flow phantom (Phantoms by Design, Inc., Bothell, WA). A slice perpendicular to the phantom's carotid bifurcation was prescribed, and through-plane velocities were measured. A cine gradient-echo 2DFT phase contrast sequence with high spatial resolution and high SNR (0.33 mm resolution, 10 averages, 80 cm/s Venc) was used as a reference. Cine spiral FVE data with $\Delta r = 3$ mm and $\Delta v = 10$ cm/s was obtained from the same scan plane. Both acquisitions were prospectively gated, and used the same TR (11.6 ms), flip angle (30°), slice profile (3 mm), temporal resolution (23.2 ms), and pre-scan settings. The total scan time was 40 minutes for phase contrast, and 12 seconds for FVE.

A simulated spiral FVE dataset was computed from the PC magnitude and velocity maps, using the convolution model described in Eq. 18. The PC-derived and FVE-measured data were registered by taking one magnitude image $m(x, y)$ from each dataset, and then using the phase difference between their Fourier transforms $M(k_x, k_y)$ to estimate the spatial shift between the images. Amplitude scaling was performed by normalizing the ℓ_2 -norm of each FVE dataset. The difference between PC-derived and FVE-measured time-velocity distributions was calculated for select voxels, and the associated signal-to-error ratios were computed. This was used as a quantitative assessment of spiral FVE's accuracy.

Figure 4 shows measured and PC-derived time-velocity FVE distributions from nine representative voxels, selected around the circumference of the vessel wall of the pulsatile carotid flow phantom's bifurcation. The signal-to-error ratio between measured and PC-derived time-velocity distributions was measured to be within 9.3–11.7 dB. Imperfect registration between the datasets, combined with spatial blurring due to off-resonance in the measured spiral FVE data, may have contributed to this moderate signal-to-error ratio. Nevertheless, the two datasets show good visual agreement, and no significant spatial variation was observed in terms of accuracy. These results show that velocity distributions

measured with spiral FVE agree well with those obtained with 2DFT phase contrast, the current MRI gold standard.

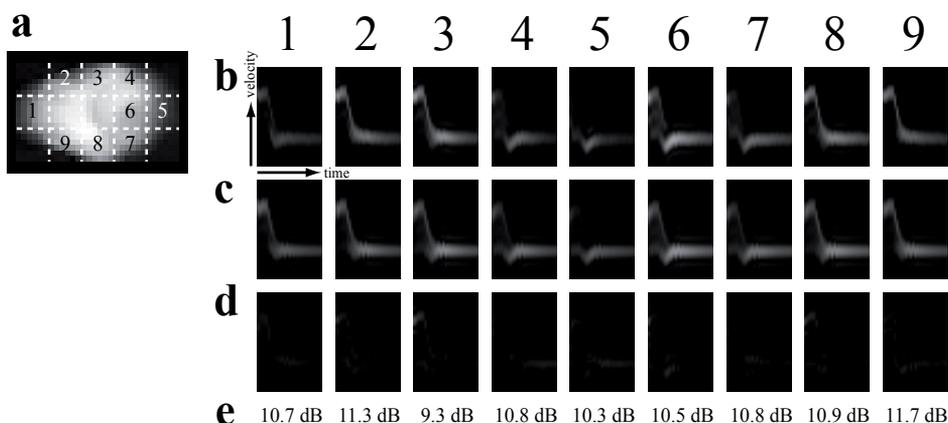


Fig. 4. *In vitro* evaluation of the accuracy of spiral FVE velocity histograms. Results are shown for nine representative voxels, selected around the circumference of the vessel wall of the pulsatile carotid flow phantom's bifurcation (a). For each voxel, it is shown: (b) time-velocity distribution derived from high-resolution 2DFT phase contrast; (c) time-velocity distribution measured with spiral FVE; (d) absolute difference between spiral FVE and 2DFT PC-derived histograms; (e) signal-to-error ratio.

3.4 *In vivo* evaluation

The spiral FVE method was evaluated *in vivo*, aiming at quantifying flow through the common carotid artery of a healthy volunteer. Doppler ultrasound was used as a gold standard and was qualitatively compared with the proposed method. We also show experiments for evaluating the most appropriate view-ordering scheme for measuring carotid flow, and we demonstrate spiral FVE's ability to measure multiple flows in a single acquisition.

3.4.1 View ordering

Experiments were performed to determine the appropriate view-ordering scheme for carotid spiral FVE studies. Flow was measured through the carotid artery of a healthy volunteer, using a 4-interleaf spiral FVE acquisition. The measurement was performed twice, and in each acquisition, a different view-ordering scheme was used. In the first acquisition, one spiral interleaf was acquired per heartbeat, and two different k_v levels were encoded throughout each R-R interval. In the second acquisition, two spiral interleaves were acquired per heartbeat, encoding one k_v level per cardiac cycle. Reconstructed velocity histograms were compared with respect to data inconsistency artifacts related to view-ordering.

The results are shown in Figure 5. Note that the ghosting artifacts that arise in the velocity histograms when acquiring two different velocity encode levels during the same heartbeat (Figure 5a) do not appear when we used interleaf segmentation instead (Figure 5b). In contrast to view-sharing along k_v , which causes ghosting in the velocity direction, view-sharing among spiral interleaves introduces swirling artifacts in image domain, reducing the effective unaliased spatial field-of-view by a factor of 2. However, only moving spins (flowing blood) experience these artifacts, and vessels on the same side of the neck are relatively close to each

other. Therefore, the unaliased field-of-view is wide enough to enclose all vessels on one side of the neck, so that quantification of a vessel of interest will not be disturbed by flow from neighboring vessels (e.g. measurement of flow in the left carotid arteries will not be disturbed by flow in the left jugular vein). In order to suppress signal from the opposite side of the neck, we separately reconstruct data from the left and right neck-coil elements. This uses the receiver coil sensitivity profile to avoid spiral view-sharing artifacts. Based on these observations, we used spiral interleaf view-sharing in all subsequent carotid studies.

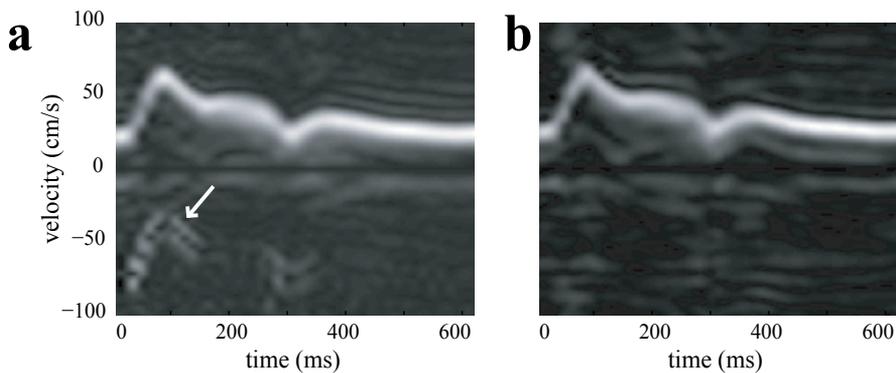


Fig. 5. Comparison of different view-orderings for multi-shot spiral FVE, in a healthy volunteer carotid study. When two or more k_v levels are acquired during the same heartbeat (a), velocity distribution changes between consecutive TRs cause ghosting artifacts along the velocity axis (arrow). This artifact is not seen if, in the same heartbeat, different spiral interleaves, but only one k_v encoding, are acquired (b).

3.4.2 Comparison with Doppler ultrasound

Representative *in vivo* results are compared with Doppler ultrasound in Figure 6. The MRI measured time-velocity histogram shows good agreement with the ultrasound measurement, as the peak velocity and the shape of the flow waveform were comparable to those observed in the ultrasound studies.

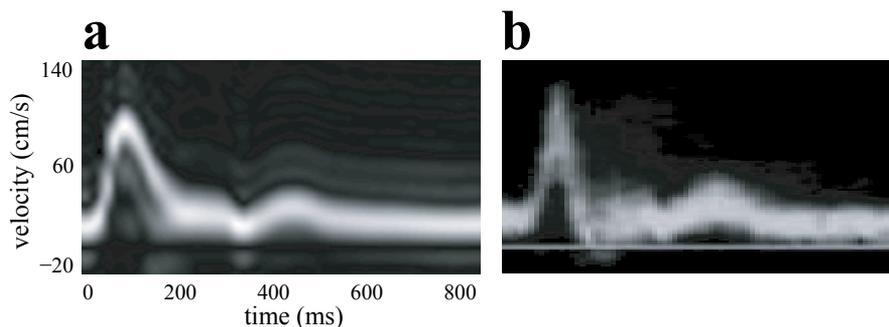


Fig. 6. Comparison of the spiral FVE method (a) with Doppler ultrasound (b), in a healthy volunteer common carotid artery study. Peak velocity and time-velocity waveforms are in good agreement.

3.4.3 Measurement of multiple flows

Figure 7 illustrates spiral FVE's ability to resolve different flows from a single dataset. A different time-velocity distribution was calculated for each voxel, and the distributions from select voxels are presented.

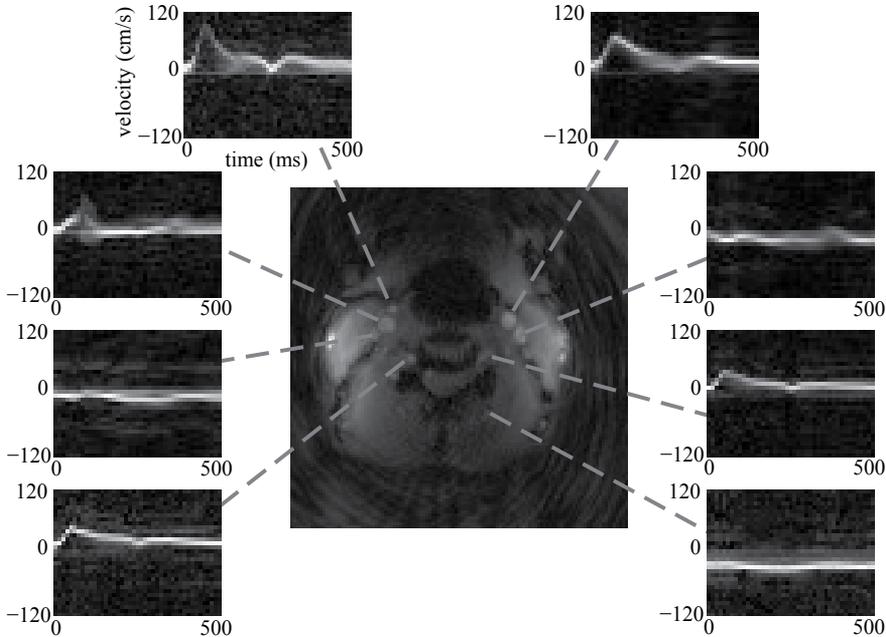


Fig. 7. Multiple flow distributions obtained from a single spiral FVE dataset. For each voxel in the image, a time-velocity distribution was calculated.

4. Wall shear rate estimation using spiral FVE

Arterial wall shear stress, the drag force acting on the endothelium as a result of blood flow, is widely believed to influence the formation and growth of atherosclerotic plaque, and may have prognostic value. According to Newton's law of viscosity, WSS can be estimated as the product of wall shear rate and blood viscosity (μ), where WSR is the radial gradient of blood flow velocity (dv/dr) at the vessel wall. Low WSS (Tang et al., 2008; Zarins et al., 1983) and highly oscillatory WSS (Ku et al., 1985) have been linked to the formation and growth of atherosclerotic plaques, and this link has been validated *in vitro* (Dai et al., 2004). High WSS has also been hypothesized as a factor responsible for the topography of atherosclerotic lesions (Thubrikar & Robicsek, 1995).

Phase contrast MRI (PC-MRI) suffers from data inconsistency, partial-volume effects (Tang et al., 1993), intravoxel phase dispersion and inadequate SNR at high spatial resolutions (Figure 8). PC-MRI is not currently capable of providing accurate absolute measurements of WSS (Boussel et al., 2009), and has been shown to underestimate blood flow velocities in the carotid bifurcation by 31–39% (Harloff et al., 2009).

As shown in section 3, spiral FVE produces accurate velocity histograms compared with those acquired with Doppler ultrasound. In addition, spiral FVE method is capable of rapid acquisition of fully localized, time-resolved velocity distributions. In this section, we propose

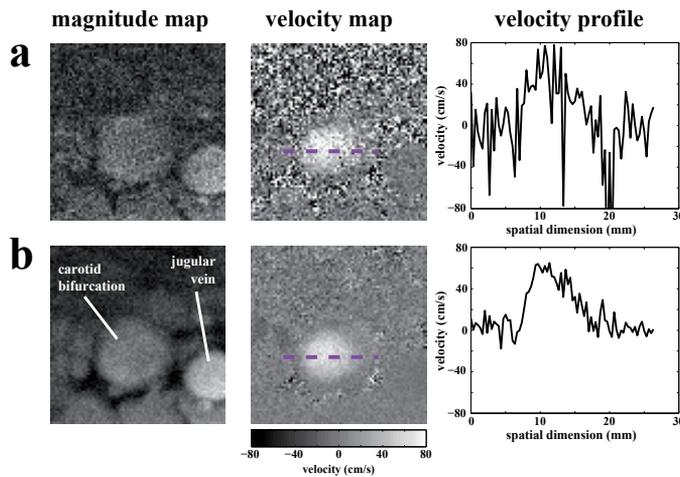


Fig. 8. Illustration of *in vivo* high-resolution 2DFT phase contrast MRI, obtained at the carotid bifurcation of a healthy volunteer at peak flow: (a) single acquisition; (b) ten signal averages. High spatial resolution is typically associated with low SNR (a). Averaging multiple acquisitions improves SNR (b), but also increases the total scan time, and may cause loss of effective resolution due to subject motion. Scan parameters: $0.33 \times 0.33 \times 3 \text{ mm}^3$ spatial resolution, 37 ms temporal resolution, 30° flip angle, 80 cm/s Venc, 2-minute scan (120 heartbeats) per acquisition.

the use of spiral FVE to estimate WSR and the oscillatory shear index. This is made possible by the method proposed by Frayne & Rutt (1995) for FVE-based WSR estimation, which is discussed next.

4.1 FVE-based WSR estimation: the Frayne method

The Frayne method for FVE-based WSR estimation (Frayne & Rutt, 1995) involves obtaining the velocity distribution for a voxel spanning the blood/vessel wall interface, and then using this distribution to reconstruct the velocity profile across the voxel, with sub-voxel spatial resolution. Assuming that signal intensity is spatially and velocity invariant, the sum of the signals from all velocities within a voxel is proportional to the total volume of material within the voxel. Furthermore, two assumptions can be made about the shape of the velocity profile: (i) the fluid velocity at the vessel wall is approximately zero, and (ii) the velocity profile within a voxel is monotonically increasing or decreasing. Using these assumptions, a step-wise discrete approximation $\tilde{v}(r)$ to the true velocity profile across a voxel can be obtained from its velocity distribution $s(v)$ by inverting the discrete function $r(v)$, which is constructed as follows:

$$r(v_i) = r(v_{i-1}) + \Delta r \cdot h(v_i), \quad \text{where} \quad h(v_i) = \frac{|s(v_i)|}{\sum_v |s(v)|}. \quad (19)$$

Note that, for each velocity bin v_i , the intra-voxel position r is incremented by a fraction of the total radial extent of the voxel (Δr). This fraction is proportional to $h(v_i)$, which is the volume fraction of the voxel that has velocity $v = v_i$. The volume fractions are

calculated by normalizing the velocity distribution. This process is demonstrated graphically in Figure 9. Spatial variations in signal intensity due to radio-frequency saturation (i.e., inflow enhancement) and due to differences in ^1H density and relaxation properties between vessel wall tissue and blood may be compensated by adjusting $s(v)$ accordingly, prior to calculating $h(v)$ (Frayne & Rutt, 1995).

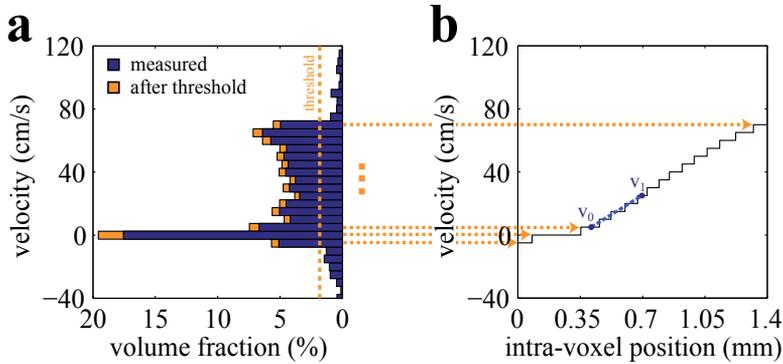


Fig. 9. Graphical illustration of the Frayne method (Frayne & Rutt, 1995). WSR is estimated from FVE velocity distributions in voxels spanning the blood/vessel wall interface. First, a threshold is applied to the velocity histogram to reduce noise sensitivity (a). Then, the volume fraction within each velocity bin is converted into a radial position across the voxel, using Eq. 19. Finally, the velocity gradient is estimated from the reconstructed velocity profile (b), within a small velocity interval $[v_0, v_1]$.

In order to reduce the effects of ringing and noise rectification due to the magnitude operation in Eq. 19, a threshold is applied to $s(v)$ before normalization (Figure 9a). The appropriate threshold level must be determined by analyzing the signal intensities in the velocity distribution for a range of velocities outside the range of expected blood flow velocities. It is assumed that signal outside this expected range is exclusively due to rectified noise (Frayne & Rutt, 1995). In our implementation, only components that are below the specified threshold *and* outside this expected range of velocities are set to zero.

The WSR can be estimated by prescribing a velocity interval $[v_0, v_1]$ and then fitting a first-order polynomial to the points of $\bar{v}(r)$ within this interval (Figure 9b). Ideally, $v_0 = 0$ and $v_1 = \Delta v$, because we wish to estimate the velocity derivative at the blood-wall interface. The SNR of shear rate estimates will increase as this velocity interval becomes larger, because of averaging across multiple velocity steps. However, as the interval becomes larger, the shear rate is averaged over a larger distance within the voxel and may deviate from the true local shear at the wall (Frayne & Rutt, 1995). Therefore, it is important to prescribe a reasonable $[v_0, v_1]$ interval. In our implementation, $v_0 = \Delta v$ and $v_1 \approx 30$ cm/s are used for an initial assessment, and then the interval is manually adjusted for selected voxels of interest. The same voxel-based approach is used with respect to the noise threshold discussed above, with a fixed threshold value being used for the initial assessment. A negative $[v_0, v_1]$ interval — e.g., $v_0 = -5$ cm/s and $v_1 = -15$ cm/s — is used when large volume fractions are measured on negative velocities, i.e., when there are large $h(v)$ values for $v < 0$. This allows measurements of negative WSR values.

4.2 *In vivo* WSR measurement

The *in vivo* measurement of carotid WSR using spiral FVE acquisitions with Frayne's reconstruction is now demonstrated. Three healthy subjects were studied. For each volunteer, five 5 mm contiguous slices (2.5 cm coverage) were prescribed perpendicular to the left carotid bifurcation. Each slice was imaged independently (separate acquisitions). Localized gradient shimming was performed, and acquisitions were prospectively ECG-gated. Several cardiac phases were acquired, spanning the systolic portion of the cardiac cycle. A time-bandwidth product 2 RF pulse was used for excitation, and the flip angle was 30° . Only through-plane velocities were measured, using 32 k_v encoding steps over a 160 cm/s velocity field-of-view (5 cm/s resolution). The velocity field-of-view was shifted from the -80 to 80 cm/s range to the -40 to 120 cm/s range during reconstruction, in order to avoid aliasing and maximize field-of-view usage. Negative velocities are encoded in order to assess negative WSR values, and also to accommodate leakage and ringing due to k_v truncation (i.e., finite velocity resolution). Spatial encoding was performed using eight 4 ms variable-density spiral interleaves (16~4 cm field-of-view, 1.4 mm resolution). The temporal resolution was 24 ms (12 ms TR, 2 views per beat). Scan time was 128 heartbeats per slice, i.e., approximately 2 minutes per slice. The subjects provided informed consent, and were scanned using a protocol approved by the institutional review board of the University of Southern California. Figures 10 and 11 show two representative sets of *in vivo* results. The WSR values are shown for manually-segmented regions-of-interest. Figure 10 illustrates the variation in WSR along all three spatial dimensions near the carotid bifurcation of subject #1. These results correspond to the cardiac phase with the highest peak velocity. Figure 11 illustrates the temporal variation of pulsatile WSR in the common carotid artery of subject #2. The results show a circumferential variation in WSR around the wall of the common carotid artery. Markl et al. (2009) recently observed a similar variation using a PC-based approach.

4.3 Estimation of oscillatory shear index

The OSI is important for the evaluation of shear stress imposed by pulsatile flow. This index describes the shear stress acting in directions other than the direction of the temporal mean shear stress vector (He & Ku, 1996; Ku et al., 1985). It is defined as the relation between the time-integral of the shear stress component acting in the direction opposite to the main direction of flow and the time-integral of the absolute shear stress. In this work, OSI was calculated as defined by He & Ku (1996), and assuming spatially and temporally invariant blood viscosity.

Measuring negative WSR with the proposed method requires voxels to be small enough to contain only reverse flow. Voxels containing both forward and reverse flow would violate the assumption of a monotonically increasing/decreasing velocity profile within the voxel. Under sufficient spatial resolution, negative WSR may be measured simply by using a negative $[v_0, v_1]$ interval.

Figure 12 presents a demonstration of *in vivo* OSI estimation, using the proposed method. Results are shown for eight representative voxels, selected around the circumference of the wall of the carotid bifurcation of subject #3. Measured velocity distributions, wall shear rates, and OSI values, corresponding to the systolic portion of the cardiac cycle, are shown for each voxel. High OSI values were observed at the wall corresponding to the carotid bulb. Non-zero OSI was also observed at the opposite wall. These findings are in agreement with PC-based OSI measurements recently reported by Markl et al. (2009). The results also suggest

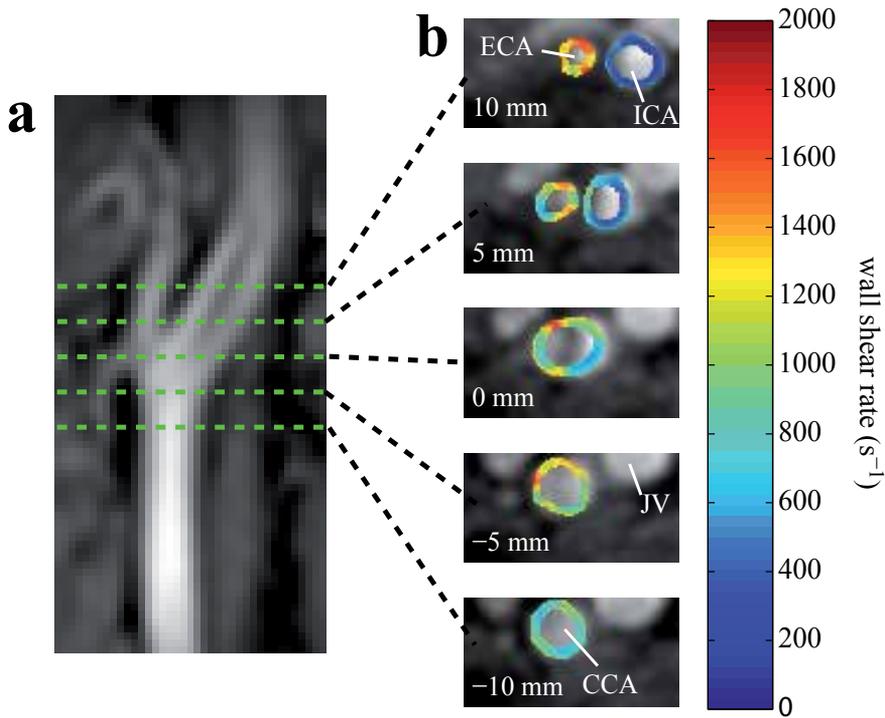


Fig. 10. Carotid WSR measured across the carotid bifurcation of subject #1: (a) slice prescription; (b) spiral FVE WSR measurements. Results correspond to the cardiac phase with the highest peak velocity (96 ms after the ECG trigger). The common (CCA), external (ECA), and internal (ICA) carotid arteries, and the jugular vein (JV), are indicated.

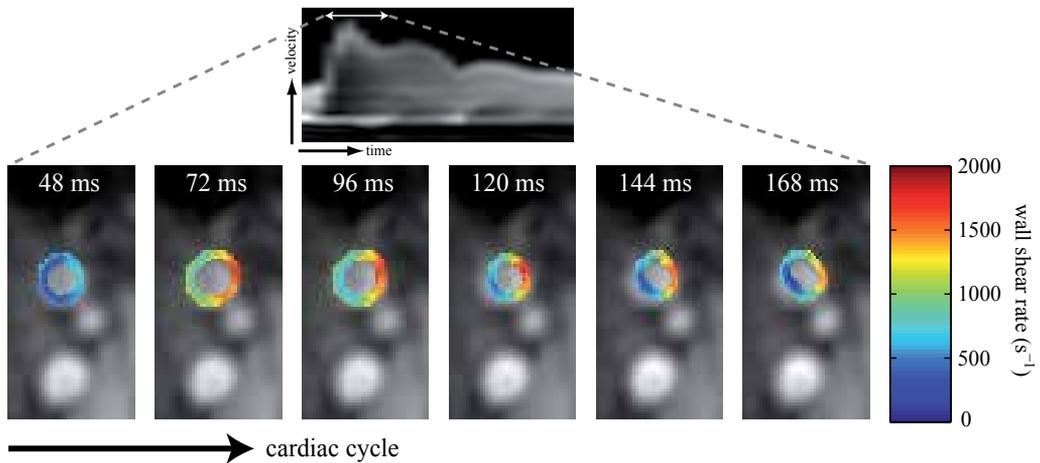


Fig. 11. Temporal variation of pulsatile WSR measured in the common carotid artery of subject #2. Results correspond to the cardiac phases acquired 48–168 ms after the ECG trigger, and to a slice prescribed 10 mm below the carotid bifurcation.

that higher spatial resolution is needed for accurately estimating OSI in some of the voxels. Notably, voxel (h) presents both positive and negative velocity components simultaneously during post-systolic deceleration. This indicates that the spatial resolution was insufficient, and the assumption of a monotonically decreasing velocity profile within the voxel was violated.

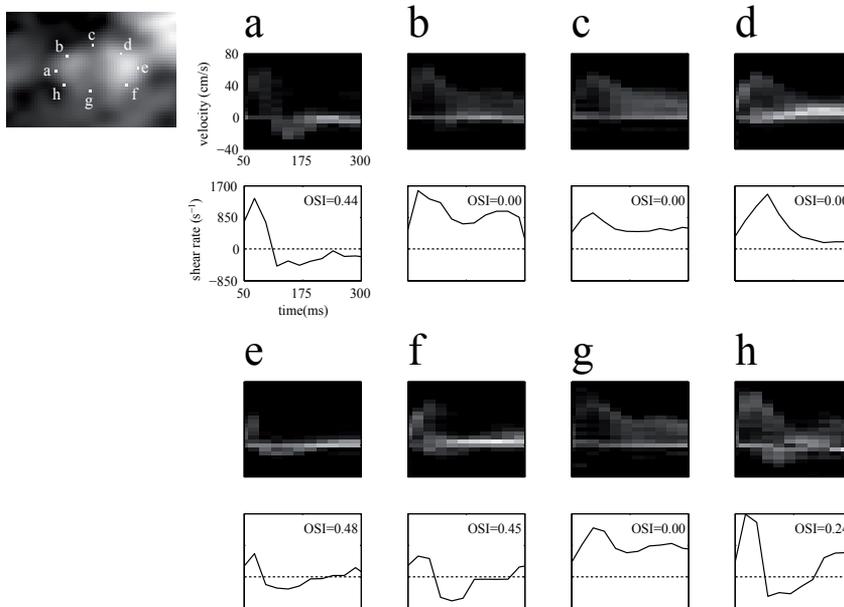


Fig. 12. *In vivo* assessment of oscillatory shear index. Results are shown for eight representative voxels (a–h), selected around the circumference of the wall of the carotid bifurcation of subject #3 (see inset). Measured velocity distributions, wall shear rates, and OSI values, corresponding to the systolic portion of the cardiac cycle, are shown for each voxel.

5. Assessment of carotid flow using CFD (and MRI)

Computational fluid dynamics methods are concerned with the approximate solution of the fluid motion as well as with the interaction of the fluid with solid bodies. Fluid dynamic problems, essentially, are based on nonlinear systems of coupled partial differential equations. Attempting to solve those systems analytically, in general, is an impracticable task. Over the years, researchers have developed methodologies, codes, schemes and algorithms to find approximate solutions, focusing on accuracy and speed of convergence, for problems involving fluid dynamics and heat transfer.

Historically, CFD started in the early 1970s, triggered by the availability of increasingly more powerful mainframes. In the beginning, the study was limited to high-technology engineering areas of aeronautics and astronautics. Nowadays, computational fluid dynamics methodologies are routinely employed in many fields such as: racing car design, ship design, meteorology, oil recovery, civil engineering, airspace engineering, and biomedical engineering.

This section is dedicated to the use of CFD for evaluation of blood flow in the carotid arteries. *In vivo* 3D blood flow patterns can be either measured directly using PC-MRI, or obtained from model-based CFD calculations. PC-MRI is accurate, but suffers from long scan times and limited spatio-temporal resolution and SNR. CFD provides arbitrarily high resolution and reduced scan times, but its accuracy hinges on the model assumptions. A numerical framework for constructing a flow field that is influenced by both direct measurements and a fluid physics model is described.

5.1 CFD in biomedical engineering

Recently, CFD is playing an important role in the analysis of blood flow. This technique of flow visualization has been widely applied in problems involving arterial diseases with reconstruction of hemodynamics in realistic models based on images generated by standard *in vivo* medical visualization tools, such as MRI. CFD can be used to improve the data obtained using MRI for estimation of the flow properties of blood vessels. CFD can also be used to compare real data obtained from patients with simulated data models using realistic geometries. These geometries may even be constructed using MRI data.

Boussel et al. (2009) compared a time-dependent 3D phase-contrast MRI sequence with patient-specific CFD models for patients who had intracranial aneurysms. The evolution of intracranial aneurysms is known to be related to hemodynamic forces such as wall shear stress and maximum shear stress. Harloff et al. (2010) used CFD image-based modeling as an option to improve the accuracy of MRI-based WSS and OSI estimation, with the purpose of correlating atherogenic low WSS and high OSI with the localization of aortic plaques. Steinman et al. (2002) used a novel approach for noninvasively reconstructing artery wall thickness and local hemodynamics at the human carotid bifurcation. Three-dimensional models of the lumen and wall boundaries, from which wall thickness can be measured, were reconstructed from black blood MRI. Along with time-varying inlet/outlet flow rates measured via PC-MRI, the lumen boundary was used as input for a CFD simulation of the subject-specific flow patterns and wall shear stress. Long et al. (2003) were concerned with the reproducibility of geometry reconstruction, one of the most crucial steps in the modeling process. Canstein et al. (2008) used rapid prototyping to transform aortic geometries as measured by contrast-enhanced MR angiography into realistic vascular models with large anatomical coverage. Visualization of characteristic 3D flow patterns and quantitative comparisons of the *in vitro* experiments with *in vivo* data and CFD simulations in identical vascular geometries were performed to evaluate the accuracy of vascular model systems.

CFD can also be used with other imaging techniques such as tomography. Howell et al. (2007) used CFD to study the temporal and spatial variations in surface pressure and shear through the cardiac cycle on models of bifurcated stent-grafts derived from computed tomography in patients who had previously undergone endovascular repair of abdominal aortic aneurysm.

5.2 MRI-driven CFD

A numerical framework for constructing a flow field that is influenced by both direct measurements and a fluid physics model is now described. The PC-MRI signal is expressed as a linear function of the velocities on the underlying computational grid, and a tunable parameter controls the relative influence of direct measurements and the model assumptions. The feasibility of the proposed approach is demonstrated in the carotid bifurcation of a healthy

volunteer. The results show that this methodology produces flow fields that are in better agreement with direct PC-MRI measurements than CFD alone.

5.2.1 Proposed approach

Here, blood is modeled as an incompressible Newtonian fluid with constant viscosity μ . This model is widely used in *in vivo* CFD analysis, and underlies the commonly held definition of wall shear stress as being proportional to dv/dr , where v is the velocity tangential to the vessel wall, and r is the perpendicular distance from the wall. The task of a CFD routine is to solve the momentum and continuity equations that the flow field — subject to the assumption of Newtonian flow — must obey. Following the control-volume formulation introduced by Patankar (1980), the momentum equation is

$$\rho \frac{D\mathbf{v}}{Dt} = -\nabla p + \mu \Delta \mathbf{v}, \quad (20)$$

where ρ is the fluid density, $\mathbf{v} = (u, v, w)$ is the velocity vector field, p is the pressure field, ∇ is the gradient operator, and Δ is the Laplacian operator. The flow field must also satisfy mass conservation (or continuity), which can be expressed as

$$\nabla \cdot \mathbf{v} = 0. \quad (21)$$

Equations (20) and (21) must be solved for the unknown scalar field variables u , v , w , and p . These equations are non-linear and coupled, and attempting to solve them directly in one step is a formidable, if not impossible, task.

The solver was built on the SIMPLER algorithm developed by Patankar (1980), which is a well-known and established numerical routine for solving Eqs. (20) and (21). SIMPLER starts with an initial estimate for the velocity field, and updates this estimate in an iterative fashion. At each iteration i , the discretized equations are linearized using the velocity estimate \mathbf{v}_{i-1} at the previous iteration, which produces a square system matrix \mathbf{A} for each velocity component. For example, for the z velocity component, we can write

$$\mathbf{A}_{i-1} \mathbf{w}_i = \mathbf{b}_{i-1}, \quad (22)$$

where \mathbf{w} is the z velocity in all voxels in the 3D calculation domain (stacked to form a 1D column vector), and \mathbf{b} is a constant column vector. Note that the pressure field is updated periodically based on the current velocity field estimate, and hence does not appear explicitly in (22).

The key step in our approach is to add additional rows to \mathbf{A} and \mathbf{b} that incorporate MRI measurements of one or more velocity components. The underlying assumption we make here is that the velocity measured with PC-MRI is equal to the average velocity within the voxel, and can hence be expressed as a linear combination of the velocities on the underlying CFD calculation grid. For example, for the z (S/I) velocity component we have

$$\mathbf{w}_{MR} \approx \mathbf{A}_{MR} \mathbf{w}, \quad (23)$$

where \mathbf{w}_{MR} is the z velocity component measured with PC-MRI.

Combining Eqs. (22) and (23), we have

$$\begin{bmatrix} \mathbf{A}_{i-1} \\ s\mathbf{A}_{MR} \end{bmatrix} \mathbf{w}_i = \begin{bmatrix} \mathbf{b}_{i-1} \\ s\mathbf{w}_{MR} \end{bmatrix}. \quad (24)$$

The weighting parameter s is an adjustable scalar that determines the influence of the MRI measurements on the solution. For example, $s = 0$ produces a conventional, unconstrained CFD solution. Here, we solve Eq. (24) in the least squares sense, using the conjugate gradient method on the normal equations. Hence, the solution \mathbf{w} at each iteration represents a weighted least squares estimate, with relative contributions of fluid physics and direct MRI measurement being controlled by the parameters after multiple interactions. The velocity field $\mathbf{v} = (u, v, w)$ converges toward the solution for a particular point t . Time is then incremented by an amount Δt , and the iterative procedure is repeated to obtain the solution at $t + \Delta t$. This way, it is possible to predict a time-dependent flow field, as long as the true velocity field \mathbf{v} at $t = 0$ is known. In this work, however, the exact velocity field is not known, since only a relatively low-resolution and noisy PC-MRI velocity field is available. Instead, we will calculate the steady flow \mathbf{v}_s subject to the measured inlet and outlet velocities. We will furthermore assume that the velocity field obtained with PC-MRI at one time-point (near peak flow) is representative of the steady flow given the inlet and outlet velocities at that time-point. We obtain \mathbf{v}_s by starting with an initial guess for \mathbf{v} , and carrying the simulations forward in time until convergence.

5.2.2 *In vivo* demonstration

PC-MRI data were obtained from four time-resolved 3DFT FGRE image volumes in the carotid artery in one healthy volunteer ($1 \times 1 \times 2.5$ mm³ voxel size; FOV $16 \times 12 \times 7.5$ cm³; TR 7.0 ms; flip angle 15°; temporal resolution 56 ms; Venc 1.6 m/s; 7 minutes per scan), on a GE Signa 3T EXCITE HD system (4 G/cm and 15 G/cm/ms maximum gradient amplitude and slew rate) with a 4-channel carotid receive coil array. The through-slab (z) axis was oriented along the S/I direction. The subject provided informed consent, and was scanned using a protocol approved by the institutional review board of the University of Southern California. PC-MRI velocity component maps \mathbf{u}_{MR} , \mathbf{v}_{MR} , and \mathbf{w}_{MR} were calculated using data from one receive coil. Residual linear velocity offsets in each velocity component map (e.g. due to eddy-currents) were removed by performing a linear fit to manually defined 3D regions containing only stationary tissue. The vessel lumen was segmented by manually outlining the vessel borders from a stack of 2D axial slices.

The solver calculations assumed a blood viscosity of 0.0027 Pa·sec and a blood density of 1060 kg/m³ (Reynolds number of order 1000), and no-slip boundary conditions. Calculations were performed on a Cartesian grid of 1 mm isotropic resolution, and all algorithms were implemented in Matlab. Low-resolution (truncated to $1 \times 1 \times 3$ mm³) measurements of the S/I velocity component \mathbf{w}_{MR} near the time-point of peak flow were incorporated into the solver, and a steady velocity vector field \mathbf{v}_s was calculated as described above. Hence, \mathbf{w} was partially constrained by the measured velocity component \mathbf{w}_{MR} , whereas \mathbf{u} and \mathbf{v} were determined solely from the fluid physics model. As in the original SIMPLER algorithm (Patankar, 1980), u , v , and w were defined on regular grids that were staggered by half a grid spacing (in different directions) with respect to the centered grid. This is done to avoid a checkerboard solution for the pressure and velocity fields (Patankar, 1980). The calculation domain was rectangular of size $20.5 \times 16.5 \times 25$ mm³.

Figure 13 compares flow fields in the carotid artery obtained with PC-MRI only, CFD only ($s = 0$), and the proposed combined solver algorithm with $s = 5$. These flow fields were calculated from 4, 1, or 2 time-resolved 3D image acquisitions, corresponding to a total scan time of 28, 7, and 14 minutes, as indicated in the figure. In the common carotid artery, flow is predominantly along z , and all methods produce comparable flow fields. In the bifurcation, the bulk flow pattern appears qualitatively similar for all three methods, which indicates that the underlying fluid physics model makes reasonable predictions regarding the transverse velocities. However, comparison of the CFD and PC-MRI results shows that CFD underestimates the velocities in the external carotid artery. The combined solver, which strikes a compromise between CFD and PC-MRI, brings the calculated flow field in closer agreement with the values measured with PC-MRI.

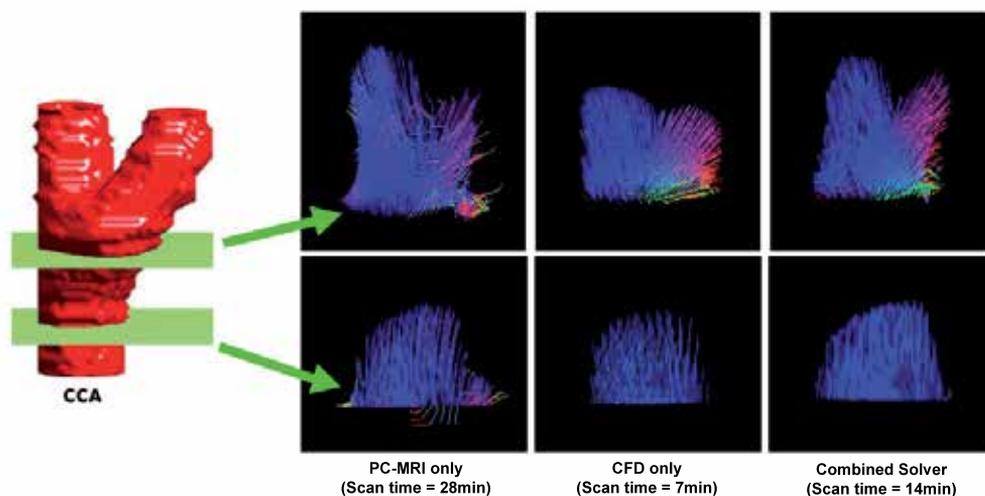


Fig. 13. Blood flow in the carotid bifurcation (top) and the common carotid artery (bottom) obtained with 4-point phase contrast MRI (left), CFD (center), and the combined solver ($s = 5$, right). The lines show the path of massless particles during the course of 60 ms (top) or 20 ms (bottom), under the assumption of a constant velocity field. Pathlines are RGB color-coded according to the local velocity direction (vertical=blue; in-plane=red and green). These velocity fields were obtained from 4 (left), 1 (center), or 2 (right) MRI scans, corresponding to total scan times of 28, 7 and 14 minutes, respectively. Compared to CFD, the combined solver produces flow fields that are in better qualitative and quantitative agreement with PC-MRI.

6. Conclusion

In this chapter, we have introduced spiral FVE, a rapid MRI method for fully-localized measurement of cardiovascular blood flow. The proposed method was shown to be capable of measuring blood flow in the carotid arteries and estimating wall shear stress and oscillatory shear index at the carotid bifurcation. In addition, we proposed a combined CFD-MRI solver that integrates the non-linear coupled system of the fluid partial differential equations using MRI data as initial data. This methodology, which has a tunable parameter, is capable of

producing flow fields that are in better agreement with direct PC-MRI measurements than CFD alone.

MRI is potentially the most appropriate technique for addressing all aspects of a complete cardiovascular disease examination. The evaluation of carotid flow will be a necessary capability in such examination. The methodologies presented in this chapter can improve the quality of the MRI measured data, reducing scan time, and improving the SNR. These techniques should improve the diagnosis and understanding of carotid diseases.

7. References

- Boussel, L., Rayz, V., Martin, A., Acevedo-Bolton, G., Lawton, M. T., Higashida, R., Smith, W. S., Young, W. L. & Saloner, D. (2009). Phase-contrast magnetic resonance imaging measurements in intracranial aneurysms in vivo of flow patterns, velocity fields, and wall shear stress: comparison with computational fluid dynamics, *Magn Reson Med* 61(2): 409–417.
- Canstein, C., Cachot, P., Faust, A., Stalder, A. F., Bock, J., Frydrychowicz, A., Kuffer, J., Hennig, J. & Markl, M. (2008). 3D MR flow analysis in realistic rapid-prototyping model systems of the thoracic aorta: comparison with in vivo data and computational fluid dynamics in identical vessel geometries, *Magn Reson Med* 59(3): 535–546.
- Carvalho, J. L. A., Nielsen, J. F. & Nayak, K. S. (2010). Feasibility of in vivo measurement of carotid wall shear rate using spiral Fourier velocity encoded MRI, *Magn Reson Med* 63(6): 1537–1547.
- Dai, G., Kaazempur-Mofrad, M. R., Natarajan, S., Zhang, Y., Vaughn, S., Blackman, B. R., Kamm, R. D., Garcia-Cardena, G. & Gimbrone Jr., M. A. (2004). Distinct endothelial phenotypes evoked by arterial waveforms derived from atherosclerosis-susceptible and -resistant regions of human vasculature, *Proceedings of the National Academy of Sciences of the United States of America* 101(41): 14871–14876.
- Frayne, R. & Rutt, B. K. (1995). Measurement of fluid-shear rate by Fourier-encoded velocity imaging, *Magn Reson Med* 34(3): 378–387.
- Glover, G. H. & Pelc, N. J. (1988). A rapid-gated cine MRI technique, *Magn Reson Annu* pp. 299–333.
- Hahn, E. L. (1960). Detection of sea-water motion by nuclear precession, *J Geophys Res* 65(2): 776–777.
- Harloff, A., Albrecht, F., Spreer, J., Stalder, A. F., Bock, J., Frydrychowicz, A., Schollhorn, J., Hetzel, A., Schumacher, M., Hennig, J. & Markl, M. (2009). 3D blood flow characteristics in the carotid artery bifurcation assessed by flow-sensitive 4D MRI at 3T, *Magn Reson Med* 61(1): 65–74.
- Harloff, A., Nussbaumer, A., Bauer, S., Stalder, A. F., Frydrychowicz, A., Weiller, C., Hennig, J. & Markl, M. (2010). In vivo assessment of wall shear stress in the atherosclerotic aorta using flow-sensitive 4D MRI, *Magn Reson Med* 63(6): 1529–1536.
- He, X. & Ku, D. N. (1996). Pulsatile flow in the human left coronary artery bifurcation: average conditions, *J Biomech Eng* 118(1): 74–82.
- Hoskins, P. R. (1996). Accuracy of maximum velocity estimates made using Doppler ultrasound systems, *Br J Radiol* 69(818): 172–177.

- Howell, B. A., Kim, T., Cheer, A., Dwyer, H., Saloner, D. & Chuter, T. A. M. (2007). Computational fluid dynamics within bifurcated abdominal aortic stent-grafts, *J Endovasc Ther* 14(2): 138–143.
- Hu, B. S., Pauly, J. M. & Macovski, A. (1993). Localized real-time velocity spectra determination, *Magn Reson Med* 30(3): 393–398.
- Kim, C. S., Kiris, C., Kwak, D. & David, T. (2006). Numerical simulation of local blood flow in the carotid and cerebral arteries under altered gravity, *J Biomech Eng* 128(2): 194–202.
- Ku, D. N., Giddens, D. P., Zarins, C. K. & Glagov, S. (1985). Pulsatile flow and atherosclerosis in the human carotid bifurcation. Positive correlation between plaque location and low oscillating shear stress, *Arterioscler Thromb Vasc Biol* 5: 293–302.
- Long, Q., Ariff, B., Zhao, S. Z., Thom, S. A., Hughes, A. D. & Xu, X. Y. (2003). Reproducibility study of 3D geometrical reconstruction of the human carotid bifurcation from magnetic resonance images, *Magn Reson Med* 49(4): 665–674.
- Markl, M., Wegent, F., Bauer, S., Stalder, A. F., Frydrychowicz, A., Weiller, C., Schumacher, M. & Harloff, A. (2009). Three-dimensional assessment of wall shear stress distribution in the carotid bifurcation, *Proc, ISMRM, 17th Annual Meeting*, Honolulu, p. 323.
- Moran, P. R. (1982). A flow velocity zeugmatographic interlace for NMR imaging in humans, *Magn Reson Imaging* 1(4): 197–203.
- Moran, P. R., Moran, R. A. & Karstaedt, N. (1985). Verification and evaluation of internal flow and motion. true magnetic resonance imaging by the phase gradient modulation method, *Radiology* 154(2): 433–441.
- Nayler, G. L., Firmin, D. N. & Longmore, D. B. (1986). Blood flow imaging by cine magnetic resonance, *J Comput Assist Tomogr* 10(5): 715–722.
- O'Donnell, M. (1985). NMR blood flow imaging using multiecho, phase contrast sequences, *Med Phys* 12(1): 59–64.
- Papathanasopoulou, P., Zhao, S., Köhler, U., Robertson, M. B., Long, Q., Hoskins, P., Xu, X. Y. & Marshall, I. (2003). MRI measurement of time-resolved wall shear stress vectors in a carotid bifurcation model, and comparison with CFD predictions, *J Magn Reson Imaging* 17(2): 153–162.
- Patankar, S. V. (1980). *Numerical Heat Transfer and Fluid Flow*, Hemisphere Publishing Corporation.
- Rebergen, S. A., van der Wall, E. E., Doornbos, J. & de Roos, A. (1993). Magnetic resonance measurement of velocity and flow: technique, validation, and cardiovascular applications, *Am Heart J* 126(6): 1439–1456.
- Singer, J. R. (1959). Blood flow rates by nuclear magnetic resonance measurements, *Science* 130(3389): 1652–1653.
- Singer, J. R. & Crooks, L. E. (1983). Nuclear magnetic resonance blood flow measurements in the human brain, *Science* 221(4611): 654–656.
- Steinman, D. A., Thomas, J. B., Ladak, H. M., Milner, J. S., Rutt, B. K. & Spence, J. D. (2002). Reconstruction of carotid bifurcation hemodynamics and wall thickness using computational fluid dynamics and MRI, *Magn Reson Med* 47(1): 149–159.
- Tang, C., Blatter, D. D. & Parker, D. L. (1993). Accuracy of phase-contrast flow measurements in the presence of partial-volume effects, *J Magn Reson Imaging* 3(2): 377–385.
- Tang, D., Yang, C., Mondal, S., Liu, F., Canton, G., Hatsukami, T. & Yuan, C. (2008). A negative correlation between human carotid atherosclerotic plaque progression and plaque wall stress: in vivo MRI-based 2D/3D FSI models, *J Biomech* 41(4): 727–736.

