

IntechOpen

# Technology, Science and Culture

A Global Vision, Volume III

*Edited by Luis Ricardo Hernández  
and Martín Alejandro Serrano Meneses*





---

# Technology, Science and Culture - A Global Vision, Volume III

*Edited by Luis Ricardo Hernández  
and Martín Alejandro Serrano Meneses*

Published in London, United Kingdom

---



## IntechOpen





*Supporting open minds since 2005*



Technology, Science and Culture – A Global Vision, Volume III  
<http://dx.doi.org/10.5772/intechopen.94928>  
Edited by Luis Ricardo Hernández and Martín Alejandro Serrano Meneses

#### Contributors

Gabriela Stefanie Solís Santos

© The Editor(s) and the Author(s) 2022

The rights of the editor(s) and the author(s) have been asserted in accordance with the Copyright, Designs and Patents Act 1988. All rights to the book as a whole are reserved by INTECHOPEN LIMITED. The book as a whole (compilation) cannot be reproduced, distributed or used for commercial or non-commercial purposes without INTECHOPEN LIMITED's written permission. Enquiries concerning the use of the book should be directed to INTECHOPEN LIMITED rights and permissions department ([permissions@intechopen.com](mailto: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 these 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 London, United Kingdom, 2022 by IntechOpen  
IntechOpen is the global imprint of INTECHOPEN LIMITED, registered in England and Wales, registration number: 11086078, 5 Princes Gate Court, London, SW7 2QJ, United Kingdom  
Printed in Croatia

#### British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Additional hard and PDF copies can be obtained from [orders@intechopen.com](mailto:orders@intechopen.com)

Technology, Science and Culture – A Global Vision, Volume III  
Edited by Luis Ricardo Hernández and Martín Alejandro Serrano Meneses  
p. cm.  
Print ISBN 978-1-83969-442-4  
Online ISBN 978-1-83969-443-1  
eBook (PDF) ISBN 978-1-83969-444-8

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

**5,800+**

Open access books available

**142,000+**

International authors and editors

**180M+**

Downloads

**156**

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 (BKCI)  
in Web of Science Core Collection™

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)







# UDLAP®

Universidad de las Américas  
Puebla

*Technology, Science and Culture:  
A Global Vision, Volume III*



# Technology, Science and Culture: A Global Vision, Volume III 2020

## Editors

Luis Ricardo Hernández  
Martín Alejandro Serrano Meneses

## Knowledge area co-editors

Aura Matilde Jiménez Garduño  
Nelly Ramírez Corona  
José Luis Sánchez Salas  
Enrique Ajuria Ibarra  
Roberto Rosas Romero

We continue with this series discussing on research topics related to the fields of Food Science, Intelligent Systems, Molecular Biomedicine, Water Science, and Creation and Theories of Culture. Our aims are to discuss the newest topics, theories, and research methods in each of the mentioned fields, to promote debates among top researchers and graduate students, and to generate collaborative works among them.

The interactions of recognized specialists in each field and graduate students, through different meetings, generated very interesting discussions, which are presented in this book. Thus, Dr. Luis A. Pardo, from the Molecular Biology of Neuronal Signals Max Planck Institute for Experimental Medicine, contributes with the article titled “Targeting the voltage-gated potassium channel Kv10.1 for cancer therapy”. Dr. Marco Carli, Associate Professor of the Department of Engineering at the Università degli Studi 'Roma TRE', Roma, Italy, explored, along with his co-author Federica Battisti, the quality of experience for immersive media with the work “QoE and immersive media: a new challenge”. Dr. Sandra Harding, Distinguished Research and Emeritus Professor of New York University, wrote the article “Strong objectivity for new social movements”. Dr. Vijay P. Singh, Distinguished Professor, Regent Professor, and Caroline and William N. Lehrer Distinguished Chair in Water Engineering at Texas A&M University contributes with the article “Challenges in flood management”. Dr. R. Paul Singh, Distinguished Professor of Food Engineering of the Department of Biological and Agricultural Engineering at University of California wrote the article “A quest for sustainability in the food enterprise”. Finally, graduate students of the Universidad de las Américas Puebla further present their key findings in a series of articles.

We believe that interactions between students and high-level researchers of different areas contribute to the creation of multidisciplinary points of view generating the advancement of science.



United Nations  
Educational, Scientific and  
Cultural Organization



**UDLAP**  
UNIVERSIDAD DE LAS AMÉRICAS PUEBLA

UNESCO Chair on  
Hydrometeorological Risks,  
Universidad de las Américas Puebla

The number and impact of water-related natural disasters has increased since the middle of last century. As result of increased climate variability and the effects of global warming, the hydrometeorological risk has increased and spread, while the resilience of societies, in many cases, is not adequate. Consequently, the risk has increased. Floods and droughts, particularly in a changing climate, require greater understanding to generate better forecasts and proper management of these phenomena. Mexico, like other countries in the world, and of course in Latin America and the Caribbean region, suffers from both weather extremes.

The UNESCO Chair on Hydrometeorological Risks, held at the University of Americas Puebla, is devoted to the analysis, measurement, modelling and management of extreme hydro-meteorological events in the context of a more urbanized world, climate change and further vulnerable regions. Focused on the development of basic and applied research for the design of adaptation and mitigation measures, dissemination and preparation of decision makers as well as the public. In its activities keeps a gender focus, directed in particular to reduce the vulnerability of women to hydrometeorological disasters.

The Chair acts in the following fields:

1. Hydrometeorological risks and climate change.
2. Modelling and forecasting of hydrometeorological risks.
3. Integrated management of hydrometeorological risks.
4. Gender and hydrometeorological risks.

A detailed description of the UNESCO Chair on Hydrometeorological Risks, members and publications, can be obtained at its Website <https://www.udlap.mx/catedraunesco/>

The Chair publish a quarterly Newsletter, in Spanish and English, that can be consulted at <https://www.udlap.mx/catedraunesco/newsletters.aspx>



# UDLAP®

## HIGHLY SPECIALIZED DOCTORAL PROGRAMS

**Doctoral Programs in:** (Doesn't require MA degree for entry)

- Molecular Biomedicine
- Food Science
- Water Sciences
- Intelligent Systems

**Doctoral Program in:** (Requires MA degree for entry)

- Creation and Culture Theory

**All accepted students have a 100% tuition scholarship and a monthly maintenance for the entire duration of the program.**

**For more information contact:**  
Research and Graduate Studies Office  
Tel: +52 (222) 2 29 20 00 exts. 2725 y 6005  
[informes.doctorados@udlap.mx](mailto:informes.doctorados@udlap.mx)

[www.udlap.mx](http://www.udlap.mx)

ACADEMIC EXCELLENCE

CHOOSE UDLAP



# Contents

Targeting the Voltage-Gated Potassium Channel Kv10.1 for Cancer Therapy <i>Luis A. Pardo</i>	1
QoE and Immersive Media: A New Challenge <i>Federica Battisti and Marco Carli</i>	9
Strong Objectivity for New Social Movements <i>Sandra Harding</i>	21
Challenges in Flood Management <i>Vijay P. Singh</i>	31
A Quest for Sustainability in the Food Enterprise <i>R. Paul Singh</i>	45
Evaluation of the Cytotoxic Activity of a Species of the Buddleja Genus in a Prostate Cancer Cell Line <i>Sofía Isabel Cuevas Cianca, Luis Ricardo Hernández and Irene Vergara Bahena</i>	57
Designing Magnetic Mesoporous Nanoparticles for Cancer Therapy <i>Jessica Andrea Flood-Garibay, Kenneth J. Balkus Jr and Miguel Ángel Méndez-Rojas</i>	65
Exoplanet Research Using Machine Learning and Multiresolution Analysis Techniques <i>Miguel Jara-Maldonado, Vicente Alarcon-Aquino and Roberto Rosas-Romero</i>	75
Network Intrusion Detection Using Dendritic Cells and Danger Theory <i>David Limon-Cantu and Vicente Alarcon-Aquino</i>	89
Automatic Terrain Perception in Off-Road Environments <i>Ethery Ramírez-Robles and Oleg Starostenko</i>	107
Analysis of Voice and Magnetic Resonance Images to Assist Diagnosis of Parkinson's Disease with Machine Learning <i>Gabriel Solana-Lavalle and Roberto Rosas-Romero</i>	121
A Systematic Review of Sensitivity Analysis of Activated Sludge Modeling <i>Rafael Andrés Borobio-Castillo, José Manuel Cabrera-Miranda, Alberto Vargas-Hidalgo and Benito Corona-Vásquez</i>	135
Microbial Photobioelectrochemical Systems: A Scoping Review <i>Luis Erick Coy-Aceves, José Luis Sánchez-Salas, Mónica Cerro-López, Miguel Ángel Méndez-Rojas and Benito Corona-Vásquez</i>	163
Methods for Persistent Organic Pollutants Removal in Wastewater: A Review <i>Valérie Pihen and Jose Luis Sanchez-Salas</i>	193

A Critical Review on Algal-Bacterial Granular Sludge Process as Potential Economical Alternative to AOPs for Textile Wastewater Treatment <i>Celina Sanchez-Sanchez, Guillermo Baquerizo and Ernestina Moreno-Rodríguez</i>	207
Isolation and Identification of Molds in Selected Dried Fruits and Seeds Sold in Bulk in México <i>David González-Albarrán, Aurelio López-Malo and Enrique Palou</i>	227



# Targeting the Voltage-Gated Potassium Channel Kv10.1 for Cancer Therapy

*Luis A. Pardo*

## Abstract

Survival and quality of life of cancer patients have improved in the last decades. However, some forms of cancer still escape current treatment options and continue to have an ominous prognosis. A plausible strategy to change this situation is to identify unexploited pathways in the cancer cell that open genuinely new therapeutic pathways. Ion channels are among such targets since they participate in all steps in the cancer process, from initiation through growth and metastasis to drug resistance. In some cases, ion channels can thus serve as therapeutic targets. Kv10.1 is particularly well suited for this purpose because the channel appears outside of the brain almost exclusively in cancer cells. Recent research showed that besides its functions as a canonical ion channel, Kv10.1 is required by dividing cells to complete division. For this, healthy cells express the channel only during a short period in the cell division cycle. Cancer cells, rather than increasing the channel's expression, maintain relatively constant levels throughout their lives, which confers a selective advantage and favors tumor progression. The mechanisms leading to abnormal expression and its consequences, and how we can take advantage of this knowledge to improve current cancer treatments will be discussed.

**Keywords:** Kv10.1, ion channels, cell cycle, cancer target

## 1. Introduction

The aim of our research is based on the design of therapeutic approaches that make use of the Kv10.1 voltage-gated channel as a target. Kv10.1 is a voltage-gated potassium channel, which was discovered in the 60s and has been the main focus of our electrophysiological and molecular biology research for many years. During our early experiments, we discovered, contrary to the typical perspective of a voltage-gated ion channel, that Kv10.1 plays a key role in the development of tumors, and even though there is still a lot to understand about how Kv10.1 helps tumor cells to survive, we have been able to unravel many of its mechanisms. An overview of those discoveries will be presented in the present work.

## 2. The importance of ion channels in oncology

Ion channels represent the second biggest protein family in the human genome after GPCRs. These proteins allow the flux of ions through the plasma membrane.

Kv10.1 (also known as Eag1) is dependent on membrane voltage, and after a depolarization, it allows the efflux of potassium ions out of the cell into the extracellular space in a time range of milliseconds. After the exit of K ions, the consequent change in membrane voltage will then function as a cellular signal. If we analyze the protein structure of Kv10.1 we identify a transmembrane region and a big intracellular domain which represents almost 50% of the whole protein [1]. Such a long cytoplasmic domain reveals the importance of the channel not only as a K ion gate but also as an interactive partner of signaling molecules. Canonical functions of voltage-gated potassium channels encompass action potential repolarization, control of resting potential and excitability and volume control. However, more than 70 different genes encode voltage-gated potassium channels and they are expressed in excitable cells as well as in non-excitable cells. This fact suggests that voltage-gated ion channels are involved in more processes beyond action potential definition, as has been demonstrated in the last decades.

### **3. Expression profile of Kv10.1 in healthy tissues and cancer**

Kv10.1 owns very distinctive electrophysiological features that allow us to identify it in different cells. Its activation depends on the membrane potential before the stimulus [2] and when we evaluate it through a whole-cell configuration of the patch-clamp technique, we can observe how the speed of its activation increases as the membrane potential before the stimulus becomes less negative. This feature endorses Kv10.1 with the ability to “remember” the previous electrical status of the membrane and regulate its gating accordingly. This phenomenon, called Cole-Moore shift can also be analyzed in single-channel experiments, where we observe the same response, channels have a delayed opening when the pre-depolarization potential is at -120 mV contrary to the immediate opening response observed when the pre-depolarization potential is at -50 mV. When the function of Kv10.1 was being unraveled some years ago, major participation during a neuronal action potential was discarded due to its slow gating. Therefore, experiments focused on its role in the synaptic membranes. Our studies on Kv10.1 knockout (KO) mice demonstrated that the channel plays a role in postsynaptic potentiation [3]. When mice cerebellar Purkinje cells were recorded after the electrical stimulus of the granule cell layer neurons (which communicate through parallel fibers to Purkinje neurons) the cell response of KO mice was unaffected to single or low-frequency stimuli. When a train of impulses is applied, the response increases progressively with successive stimuli, but only to a certain point in the wildtype, becoming then constant even if further impulses arrive. On the other hand, when Knockout (KO) mice were recorded, the response of Purkinje cells does not become controlled and continued increasing during stimulation. This effect was only associated with mild behavioral alterations of mice under stress, and therefore, the channel seems to play roles that can be compensated by other channels under less demanding conditions.

Anyhow, during the studies on Kv10.1, we found that its expression is almost exclusively confined to the central nervous system, although our first molecular and functional studies had been made on cancer cells.

### **4. A selective advantage for cancer cells**

Therefore, we looked for the expression of Kv10.1 on a wide variety of human cell lines and cancer samples and we found that it was expressed in 72% of all tumor samples, whereas the healthy tissues where the tumor originates did not

express it [4]. This means that we were in front of a tailored designed cancer target, a protein absent in healthy non-central nervous system tissues but expressed in a vast majority of tumors. In addition, our studies demonstrated that tumors expressing Kv10.1 have a worse clinical behavior compared to tumors negative for the channel. Acute myeloid leukemia showed that mortality increased for Kv10.1 positive leukemias [5]. Also, other authors have reported its potential use for bad prognosis in the ovary, gastric, colon, esophagus and cervix tumors [6].

Moreover, we know that imipramine can specifically block the function of Kv10.1, and when the outcome of patients with brain metastatic tumors taking or other tricyclic antidepressants was compared to patients with similar tumors taking a non-Kv10.1 blocker as an antidepressant, we observed that survival was higher in patients with the Kv10.1 blocking treatment [7]. This result suggests that we could be able to delay tumor growth by blocking Kv10.1 and evidences the biological advantage that cells acquire when its expression begins.

If we look for the phylogeny of Kv10.1 we can identify the whole EAG family in species such as *Trichoplax adhaerens*, long before the appearance of neurons. Therefore, there must be an ancestral function of Kv10.1 that does not involve neuronal activity and excitability [8].

One of the most ancient processes of life is cell division regulation. All cells either divide at least once or descend from the division of another cell, cell division is a very universal process in cells. To divide, a cell must pass through a series of phases that have been well characterized by researchers. The S phase is characterized by the duplication of the DNA content. The M phase is Mitosis when cell division occurs. In between those phases, we find two Gap phases called G1 and G2. G1 is a growth phase when cells prepare to divide and G2 is a checkpoint after the S phase to screen for errors during DNA duplication and if absent, proceed to Mitosis. The role of membrane potential in the process was known for at least fifty years. Clarence Cone showed that the membrane potential of a cell needs to oscillate during cycles of replication [9]. If the dynamics of the membrane potential are blocked then, cell division stops. We generally accept that at the end of G1 a hyperpolarization occurs and then from the S phase to the M phase a depolarization takes place. Those changes are completely dependent on ion channels [10]. Bijlenga et al. have already demonstrated that myoblast express Kv10.1 to fusion, which is a cell cycle-dependent process, however insights into the details of its precise role during the cell cycle were still lacking [11]. Our group evaluated synchronized cancer cells and we could demonstrate that the expression of Kv10.1 changes during the cell cycle and is maximal during the G2 phase of cells, which can be identified by the enrichment of other G2 protein markers [12].

## 5. Mechanisms of action

This elemental role of Kv10.1 made us propose the following question: Is Kv10.1 expressed only in some cells all the time, or in all cells but only for some time? If we assume that only a very small fraction of cells will be at G2 at any given time it is possible that we simply missed expression because it occurs for very short periods. So, when analyzing with more detail healthy tissues in their replicative zones, such as the bottom of colon crypts that have stem cells, or testis that contain G2 arrested cells, we were able to demonstrate that Kv10.1 was expressed in those healthy cells together with G2 markers such as Cyclin B [4, 12]. Therefore, tissues do express Kv10.1 for short periods during replication. When looking for its expression during G2, our group showed that when cells lose expression of Kv10.1 using RNA interference, G2 phases last more time compared to controls. This means that

Kv10.1 somehow accelerates G2 phases and therefore, replication. But how exactly is Kv10.1 speeding up cell division? Well, one of the most important processes during cell division is cytoskeleton rearrangement, specifically, microtubule organization. When Kv10.1 is eliminated from cells, the dynamics of microtubule rearrangement is accelerated, with longer growth periods [13]. These changes correlated with changes in calcium oscillations when analyzed by fluorescent calcium sensors. Cells without Kv10.1 have higher calcium oscillation frequencies. Calcium enters the cells in a voltage-dependent manner, so it makes sense that Kv10.1, which hyperpolarizes the cell stabilizes the entry of calcium, making the oscillations less frequent [13].

Other groups have demonstrated that Kv10.1 functionally interacts with Orai1, a calcium channel. So, we looked for physical proximity between Kv10.1 and Orai1 and found a higher amount of interaction in tumoral cells demonstrated by proximity ligation assays. This would mean that Kv10.1 controls calcium entrance by regulating Orai1 and therefore improves the microtubule dynamics during cell division [13, 14].