- Thubrikar, M. J. & Robicsek, F. (1995). Pressure-induced arterial wall stress and atherosclerosis, *Ann Thorac Surg* 59(6): 1594–1603.
- van Dijk, P. (1984). Direct cardiac NMR imaging of heart wall and blood flow velocity, *J Comput Assist Tomogr* 8(3): 429–436.
- Winkler, A. J. & Wu, J. (1995). Correction of intrinsic spectral broadening errors in Doppler peak velocity measurements made with phased sector and linear array transducers, *Ultrasound Med Biol* 21(8): 1029–1035.
- Zarins, C. K., Giddens, D. P., Bharadvaj, B. K., Sottiurai, V. S., Mabon, R. F. & Glagov, S. (1983). Carotid bifurcation atherosclerosis. Quantitative correlation of plaque localization with flow velocity profiles and wall shear stress, *Circulation Research* 53(4): 502–514.

Numerical Simulation for Intranasal Transport Phenomena

Takahisa Yamamoto¹, Seiichi Nakata², Tsutomu Nakashima³
and Tsuyoshi Yamamoto⁴

¹*Gifu National College of Technology*

²*Fujita Health University*

³*Nagoya University*

⁴*Kyushu University*

Japan

1. Introduction

More than 10 million people in Japan suffer some of nasal diseases every year (Haruna (2003)); the paranasal sinusitis (so-called the empyema), hypertrophic rhinitis and inferior concha inflammation. Nebulizer treatment has been used for the nasal diseases. The effectiveness of the nebulizer treatment has been confirmed from clinical view points until now. However there are a few researches that evaluate the effect of the nebulizer treatment theoretically and quantitatively, i.e., the transport characteristics of medicinal droplets and their deposition on the inflammation areas of nasal wall.

The development of medical image processing technique is in progress and now gives us exquisitely detailed anatomic information. Some researchers calculated blood flow inside vital arteries and intranasal airflow characteristics by means of the medical image processing technique as well as Computed Fluid Dynamics (CFD) analysis. As for the CFD analysis of intranasal flow, Weinhold et al. constructed both a transparent resin model and numerical three-dimensional anatomy model with nasal cavities using a patient's CT data (Weinhold & Mlynski (2004)). They subsequently made clear airflow characteristics in the nose experimentally and numerically, and found that pressure drop was a main factor of nasal airflow. Lindemann et al. focused at a case which underwent radical sinus surgery (Lindemann et al. (2004; 2005)). In their case, both the lateral nasal wall and the turbinates, inhibiting physiological airflow, were removed by the surgery to realize the enlargements of the nasal cavity volumes and to increase the ratio between nasal cavity volume and surface area. However the researches mentioned here dealt with only a few patient-cases even though there are individual differences in the shape of human nasal cavity and in the grade of medical conditions. The past researches considered the individual differences insufficiently. The authors analyzed intranasal transport phenomena for several patient cases and compared each others in the past study (Monya et al. (2009); Yamamoto et al. (2009)). From the results characteristics of airflow and medicinal droplet transportation strongly depend on inflow conditions such as inflow angle, velocity and size of particle even if there are the individual differences for the shape of patient's nasal cavity.

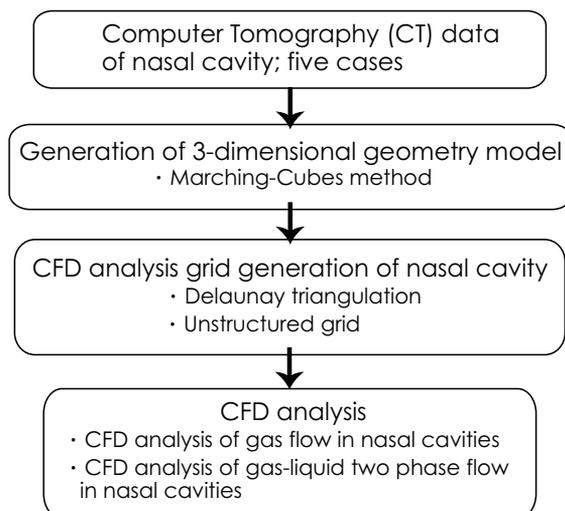


Fig. 1. Flowchart diagram of CFD analysis for heat and mass transport phenomena in nasal cavity

Figure 1 indicates the flowchart of the CFD analysis for intranasal transport phenomena. This chapter shows you detail features of nasal cavity and some nasal diseases, subsequently how to yield three-dimensional model of nasal cavity and calculation mesh for CFD analysis. Finally the characteristics of intranasal transport phenomena are presented in this chapter.

2. Anatomy of nasal cavity

2.1 Structure of nasal cavity and nasal sinuses

Figure 2 shows front and side views of human head, nasal cavity, nasal conchae and nasal sinuses. The nasal cavity is a large air filled space, and the pathway of respiration flow. The nasal cavity conditions the air to be received by the other areas of the respiratory tract. The nasal conchae, which has large surface area, warms and cools passing air through the nasal cavity. At the same time, the passing air is humidified and cleaned by nasal conchae and nasal hairs; most parts of dusts and particulate matters are removed there. The nasal sinuses are four caviums existing in the ossa faciei as shown in Fig.2; maxillary sinus, frontal sinus, ethmoidal sinus and sphenoidal sinus. These sinuses connect to the nasal cavity via very small ducts, ostia; both diameter and streamwise scales are several millimeters. Figures 3 represents coronal and sagittal cross-sectional CT data for nasal cavity and nasal sinuses. The white areas in the figure represent bones of the patient's head, and the gray areas indicate muscles and fat. The nasal cavity and sinuses are represented as dark area in the CT data.

In paranasal sinusitis, both the sinuses and the ostia are infected by bacillus and the virus. The details of such the diseases are stated in following subsections.

2.2 Nasal diseases

The nose often suffers some injuries such as fractures because the nose is protruded forward from the human face. Infection, epistaxis, as well as polyps are also found in the nasal diseases. Almost all people have contract rhinitis which is caused by the inflammation of nasal mucosa. Sinusitis will develop when the inflammation of the rhinitis spread to the sinuses via ostia.

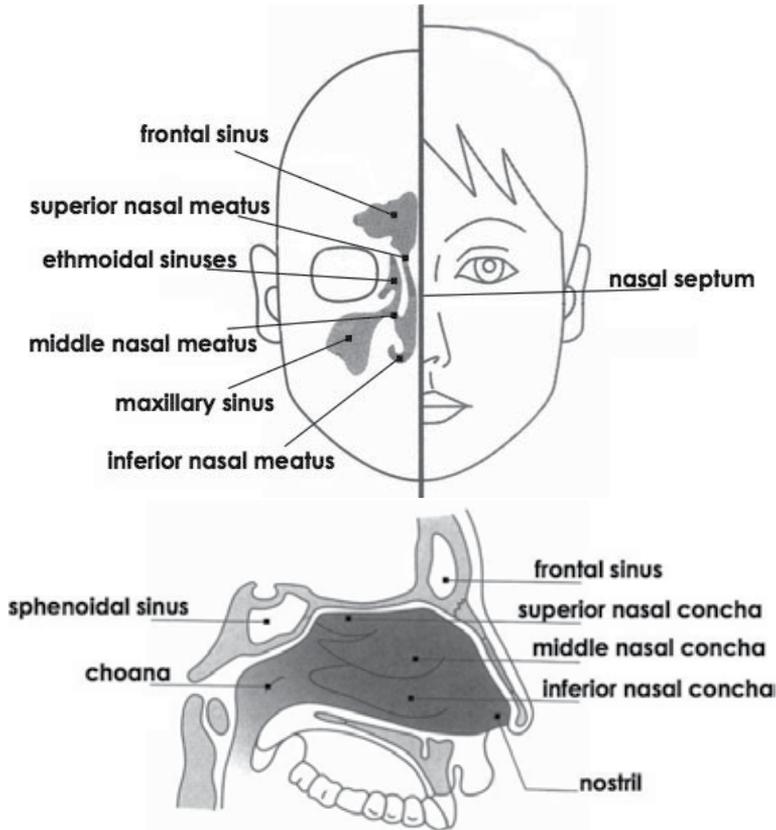


Fig. 2. Front and side view of human head and nasal sinuses (adapted from Haruna (2003))

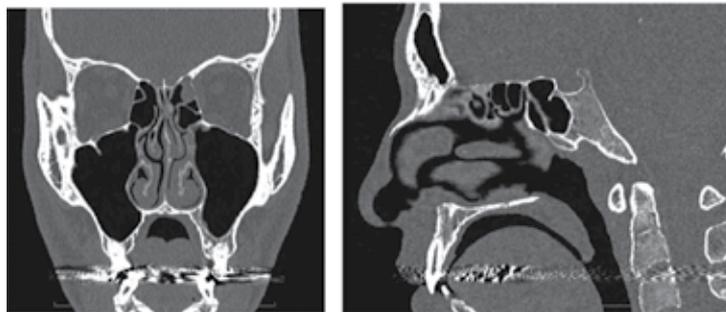


Fig. 3. Cross-sectional CT data for nasal cavity and nasal sinuses: left) coronal and right) sagittal sections

2.2.1 Deviated nasal septum

Nasal septum locates on the middle of left and right nasal cavities, being almost straight stretches. Many people has a slightly-curved nasal septum as shown in the left-side figure of Fig.5. There are small size different between the left and right nostrils. The right figure of Fig.5 indicates a CT image that is the case of deviated nasal septum. The deviated nasal septum is caused by external nasal injuries and/or congenital origin. As for the cases, one of the nasal

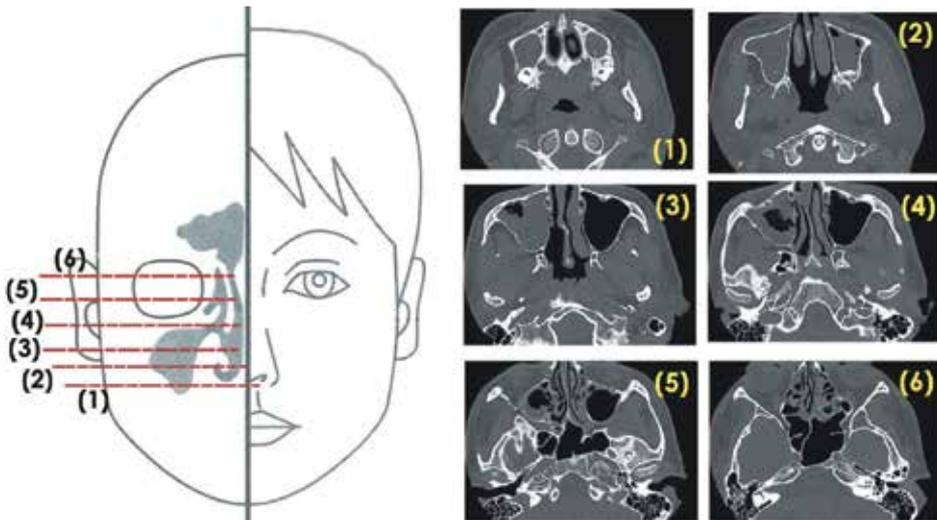


Fig. 4. Series of axial cross-sectional CT data for nasal cavity and nasal sinuses (adapted from Haruna (2003))

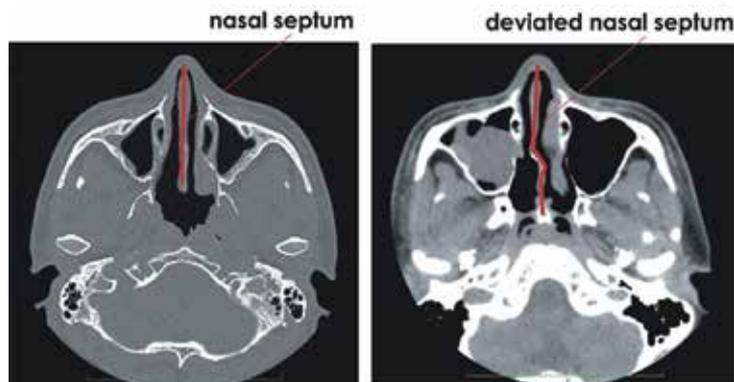


Fig. 5. Axial cross-sectional CT data for nasal septum: left) healthy nasal septum and right) deviated nasal septum

cavity is significantly smaller than the other. Then the patient suffers nasal congestion and drying nasal cavity, consequently epistaxis. In the sever cases operative treatment is required.

2.2.2 Rhinitis (nasal inflammation)

The nasal cavity is the part that easily causes infection in the upper respiratory tracts. Rhinitis (nasal inflammation) is mainly classified into two cases; the one is acute rhinitis and the other is chronic rhinitis. The acute rhinitis is generally caused by the infection of virus and allergies of pollen and house dust. The difference point between both rhinitis is the term of these diseases; the former occurs in a short term and the latter in a long term (two weeks and more). In some cases, the chronic rhinitis occasionally accompanies the chronic sinusitis.

2.2.3 Sinusitis

Figure 6 shows cross-sectional CT data for both healthy maxillary sinus and sinusitis one. The sinusitis happens anywhere about four kinds of the sinus paranasalis such as the maxillary sinus, sinus ethmoidales, the master wrestlers caves, and the sphenoidal sinuses. The sinusitis is classified into the acute sinusitis (in a short term) and the chronic sinusitis (in a long term).

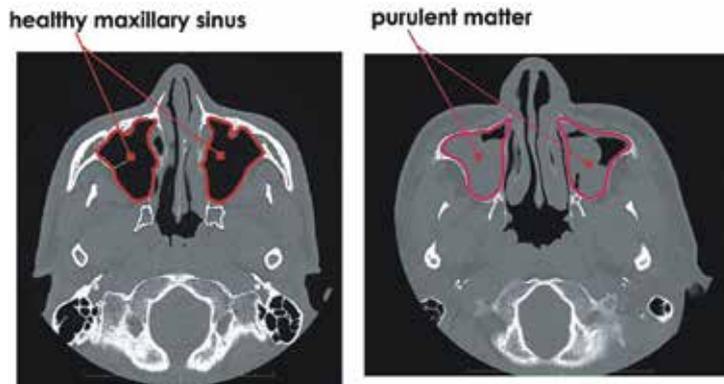


Fig. 6. Cross-sectional CT data for maxillary sinuses: left) healthy maxillary sinus and right) sinusitis

The acute sinusitis is caused by a variety of bacteria and virus, often develops after the blockage of ostia which are opening of sinuses connecting to nasal cavity. The blockage normally occurs as a result of virus infection in the upper respiratory caused by a cold. The cold brings on the swelling of the mucous membrane of nostril, and then the ostia are easily obstructed. In the blocked sinus, the air inside the cavity are absorbed to the bloodstream, and subsequently the pressure inside the sinus decreases. This pressure drop produce sever pains and the sinus fills with secretory fluid. The secretory fluid becomes a breeding ground for viruses. In order to attack the viruses, white blood cell and other matters such as the secretory fluid are aggregated in the blocked sinus, then the pressure in the sinus is increased and pain becomes more sever.

The case that a symptom of the sinusitis continues more than 8-12 weeks is called the chronic sinusitis. The developing mechanism of the chronic sinusitis has not been clear, however it has been confirmed that the chronic sinusitis develops after the infection of the virus, severe allergies and the influences of the environmental pollution material. A genetic contributor is also regarded as one of the factors concerned with development of the chronic sinusitis.

3. Construction of three-dimensional nasal cavity model

In order to construct a three-dimensional model for nasal cavity, data conversion algorithm from two-dimensional CT/MRI data to three-dimensional geometric model as well as smoothing algorithm for the three-dimensional model are required. Furthermore, mesh generation models are needed to execute CFD analysis of intranasal transport phenomena. This section describes the fundamental theories of these algorithms.

3.1 Data conversion algorithm from 2-D to 3-D: marching cubes algorithm

Marching cubes algorithm is one of the latest models of surface construction used for viewing three-dimensional data (Lorenson & Cline (1987)). This algorithm extracts a polygonal mesh

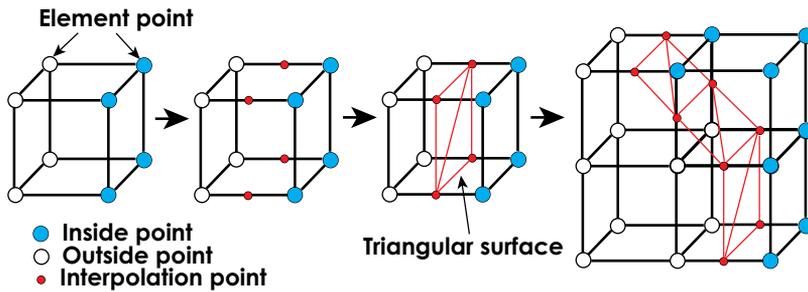


Fig. 7. Configurations of fourteen unique cubes for marching cubes algorithm

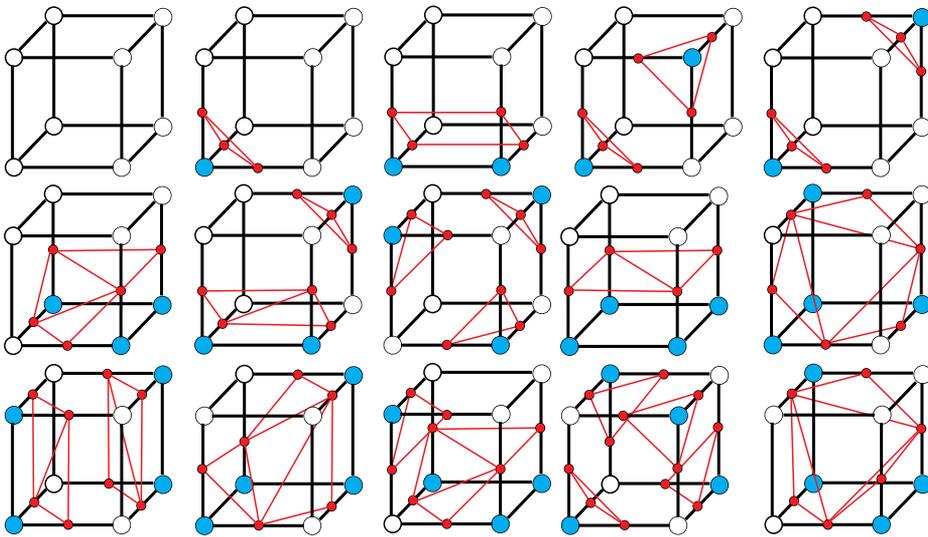


Fig. 8. Triangular surface patterns for the marching cubes algorithm

of isosurface from a three-dimensional scalar field. The algorithm proceeds through the scalar field, taking eight neighbor locations at a time, then determining the polygons needed to represent the part of the isosurface that passes through this cube. The individual polygons are then fused into the desired surface as shown in Fig.7. This is done by creating an index to a pre-calculated array of 256 possible polygon configurations ($2^8 = 256$) within the cube, by treating each of the 8 scalar values as a bit in an 8-bit integer. Finally each vertex of the generated polygons is placed on the appropriate position along the cube's edge by linearly interpolating the two scalar values that are connected by that edge. The precalculated array of 256 cube configurations can be obtained by reflections and symmetrical rotations of 14 unique cases as shown in Fig.8.

The applications of this algorithm are mainly concerned with medical visualizations such as CT and MRI scan data images, and special effects on three-dimensional modeling.

3.2 Smoothing algorithm

In the image processing and image transmission, noises are physically captured and fed into the processes due to various factors. For instance, noises are generated caused by the image resolution and image slice thickness of CT and MRI in the field of medical engineering. When the noise generation mechanism is clear and mathematically modeled, we can remove the noise using the optimized filter. However, it is difficult to modeled the noise mathematically in almost all situations. The smoothing processes are needed to reduce and minimize the influence of the noise. The noise makes patterned indented surfaces in the three-dimensional model, therefore affects the quality of the constructed nasal cavity model significantly. There are two smoothing algorithms that often uses in the medical image processing; moving-average algorithm and median-average algorithm. This subsection presents both algorithms as follows.

3.2.1 Moving-average algorithm

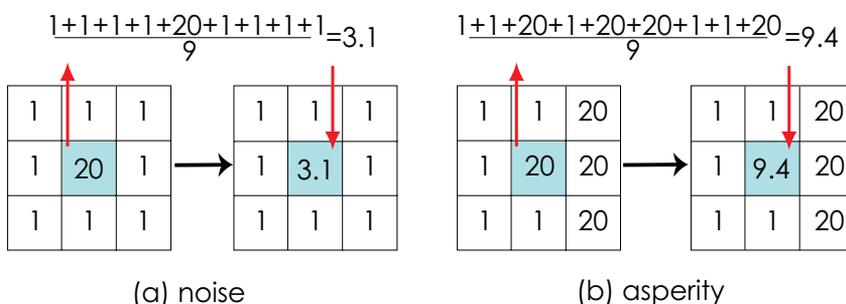


Fig. 9. Procedure of Moving-Average Algorithm for 3×3 Cubes

Figure 9 shows the application of the moving-average algorithm for 3×3 cubes (Savitzky & Golay (1964); Sun (2006); Tamura (2002)). The number inside a cube expresses a scalar variable, for example the concentration of a chemical specie. The case of Fig.9(a) presents a noise included in the data and that of Fig.9(b) expresses an asperity in the data. The moving-average algorithm calculates an average of surroundings of the concentration $f(i, j)$ in the input image, and converts to the concentration $g(i, j)$ in the output image.

$$g(i, j) = \frac{1}{n^2} \sum_{k=-[n/2]}^{n/2} \sum_{l=-[l/2]}^{n/2} f(i + k, j + l) \tag{1}$$

The moving-average algorithm feathers the edge and the boundary of the input image, consequently smoothes the object surfaces.

3.2.2 Median-average algorithm

Figure 10 indicates the application of the median-average algorithm for 3×3 cubes (Boyle & Thomas (1988); Tamura (2002)). The cases of Fig.10(a) and (b) represent a noise and an asperity in the data, respectively. This algorithm sorts both a concentration value $f(i, j)$ at a target position of the input image and its surroundings in ascending order, subsequently converts the output image $g(i, j)$ using their median value. As for Fig.10(a), the concentration values in input image are rearranged like "1, 1, 1, 1, 1, 1, 1, 1, 20" in ascending order, and then the fifth value "1" is picked up as a median value and used as the concentration $g(i, j)$ in the

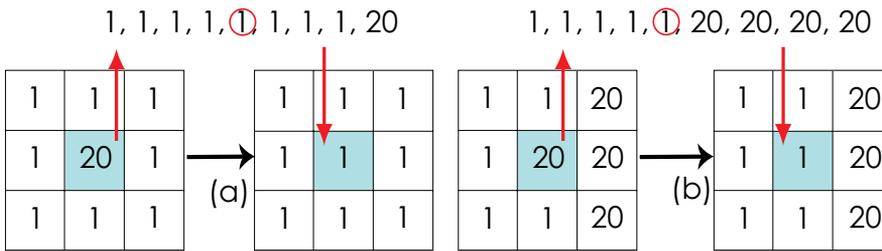


Fig. 10. Procedure of Median-Average algorithm for 3×3 cubes

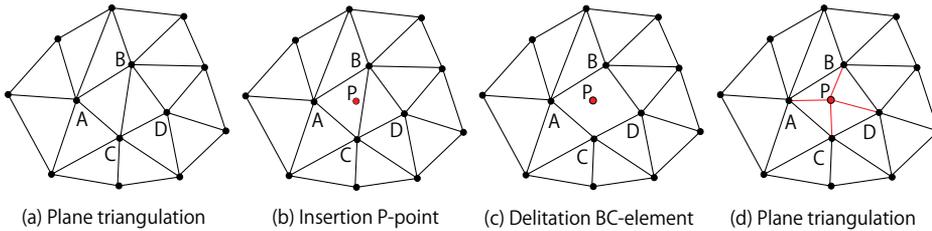


Fig. 11. Schematic drawing of the Delaunay triangulation model

output image. On the other hand, the concentration values in input image are rearranged like "1, 1, 1, 1, 1, 20, 20, 20, 20" in ascending order as shown in Fig.10(b). Finally, the fifth value "1" is selected as the concentration $g(i, j)$ in the output image. The median-average algorithm can remove the noise effectively compared with the moving-average algorithm. At the same time, this algorithm involves a risk that the magnitude of rearranging the original CT/MRI data is larger than the moving-average algorithm, consequently could yield different three-dimensional model from the real target.

3.3 Mesh generation: Delaunay triangulation model

The Delaunay triangulation model is one of the most-used mesh generation technique. This model is applicable for a wide variety of very complicate geometries. This model mathematically ensures that the circumcircle associated with each triangle contains no other point in its interior. This definition extends naturally to three-dimensional problems. Figure 11 shows a schematic drawing of the Delaunay triangulation model (Nakahashi (1995)). The triangles as shown in Fig.11(a) contains the edges and the edge's endpoints. A new point "P" is added inside the triangle "ABC", then the mathematical theory of the Delaunay triangulation is not satisfied in the triangle "ABC" and its surrounding triangle "BCD". The model can eliminate the edge "BC" which is a common edge for both triangles "ABC" and "BCD", subsequently reconstruct triangles using the added point "P". Consequently, as shown in Fig.11(d), the model can yield four triangles "ABP", "ACP", "BDP" and "CDP". The Delaunay triangulation constructs and reconstructs triangles in target domain simultaneously and give us fine meshes for CFD analysis.

4. Theory of computational fluid dynamics

In the nasal cavity, particulate matters such as pollens, house dust as well as medicinal droplets are transported via airflow of the respiration and the nebulizer treatment. In order

to execute CFD analysis for such the transport phenomena, we have to consider multi-phase flow. This section describes the basic theory of CFD analysis for two-phase flow.

4.1 Governing equations for continuous phase

The Reynolds number of intranasal flow is a few hundred. Therefore the flow is assumed as incompressible and laminar flow.

The equation for conservation of mass, or continuity equation, can be written as follows:

$$\frac{\partial \rho_F}{\partial t} + \nabla \cdot \rho_F \mathbf{u}_F = 0 \quad (2)$$

here ρ_F and \mathbf{u}_F are density and velocity of continuous phase. Equation 2 is the general form of the mass conservation equation and is valid for incompressible as well as compressible flows. Conservation of momentum in an inertial reference frame is described by

$$\frac{\partial \rho_F \mathbf{u}_F}{\partial t} + \mathbf{u}_F \cdot \nabla \rho_F \mathbf{u}_F = -\nabla p + \mu_F \nabla^2 \mathbf{u}_F + \rho_F \mathbf{g} + \mathbf{F} \quad (3)$$

where p is the static pressure, μ_F is viscosity of the fluid, and $\rho_F \mathbf{g}$ and \mathbf{F} are gravitational body force and external body forces, respectively. \mathbf{F} contains interaction with the dispersed phase in multi-phase flow and other model-dependent source terms.

The thermal energy transport equation is meant to be used for flows which are low speed and close to constant density.

$$\frac{\partial \rho_F c_p T_F}{\partial t} + \mathbf{u}_F \cdot \nabla \rho_F c_p T_F = \nabla \cdot (\lambda_F \nabla T_F) + \mathbf{S}_{FP} \quad (4)$$

where c_p and λ_F are specific heat at constant pressure and heat conductivity, respectively. \mathbf{S}_{FP} denotes a source term regarding the interaction between continuous phase and disperse phase.

4.2 Governing equations for disperse phase

Consider a discrete particle traveling in a continuous fluid medium. The forces acting on the particle which affect the particle acceleration are due to the difference in velocity between the particle and fluid, as well as to the displacement of the fluid by the particle. The equation of motion for such a particle was derived by Basset, Boussinesq and Oseen for a rotating reference frame (Akiyama (2002)):

$$m_p^i \frac{d\mathbf{u}_p}{dt} = \mathbf{F}_{\text{Drag}}^i + \mathbf{F}_{\text{Pressure}}^i + \mathbf{F}_{\text{Basset}}^i + \mathbf{F}_{\text{Buoyancy}}^i + \mathbf{F}_{\text{Rotation}}^i + \mathbf{F}_{\text{Others}}^i \quad (5)$$

where, superscript i denote the i -th particle, and \mathbf{F}_{Drag} , $\mathbf{F}_{\text{Pressure}}$, $\mathbf{F}_{\text{Basset}}$, $\mathbf{F}_{\text{Buoyancy}}$, $\mathbf{F}_{\text{Rotation}}$ express drag force acting on the particle, pressure gradient force, Basset force, buoyancy force due to gravity and forces due to domain rotation, respectively. The pressure gradient force term applies on the particle due to the pressure gradient in the fluid surrounding the particle caused by fluid acceleration. The Basset force term accounts for the deviation in flow pattern from a steady state.

The aerodynamic drag force on a particle is propotional to the slip velocity between the particle and the fluid velocity

$$\mathbf{F}_{\text{Drag}}^i = \frac{1}{8} \pi (d^i)^2 \rho_F C_D \mathbf{u}_F - \mathbf{u}_p (\mathbf{u}_F - \mathbf{u}_p^i) \quad (6)$$

where \mathbf{u}_F and \mathbf{u}_P are the velocities of the fluid flow and the particle. C_D denotes the drag coefficient of the particle.

$$C_D = \frac{24}{Re^i(1 + 0.15(Re^i)^{0.687})} \quad (7)$$

Re^i is the Reynolds number based on the slip velocity between the fluid flow and the i -th particle velocities.

$$Re^i = \frac{d^i(\mathbf{u}_F - \mathbf{u}_P^i)}{\nu_F} \quad (8)$$

The pressure gradient force results from the local fluid pressure gradient around the particle and is defined as

$$\mathbf{F}_{\text{pressure}}^i = -\frac{1}{6}\pi(d^i)^3\nabla p \quad (9)$$

This force is only important if large fluid pressure gradients exist and if the particle density is smaller than or similar to the fluid density.

$$\mathbf{F}_{\text{Basset}}^i = \frac{3}{2}(d^i)^2\sqrt{\pi\rho_F\mu_F}\int_{t_0}^t\frac{d(\mathbf{u}_F - \mathbf{u}_P^i)}{d\tau}(t - \tau)^{-\frac{1}{2}}d\tau \quad (10)$$

The buoyancy force is the force on a particle immersed in a fluid. The buoyancy force is equal to the weight of the displaced fluid and is given by

$$\mathbf{F}_{\text{Buoyancy}}^i = \frac{3}{2}(d^i)^3(\rho_P^i - \rho_F)\mathbf{g} \quad (11)$$

here \mathbf{g} expresses the gravity vector.

In rotating frame of reference, the rotation term is an intrinsic part of the acceleration in and is the sum of Coriolis and centripetal forces.

$$\mathbf{F}_{\text{rotation}} = -\frac{1}{3}(d^i)^3\rho_P(\boldsymbol{\omega} \times \mathbf{u}_P) - \frac{1}{6}\rho_P\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}_P) \quad (12)$$

where, $\boldsymbol{\omega}$ denotes rotation speed of the reference.

In the intranasal flow and heat and mass transfer, the influence of both the pressure gradient term and the Basset force term on the transport phenomena in the nasal cavity is extremely small. The momentum conservation equation for the particle is shown as

$$m_P^i\frac{d\mathbf{u}_P^i}{dt} = \frac{1}{8}\pi(d^i)^2\rho_FC_D\mathbf{u}_F - \mathbf{u}_P^i(\mathbf{u}_F - \mathbf{u}_P^i) \quad (13)$$

When the Lagrangian method is applied to the particle movement, the location of the particle is given by

$$\frac{d\mathbf{x}_P^i}{dt} = \mathbf{u}_P^i \quad (14)$$

As for the almost all particulate matter transported in nasal cavity, the mass fraction of the disperse phase are extremely small compared with the continuous phase. Therefore it is only necessary to consider the one-way interaction between both phases, namely only the continuous phase influences the disperse phase.

Heat transfer concerning the disperse phase is governed by three physical processes, namely convective heat transfer, latent heat transfer associated with mass transfer, and radiative heat transfer. Since the heat transfer of intranasal flow is not so high temperature condition, we

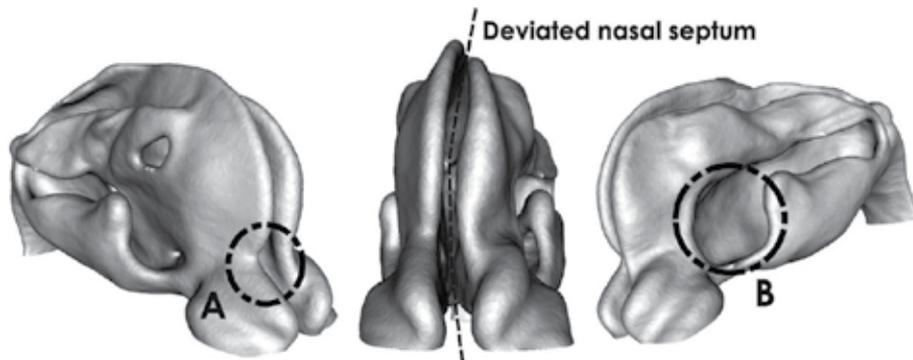


Fig. 12. Three-dimensional nasal cavity model (case-1)

Case No.	Deviated septum	nasal	Chronic sinusitis	hypertrophic rhinitis	Bloating inferior concha
1	○		Left	Left	
2			Left & Right		
3	○			Left & Right	Left
4	○			Left & Right	
5	○			Left & Right	Left

Table 1. Detail patient’s case data

can neglect both the latent heat transfer and the radiative heat transfer. The convective heat transfer Q_C is given by

$$Q_C^i = \pi d^i \lambda_F Nu (T_F - T_P^i) \tag{15}$$

where T_P^i is the temperature of the i -th particle, and Nu is the Nusselt number given by

$$Nu = 2 + 0.6Re^{0.5} \left(\mu_F \frac{C_p}{\lambda_F} \right)^{\frac{1}{3}} \tag{16}$$

The convective heat transfer Q_C is calculated and used in the source term of the thermal energy transport equation, Eq.4.

5. Numerical simulations of intranasal transport phenomena

5.1 Case data and three-dimensional model

Three-dimensional models of nasal cavity were constructed by means of five actual patient’s CT data. These cases suffer several grades of deviated nasal septum symptom, hypertrophic rhinitis, chronic sinusitis and inferior nasal concha swelling. The case data are summarized in Table 1. Since head CT data involves not only nasal cavity but also skin, fat, bone and purulent matter, it is required to eliminate them. This study used three-dimensional medical image processing software "Mimics (Materialize Inc.)" and the marching cubes algorithm to determine the surfaces of nasal wall. The moving-average algorithm is also applied to smoothing the surface. Figure 12 shows the three-dimensional nasal model of case-I constructed in this study. The anterior and the posterior of the figure show anterior nares and upper pharynx, respectively. The three-dimension model has curved nasal septum which corresponds to the curvature as shown in Fig.12. The cross-sectional areas of nasal cavities

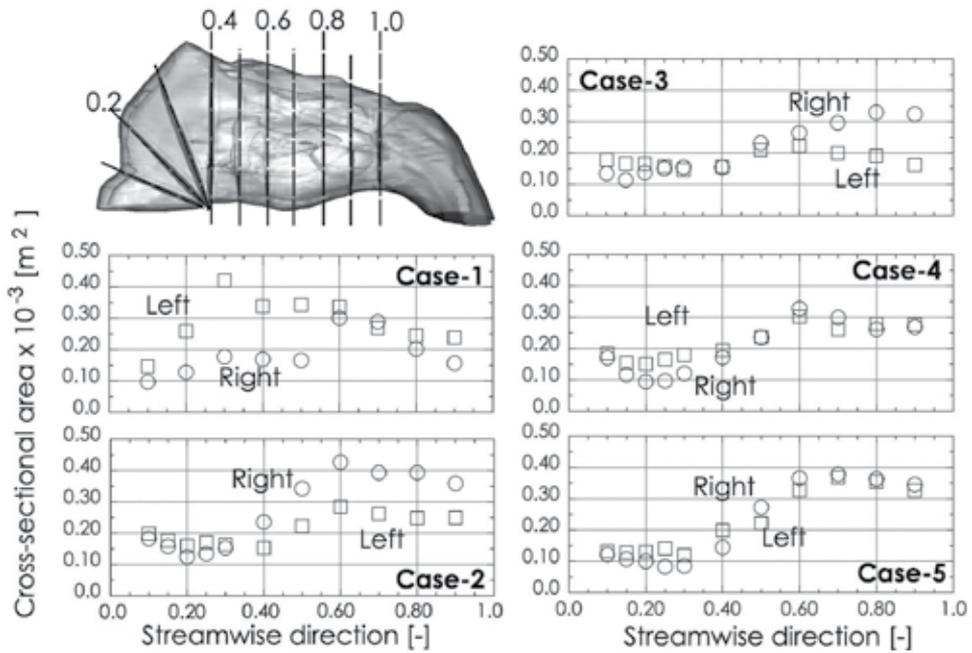


Fig. 13. Summaries of cross-sectional areas of nasal cavities for all cases

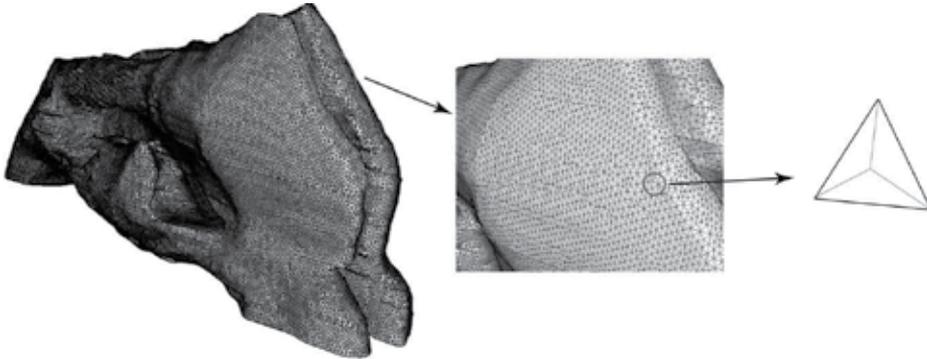


Fig. 14. Schematic drawing of CFD analysis grid (unstructured grid)

which significantly affect intranasal flow are summarized in Fig.13. The lumen and the wall of three-dimensional models are meshed using commercial meshing software ICFM CFD (ANSYS, Inc.) and the delaunay triangulation algorithm as shown in Fig.14.

Figure 15 indicates a example of the influence of the mesh quality on the accuracy of CFD analysis. In this figure, pressure drop between inlet and outlet and the number of grid nodes are used as the factors of the accuracy and the mesh quality, respectively. The pressure drop significantly depends on the number of grid nodes; the pressure drop shows occasional ups and downs in the case of small number of grid nodes, on the contrary it becomes stable in the case of large number of grid nodes.

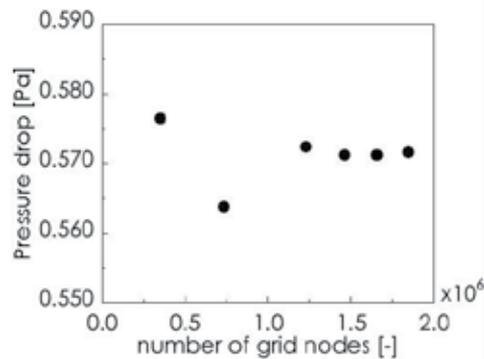


Fig. 15. Dependency between pressure drop and number of grid nodes in the CFD analysis for case-I

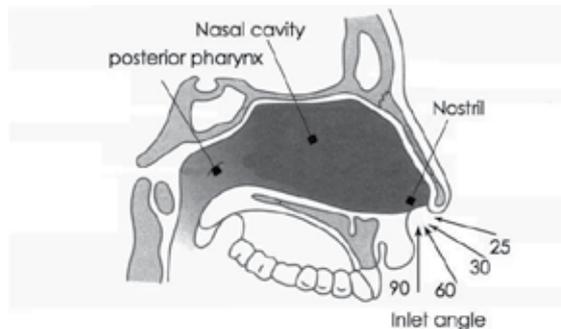


Fig. 16. Schematic figure of inlet nebulizer angles; 90, 60 and 30 degree

5.2 Computational fluid dynamics model

This study used CFD analysis code "CFX-11 (ANSYS, Inc.)" to analyze the intranasal transport phenomena of the nebulizer treatment. This code is able to solve partial differential equations for mass, momentum and energy transportations with appropriate boundary conditions. Since the Reynolds number of intranasal flow was several hundred, normal laminar flow analysis was used in this study. The Lagrange approach is adopted to solve transportation of the droplets of medicinal mist. Then the model employs assumptions that a droplet is sensitive with fluid flow, accelerating by velocity difference between fluid flow and the droplet. Air inlet velocity is set at 0.5m/s. This velocity condition corresponds to the condition of actual nebulizer treatment. As shown in Fig.16, four air inlet angles are considered in this study; 90, 60, 30 and 25 degree. In the nebulizer treatment, medicinal mist is atomized and transported by dry air. This study considers dry air (298K, 0.1MPa) as a fluid medium. No-slip condition, where fluid velocity on the wall surface is set at zero, is adopted to the nasal wall. Droplet diameter of medicinal mist is uniformly 50 micro meters. This value also corresponds to the condition of actual nebulizer treatment.

5.3 Results and discussion

Figures 17-19 shows stream lines in case-I left nasal when inlet velocity is 0.5m/s and nebulizer angles are 90, 60 and 30 degree. As we can see from the stream lines, changing nebulizer angle 60 to 30 degree changes the main flow of nebulizer from the superior nasal concha (upper-area of nasal cavity) to the inferior nasal concha (lower-area). In the all

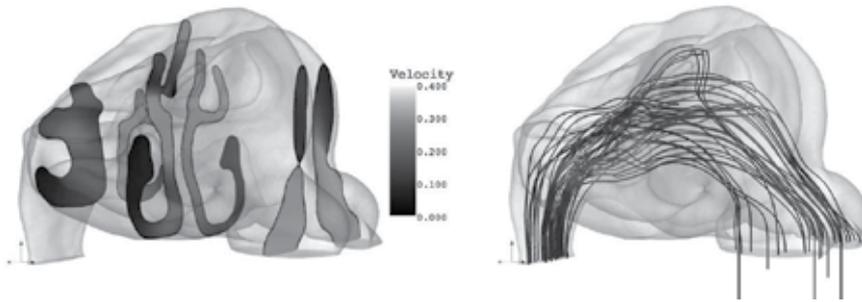


Fig. 17. Results of intranasal flow of the case-I; velocity magnitude distribution in each cross-section and streamlines (inflow velocity of 0.5m/s and inlet angle of 90 degree)

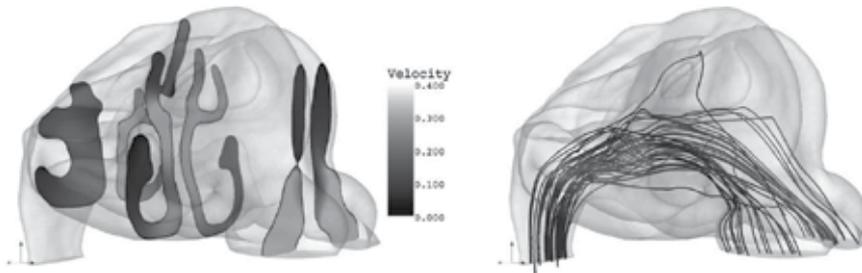


Fig. 18. Results of intranasal flow of the case-I; velocity magnitude distribution in each cross-section and streamlines (inflow velocity of 0.5m/s and inlet angle of 60 degree)

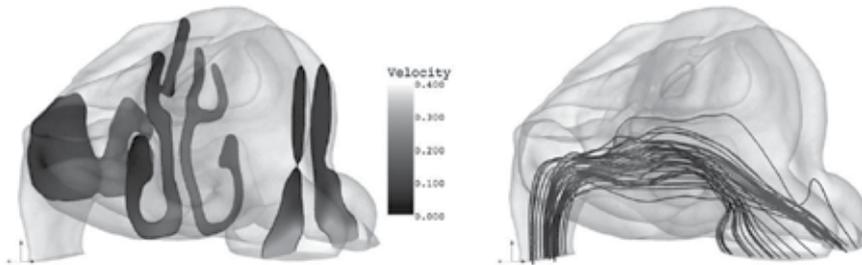


Fig. 19. Results of intranasal flow of the case-I; velocity magnitude distribution in each cross-section and streamlines (inflow velocity of 0.5m/s and inlet angle of 30 degree)

nebulizer angle conditions flow velocity increases near the nares. Especially, the increasing flow velocity is most obvious in the inflammation regions. Circulation flow is found in the lower inlet angle conditions. This circulation flow will be able to extend residence times of airflow and medicinal droplet in nasal cavity, consequently enhance the effect of the nebulizer treatment.