Even if many mechanisms by which Kv10.1 promotes cell division are still to be explained, we are certain that blocking its conductive function impacts significantly on tumor growth, therefore, approaches towards drugs and strategies to block Kv10.1 in animal models are also a priority of our lab.

## **6. Therapeutic approaches**

In mice models where MDA-MB435S (melanoma) cells are implanted, cells form tumors that can be easily studied. When we compare the effect of Astemizole (a non-specific Kv10.1 blocker, which has antihistaminic properties) vs. Cyclophosphamide, a known chemotherapeutic, we observed that both drugs can diminish tumor size after 40 days of implantation [15]. Moreover, if we analyze the effects of a non-blocking Kv10.1 antibody compared to a blocking Kv10.1 antibody and to cyclophosphamide, we observe that the blocking Kv10.1 antibody again can reduce tumor growth at the same rate as cyclophosphamide in some models. If we implant patient-derived cancer cells in those mice, and we test for the antibodies against Kv10.1 we observed a less potent effect, compared to cyclophosphamide [16].

Nowadays chemotherapeutic treatment schemes use the combined effect of synergic drugs. One recent observation in Kv10.1 knock-down cells was the change in mitochondrial structure, generating a more fragmented pattern. Mitochondria are essential organelles for cancer cells due to the high metabolic rate they sustain. Mitochondrial fragmentation sensitizes cells for the additional use of antimetabolic drugs. We could demonstrate that blockage of Kv10.1 increases sensibility for antimetabolic drugs proportionally to their basal Kv10.1 expression, demonstrating that this effect is Kv10.1 expression-dependent [17].

Another approach currently under study by our group is the use of Kv10.1 attached to a more potent cytotoxic molecule such as TRAIL (TNF-related apoptosis-inducing ligand) which can induce apoptosis specifically in cancer cells [18]. We have now an improved design of such molecule using a single domain antibody (nanobody) against Kv10.1 bound to a single-chain TRAIL, which can induce apoptosis in the central region of the tumor in only 24 h at a dose of 3 ng/ml in tumor spheroids.

In conclusion, we believe that Kv10.1 represents one of the best oncological targets ever known, due to their selective expression in normal tissue. Therefore, we hope that in a near future the best anti-cancer strategy can be developed taking advantage of the Kv10.1 expression.

## Author details

Luis A. Pardo  
Oncophysiology Group, Max Planck Institute of Experimental Medicine,  
Göttingen, Germany

\*Address all correspondence to: [pardo@em.mpg.de](mailto:pardo@em.mpg.de)

## IntechOpen

---

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Whicher JR, MacKinnon R. Structure of the voltage-gated K<sup>+</sup> channel Eag1 reveals an alternative voltage sensing mechanism. *Science*. 2016;**353**(6300):664-669
- [2] Terlau H, Ludwig J, Steffan R, Pongs O, Stühmer W, Heinemann SH. Extracellular Mg<sup>2+</sup> regulates activation of rat eag potassium channel. *Pflügers Archiv*. 1996;**432**(2):301-312
- [3] Mortensen LS, Schmidt H, Farsi Z, Barrantes-Freer A, Rubio ME, Ufartes R, et al. Kv10. 1 opposes activity-dependent increase in Ca<sup>2+</sup> influx into the presynaptic terminal of the parallel fibre–Purkinje cell synapse. *The Journal of Physiology*. 2015;**593**(1):181-196
- [4] Hemmerlein B, Weseloh RM, de Queiroz FM, Knötgen H, Sánchez A, Rubio ME, et al. Overexpression of Eag1 potassium channels in clinical tumours. *Molecular Cancer*. 2006;**5**(1):1-13
- [5] Agarwal JR, Griesinger F, Stühmer W, Pardo LA. The potassium channel Ether a go-go is a novel prognostic factor with functional relevance in acute myeloid leukemia. *Molecular Cancer*. 2010;**9**(1):1-16
- [6] Ouadid-Ahidouch H, Ahidouch A, Pardo LA. Kv10. 1 K<sup>+</sup> channel: From physiology to cancer. *Pflügers Archiv-European. Journal of Physiology*. 2016;**468**(5):751-762
- [7] Martínez R, Stühmer W, Martin S, Schell J, Reichmann A, Rohde V, et al. Analysis of the expression of Kv10. 1 potassium channel in patients with brain metastases and glioblastoma multiforme: Impact on survival. *BMC Cancer*. 2015;**15**(1):1-9
- [8] Li X, Martinson AS, Layden MJ, Diatta FH, Sberna AP, Simmons DK, et al. Ether-à-go-go family voltage-gated K<sup>+</sup> channels evolved in an ancestral metazoan and functionally diversified in a cnidarian–bilaterian ancestor. *Journal of Experimental Biology*. 2015;**218**(4):526-536
- [9] Cone CD Jr, Tongier M Jr. Control of somatic cell mitosis by simulated changes in the transmembrane potential level. *Oncology*. 1971;**25**(2):168-182
- [10] Urrego D, Tomczak AP, Zahed F, Stühmer W, Pardo LA. Potassium channels in cell cycle and cell proliferation. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2014;**369**(1638):20130094
- [11] Bijlenga P, Occhiodoro T, Liu JH, Bader CR, Bernheim L, Fischer-Lougheed J. An ether-à-go-go K<sup>+</sup> current, Ih-eag, contributes to the hyperpolarization of human fusion-competent myoblasts. *Journal of Physiology*. 1998;**512**(2):317-323
- [12] Urrego D, Movsisyan N, Ufartes R, Pardo LA. Periodic expression of Kv10. 1 driven by pRb/E2F1 contributes to G2/M progression of cancer and non-transformed cells. *Cell Cycle*. 2016;**15**(6):799-811
- [13] Movsisyan N, Pardo LA. Kv10. 1 regulates microtubule dynamics during mitosis. *Cancers*. 2020;**12**(9):2409
- [14] Peretti M, Badaoui M, Girault A, Van Gulick L, Mabilille MP, Tebbakha R, et al. Original association of ion transporters mediates the ECM-induced breast cancer cell survival: Kv10. 1–Orai1–SPCA2 partnership. *Scientific Reports*. 2019;**9**(1):1-14
- [15] Downie BR, Sánchez A, Knötgen H, Contreras-Jurado C, Gymnopoulos M, Weber C, et al. Eag1 expression interferes with hypoxia homeostasis and induces angiogenesis in tumors. *Journal*

of Biological Chemistry.  
2008;**283**(52):36234-36240

[16] Gómez-Varela D, Zwick-Wallasch E, Knötgen H, Sánchez A, Hettmann T, Ossipov D, et al. Monoclonal antibody blockade of the human Eag1 potassium channel function exerts antitumor activity. *Cancer Research*. 2007;**67**(15): 7343-7349

[17] Hernández-Reséndiz I, Pacheu-Grau D, Sánchez A, Pardo LA. Inhibition of Kv10. 1 channels sensitizes mitochondria of cancer cells to antimetabolic agents. *Cancers*. 2020;**12**(4):920

[18] Hartung F, Pardo LA. Guiding TRAIL to cancer cells through Kv10. 1 potassium channel overcomes resistance to doxorubicin. *European Biophysics Journal*. 2016;**45**(7):709-719





# QoE and Immersive Media: A New Challenge

*Federica Battisti and Marco Carli*

## Abstract

New real-world capture and rendering systems are flooding the market. Mobile phones are now equipped with more than one camera, thus creating multi-view portable systems. Virtual reality rendering equipment is now within the reach of the consumer and many applications are available to the user. A big effort is being made by industrial and research bodies for spreading the new technologies. In this contribution, an overview of the main issues related to the quality evaluation of immersive media is presented.

**Keywords:** virtual reality, immersive media, quality of experience, multiple views, computer-generated data

## 1. Introduction

Recent years have witnessed an overwhelming rise in multimedia technologies. Their impact on the consumer is very high. The terms immersivity, virtual reality, augmented reality, and 3D content have now become familiar even to non-professionals. Under the boost of the entertainment sector and, more generally, of multimedia interaction, many novel services have been proposed. Immersive media can be defined as technologies that attempt to produce or imitate the physical world by exploiting computer-generated data. This status is achieved by techniques, both aural and visual, able to completely engage the user [1]. As stated by Dale Lovell in [2], “engagement is great, but immersion is the future. Immersion is when you forget the message entirely, forget you are the audience even, and instead fall into a newly manufactured reality”.

One of the first approaches in the direction of providing the user with the feeling of immersion was the Sensorama system in 1957. It is a mechanical device, which includes a stereo color display, fans to generate the sensation of the wind, odor emitters, a stereo sound system, and a chair mounted on a moving platform. The experience shown to users consisted of a motorcycle tour through the streets of New York. The user, sitting on the chair, was able to relive the riding experience through sounds, chair movements and pre-recorded images. The smell of the city (gasoline vapors and snack bar pizza) has been recreated by chemicals. According to the situation surrounding the user, different effects are rendered (i.e., when the rider approaches a bus, the typical bus noise and gasoline smell are sent to the user). However, the user interaction was quite limited.

Nowadays different devices are available for acquiring, processing, and rendering information in the best interactive way. They are the basic elements of immersive media, such as virtual reality, augmented reality, and mixed reality.