Figure 20-22 shows trajectories of medicinal droplets in the case-III when inlet velocity is 0.5m/s and inlet angles are 90, 60 and 30 degree. Almost all droplets impinge and deposit near the nares and middle nasal concha where the cross-section areas are narrow caused by deviated nasal septum and inflammation. From these results, the nebulizer angle is able to control the impingement and deposition characteristics of medicinal droplets. By the control

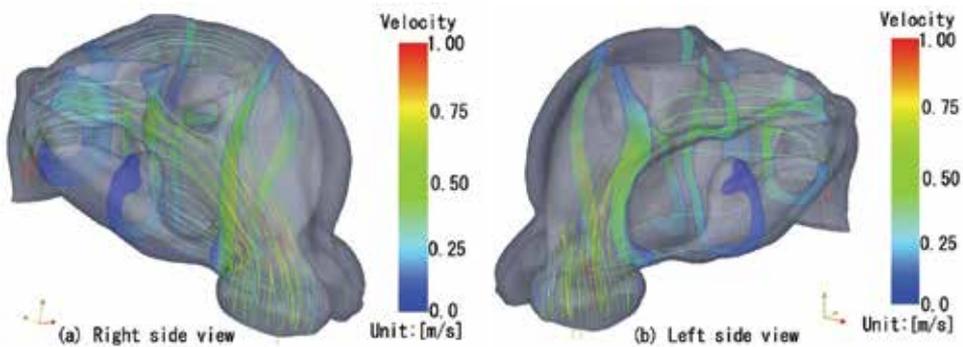


Fig. 20. Trajectories of nebulizer droplets of the case-III; inflow velocity of 0.5m/s and inlet angle of 90 degree)

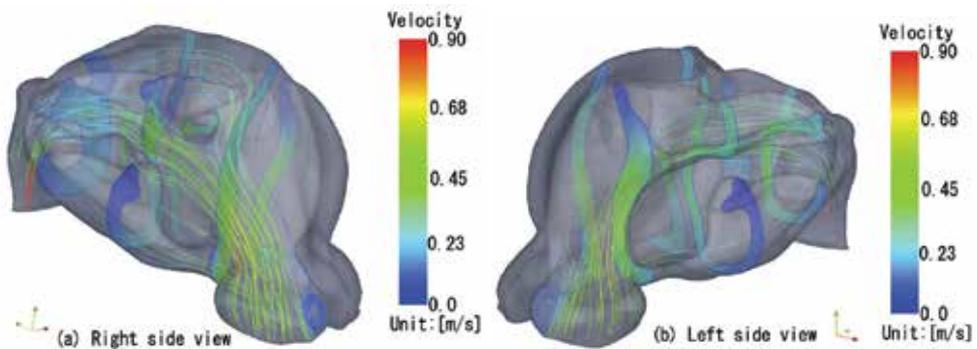


Fig. 21. Trajectories of nebulizer droplets of the case-III; inflow velocity of 0.5m/s and inlet angle of 60 degree)

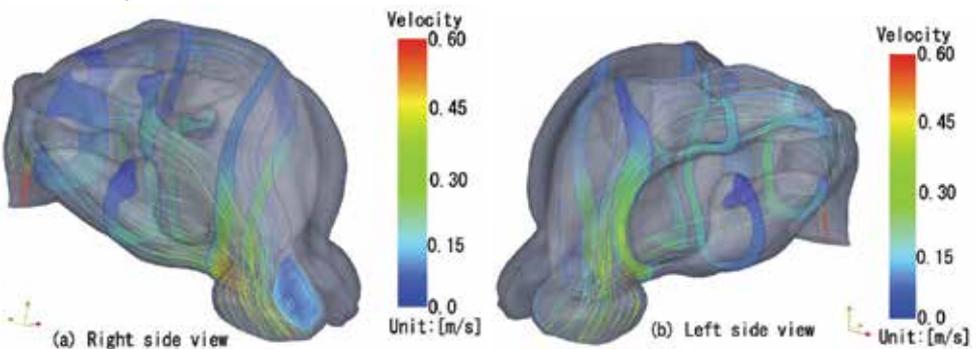


Fig. 22. Trajectories of nebulizer droplets of the case-III; inflow velocity of 0.5m/s and inlet angle of 30 degree)

of nebulizer angle, we can reduce the amount of medicinal droplet's deposition on nares and middle nasal concha, consequently transport wide area of nasal cavity.

This study focuses and uses three parameters to organize the results of each case; pressure drop of nebulizer flow, maximum velocity of intranasal flow and deposition ratio which expresses how many droplets impinge and deposit on nasal wall. Figures 23-25 are correlations among them for all patient's cases. The correlations of each case show similar

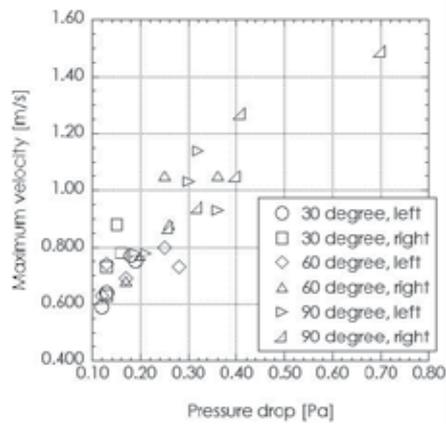


Fig. 23. Correlation analysis between maximum velocity and pressure drop

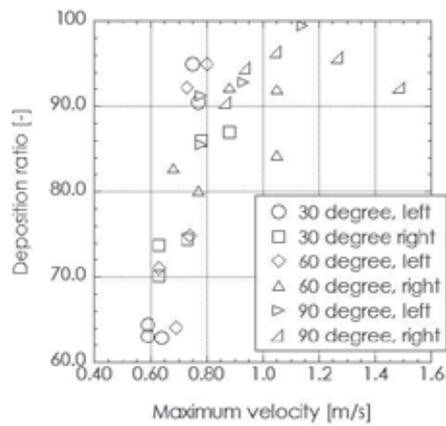


Fig. 24. Correlation analysis between deposition ratio and maximum velocity

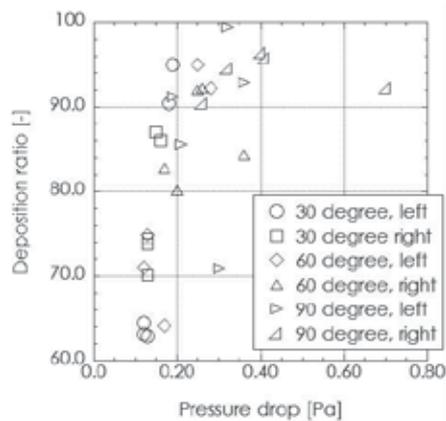


Fig. 25. Correlation analysis between deposition ratio and pressure drop

tendencies even though there are the individual differences such as the shape of nasal cavity and the state of nasal disease. Therefore these parameters are useful to organize the results mentioned above. The maximum velocity has direct proportional relationship with pressure drop. Meanwhile pressure drop and maximum velocity have strong correlation with deposition ratio. These results lead us to conclude that the nebulizer conditions have same-level influence on the intranasal transport phenomena for all cases.

6. Conclusion

This chapter presents not only the procedure of CFD analysis for intranasal transport phenomena but also medicinal droplets transport characteristic in patient's nasal cavity using actual CT/MRI data. Furthermore, the result of correlation analysis for the nebulizer treatment condition is also presented here. As for the results of CFD analysis, nebulizer angle, inlet velocity and size of droplet significantly affect on the intranasal transport phenomena. The condition of low nebulizer angle achieves that the medicinal droplets are transported all over the nasal cavity, that is this condition will increase therapeutic response to the nasal diseases. The correlation among velocity of intranasal flow, pressure drop and deposition ratio, which are main parameters for intranasal transportation of medicinal droplets in the nebulizer treatment, indicates that the basis of the intranasal transportation show similar characteristics in all patient's cases. These results mean that there is an optimum condition of nebulizer treatment for all patients who suffer nasal diseases. Further research of the intranasal transport phenomena for many patient's cases is required to survey the best treatment condition.

Since the development of the medical image processing technique has been in progress, the CFD analysis for human body, such as blood flow from the heart, respiration flow in the pharynx and the lung, will be widely spread and achieve tangible results in clinical practices within the next several years.

7. Acknowledgement

This work was supported by the grant from JSPS KAKENHI 19760114.

8. References

- Akiyama, M. (2002). *Analysis of Two-Phase Flow Dynamics: Multi-Physics Flow Analysis*, Korona Inc., Tokyo.
- Boyle, R. & Thomas, R. (1988). *Computer Vision: A First Course*, Blackwell Scientific Publications, New York.
- Haruna, S. (2003). *An endoscope surgical operation of chronic sinus infection*, Hokendojinsha Inc., Tokyo.
- Lindemann, J., Keck, T., Wiesmiller, K., Sander, B., Brambs, H.-J., Rettinger, G. & Pless, D. (2004). A numerical simulation of intranasal air temperature during inspiration, *The Laryngoscope* Vol.114(No.1): 1037–1041.
- Lindemann, J., Keck, T., Wiesmiller, K., Sander, B., Brambs, H.-J., Rettinger, G. & Pless, D. (2005). Numerical simulation of intranasal airflow after radical sinus surgery, *American Journal of Otolaryngology* Vol.26(No.1): 175–180.
- Lorenson, W. E. & Cline, H. E. (1987). Marching cubes:, *Computer Graphics: A high Resolution 3D Surface Construction Algorithm* Vol.21(No.4): 163–169.

- Monya, M., Yamamoto, T., Nakata, S., Nakashima, T., Yamamoto, T. & Suzuki, T. (2009). Cfd analysis for intranasal mass transportation in deviated septum case (japanese), *JSME Journal Series B* Vol.75(No.751): 175–180.
- Nakahashi, K. (1995). *Grid Generation and Computer Graphics*, Univ. of Tokyo, Tokyo.
- Savitzky, A. & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures, *Anal. Chem.* Vol.36(No.8): 1627–1639.
- Sun, C. (2006). Moving average algorithms for diamond, hexagon, and general polygonal shaped window operation, *Pattern Recognition Letters* Vol.27(No.6): 1627–1639.
- Tamura, H. (2002). *Computer Image Processing*, Ohm Inc., Tokyo.
- Weinhold, I. & Mlynski, G. (2004). Numerical simulation of air flow in the human nose, *Eur Arch Otorhinolaryngol* Vol.261(No.1): 452–455.
- Yamamoto, T., Nakata, S., Nakashima, T. & Yamamoto, T. (2009). Computational fluid dynamics simulation for intranasal flow (japanese), *Foresight of Otorhinolaryngology* Vol.52(No.1): 24–29.

Part 5

Additional Important Themes

Fluid-Dynamic Characterization and Efficiency Analysis in Plastic Separation of the Hydraulic Separator Multidune

Floriana La Marca, Monica Moroni and Antonio Cenedese
*DICEA - Sapienza University of Rome, Rome
Italy*

1. Introduction

Recovery of useable plastics from post-consumer and manufacturing waste remains a major recycling challenge. The global consumption of plastics was reported to be 230 million tonnes in 2005 (PlasticsEurope 2007a) of which 47.5 million tonnes were produced in Europe (25 European Union countries + Norway and Switzerland). Of the European production, only 22 million tonnes were reported as having been collected. Of this collected waste, 4 million tonnes were recycled as a manufacturing feedstock (18%) and 6.4 million tonnes went into energy recovery (29%), with the balance (11.6 million tonnes) probably being disposed in landfills (PlasticsEurope 2007b).

The recycling of plastics is a process essential to reduce the efflux of materials to landfills and to decrease the production of raw materials. In recent years awareness of the importance of environmental protection has led to the development of different techniques for plastic recycling. One issue related to the recycling of this material is the presence in the market of many types of plastics (polymers with additives), often with similar characteristics that make them difficult to differentiate in the recovery phase.

The separator "Multidune" is a hydraulic separator by density. Its name derives from the characteristic undulate profile of the channel where separation occurs. The channel is constructed from a sequence of closed parallel cylindrical tubes welded together in plane which are then sliced down the lateral mid-plane and the lower complex is laterally shifted relative to the upper complex. The Multidune allows solid particle separation according to their specific weight and the velocity field establishing within the apparatus.

Previous investigations (De Sena et al., 2008; Moroni et al. 2008) suggested the flow within the Multidune is organized into three main patterns. Principally, a longitudinal transport flow takes place, where the velocity is high. A particle belonging to this region can move from one camera to another. The second region is the lower recirculation zone with high values of the vorticity field. Particles belonging to this region undergo the vertical impulse of the fluid. The thrust is proportional to the vertical velocity component and, in conjunction with gravity and buoyancy, determines the destiny of a particle. If the thrust is larger than the net weight of the particle, an interaction with the principal transport flow occurs and, consequently, the particle will move to the following chamber. The third region is the upper recirculation zone whose dimensions are smaller than the other recirculation zone. If a

particle moves from the principal flow to the secondary vorticity zone, it will have the chance to come back to the previous chamber, assuming the principal transport flow thrust does not prevent it from falling out.

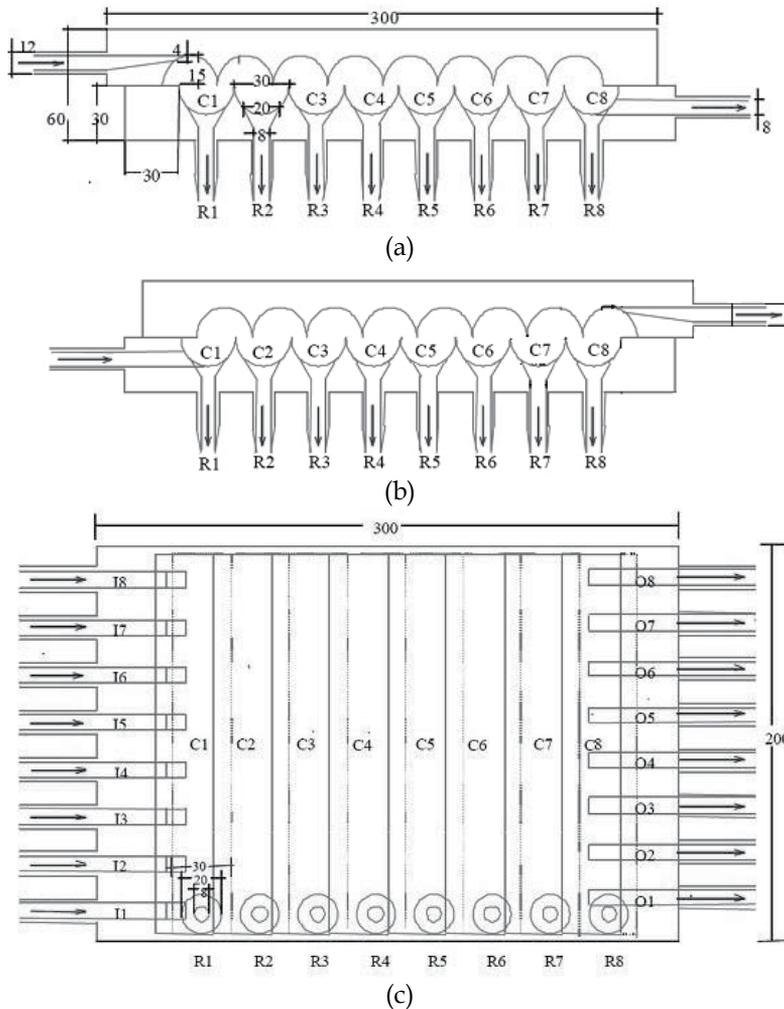


Fig. 1. Longitudinal section (dimensions in mm) for Arrangement (a) A and (b) B; (c) top view of the apparatus. C indicates the 8 chambers, R the 8 collection nozzles, I the 8 inlet nozzles and O the 8 outlet nozzles. Dimensions for Arrangement B are the same as for Arrangement A

Because of the role played by velocity, the fluid dynamic investigation of the Multidune apparatus is a preliminary step to carry out in order to investigate its capability in separating solid particles. For this reason a novel experimental campaign was set up and image analysis was employed to detect the velocity field within the apparatus. Two different arrangements of the Multidune were employed, hereinafter Arrangement A and

Arrangement B. Figure 1 presents a sketch of the longitudinal sections for the two arrangements, the top view and the apparatus dimensions in mm.

Two sets of experiments were run. The camera imaged all chambers of Arrangement A and B respectively and a passive tracer, i.e., a neutrally buoyant seeding particulate, was used to seed the fluid and reconstruct the fluid velocity field.

The fluid-dynamic investigation deals with the detection of several quantities of interest:

- mean velocity field and velocity profiles along the apparatus
- two-dimensional turbulent kinetic energy
- vorticity field

Generally, the detection of vortical structure (or coherent structure) immersed within a turbulent flow field is not straightforward. This book chapter will deal with this issue as well. Starting from the velocity gradient tensor, a topological analysis will be carried out with the aim of determining the critical points within the domain and to better represent areas with elevated values of the vorticity field.

The fluid-dynamic characterization will be correlated with the results of separation tests, carried out at lab-scale to separate particles of different plastic materials, characterized by low values in specific weight (about 1 g/cm³, close to water specific weight).

2. Topological description of a fluid-dynamic field

It is usually difficult to individuate vortices, or coherent structures, within a statistical field which represents the background turbulent phenomenon. The detection of those structures is strictly related to their definition and it is influenced by the fact that the flow field can not be monitored at all scales. If an experimental monitoring technique is employed, a further limit is related to the availability of point, mono-dimensional, or two-dimensional measures.

Among the principal methods employed to detect vortices, visualization is the first and the currently most employed. The fluid motion can be visualized through the injection of dye or tracer particles. Remarkably, the surface where vorticity assumes a constant value does not necessarily match the surfaces where the dye concentration is constant. Even if the Schmidt number ($Sc = \nu/D$, where ν is the kinematic viscosity and D the dye molecular diffusivity) is unitary, unlike the scalar behavior, vortices are subject to vortex stretching. Tracers would then present lower concentrations in areas of larger vorticity, and will not highlight fundamental behaviors of vortical dynamics. This inconvenient can be avoided seeding the fluid locally and detecting its behavior at distance and/or time intervals small respect to the characteristic scales (Joeng and Hussain, 1995).

Although the concept of vorticity is one of the most widely used in fluid mechanics, even today there is no generally accepted definition of vortex. A vortex is usually defined as the region of space enclosed by a surface formed by swirling lines (tangent at every point of the velocity curl). This definition is ambiguous as it would lead to identify a vortex within a laminar channel or in a Poiseuille flow.

The conditions necessary for a proper definition of the vortex are (Hussain, 1986):

- inside the vortex, the vorticity ($\nabla \times \mathbf{u}$) must assume non-zero values, it is a necessary but not sufficient condition;
- the existence of the vortex must be identified by a scalar quantity;
- identification criteria must be invariant under Galilean transformations.

Since the concept of vortex is associated with rotation of the fluid, some authors have proposed to identify a vortex where the pressure reaches the minimum required to balance the centripetal force. This definition falls in some situations, for example, in a convergent a minimum of pressure exists unless the fluid is characterized by a rotating motion.

Similarly, the use of trajectories and streamlines to identify a vortex does not appear appropriate since both are not invariant with respect to a Galilean transformation.

Some criteria for identifying vortical structures that meet the requirements above use scalar quantities derived from the velocity gradient tensor, $\nabla \mathbf{u}$ or A_{ij} which is decomposed in its symmetric S_{ij} and antisymmetric W_{ij} parts. In particular, the eigenvalues of the velocity gradients are examined. In particular, the three criteria are based on:

- identification of regions where eigenvalues of $\nabla \mathbf{u}$ are complex numbers
- identification of regions where the second eigenvalue of tensor $S_{ik}S+W_{ik}W_{ik}$ is negative (eigenvalues sorted in descending order)
- identification of the regions where the second invariant of the velocity gradient tensor is positive.

Those three criteria are equivalent in a two-dimensional case, as the case under investigation.

3. Experimental set-up

The Multidune apparatus, $0.30 \times 0.20 \times 0.06 \text{ m}^3$ in size, is composed by a sequence of closed parallel cylindrical tubes welded together in plane. The device is sliced down its lateral mid-plane and the lower half is shifted laterally and then fixed relative to the upper half. Arrangements A and B (Figure 1) are shaped by two different positions of the lower half. In Arrangement A, the lower half is shifted 0.012 m right, while it is shifted 0.012 m left for Arrangement B. Each resulting chamber is labelled according to its position along the flow direction [C1 (first chamber), C2, ..., C8 (last chamber)]. The first chamber has eight round input nozzles (I1, I2, ..., I8) held at constant head ranging between 0.84 m and 2.69 m. The last chamber has 8 round output nozzles (O1, O2, ..., O8). Flow is induced in the lateral direction normal to the axis of the tubes by cutting inlet nozzles on one side of the device and outlet ones on the opposite side. Settled materials may be collected in the lower part of each half-cylinder.

The Multidune feeding occurs through a tank whose output pipe is split into 8 to distribute water within the apparatus through the 8 inlet nozzles. An overflow exit allows controlling the water level in the tank. The average flowrate within the Multidune apparatus will rely both on hydraulic head at the inlet nozzles and on the number of open outlet nozzles. With both arrangements, the experiments were run for three elevations of the tank (referred to the inlet nozzles height): 0.84 m, 1.84 m, 2.69 m. Three outlet nozzles (O2, O4 and O6) were open when Arrangement A was adopted. Table 1 reports the difference in fluid elevation between the tank and the middle of inlet nozzles and the corresponding flowrates. Preliminary tests conducted with Arrangement B suggested the flowrate determined by opening three outlet nozzles was too large to allow plastic particles separation. Then experiments have been run with only one opened outlet nozzle (O3). In this case, the flowrates establishing within the apparatus are consistently lower.

Experimental results of separation of both mono- and multi-material samples (Moroni et al. 2011; La Marca et al. 2011) suggest the behaviour of particles introduced within the Multidune apparatus strongly depends on the characteristic velocity within the apparatus,

dimension and density of plastic particles. The geometry of the channel and the type of working fluid are considered to play a key role in the process of separation. We then introduce the non-dimensional parameters:

	Difference in fluid elevation between tank and Multidune	Average flowrate for Arrangement A	Re for Arrangement A	Average flowrate for Arrangement B	Re for Arrangement B
	(m)	(l/min)	-	(l/min)	-
Q1	0.84	9.35	584.38	4.57	380.58
Q3	1.84	11.82	738.75	5.00	416.83
Q5	2.69	13.53	845.63	6.55	545.58

Table 1. Difference in fluid elevation between the tank and the middle of the Multidune inlet nozzles and corresponding flowrates for both Arrangements.

- particle Froude number

$$\theta = \frac{V^2}{\frac{\rho_s - \rho}{\rho} g d} = \frac{V^2}{g' d} \quad (1)$$

where:

V: working fluid (water in the present case) characteristic velocity

ρ_s : particle density

ρ : water density

g: acceleration of gravity

d: particle diameter

g' : buoyancy parameter

- Reynolds number

$$Re = \frac{V H}{\nu} \quad (2)$$

where:

H: apparatus characteristic dimension, set equal to the cylinder radius, i.e. 1.5 cm

ν : working fluid kinematic viscosity

Re for the three tank heights and both arrangements are reported in Table 1.

The following procedure was adopted to carry out experimental tests aimed at reconstructing the velocity field and separating the plastic materials:

1. set up of water tank height and hydraulic head;
2. water feeding in the apparatus through the 8 input nozzles;
3. saturation of the apparatus with water;
4. opening of the chosen output nozzles (O2, O4 and O6, in Arrangement A; O3 in Arrangement B);
5. passive tracer or sample feeding in the Multidune apparatus through the I3 e I4 input nozzles;
6. test execution, about three minutes;

7. only for separation tests, recovery of material expelled from the output nozzle(s);
8. closing of the output nozzle(s);
9. only for separation tests, recovery of material settled in each chamber;
10. weighting of recovered materials after 24-h drying.

4. Fluid-dynamic characterization by image analysis

The aim of the fluid-dynamic characterization of the Multidune apparatus was to reconstruct trajectories of tracer particles seeding the fluid under investigation and the velocity field evolution with time by means of image analysis techniques.

The need for measuring velocity fields has historically led researchers to develop experimental techniques and related instrumentation. An 'ideal' measurement system should be non intrusive to avoid flow field perturbations, should not require calibration, and should be suitable for obtaining the velocity field with a time and space resolution smaller than the characteristic time and length scales (i.e. Kolmogorov scale for turbulent flows). Velocity measurements based on optical methods capable of providing the velocity of tracer particles illuminated by a light source represent the best approximation of this 'ideal' system. The requirement is the working fluid to be seeded with neutrally buoyant particles which are assumed to follow the flow. The velocity vector is evaluated from the ratio of the tracer displacement, Δs , and the time interval Δt required for the displacement to take place. The time interval must be small enough for the approximation to be reasonable.

Digital images were acquired using a high-speed high-resolution (1280×1024 pixels) camera at a rate of 250 frames per second and stored for analysis. The camera axis was set perpendicular to the Multidune lateral face. A high powder lamp produces a light sheet for illuminating the interior of the channel. Green plastic powder (mean diameter of about 200 μm) preconditioned with a solution of water and sodium hydroxide was used as the tracer for particle tracking. Preconditioning was used to neutralize the electrostatic charge on the particles. Tracer injection within the apparatus and data collection were started after the fluid reached steady state.

Feature Tracking (FT) was employed as the image analysis technique, i.e. a Particle Tracking algorithm which allows ignoring the constraint of low seeding density, being able to provide accurate displacement vectors even when the number of tracer particles within each image is very large (Moroni and Cenedese, 2005). FT reconstructs the displacement field by selecting image features (image portions suitable to be tracked because they remain almost unchanged for small time intervals) and tracking these from frame to frame. The matching measure used to follow a feature (and the $L \times H$ window around the feature, where L and H are the horizontal and vertical dimensions respectively) and its "most similar" region at the successive times is the "Sum of Squared Differences" (SSD) among intensity values: the displacement is defined as the one that minimizes the SSD. In Feature Tracking one applies the algorithm only to points where the solution for the displacement exists: those points are called "good features to track" (Shi and Tomasi, 1994). FT allows a Lagrangian description of the velocity field providing sparse velocity vectors with application points coincident with large luminosity intensity gradients (likely located along tracer particles boundaries). Lagrangian data are then used to reconstruct instantaneous and time-averaged Eulerian velocity fields through a resampling procedure.

Fig. 2a and Fig. 2b show the trajectories reconstructed by the FT algorithm within chambers C3 and C4 for experiments run with Arrangements A and B and difference in fluid elevation between the tank and the center of the Multidune inlet nozzles set to Q3.

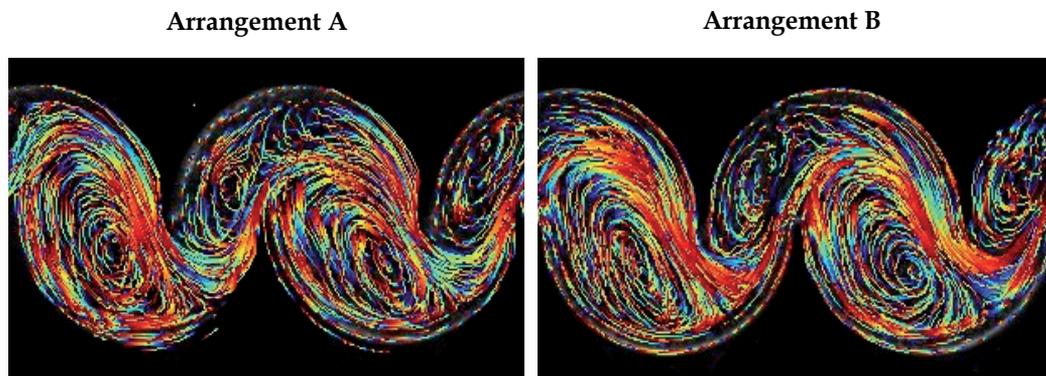


Fig. 2. Trajectories reconstructed by the Feature Tracking algorithm within chambers C3 and C4 for experiments run with Arrangement A (a) and Arrangement B (b) and difference in fluid elevation between the tank and the center of the Multidune inlet nozzles set to Q3

The trajectories are visualized overlapping 32 consecutive positions of the tracer particles. The colours range from blue (associated with the first time shown) to red (associated to the last one). The comparison between the two images evidences the increased length of the trajectories reconstructed when three nozzles are open instead of one. At the higher flowrate, the movement of tracer particles from one frame to the next is expected to be larger. The trajectories qualitatively describe the velocity field within the apparatus and confirm the presence of three sectors, i.e. the principal current which appears more consistent with Arrangement B, the lower recirculation area, larger with Arrangement A, and the upper recirculation area bigger with Arrangement B.

Figure 3 and Figure 4 present the velocity vectors overlapped to the colormap of, respectively, the horizontal and vertical velocity components within chambers C3 and C4, both arrangements and differences in fluid elevation between the tank and the center of the Multidune inlet nozzles set to Q1, Q3 and Q5.

The analysis of both velocity fields and streamlines (Figure 5) suggests the fluid-dynamic behaviour of the Multidune apparatus for both arrangements is characterized by three predominant areas.

The principal transport flow. It is characterized by a positive value of the velocity component along the x axis in the entire longitudinal section. The principal current is responsible for the transport of particles from one chamber to the next one and eventually it drives material to the outlet nozzles without separation. In Figure 5, it is indicated with red streamlines.

The lower recirculation zone. It is visible below the principal current, in each of the height chambers. The clock-wise rotating motion is suitable for subtracting particles from the principal current. It is expected that captured plastic particles will behave in one of the following ways:

- settle within the chamber if sufficiently heavy;
- follow the upward portion of the rotating motion without reaching again the principal current being too heavy to perform a complete rotation;
- execute a complete rotation and being captured again by the principal transport flow to settle in the next chamber or be expelled.

The particle behavior within the recirculation zone is influenced by its density and dimension as well as by the presence of a vortex. In Figure 5, it is indicated with green streamlines.

The upper recirculation zone. It develops behind the principal transport current. The purpose of the upper recirculation zone is to subtract particles from the principal current and to transfer them backward. To do so they have to pass across the principal current and to settle in the chamber. The particle physical attributes and the characteristic velocity of the principal current will influence the efficacy of this process. In Figure 5, it is indicated with blue streamlines.

The fluid vein characterizing the principal current is constituted by an inner part with velocity increasing with the hydraulic head at the apparatus inlet. Nevertheless, the characteristic dimension of the principal current remains practically constant even varying the flowrate. It appears more consistent with Arrangement B. At each flowrate, the upper recirculation zone developing in the Multidune- Arrangement A presents values of the both components of velocity field significantly lower than the characteristic velocity of the principal current. For this reason, this zone results ineffective in capturing particles from the principal current and then in the separation process. It was expected that reducing the flowrate and/or modifying the inner geometry of the apparatus, the upper recirculation zone would have played a role in the separation process. This is confirmed by the experiments with the Multidune- Arrangement B. Reducing the flowrate, the velocity of the principal current decreases whereas it remains practically constant within the recirculation areas becoming comparable. Furthermore, the modified geometry favours the formation of a larger upper recirculation area in each chamber incrementing the contact surface with the main current and the probability of material exchange between the two zones.

The velocity field in the lower recirculation area is in each case significantly lower than the characteristic velocity in the principal current. With the Multidune- Arrangement A, the velocity within the lower recirculation zone increases rising the flowrate, especially in the ascending portion of the vortex. This is actually counterproductive in terms of separation, because of the drag of settled particles toward the principal current. A similar behaviour with analogous consequences characterizes the descending portion of the upper recirculation area establishing in the Multidune- Arrangement B.

Noticeably, the analysis of the fluid-dynamic field developing within the Multidune apparatus with both arrangements suggests the increase of the hydraulic head augments the transport effectiveness of the main current without improving the capture feasibility of both the upper and lower recirculation zones. The apparatus will then lose its effectiveness in separating plastic particles increasing both the hydraulic head and the transiting flowrate. Furthermore, with both arrangements, C3 and C4 present analogous recirculation areas. The same circumstance occurs at each flowrate establishing in the apparatus. Furthermore, an analogous behavior characterizes C5 and C6. Then, the presence of eight chambers assures plastic particle separation even if a chamber should be filled with the settled material allowing the following chambers to become effective in the separation process.

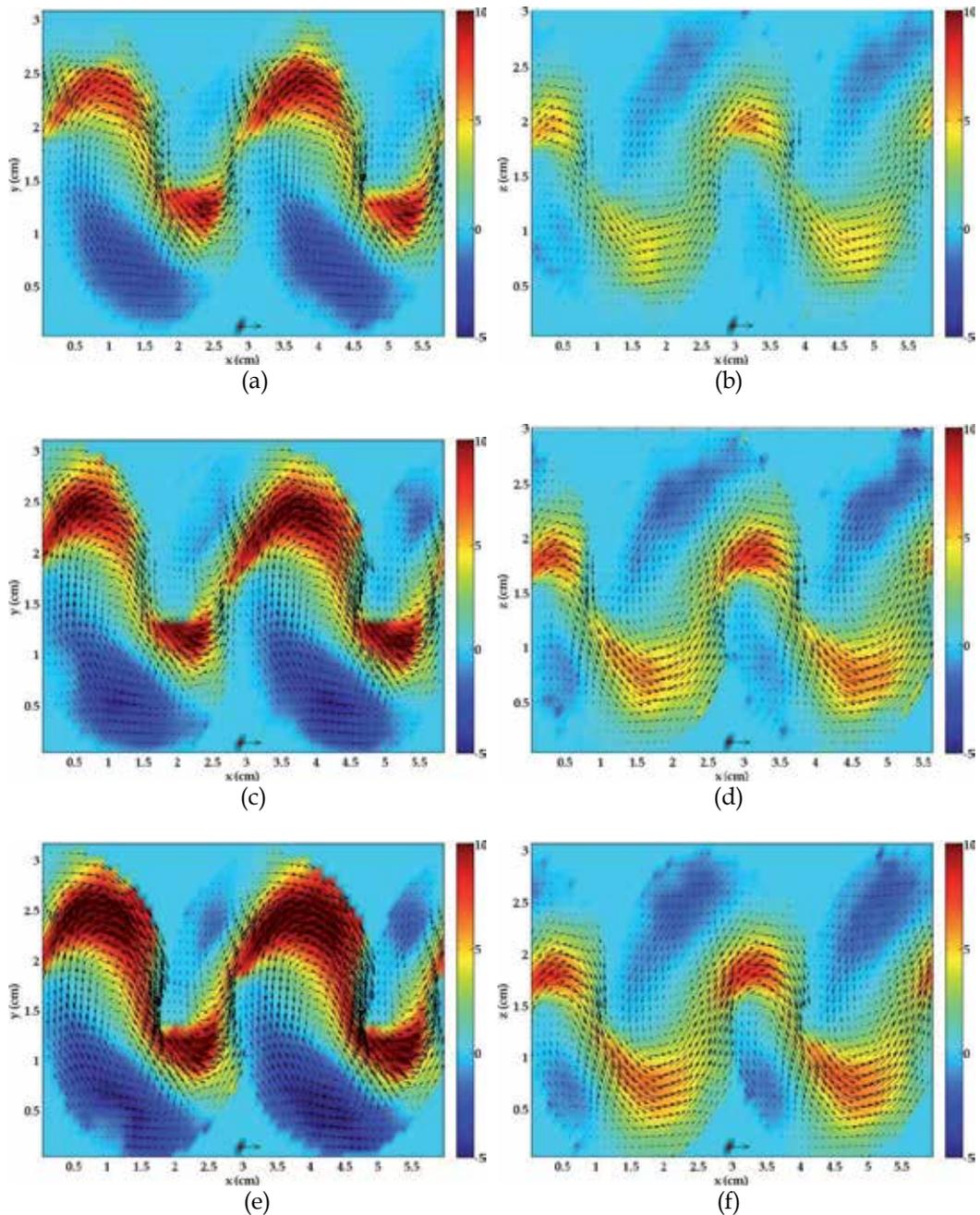


Fig. 3. Eulerian velocity field within C3 and C4 overlapped to the horizontal velocity component represented as colormap for Arrangement A at (a) Q1, (b) Q3, (c) Q5 and Arrangement B at (d) Q1, (e) Q3, (f) Q5. The reference velocity vector is equal to 10 cm/s

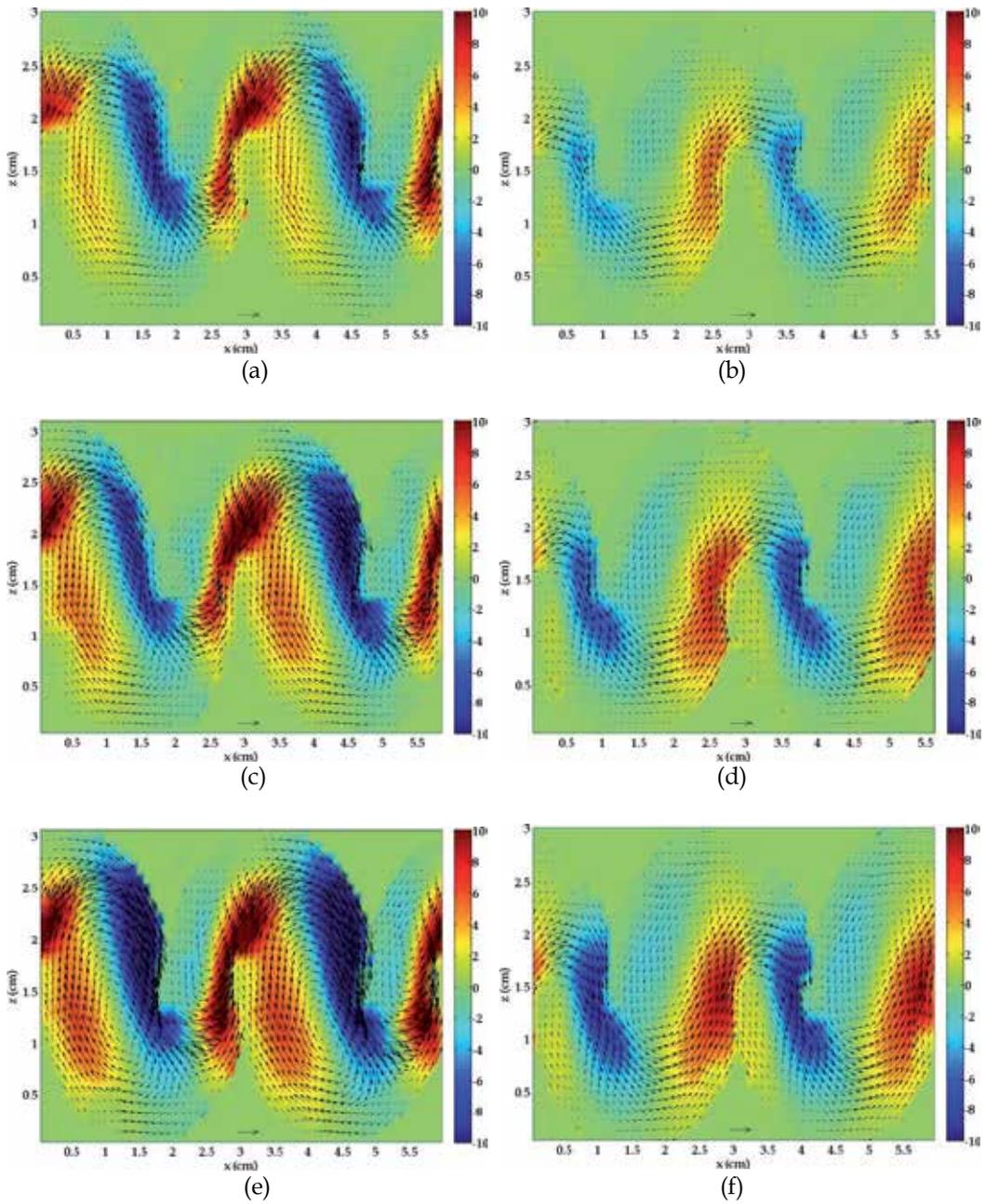


Fig. 4. Eulerian velocity field within C3 and C4 overlapped to the vertical velocity component represented as colormap for Arrangement A at (a) Q1, (b) Q3, (c) Q5 and Arrangement B at (d) Q1, (e) Q3, (f) Q5. The reference velocity vector is equal to 10 cm/s

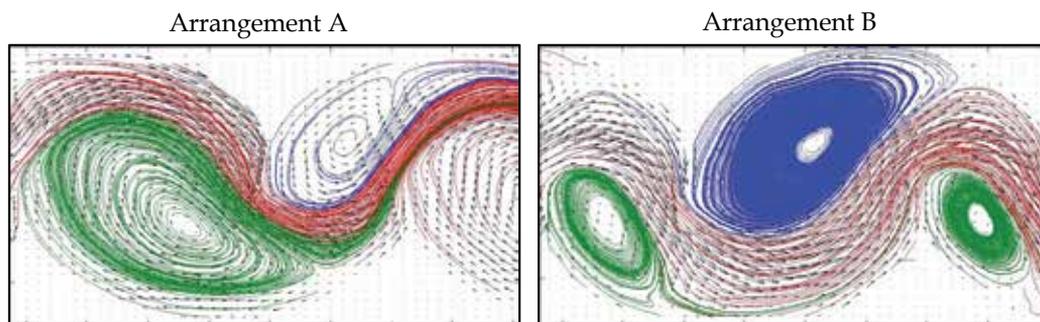


Fig. 5. Velocity field overlapped to the streamlines for Arrangement (a) A and (b) B. The difference in fluid elevation between the tank and the center of the Multidune inlet nozzles is set to Q3

Figure 6 presents the vorticity field computed for both Arrangements and for Q1, Q3 and Q5. The difficulty in identifying vortical structured by examining just $\nabla \times \mathbf{u}$ should be clear from those figures. On the other hand, Figure 7 presents the complex part of the first eigenvalue of the velocity gradient tensor for both Arrangements and the difference in fluid elevation between the tank and the center of the Multidune inlet nozzles is set to Q5. It should be noted as the areas where the first (and the second as well, not shown) presents a complex value is within recirculation areas. It is then confirmed as the criteria previously mentioned for identifying vortical structures can be used for this experimental investigation.

5. Separation efficacy

Experimental tests were executed utilising four samples of plastics, in order to investigate efficiency and capability in differentiating trajectories according to plastic typologies and fluid-dynamic properties of the Multidune with both Arrangement A and B (Figure 1). The samples were composed with particles of different plastic material and size, as described in Table 2.

The tank heights and the corresponding flowrates utilized in the experimental tests were Q1, Q3 and Q5, as reported in the previous Table 1.

	Sample material	Mean density	Particle size distribution	Weight
Sample		(g/cm ³)	(mm)	(g)
MONO.2	Brown plastic	1.416	0.85-1.00	1.5
MONO.3	Green plastic	1.353	0.85-1.00	1.5
MONO.4	Green plastic	1.353	1.70-2.00	2.5
MONO.5	Red plastic	1.143	1.70-2.00	2.5

Table 2. Composition of samples utilized in separation tests with Arrangements A and B.

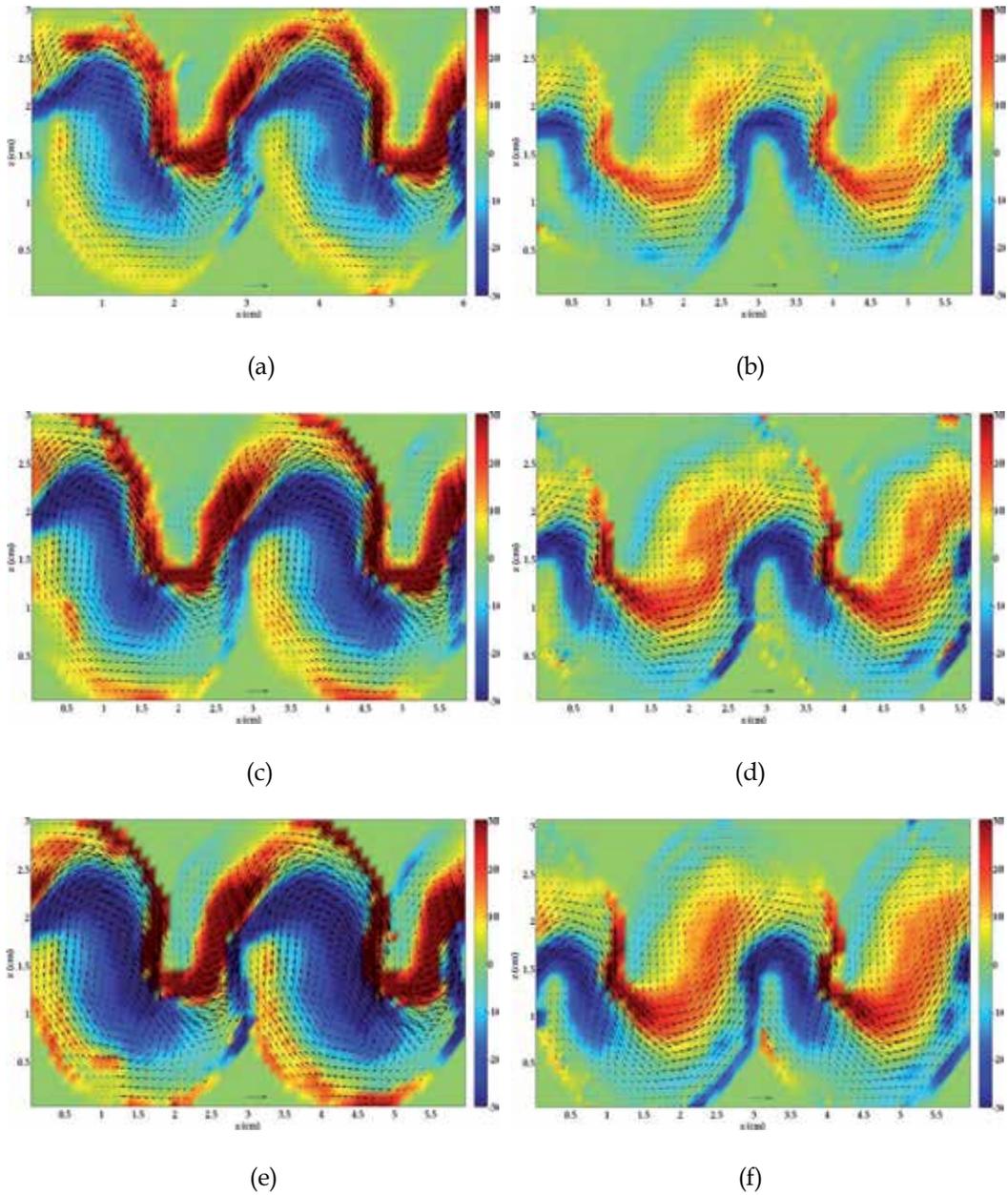


Fig. 6. Eulerian velocity field within C3 and C4 overlapped to the vorticity field represented as colormap for Arrangement A at (a) Q1, (b) Q3, (c) Q5 and Arrangement B at (d) Q1, (e) Q3, (f) Q5. The reference velocity vector is equal to 10 cm/s

The results of the tests on plastic samples suggested that the system may be able to separate different types of plastics in a mixture imposing an appropriate hydraulic head. In the following further remarks are given about the results referring to the adopted Arrangement.

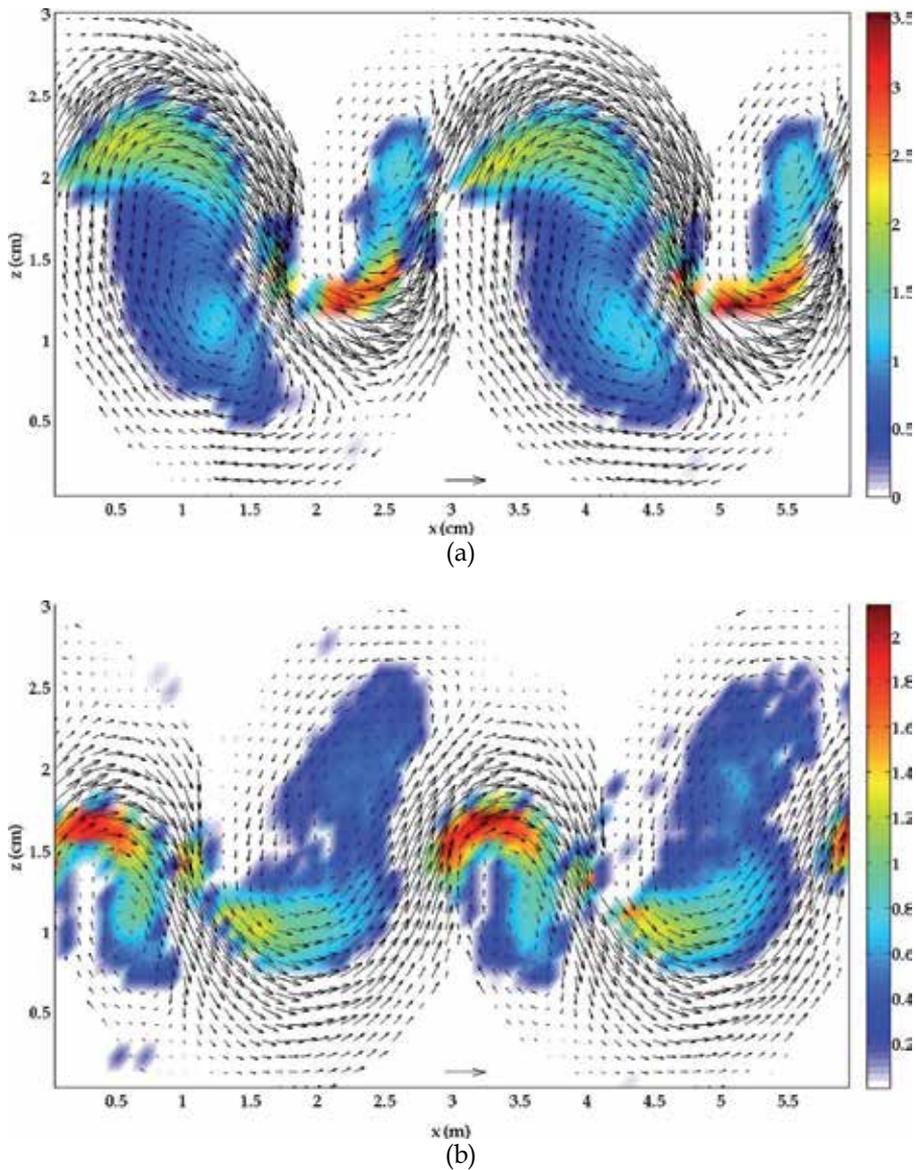
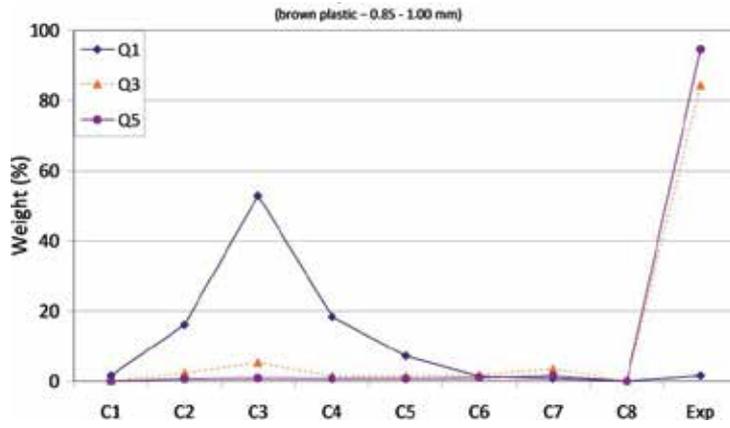


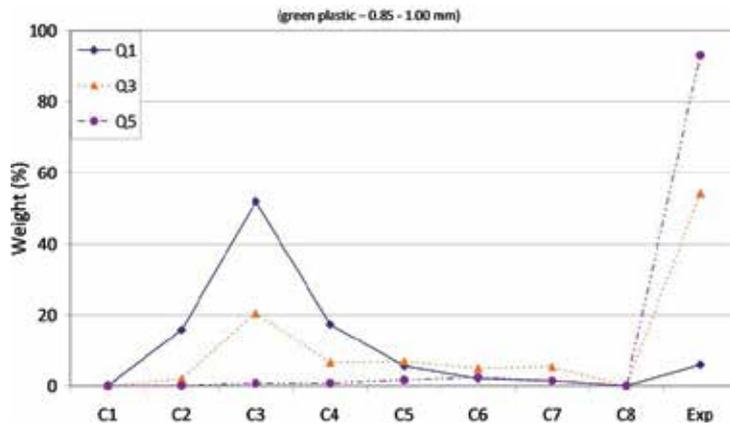
Fig. 7. Velocity field overlapped to the complex part of the first eigenvalue of the velocity gradient tensor for Arrangement (a) A and (b) B. The difference in fluid elevation between the tank and the center of the Multidune inlet nozzles is set to Q_5

5.1 Results with arrangement A

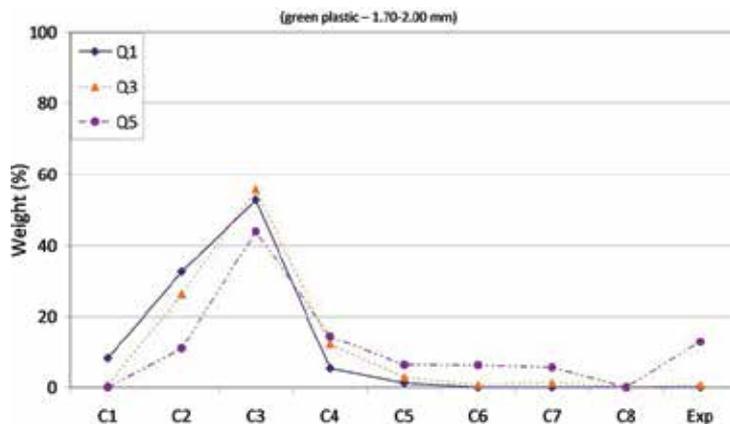
Figure 8 shows the results of the experimental tests on plastic samples at each adopted flowrate, reporting the percentage in weight of the materials recovered from each chamber (C1-C8) and expelled through the output nozzles (Exp).



Sample MONO.2



Sample MONO.3



Sample MONO.4

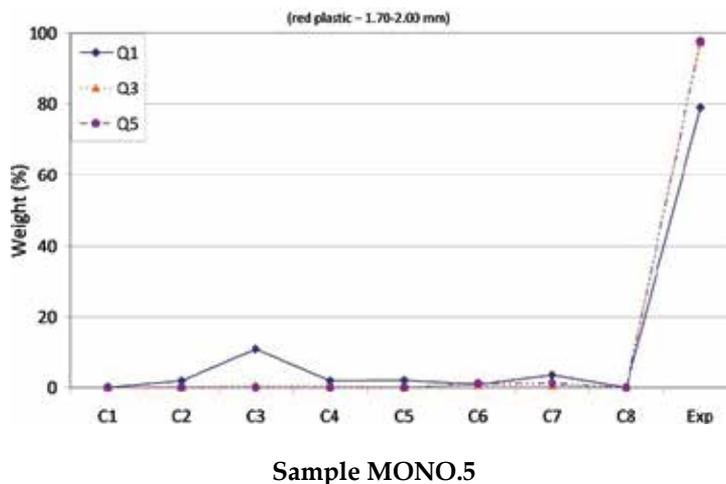


Fig. 8. Results of the experimental tests on mono-material samples at each adopted hydraulic head with Arrangement A

It is evident as the highest hydraulic heads (Q3 and Q5) involve high flowrates, and, therefore, high velocities that sweep away all the small-size particles (0.85-1.00 mm, samples MONO.2 and MONO.3), without achieving satisfactory differences in respective trajectories.

On the other hand, with the lowest hydraulic head (Q1), the heaviest particles (samples MONO.2, MONO.3 and MONO.4) tend to settle into the apparatus, while the lightest ones (sample MONO.5) tend to be expelled through the output nozzle, regardless of particle size. In this case, a separation of the different types of plastics is possible, thanks to the differentiation of trajectories.

The more selective chamber is the third one, in which the greatest amount of settled material is accumulated, independently from applied flowrate. In the chambers C1 and C8, only few particles settled due to the presence of input and output nozzles, respectively.

Green and brown plastic particles sized 0.85-1.00 mm (samples MONO.2 and MONO.3) show the same behaviour, so a separation seems not possible for each tested operative condition.

Also in the case of coarser particles (samples MONO.4 and MONO.5), the particle settling in the chambers shows the same trend, being C3 the chamber in which mainly the plastic particles settle regardless of the applied hydraulic head, while C1 and C8 remain almost empty.

The green plastic particles sized 1.70-2.00 mm (sample MONO.4) tend to settle in the apparatus, with slight differences in the distribution in the chambers by increasing the hydraulic head; in particular, the particles shift from C2 to C3, and then from C3 to the following chambers.

The red plastic particles sized 1.70-2.00 mm (sample MONO.5) tend to be expelled from the Multidune. Only at the lowest hydraulic head, a significant amount of material settles in the C3.

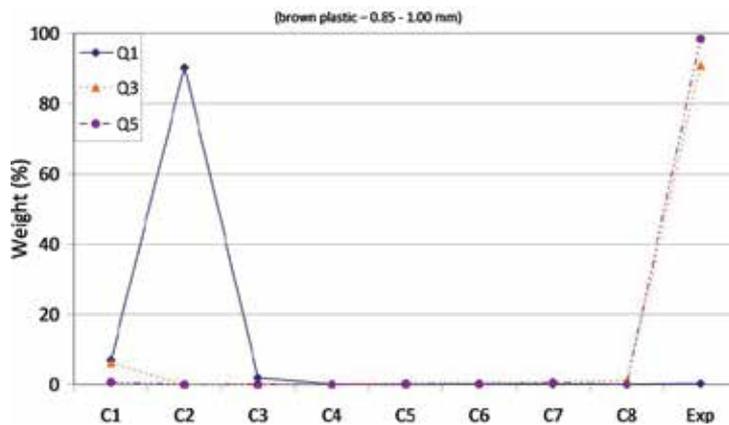
Therefore, a different behaviour of the samples sized 1.70-2.00 mm for each tested hydraulic head was observed, so a separation between them seems possible to be achieved.

5.2 Results with arrangement B

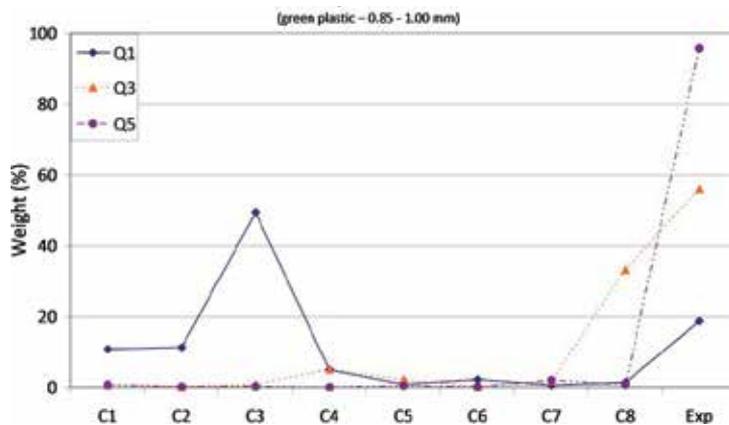
Analogous experimental tests have been executed with the Arrangement B of the Multidune (Figure 1). As previously commented, the only difference in the operative procedure was the choice of the output nozzles to be opened. Firstly, the same three output nozzles as in Arrangement A were opened (O2, O4 and O6), but all the samples were expelled out because the high flowrate determined too high flow velocity, in each operative condition. Considering such results, the experiments were carried out by opening only one output nozzle (O3), to lower flowrate and velocity.

Figure 9 shows the results of the experimental tests at each adopted flowrate, in terms of percentage in weight of the materials recovered from each chamber (C1-C8) and expelled through the output nozzle (Exp).

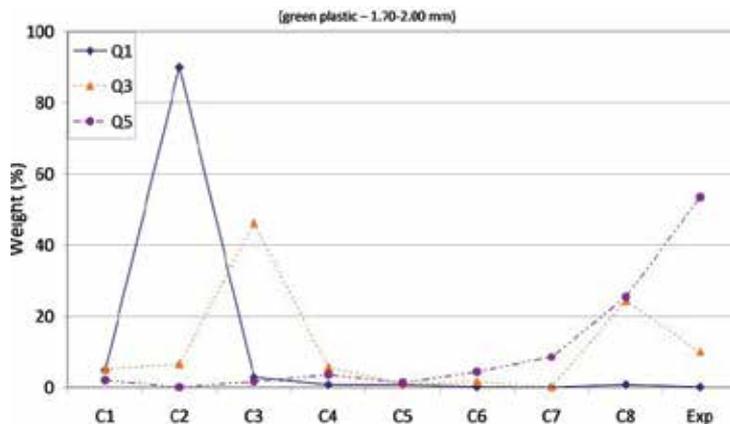
In Arrangement B, it is confirmed as the small-size particles (0.85-1.00 mm, samples MONO.2 and MONO.3) are quite totally expelled from Multidune imposing the highest hydraulic heads (Q3 and Q5): so, in these operative conditions the trajectories of the two types of plastics are similar.



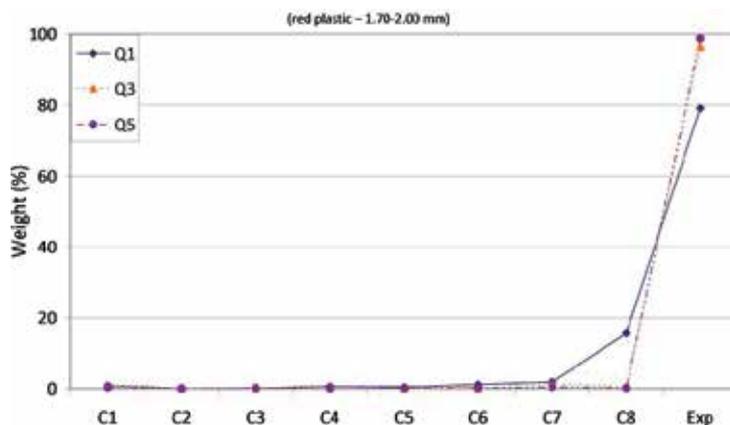
Sample MONO.2



Sample MONO.3



Sample MONO.4



Sample MONO.5

Fig. 9. Results of the experimental tests on mono-material samples at each adopted hydraulic head with Arrangement B

A separation seems to be possible at the lowest hydraulic head (Q1): actually, the heaviest particles (samples MONO.2, MONO.3 and MONO.4) tend to settle into the apparatus, while the lightest ones (sample MONO.5) tend to be expelled through the output nozzle, regardless of particle size.

The distribution of settled particles into the apparatus is more complex in Arrangement B than in Arrangement A. Particles path is in relation with the hydraulic head: as the hydraulic head increases, the particles tend to settled in the chambers closer to the output nozzles, consequently there is not a specific chamber (as C3 in Arrangement A) where particles are mainly accumulated.

The results of tests with Arrangement B confirm that it is possible to separate brown and green particles sized 0.85-1.00 mm (samples MONO.2 and MONO.3) imposing low flowrates, while green and red plastic particles of larger size (samples MONO.4 and MONO.5) can be separated at higher flowrates.

6. Conclusion

The high temporal and spatial resolution technologies employed for the reconstruction of the fluid-dynamic field inside the Multidune allows recognizing the flow field within the apparatus with both arrangements is characterized by three areas: the main transport current and, in each chamber, recirculation areas above and below. The fluid-dynamic behaviour is substantially similar in each chamber but in the first and the last one (C1 and C8), where the inlet and outlet nozzles prevent the formation of similar vortical structures. With Arrangement A, the characteristic velocity of the principal current appears significantly larger than the velocity within the upper and lower recirculation areas; this aspect is amplified with increasing hydraulic head at the apparatus inlet. With Arrangement B they appear more comparable. With both arrangements, the increase of the hydraulic head augments the transport effectiveness of the main current without improving the capture feasibility of both the upper and lower recirculation zones. The apparatus will then lose its effectiveness in separating plastic particles increasing both the hydraulic head and the transiting flowrate. Furthermore, with both arrangements, C3, C4, C5 and C6 present analogous recirculation areas. The same circumstance occurs at each flowrate establishing in the apparatus. Then, the presence of eight chambers assures plastic particle separation even if a chamber should be filled with the settled material allowing the following chambers to become effective in the separation process.

7. Acknowledgments

The authors would like to acknowledge Dr. Leonardo Cherubini and Dr. Emanuela Lupo for their contribution during the experiments.

8. References

- De Sena G., Nardi C., Cenedese A., La Marca F., Massacci P., Moroni M. (2008). The Hydraulic Separator Multidune: Preliminary Tests on Fluid-Dynamic Features and Plastic Separation Feasibility. *Waste Management* 29(9), 1560-1571.
- Hussain F. (1986). Coherent structures and turbulence, *J. Fluid Mech.* 173, 303.
- Joeng G., Hussain F. (1995). On the identification of a vortex, *J. Fluid Mech.* 285, 69.
- La Marca F., Moroni M., Cherubini L., Lupo E., Cenedese A. (2011). Recycling of plastic waste via the hydraulic separator Multidune. *Waste Management*, ISSN: 0956-053X, (submitted).
- Moroni M., Cenedese A. (2005). Comparison among feature tracking and more consolidated velocimetry image analysis techniques in a fully developed turbulent channel flow. *Measurement Science and Technology* 16, 2307-2322.
- Moroni M., Kleinfelter N., Cushman J.H. (2008). Alternative Measures of Dispersion Applied to Flow in a Convolutated Channel. *Advances in Water Resources* 32(5), 737-749.
- Moroni M., La Marca F., Cherubini L., Cenedese A. (2011). Recovering plastics via the hydraulic separator Multidune: flow analysis and efficiency tests. *International Journal of Environmental Science and Technology*, ISSN: 1735-1472, (submitted).
- PlasticsEurope (2007a). An analysis of plastics production, demand and recovery for 2005 in Europe. *PlasticsEurope*, 21 pp.
- PlasticsEurope (2007b). Press Release 9 May 2007, 1 p. Association of Plastics Manufacturers in Europe (AISBL), Brussels, Belgium.

Optimization of Pouring Velocity for Aluminium Gravity Casting

Y. Kuriyama, K. Yano and S. Nishido
Gifu National College of Technology
Mie University
AISIN TAKAOKA CO., LTD
Japan

1. Introduction

In current casting factories, tilting type automatic pouring machines are often used to pour the molten metal into the mold, with the operator relying on experience, perception and repeated testing to manually determine the pouring velocity. However, seeking an optimum multistep pouring velocity through trial and error results in an enormous number of combinations and is very difficult. For this reason, it cannot be said that suitable casting that realizes a high-quality cast is being carried out, inviting a decline in the yield rate due to product defects.

Furthermore, the extension of the production preparatory phase and increase in costs due to this kind of trial operation also become a significant problem.

Until now, much research relating to product defects in aluminum gravity molding has been conducted (Yutaka et al., 2001)(Takuya, 2004). Meanwhile, research applying casting CAE for the purpose of improving the quality of castings and production efficiency is coming to attention (Itsuo, 2006). Furthermore, in recent years, optimization of the casting problem has begun to be carried out in accompaniment with developments in computers (Takuya et al., 2007)(Ken'ichi et al., 2008). However, these all target comparatively short calculation time die-casting and adoption of optimization technology in sand mold casting and gravity casting is lagging.

Accordingly, in this research, the objective is to stabilize the fluid speed in the mold gate and derive an optimum pouring velocity to realize a mitigation of defects such as pin holes and blow holes in aluminum gravity casting through invoking a fluid behavior simulator, swiftly filling the sprue cup and controlling the liquid level at a fixed high level of liquid. For the automatic pouring machine, the multistep velocity input is designed for actual product of an intake manifold. The effectiveness of this research is shown by a fluid analysis simulation and an actual test.

2. Experimental apparatus

An overview of the automatic pouring machine is shown in Fig. 1. It is a tilting type automatic pouring machine with one degree of freedom in the forward and back direction of the ladle.

In the pouring machine, the tilting angular velocity and velocity switching angle are configured by teaching pendant and the pouring velocity is determined. The setup enables

four steps of tilting angle velocity and three of switching angle for a total of seven variables as shown in Fig.2. The tilting angular velocity command is assigned in the form of raised trapezoidal shapes as shown in Fig.3.

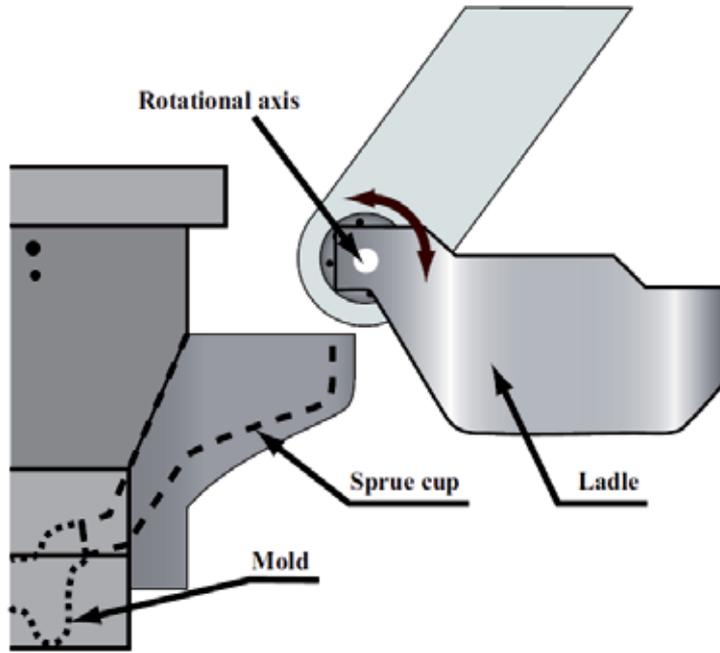


Fig. 1. Overview of the pouring machine

Switching angle [deg]	Pouring speed [%]
10	} v_1 : 1st speed } v_2 : 2nd speed } v_3 : 3rd speed } v_4 : 4th speed
θ_1 : Switching angle 1	
θ_2 : Switching angle 2	
θ_3 : Switching angle 3	
63	
10 ~ 63	1 ~ 100

Fig. 2. Input setting of the automatic pouring machine.

As the tilting angular velocity setting is displayed as a percentage, each step of tilting angular velocity is displayed by the following formula using the maximum angular velocity.

$$V_n = \frac{v_n}{100} V_{\max} \tag{1}$$

Here, it is necessary to carry out parameter identification, since maximum angular velocity V_{\max} and the angular acceleration a are unknown parameters.

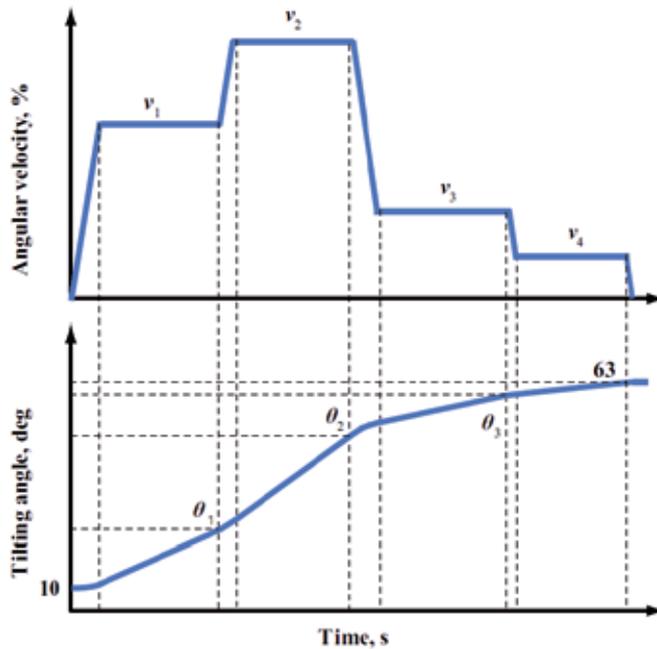


Fig. 3. Tilting velocity of pouring machine.

3. Identification the motion of the pouring machine

In this pouring machine, the tilting movement to input is unknown because the machine doesn't have the output device for the angular velocity or the angle. At the simulation of the fluid of molten metal, the identification of the tilting movement to input is needed. Thus, to get the unknown parameter of V_{max} and a , the movement to input of the actual pouring machine is filmed with the motion capture system. Table 1 shows the input of the pouring machine using analysis the motion.

	Switching angle (deg)	Pouring velocity (%)
1	22	10
2	32	30
3	42	50
4	--	30

Table 1. Setting of the tilting input for identification the motion.

The unknown parameter of V_{max} and a , is identified by using the data of the angular velocity from the motion capture. The angular acceleration and maximum angular velocity are $a=200$ (deg/s²), $V_{max} = 51.9$ (deg/s) respectively. Fig.4 shows the result of identification of actual pouring machine, where the solid line is the path of actual pouring machine and the broken line is the path of plant model.

From the Fig.4, it can be seen the motion is accorded very well. In this result, the plant model can be recreated the motion of the actual movement.

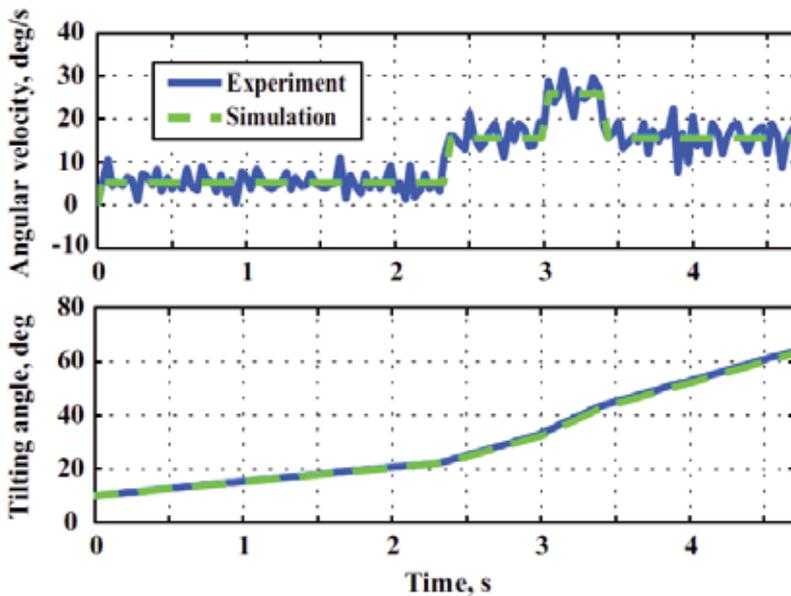


Fig. 4. Result of identification of actual pouring machine.

4. Construction of fluid behavior simulator and flow evaluation

The fluid analysis software FLOW-3D was used in this research. AC2B is taken as the subject fluid. The cast quantity is $1.863 \times 10^{-3} (\text{m}^3)$ and the product part volume is $1.429 \times 10^{-3} (\text{m}^3)$. The physical values of AC2B for the analysis were set as in Table 2. An outline map of the mesh in the simulation domain is shown in Fig. 5 and the mesh parameters are shown in Table 3.

Fluid parameters	AC2B aluminum alloy
Density	2550 kg/m ³
Viscosity	0.00125 Pa·s
Temperature of fluid	993 K
Specific heat	1071 J/(kg·K)
Thermal conductivity	100 W/(m·K)

Table 2. Fluid parameters

Mesh block	Cell size (m)	Number of cell
X-direction	0.004~0.080	108
Y-direction	0.005	54
Z-direction	0.004~0.080	65
Total number of cell		379080

Table 3. Mesh parameters

In an actual plant, a molten metal filter (wire mesh) is installed in the sprue runner with the purpose of removing slag as shown in Fig. 5. The thickness of the wires in the mesh is 0.5×10^{-3}

(m) and the number of wires is 50 in the vertical and 55 in the horizontal. In this research a porous baffle (hereinafter, baffle) was used to reproduce this filter in the CFD simulator. The air porosity b_p , the linear velocity drop coefficient b_l and the two-dimensional velocity drop coefficient b_q are assigned as setting parameters for the baffle. b_l and b_q are defined by an equation for baffle flow loss shown by Eq. (2).

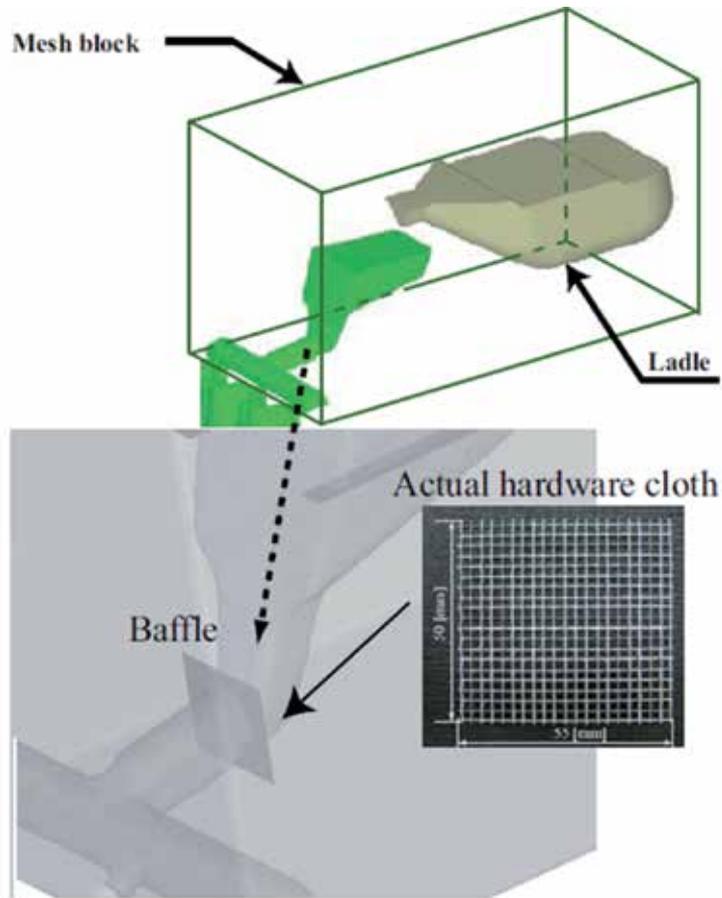


Fig. 5. Mesh setting of CFD

$$B = \frac{1}{L}(b_l u + 0.5b_q u|u|) \quad (2)$$

Here, B denotes the baffle flow loss, u the flow speed within the baffle and L the length in which the flow loss is produced. The air porosity of the wire mesh can be calculated from the area ratio of the metal wires and opening between them. As the linear velocity drop is dominant for the baffle flow loss, the 2-D velocity drop coefficient is set at $b_q=0$ and the results of searching for the linear velocity drop coefficient b_l are shown in Table 4. Furthermore, the search range was carried out in 0.05 increments over $b_l=0.00-1.50$. Simulation results considering the baffle loss are shown in Fig. 6.

Void	0.655
Linear loss coefficient : b_l	0.90
Quadratic loss coefficient : b_q	0.00

Table 4. Parameters of porous baffle

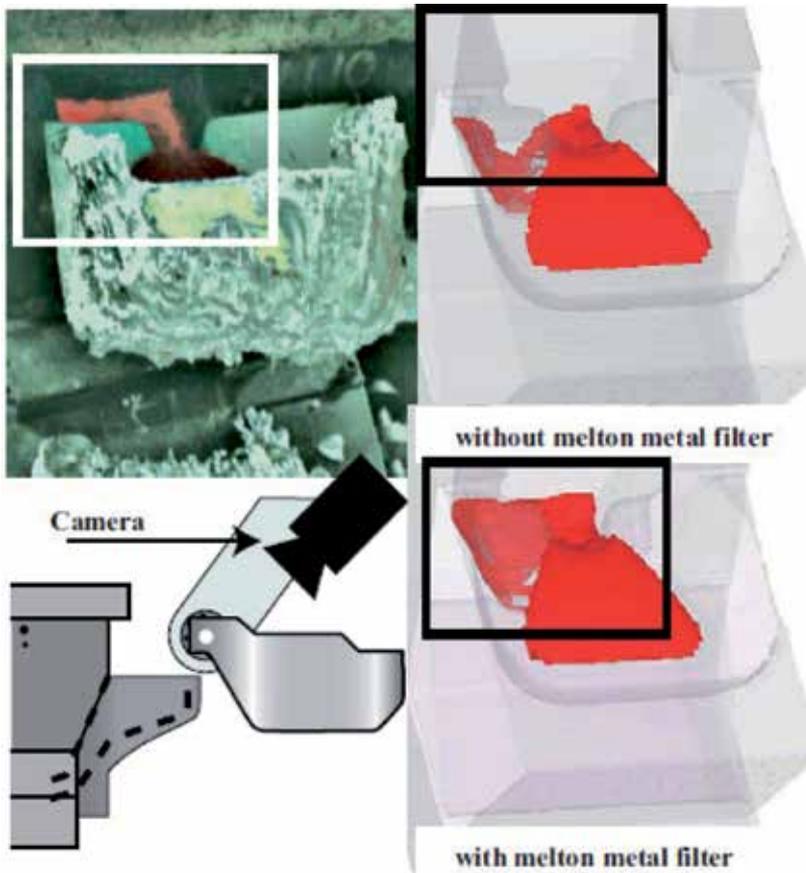


Fig. 6. Comparison of flow in sprue cup between simulation with molten metal filter and without molten metal filter.

Comparing these with the pouring test results, it can be seen that a satisfactory reproduction of the molten metal behavior inside the sprue is achieved.

5. Derivation of optimum pouring input using a genetic algorithm

Through swiftly filling the inside of the sprue cup with molten metal and controlling the liquid level to preserve a liquid level with high uniformity, the flow velocity in the mold gate is made constant.

This is regarded as making possible the production of high-quality casting that mitigates product defects. However, in the case the pouring velocity is determined by operator trial and error, problems occur such as the overflowing of molten metal from the melt due to

improper velocity setting and unstable liquid levels inside the sprue cup etc. Accordingly, in this research, liquid level control is realized through optimizing the pouring input using a genetic algorithm (GA) (Thomas et al., 1993).

GA is an algorithm that models natural selection and mutation in the processes of inheritance and evolution in biological groups in the processes of evolution and inheritance for populations in the natural world and is a probabilistic optimization method.

The three steps of switching angle and four steps of pouring velocity for a total of seven set parameters are taken as variables that are the settings input for the pouring machine, and an optimum tilting velocity pattern is calculated within the limitations of the real machine. In order to swiftly fill the sprue cup and stabilize the liquid surface at a uniform level, the tilting end time is taken as a performance function and the optimization problem is expressed by Eq. (3) with the liquid level inside the sprue as a limiting condition. Through taking the tilting end time as a performance function, the filling time is reduced and through already taking the liquid level as a limiting condition, liquid level control becomes possible.

$$\text{minimize : } J = t_p + J_p \tag{3}$$

$$h \geq 0.025$$

Here, t_p denotes the tilting end time, J_p a penalty function denoted by Eq. (4), and h the displacement from the sprue exterior to the molten metal.

$$J_p = w_1 + w_2 \tag{4}$$

In Eq. (4), the penalty clause $w_1, w_2=100$ is imposed in the case the displacement from the sprue exterior to the sprue interior liquid level drops to below 0.025(m) ($h < 0.025$).

6. Optimization with CFD simulator

6.1 Optimization result and melt flow analysis

Optimization by GA was carried out using the calculation parameters shown in Table 5 Forty-eight hours was required for optimization with 41st generation using an Intel Core2 Quad CPU equipped PC.

Number of variable	7
Number of population	10
Number of elite preservation	1
Mutation fraction	0.01
Crossover fraction	0.80

Table 5. Parameters of genetic algorithms

At that time the evaluation value was $J=4.668$. The tilting angles and velocities obtained from the optimum parameters are shown in Fig. 7. And the simulation results of the liquid level control is shown in Fig. 8.

As a result, it can be seen that the objective liquid level is not reached in the case of any control and a satisfactory liquid surface is not maintained. Conversely, it can be seen that a satisfactory liquid level control that swiftly fills the sprue cup is realized through the optimization of pouring control input.

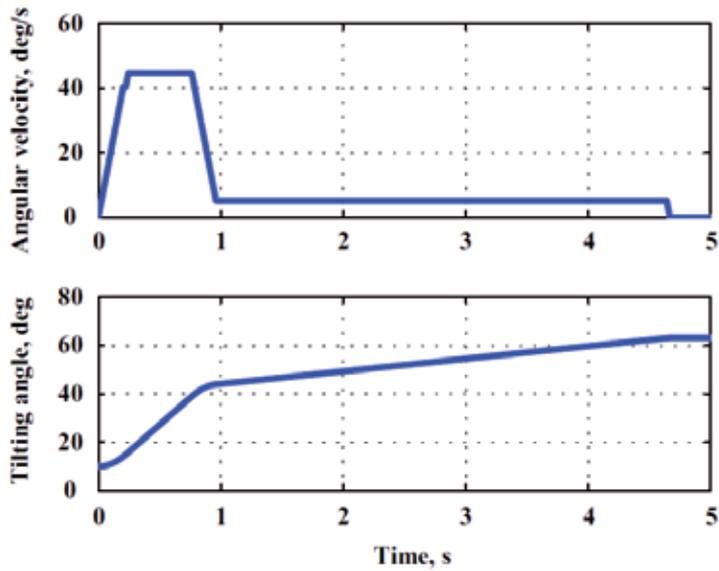


Fig. 7. Tilting input of optimization result.

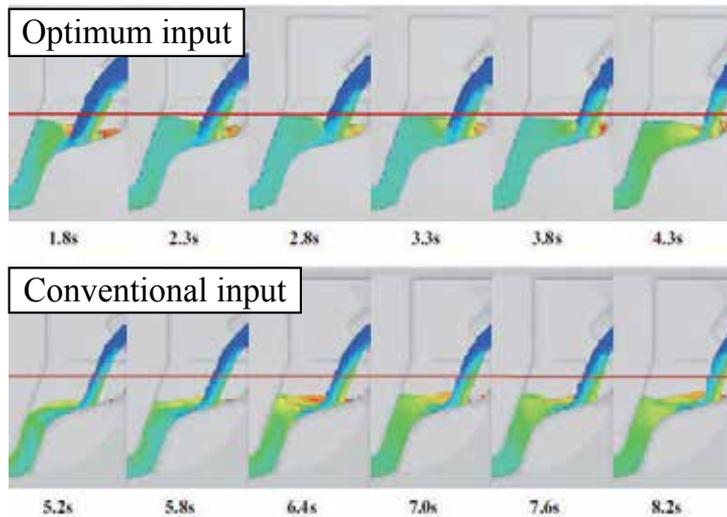


Fig. 8. Tilting input of optimization result.

6.2 Evaluation of the optimum input

Air entrainment is one of the defect origins of such as pin holes and blow holes as shown in Fig.9. Thus, using the evaluation function of air entrainment which is one of the functions of *Flow-3D*, the optimum input is evaluated.

The air entrainment at the liquid surface is based on the concept that turbulent eddies raise small liquid elements above the free surface that may trap air and carry it back into the body of the liquid. The extent to which liquid elements can be lifted above the free surface

depends on whether or not the intensity of the turbulence is enough to overcome the surface-stabilizing forces of gravity and surface tension as shown in Fig. 10.



Fig. 9. Photograph of blow hole

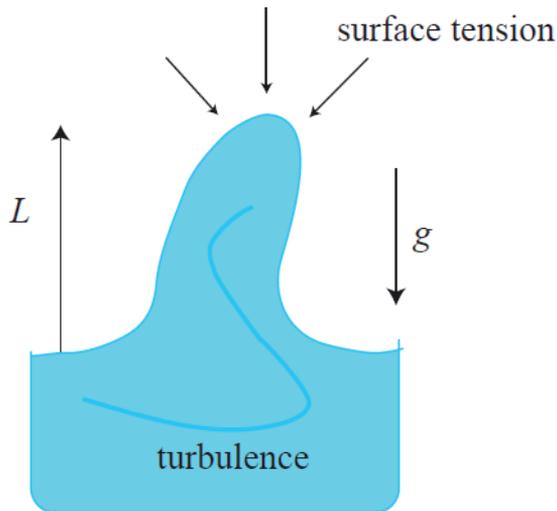


Fig. 10. Model of air entrainment

Turbulence transport models characterize turbulence by a specific turbulent kinetic energy Q and a dissipation function D . The characteristic size of turbulence eddies is then given by Eq. (5).

$$L = 0.1 \frac{\sqrt{Q^3}}{D} \quad (5)$$

This scale is used to characterize surface disturbances. The disturbance kinetic energy per unit volume (i.e., pressure) associated with a fluid element raised to a height L , and with surface tension energy based on a curvature of L is given by Eq. (6).

$$P_d = \rho g L + \frac{\sigma}{L} \tag{6}$$

where ρ is the liquid density, σ is the coefficient of surface tension, and g is the component of gravity normal to the free surface.

For air entrainment to occur, the turbulent kinetic energy per unit volume, $P_t = \rho Q$, must be larger than P_d ; i.e., the turbulent disturbances must be large enough to overcome the surface-stabilizing forces. The volume of air entrained per unit time, V_a , is given as Eq.(7).

$$\frac{\partial V_a}{\partial t} + \nabla V_a = R t (1 - V_a) \tag{7}$$

where $R = C_{air} \sqrt{2(P_t - P_d) / \rho}$, u is fluid velocity, t is time, and C_{air} is a coefficient of proportionality : $C_{air} = 0.5$; i.e., assume on average that air will be trapped over about half the surface area. If P_t is less than P_d then V_a is zero. Furthermore, Fig.11 is shown the measurement point of out flow rate and V_a .

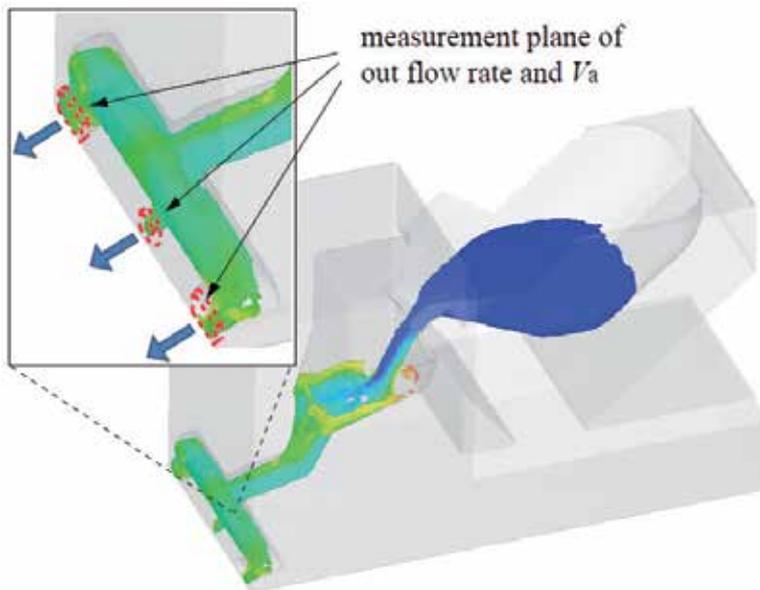


Fig. 11. Measurement plane of out flow rate and V_a

The air entrainment is expressed in Eq.(8).

$$A = \sum_{k=1}^n V_{ak} F_{fk} V_{fk} V_{ck} \tag{8}$$

where A is the quantity of air entrainment, V_a is the volume of air entrained per unit time, F_f is the fluid fraction, V_f is the volume fraction, V_c is the volume of the mesh cell, and n is the aggregate number of mesh cells.

Fig.12 shows the flow of molten metal with the mold. It can also be seen the satisfactory liquid level control that swiftly fills the sprue cup is realized. Fig.13 shows the air entrainment at the machining surface. Upper figure shows using the conventional input,

lower figure shows the optimum input, where the color of this figure is indicated the value of the air entrapment, and red color is the most air entrapment point. From the figure, the optimum input can be decrease the air entrapment, and it can be expect to realize a mitigation of defects such as pin holes and blow holes in experiment.

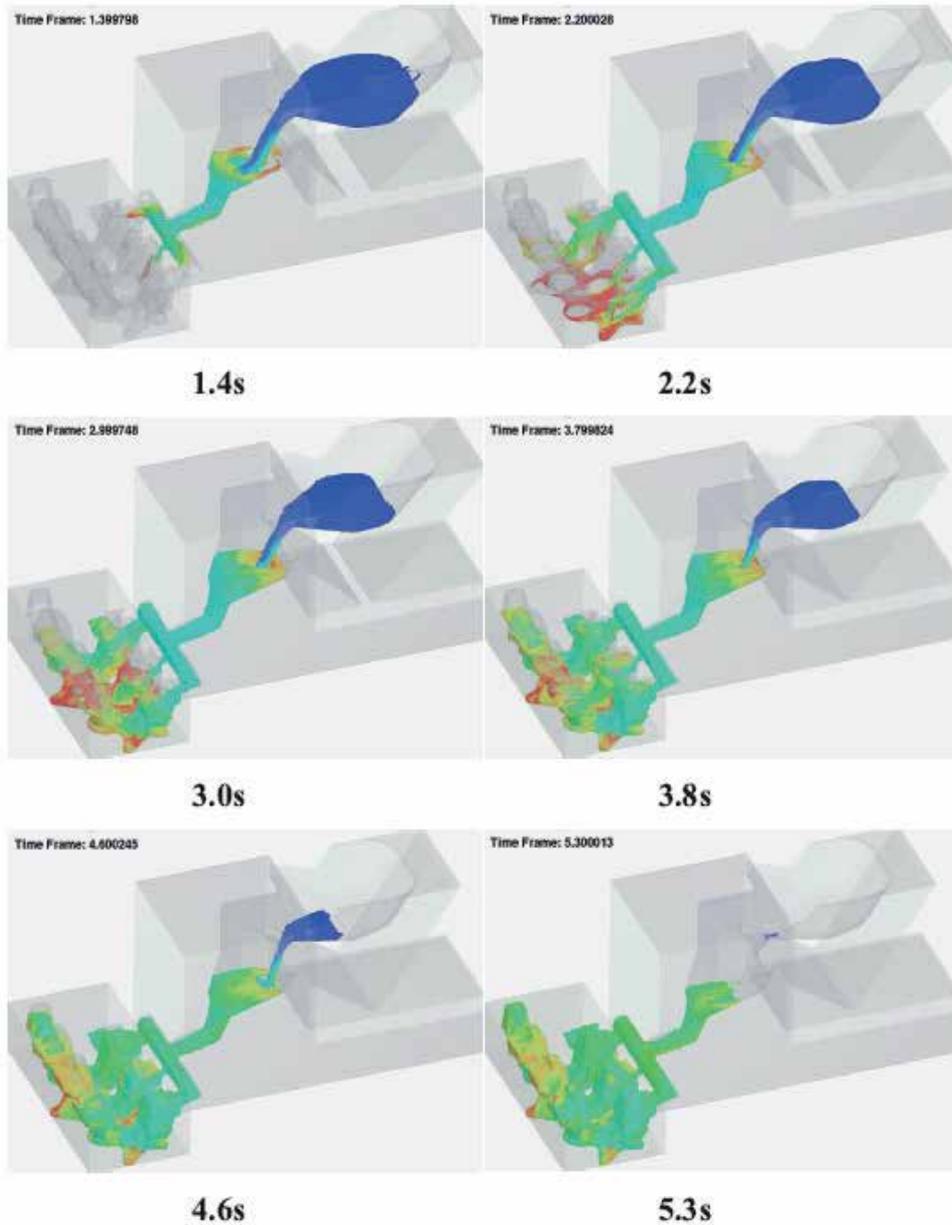


Fig. 12. Flow of molten metal with the mold

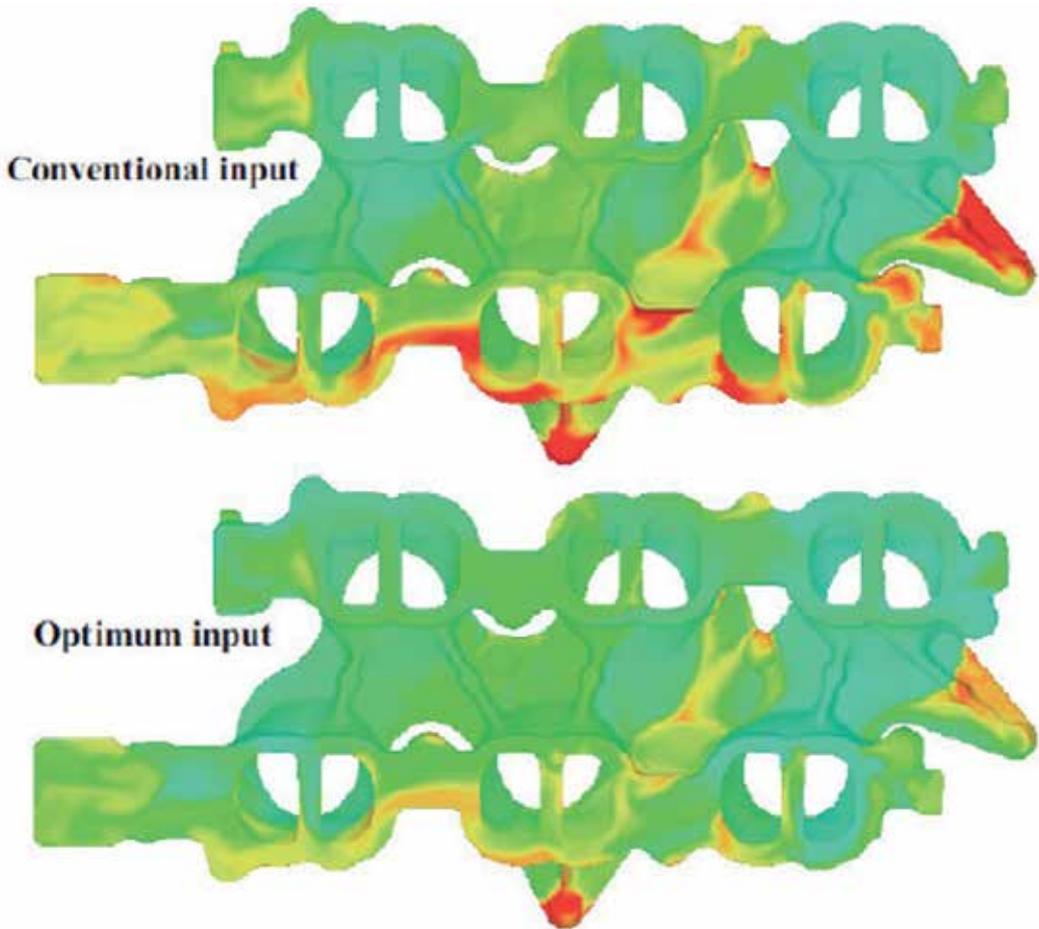


Fig. 13. Air entrainment at machining surface

7. Pouring control experiment using optimum velocity input

The desired optimum velocity input was applied to an actual machine and a pouring experiment was conducted. The conventional inputs were used to compare the optimal solution. Eleven trial runs were carried out and the frequency of defects appearing in the machined surface was evaluated after machining the product part from which the sprue, gating system and gate riser part had been removed. Fig. 14 shows the experiment results using each input and Fig. 15 shows the defect frequency for each.