*Virtual Reality* replaces the user's physical environment (including surrounding sound) with a computer generated, interactive, 3D environment in which a person is immersed. One of the identifying marks of a virtual reality system is the use of head-mounted displays worn by users. These displays block out all the external world and present to the wearer a view that is under the complete control of the computer. This allows a scene to be seen in any direction from one viewpoint. When using a head-mounted display to watch such content, the viewing direction can be changed by head movements. A less immersive effect can also be obtained by rendering virtual reality content with different devices. On smartphones and tablets, the viewing direction can be changed by touch interaction or by moving the device around, thanks to built-in sensors. On a desktop computer, the mouse or keyboard can be used for interacting with omnidirectional video.

*Augmented Reality* combines the real-world with computer-generated data. Most of the AR research is currently concerned with the use of video imagery which is digitally processed and augmented by adding computer-generated graphics. The goal is to enhance instead of recreating the real scenario. A commonly used example of augmented reality is the Snapchat photo filtering tool.

*Mixed Reality* fuses the information collected by the real world with ad-hoc created digital ones. In this case the user may interact seamless with both. The user is generally equipped with a semi-transparent head-mounted display or with smart glasses. In mixed reality, the user must still be aware that he or she is present in the "real world." There are three components needed to make an augmented-reality system work: 1) the see-through rendering system, 2) the tracking system, and 3) mobile computing power. All these components are fundamental, and their performances highly affect the perceived Quality of Experience (QoE).

Many sectors will benefit from these technologies. In the following, a few examples are reported.

- Automotive industry: virtual reality technology allows to design of a vehicle or its constituent parts in a simple and inexpensive way before proceeding with the construction of expensive prototypes. At the same time, virtual reality and augmented reality may improve the maintenance services by showing how the situation should be and giving indications on the spot.
- Tourism: in this case, the tourist can have the feeling of the trip while not traveling. Immersive media can be used by travel agencies to go beyond the classical booklet of images by showing to the customer virtual guided tours around the world that can improve the final user satisfaction or the QoE.
- Healthcare: the use of these technologies is already in place for both training and patient care. Students and healthcare professionals can train in a low-risk 3D environment before working on real scenarios.

To achieve immersive goals, sophisticated media acquisition devices, new rendering systems, compression techniques, have been designed and, consequently, new application areas. Among others, 360° camera, light field camera, multiview camera setup, virtual reality equipment (audio and video), AR equipment, Tactile tools.

Especially when human subjects are involved, the impact of new technology on the perceived experience is a fundamental issue. If the human-in-the-loop factor is not properly addressed, the novel technology may not be successful. The negative trend of stereo content, especially in a home environment, is probably because the actual 3D content production, delivery, and presentation, are not compliant with 3D QoE. The success of the immersive imaging market relies on the ability of 3D

systems to provide added value compared to conventional monoscopic imaging (i.e., depth feeling or parallax motion) coupled with high-quality image contents. Dealing with these issues can result in the creation of perceivable impairments in the 3D content that may be originated in different points of the 3D chain, from content creation to display techniques. Many artifacts are common to the 2D imaging systems. However, novel distortions typical of the 3D structure should be considered (i.e., crosstalk or keystone) especially because their presence highly impacts the perceived quality (i.e., compression artifacts due to coding). Subjects are prone to prefer 2D contents to 3D ones, as soon as fatigue and discomfort are induced during the content presentation. The understanding of the quality of the experience is mandatory. However, this task is quite challenging. In the following, a brief overview of quality and the related issues is reported.

## 2. Quality of immersive media

The word quality is widely used in the most diverse fields. However, the agreement on the idea of quality is very hard and depends on several aspects: the application, the historical period, or even the background of each person. The concept of quality is something everybody understands but can hardly define.

Going back to ancient times, Aristotle classified every object of human apprehension into 10 categories: Substance, Quantity, Quality, Relation, Place, Time, Position, State, Action, Affection. Qualities are hylomorphically-formal attributes, such as “white” or “grammatical”. Always remaining in antiquity, Quality in ancient Egypt was “a sign of perfection.”

Nowadays, scientists have tried to better define this concept to be able to measure it. Among others, relevant ones are:

- General: Measure of excellence or state of being free from defects, deficiencies, and significant variations.
- ISO 8402-1986 standard defines quality as “the totality of features and characteristics of a product or service that bears its ability to satisfy stated or implied needs”.
- Google: the standard of something as measured against other things of a similar kind; the degree of excellence of something.
- Manufacturing: strict and consistent adherence to measurable and verifiable standards to achieve uniformity of output that satisfies specific customer or user requirements.
- ISO 9000: a family of standards for quality management systems.

To summarize, quality is a relative concept: it can rather be expressed as a degree of quality. We can agree with this statement: “the quality of something can be determined by comparing a set of inherent characteristics with a set of requirements. We will have high quality if characteristics meet requirements, and low quality, if characteristics do not meet all requirements. Nowadays, the research is devoted to the QoE evaluation, that is “The degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations for the utility and / or enjoyment of the application or service in the light of the user’s personality and current state” [3].

### **3. Measuring quality**

Quality evaluation of digital content is critical in all applications of information delivery. This is particularly true in case of digital image and video. Each stage of processing, storing, compression, and enhancement, may introduce perceivable distortions. For example, in image and video compression, the use of lossy schemes for reducing the amount of data may introduce artifacts as blurring and ringing, which leads to quality degradation. Similarly, during the transmission phase, due to the limited bandwidth available and to the channel noise, data might be lost or be modified, thus resulting in quality degradation of the received content.

The visibility and annoyance of these impairments are directly related to the quality of the received/processed data. The possibility of measuring the overall perceived quality to maintain, control, or enhance the quality of the digital data is fundamental. During the last two decades, many efforts have been directed by the scientific community to the design of quality metrics. The choice of an adequate metric usually depends on the requirements of the considered application.

There are two main methods of assessing media quality: subjective or objective. The first is carried out by human observers, while the second consists of the definition of models for predicting subjective evaluation.

#### **3.1 Objective metrics**

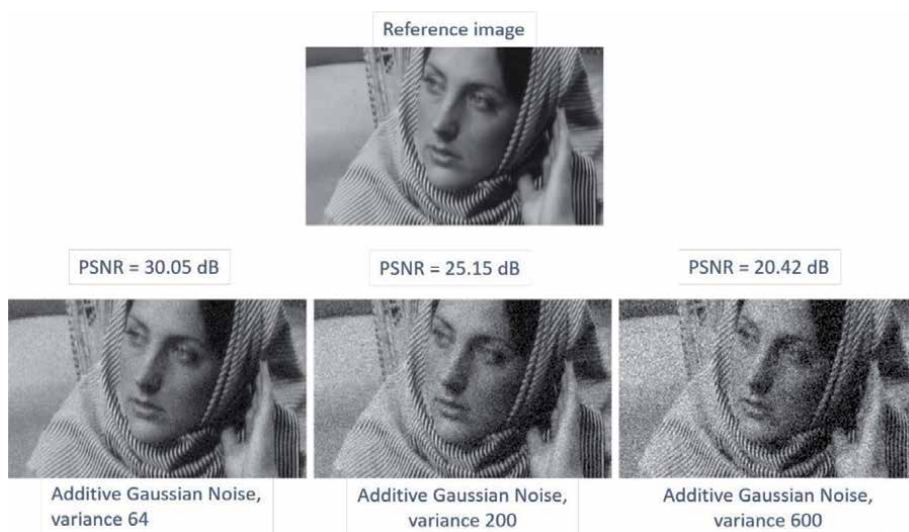
In objective measurements of the performances of an imaging system, image quality and quality losses are determined by evaluating some parameters based on a given general mathematical, physical or psycho-psychological model. That is, the goal is to obtain a measurable and verifiable aspect of a thing or phenomenon, expressed in numbers or quantities, such as lightness or heaviness, thickness or thinness, softness or hardness.

Objective quality metrics can be classified according to the amount of side information required to compute a given quality measurement. Using this criterion, three generic classes of objective metrics can be classified as Full Reference (FR) when the original and the impaired data are available, Reduced Reference (RR) when some side information regarding the original media can be used, and No-Reference (NR) if only the impaired image is available.

To make an objective assessment, one can use measuring devices to obtain numerical values; another method is to use image or video quality metrics. These metrics are usually developed to consider the human visual system and try to better match the subjective assessment.

To the first class belong the FR quality metrics. Among the most widely adopted FR objective metrics are the Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR). Both are pixel-wise measures of the difference between the original and of the impaired media. In particular, the PSNR is a measure of the peak error between the compressed image and the original image. PSNR is given as  $PSNR = 20 \log_{10} \text{MAX}(I) / \sqrt{\text{MSE}}$ , where  $\text{MAX}(I)$  represents the maximum possible value of the media. The higher the PSNR, the better the quality of the reproduction. PSNR has usually been used to measure the quality of a compressed or distorted image. It is also applied, frame by frame, to video as the first information about video degradation. Other metrics are SSIM [4], MS-SSIM [5], VIF [6], MAD [7], FSIM [8], etc.

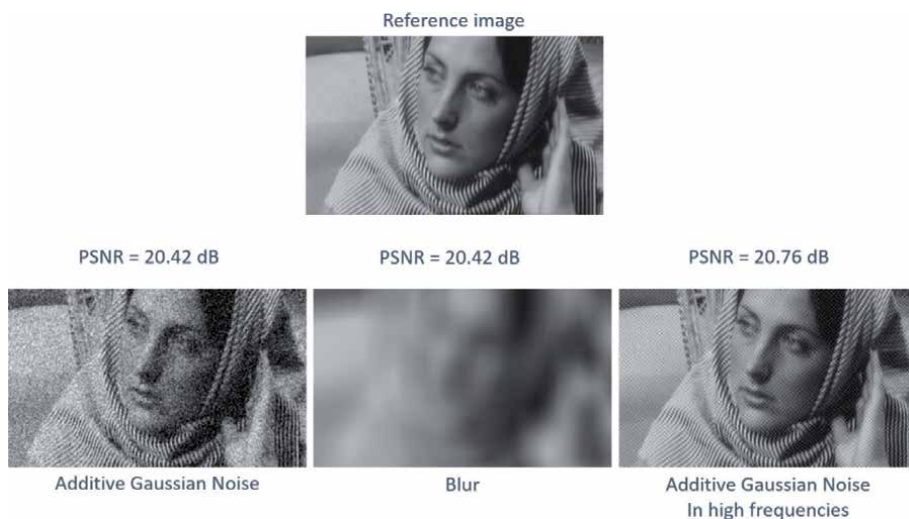
Objective metrics have low computational cost, physical meanings, and are mathematically easy to deal with for optimization purposes. However, they have been widely criticized for not being well correlated with the perceived quality measurement.



**Figure 1.**  
*Additive Gaussian noise of increasing variance.*

**Figure 1** shows an original image and its version deteriorated by additive Gaussian noise with increasing intensity. **Figure 2** shows the same original image and three versions of the same image in which different distortions are introduced. As can be noticed, in the first case the value of the objective metric agrees with the perceptual judgment. In the second case, the objective metric returns the same score, thus indicating an equal level of distortion. However, from a perceptual point of view, the images are perceived as of different quality.

To overcome such problems, HVS inspired objective quality metrics have been introduced e.g., PSNR-HVS and PSNR-HVS-M. The main difference between these metrics and the mathematical ones (MSE, PSNR) is that they are more heuristic. It is more difficult to perform a mathematical comparison of their performances. Thus, to adequately evaluate the quality of such metrics statistical experiments are needed [9, 10].



**Figure 2.**  
*Different distortions on the same image. The objective score is 20.42 dB.*

### 3.2 Subjective metrics

In subjective tests, the digital content quality is assessed by performing subjective psychological tests. In this case, the goal is to find attributes, characteristics, or properties that can be observed and interpreted, and maybe approximated (quantified) but cannot be measured, such as beauty, feel, flavor, or taste. The quality score is generated by averaging the result of a set of standards, subjective tests and it can be considered as an indicator of the perceived media quality. A pool of subjects evaluates a set of images (or videos) ranking the perceived quality according to a specific scale [11]. **Table 1** is reported the most used ranking scale in which the score 1 should be given to the media perceived ‘bad’ since it is affected by a ‘very annoying’ artifact. Similarly, the score of 5 should be given to the media showing excellent quality, in which no impairments are perceivable.

Contrary to what it may seem, the subjective evaluation methodology is complex and time-consuming since, to be reliable, it requires to be properly designed and a large number of subjects is needed.

In more details, the subjective test depends on the test environment (i.e., type of monitors/speakers and other test equipment, lighting/acoustic conditions, laboratory architecture, background, ...), the test material (i.e., meaningful content for the envisaged scenario/application, best, typical, worst cases, ...), the test methodology (i.e., viewing distance/hearing position, subject selection, instruction phase, opinion or judgment collection, training - presentation – grading scale), and the carried-out analysis of the data.

### 3.3 Test material

To verify the performances of an objective metric, as well as for collecting the subjective score, a large database of distorted test images is usually prepared, and the Mean Opinion Score (MOS) from many human observers is collected. Then, the subjective results are compared with the objective scores of the tested metrics to identify the metric which metric shows the highest correlation with the subjective scores. However, some drawbacks have to be considered: usually, the size of the test database is not big enough, the number of different distortions is limited [12, 13], and methodological errors in planning and execution of the experiments can occur. Since in most applications humans are the ultimate receivers of digital data, the most accurate way to determine its quality is to measure it directly using psychophysical experiments with human subjects. One of the most intensive studies in this field has been carried out by the Video Quality Expert Group (VQEG). In the image quality framework, many datasets have been created as LIVE [4, 14] or TID2013 [15]. Relevant efforts have also been devoted to the design and test of video quality datasets. In this direction, among others, LIVE Video Quality Assessment Database [16] and the EPFL-Polimi [17] video databases have been extensively adopted.

MOS	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

**Table 1.**  
*Mean opinion score assessment table.*

### 3.4 Concluding remarks

As can be deduced from the brief and non-exhaustive analysis made in the previous paragraphs, the goal of obtaining a general-purpose objective metric is far from being achieved.

There are many difficulties, such as the availability of well-designed test datasets or the need for extensive subjective tests for collecting subjective opinion. In the frameworks of virtual, augmented, and mixed reality, quality evaluation is even more complex. In fact, up to now no standardized guidelines for running subjective tests have been defined. Even the use of ACR (Absolute Category Rating), ACR-HR (Absolute Category Rating with Hidden Reference), DSIS (Double-Stimulus Impairment Scale), DSCQS (Double Stimulus Continuous Quality-Scale) for quality assessment is not well defined, since it mainly depends on the target application and the rendering device. The situation is not different for the virtual reality case.

In the next paragraphs, we will make a quick introduction of one of the systems currently among the most used to acquire a scene from multiple points of view.

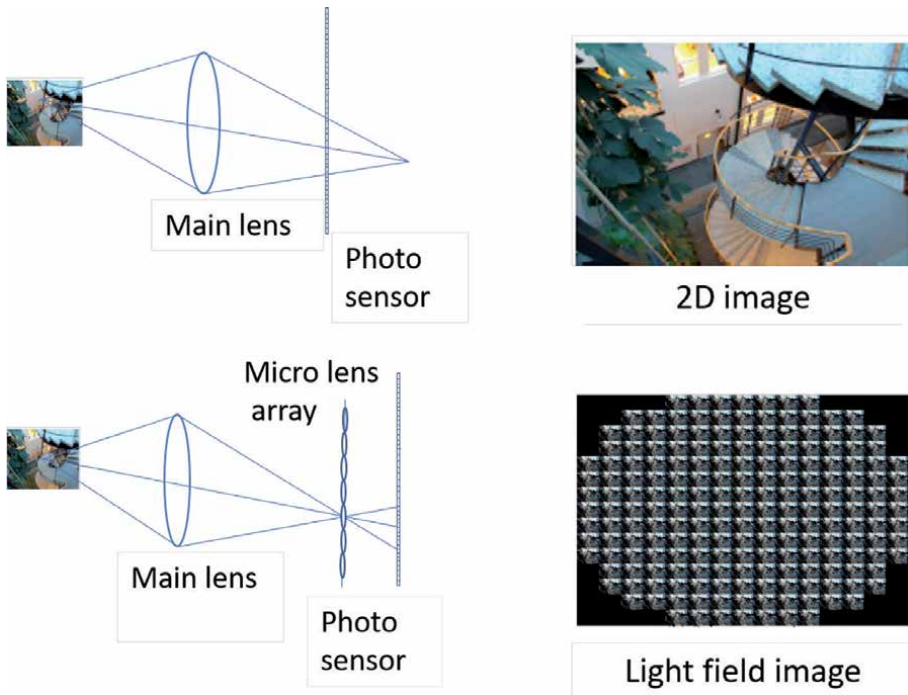
## 4. Light field

The Light Field (LF) expresses the radiance as a function of position and direction in regions of free space [18, 19]. In other words, it represents the number of light rays within a specific area. The capturing of all light rays in a scene allows generating a perspective view from any position. Therefore, LF technology can be effectively used in many applications: from accurate passive depth estimation to change of viewpoint or view synthesis, which can be useful in augmented reality content capture or movie postproduction.

The capturing of an LF is a quite complex procedure from the technological point of view; in fact, the light field represents rays with varying positions and angles, and, to obtain this information, it is necessary to record the scene from multiple positions.

To this aim, different techniques can be adopted: the use of camera arrays, camera gantry, or plenoptic cameras. By spatially locating multiple cameras into an array, the entire LF may be collected at once. This approach is used up to a planar array of 128 cameras. A different system is based on moving a single camera while capturing a stationary scene to measure the incident light rays. The basic idea behind the plenoptic imaging systems is the use of a micro-lens array positioned on the focal point of the camera lens, in front of the imaging sensor as shown in **Figure 3**.