As a result, it was possible to reduce the manufacturing defect to 1/6 rate by using an optimum pouring velocity as shown in Fig. 14 and Fig. 15. From this it was seen that the optimum input parameter derived using the proposed technique was valid for mitigating the occurrence of product defects.

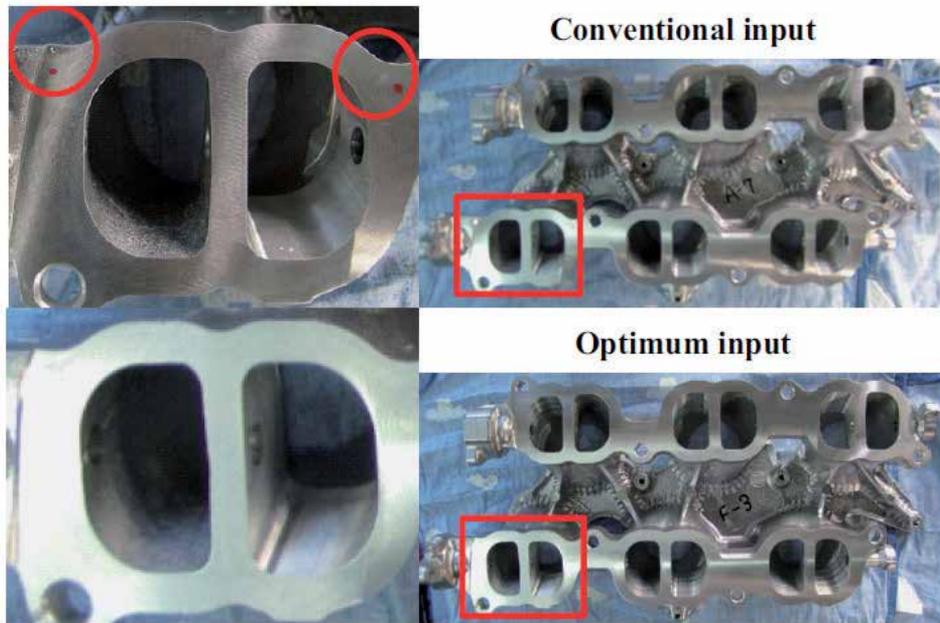


Fig. 14. Experimental results with conventional input and optimum input

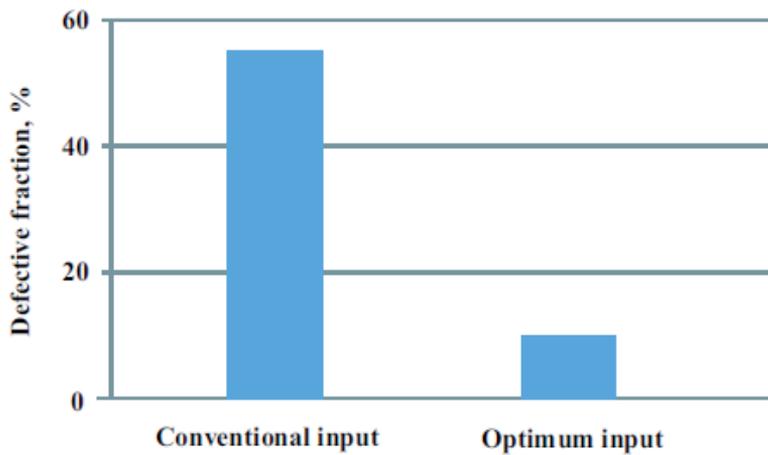


Fig. 15. Defect fraction of machining surface

8. Conclusion

In this research, an analysis technique that enables a reduction in calculation time in fluid analysis simulations was proposed and pouring control input was optimized with the purpose of reducing occurrence of defects such as blow holes and pin holes in aluminum gravity casting.

As a result of optimizing pouring control input using GA, optimum pouring control input realized liquid level control and its validity in mitigating product defects was seen through a real machine pouring experiment.

9. References

- Yutaka K., Nobuaki N. & Hideaki O., (2001). Classification of Pin Hole Defect of Iron Casting. *Journal of Japan Foundry Engineering Society*, Vol.,73, No.,4, (2001 Apr.) pp. 258-263, 1342-0429
- Takuya S.,(2004). Prediction of Gas Defects by Mold-Filling Simulation with Consideration of Surface Tension. *Journal of Japan Foundry Engineering Society*, Vol.,76, No.,7, (2004) pp. 562-569, 1342-0429
- Itsuo O., (2006). State of the Art of Simulation of Casting. *Journal of Japan Foundry Engineering Society*, Vol.,78, No.,12, (2006 Dec.) pp.602-608, 1342-0429
- Takuya S. & Ichiyo N., (2007). Optimization of Injection Speed for Reduction of Cold Shut in Die Casting. *Journal of Japan Foundry Engineering Society*, Vol.79, No.10,(2007 Oct.), pp.592-600, 1342-0429
- Ken'ichi Y., Koutarou H., Yoshifumi K.& Seishi N., (2008). "Optimum Velocity Control of Die Casting Plunger Accounting for Air Entrapment and Shutting," *International Journal of Automation Technology*, Vol.2, No.4,(2008 Jun.) pp. 259-265, 1881-7629
- Thomas B. & Hans- Paul S.,(1993). "An Overview of Evolutionary Algorithm for Parameter Optimization, " *Evolutionary Computation*, Vol.1, No.1,(1993 Apr.) pp. 1-23, 1063-6560

Fluid Dynamics Without Fluids

Marco Marcon

*Politecnico di Milano, Dipartimento di Elettronica e Informazione,
Milano
Italy*

1. Introduction

This chapter will discuss some interesting real applications where Fluid Dynamics equations found fruitful applications without dealing with "strictly speaking" fluids. In particular, thanks to the large set of analyses performed over different kinds of fluids in different operating and boundary conditions, a wide range of Computational Fluid Dynamics algorithms flourished tackling different aspects, from convergence rate, to stability according to the discretization, to multigrid and linearization problems. This robust and thorough background, both on theoretical and on practical aspects, made Computational Fluid Dynamics (CFD) appealing also to other sciences and applications where Fluid Dynamics equations, or similar equations very close to them, can be useful in describing complex phenomena not related to fluids. Some applications that will be discussed concern, e.g., Geometry of liquid snowflakes whose contour is growing steered by curvature, starting from a circle. Furthermore Image Restoration and Segmentation can also benefit from CFD since a set of evolutionary algorithms, based on level-set curvature flow equations, plays a fundamental role in steering active contours or snakes through the noise present in the image till the complete warping of the desired framed object. Also in this case advanced techniques like Ghost Fluids Method for two competing fluids dynamics can be used to separate different objects in images. Other interesting applications that will be described concern applicability of CFD to surface extraction from cloud of points. This is a common problem when complex clouds of points, representing 3D objects or scenes are obtained by laser scanners or multi-camera vision systems. These points represent unambiguous features from corners or sharp edges and the final 3D closed surface must fit on these points smoothly interpolating empty space between them. Also in this case CFD can provide useful tools to define the evolution of a 3D surface representing the border between two competing fluids, one representing the "inside" and the other the "outside" of the object itself. The two fluids evolution will stop when surface sticks on all the 3D points: the viscosity of the two fluids will control the smoothness of this surface that will wrap the cloud and turbulence is used to model injection into grooves or narrow holes. This chapter will also discuss another interesting application of CFD to robotic navigation in complex environments where we are looking for the best path, both in terms of length and distance from objects, through a set of obstacles, different terrains traversability or path slope. Also in this case an imaginary fluid with a predefined viscosity floods from the robot position through the whole environment, its front

evolution speed, accordingly to CFD, will be slower in narrow passages and, once it reaches the target, it will define the easiest way.

2. Snowflakes and phase interfaces

In many physical problems outlines of a considered fluid have different speeds at different points, and in many cases local speed is directly connected to the curvature. In the following we will show how a geometrical contour, represented by a 2D line or a 3D surface of the evolving shape changes from point to point accordingly to the local geometrical properties of the contour itself. The presented approach can be applied to a generic N-dimensional case and we can refer to the moving contour as a general hypersurface, anyway in the following, we will limit without loss of generalization our considerations to a 3D space where fluid boundary is represented by a closed surface.

In particular Sethian (1989) showed how algorithms based on direct parametrization of the surface evolution could present ambiguous and error-prone solutions due to the local error propagation while global approaches based on implicit representations can result in a much more robust solution. This kind of approaches usually relies on higher dimensional functions for which the considered surface is just represented by a level-set. The contour motion can then be described applying the proper Hamilton-Jacobi equation related to a hyperbolic conservation law to the implicit function and tracking the particular level-set representing the contour itself.

Implicit function formulation is much more robust to numerical techniques with respect to direct methods and can be implemented starting from an original closed and non-intersecting surface: two famous physical examples are the combustion model of a flame and the grow of a snowflake: Markstein (1951) assumed that thin flame fingers close to their ends are cooler than their inner part and move slower than hot nonconvex regions: the proposed evolution equation assumes that flame contour speed is inversely proportional to its curvature κ , (where κ is positive for convex regions and negative for nonconvex ones). This assumption comes from the fact that in highly convex regions, flame particles collide with a high number of slower and cooler surrounding particles that slow down their motion. The symmetrical effect is present in spikes growing in snowflakes. In particular these two last examples, flames and snowflakes, represent two very good examples where simple equations can explain very well complex phenomena. For a comprehensive description of ice crystal growth we suggest Libbrecht (2005), while in the following we describe how solid-liquid interface dynamics are steered by curvature in liquid snowflakes (or Tyndall figures). We refer to some recent studies (Hennessy (2010), Hobbs (2010)) on cylindrical discs of liquid in superheated crystals of ice. In this case the snowflake is made of water inside ice and its geometrical shape evolution is determined by boundary condition between solid and liquid phase. In Fig. 1 there is a typical snapshot of a Liquid snowflake.

2.1 Liquid snowflakes surface evolution

The typical modelization of snowflakes surface growing is the Gibbs-Thomson equation where the interface temperature T_I , between the solid and the liquid part is ruled by the equation:

$$T_I(\kappa) = T_M \left(1 - \kappa \frac{\gamma}{\rho L} \right) \quad (1)$$

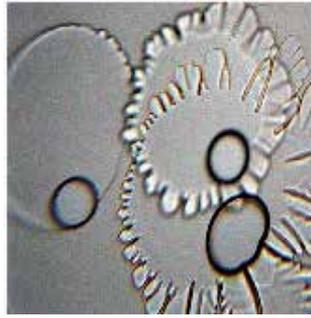


Fig. 1. A typical Tyndall figure (Liquid snowflake)

Where T_M is the equilibrium temperature of a planar solid-liquid interface (273.15K for water), κ is the curvature, γ is the surface energy per unit area for the solid-liquid interface, ρ is the density of a specific phase but in this case we are assuming equal density both for ice and water and L is the latent heat of fusion. The steering interface force is then due to the pressure variation across the interface and depends from curvature:

$$p_s - p_l = \rho (s_l - s_s) (T_I - T_M) = \rho L \left(\frac{T_I(\kappa)}{T_M} - 1 \right) \quad (2)$$

Where the subscript l and s refer to liquid and solid state respectively and s indicates the entropy per unit mass, for further details refer to (Hills & Roberts (1993)).

A further improvement in the Gibbs-Thomson model considers surface energy anisotropy due to crystal ice structure (six fold symmetry), this will add a dependence of the surface energy γ from the relative orientation of the surface with respect to the underlying crystal. Accordingly to (Davis (2001)) a six fold symmetrical crystal will have a:

$$\gamma_6 = \gamma_0 \left[1 + \frac{\sigma_n}{n^2 - 1} \cos(n\theta) \right] \Big|_{n=6} \quad (3)$$

Where σ_n represents the anisotropy degree and the complete interface temperature will be, accordingly to (Pimpinelli & Villain (1999));

$$T_I(\kappa) = T_M \left(1 - \kappa \frac{\gamma_6 + \frac{d^2\gamma_6}{d\theta^2}}{\rho L} \right) \quad (4)$$

In particular, accordingly to equations 3 and 4, whenever $|\sigma_n| < 1$ we are in a weak anisotropy condition, i.e. the surface energy γ is always positive, and, at equilibrium, the surface will resemble a smooth line, while, for anisotropy values greater than one, as for the snowflakes water example, the thermodynamic equilibrium implies a closed polygonal made of straight line segments (Dobrushin et al. (1992)).

Once the Temperature of the interface is determined the shape of the boundary between the liquid and solid part and its time evolution can be recovered from the energy conservation, in particular, in order for the ice-water interface to advance inside the solid part, two energy contributions must be provided: one for solid melting and one for interface stretching. The melting energy that must be provided per unit of time for a surface A that is moving inside

the solid at a normal speed v_n is $l_{mel} = \rho L A v_n$ while the energy contribution per unit of time due to surface stretching is $l_{sur} = \gamma \frac{dA}{dt}$; which, accordingly to the differential geometry can be written as:

$$l_{sur} = \gamma \frac{dA}{dt} = A \gamma \kappa v_n \quad (5)$$

The surface advancement is steered by two opposite heat diffusive fluxes: one inside the liquid part and one inside the solid part; respectively: $-k_l \nabla T_l$ and $k_s \nabla T_s$ where k indicates the thermal conductivity of the considered phase.

The equilibrium of energy equations per area and time unit can be written as:

$$\frac{l_{mel}}{A} + \frac{l_{sur}}{A} = (-k_l \nabla T_l + k_s \nabla T_s) \cdot \mathbf{n} \quad (6)$$

i.e.:

$$(\rho L + \gamma \kappa) v_n = (-k_l \nabla T_l + k_s \nabla T_s) \cdot \mathbf{n} \quad (7)$$

Where \mathbf{n} is surface normal oriented from the liquid part towards the solid one; this equation is called the Stefan condition. The numerical solution for the liquid-solid interface evolution, accordingly to the Stefan condition, requires some particular precaution due to accuracy that must be placed in surface tracking. Since the analyzed problem is far from thermodynamic equilibrium due to the superheated temperature of the solid, some typical approaches, e.g. the Enthalpy model (Shastri & Allen (1998)), are not accurate. In fact, Enthalpy approach, even if requires only the evaluation of a single parabolic PDE at each time step, assumes that the interface is within a mushy solid-liquid region and the evolving interface is just approximated. A possible extension of Enthalpy approach to account for superheated solid state could be phase-field models; anyway an accurate interface tracking requires a very small grid spacing resulting in a high computational cost (Biben et al. (2005)). One of the method that in recent years is getting growing attention is Level-set: in the following of this chapter we will show its versatility and accuracy in many CFD applications and in liquid snowflakes shape definition.

3. Level-set methods

For most of real CFD systems symbolic formulation that easily yield to symbolic mathematical solution is unfeasible, and the goal becomes to find the system modelization with the best trade-off between accuracy and computational complexity accounting for local, global and independent properties of the moving surface. In particular we will refer to "local" as properties associated to local surface features (e.g. curvature, local density or temperature), "global" as properties associated to overall surface shape or extension, while independent properties are not linked to the surface itself, like a flow below the surface. The Level-set method aims to track a thin propagating interface over time, it outperforms many other approaches handling topological complexities such as corners and cusps, and in handling complexities in the evolving interface such as entropy conditions and weak solutions. A detailed description of level-set can be found in (Sethian (1999)).

The general idea is to define a function $\phi(p, t)$ over the whole space-time domain, where p indicates a point in the space and t represents the time variable. The initial surface is the zero level-set of an implicit function for $t = 0$: i.e. the starting surface is the set of points p so that $\phi(p, t) = 0$ the value of the ϕ function for non-interface points is, the signed distance from the surface where 'signed' means that different signs indicate internal or external position with

respect to the closed surface. The discrete formulation requires that the value of ϕ is initially evaluated over all the mesh points and the steering equations would require the update of each point at each time step. This computational complexity of order $O(n^3)$ at each time step could be unacceptable in many practical cases so many faster update solutions have been proposed; the most common are:

- Narrow band: the updated mesh points are only those in a small stripe surrounding the interface, this will reduce computation to $O(n^2)$ but as soon as the interface is approaching the stripe boundary all values must be recomputed (more details in (Adalsteinsson & Sethain (1995))
- Octree: relevant mesh points and high detailed surface regions can be divided in sub-grids, e.g. a cube can be split in eight sub-cubes (the method name come from here) and the procedure can be iterated for higher precision. More details can be found in (Losasso et al. (2004))
- Sparse block grid: The whole domain is subdivided in blocks representing clusters of the original grid points (they can be overlapped or not) then the update is evaluated only on blocks containing the surface narrow band. (Further details in (Bridson (2003))

3.1 Level-set evolution

Accordingly to the previous description the set of points belonging to the surface will satisfy the zero level-set condition: $\phi(\mathbf{p}(t), t) = 0$ where $\mathbf{p}(t)$ represents the path of a surface point. The first temporal derivative will become:

$$\frac{d\phi(\mathbf{p}(t), t)}{dt} = \frac{\partial\phi}{\partial t} + \nabla\phi \cdot \frac{d\mathbf{p}(t)}{dt} = 0 \quad (8)$$

Considering the normal unitary vector to a generic point of ϕ for a generic level-set (contour), $\mathbf{n} = \frac{\nabla\phi}{|\nabla\phi|}$ the speed function v along this direction is then:

$$v(t) = \frac{d\mathbf{p}(t)}{dt} \cdot \mathbf{n} \quad (9)$$

obtaining:

$$\frac{d\phi(\mathbf{p}(t), t)}{dt} = \frac{\partial\phi}{\partial t} + v(t) |\nabla\phi| = 0 \quad (10)$$

In the cases considered in the following and in many physical cases the interface evolution is ruled by a Hamilton-Jacobi equation and the local velocity can be assumed proportional to the local curvature of the surface itself:

$$v = v_0 - \epsilon\kappa \quad (11)$$

where ϵ is a constant and

$$\kappa = \frac{\phi_{xx}\phi_y^2 - 2\phi_y\phi_x\phi_{xy} + \phi_{yy}\phi_x^2}{(\phi_x^2 + \phi_y^2)^{3/2}} \quad (12)$$

We can rewrite ϕ as follows:

$$\frac{\partial\phi}{\partial t} = -v_0 |\nabla\phi| + \epsilon\kappa |\nabla\phi| \quad (13)$$

3.2 Liquid snowflakes surface evolution

Assuming that the ice is radiated with an external electromagnetic radiation normal to its surface (a focused beam with a gaussian shape) the heat will follow an exponential decay penetration, accordingly to the following heat propagation equation:

$$I(r, z) = I_0 e^{-(r/r_b)^2 - \alpha z} \quad (14)$$

where I_0 is the central beam intensity, r_b is its standard deviation while r is the in-plane coordinate for the ice surface. Furthermore z is the coordinate orthogonal to the ice surface and α is a penetration coefficient. The formation of the liquid snowflake is then in the in-plane direction. A level-set function can be defined where ϕ is the normalized temperature that follows heat PDE:

$$\begin{cases} \frac{\partial \phi}{\partial t} = \nabla^2 \phi + e^{-(r/\beta)^2} & \text{for the liquid} \\ c_{pr} \frac{\partial \phi}{\partial t} = k_r \nabla^2 \phi + \alpha_r e^{-(r/\beta)^2} & \text{for the solid} \end{cases} \quad (15)$$

where $c_{pr} = \frac{c_{ps}}{c_{pl}}$ is the relative measure of heat capacities, $k_r = \frac{k_s}{k_l}$ is the relative thermal conductivity and $\alpha_r = \frac{\alpha_s}{\alpha_l}$ is the relative absorption coefficient.

The liquid/solid interface is then associated to the zero level-set in a sort of rescaled Celsius scale. Experimental observations show that liquid snow flakes often begin as circular discs which then grow outwards in a radially symmetric manner. However, after a certain amount of time, the interface becomes unstable and small, sinusoidal perturbations with a well defined wave-number appear. In Fig. 2 it is possible to see the growing of a liquid snowflake starting from a circle and then evolving with equally spaced rippling.

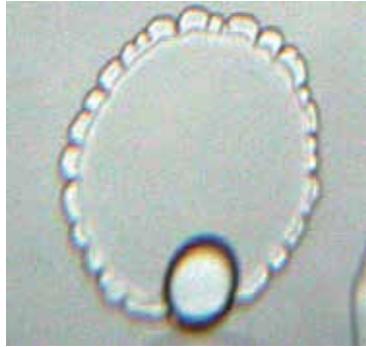


Fig. 2. Liquid snow flakes formation starting from a circle and the evolving, accordingly to linear instabilities with sinusoidal perturbations

The eq. 15, accordingly to the problem symmetry can be written in polar coordinates for the quasi-steadily condition:

$$\begin{cases} \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \phi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \phi}{\partial \theta^2} + e^{-(r/\beta)^2} = 0 & \text{for the liquid} \\ \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \phi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \phi}{\partial \theta^2} + \frac{\alpha_r}{k_r} e^{-(r/\beta)^2} = 0 & \text{for the solid} \end{cases} \quad (16)$$

Even small linear perturbations must satisfy the equations above and an integer number of oscillations (first order approximation) will constitute the zero level-set. Assuming a starting

contour equal to a circle of radius R the curvature will be equal to $\kappa = -\frac{1}{R}$. The time evolution of the level-set accordingly to the aforementioned equations is represented in Fig. 3 where a very good agreement of the zero level-set with real data is reached. For further details about analytical integration of eq. 16 refer to (Hennessy (2010)).

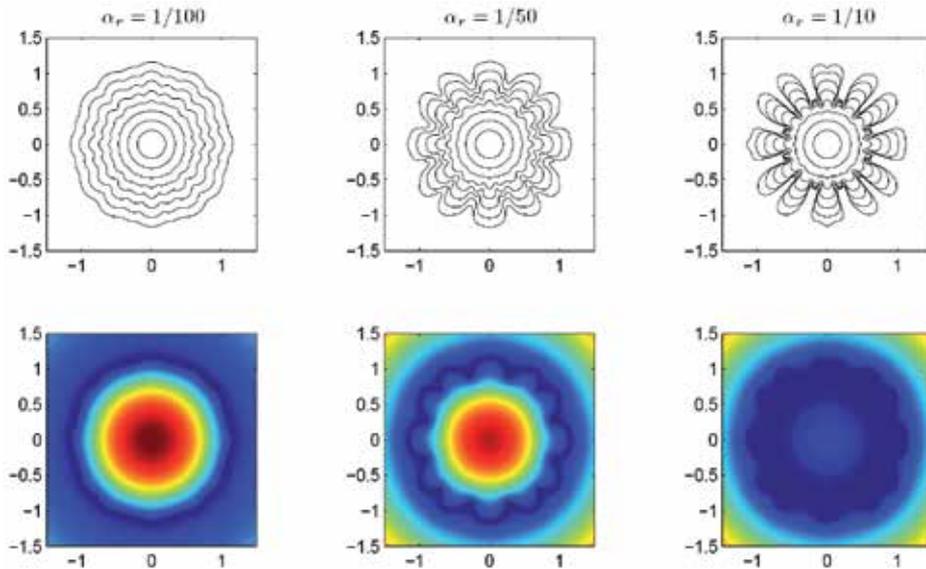


Fig. 3. on the top row: different snapshots of the zero level set evolution varying the relative absorption α_r ; instead on the bottom line different level set values (different temperatures) are presented for the larger curve of the row above

4. Motion planning for traveling robots in complex environments

Fluid dynamics and Level-set formulation can also be fruitfully adopted in Motion planning Algorithms. In particular the flexibility of the presented technique for time optimal motion planning with moving obstacles can be easily integrated with further constraints like different terrains traversability, path narrowing or fuel economy. Many classical techniques are present in literature for shortest path search with stationary obstacles, a good survey can be found in Latombe (1991). The space-time representation of obstacles moving in a 2D environment, assuming that one axis will be the time variable will be a solid structure. In Fig. 4 there is the representation of a circular obstacle moving in the x - y plane along x direction with an abrupt slow down.

As shown in Kimmel et al. (1998) the search for the time-optimal path within the aforementioned 3D time-space can be reduced to the search in a reduced set of 2D regions forming a Multivalued Distance Map (MDM). A Distance Map from a point \mathbf{p} is a function that assigns to each point \mathbf{q} in the considered domain a value corresponding to the "minimal length" (geodesic). The simplest approach uses a contour with a constant speed propagating from a small circle around \mathbf{p} and the distance assigned to each crossed point \mathbf{q} is proportional to the propagation time. If obstacles are moving the optimal path will be the minimal geodesic only when it does not cross any moving obstacles in the scene. For all the other cases a

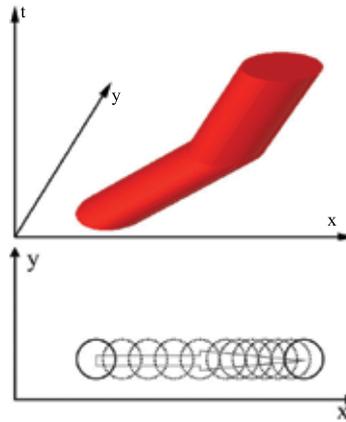


Fig. 4. a space-time representation for a circular obstacle moving in the x - y plane along the x direction that slow down its speed

more sophisticated approach is required: MDM will keep track of multiple crossing of the propagating contour when moving obstacles are present: in particular $MDM_p(\mathbf{q}, 1)$ will indicate the first time at which the point \mathbf{q} is reached from \mathbf{p} while $MDM_p(\mathbf{q}, 2)$ will indicate the second time at which, accordingly to the moving objects the propagating contour can reach \mathbf{q} , and so on. We will call $\mathcal{N}(\mathbf{q})$ the number of times that the propagating front will cross that point. Once the MDM is built for all points in the domain, the best path is obtained using a back-track procedure.

Consider now, given the source point \mathbf{p} , a continuous surface \mathcal{S} in the space-time domain, centered in \mathbf{p} indicating the wavefront contour as a function of time: $\mathcal{S} = \{(x, y, t(x, y))\}$, where a contour at a specific geodesic distance \bar{d} , $\alpha(\bar{d})$ is given by:

$$\alpha(\bar{d}) = \{\mathbf{q} | d_g(\mathbf{p}, \mathbf{q}) = \bar{d}\} \quad (17)$$

where $d_g(\mathbf{p}, \mathbf{q})$ is the minimal geodesic distance along the surface \mathcal{S} . Accordingly to Kimmel et al. (1995), due to the continuity and smoothness of \mathcal{S} we can define a curvilinear abscissa u for $\alpha(u, \bar{d})$.

If we assume that the motion velocity is constant, V without loss of generality we can also assume that $V = 1$ then the geodesic distance $d = d_g(\mathbf{p}, \mathbf{q})$ coincides with time t and $\alpha(u, d) = \alpha(u, t)$ the equal distance contour evolution rule is:

$$\frac{\partial \alpha(u, t)}{\partial t} = \mathbf{N} \times \mathbf{T} \quad (18)$$

where \mathbf{T} is the unit vector tangent to α while \mathbf{N} is the unit vector normal to the surface \mathcal{S} .

To avoid point discontinuities the starting contour $\alpha(u, 0)$ is a small circle around the point \mathbf{p} and the projection of equation 18 on the x - y plane will be the velocity vector orthogonal to the actual front, V_N :

$$V_N = \mathcal{P}(\mathbf{N} \times \mathbf{T}) \cdot \mathbf{n} \quad (19)$$

where $\mathcal{P}()$ is the projector on the x - y plane and $\mathbf{n} = (n_x, n_y)$ is the unit normal to the planar curve in the x - y plane.

The evolution of the planar contour will then be:

$$\frac{\partial \mathcal{P}(\alpha(t))}{\partial t} = -V_N \mathbf{n} \tag{20}$$

In Fig. 5 there is a graphical representation of the aforementioned quantities.

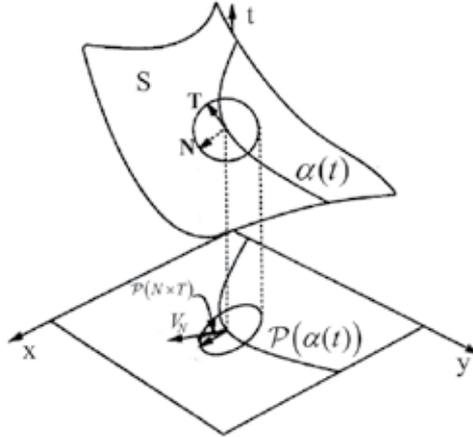


Fig. 5. The time evolution of variable α , and the definition of the moving front velocity V_N

Calling $p = \frac{\partial t(x,y)}{\partial x}$ and $q = \frac{\partial t(x,y)}{\partial y}$, accordingly to (Kimmel et al. (1995)) we can write:

$$\mathbf{N} = \frac{(-p, -q, 1)}{\sqrt{1 + p^2 + q^2}} \tag{21}$$

and:

$$V_N = \sqrt{\frac{(1 + q^2) n_x^2 + (1 + p^2) n_y^2 - 2pq n_x n_y}{1 + p^2 + q^2}} \tag{22}$$

Since the minimal geodesics are perpendicular to the equal distance contours on the surface, we can track the optimal trajectory starting from the destination point and, considering the equal distance contour α crossing it we have to move backward thanks to the orthogonality between the optimal path and the equal distance contour:

$$\eta = -(\mathbf{N} \times \mathbf{T}) \tag{23}$$

Using an arc length parametrization for the curve α we can write $r = \frac{\partial MDM_p}{\partial x}$, $l = \frac{\partial MDM_p}{\partial y}$ and using a normalization function

$$\psi(p, q, r, l) = \frac{p^2 + q^2 + 1}{\sqrt{l^2(1 + p^2) + r^2(1 + q^2) - 2pqr l}} \tag{24}$$

and the backtracking rule will then consist in tracking, from destination point \mathbf{q} over the surface S , the best path β so that:

$$\frac{\partial \beta}{\partial t} = \eta = \psi(p, q, r, l) [r(1 + q^2) - pql, l(1 + p^2) - pqr, pr + ql] \tag{25}$$

If there are moving obstacles along the path the MDM_p becomes multivalued and it could be possible that some points change their index along the optimal path, to account for this all distance values must be evaluated at a candidate point.

4.1 The Level-set approach

Moving obstacles introduce new boundary constraints and dealing with them could be very complex following the previously described approach, a simpler way could be using the level-set approach where the projection of α on the x-y plane is the zero level set $\phi = 0$ and ϕ will be negative in the interior and positive in the exterior of the zero level set.

$$\phi(\mathcal{P}(\alpha(t)), t) = 0 \tag{26}$$

A graphical representation of this process can be seen in Fig. 6.

The next step requires the definition of the evolution law for ϕ , that, accordingly to the chain rule can be written as:

$$\nabla\phi(\mathcal{P}(\alpha(t)), t) \frac{\partial\mathcal{P}(\alpha(t))}{\partial t} + \frac{\partial\phi(\mathcal{P}(\alpha(t)), t)}{\partial t} = 0 \tag{27}$$

Furthermore, accordingly to this level-set formulation, the in-plane normal vector can be written as $\mathbf{n} = \frac{\nabla\phi}{\|\nabla\phi\|}$. This relation together with the Lagrangian evolution equation:

$$\frac{\partial\phi(\mathcal{P}(\alpha(t)), t)}{\partial t} = -V_N \|\nabla\phi\| = \sqrt{\frac{(1+q^2)\phi_x^2 + (1+p^2)\phi_y^2 - 2pq\phi_x\phi_y}{1+p^2+q^2}} \tag{28}$$

This is the Eulerian formulation for curve evolution.

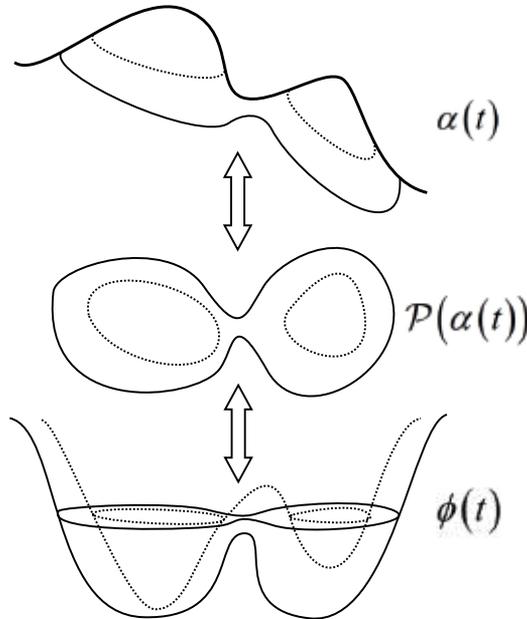


Fig. 6. Implicit representation of the projection of the α function through the level-set.

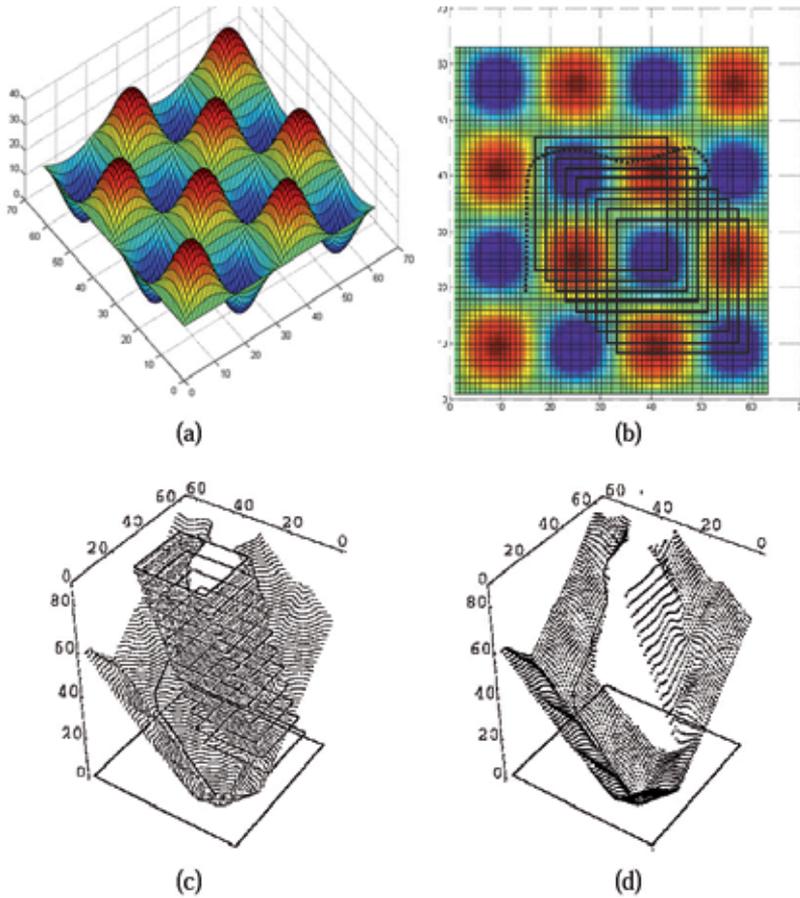


Fig. 7. (a) the terrain definition, (b) a squared object moves across the scene. (c) $\alpha(u, t)$ for different values of t representation (d) Back tracking from the final point toward the starting one along the estimated geodesic path

The steps for the best path definition can then be summarized as:

1. Evolve the PDE (eq.28) till the destination point is reached.
2. back-track along the optimal path till the origin.

In Fig. 7 there is the representation of level-set evolution, accordingly to moving objects, in an "egg-box" terrain where a square is moving across the scene.

Terrain traversability can be easily integrated in the presented approach since it is possible to define a map $F(x, y)$ describing the terrain traversability according to continuous values from 1 (optimal traversability) to 0 (impassable) and the optimal path can be obtained just changing eq. 18 into

$$\frac{\partial \mathcal{P}(\alpha(t))}{\partial t} = -F(x, y) V_N \mathbf{n} \tag{29}$$

Further velocity dependencies, e.g. from the vehicle weight reduction due to fuel combustion, can be accounted introducing a time dependance in V_N . More complex conditions can be

easily added just acting on V_N , e.g. if the Robot has a specific shape the velocity can reflect the ability to walk through a narrow path or to turn at a specific point. In some cases the $F(x, y)$ correcting function is called viscosity for its similarity to fluid property that can prevent or slow down motion in particular circumstances.

5. Surface reconstruction from cloud of points

The temporal evolution of a volumetric implicit function can also be fruitfully applied to reconstruct 3D surfaces from large set of unorganized sample points. The evolving front can be thought as the surface that separates two different fluids obeying specific fluid dynamics laws. One remarkable feature of this approach is its ability to model complex topologies using a set of intuitive tools derived from fluid physics: Global and local surface descriptors are used allowing the parallelization of the algorithm on different object parts working with different resolutions and accuracies. The problem of building surfaces from unorganized sets of 3D points has recently gained a great deal of attention. In fact, in addition to being an interesting problem of topology extraction from geometric information, its applications are becoming more and more numerous. For example, the acquisition of large sets of 3D points is becoming easier and more affordable using, for example, 3D-scanners (*Registration and fusion of intensity and range data for 3d modelling of real world scenes* (2003)). The presented approach is particularly important in applications where objects are better described by their external surface rather than by simple unorganized data (clouds of points, data slices, etc.). For example, in medical applications based on Tomography scans or NMRs it is often necessary to visualize some specific tissues such as the external surface of an organ starting from the acquired 3D points. This can be achieved by selecting the points that belong to a specific class (organ boundary, tissue, etc.) and then generating the surface from their interpolation. In most cases the definition of this surface is an ill-posed problem as there is no unique way to connect points of a dataset into a surface, therefore it is often necessary to introduce constraints for globally or locally controlling the surface behavior. As a matter of fact, the resulting surface often turns out to exhibit a complex topology due to noise in the acquired data or ambiguities in the case of non-convex objects Hoppe (1994). In order to overcome such problems, surface wrapping algorithms need to incorporate specific constraints on the quality of the data fitting (surface closeness to the acquired points), on the maximum surface curvature and roughness, on the number of resulting triangles, etc. The existing surface reconstruction methods can be classified into two broad categories: the former describes the surface as an implicit function while the latter describes it in an explicit form. Explicit (boundary) representations describe the surface in terms of point connections, and traditional approaches are based on Delaunay triangulation, Voronoi diagrams (Amenta et al. (1998)) or NURBS (Piegel & Tiller (1996)).

Here we will show how Level-set technique together with Volume of Fluid and Fluid Dynamics equations can be fruitfully applied to obtain a wrapping surface with intuitive parameters tuning. To evaluate quality of the final reconstruction an important parameter that can also be used as stopping condition for the surface evolution is an energy functional defined as:

$$E(\Gamma) = \left[\int_{\Gamma} d^p(\mathbf{x}) ds \right]^{1/p}, \quad 1 \leq p \leq \infty \quad (30)$$

where Γ represents the surface, \mathbf{x} is the set of 3D points, d is the euclidean distance between the point \mathbf{x} and the closest point on the surface Γ and $1 \leq p \leq \infty$ is the order of the functional distance: for $p = 2$ it gives a root mean square error while for $p = 1$ we just focus on the point with the maximum distance from Γ . Looking for the surface that minimizes the functional (30) is similar to enveloping the data set with a membrane with certain elastic properties and have this membrane to evolve in time until it comes to rest. This time evolution paradigm can be described as:

$$\frac{\partial \Gamma}{\partial t} = - \left[\int_{\Gamma} d^p(\mathbf{x}) ds \right]^{\frac{1}{p}-1} d^{p-1}(\mathbf{x}) \left[\nabla d(\mathbf{x}) \cdot \mathbf{n} + \frac{1}{p} d(\mathbf{x}) \kappa \right] \mathbf{n} \quad (31)$$

Euler-Lagrange equation can be used to find its minimum:

$$d^{p-1}(\mathbf{x}) \left[\nabla d(\mathbf{x}) \cdot \mathbf{n} + \frac{1}{p} d(\mathbf{x}) \kappa \right] \quad (32)$$

where \mathbf{n} points accordingly to the evolving point direction, κ indicates the mean curvature, $\nabla d(\mathbf{x}) \cdot \mathbf{n}$ represents the surface attraction term and $d(\mathbf{x}) \kappa$ is the surface tension. An important consideration is that function $d(\mathbf{x})$ will automatically impose that the final surface is more flexible where the 3D points dataset is denser and more rigid in low density regions: this will allow smoother regions where we have less samples and accurate fitting in highly sampled regions. Fluids, with their properties like surface tension and viscosity represent an optimal modelization for the warping surface providing us with a flexible tool for this goal. Also in this case we present a Level-set approach, in particular we will adopt an implicit function ϕ over the spatial 3D domain so that $\phi(\mathbf{x}, t) = 0$ represent the warping surface $\Gamma(t)$ at time t , for the initialization we use $\phi(\mathbf{x}, t) |_{t=0} = D$ where D is the signed distance of \mathbf{x} from $\Gamma(t)$ (positive values indicate external points while negative ones represent internal points). One possible approach to make surface Γ to converge towards the cloud of points is, accordingly to Zhao et al. (2001), to use the fluids convection defining a velocity field $\mathbf{v}(\mathbf{x})$ accordingly to:

$$\frac{\partial \Gamma(t)}{\partial t} = \mathbf{v}(\Gamma(t)) \quad (33)$$

In particular we adopted, as a steering potential field, the distance function $d(\mathbf{x})$ so that the velocity field \mathbf{v} is oriented towards the minima of the potential field: $\mathbf{v} = -\nabla d(\mathbf{x})$. So, accordingly to eq. 8, the level-set evolution in this case can be defined as:

$$\frac{\partial \phi}{\partial t} = \nabla d(\mathbf{x}) \cdot \nabla \phi \quad (34)$$

A major computational drawback comes from the fact that the ϕ function at each time iteration becomes more and more compressed presenting a growing slope across the zero level-set which could introduce numerical imprecisions. To solve this problem some solutions are present in literature, for example a level-set re-initialization in a region around the zero level: this is particularly useful for Fast Marching methods which, as described in section 3 for the Narrow Band approach Sethian (1999). Other approaches, instead, suggest a different implementation for the Hamilton Jacobi PDE like Gomes & Faugeras (1999). In particular in

Marcon et al. (2008) a quite different interpretation for the level-set function is presented that aims to conciliate Level-set upgrade, narrow band approach, discretization issues and Volume of Fluid (VoF) method: in particular, the VoF method (see Noh & Woodward (1976)) has been formulated in a variety of forms and re-introduced under different names such as the "cell method" and the "partial fractions method". The key idea behind this technique is to define a fixed computational grid and assign to each grid cell a value that describes the relative proportions of two materials contained in that cell. Our particular modeling metaphor is based on two immiscible fluids of opposite densities $\rho = 1$ and $\rho = -1$, which identify the surface's outside and inside, respectively. More specifically, cells completely filled with outer fluid take on the value +1 while cells that are only filled with inner fluid will take on the value -1. When filled with a mix of both fluids, the cell is assigned an intermediate value in the $] - 1, 1[$ range, which measures the relative density of the mix (i.e. a value of zero indicates the same amount, half and half of the two fluids inside the cell). As a first step we define a bounding box that completely encloses the point cloud. We assume that this box is completely filled with inner fluid and that the whole space outside of the box is filled with outer fluid. We define the rules of evolution so that both fluids are set in motion by the presence of data points, which act as attractors. This fluid migration takes place in compliance with laws of conservation. According to such rules, the front starts propagating inward (towards the cloud of points) and stops only when the inner fluid is confined "inside" the cloud of points and the interface between the two fluids wraps the point-cloud. At each iteration step we will replace the value of points on the domain boundary \mathcal{B} with 1 (this is equivalent to place outer fluid sources on the domain boundary), the cells with values greater than 1 will be clipped with value '1' while cells with values lower than -1 will be clipped to -1: this will be equivalent to remove excess fluids from cells without any need to introduce further fluid sources or drains. Two main contribution to fluid motion are considered: convective and diffusive: the first one is important for steering the zero level-set toward the cloud of points while the second has an important role in smoothing evolving front. Accordingly to the Gauss theorem when no sources or drains are present the fluid density will follow eq. (35):

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (\mathbf{v}\rho) + \nabla \cdot (\xi \nabla \rho) \quad (35)$$

where ρ indicates the fluid density, $\mathbf{v}\rho$ is the convective term indicating the amount of fluid per volume unit that is transported by the velocity field while $\xi \nabla \rho$ represents the diffusive part where ξ is the diffusive term (further details in (Gueyffier et al. (1999))). In this approach a point will be assigned a velocity that is proportional to the vector \mathbf{v} that joins that point to the closest data point. This way data points act as attractors for both internal and external fluids and, therefore, for the evolving front: each sample point represents a central vector field. Even if this attraction field has no physical equivalent, it is physically consistent thanks to its irrotational characteristic (see Marcon et al. (2008) for further details). This fact guarantees algorithmic convergence. In Figures 8 and 9 it is possible to see the direction and intensity of the steering field toward the cloud of points and different level-set curves; in particular in the magnification of zones A and B is possible to see how the steering field automatically stops external liquid from entering the object (all the internal arrows point outside) while, in the convex part, the steering field force the outer fluid to enter the convex region. In Fig. 10 we represent the final state for the wrapping of a two dimensional cloud of points.

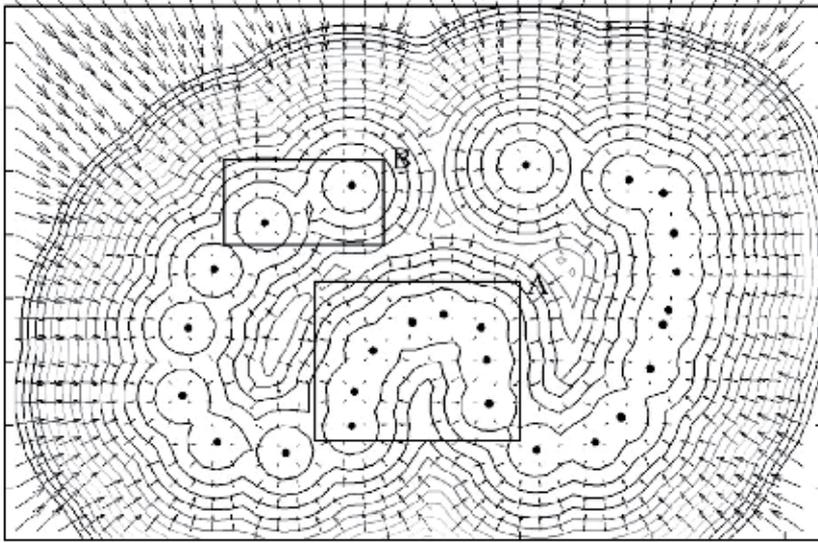


Fig. 8. a 2D non-convex example, arrows indicate the steering velocity field toward the cloud of points and different level-sets are also represented. Zones A and B are zoomed in Fig.9

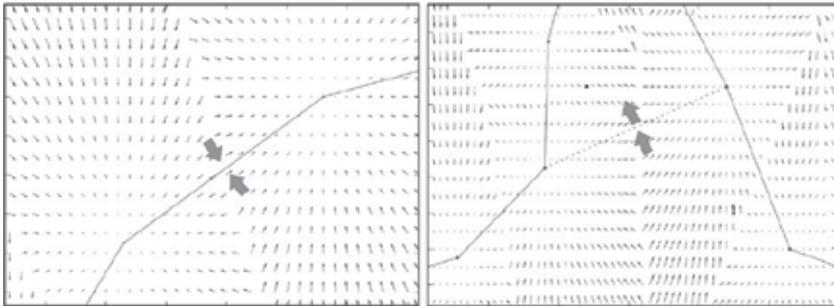


Fig. 9. zoom of zone A and B of Fig. 8 in particular different behavior of the steering field prevents external fluid from entering the convex part in B while favor its penetration in A (nonconvex region)

Further aspects from fluid dynamics can be fruitfully applied, e.g. the diffusive parameter ζ in eq. 35 is a delicate parameter as it is strictly connected to fluid viscosity. High values of ζ result in a smooth but often inaccurate surface while low values result in a surface that strictly honors the data but is sensitive to acquisition noise. As a matter of fact for laser scanner acquisitions the portions of the object with high samples density (usually due to multiple acquisitions from different viewpoints) requires high accuracy and/or present complex topologies, while smoother or less critical areas usually exhibit a lower sample density. In order to account for that, we can establish an inverse dependency between the diffusive constant ζ and the local point density:

$$\zeta = \frac{\alpha}{1 + \rho_p} \tag{36}$$

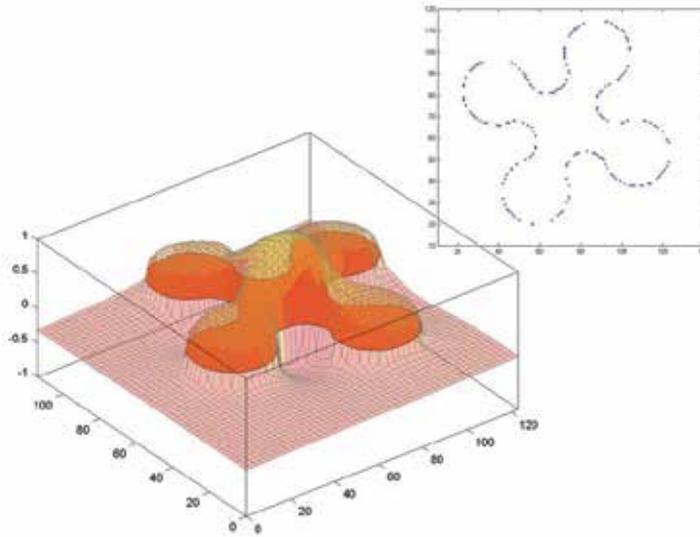


Fig. 10. the wrapping of a set of 2D points (top right), the zero level-set at the end of the evolution represent the joining line, inner points have positive ψ while external ones have negative values

where α is a parameter that indicates the estimated noise in the point location (the higher the value the smoother the surface). The density ρ_p at a given point is defined as the number of samples that fall into a cube of volume V/N around that point, where V is the volume of the point cloud (approximated by the volume of the smallest parallelepiped that contains the point cloud) and N is the total number of samples. This means that if all the samples would be uniformly distributed in the considered volume (like in a cubic crystal) in each cubic cell of volume V/N there would be just one sample. In our case, since we have to define the ρ_p variable for each voxel, we define it as the sum of all the points falling inside a cube of volume V/N centered at the voxel center. In surface regions with a low point density the diffusive constant ζ corresponds to a smooth surface, while in high-density zones ζ becomes very low, to guarantee that the surface will strictly honor the sample points.

The level-set evolution equation is then the following:

$$\begin{cases} \frac{\partial \phi(\mathbf{x}, t)}{\partial t} = \nabla \cdot (\mathbf{v}\phi(\mathbf{x}, t)) + \zeta \nabla^2 \phi(\mathbf{x}, t), & t \geq 0 \\ \phi(\mathbf{x}, t) = +1, & t \geq 0, \mathbf{x} \in \mathcal{B} \\ \phi(\mathbf{x}, 0) = -1 & \mathbf{x} \notin \mathcal{B} \end{cases} \quad (37)$$

where the second row indicates that the boundary \mathcal{B} of the considered domain is always kept equal to +1 assuming that those points are sources of external fluid. The third row indicates that at the starting point all the points inside the domain are filled with the internal fluid. In Fig. 11 the level-set evolves toward the cloud of points which is a laser scanner acquisition of a plaster wolf.

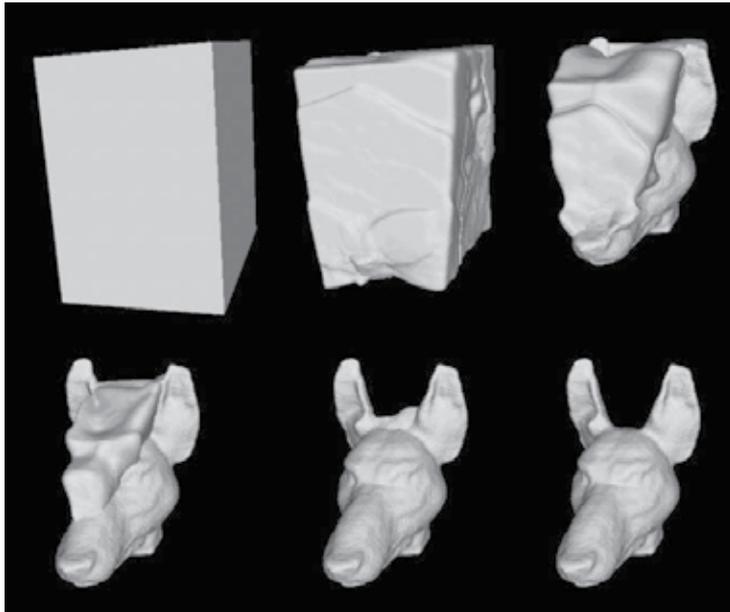


Fig. 11. Evolution of the zero level-set toward a 3D cloud of points acquired by a laser scanner. The level-set starting from the boundary of the considered volume (top left) evolves towards the final cloud of points.

5.1 Turbulence and convex regions

In some particular cases, where high curvature is present in nonconvex regions with a small entrance, external fluid could be inhibited from entering (e.g. consider the cloud of points of Fig. 12).

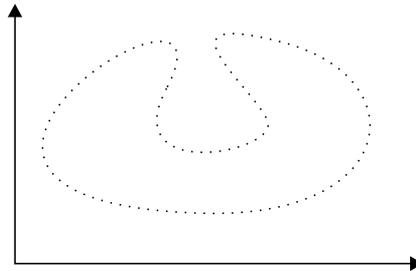


Fig. 12. a set of points presenting a high curvature region within a concavity

For cases like these a further help could come from vorticity. In particular, since our steering field is based on the distance from the closest point it will be a conservative irrotational vector field and when the outer fluid will approach the concavity mouth it will be pushed to its borders without the opportunity to get inside (we underline that this condition is different from the one in Fig. 8 where the concavity of the region A does not present high curvature and external fluid is correctly steered inside the region). A possible solution consists in introducing a rotational component in the steering field, this could be obtained adding periodically a small random displacement for each point around its original position: it will result in a vorticity

contribution: $\zeta = \nabla \times \mathbf{v}(\mathbf{x})$ the resulting effect is that inner fluid in convex regions is pushed inside and outside with the same probability while in nonconvex regions inner fluid that went inside stays there and progressively fills the region. Particular care must be placed in introducing the amount of random points displacement since convergence of the algorithm is no more guaranteed by the conservative vector field.

6. Image processing and viscous noise removal

Fluid dynamics applied together with level-set method can also provide a powerful tool for noise removal in complex images (i.e. when noise is not simply removable using intensity thresholds of histogram analysis), a Total Variation denoising approach is presented in (Osher & Fedkiw (2002)) where the regularization algorithm is based on this PDE:

$$Lu = \lambda Ru, \quad (38)$$

where u is the image represented as a 2D function ($u(x, y)$) that indicates intensity value for each point; R is the regularization operator and L is the time-space operator. A very common energy functional $E(u)$ whose minimization provide the denoised image is the Mumford-Shah multiscale segmentation model:

$$\min_u E(u) = \frac{1}{2} \int_{\Omega} (u - u_0)^2 dx + \mu \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx + \nu \mathcal{H}^2(\Gamma) \quad (39)$$

where u_0 is the given noisy image, Ω is the image function domain and μ is a weighting coefficient for the average smoothness of regions divided by the contour Γ . ν is the weight coefficient for the total contour length expressed by the Hausdorff measure \mathcal{H} .

The minimization of $E(u)$ will provide an piecewise-smooth approximation of the initial u_0 image, Γ has the role of approximating the edges in the image u_0 and u will be smooth only in $\Omega \setminus \Gamma$; Accordingly to the level-set formulation the implicit function will be ϕ and in particular $\phi(\mathbf{x}) = 0$ represents the contour Γ separating region of positive values of ϕ from the region of negative values. In particular $u^+(\mathbf{x})$ will be the intensity function representing values of the region where $\phi(\mathbf{x}) > 0$ while $u^-(\mathbf{x})$ will represent the intensity of points in the $\phi(\mathbf{x}) < 0$ region. Eq. 39 then become:

$$\begin{aligned} \min_u E(u^+, u^-, \phi) &= \int_{\Omega} (u^+ - u_0)^2 H(\phi) dx + \int_{\Omega} (u^- - u_0)^2 (1 - H(\phi)) dx + \\ &+ \mu \int_{\Omega} |\nabla u^+|^2 H(\phi) dx + \mu \int_{\Omega} |\nabla u^-|^2 (1 - H(\phi)) dx + \nu \int_{\Omega} |\nabla H(\phi)| \end{aligned} \quad (40)$$

Where $H()$ represents the Heaviside step function and the last term, $\int_{\Omega} |\nabla H(\phi)|$ is the integral in the sense of geometric measure, i.e. for a 2D case, is the length of contour Γ .

The minimization can be done iteratively on u^+ , u^- and on ϕ , in particular, minimizing with respect to u^+ and u^- , i.e. deriving eq. 40, will give:

$$u^+ - u_0 = \mu \nabla^2 u^+ \quad (41)$$

and

$$u^- - u_0 = \mu \nabla^2 u^- \tag{42}$$

which correspond to the diffusive part inside the borders defined by Γ , since $\frac{\partial u^{+,-}}{\partial n} = 0$ (i.e. no diffusion will take place across the boundary keeping edges sharp). While, deriving with respect to ϕ will give:

$$\frac{\partial \phi}{\partial t} = \delta_\epsilon(\phi) \left[-|u^+ - u_0|^2 - \mu |\nabla u^+|^2 + |u^- - u_0|^2 + \mu |\nabla u^-|^2 + \nu \nabla \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right] \tag{43}$$

Accordingly to a steepest descendent approach. The δ_ϵ indicates that the delta of Dirac, in a discrete environment, must be calculated in a narrow band and, to prevent discontinuities u^+ has to be extended, in the narrow band, for $\phi < 0$ values and also u^- has to be extended in the narrow band for $\phi > 0$. A possible approach to these extension could be the "Ghost fluid Method" (Fedkiw et al. (1999)) where the interface is treated as a moving boundary. Solving the governing equations (Terashima & Tryggvason (2009)) with the extension of the discontinuous variables across the fluids interface, (for real CFD applications this is usually entropy), will reduce smearing of discontinuities providing smooth convergence to the final image without oscillations. In Fig. 13 is represented how the contour between the two image parts is steered accordingly to the Ghost Fluid Method in the narrow band while in Fig. 14 it is possible to see the evolution of the contour across the noisy image: diffusive terms smooth uniform regions reducing noise while the contour evolution smoothly warp the objects.

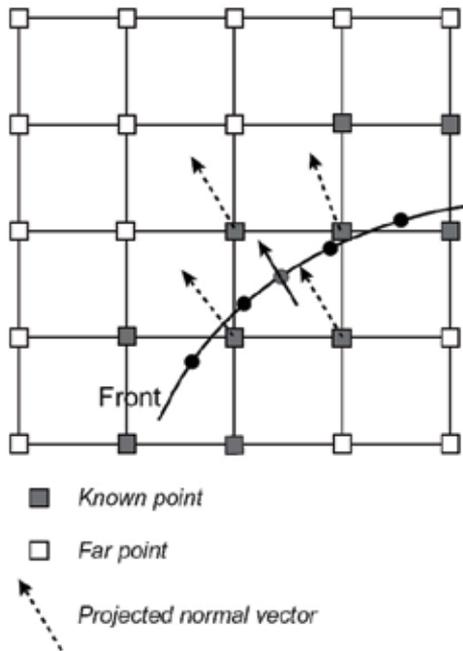


Fig. 13. update procedure in the narrow band points

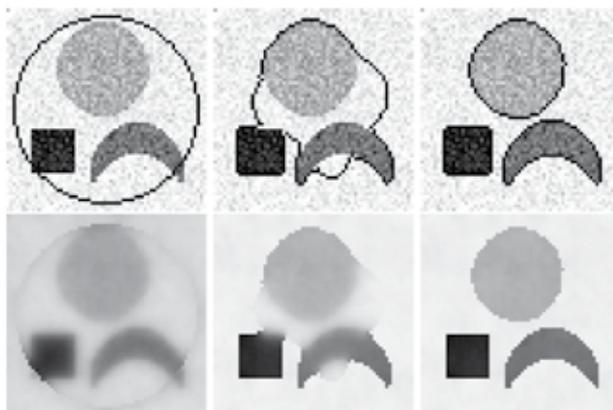


Fig. 14. zero level set evolution accordingly to the Mumford-Shah energy minimization for object segmentation and image denoising

7. Conclusion

In this chapter we have shown some significant implementations of Computational Fluid Dynamics equations in completely different contexts ranging from image denoising to optimal path research for moving robots. In recent literature it is also possible to find different applications to many other research fields: The main reason for this large diffusion and reuse of CFD equations can be mainly addressed to:

- the actual deep insight and knowledge both into the theoretical and computational aspect of fluids.
- The accuracy and stability of modelling equations that can grant convergence in a wide variety of applications.
- Versatility and adaptability of different models from multiphase, to viscous or compressible fluids that provide the user with many intuitive and easy-to-tune parameters allowing an easy adaptation to different contexts.
- The close relation to physical quantities that in many cases can automatically provide the minimum energy solution.

All these aspects concurred in making CFD a science that goes well beyond fluids analysis providing a self consistent, well known, spread of tools to find fruitful and unlimited applications in engineering, computer science and other scientific research fields.

8. References

- Adalsteinsson, D. & Sethain, J. A. (1995). A fast level set method for propagating interfaces, *Journal of Computational Physics* 118(2): 269–277.
- Amenta, N., Bern, M. & Eppstein, D. (1998). The crust and the β -skeleton: combinatorial curve reconstruction, *Graph. Models Image Process* 60(2): 125–135.
- Biben, T., Klaus, K. & Chaouqi, M. (2005). Phase-field approach to three-dimensional vesicle dynamics, *Phys. Rev. E* 72(4).

- Bridson, R. (2003). *Computational aspects of dynamic surfaces*, PhD thesis, University of British Columbia.
- Davis, S. H. (2001). *Theory of solidification*, Cambridge University Press.
- Dobrushin, R. L., Kotecký, R. & Shlosman, S. (1992). *Wulff Construction: A Global Shape from Local Interaction*, American Mathematical Society.
- Fedkiw, R. P., Aslam, T., Merriman, B. & Osher, S. (1999). A nonoscillatory eulerian approach to interfaces in multimaterial flows (the ghost fluid method), *Journal of Computational Physics* 152(2): 457–492.
- Gomes, J. & Faugeras, O. (1999). Reconciling distance functions and level sets, tech. rep. 3666, *Technical report*, INRIA Sophia Antipolis.
- Gueyffier, D., Li, J., Nadim, A., Scardovelli, R. & Zaleski, S. (1999). Volume-of-fluid interface tracking with smoothed surface stress methods for three-dimensional flows, *J. Comput. Phys.* 152: 423–456.
- Hennessy, M. G. (2010). *Liquid Snowflake Formation in Superheated Ice*, PhD thesis, University of Oxford.
- Hills, R. N. & Roberts, P. H. (1993). A note on the kinetic conditions at a supercooled interface, *Int. Comm. Heat Mass Transfer* 20: 407–416.
- Hobbs, P. V. (2010). *Ice Physics*, Oxford University Press.
- Hoppe, H. (1994). *Surface reconstruction from unorganized points*, PhD thesis, University of Washington.
- Kimmel, R., Amir, A. & Bruckstein, A. M. (1995). Finding shortest paths on surfaces using level sets propagation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17(6): 635–640.
- Kimmel, R., Kyriati, N. & Bruckstein, A. M. (1998). Multivaueled distance maps fro motion planning on surfaces with moving obstacles, *IEEE Trans. on Robotics and automation* 14(3): 427–436.
- Latombe, J. C. (1991). *Robot Motion Planning*, Kluwer.
- Libbrecht, K. G. (2005). The physics of snow crystals, *Rep. Prog. Phys.* 68, Norman Bridge Laboratory of Physics, California Institute of Technology.
- Losasso, F., Gibou, F. & Fedkiw, R. (2004). Simulating water and smoke with an octree data structure, *ACM Transactions on Graphics* 23(3): 457–462.
- Marcon, M., Piccarreta, L., Sarti, A. & Tubaro, S. (2008). Fast pde approach to surface reconstruction from large cloud of points, *Computer Vision and Image Understanding* 112: 274–285.
- Markstein, G. H. (1951). Experimental and theoretical studies of flame front stability, *J. Aero. Sci.* 18: 199.
- Noh, W. & Woodward, P. (1976). A simple line interface calculation, in Springer-Verlag (ed.), *Proceedings of 5th International Conference on Fluid Dynamics*.
- Osher, S. & Fedkiw, R. (2002). *Level Set Methods and Dynamic Implicit Surfaces*, Springer-Verlag.
- Piegel, L. & Tiller, W. (1996). *The NURBS Book*, second ed. edn, Springer-Verlag, Berlin.
- Pimpinelli, A. & Villain, J. (1999). *Physics of Crystal Growth*, Cambridge University Press.
- Registration and fusion of intensity and range data for 3d modelling of real world scenes* (2003). Vol. Fourth International Conference on 3-D Digital Imaging and Modeling, 3DIM Proceedings.

- Sethian, J. A. (1989). A review of recent numerical algorithms for hypersurfaces moving with curvature-dependent speed, *Journal Differential Geometry* 31: 131–161.
- Sethian, J. A. (1999). *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*, Cambridge University Press.
- Shastri, S. S. & Allen, R. M. (1998). Method of lines and enthalpy method for solving moving boundary problems, *International Communications in Heat and Mass Transfer* 5(4): 531–540.
- Terashima, H. & Tryggvason, G. (2009). A front-tracking/ghost-fluid method for fluid interfaces in compressible flows, *Journal of Computational Physics* 228: 4012–4037.
- Zhao, H., Osher, S. & Fedkiw, R. (2001). Fast surface reconstruction using the level set method, *Proceedings of IEEE Workshop on Variational and Level Set Methods in Computer Vision (VLSM 2001)*.

Fluid Dynamics in Space Sciences

H. Pérez-de-Tejada
Institute of Geophysics, UNAM
Mexico

1. Introduction

Much of what it is understood in nature and that it is also inherent to our common activities can be appropriately interpreted as representing evidence of a collective response that substantiates the basis of fluid dynamics. Such is the case for the behavior for a group of particles or individuals that interact with each other as they follow common trajectories. Under most circumstances their interaction takes place across distances (mean free path) that are much smaller than the size of the region where they move and, as a whole, they exhibit properties (density, speed, temperature) that represent average local values of the conglomerate. Besides the application of this view to standard problems in physics and engineering it has been intuitively suggested that it could also be relevant to describe the motion of the solar wind which at large distances from the sun travels with its particles barely executing any collisions among them. Even though the solar wind is, in fact, a (collisionless) ionized gas it maintains a collective response as it expands through interplanetary space and interacts with the planets of the solar system. The text below describes the manner in which the fluid dynamic interpretation of the solar wind was initiated and how it has expanded to examine its interaction with the planets of the solar system.

2. The solar wind as a continuous expanding gas

The over one million degree temperature of the solar corona is significantly larger than the nearly six thousand degrees of the solar surface and thus provides the energy source that ultimately leads to its strong outward expansion. In fact, the large amounts of energy that are delivered to the solar corona from the solar interior can be mechanically used for producing the solar wind. The prediction and theoretical description of this phenomenon was provided by E. N. Parker¹ at the University of Chicago who over 50 years ago devised a fluid dynamic interpretation of the manner in which the ionized coronal gas (plasma) is forced to expand outward reaching speeds of the order of 300-400 km/s. A remarkable aspect of this concept is that the solar wind is predicted to be an outflowing continuum gas that expands away from the sun as its density decreases to values that by the earth's orbit are very small ($\sim 10 \text{ cm}^{-3}$). As a result, collisions among its particles (mostly protons and electrons) become extremely rare and the solar wind rapidly becomes a collisionless plasma with an effective mean free path for the collisions of its particles that is comparable to one astronomical unit (this is the distance between the earth and the sun). Despite this constraint observations made with various spacecraft in the interplanetary medium have confirmed the existence of the solar wind and its overall collective response when it interacts with the

planets of the solar system. Much research has been conducted to investigate the processes that allow the solar wind to explain its collective behavior as it moves through space, interacts with the magnetic fields and the ionospheres of planets, and reaches its outer boundary (the heliopause) located at nearly 80 astronomical units.

A fundamental property of the solar wind is that it is rapidly accelerated to reach supersonic speeds and that the process that produces it was incorporated by applying conditions similar to those employed in the fluid dynamic de-Laval nozzle theory. The latter problem serves in rocket engine turbines and is designed to produce supersonic speeds in a streaming flow that is subject to a pressure gradient along its direction of motion. When a subsonic flow is compressed by decreasing the cross-sectional area of the region where it moves its speed increases but remains subsonic. In order to make the flow supersonic it is necessary to introduce a nozzle (converging/diverging) geometry so that the gas first becomes compressed and later expands upon streaming through a "throat" or region of minimum cross-sectional area as it is depicted in Figure 1. The reason here is that when a flow moves into the region of smaller cross-sectional area the imposed pressure gradient can increase the flow speed only up to the sound speed, but a further increase requires an

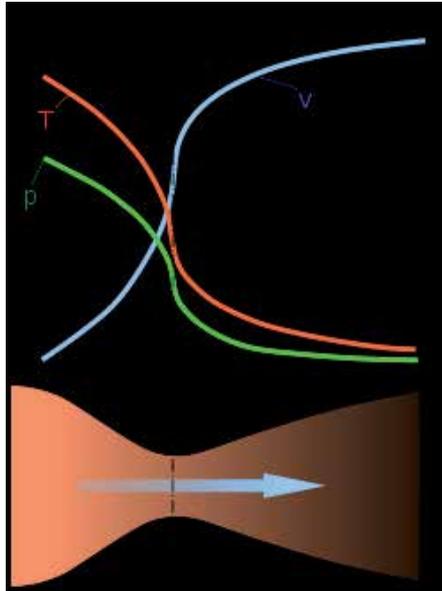


Fig. 1. Scheme of flow through a de-Laval nozzle. The speed profile (blue curve) shows how the flow reaches the sound speed at the throat and becomes supersonic afterwards. The temperature and pressure profiles (labeled T and P) indicate how the gas cools off and expands when it becomes supersonic ².

expansion, that is, that the flow proceeds in a diverging geometry. These considerations are better illustrated by the de-Laval equation in hydrodynamics²:

$$ds/s = (v^2/c^2 - 1) dv/v \quad (1)$$

where v is the speed of the flow, s is the cross-section of a duct or region where the flow moves and c is the speed of sound (dv and ds are changes of those variables along the direction of motion). This equation is derived from the continuity and the momentum

equations of the flow and states the correlation that there exists between a change in the speed of the flow and that of its cross-sectional area. Most notable is that when the duct is converging (negative $ds < 0$ values) and the speed of the flow increases (positive $dv > 0$ values) then v must be less than c so that the value of the parenthesis is negative. Such conditions correspond to subsonic flows where the Mach number $M = v/c$ is smaller than one. When $v = c$, that is if $M = 1$, we now have $ds = 0$ and the duct should stop converging. Finally, if v is to exceed c (supersonic flows) then ds must increase ($ds > 0$) for positive changes in the value of the flow speed ($dv > 0$), thus implying that the duct should now be diverging. Such flow conditions of the de-Laval nozzle require, in addition, a large pressure difference along the direction of motion. If that difference is not large, implying that the speed of the flow is slow and does not reach the sound speed at the narrowest part of the duct, then it will only speed up the flow as a Venturi tube achieving only subsonic speeds.

The coordinated dependence between changes in the speed of the flow and those in the cross-section of the duct that was derived from equation 1 for the converging/diverging geometry of Figure 1 is a remarkable property that characterizes subsonic and supersonic flows. Its origin lies on the simultaneous consideration of pressure forces that balance the momentum convective term in the momentum equation of the flow from which Equation 1 was derived. On the other hand, even though the presence of the gravitational force and the outward radial expansion of the solar wind in the solar corona leads to flow conditions that are entirely different from those of the two-dimensional flow that streams in the duct of Figure 1, it is significant that a relationship similar to that in Equation 1 can also be derived by employing the continuity and the momentum equations that apply to the solar wind. Following Parker it is possible to replace equation 1 by ³:

$$[2 - MG/(r c^2)] dr/r = (v^2/c^2 - 1) dv/v \quad (2)$$

where M is the solar mass and G the gravitational constant in this relation which holds along the distance r away from the sun. This equation was derived by considering that the gravitational force constricts or chokes the solar wind in a manner similar to the converging section of the de-Laval nozzle in Equation 1; that is, the gravitational force replaces the geometric throat that is not present in the region where the solar wind expands. All in all it is very significant that in the derivation of Equation 2 the gravitational force plays a role similar to that of the geometric throat in Equation 1. The similitude between equations 1 and 2 serves to substantiate the fluid dynamic interpretation of the motion of the solar wind. For example, if the expansion velocity of the solar wind increases outward ($dv > 0$) then for subsonic speeds ($v < c$) we need $MG/(c^2r) > 2$ so that both sides of the equation become negative, and for supersonic speeds ($v > c$) we need $MG/(c^2r) < 2$ in order to make them positive. These conditions imply that there will be subsonic speeds at small r values and supersonic speeds at large r values. The speed of the solar wind will reach the speed of sound at the critical distance $r_c = MG/2c^2$ implying that the solar wind will gradually become supersonic as it moves away from the sun. Measurements conducted with various spacecraft have confirmed this result and, in fact, its observed speed is supersonic being about eight times the speed of sound. It should be noted here that the solar wind will become supersonic at a distance between three and four solar radii, or about two million kilometers above the sun surface. In addition, since the critical distance decreases with increasing the temperature T of the coronal gas, that is by increasing the speed of sound $c = \sqrt{\gamma KT/m}$ where γ is the specific heats of the gas, K the Boltzmann constant and m the mass of the gas particles, then it may occur that the critical distance becomes located below one solar radius if the gas temperature is very high (larger than four million degrees). In that case the solar wind will not reach the sound speed but will

remain subsonic as it expands. Under standard conditions the solar wind becomes supersonic and flows as a continuum gas whose motion is also related to the magnetic field fluxes that are brought up from the sun. As it will be shown below the effects of motion of the solar wind particles in the presence of a magnetic field can be better described by considering the interaction of the solar wind with the earth's magnetic field.

3. The gas dynamic analogue of the solar wind around the earth's magnetosphere

An important contribution to the continuum flow interpretation of the solar wind became available from studies directed to describe its interaction with the earth's magnetic field. That field is, in fact, the effective obstacle that the solar wind encounters as it approaches the earth since its ionized populations (protons and electrons) can be directly influenced by the planetary magnetization. The effect of their interaction is that as a whole the solar wind compresses and deforms the distribution of the earth's magnetic field which becomes bounded in the dayside and is deviated in the nightside forming a magnetic tail that extends far downstream from the earth as it is illustrated in Figure 2. As a result the earth's magnetic

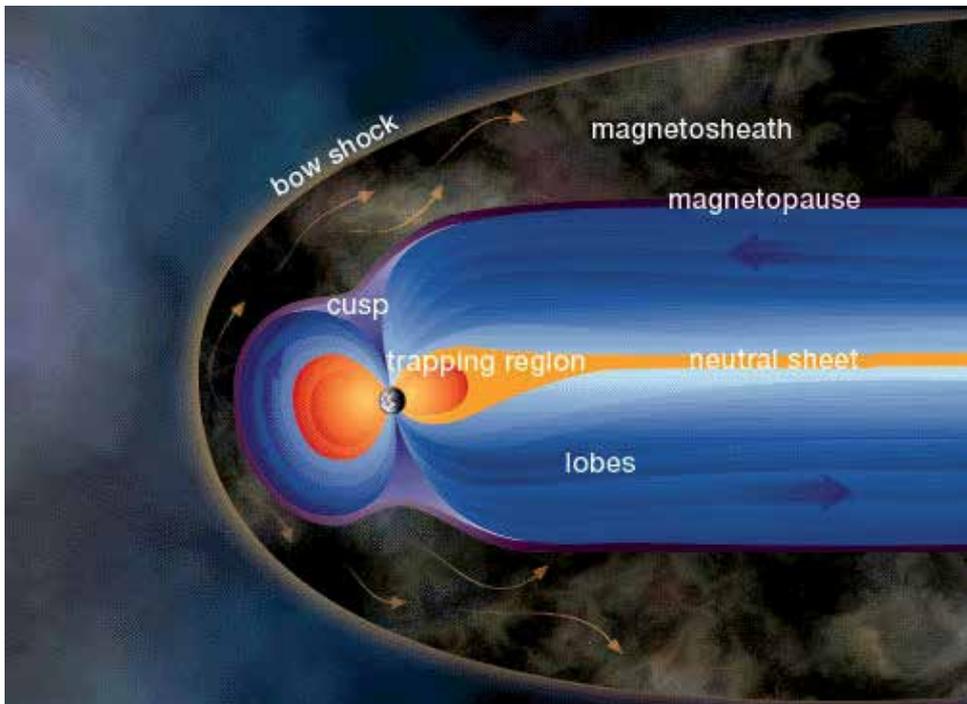


Fig. 2. A schematic view of the earth's magnetosphere formed by the supersonic solar wind that compresses the magnetic field and streams around it as an obstacle after crossing a bow shock front ^{3,4}.

field becomes present only within a large confined region in the form of a cavity (called the magnetosphere) whose outer boundary may reach seven earth radii in the direction of the sun. That boundary (the magnetopause) results from a local pressure balance between the solar wind pressure and the compressed pressure of the earth's magnetic field, and implies a

collective response despite the fact that there are no particle collisions in the solar wind population. The mechanism that replaces collisions to allow the transport of information upon encountering the earth's magnetic field is the gyration that the protons and electrons of the solar wind execute when they enter the region occupied by the magnetic field. The process is related to the magnetic (Lorentz) force produced by the magnetic field and the velocity of the particles, and that drives them in a circular (Larmor) motion in a plane transverse to the magnetic field direction. Implied by this motion there is an electric current along the boundary of the magnetosphere that is generated by the different trajectory of the protons and the electrons of the solar wind as they interact with the earth's magnetic field. Their accelerated motion around the magnetic field lines is important because it produces an induced magnetic field whose presence influences the motion of other particles. Thus, despite the fact the solar wind particles are not interacting through collisions with each other there is an ample communication among them which is produced by their motion around the earth's magnetic field lines and that, in the end, justifies its collective response to the presence of that field.

This general idea to view the encounter of the solar wind with earth's magnetic field led J. Spreiter ⁴ and M. Dryer ⁵ to apply a fluid dynamic model available from studies of the interaction of a streaming object in the earth's atmosphere. The purpose of their studies was to describe the manner in which the solar wind encounters the earth's magnetic field and then streams around the magnetospheric cavity. The gas dynamic description of the fluid properties of the solar wind as it reaches the earth's magnetic field and then streams around the magnetosphere is dominated by various features. Most important is a bow shock produced by the supersonic speed of the solar wind and that is located upstream from the magnetosphere in agreement with the common response of a supersonic flow that encounters an obstacle. As it occurs in standard fluid dynamic problems the solar wind decelerates across the bow shock where it also becomes compressed and thermalized reaching values for these properties that are in agreement with those expected from the Rankine-Hugoniot equations of fluid dynamics. As the solar wind streams behind the bow shock and moves around the magnetosphere it gradually recovers its freestream conditions by expanding, cooling and increasing its flow speed as it has been observed with various spacecraft that have probed those regions. As a whole the changes that the solar wind experiences as it interacts with the earth's magnetic field fit adequately with those expected in fluid dynamics.

An evolution of the solar wind similar to that observed around the earth's magnetosphere also occurs when it encounters the magnetic field of other planets. Such is the case for Jupiter, Saturn Uranus and Neptune where their strong intrinsic magnetization leads to large scale magnetospheric cavities and that may reach up to 100 Jovian radii along the sun direction upfront from Jupiter. While comparable large values are also measured for the other planets conditions at Mercury are different where its magnetospheric cavity is very small and only reaches a few thousand kilometers above the surface in the direction of the sun. The solar wind that streams around the magnetic cavity of all these planets bears a similar response in the interaction process.

4. The fluid dynamic interaction of the solar wind with the upper ionosphere of Venus and Mars

4.1 Transport of solar wind momentum to the Venus ionosphere

Unlike magnetized planets Venus and Mars present obstacles to the solar wind that lead to phenomena that are not related to its interaction with a planetary magnetic field. In particular, in the absence of an appreciable intrinsic planetary magnetization the solar wind

does not stream around a magnetic cavity but reaches the upper atmosphere of the planets. The flow conditions are more adequately described as an interaction between two separate bodies of plasma namely, the solar wind and the upper ionosphere. Consequently, the dynamics of the particles that are involved in their interaction is entirely different from that occurring at the boundary of a planetary magnetosphere, and to a large extent it is far more complicated. For example, the access that the solar wind acquires when interacting with particles in a planetary upper ionosphere produces phenomena that do not occur when the solar wind interacts with a planetary magnetic field and suggest the existence of turbulent flow conditions. Information obtained from measurements conducted by spacecraft that have probed Venus and Mars indicate that the interaction between both ion populations seems to be more strongly influenced by turbulent fluid processes rather than through a plasma dynamics similar to that occurring at the boundary of the earth's magnetosphere. There is, in fact, evidence obtained from different sources that suggest the existence of phenomena that strongly resemble those occurring in fluid dynamic problems. A collected view of the various features that have been detected in the Venus plasma environment is presented in Figure 3 from observations conducted with the Pioneer Venus spacecraft (PVO) ⁶. They include, in

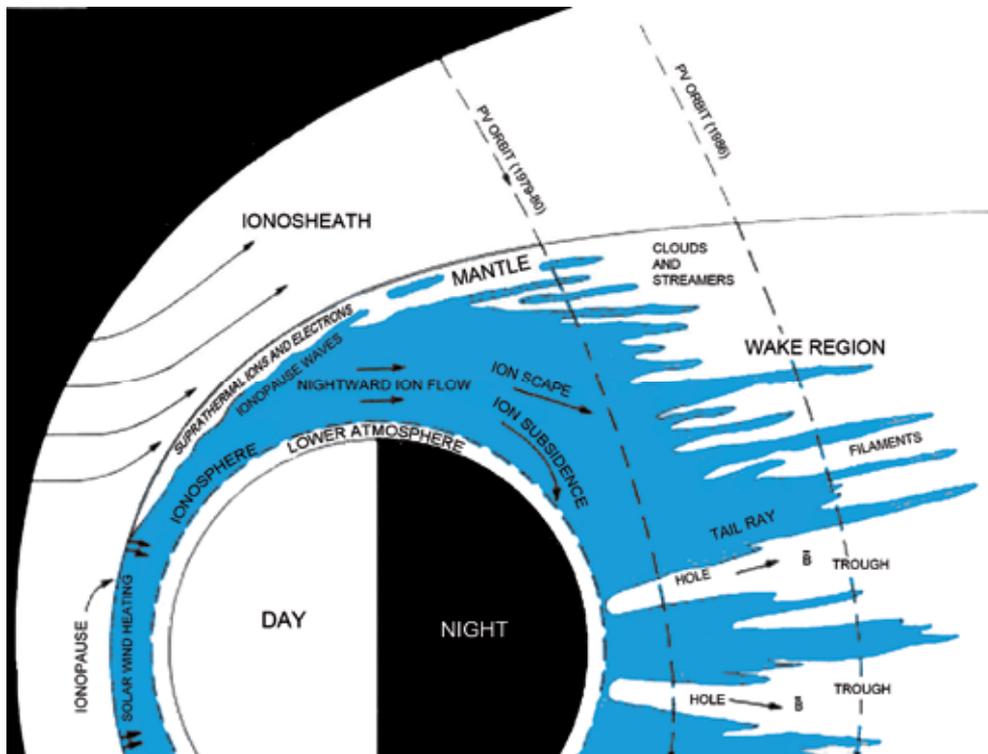


Fig. 3. Schematic diagram of plasma features (bow shock, ionopause, nightward ion flow, tail rays, ionospheric holes) that were identified near Venus with instruments in the Pioneer Venus Orbiter spacecraft (the altitude scale has been enlarged to better describe the geometry of the plasma features)⁶.

addition to a bow shock, plasma clouds or regions of separated ionospheric plasma that imply, in fact, crossings through structures or filaments that extend downstream from Venus. At the

same time there are regions within the nightside ionosphere where there is a large deficiency in the plasma density (ionospheric holes), and equally important is a generalized nightward directed flow of the dayside ionospheric plasma across the terminator.

Studies of the origin of the latter phenomenon (trans-terminator ionospheric flow) have been directed to examine measurements conducted in the region where it has been observed. From the early Mariner 5 spacecraft mission that probed the Venus environment it was learned that the solar wind that flows around the ionosphere exhibits a significant loss of its momentum; that is, its velocity and density show a strong deficit with respect to the freestream values ⁷. That variation is described in the velocity and in the density profiles of the Mariner 5 spacecraft that are shown in the upper panel of Figure 4, and imply that a large fraction of the oncoming energy (shaded regions along the wake) has been removed

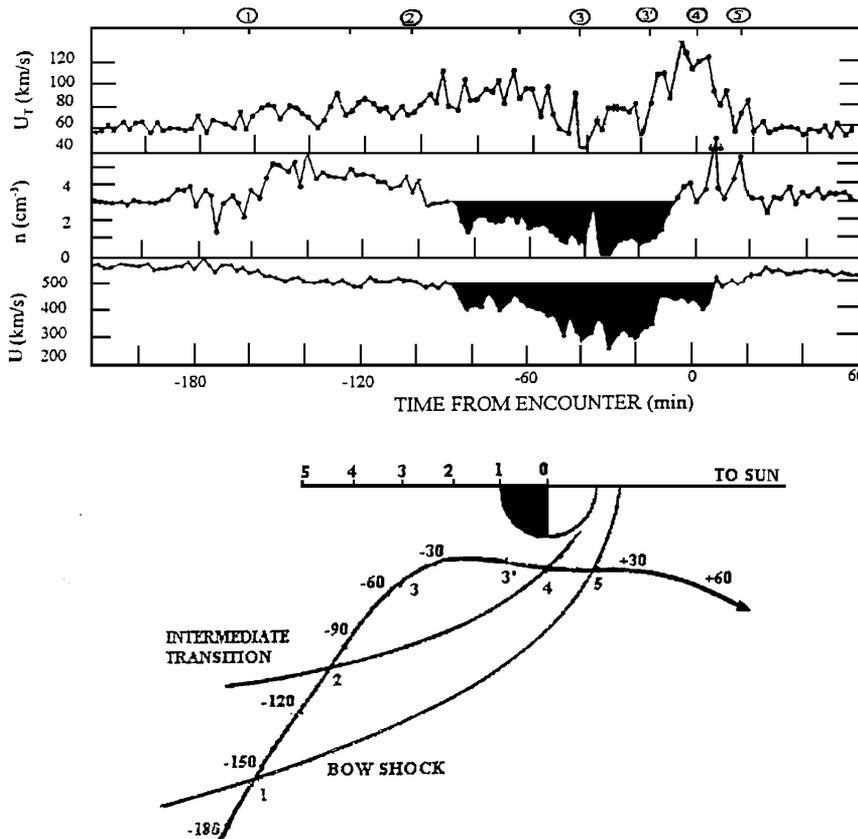


Fig. 4. Thermal speed, density, and bulk speed of the solar wind measured with the Mariner 5 spacecraft (its trajectory projected in cylindrical coordinates is shown in the lower panel). The labels 1 through 5 along the trajectory and at the top of the upper panel mark important events in the plasma properties (bow shock, intermediate plasma transition) ⁷.

from the solar wind. Related information was obtained later from measurements conducted with the Pioneer Venus Orbiter which revealed the ionospheric flow across the terminator that is indicated in the upper panel of Figure 5 and that reaches 3-4 km/s speeds as it is shown in the speed profile in the lower panel of that figure ^{8,9}. The momentum implied by

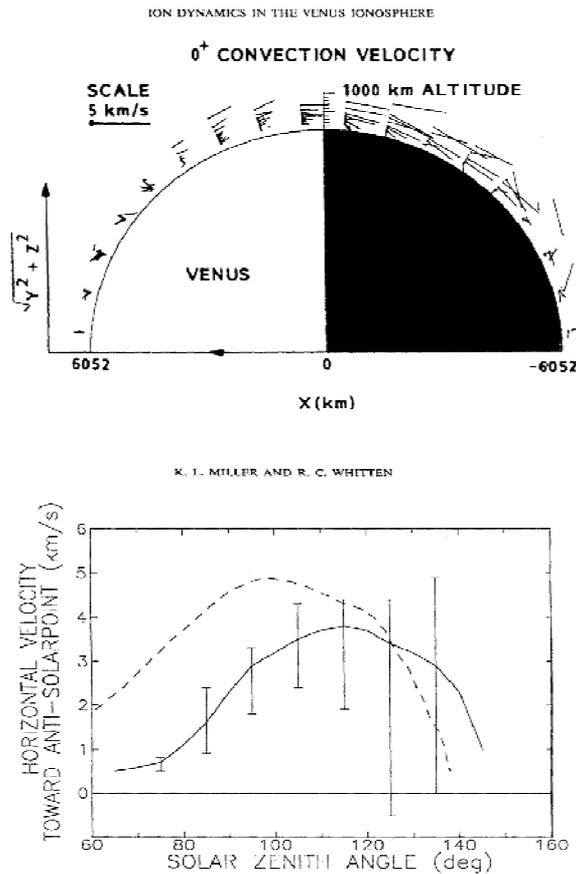


Fig. 5. (upper panel) Average velocity vectors measured in the Venus upper ionosphere with the PVO spacecraft and projected in cylindrical coordinates (the solar wind arrives from the left)⁸. (lower panel). Speed values of the ionospheric particles that move across the terminator at 400 km altitude⁹.

that ionospheric flow is related to the deficit of momentum of the solar wind measured outside the flanks of the ionosphere that is shown in Figure 4. In fact, its kinetic energy density (momentum flux) is comparable to the quantity that was removed from the solar wind; that is, it implies an efficient transport of solar wind momentum to the Venus upper ionosphere¹⁰. Calculations of a momentum conservation equation applied to this problem led to such result which is unrelated to pressure gradient forces across the terminator that have also been suggested to account for the Venus ionospheric flow⁹. The (supersonic) speed of the ionospheric flow and its asymmetry in latitude which is indicated by differences in the extent of the region where the ionospheric plasma is displaced downstream from Venus argues more favorably in terms of momentum transport which mostly occurs by the magnetic polar regions of the Venus ionosphere¹¹. It should be noted in this regard that the "magnetic polar regions" are related to the position where the solar magnetic field fluxes that are convected by the solar wind, and that first pile up around the dayside ionosphere to form a magnetic barrier¹², slip over the planet to take the shape of a hairpin at those regions and then enter the wake¹³. Thus, the outcome of the transport of solar

wind momentum to the Venus upper ionosphere stresses the value of a fluid dynamic view to interpret the interaction between both bodies of plasma despite the fact that it is still necessary to define the mechanisms that through turbulent flow conditions support that behavior.

4.2 Plasma transitions outside the Venus ionosphere

Together with the transport of momentum into the Venus ionosphere the presence of boundaries in the plasma data shown in Figure 4 is indicative of fluid dynamic processes that are operative throughout the region of interaction with the solar wind. As the Mariner 5 spacecraft moved near Venus in its flyby transit from the wake to the dayside (its trajectory projected in cylindrical coordinates is shown in the lower panel of Figure 4) there is clear evidence of two crossings of a bow shock at events labeled 1 and 5. Across the bow shock the speed of the solar wind is smaller in the downstream side (bottom profile in the upper panel), and there is also evidence of heating (larger values in the thermal speed profile labeled U_T) and higher density values (all these variations are similar to those encountered across the bow shock present upstream from the earth's magnetosphere in Figure 2). Equally informative is the observation of a different plasma transition at events labeled 2 and 4 which indicate changes that are not related to a bow shock crossing. In particular, while the solar wind speed further decreases in the downstream side of that transition the density does not increase but exhibits lower values (after event labeled 2 and before event labeled 4). Thus, rather than a bow shock crossing where the solar wind becomes compressed the changes at that plasma transition indicate the presence of a region where the plasma becomes expanded. That transition is not the upper boundary of the ionosphere (the ionopause) which was not crossed by the Mariner 5 spacecraft but a boundary that is innovative since it has not been detected downstream from the bow shock of planetary magnetospheres.