This system allows recording multiple views of a scene in a single shot, thus reducing issues related to calibration and camera synchronization. The micro-lens array records the information on the incident light direction at different positions, i.e., it records the LF. The availability of low-cost acquisition devices allows novel applications for these imaging systems. The exploitation of the LF redundancy in the post-processing and editing phases brings photographers and art directors new opportunities. One of the main issues of this technology is related to the rendering modality. Many efforts are being devoted to the design of dedicated displays (e.g., an array of video projectors aimed at a lenticular sheet, 3D Displays, up to recently proposed tensor displays) or devices (e.g., head-mounted systems for virtual reality applications). However, up to now, these systems are very expensive and there are many challenges to be addressed (e.g., the reduced angular resolution of an LF cinema). The simplest and cheapest solution is the rendering of the LF data on conventional 2D screens. Since the LF allows rendering the scene from several



**Figure 3.**  
*Light field vs. 2D imaging system.*

points of view and focus points, the questions of what and how to render the scene on a 2D display arise. To solve this issue, recent works have been devoted to an in-deep analysis of the impact of different visualization techniques of LF images on a 2D display [20].

Datasets	Year	SRCs (Acquisition)	Artifacts (HRCs)	Protocol	Rendering Visualization	Stimuli
SMART	2017	16 (Lytro Illum)	Coding: SSDC, HEVC-Intra, JPEG, JPEG2K	PC	EDoF images (all-in-focused view)	256
MPI-LFA	2017	14 (synthetic and captured)	3D-HEVC, Linear Nearest Interpolation	ACR	stereoscopic viewing	336
VALID dataset	2018	5 (Lytro Illum)	Compression HEVC and VP9	DSIS	2D displays	140
Win5-LIDt	2018	10 (Lytro Illum and Synthesis)	HEVC, JPEG2000, Linear Nearest Interpolation	extended DSCQS	stereo display	200
LF Dataset	2019	8 (Lytro Illum)	Gaussian blur, JPEG2000, JPEG, motion blur, white noise	DSCQS	pseudo-sequence	240
LFDD	2020	8 (synthetic)	Image-based compression, Video-codecs, Geometric distortion, Noise	DSIS	pseudo-sequence	480

**Table 2.**  
*Annotated light field dataset.*



The research community is also trying to define quality metrics and test datasets specifically designed for LF data. In **Table 2**, a list of available LF datasets, annotated with the corresponding subjective scores, is reported.

In the direction of the definition and assessment of quality, some efforts have been made. Methodologies for performing subjective quality assessment experiments were investigated in [20] and the impact of compression systems in [21]. The study was conducted by designing the SMART LF image quality dataset consisting of source images, compressed images, and subjective scores. The impact of the compression, reconstruction, and visualization phases was studied in [22] together with the definition of the Dense Light Fields dataset. The applicability and perceptual impact of existing and specifically designed compression techniques have been studied in [23]. A tentative to assess the subjective quality of experience of decoded LF images was performed in [24]. A reduced reference LF image quality metric based on the relationship between the distortion of the estimated depth map and the LF image quality was presented in [25]. Full reference metrics based on multi-order derived characteristics (MDFM) [26] and EPI [27] were presented. More recently, the log-Gabor feature-based light field coherence (LGF-LFC) feature has been proposed for a full reference metric in [3].

## 5. Conclusions

Defining objective quality metrics for immersive media is a very challenging task.

It needs a good understanding of both acquisition and rendering devices and subjective perception.

It depends on several parameters, which are difficult to be identified. The knowledge acquired in these decades from the use of images and videos has led to the definition of objective metrics, methodologies, and sufficiently defined test materials.

In the case of new media, the direct transfer of this knowledge is not possible. It is necessary to understand the possibilities of the application of new media and their limitations. The definition of use cases and the identification of significant parameters is needed. In addition, there is a need for databases annotated with subjective data, such as Mean Opinion Score, eye-tracking information, content definition. Open research questions are related to the understanding of the impact of the content on the quality of experience, to the definition of specific assessment protocols, and the definition of effective quality metrics. It is worthwhile to underline that what is needed in the evaluation of the Quality of the Experience, rather than the 'simple' quality of the media. Therefore, the human factor must be included in all the phases of the design of the immersive system.

## **Author details**

Federica Battisti<sup>1\*</sup> and Marco Carli<sup>2\*</sup>

1 Department of Information Engineering, University of Padova, Padova, Italy

2 Department of Engineering, Roma Tre University, Rome, Italy

\*Address all correspondence to: federica.battisti@unipd.it  
and marco.carli@uniroma3.it

## **IntechOpen**

---

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Domański M, Stankiewicz O, Wegner K, Grajek T. Immersive visual media — MPEG-I: 360 video, virtual navigation and beyond. In: 2017 International Conference on Systems, Signals and Image Processing (IWSSIP). USA: IEEE; 2017. pp. 1-9. DOI: 10.1109/IWSSIP.2017.7965623
- [2] Lovell D. Immersive Media, Virtual Reality and a New Form of Media. UK: Internet Advertising Bureau; 2016. Available from: <https://www.iabuk.com/opinions/immersive-media-virtual-reality-and-new-form-media>
- [3] Tian Y, Zeng H, Hou J, Chen J, Ma K-K. Light field image quality assessment via the light field coherence. *IEEE Transactions on Image Processing*. 2020;29:7945-7956
- [4] Wang Z, Bovik AC, Simoncelli EP. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*. 2004;13:600-612
- [5] Wang Z, Simoncelli EP, Bovik AC. Multi-scale structural similarity for image quality assessment. In: *IEEE Asilomar Conf. on Signals, Systems, and Computers*. USA: IEEE; 2003
- [6] Sheikh HR, Bovik AC. Image information and visual quality. *IEEE Transactions on Image Processing*. 2006;15:430-444
- [7] Larson EC, Chandler DM. Most apparent distortion: full-reference image quality assessment and the role of strategy. *Journal of Electronic Imaging*. 2010;19:011006:1-011006:21
- [8] Zhang L, Zhang L, Mou X, Zhang D. FSIM: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*. 2011;20:2378-2386
- [9] Egiazarian K, Astola J, Ponomarenko N, Lukin V, Battisti F, Carli M. Two new full-reference quality metrics based on HVS. In: *Proceedings of the Second International Workshop on Video Processing and Quality Metrics for Consumer Electronics*. Arizona, USA; 22-24 January 2006. 4 p. Available from: <http://sp.cs.tut.fi/pubdl/Egiazarian2006-Two.pdf>
- [10] Ponomarenko N, Silvestri F, Egiazarian K, Carli M, Astola J, Lukin V. On between-coefficient contrast masking of DCT basis functions. In: *Proceedings of the Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics*. VPQM 2007, Scottsdale, Arizona, USA; 25-26 January 2007. 4 p. Available from: <http://sp.cs.tut.fi/pubdl/Ponomarenko2007-On.pdf>
- [11] ITU-R. Methodology for the subjective assessment of the quality of television pictures. ITU-R Rec. BT. 2012;1:500-513
- [12] Ponomarenko N, Lukin V, Egiazarian K, Astola J, Carli M, Battisti F. Color image database for evaluation of image quality metrics. In: Feng D, editor. *Proceedings of the 2008 IEEE 10th Workshop on Multimedia Signal Processing*. MMSP 2008, Cairns, Australia; 8-10 October 2008. pp. 403-408. DOI: 10.1109/MMSP.2008.4665112
- [13] Ponomarenko N, Battisti F, Egiazarian K, Astola J, Lukin V. Metrics performance comparison for color image database. In: *Proceedings of the Fourth International Workshop on Video Processing and Quality Metrics Consumer Electronics*. VPQM 2009, Scottsdale, Arizona, USA; 14-16 January 2009. 6 p
- [14] Sheikh HR, Sabir MF, Bovik AC. A statistical evaluation of recent full reference image quality assessment

- algorithms. *IEEE Transactions on Image Processing*. 2006;**15**(11):3440-3451
- [15] Ponomarenko N, Jin L, Ieremeiev O, Lukin V, Egiazarian K, Astola J, et al. Image database TID2013: Peculiarities, results and perspectives. *Signal Processing: Image Communication*. 2015;**30**:57-77. DOI: 10.1016/j.image.2014.10.009
- [16] LIVE Video Quality Assessment Database. Austin: The University of Texas; 2009. Available from: [http://live.ece.utexas.edu/research/quality/live\\_video.html](http://live.ece.utexas.edu/research/quality/live_video.html)
- [17] De Simone F, Tagliasacchi M, Naccari M, Tubaro S, Ebrahimi T. H.264/AVC video database for the evaluation of quality metrics. In: *IEEE International Conference on Acoustics, Speech, and Signal Processing*. USA: IEEE; 14-19 March 2010
- [18] Lippmann G. Sur la théorie de la photographie des couleurs simples et composées par la méthode interférentielle. *Journal of Physics: Theories and Applications*. 1984;**3**(1): 97-107
- [19] Gabriel L. Epreuves reversibles photographiques intégrales. *Comptes-Rendus Academie des Sciences*. 1908;**146**(8):446-451
- [20] Paudyal P, Battisti F, Carli M. Effect of visualization techniques on subjective quality of light field images. In: *Proc. IEEE Int. Conf. Image Process*. USA: IEEE; 2017. pp. 196-200
- [21] Paudyal P, Battisti F, Sjostrom M, Olsson R, Carli M. Towards the perceptual quality evaluation of compressed light field images. *IEEE Transactions on Broadcasting*. 2017;**63**(3):507-522
- [22] Adhikarla VK et al. Towards a quality metric for dense light fields. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA; 2017. pp. 3720-3729. DOI: 10.1109/CVPR.2017.396
- [23] Viola I, Rerabek M, Ebrahimi T. Comparison and evaluation of light field image coding approaches. *IEEE Journal of Selected Topics in Signal Processing*. 2017;**11**(7):1092-1106
- [24] Perra C. Assessing the quality of experience in viewing rendered decompressed light fields. *Multimedia Tools and Applications*. 2018;**77**(16): 21771-21790
- [25] Paudyal P, Battisti F, Carli M. Reduced reference quality assessment of light field images. *IEEE Transactions on Broadcasting*. 2019;**65**(1):152-165
- [26] Tian Y, Zeng H, Xing L, Chen J, Zhu J, Ma K-K. A multi-order derivative feature-based quality assessment model for light field image. *Journal of Visual Communication and Image Representation*. 2018;**57**:212-217
- [27] Fang Y, Wei K, Hou J, Wen W, Imamoglu N. Light field image quality assessment by local and global features of epipolar plane image. In: *Proc. IEEE 4th Int. Conf. Multimedia Big Data*. USA: IEEE; 2018. pp. 1-6