Arguments leading to the presence of the latter transition do not follow from standard views of the acceleration of planetary ions that are generated through the ionization of neutral particles of the Venus atmosphere and that are immersed in the solar wind. In those views the neutral particles that form the planet's exosphere become ionized by the ultraviolet and the x-ray solar radiation and as a result should be rapidly accelerated by an electric field that arises from the relative speed between them and the solar wind¹⁴. The process, named mass loading, is expected to account for the population of planetary ions that is carried by the solar wind and whose contribution should gradually increase with the larger density of exospheric neutral particles that the solar wind encounters when it approaches the Venus atmosphere. While there is clear evidence for the existence of such process there is no indication that it should produce a sharp plasma transition since the accumulation of incorporated planetary ions to the solar wind through mass loading should proceed in a gradual manner¹⁵. Instead, it is necessary to consider that the transition detected downstream from the bow shock in Figure 4 derives from other phenomena, and since it defines the outer boundary of a region adjacent to the ionopause where lower values of the speed and density of the solar wind are measured, it should be related to transport of momentum to the Venus upper ionosphere. In particular, it has been suggested that such plasma transition represents the upper boundary of a region where solar wind momentum is transferred to the Venus ionosphere through viscous forces as it would occur when a streaming flow moves over an obstacle¹⁰. Its position should, therefore, indicate the width of a viscous boundary layer which increases with the downstream distances from Venus as it occurs in comparable fluid dynamic problems.

Further information on the presence of that plasma transitions in the Venus plasma environment has been obtained from spacecraft that have orbited the planet. This is the

case for the Pioneer Venus and more recently for the Venus Express spacecraft whose data have strongly supported the presence of both the bow shock and the plasma transition located downstream from it¹⁶⁻¹⁸. A useful description of recent data is reproduced in Figure 6 from the Venus Express measurements whose trajectory, projected in cylindrical coordinates, is presented in Figure 7. The top and the second panels in Figure 6 show the

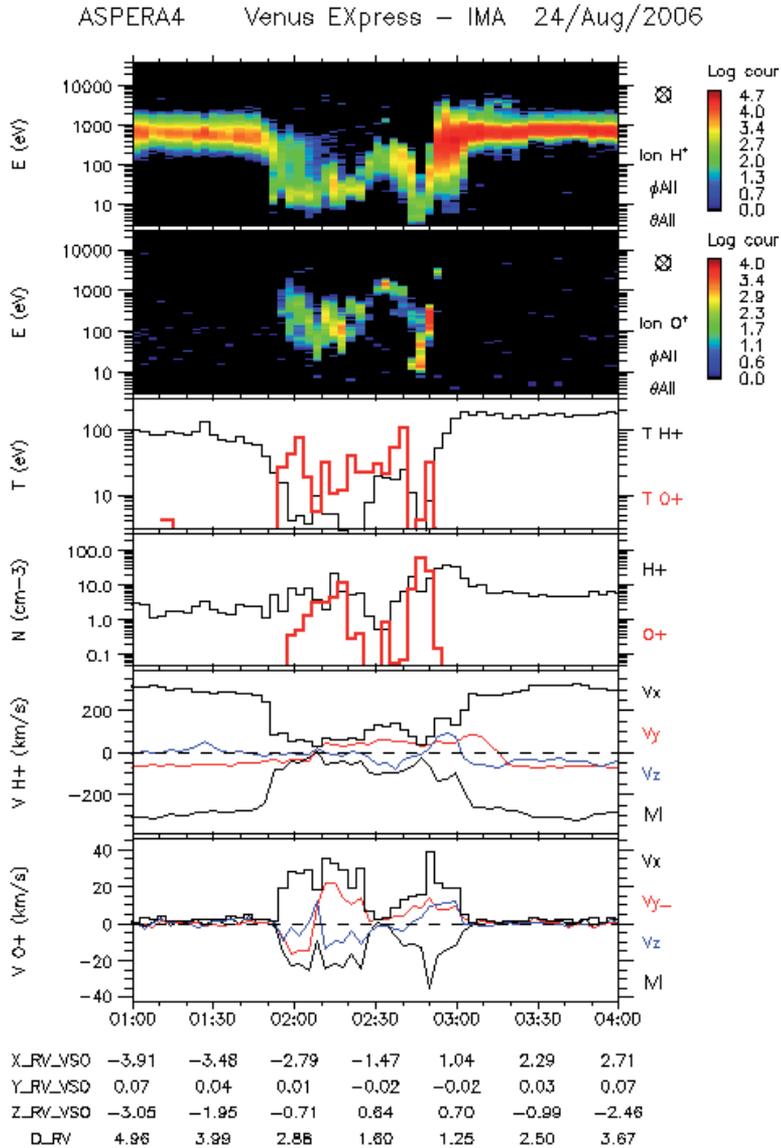


Fig. 6. Energy spectra of H⁺ and O⁺ plasma fluxes (first and second panels) measured in the Venus wake and over the north polar region during orbit 125 of the Venus Express (VEX) spacecraft in August 24, 2006. Temperature and density profiles of both plasma components (third and fourth panels) and their speed V, together with their V_x, V_y, and V_z, velocity components (fifth and sixth panels).

not been expected if the planetary O⁺ ions were accelerated by the electric field that, as noted above, arises in the mass loading process (when exospheric neutral particles are ionized their speed is different from that of the solar wind). Rather than reaching solar wind speeds the planetary O⁺ ions maintain small speed values thus suggesting that their motion does not follow gyration around the magnetic field lines as it occurs when they enter a planetary magnetosphere. Instead, their response to local conditions is that they are slowly accelerated through processes that resemble those occurring in turbulent fluid dynamics. The data presented in Figure 6 thus indicates that the interaction of the solar wind with planetary ions in the Venus plasma environment can be viewed in the context of fluid dynamics with processes related to local turbulence and that provide the basis of that approach.

4.3 Plasma channels over the magnetic polar regions of the Venus ionosphere

Observations related to the geometry of the Venus nightside ionosphere and the plasma distribution in the wake are also supportive of a fluid dynamic view since they refer to well defined features that in many instances are present downstream from the planet. Such observations describe plasma clouds and ionospheric holes that were detected as the Pioneer Venus spacecraft moved through the nightside hemisphere within and in the vicinity of the ionosphere. As depicted in Figure 3 plasma clouds are regions that are separated from the main ionosphere and that in fact should be part of ionospheric streamers or filaments that extend downstream from Venus¹⁹. Ionospheric holes are, on the other hand, regions of very low plasma density that in some orbits of the PVO are detected as the spacecraft moves within the nightside ionosphere²⁰. A large fraction of them are measured in the inner wake near the midnight plane but their location is displaced towards the dawn side. A profile of the electron density measured as the PVO moved through the nightside ionosphere in the PVO orbit 530 is shown in the upper panel of Figure 8. That profile describes the ionospheric holes that were detected at 09:30 UT and at 09:40 UT in the northern and in the southern hemisphere along the near polar oriented trajectory of the spacecraft (As it is shown in the lower panel of Figure 8 the lower values of the electron density within the holes are accompanied by enhanced magnetic field intensities)²¹. In some orbits the PVO only detected one ionospheric hole and in others none at all. Their origin and position have been the source of much research and, in particular, they have been interpreted as resulting from plasma channels (or ducts) that the solar wind produces as it moves over the magnetic polar regions of the Venus ionosphere²¹. A schematic view of such plasma channels is illustrated in Figure 9 to describe how they could be sometimes encountered along the PVO trajectory and thus be detected as ionospheric holes. The low density values of the ionospheric plasma within the plasma channels, and hence observed as ionospheric holes, could be due to the erosion that the solar wind produces upon reaching the (magnetic) polar regions. Favorable conditions for a turbulent interaction between the solar wind and the ionospheric plasma in those regions will arise since the local magnetic field intensity is not significantly enhanced as it occurs in the magnetic barrier that builds up around most of the dayside ionosphere¹².

The results of recent analysis of the Venus Express plasma data are also compatible with the presence of plasma channels that extend downstream from the polar regions and thus stress that further research is required to examine its geometric properties such as width, depth, and their evolution along the downstream distance from Venus. Independent of these characteristics it is to be noted that plasma channels in the polar ionosphere provide an important lead on the fluid dynamic interpretation of the data (plasma clouds,

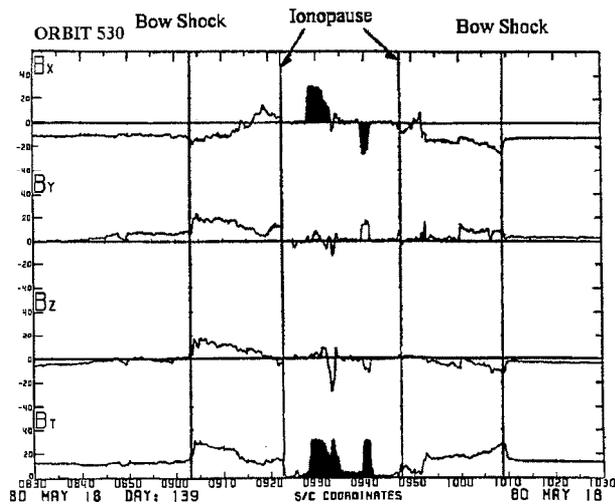
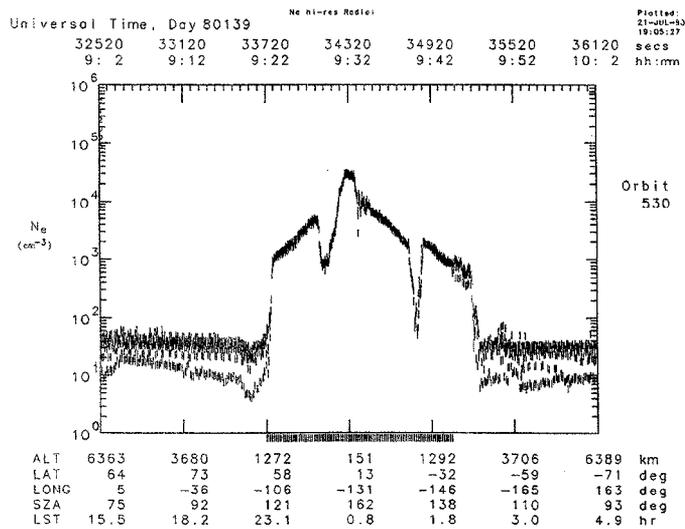


Fig. 8. (upper panel) Electron density profile measured across the nightside ionosphere in orbit 530 of the PVO. The ionospheric holes at 09:30 UT and at 09:40 UT were detected as the spacecraft moved through the northern (inbound pass) and southern (outbound pass) hemispheres²⁰. (lower panel) Magnetic field components and magnetic field intensity B_T measured across the nightside ionosphere in orbit 530 (the shaded areas represent the position of the ionospheric holes)²¹.

ionospheric holes). In particular, while much of the discussion presented here has been directed to describe those features the physical processes that produce them should still be addressed, namely, the manner in which turbulent flow conditions in a collisionless plasma lead to the fluid dynamic behavior.

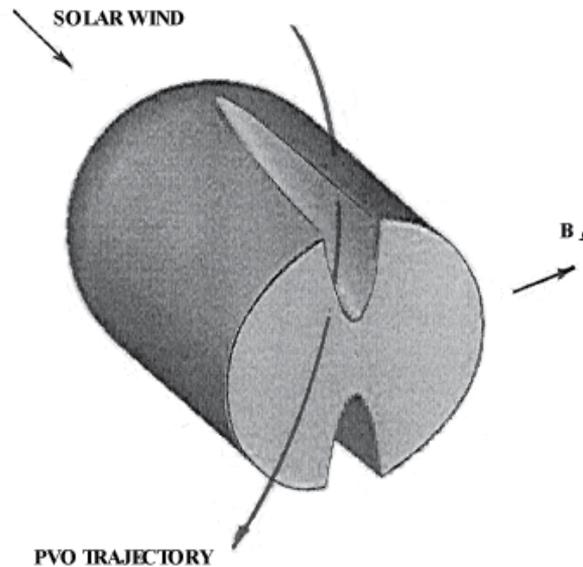


Fig. 9. Schematic view of plasma channels that form by and behind the magnetic polar regions of the Venus Ionosphere ²¹.

4.4 Fluid dynamic Magnus force derived from the rotation motion of the Venus ionosphere

Unrelated to the erosion that the solar wind produces in the upper polar ionosphere there are other observations that also fit within the context of a fluid dynamic interpretation. These refer to a dawn-dusk asymmetry in the distribution of the ionospheric trans-terminator flow as it streams into the nightside hemisphere. Measurements conducted with the PVO show that the flow of ionospheric plasma behind the planet is not seen to converge towards the midnight plane but is oriented to the dawnside of the wake ²². A general description of the manner in which the velocity vectors of the ionospheric flow are traced around and behind the planet (projected on the equatorial plane) is reproduced in the upper panel of Figure 10. As it was noted in Figure 5 the ionospheric flow is mostly measured by and downstream from the terminator but the distribution of the velocity vectors in the upper panel of Figure 10 indicate that the plasma experiences a dawn directed deviation at a $\sim 15^\circ$ angle away from the midnight plane. It is to be noted that the downstream displacement of the trans-terminator flow occurs mostly in the upper ionosphere (between ~ 400 km up to the ~ 1000 km altitude boundary of the ionosphere) and that it is superimposed on a different flow motion that carries the lower ionospheric plasma in the same retrograde rotation of the Venus atmosphere (measurements conducted with the PVO and the Venus Express spacecraft have shown that the neutral Venus atmosphere completes a steady rotation in a 4-5 earth day time span and that this motion is directed in the sense contrary to that of the earth's rotation ²³). Figure 10 shows that the retrograde rotation of the Venus atmosphere is also applicable to the lower ionosphere and hence produces a velocity field that, as it is indicated in the dawn side, is superimposed on the trans-terminator flow. The kinetic energy of the latter flow provides an important contribution to drive the retrograde rotation of the atmosphere ²⁴.

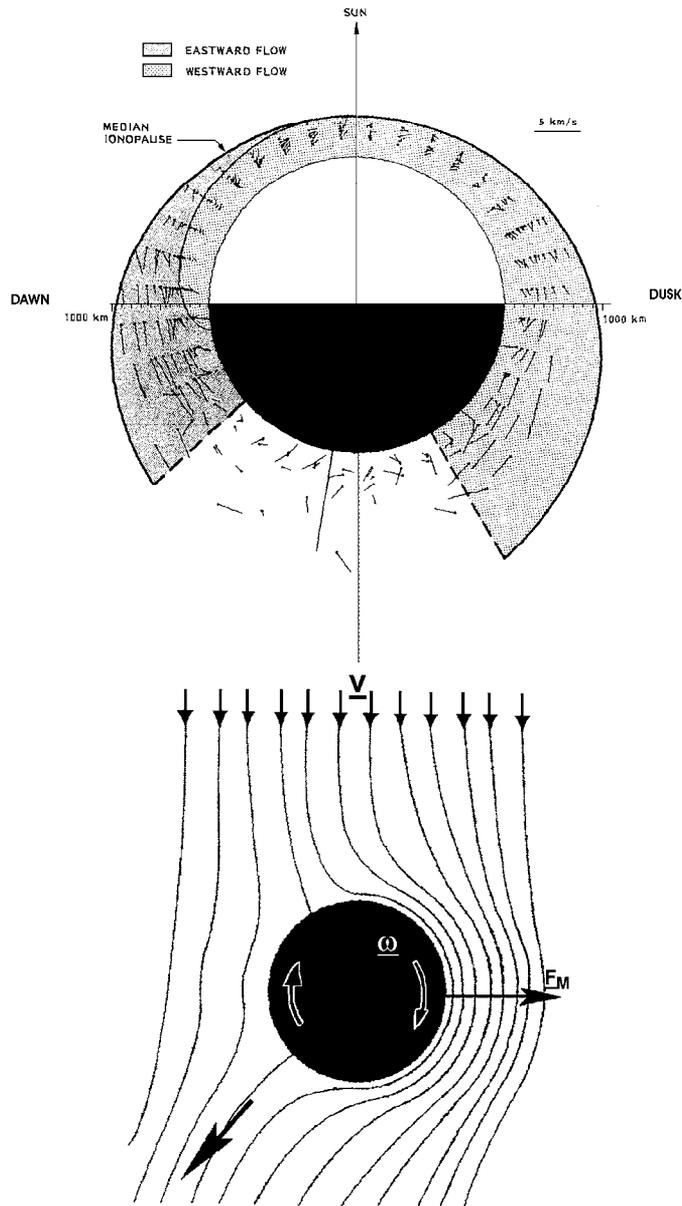


Fig. 10. (upper panel) Velocity averages of the trans-terminator flow and of the rotation of the low altitude ionosphere (which is better seen in the dawn side) that were measured in the Venus ionosphere with the PVO spacecraft and that are projected on the equatorial plane. The velocity vectors of the trans-terminator flow are directed away from the sun (the altitude scale has been increased four times in the figure) ²². (lower panel) Streamline distribution around an obstacle rotating with a frequency ω immersed in a streaming fluid of velocity V . The Magnus Force F_M is directed transverse to the oncoming velocity direction and to the rotation vector ω that points out of the page. The arrow denotes the direction of the deviated wake ²⁶.

The concurrent presence of the rotation motion of an object immersed in a directional flow as that inferred for Venus; namely, the nightward flow of the upper ionosphere and the rotational motion of the lower ionosphere is common in fluid dynamic problems and yields a deflection of the wake that can be used to account for the observations at Venus. A schematic description of the problem is illustrated in the lower panel of Figure 10 which reproduces conditions similar to those presented in the upper panel. The rotation motion of the obstacle produces a velocity field that yields velocity vectors parallel and anti-parallel to those of the directional motion. When they are parallel (right side in the lower panel of Figure 10) the added velocity is large, but becomes small when they are anti-parallel (left side). An important implication derived from the different speed values across the obstacle can be inferred from the Bernoulli's equation:

$$P + \rho v^2 = \text{cst} \quad (3)$$

which implies an inverse relation in the value of the pressure P across the obstacle (ρ is the mass density of the flow) and the kinetic energy of the flow ρv^2 ; that is, low pressure values should be expected in the region where the added flow speed is large (right side in the lower panel of Figure 10) and large pressure values will occur where the added speed is small (left side). In turn, the pressure difference across the obstacle derived from Bernoulli's equation leads to a force (Magnus force) that is responsible for the lateral deviation of rotating baseballs and golf balls in sport activities (curve force)²⁵. The dawnward directed deflection of the Venus trans-terminator flow indicated in the upper panel of Figure 10 results from the response force that the rotating Venus ionosphere applies on the trans-terminator flow (likewise the deflection of the wake in the lower panel of Figure 10 is produced in response to the Magnus force that the streaming flow applies on the rotating obstacle). Calculations were made to estimate the dawnward directed displacement of the trans-terminator flow produced by the Magnus force when using its 3-4 km/s flow speed and the rotation frequency of the lower ionosphere. The predicted displacement value is of the order of ~ 1000 km for the plasma located in the midnight ionosphere, which compares well with the measurements presented in the upper panel of Figure 10 (the dawnward deflection of the symmetry axis of the trans-terminator flow leads to that value at nightside ionospheric altitudes). In summary, the dawn-dusk asymmetry in the distribution of the trans-terminator flow around Venus is compatible with effects produced by the fluid dynamic Magnus force that applies to a rotating obstacle immersed in a directional flow²⁶.

4.5 Solar wind erosion of the Mars ionosphere

Much of the interaction of the solar wind with Mars leads to conditions that are similar to those observed at Venus. In the absence of a global intrinsic magnetization in both planets the solar wind reaches their atmosphere and produces phenomena that result in the erosion of their upper ionosphere. This view is applicable around most of Mars but it is modified in the vicinity of regions where there are fossil magnetic fields possibly remnants of an early global field²⁷ (in those regions the solar wind interacts with small scale magnetospheres of the type generated by the earth's magnetic field). Conditions over the polar regions of the Mars ionosphere resemble those at Venus with ionospheric plasma being removed by the direct access of the solar wind. Information that supports the erosion of the Mars polar ionosphere was obtained from remote observations conducted with the XMM Newton

satellite in orbit around the earth. A long time span image of Mars observed in the x-ray spectrum is presented in Figure 11 to describe the configuration of the region that has been eroded by the solar wind ²⁸. Most notable is the prominent plasma bulge that extends several thousand kilometers downstream from the Mars over the polar regions, and that is

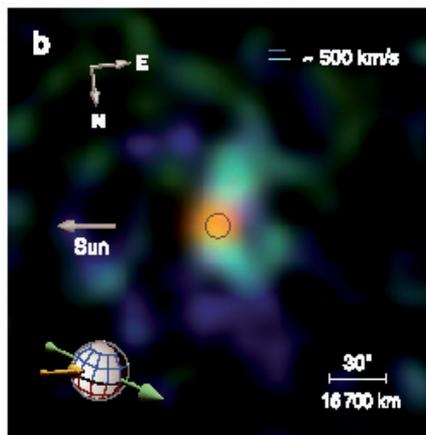


Fig. 11. View of the x-ray emission line halo around Mars measured with the reflecting grating spectrometer (RGS) of the XMM-Newton satellite. The halo is most prominent above the poles and is tilted away from the Sun. The sphere at the lower left indicates the direction of motion of Mars (green arrow) and the velocity of the solar wind particles (yellow arrow) ²⁸.

similar to those occurring at the Venus polar regions. At Mars the solar wind carves out plasma channels with ionospheric plasma being continuously removed within a funnel cross-section that expands away from the wake. A comparable erosion of plasma also occurs by the polar regions of the Venus ionosphere as the channels are maintained open. However, because of the smaller gravitational force of Mars the density of the eroded plasma in that planet should be larger than at Venus, thus yielding evidence of the notable asymmetric plasma halo that is seen in its x-ray spectrum. The x-ray emission in the Mars spectrum corresponds to high energy electron transitions that imply very high temperatures. Calculations made from the emission line distribution in the x-ray Mars spectrum lead to temperature values of the order of one million degrees which may be produced through dissipation processes associated with the transport of solar wind momentum to the polar ionospheric plasma, as it was also suggested for Venus ²⁹. The remarkable shape of the Mars plasma halo shown in Figure 11 represents visible evidence of phenomena that substantiate the value of fluid dynamics in the interpretation of the solar wind interaction with planetary ionospheres. The collective removal of ionospheric plasma from the planets Venus and Mars through erosion processes is a useful byproduct of turbulent flow conditions that allow fluid dynamics to become applicable.

5. Concluding remarks

Much of what has been discussed in regard to the fluid dynamic view of the origin of the solar wind, its expansion through interplanetary space, and its interaction with the planets of the solar system leads to processes that are related to the motion of charged

particles through magnetic fields and/or are subject to turbulent flow conditions. In the first case the solar wind particles enter the earth's magnetic field and produce electric currents that bound a cavity in which the magnetic field is compressed and is deformed to produce a magnetic tail. In the latter case the solar wind reaches planetary ionospheres where instead of gyration motion around the magnetic field lines the plasma particles describe stochastic trajectories produced by local turbulence and that suggest the existence of wave-particle interactions. The outcome of this latter activity is the onset of phenomena that can be interpreted in terms of those occurring in fluid dynamic problems. A process that has been applied in this context is the transport of solar wind momentum that occurs along the flanks of planetary ionospheres as it is shown in Figure 4 (the boundary traced between events 2 and 4 represents the outer extent of a viscous layer). For those conditions it has been possible to estimate the viscosity coefficient of the solar wind that is suitable to account for the observations. The calculated values of the kinematic viscosity of the solar wind suggest that they are compatible with the plasma properties of the solar wind; namely, an efficient transport of momentum across a thick ($\sim 10^3$ km) viscous boundary layer measured by the terminator and that is produced under its very low particle densities. In fact, since the value of the kinematic viscosity coefficient ν describes the ability of a flow to modify a velocity profile³⁰ the small density of the solar wind leads to large $\nu \sim 10^4$ km²/s values when compared with those of comparable fluid dynamic problems³¹. Empirical values of the kinematic viscosity of the solar wind are derived on the basis that local turbulence produced by wave-particle interactions provide the origin of the fluid dynamic interpretation^{7, 32}. However, while there are various phenomena that substantiate this view it is still necessary to determine the mechanisms that produce such interactions which lead to the collective response of the solar wind. A required source of information would be obtained from a statistical mechanics applicable to wave-particle interactions and that could provide in mathematical form expressions for the transport coefficients. Much theoretical research should be conducted to improve the information that is now available.

6. References

- [1] Parker, E. N., *Interplanetary Dynamic Processes*, John Wiley, 1963.
- [2] Liepmann, H. W. and A. Roshko, *Elements of Gas dynamics*, John Wiley, 1967 (Chapter 2, p. 52).
- [3] Dessler, A. J., *Solar Wind and Interplanetary Magnetic Field*, *Reviews of Geophysics*. 5,1, 1967.
- [4] Spreiter, J., et al., *Hydromagnetic flow around the magnetopause*, *Planetary Space Sciences*, 14, 223, 1966.
- [5] Dryer, M., et al., *Magnetogasdynamic conditions for a closed magnetopause*, *American Institute Aeronautics Astronautics Journal (AIAA)*., 4, 246, 1966.
- [6] Brace, L., et al., *The Ionotail of Venus: Its configuration and evidence for ion escape*, *Journal of Geophysical Research*, 92, 15, 1987.
- [7] Bridge H., et al., *Plasma and magnetic fields observed near Venus*, *Science*, 158, 1669, 1967.
- [8] Knudsen, W., et al., *Improved Venus ionopause: Comparison with measurements*, *Journal of Geophysical Research*, 87, 2246, 1982.

- [9] Whitten, R. et al., Dynamics of the Venus ionosphere: A 2D model study, *ICARUS*, 60, 317, 1984.
- [10] Pérez-de-Tejada, H., Fluid dynamic constraints of the Venus ionospheric flow, *Journal of Geophysical Research*, 91, 6765, 1986.
- [11] Pérez-de-Tejada, H. Friction layer in the plasma channels of the Venus ionosphere, *Advances of Space Research*, 36, 2030, 2005.
- [12] Zhang, T., J. Luhmann, and C. Russell, The magnetic barrier at Venus, *Journal of Geophysical Research*, 96, 11145, 1991.
- [13] Pérez-de-Tejada, H., et al., Plasma distribution in the Venus near wake, *Journal of Geophysical Research*, 88, 9109, 1983.
- [14] Breus, T. et al., Solar wind mass-loading at comet Halley, *Geophysical Research Letters*, 14, 499, 1987.
- [15] Intriligator, D., Observation of mass addition in the Venusian ionosheath, *Geophysical Research Letters*, 9, 727, 1982.
- [16] Slavin, J., et al., The solar wind interaction with Venus: PVO-Venus bow shock, *Journal of Geophysical Research*, 85, 7625, 1980.
- [17] Pérez-de-Tejada, H., et al., Intermediate transition of the Venus ionosheath, *Journal of Geophysical Research*, 100, 14523, 1995.
- [18] Pérez-de-Tejada, H. et al, Plasma transition at the flanks of the Venus ionosheat: Evidence from the Venus Express, *Journal of Geophysical Research*, 116, A01103, 2011.
- [19] Brace, L., et al., Plasma clouds above the ionosphere of Venus and their implications *Planetary Space Sciences*, 30, 29, 1982.
- [20] Brace, L., et al., Holes in the nightside ionosphere of Venus, *Journal of Geophysical Research*, 87, 199, 1982.
- [21] Pérez-de-Tejada, H., Plasma channels in the Venus nightside ionosphere, *Journal of Geophysical Research*, 109, A04106, 2004.
- [22] Miller, K., and R. Whitten, Ion dynamics in the Venus ionosphere, *Space Science Reviews*, 55, 165, 1991.
- [23] Schubert, G., et al., Structure and circulation of the Venus atmosphere, *Journal of Geophysical Research*, 85, 8007, 1980.
- [24] Lundin, R. et al., Ion flow and momentum transfer in the Venus environment, *ICARUS* (in press, 2011).
- [25] Rouse, H., *Elementary Mechanics of Fluids*, chap. IX, Dover, Mineola, N.Y. 1978.
- [26] Pérez-de-Tejada, H., The Magnus force in the Venus ionosphere, *Journal of Geophysical Research*, 111, A11105, 2006.
- [27] Acuña, M., et al., Global distribution of crustal magnetization at Mars, *Science*, 184, (5357), 790, 1999.
- [28] Dennerl, K., et al., First observation of Mars with XMM-Newton, *Astronomy and Astrophysics*, 451, 709, 2006.
- [29] Pérez-de-Tejada, H. et al, Solar wind erosion of the Mars polar ionosphere, *Journal of Geophysical Research*, 114, A02106, 2009.
- [30] Batchelor, G. K. *An Introduction to Fluid Dynamics*, Cambridge University Press (p. 36), 1979.

- [31] Pérez-de-Tejada, H., Viscous forces in a boundary layer at the Venus ionosphere, *Astrophysical Journal*, 525, L65, 1999.
- [32] Vörös, Z., et al., Intermittent turbulence, noisy fluctuations, and wavy structures in the Venusian magnetosheath and wake, *J. Geophys. Res.-Planets*, 113, E00B21, doi:10.1029/2008, 2008.

Aero - Optics: Controlling Light with Air

Cosmas Mafusire^{1,2} and Andrew Forbes^{1,2}

¹*Council for Industrial Research National Laser Centre*

²*School of Physics, University of KwaZulu-Natal
South Africa*

1. Introduction

It is possible to control light in a medium comprising nothing more than a fluid, whether in gas or liquid phase. Using air as an example, it is possible to bend light by a process of continuous refraction as opposed to stepwise refraction in the case of solid state optics. This is achieved by effecting a continuous change of the refractive index from a high value for lowest temperature (or high density) to a low value for the highest temperature (or lowest density). Since the density changes gradually from a low to a high value, the refractive index is graded as well. In fact, this effect is well known in nature: the mirage effect where light refracts away from a hot surface, usually a road or desert sand, is precisely due to continuous refraction of light. If the refractive index can be customised and controlled, then one has the possibility to create aero-optics: control of light with air. Such optics are a special subset of GRIN (Graded Refractive INdex) optics, with the most studied example being the simple lens, so-called gas lenses. The efficiency at which these lenses operate is dependent on the effectiveness of the refractive index gradient – the greater the gradient, the stronger the lens. There are advantages to such optics: they are not dispersive, and hence suitable for broad bandwidth laser pulses, and have no practical damage threshold as compared to most solid-state optics (e.g., glass). This latter point remains to be exploited for high power laser beam delivery in such applications as laser fusion and peta-watt high intensity laser beam delivery.

In the remainder of this chapter we will consider such gas lenses in more detail. We outline a simple approach to controlling light through a medium comprising nothing more than air in a spinning pipe that is heated along its boundary. We outline the fluid dynamics of this simple system and through the use of computational fluid dynamics (CFD), explain the mechanisms that give rise to a plethora of interesting properties, from a lensing action resulting in the focussing of light, to a simulator of optical turbulence in the laboratory. The CFD models are supported in most cases by intuitive analytical expressions, teasing out fundamental physics from the fluid dynamics of this system, while the predicted impact of the fluid dynamics on the optical field is verified experimentally. Despite the simplicity of the system – merely a spinning, heated pipe – the fluid dynamics of the system is shown to be rather complex, comprising regions of slow and fast flow, symmetric and asymmetric density gradients, and laminar and turbulence regions. We conclude the study by outlining how such aero-optics, using aerodynamics and fluid flow to control light, may be exploited in applications ranging from long range telescopes to overcoming the damage thresholds of conventional solid state glass optics for delivery of high power lasers.

2. Gas lenses

The principle of operation of the spinning, heated, pipe as a gas lens, an example of which is shown in Fig. 1, is based on the concept that the shear viscosity of air increases when heated, so that when the outside wall of the pipe is heated the layer of air next to the pipe's inner wall is heated as well. During rotation this inner layer is centrifugally expelled and replaced by cold air from the surroundings, inhaled along the axis from both ends, resulting in an axial (air) region that is cooler compared to the air at the wall. This temperature gradient creates a density gradient resulting in a refractive index gradient conducive to focusing, thus the name: gas lens. Rotating the heated pipe converts the density with a vertical gradient due to gravity to one that is rotationally symmetric about the axis. In order to customise the aero-optic for a lensing action, it is necessary to create a parabolic density gradient (and hence temperature gradient) across the pipe, and preferably everywhere along its length. A purely parabolic gradient in the density would be an ideal GRIN lens: the perfect gas lens.

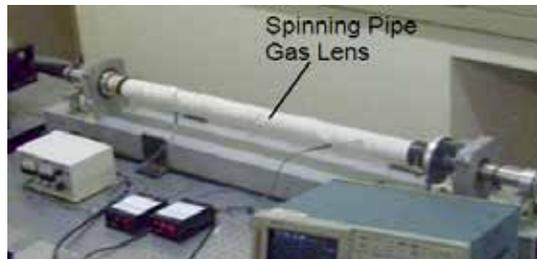


Fig. 1. An image of the spinning pipe gas lens (SPGL), comprising a steel pipe that is heated and spun around its axis. The central heated section is evident by the white heater tape surrounding it.

Previous studies have shown that the SPGL imitates a graded index (GRIN) medium with a refractive index which is a maximum along the axis and decreases parabolically with radial distance r , towards the walls (Michaelis, 1986; Forbes, 1997; Mafusire, 2006)

$$n(r) = n_0 - \frac{1}{2}\gamma^2 r^2 \quad (1)$$

The radial refractive index parameter, γ is a measure of the power of the SPGL, while n_0 is the refractive index along the axis.

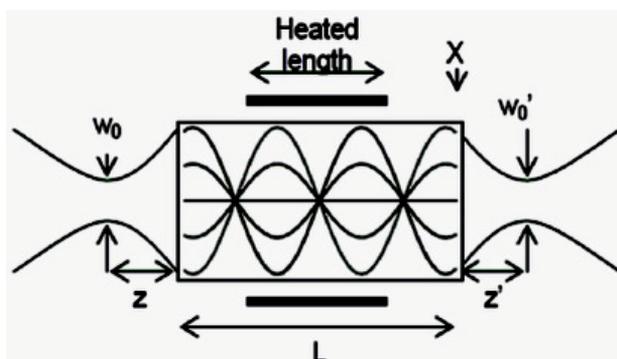


Fig. 2. A laser beam being converged on exit from a spinning pipe gas lens.

The focal length of the lens in terms of the input and output beam waist positions, w_0 and w_0' , located at respective distances, z and z' with respect to the SPGL (Fig. 2) is given by (Mafusire, 2008b)

$$f = \frac{1}{2(z' - z)} \left[z_0^2 + \frac{1}{\gamma^2} + 2zz' - z \pm \sqrt{z_0^4 + 2z_0^2 \left(\frac{1}{\gamma^2} + 2z' - z^2 \right) + \left(\frac{1}{\gamma^2} + z^2 \right)^2} \right] \quad (2)$$

where z_0 is the beam's Rayleigh range in front of the lens.

The earliest gas lenses were made at Bell Laboratories in an early attempt at producing waveguides for long range communication as a forerunner to the optical fibre. They were operated by heating the walls to a suitably high temperature then cold air was injected along the axis, creating a graded index between the lower density of the hot marginal air and increasing radially towards the cooler axis. Any laser beam propagating along the axis is refracted towards the denser axis hence were referred to as the tubular gas lenses. Early work was done by Berreman (Berreman, 1965) and Marcuse (Marcuse, 1965). In fact Marcuse showed that these lenses approximate a thin lens since both their principal planes coincide. Beyond the potential use for waveguiding (Kaiser, 1968), they were also used as telescope objectives (Aoki, 1967).

The main disadvantage with this design was that convection currents were clearly present and so introduced aberrations due to gravity on the laser beam wavefront. Gloge (Gloge, 1967) showed that the optical centre of such lenses is displaced vertically downwards, an effect which increased with tube diameter. Kaiser (Kaiser, 1970) circumvented this problem by designing a gas lens where hot gas could be exhausted radially, with the cold gas injected as before, resulting in improved performance. Meanwhile, in the former Soviet Union, another improvement was made by replacing the gas injection with rotating a heated pipe about its axis (Martylenko, 1975). It was demonstrated, theoretically and experimentally, that the rotation removed distortions due to gravity by creating a rotationally symmetric density distribution which completely overcame the convection currents of the tubular gas lens.

The investigation into this version of this gas lens design, now called the spinning pipe gas lens, or SPGL for short, was taken over by the then University of Natal without the knowledge of the research in the former USSR. They worked with a vertical lens which was used for focusing a high power laser for drilling holes into metal sheets (Michaelis, 1986). Another improvement was to combine both rotation and gas injection to create a steady focus (Notcutt, 1988). An interesting application was when the SPGL was used as a high quality telescope objective to take images of sun spots and moon craters, illustrating the power of such devices (Michaelis, 1991). Further improvements included operating the SPGL at pressures higher and lower than atmospheric pressure (Forbes, 1997) in order to control the focal length, and careful characterisation of the temperature distribution inside the pipe (Lisi, 1994).

Recently a more modern approach to understanding such devices has been completed, considering the interface of fluid dynamics with physical optics propagation and characterisation of laser beams (Mafusire et al, 2007; Mafusire et al, 2008a; Mafusire et al, 2008b; Mafusire et al, 2010a; Mafusire et al, 2010b). Using computational fluid dynamics (CFD) to simulate the density and velocity distributions inside the SPGL, one can deduce the phase change imparted to the light through the Gladstone-Dale law, thus making it possible

to calculate optical aberrations at any plane along the beam's path. To appreciate this, consider the gas medium where the density at each point in the medium fluctuates about a certain mean density, σ . These density fluctuations result in refractive index fluctuations also about a mean refractive index, N . These two parameters are related by the Gladstone-Dale law given by (Zhao et al, 2010)

$$N = G(\lambda)\sigma + 1 \quad (3)$$

where $G(\lambda) = 2.2244 \times 10^{-4} [1 + (6.7132 \times 10^{-8} / \lambda)^2]$, the wavelength dependant Gladstone -Dale constant which has a value of about 2.25×10^{-4} m³/kg in air for HeNe laser radiation ($\lambda = 6.328 \times 10^{-7}$ m). If the geometrical length of the laser beam path in the gas lens is given by, l , then the phase change of the laser beam is given by

$$\phi = k l N \quad (4)$$

where $k = 2\pi / \lambda$ is the laser beam wavenumber for radiation of wavelength, λ . The implication is that the optical distortions in the gas lens, due to imperfect control of the fluid to create the lens, can be measured by observing the phase changes in the laser beam. Distortions to the phase are called optical aberrations. The measurement of the optical aberrations of this particular aero-optic has introduced a very interesting paradox: rotation of the gas lens removes distortions due to gravity, but as we will see later, the other aberrations increase in magnitude as the rotation speed and the wall temperature increase. Therefore such optics are not perfect. To understand where the imperfections originate from, we start with a computational fluid dynamics (CFD) model of the gas lens.

2.1 Computational fluid dynamics model of the gas lens

To confirm the theoretical analysis of the SPGL given above and the optical aberrations measured from its performance, a computational fluid dynamics (CFD) simulation of a simplified test system was executed using the commercial CFD code, STAR-CD® using the $k-\varepsilon$ model which involves the numerical solution of two coupled equations, a turbulent kinetic energy (k) equation and an energy dissipation rate (ε) equation [Blazek, 2001; Davison, 2011]. The purpose of this study was to show the effect of the heat and mass transfer on the velocity distribution and density. Assumptions included the removal of the mounts and other three-dimensional geometry features that would complicate the geometric model. The pipe used in the model was based on the dimensions of the actual SPGL we have in the laboratory: 1.43 m in length and with an internal diameter of 0.0366 m. The heated section of the pipe was approximately 0.91 m long, leaving two unheated end lengths of length 0.25 m each. The pipe was accurately reproduced with the further assumption that the mounts act as a heat sink and thus the pipe ends are unheated. A fully transient solution is presented in which the pipe is spun up from a heated (at a temperature of 100 °C) steady-state buoyancy-driven solution, and held at fixed speed of 20 Hz until a steady state was reached.

The mesh used in the solution consisted of 350 transverse slices regularly arranged along the pipe's length. Each slice consisted of 512 nodes distributed across the section with more nodes concentrated along the boundary (see Fig. 3). The results extracted from the model included density for each data point in the centre of the cell in the mesh and animations

showing the evolution of the velocity and density at selected transverse and longitudinal cross-sections. Temperature distribution was extracted from the density data.

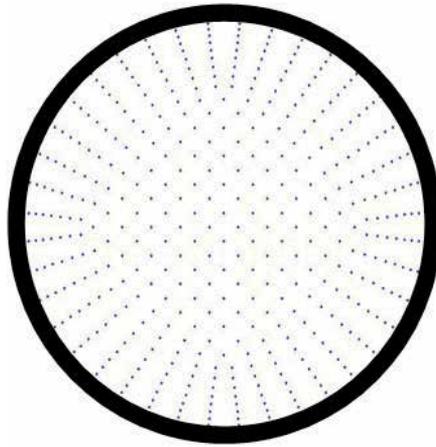


Fig. 3. The SPGL mesh transverse cross-section used in the CFD model.

2.1.1 Velocity distribution

Fig. 4 is a longitudinal cross-section of the gas lens showing the velocity distribution for various cuts of the pipe: (a) left third of the pipe, (b) the mid-section of the pipe, and (c) the right third of the pipe. The velocity distribution in the unheated sections is dominated by transverse movement of air towards the centre from both sides in the inviscid section

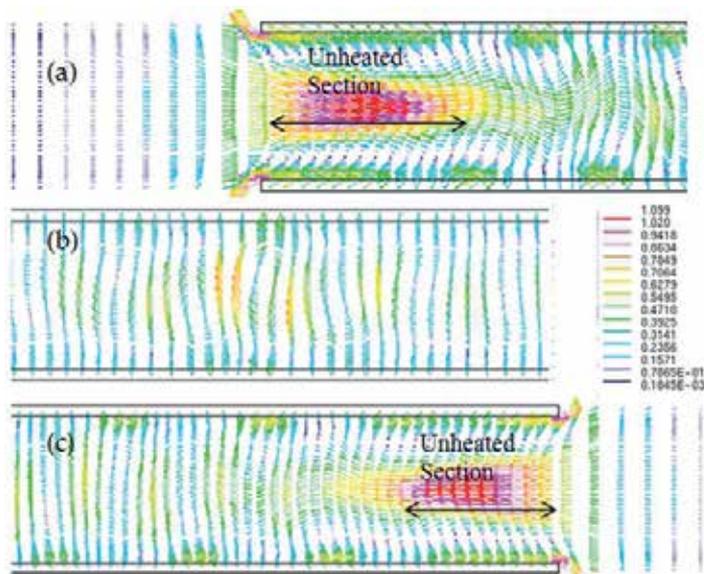


Fig. 4. Velocity profiles in the SPGL for the (a) left end section, (b) the mid-section of the pipe, and (c) the right end section.

outwards along the boundary section, with rotation playing very little or no part. As one moves towards the centre of the pipe length, the transverse speed of incoming air decays (as expected). In the boundary layer, the expelled air is fastest as it exits from the pipe. As the air along the boundary approaches the pipe end, it increases in both translation and rotational velocity, thus the boundary expels hot air in a spiral motion at both ends of the pipe. From this velocity distribution, we can infer the heat distribution in the inviscid region since fast moving air accumulates transversely injected heat much more slowly compared to stationary air. This means the incoming air is cooler but accumulates heat as it slows down and approaches the centre. This might point to a situation where the temperature along the axis at the centre is higher than that of the walls: because of the particles' slow velocity, they accumulate heat over time.

A closer look at the velocity profile of Fig. 4 (b) shows the complicated velocity distribution away from ends of the pipe. This could be evidence of multi-cellular flow, the irregularity of which is most likely responsible for the presence of some of the optical aberrations. This oscillatory activity is expected to increase with increase in wall temperature or rotation speed, which in turn, increases the magnitude of the optical aberrations.

2.1.2 Density distribution

The initial state of the gas in a heated stationary heated SPGL is a result of natural convection. A CFD simulation of this is shown in Fig. 5 (a): it shows a density gradient that decreases vertically so that the gas is layered in horizontal bands within the pipe. When the pipe is spun the image changes radically: the vertical density gradient gives way to a density gradient that has rotational symmetry due to the forced convection of the rotating system, with the highest density along the axis and decreases towards the edges of the pipe (see Fig. 5 (b)). This is an example of a customised aero-optic: such a density distribution is conducive to focusing, since it results in a refractive index profile given by Eq. (1).

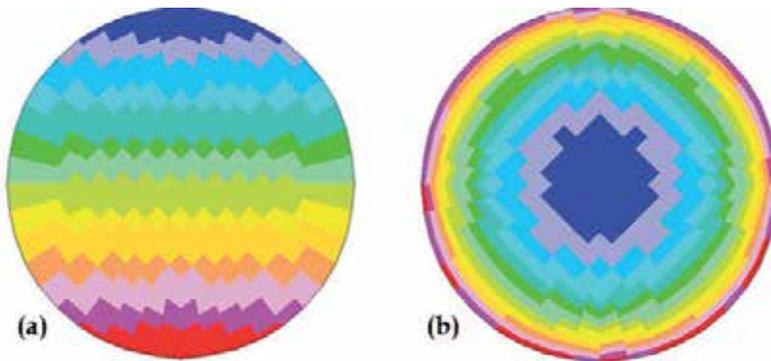


Fig. 5. Cross-sectional density profiles of an SPGL showing: (a) the initial state after heating, and (b) the rotating steady-state near the end face of the pipe, with high density centre (blue) and low density edges (red).

The other results from the density calculations confirm the velocity profile from the previous sub-section. The temperature (T) distribution was extracted from the density (ρ) data using the relation, $\rho = \rho_o(1 - a(T - T_o))$ where $a = 2.263123 \times 10^{-4} \text{ K}^{-1}$ is the coefficient of volume expansion of air and ρ_o is the known density at a known temperature, T_o which we

take to be room temperature. Fig. 6 (a) shows a 3D plot of the density profile of the air everywhere along the length of the pipe for a particular (but arbitrary) cross-section. From this one can visualise the density distributions in transverse sections at critical planes, and a zoomed in view of the centre of the pipe is shown in Fig. 6 (b). The corresponding temperature profile shows that the unheated section is at room temperature (27 °C) and the centre of the pipe is at about 100 °C, the temperature of the wall.

From Figs. 4 and 6 we can start to understand how this particular device works: the intake of cold and the expulsion of hot air, which takes place at the ends of the pipe, results in parabolic-like density profiles near the pipe ends, which results in the lensing effect. As can be observed in the transverse profiles, a significant parabolic density distribution is only evident in a short section of the entire length, possibly less than half. As the pipe is spun faster, so this “lensing length” increases, making the lens stronger. A closer look at the cross-section profiles in Fig 6 (a) shows that some of them have a quartic density distribution, albeit for a very short length compared to the purely parabolic density profile, and this is the source of spherical aberration which has been shown to increase with rotation speed and/or temperature (Mafusire et al, 2008a).