# Strong Objectivity for New Social Movements

*Sandra Harding*

## Abstract

Standpoint methodology and its strong objectivity standard emerged four decades ago in the context of social justice movements of the 1960s and 1970s. Movements for poor people, African Americans, women, LGBTQ, and the disabled differed in many ways. Yet, all were firmly anti-authoritarian, criticizing the top-down policies and practices of governments and international agencies, as well as the natural and social sciences that served the interests of such institutions. The social justice movements argued that dominated groups would continue to be oppressed by research methodology, epistemology, theory, and public policy that ignored how the conditions of their marginalized lives differed from the living conditions of elite white men. They all insisted that the questions arising from their daily lives provided more effective starting points for maximally objective research results and the democratic public policies that such research was supposed to direct. This essay will focus on how, several decades later, newer social justice movements are demanding additional attention to the research practices that have bad effects on the public policy that shapes the everyday lives of peoples in such groups and elites. How do standpoint research strategies and their strong objectivity standards fare in these new social justice movements?

**Keywords:** new social movements, social justice, oppressed groups, public policy, standpoint research strategies

## 1. Introduction

It is now more than 60 years since C.P. Snow's [1] *Two Cultures* pointed out that scholars in the humanities and those in the sciences lived in two different worlds. They rarely encountered each other in scholarly contexts and were mostly entirely ignorant about each other's projects. What interested Snow was the scientifically illiterate humanists.

New groups have joined the ranks of the scientifically illiterate, in the eyes of their critics: namely scientists themselves and the educated classes, as well as the policymakers who depend on scientific findings. These newest groups accuse the scientists of ignorance about androcentrism, racism, coloniality, and Eurocentrism that damages the reliability of their results of research. Science is a fully social process, they argue. What we know and do not know is shaped not only by "nature herself," but also by what the most powerful corporations and governments want to know. They point, for example to widespread ignorance about climate change, which has held back effective public policy in this area. And an increasing number of such critics point to the effective non-modern knowledge systems of

non-Western cultures, which have served those cultures well in their distinctive social and natural environments.

Obviously, nobody wants biased research that produces inaccurate accounts of nature and social relations. We want reliable accounts on which to base public policies and our practices. However, this can seem to be a dangerous moment even to take up this question in light of the constant barrage of false claims and “science-bashing” that issues daily, as I write, from the outgoing U.S. president and other authoritarian regimes around the globe.

Yet today, as the front pages of our newspapers have revealed, who catches COVID-19 and who dies from it indicate that in important respects, maximally objective environmental and medical/health assumptions and practices have not been guiding public policy. COVID-19 is an equal-opportunity virus, but the conditions of life for poor people and peoples of color ensure that they are more likely to catch it and have fewer resources to deal with it. Moreover, in the related economic crisis, who falls into poverty and who does not reveal similar faulty assumptions that shape economic policy.

In response to the earlier complaints, the sciences have corrected their processes in significant ways. As most physicians now recognize, the bodies of members of these other groups are not in all respects exactly like the stereotypical model of the human as the idealized elite white man. Women’s bodies are not immature or defective versions of men’s bodies, with simply a different reproductive system characterizing them, as the old, pre-1970s accounts claimed.

Engineers also got the message. Automobile designers created the possibility of adjusting the height and position of drivers’ seats so that even small drivers, such as many white women and most people in some other ethnic groups, could see out of the front window and at the same time reach the gas or brake pedals. Yet this morning, a story on NPR gave voice to women farmers, who were complaining that tractors and other farming equipment are not user-friendly to anyone but big, very strong men. Manufacturers need to design such items for use by the full array of peoples who farm, including white women as well as men and women in ethnic groups that characteristically are shorter and less heavily muscled.

## **2. Calls for greater objectivity**

Thus, accommodation to servicing the needs of physically and socially diverse groups has produced economic, political, and educational revisions of our policy worlds and our daily experiences in them. These groups want to research that is more objective than the conventional supposedly universally valid research that was grounded only in dominant groups’ experiences. They do not want “subjective” research, as their critics often claim. It is the dominant models of the human and their standards that have been only subjective, the critics counter, representing only elite groups’ experiences and interests. Rather, they want “stronger objectivity” that can more accurately chart all of our naturally and socially different lives in the worlds that we share.

Some sciences are more liable to such charges than others. High-energy physics certainly seems reasonably resistant to such charges. It does not seem to be at all about people as social beings. Yet one can still ask questions about why it is that these sciences’ projects are so highly funded by the U.S. Department of Defense. Could this have something to do with U.S. military politics rather than only with the objective desire to understand “pure nature”? Why do not sciences that could effectively prepare for a pandemic—one that on last Friday alone newly infected 99 thousand U.S. citizens and killed 1000—receive equal federal funding? It is becoming clear that today we live in a historically extraordinary moment in which deeply

anti-democratic infrastructures have become increasingly visible. Such infrastructures ensure that scientific research will not be maximally objective; it will continue to serve the desires of the powerful at the expense of the needs of the vast majority of the world's citizens. Our standards for objective research that were produced as a result of the earlier social justice movements did not go far enough.

### **3. The invention of standpoint methodology and its strong objectivity standard**

Standpoint methodology was the name given to the research methodology intended to address such problems. It calls for "strong objectivity," that can provide a more reliable standard for universally valid research. Though it emerged from all of the social justice movements of the 1970s, it was not so named at the time. Each of those movements proposed that reliable research to guide policy about their lives should start off its projects in a different way. It should not be addressing the standard issues that were the focus of mainstream natural and social sciences, but instead, start off from questions arising from the everyday lives of members of groups that experienced oppression and discrimination. Health, environmental, and social science research must take the "standpoints" of the everyday lives of marginalized groups to produce maximally reliable results of research. Through the efforts of marginalized members of the sciences, as well as of many non-marginalized scientists who immediately recognized the importance of the issue, this practice rather quickly became the strong objectivity standard for good research across most of the social sciences as well as health and environmental sciences that are a mix of natural and social science projects. Of course, there persist today continuing cases of both ignorance of and resistance to such practices.

Feminists were the first of these groups to call it standpoint theory. This began with a half dozen such political scientists, sociologists, and philosophers. Interestingly, they were almost entirely working independently of each other in the U.S., Canada, and the U.K. These included the sociologist of science Hilary Rose [2] in the U.K., sociologist of knowledge Dorothy Smith [3] in Canada, political scientists Nancy Hartsock [4] and myself, a philosopher of science, in the U.S. We all began asking such questions in the 1970s. Soon, sociologist Patricia Hill Collins [5] and many more African Americans and other feminists of color also began to refer to it as standpoint methodology, epistemology, and theory.

### **4. The beginning of the end of Western modernity?**

Now newer social justice movements are raising additional issues. The sciences today are beginning to realize that if they want to understand how COVID-19, the associated economic crisis, and climate change actually work, they have to start off their research from the daily lives of the peoples least advantaged by such phenomena. Everyone is affected by what happens to everyone else in our shared world, but we are affected in different ways depending on the circumstances of our daily lives.

As Sheila Jasanoff [6] argued, sciences and their societies co-create and co-constitute each other. Early modern science was co-created and co-constituted with the new economic, political, social, and technical forms of life emerging in early modern Europe [7]. These sciences bore the imprint of the still existing residues of medieval European societies and were directed by the desires of the new social classes coming into power at that time. Today we may well be experiencing the beginnings of a similarly big shift in economic, political, social, and technical forms

of life as electronic advances now permit both good and bad news to travel rapidly around the globe, and apparently beyond the kinds of federal controls permissible in democratic societies, and as our existing institutions appear unable to act effectively for the linked phenomena of the pandemic, the economic collapse, and climate change. The gap between the rich and the poor has rapidly escalated over the last four years, but it was well underway before this disastrous period in U.S. and international life. Traditional Liberal governments seem unable to organize the resources necessary to block the anti-democratic effects of such processes. Are we experiencing the beginning of the end of Western modernity its Liberal form of democracy and its philosophy of science?

Standpoint methodologies were developed for the political projects of the 1960s and 70s social justice movements in the global North, as noted earlier. Can they be adapted to these changing circumstances of peoples' everyday lives, as these are represented in the new global South social justice movements?