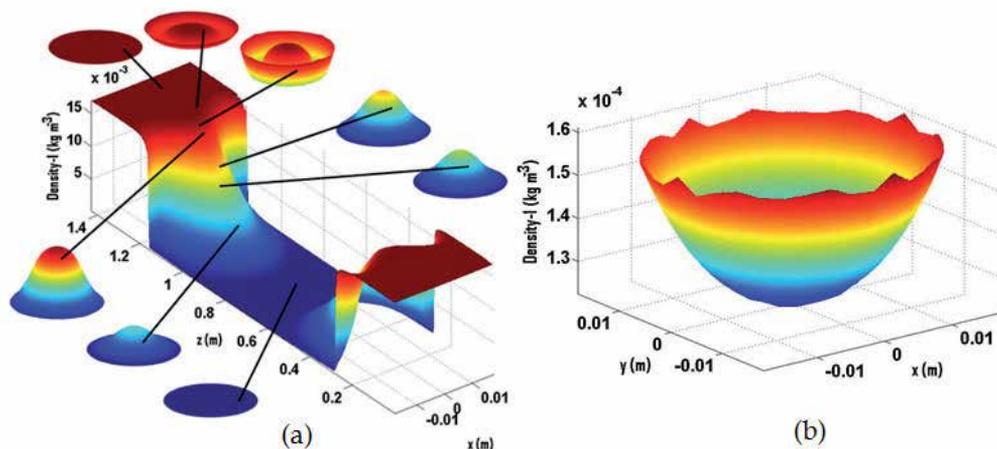


Fig. 6. (a) A 3D longitudinal density profile for the SPGL showing transverse sections at critical planes, and (b) the transverse density profile at the centre of the SPGL.

Another interesting observation from the CFD model of the SPGL is the transverse temperature profile in the centre of the pipe, i.e. away from the face ends, which reveals that the temperature along the axis is just over 0.16 K higher than at the walls, corresponding to a density $4 \times 10^{-5} \text{ kg m}^{-3}$ lower than at the walls (Fig 6 (b)). The explanation for this is that because air particles at this part of the pipe are moving very slowly and thus accumulate heat faster due to both convection and conduction. Thus, somewhat counter intuitively, the central heated part of the gas lens actually results in a decrease in the lensing action.

2.2 Experimental verification of the gas lensing

The experimental setup used is shown in Fig. 7. When the pipe is not rotated and not heated, no lensing takes place, and the phase of the light is taken as a reference. One expects

that if the pipe acts as a lens, the parabolic refractive index will result in curvature on the phase of the light, indicative of a lens. To test this, the pipe was heated to wall temperatures of 348 K, 373 K, 398 K and 423 K, and spun at various rotation speeds (limited to 20 Hz in our experiments). As the pipe is spun, so the focal length of the lens decreased, suggesting a stronger lens. Similarly, when the wall temperature of the pipe was increased, the lensing became stronger. Experimental results of this lensing is shown in Fig. 8.

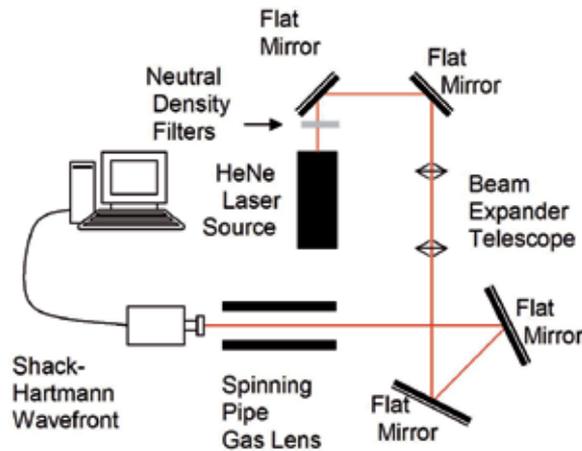


Fig. 7. Experimental setup used to test the lensing properties of the gas lens.

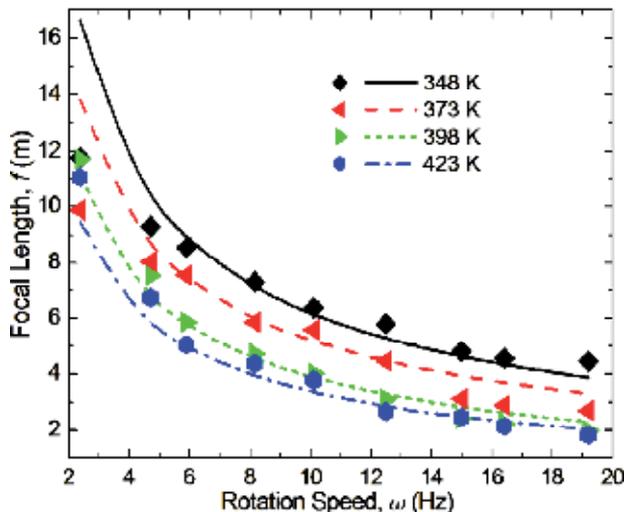


Fig. 8. The strength of the lensing action of the gas lens increases with both pipe rotation speed and pipe wall temperature. Short focal lengths suggest a strong lens, while long focal lengths suggest a weak lens.

3. Optical aberrations

We have seen earlier that the gas lens is aberrated due to the inability to completely control the fluid as desired. In order to describe just how aberrated the lens is, we introduce the concept of optical aberrations, and their description by Zernike polynomials, the coefficients of which are correlated to the magnitude of the optical aberrations. Once these coefficients are known, we can calculate the rms phase error (the deviation from flatness) and the Strehl ratio – how good the focus is for retaining power in a small spot.

3.1 Zernike polynomials

The phase of a laser beam can be written as a linear combination of Zernike polynomials. Zernike polynomials are unique in that they are the only polynomial system in cylindrical co-ordinates which are orthogonal over a unit circular aperture, invariant in form with respect to rotation of the co-ordinate system axis about the origin and include a polynomial for each set of radial and azimuthal orders [Born & Wolf, 1998; Dai, 2008; Mahajan, 1998; Mahajan, 2001]. More importantly, the coefficients of these polynomials can be directly related to the known aberrations of laser beams, making them invaluable in the description of phase errors. The transverse electric field of a laser beam, U , in cylindrical coordinates, ρ and θ can be represented by the product of the amplitude ψ and phase φ as shown by

$$U(\rho, \theta) = \psi(\rho)e^{i\varphi(\rho, \theta)} \quad (5)$$

The expansion of an arbitrary phase function, $\varphi(\rho, \theta)$, where $\rho \in [0, 1]$ and $\theta \in [0, 2\pi]$, in an infinite series of these polynomials will be complete. The circle polynomials of Zernike have the form of an angular function modulated by a real radial polynomial. We can represent each Zernike term by

$$Z_{nm}(\rho, \theta) = C_{nm}R_{nm}(\rho)\Theta_m(\theta) \quad (6)$$

The angular part is defined as

$$\Theta_m(\theta) = \begin{cases} \cos m\theta, & m > 0 \\ \sin m\theta, & m < 0 \\ 1, & m = 0 \end{cases} \quad (7)$$

whereas the radial part is a polynomial given by

$$R_{nm}(\rho) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k (n-k)! \rho^{n-2k}}{k! (\frac{n+m}{2} - k)! (\frac{n-m}{2} - k)!} \quad (8)$$

where n and m are the non-negative order and ordinal numbers respectively which are related such that $m \leq n$ and $n - m$ is even. C_{nm} is the respective coefficient for a particular aberration and can be either A_{nm} and B_{nm} depending whether the aberration is either even or odd, respectively.

This means that a laser beam wavefront described by a phase function φ can be expanded as a linear combination of an infinite number of Zernike polynomials, using generalized coefficients as follows

$$\varphi(\rho, \theta) = 2\pi \sum_{n=0}^{\infty} A_{n0} R_{n0}(\rho) + 2\pi \sum_{n=1}^{\infty} \sum_{m=1}^n R_{nm}(\rho) [A_{nm} \cos m\theta + B_{nm} \sin m\theta] \quad (9)$$

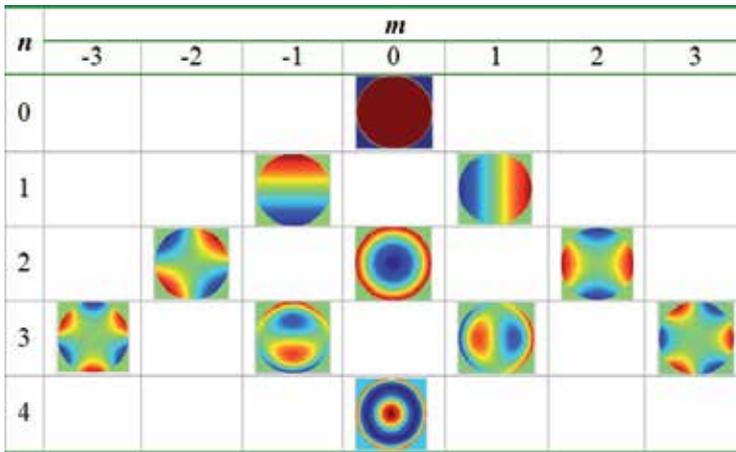


Table 1. Contour plots of the Zernike primary aberration polynomials.

n	m	Description and symbol	Polynomial
0	0	Piston, A_{00}	1
1	-1	y-Tilt, B_{11}	$\sqrt{2}\rho \sin \theta$
	1	x-Tilt, A_{11}	$\sqrt{2}\rho \cos \theta$
2	-2	y-Astigmatism, B_{22}	$\sqrt{6}\rho^2 \sin 2\theta$
	0	Defocus, A_{20}	$\sqrt{3}(2\rho^2 - 1)$
	2	x-Astigmatism, A_{22}	$\sqrt{6}\rho^2 \cos 2\theta$
3	-3	y-Triangular Astigmatism, B_{33}	$\sqrt{8}\rho^3 \sin 3\theta$
	-1	y-Primary Coma, B_{31}	$\sqrt{8}(3\rho^3 - 2\rho) \sin \theta$
	1	x-Primary Coma, A_{31}	$\sqrt{8}(3\rho^3 - 2\rho) \cos \theta$
	3	x-Triangular Astigmatism, A_{33}	$\sqrt{8}\rho^3 \cos 3\theta$
4	0	Spherical Aberration, A_{40}	$\sqrt{5}(6\rho^4 - 6\rho^2 + 1)$

Table 2. The names of the Zernike primary aberration coefficients.

The A and B terms are referred to as the symmetric and non-symmetric coefficients, respectively, in the units of waves. If the phase is known as a function, then the rms Zernike coefficients A_{nm} and B_{nm} can be calculated using

$$A_{nm} = \frac{1}{\pi} \sqrt{\frac{2(n+1)}{1+\delta_{m0}}} \int_0^{2\pi} \int_0^1 \varphi(\rho, \theta) R_{nm}(\rho) \cos m\theta \rho d\rho d\theta \quad (10a)$$

$$B_{nm} = \frac{1}{\pi} \sqrt{\frac{2(n+1)}{1+\delta_{m0}}} \int_0^{2\pi} \int_0^1 \varphi(\rho, \theta) R_{nm}(\rho) \sin m\theta \rho d\rho d\theta \quad (10b)$$

where δ_{m0} is the Kronecker delta function and the integrals' coefficients are normalizing constants. The names of the coefficients of primary aberrations which we are going to discuss in this paper are given in Table 1. From the aberration coefficients, we can calculate the wavefront error, the Strehl ratio (Mahajan, 2005), focal length (Mafusire and Forbes, 2011b) and the beam quality factor (Mafusire & Forbes, 2011a). The Strehl ratio is defined as the ratio of the maximum axial irradiance of an aberrated beam over that of a diffraction limited beam with the same aperture size. The general definition of the Strehl ratio is given by (Mahajan, 2001; Mahajan, 2005).

$$S = \left(\frac{\left| \int_0^{2\pi} \int_0^1 \psi(\rho) e^{i\phi(\rho, \theta)} \rho d\rho d\theta \right|^2}{\int_0^{2\pi} \int_0^1 \psi(\rho) \rho d\rho d\theta} \right)^2 \quad (11)$$

$$\approx e^{-(\Delta\varphi^2)}$$

In the last result, we have used an approximation of the Strehl ratio for small aberrations less than the wavelength of the radiation. For a better understanding on the wavefront error and Strehl ratio, the interested reader is referred to Mahajan (Mahajan, 2001; Mahajan, 2005).

3.2 Calculation of Zernike primary aberrations from CFD density data

A practical way to extract Zernike coefficients from CFD density data is to generate phase data in layers normal to the propagation direction of the beam using Eq. 4 with distance l between layers. A Taylor polynomial fit given by Eq. 12 is made to each plane to create a phase function in Cartesian form.

$$\begin{aligned} \varphi(x, y) = & p_{00} + p_{10}x + p_{01}y + p_{02}y^2 + p_{11}xy + p_{20}x^2 + p_{03}y^3 + p_{12}xy^2 + p_{21}x^2y \\ & + p_{30}x^3 + p_{04}y^4 + p_{13}xy^3 + p_{22}x^2y^2 + p_{31}x^3y + p_{40}x^4 + p_{05}y^5 + p_{14}xy^4 \\ & + p_{23}x^2y^3 + p_{32}x^3y^2 + p_{41}x^4y + p_{50}y^5 \end{aligned} \quad (12)$$

We can then convert the Cartesian coordinates to cylindrical coordinates using $x = r \cos\theta$ and $y = r \sin\theta$, replace r with ρ/a then reduce the resultant function. If we substitute this function into Eq. 10 we can then extract the *rms* primary Zernike coefficients which are given by Eq. 13. These equations enable us to calculate the Zernike coefficients from the phase distribution data. With these data, it is possible to work out the phase change experienced by the laser beam as it propagates through each layer. If we have data for the next layer, and the one after that, then it is possible to work out the impact on the laser beam as the output from one layer becomes the input for the next layer. If the density of each layer is known from a CFD model, then propagating the laser beam through the medium in this manner will approximate the total aberrations imparted to the laser beam, assuming the propagation distance is small. In our simulations we have restricted our description of the aberrations to the fourth order, but of course one can make the expansion as accurately as one desires.

$$\begin{aligned}
A_{00} &= \frac{1}{24}(a^2(a^2(3p_{04} + p_{22} + 3p_{40}) + 6p_{02} + 6p_{20}) + 24p_{00}) \\
A_{11} &= \frac{1}{96}a(a^2(3(a^2(p_{14} + p_{32} + 5p_{50}) + 8p_{30}) + 8p_{12}) + 48p_{10}) \\
B_{11} &= \frac{1}{96}a(a^2(3a^2(5p_{05} + p_{23} + p_{41}) + 24p_{03} + 8p_{21}) + 48p_{01}) \\
A_{20} &= \frac{1}{16\sqrt{3}}(a^2(a^2(3p_{04} + p_{22} + 3p_{40}) + 4p_{02} + 4p_{20})) \\
A_{22} &= \frac{1}{8\sqrt{6}}(a^2(-3a^2(p_{04} - p_{40}) - 4p_{02} + 4p_{20})) \\
B_{22} &= \frac{1}{16\sqrt{6}}(a^2(3a^2(p_{13} + p_{31}) + 8p_{11})) \\
A_{31} &= \frac{1}{120\sqrt{2}}(a^3(3(a^2(p_{14} + p_{32} + 5p_{50}) + 5p_{30}) + 5p_{12})) \\
B_{31} &= \frac{1}{120\sqrt{2}}(a^3(3a^2(5p_{05} + p_{23} + p_{41}) + 15p_{03} + 5p_{21})) \\
A_{33} &= -\frac{1}{40\sqrt{2}}((a^3(a^2(3p_{14} + p_{32} - 5p_{50}) + 5p_{12} - 5p_{30})) \\
B_{33} &= \frac{1}{40\sqrt{2}}(a^3(a^2(-5p_{05} + p_{23} + 3p_{41}) - 5p_{03} + 5p_{21})) \\
A_{40} &= \frac{1}{48\sqrt{5}}(a^4(3p_{04} + p_{22} + 3p_{40}))
\end{aligned} \tag{13}$$

3.3 Piston as a measure of average density

It is well known that piston is the average phase of a wavefront (Mahajan, 1998; Dai, 2008). In our formulation, Zernike coefficients are in the units of waves and phase is in radians, so that the piston is related to the average phase by $A_{00} = \frac{1}{2\pi} \bar{\varphi}$. Substituting for the phase as defined in Eq. 4 and introducing a new term, $L = l / \lambda$, the piston becomes $\bar{\varphi} = L\bar{N}$ where \bar{N} is the average refractive index. Finally, combining these two equations and then substituting for the average refractive index as defined by Eq. 3, the result is

$$A_{00} = L(G\bar{\sigma} + 1) \tag{14}$$

The equation implies that the piston measured for a beam having passed through a medium of length, l , is directly proportional to average density in that medium. This means that if you have divided a propagation path of a laser beam in an aero-optic medium to steps each of length, l , then the variation of piston for each step is actually the variation in average density. This equation can also mean that if the piston is known after the beam has passed through such a medium, then we can use it to calculate the total average density of that path. If the density in the path is a constant, or $\sigma = \bar{\sigma}$, measuring piston can now be used to measure density. This would be true if all aberrations, except for piston, are zero, or at the very least, very small compared to piston. We can also conclude that the greater piston is compared to other aberrations, the more uniform the medium is.

3.4 Extraction of the optical parameters from the CFD model

Returning to our gas lens, CFD density data was extracted from 351 planes equally spaced along the length of the pipe, so that we model the pipe as 350 layers, or "individual gas lenses" placed one after another, each of length 4.1 mm. Density data is converted to refractive index data using the Gladstone-Dale's law and from this, the phase, φ is calculated using Eq. 4 with l set to 4.1 mm. A Taylor polynomial surface fit was carried out for each plane to create a phase function in Cartesian form. The coefficients thus acquired

were used to calculate the Zernike coefficients, using Eq. 13, resulting in 350 sets of primary Zernike coefficients for the spinning pipe gas lens. It is instructive to partition the SPGL data into sections based on the known fluid behaviour discussed earlier, as illustrated in Fig. 9. These are the end sections, labelled A, which are just over 20 cm each, the central section, C, of about 20 cm. This leaves the two sections in-between, B, each about 40 cm. Sections A are unheated. Let us assume that the laser beam is propagating through the pipe with its axis coinciding with that of the pipe. The beam should be small enough not to experience any diffraction with the pipe walls. For that reason we choose a , of size 0.371 cm against a pipe of radius 1.83 cm. The size of the beam was set at 0.548 cm, the same beam size we used as in the experiment (Mafusire, 2008a). We further assume that the only aberrations the beam experiences are from the medium and not from diffraction due to the propagation itself.

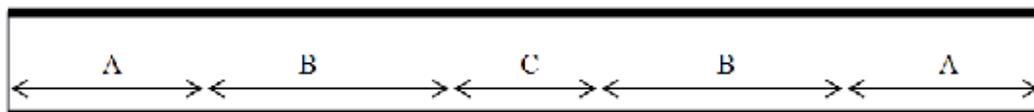


Fig. 9. Sections for analysing the spinning pipe gas lens based on the fluid behaviour discussed earlier.

We now present the graphs of all the primary aberrations as shown in Fig. 10. The aberrations have been organised starting with piston, followed by defocus, spherical aberration, tilt, coma and then astigmatic aberrations, i.e., astigmatism and triangular astigmatism (see Figure 10 (a)-(e)). Piston, Fig. 10 (a), has a characteristic curve which shows local average density, related to the overall phase delay experienced by the beam (Eq. 14). It is maximum in section A of the gas lens and minimum in section C. Section A is where the pipe is not heated and section C is the hot section where rotational motion is dominant. Section B is dominated by a phase gradient. This is the section in which the hot outgoing air mixes with the cool incoming air. We might call it the mixing length. This is the turbulent section of the pipe; the source of aberrations. At the same time, piston is much larger compared to other aberrations. Local piston has an average size of about 3.76λ whereas the second most dominant one is defocus which has an average less than -0.005λ , a factor of about 700. This implies that the density in each slice is almost uniform. From this, we can tell that the lens is very weak. Considering defocus, Fig. 10 (b), we notice immediately that focusing takes place in two parts of the pipe: the sections labelled B, with a large contribution from the region interfacing with section A, reaching a local maximum of -0.015λ . Along its length, the SPGL has two centres of focus (sections B), thus making the SPGL very difficult to align. This confirms the 3D profile in Fig. 6 (a), that the lensing action of the SPGL comes from the mixing of hot and cold air.

The higher order aberrations increase dramatically in section B, suggesting that it is the mixing that gives rise to lensing also has a deleterious effect on the laser beam. This is because the mixing of hot and cold air creates local random varying density, which generates aberrations, the effect of which should increase with temperature and/or rotation speed. Spherical aberration reaches a local maximum of 0.0085λ a magnitude half the size of defocus. Tilt also increases in the same region with a maximum of around 0.00002λ . This behaviour is similar to the behaviour exhibited by coma (Fig. 10 (e)) though the values are about 10 times smaller.

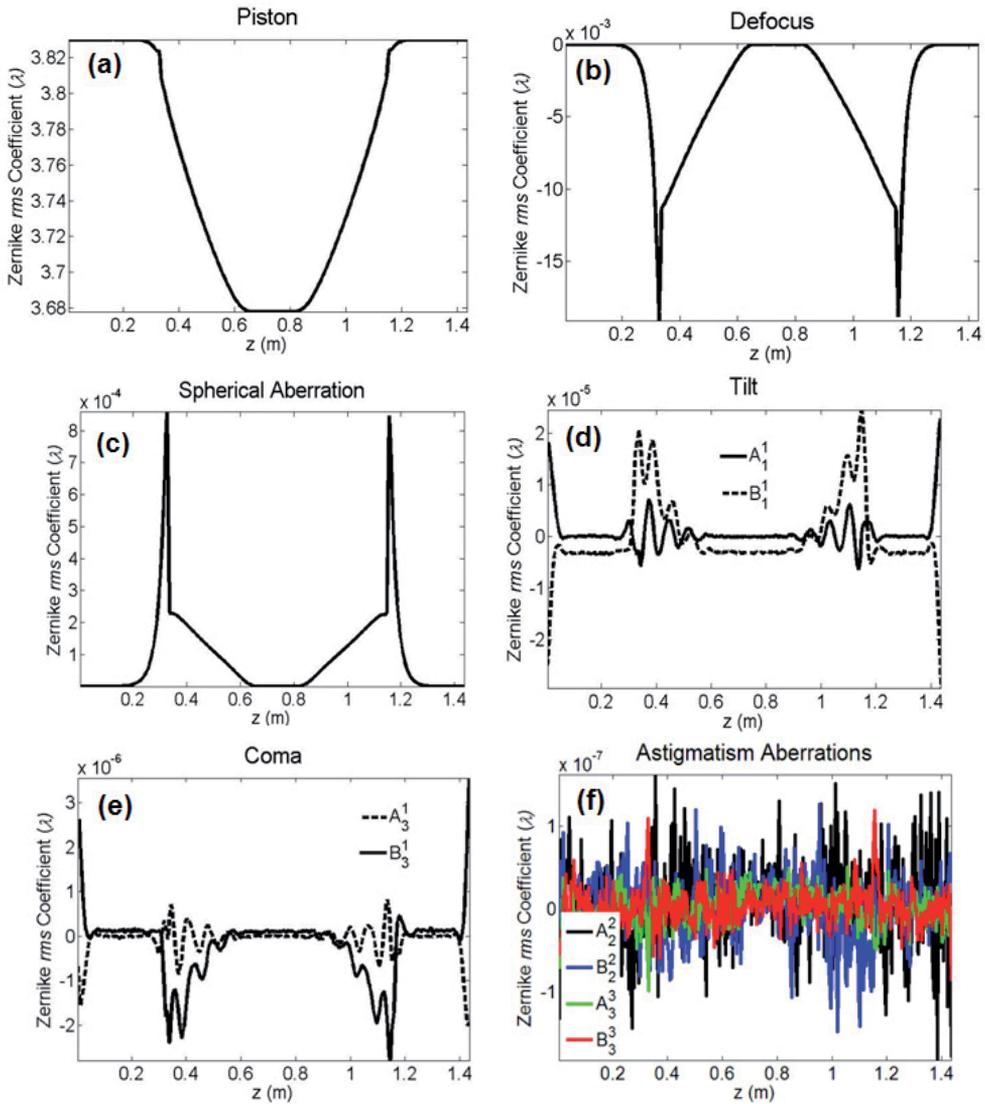


Fig. 10. Local optical aberration distribution along the SPGL calculated from the CFD density data

The beam quality factor distribution (Figure 11 (a)) in the SPGL confirms the aberration distribution. The beam quality factor is highest (suggesting a poor beam) at the same points where spherical aberration is highest. Of all the aberrations, spherical aberration has the largest coefficient. This confirms that spherical aberration is biggest contributor to beam quality deterioration in a gas lens. The wavefront error and Strehl ratio provide further proof of this. However, the important thing to note is that these parameters prove that the gas lens does not cause deterioration of the laser beam by that much. An unaberrated Gaussian beam has an M^2 of 1, whereas the model shows local values of M^2 of about 1.57. This results in a very low local wavefront error of about $0.0001 \lambda^2$ on average. The Strehl ratio (Fig. 11b) is almost always 1 throughout the SPGL except in the mixing length where it

drops by an infinitesimally small amount. The overall beam quality factor for the entire gas lens was found to be 2.5071 in both axes. The only disappointing aspect of the lens was its focal length which was found to be 5.03 m in both axes. This confirms that the gas lens is a very weak lens.

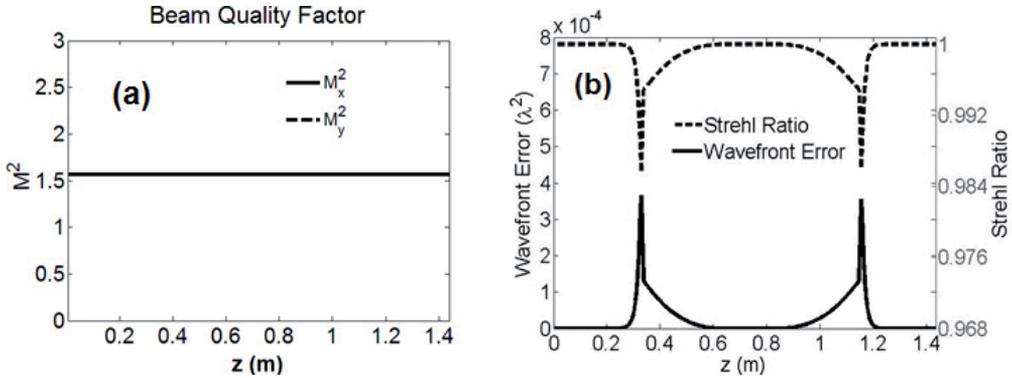


Fig. 11. The local beam quality factor (a), wavefront error and Strehl ratio (b) distributions along the pipe calculated from the SPGL CFD density data

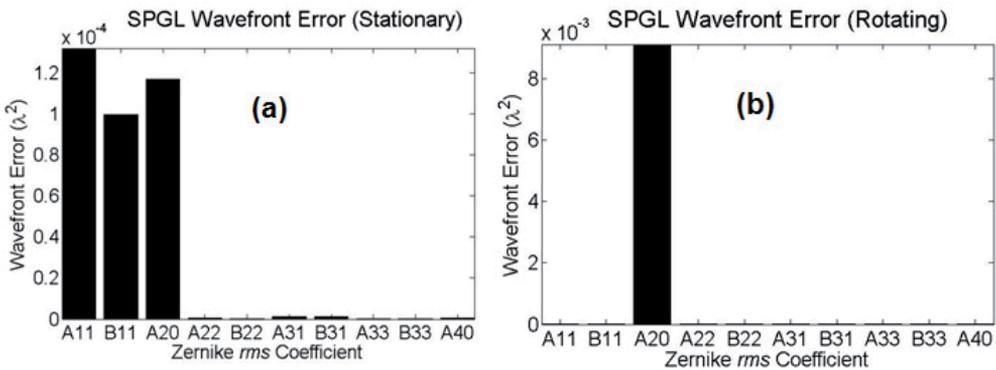


Fig. 12. Global primary optical aberrations except piston calculated from the SPGL CFD density data summarising the overall SPGL for a stationary (a) and rotating (b) heated pipe.

Now let us summarise the performance of the SPGL by looking at the CFD calculated global wavefront error for a heated stationary (Figure 12 (a)) and a rotating (Figure 12 (b)). The results for a stationary SPGL show the dominance of tilt in both axes, both with have values of at least $10^{-4} \lambda^2$. Defocus has a value around the average of the two. This confirms that there is, indeed some focusing before rotation begins, though it is still very small, about $1.2 \times 10^{-4} \lambda^2$. However, as rotation commences, there is a significant increase in defocus, to $9.2 \times 10^{-3} \lambda^2$, which dominates other aberrations, including tilt, by a large amount. On the other hand, x - and y -tilt drop slightly from $1.3 \times 10^{-4} \lambda^2$ and $10^{-4} \lambda^2$ to just below $10^{-5} \lambda^2$. This confirms the fact that before rotation, the SPGL is dominated by tilt, due to gravitational distortion, but the effect is completely reduced under steady state rotation leaving defocus completely dominant. In other words, we have customised the density gradient to produce a lens.

3.5 Experimental verification of the SPGL model

Since density is directly proportional to refractive index, the phase of a laser beam propagating through a gas lens will be altered depending on the refractive index distribution. For completeness, we present a summary of the optical investigation on the aberrations generated by an actual SPGL that was characterised with a Shack-Hartmann wavefront sensor (Mafusire et al, 2008a). An expanded HeNe laser beam steered by flat mirrors is made to propagate through the lens. A Shack-Hartmann wavefront sensor was placed just behind the lens and used to measure the beam's quality and phase aberrations for rotation speeds up to about 17 Hz for wall temperatures 351, 373, 400 and 422 K.

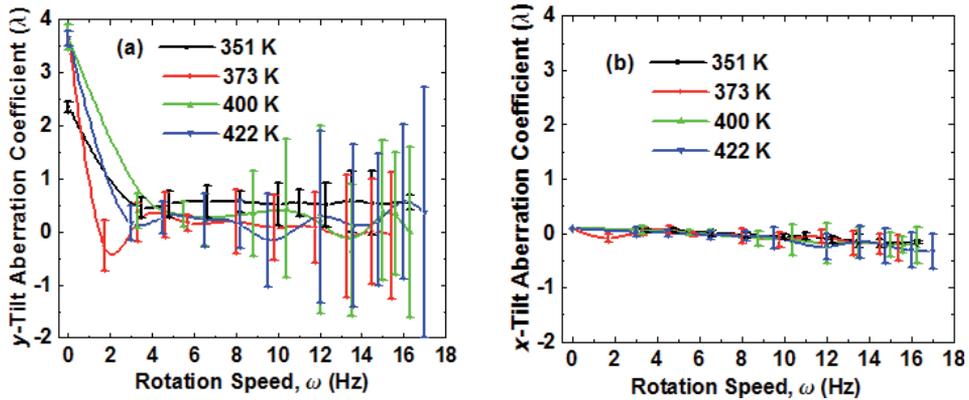


Fig. 13. y -Tilt (a) and x -tilt (b) generated by a spinning pipe gas lens at selected wall temperatures and rotation speeds.

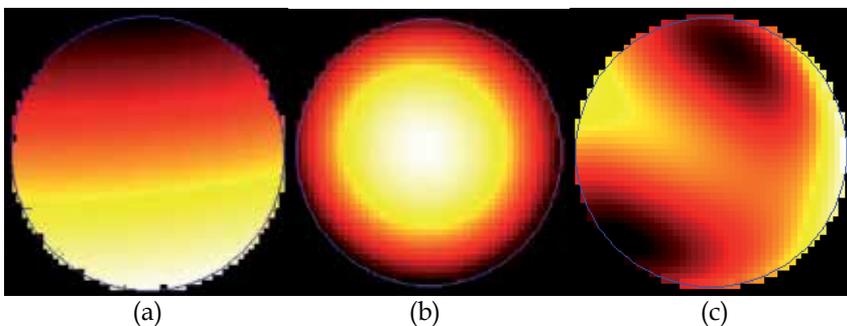


Fig. 14. The phase distribution of the laser beam with: (a) no rotation but heated to 422 K, showing tilt; (b) after rotating the SPGL at 17 Hz, showing significant curvature on the wavefront; and (c) same conditions as in (b) but with defocus and tilt removed, revealing the higher order aberrations.

The first result confirms the fact that rotation removes distortions which are caused by gravity: y -tilt, which is induced by gravity is reduced to a bare minimum as soon as rotation commences (Fig. 13 (a)) whereas x -tilt remains very small (Fig. 13 (b)) throughout. We can observe the same effect by looking the phase distribution before and during rotation. Figs. 14 (a) and (b) show the phase distribution before and during rotation, respectively. The phase maps are dominated by y -tilt (Fig. 14 (a)) before rotation, and defocus (Fig. 14 (b))

during rotation. However, digital removal of defocus and tilt reveals the presence of higher order aberrations, Fig. 14 (c). This phase map helps illustrate the other result observed during the experiment, the effect of the SPGL on the beam quality factor.

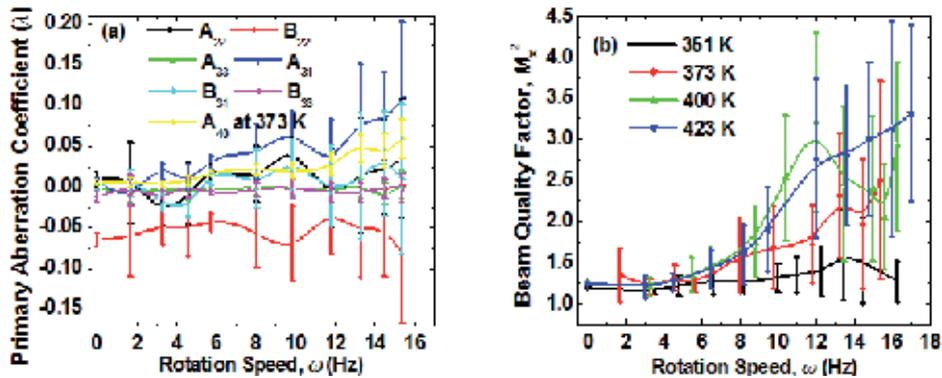


Fig. 15. (a) Higher order aberrations introduced by the SPGL; (b) increase in M_x^2 with rotation speed and temperature as a direct result of the aberrations in (a).

The other aberrations increase in magnitude as the rotation speed and/or wall temperature is increased (Fig. 15 (a)) thereby increasing the beam quality factor, M^2 (Fig. 15 (b)). This confirms the fact that the gas lens also generates aberrations which are increasing in power as the lens becomes stronger.

4. Optical turbulence

We now ask if the SPGL may be used as a controlled turbulence medium in the laboratory for the study of the propagation of optical fields through the atmosphere. The basis of this question is the fact that the SPGL introduces aberrations, and further that these aberrations may be controlled through the rotation speed and temperature of the pipe. Optical turbulence may also be described by aberrations, except that the weighting of the aberrations should take on a particular form. Characterising the aberrations in the SPGL from a turbulence perspective shows that the turbulence is uniform and isotropic near the pipe axis (about which it is spun), becoming non-uniform and anisotropic at the pipe boundary. A modified von Karman turbulence model (Andrews & Phillips, 2005) is used to analyse the turbulence strength along the pipe axis, and we find that the turbulence strength increases with rotation speed and pipe-wall temperature allowing for 'controlled' turbulence in the laboratory: our simple system allows for a controlled adjustment of the refractive index structure constant by more than two orders of magnitude.

The results of Fig. 16 illustrate this: in Fig. 16 (a) the log of the structure constant, which is a measure of the turbulence strength, is adjusted by two orders of magnitude as the pipe parameters of rotation speed and wall temperature are adjusted. Other supporting data (not shown here) confirms that the aberration weighting is correct for a particular atmospheric turbulence model – the modified von Karman turbulence model. This model has an inner and outer scale, the smallest and largest scales of the turbulence in the atmosphere respectively, relating directly to the smallest and largest fluid flow structures. These parameters can be measured in the pipe, and are shown in Fig. 16 (b). The pertinent point is that all the characteristics of the turbulence can be measured and therefore simulated in this simple aero-optic device, allowing for easy experiments of atmospheric turbulence in the laboratory.

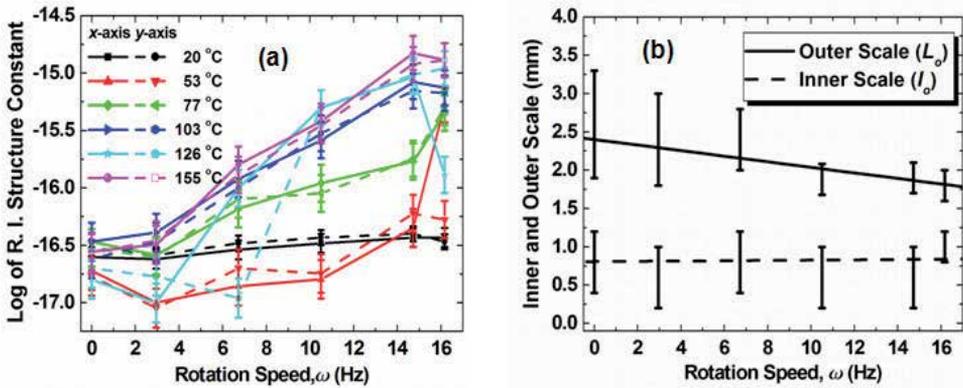


Fig. 16. (a) Refractive index structure constant at selected rotation speeds and wall temperature. (b) Inner and outer scales at the selected rotation speeds. These are standard parameters used to describe optical turbulence, and illustrate that our aero-optic may be used as a simulator of turbulence.

We also consider the optical aberrations imparted to the field when propagating near the boundary layer (Fig. 17), and find the phase distortions to the laser beam to be dominated by x -astigmatism with y -astigmatism and tilt, increasing dramatically in magnitude at the highest rotation speed and temperature. It is apparent that, in the spinning pipe gas lens, the parent flow is derived from the rotation of the pipe.

It is this rotation, together with the physical size of the lens, which limit the outer scale of the turbulence in the pipe. On the other hand, the outer scale is not much larger than the inner scale for a small inertial sub range, the range of scales between which the turbulence is isotropic, homogenous and independent of the parent flow.

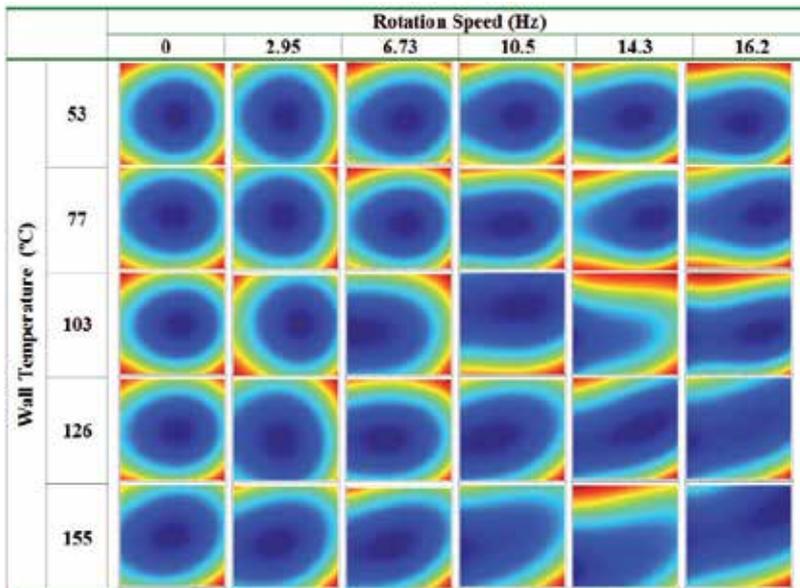


Fig. 17. Phase maps of the beam as it propagates near the boundary layer of the flow, for various wall temperatures and rotation speeds.

5. Summary

Aero-optics has found some novel applications of late, recently reviewed by Michaelis et al (Michaelis et al, 2006). These include long range telescoping elements, replacing high power laser windows to overcome damage threshold problems, adaptive lenses for delivery of high power laser beams in space propulsion experiments, and potential applications in laser fusion, control of peta-watt laser beams, photo-lithography with virtual capillaries and possibly novel guiding media for laser accelerators. Certainly if the quality of gas lenses could be improved, then the virtually limitless damage threshold of such lenses would make them ideal for most high power laser applications. The only drawback of most aero-optical devices is the distortions introduced to the laser beam due to imperfect control of the fluid, but as we have shown here, even this property may be exploited to simulate atmospheric turbulence in the laboratory.

In this chapter we have shown that it is possible to control the fluid flow inside a spinning heated pipe such that the density gradient of the air inside the pipe acts as a lens. As the focal length of this lens is a function of the rotation speed of the pipe and the temperature of the pipe wall, one has a variable focal length lens. We have shown focal lengths from infinity down to a couple of meters. We have shown that the lens is unfortunately aberrated, but highlighted that such aberrations in fact match atmospheric turbulence, so that the system may also be used as a simulator of atmospheric turbulence in the laboratory – again in a controlled and adjustable manner. It is this property – that such devices may be controlled – that makes aero-optics such an attractive possibility for future optical devices.

6. Acknowledgment

We would like to thank M.M. Michaelis for significant advice and for providing the original motivation for studying this field, and G. Snedden for his invaluable assistance in executing the CFD commercial code.

7. References

- Andrews, L. C. & Phillips, R. L. (2005). *Laser Beam Propagation through Random Media*, SPIE Press.
- Aoki, Y. & Suzuki, M. (1967). Imaging Property of a Gas Lens", *IEEE Trans. on Microwave Theory & Techniques*, Vol. 15, No. 1.
- Blazek, J., (2001). *Computational Fluid Dynamics - Principles and Applications*, Ch 7 Elsevier, Oxford.
- Berreman, D. W., (1965). Convective Gas Light Guides or Lens Trains for Optical Beam Transmission, *J. Opt. Soc. Am*, Vol. 55, No. 3, pp. 239-247.
- Born, M. & Wolf, E. (1998). *Principles of Optics: Electromagnetic theory of propagation, interference and diffraction of light 7th Ed* Cambridge University Press, Cambridge 517-553.
- Dai, G-m, (2008). *Wavefront Optics for Vision Correction*, SPIE Press.
- Davison, L., (2011). An Introduction to Turbulence Models. *Chalmers University of Technology*. Available from the University of Chalmers website: www.tfd.chalmers.se/~lada/postscript_files/kompendium_turb.pdf
- Forbes, A. (1997). *Photothermal Refraction and Focusing*, PhD Thesis, University of Natal, Durban, South Africa.
- Gloge, D. (1967). Deformation of Gas Lenses by Gravity, *Bell Sys. Tech. J.*, Vol. 46, No. 2, 357-365.
- Kaiser, P. (1967). Measured Beam Deformations in a Guide Made of Tubular Gas Lenses, *Bell Sys. Tech. J.*, Vol. 47, 179-194.

- Kaiser, P. (January, 1970). An improved thermal gas lens for optical beam waveguides, *Bell Sys. Tech. J.* Vol. 49, (January 1970) pp 137-153
- Kellet, B.J.; Griffin, D.K.; Bingham R.; Campbell R.N.; Forbes, A. & Michaelis, M.M. (2008). Space polypropulsion, *Proc. SPIE* Vol. 7005, pp. 70052W-1.
- Lisi, N.; Bucellato, R. & Michaelis, M. M. (1994). Optical quality and temperature profile of a spinning pipe gas lens, *Optics and Laser Technology*, Vol. 26, 25-27.
- Mahajan, V. N., (1998). *Optical Imaging and Aberrations, Part I: Ray Geometrical Optics*, SPIE Press.
- Mahajan, V. N. (2001). *Imaging and Aberrations, Part 2: Wave Diffraction Optics*, SPIE Press.
- Mahajan, V. N., (2005). Strehl ratio of a Gaussian beam, *J. Opt. Soc. Am. A* Vol. 22, No. 9.
- Martynenko, O. G., (1975). Aerothermooptics, *International J. of Heat and Mass Transfer*. 18, 793-796.
- Michaelis, M. M.; Notcutt, M. & Cunningham, P. F. (1986). Drilling by a Gas Lens Focused Laser, *Opt. Comm.* Vol. 59, 369-374.
- Michaelis, M. M.; Dempers, C. A.; Kosch, A. M.; Prause, A.; Notcutt M.; Cunningham, P. F. & Waltham, J., (1991). Gas lens telescopes, *Nature* Vol.353, 547-548.
- Michaelis, M.M.; Forbes, A.; Conti, A.; Nativel, N.; Bencherif, H.; Bingham, R.; Kellett, B. & Govender, K. (2006). Non-solid, non-rigid optics for high power laser systems, *Proc. SPIE* Vol. 6261, pp. 15.1-15.13.
- Mafusire, C. (2006). *Gas Lensing in a Heated Rotating Pipe*, MSc Thesis, University of Zimbabwe, (2006).
- Mafusire, C.; Forbes, A.; Michaelis, M. M. & Snedden, G. (2007). Characterization of a spinning pipe gas lens using a Shack-Hartmann wavefront sensor", *Laser Beam Shaping VIII, Ed. F. Dickey, Proc. SPIE.*, Vol. 6663, 6663H.
- Mafusire, C.; Forbes, A.; Michaelis, M. M. & Snedden, G. (2008a) Optical aberrations in a spinning pipe gas lens", *Opt. Exp.*, Vol. 16, No. 13, 9850-9856.
- Mafusire, C.; Forbes, A.; Michaelis, M. M. & Snedden, G. (2008b). Spinning pipe gas lens revisited, *SA. J. Sc.*, Vol. 104.
- Mafusire, C.; Forbes A.; Michaelis M.M. & Snedden, G. (2010a). Optical aberrations in gas lenses, *Proc. SPIE* 7789, pp. 778908-1.
- Mafusire, C., Forbes, A. & Snedden, G. (2010b). A computational fluid dynamics model of a spinning pipe gas lens, *Proc. SPIE* 7789, pp. 77890Y-1.
- Mafusire, C. & Forbes, A., (2011a). The Beam Quality Factor of Truncated Aberrated Gaussian Laser Beams, *J. Opt. Soc. Am. A*, Vol. 28, No. 7, 1372-1378.
- Mafusire, C. & Forbes, A., (2011b). The Mean Focal Length of an Aberrated Lens, *J. Opt. Soc. Am. A* Vol. 28, No. 7, 1403-1409 (2011).
- Mafusire, C. & Forbes, A., (2011c) Controlling optical turbulence in the laboratory, *Appl. Opt.* (submitted for publication).
- Marcuse, D. (1965). Theory of a Thermal Gradient Gas Lens, *IEEE Trans. on Microwave Theory & Technology*, Vol. MMT-13, No. 6, 734-739 (1965).
- Notcutt, M; Michaelis, M. M.; Cunningham, P. F. & Waltham, J. A. (1988). Spinning Pipe Gas Lens, *Optics and Laser Technology*, 20(5), 243-250.
- Siegman, A. E. (1999). Laser Beams and Resonators: The 1960s, *IEEE Jour. of special topics in Quant. Elec.* Vol. 20, No. 5, 100-108.
- Steier, W. H. (1965) Measurements on a Thermal Gradient Gas Lens, *IEEE Trans. on Microwave Theory & Technology*, Vol. MMT-13, No. 6, 740-748.
- Zhao, Y. X., Yi, S. H., Tian, L. F., He, L & Cheng, Z. Y., (2010) An experimental study of aero-optical aberration and dithering of supersonic mixing layer via BOS, *Science China: Physics, Mechanics & Astronomy* Vol. 53, No. 1, 89-94.

Edited by L. Hector Juarez

The content of this book covers several up-to-date topics in fluid dynamics, computational modeling and its applications, and it is intended to serve as a general reference for scientists, engineers, and graduate students. The book is comprised of 30 chapters divided into 5 parts, which include: winds, building and risk prevention; multiphase flow, structures and gases; heat transfer, combustion and energy; medical and biomechanical applications; and other important themes. This book also provides a comprehensive overview of computational fluid dynamics and applications, without excluding experimental and theoretical aspects.

Photo by woottigon / iStock

IntechOpen