The Latin American theorists of recovering ancestral knowledges provide one of the major critical forces developed in the global South that are calling for new scientific epistemologies and ontologies. They claim to offer radically different accounts of how nature and social relations work in our everyday lives and consequently point toward the need for new political resources to advance pro-democratic outcomes. And they insist that this recovery project is necessarily entangled with gender issues. What is the relation between these projects and those directed by standpoint methodologies?

## **5. Recovering ancestral knowledges: Latin America**

In Latin America, social studies of knowledge production have been constructed in opposition to its distinctive history of primarily Spanish and Portuguese colonialism.<sup>1</sup> The very modernity that was co-constituted with European sciences is itself both a product of and a contributor to colonialism. Yet the Latin American opposition to its colonial history is articulated also as an opposition to the postcolonial theory with which the North has reevaluated its mostly British colonial history with Asia (e.g., [12]). A significant group of these theorists has named themselves the modernity/colonial/decolonial group, or Decolonial for short (e.g., [13–15]).

Decolonial analyses occur in significantly different historical contexts than those in which the more familiar postcolonial accounts were generated. First, there are important chronological differences marked especially by the MCD scholars. Colonial relations in the Americas began in 1492—more than two and a half centuries before the British began to establish their colonies in India and the Middle East. For the Decolonial scholars, it is no accident that the so-called discovery of the Americas coincides with the emergence of modernity in Europe, though standard Northern histories tend not to link these two phenomena. “Modernity appears when Europe organizes the initial world-system and places itself at the center of world history over against a periphery equally constitutive of modernity” ([16], pp. 9–10). So, for Latin American theorists, modernity and Iberian colonialism co-produce and co-constitute each other. This not only shifts the beginning of modernity to a much earlier date but also inserts Iberian colonialism centrally into the history of modernity, which is something that has been largely denied by North Atlantic scholars.

Another chronological difference is that formal independence from European rule began much earlier in the Spanish, Portuguese, and French colonies in the

---

<sup>1</sup> Earlier versions of Section 5 have appeared in [8–11].



Americas than in the British colonies (except for the United States). Most of the other colonies in the Americas achieved formal independence from Spain, Portugal, and France by 1830, except Cuba, which gained independence in 1898.<sup>2</sup> Moreover, for the anti-colonial scholars, 1492 is the starting date of anti-colonial thinking. The Amerindians whom Cortes encountered, as well as Nahua and Quechua intellectuals in the early sixteenth century, clearly resisted both the idea and the reality of Iberian colonization [17, 18]. Anti-colonial thought has a longer and different history in Latin America than the familiar British postcolonial accounts.

Second, the origins of the Scientific Revolution are broader than assumed in conventional philosophies and histories of science, and they have roots in colonialism. Colonization of the Americas required that the conquerors interact effectively with physical worlds different from those familiar to them. Yet they lacked astronomy of the Southern hemisphere with which to navigate back to Europe across the South Atlantic. The cartography of the South Atlantic and their environments in the Americas had to be created. They also needed climatology, oceanography, and better engineering to secure the safe travels of their crews and their precious cargoes. In the Americas, they needed knowledge of the unfamiliar geographies and flora and fauna that they encountered. They needed better geology, mining, and engineering, even though they soon appropriated from the Amerindians sophisticated forms of these technologies which they improved to extract the gold and silver that they found in Mexico and Peru. In 1492, the Europeans were behind the Amerindians in these kinds of scientific and technical knowledge: they were the backward ones. Europe's colonial projects in the Americas turned a huge part of the globe into a laboratory for European sciences [19, 20].

Third, in addition to the scientific and technical needs created by the different chronologies and geographies, the Iberian colonizers lived in social worlds different from those that shaped the coloniality of the British Empire. For the Europeans, the "discovery" of new lands across the Atlantic appeared as a solution to some of their most vexing social problems. Europeans welcomed the thought of being able to leave behind the economic and political challenges of the continual religious and political wars, as well as of overpopulation and famines. The Europeans imagined that they could start over in the "Garden of Eden" that had been "discovered" across the Atlantic.

Fourth, yet those peoples that the Spanish and Portuguese colonized were culturally different from those the British colonized centuries later. For the Amerindians, the arrival of the Europeans was a cataclysmic event. It meant the destruction of their cultural and physical worlds, the loss of sovereignty over their lands, the loss of their freedom, and the destruction and devaluation of their forms of knowledge and spirituality.

It is only relatively recently that demographic, historical, and environmental research undermined long-held assumptions that the Americas were only sparsely inhabited in 1491, and that those inhabitants were at a much more primitive stage of social and scientific development than were Europeans. In 1491, there were probably more people living in the Americas than in Europe (e.g., [21, 22]). Estimations of the actual numbers in the Americas vary hugely, from 10 million to over 100 million. Some of the world's largest cities at the time were in the Americas [22]. Inca, Aztec, and Mayan architecture, engineering, and road systems were among the most advanced of ancient civilizations, and in some respects superior to those of the Europeans. Amerindians had extensive agricultural techniques, such as

---

<sup>2</sup> However, some would argue that Puerto Rico remains a colony of the United States: residents of Puerto Rico are residents of the United States, but they are not citizens of the United States. They cannot vote in national elections.

controlled fires to clear the land and increase the nutrients in the soil, and were able to preserve food that could last for years through processes of freezing, dehydration, and rehydration.

What did the Amerindians know in 1491 in addition to their agricultural, environmental, and spiritual-philosophical knowledge? The Nahua effectively mined silver and gold, as indicated, and drained the swamps, and then engineered the hanging gardens of the town that became Mexico City. Moreover, the Europeans had no way to project dates into BC eras, and no precise way to measure a solar year. The Nahua, Mayans, and Incas could do both. Amerindians also learned that they could locate their calendars on the European Christian calendars; Aztec and Inca events could be celebrated to coincide with Christian events, unbeknownst to the Europeans. And there was more knowledge production in the realm of medicine, pharmacopeia, and botany. Today indigenous knowledges are being reconstructed and are experiencing a boom perhaps never seen since the conquest.

Indigenous philosophies appeared dormant or invisible to the non-indigenous outsider until very recently and are still largely unknown to Anglo academics. Yet they have existed underground and persisted throughout the centuries inside indigenous communities. Today, indigenous peoples see as their task not only to reconstruct ancestral knowledges for their own survival but also for the survival of the rest of humanity that seems unable to halt the most pernicious aspects of modernity such as infinite economic growth, the destruction of the planet or “Pachamama”—the earth mother, and modern science and technology at the service of profit and constant wars [23]. What is especially remarkable about this process is that for the first time in colonial history, indigenous women’s voices can now be heard. Indigenous women had in the pre-intrusion era occupied important social and political positions that were undermined with colonization. Equally important was the place women occupied in indigenous cosmogonies and ontologies. These positioned women in a parallel, but not always equal position with men. It is this last point that not only defines the particularity of indigenous epistemologies, cosmogonies, and ontologies but also gives rise to one of the most contentious points in today’s feminist debates around gender.

The recuperation of ancestral knowledges is necessarily a contested terrain. The difficulty of recovering them lies not only in their fragmented and dispersed state after centuries of colonization; they also have fused with Western, Christian elements which have altered not only the collective memory but also their existence in the present. It is not always clear what remains of the past and what is a recent invention. To complicate matters, the process of recuperation is often manipulated by present-day political interests of both indigenous and mestizo men, but also of women.

Yet no matter how important it is to keep in mind these contradictions in the process of recuperation of ancestral knowledges, such knowledges do pose serious challenges to Western totalitarian knowledge that sees itself as the only valid knowledge. It is perhaps in the discussions about gender where the disparities seem to be the greatest. Gender permeates the entire recomposition of indigenous cosmovisions.

## **6. Indigenous conceptions of gender/sexuality**

Indigenous conceptions of gender in both the Mesoamerican and Andean regions are based on a cosmic vision of life that is entirely different from the West.<sup>3</sup> Cartesian dichotomies that separate mind and body, humans and nature, nature

---

<sup>3</sup> An extended version of Section 6 appears in [11].

and society are foreign in these cultures. In their cosmic vision, all of these elements are interdependent; they must maintain an equilibrium for a harmonious existence. There is a fluidity that runs through the earth, heavens, water, wind, and the humans and non-humans that fuses them together. The cosmos is itself constituted by dualistic forces that are fluid, but not hierarchical as in Cartesian precepts, nor gendered. Thus, the feminine and masculine forces are complementary, of equal importance to the cosmos, and must maintain an equilibrium to guarantee the perpetuity of life.

Sociologically, this gendered division of the cosmos translates into gender complementarity, gender parallelism, or what the Aymara call *chachawarmi*. Man-woman constituted a unit of pairs. A married couple of man and woman were the basic unit of the community. Their work in tandem, although differentiated, was of equal worth. Women were not economically dependent on men. In gender parallel structures women constituted a lineage where inheritance was passed down to their daughters.

And yet, historically we can see those elements of gender hierarchies were present. Gender differentiation increased as empire and state-building advanced both among the Mexicas and the Incas [24, 25]. Men as soldiers and warriors had a public face that women lacked. Men were the representatives of the community before the ruler. While noble women had the class privilege and could occasionally occupy positions of power, the highest positions of power were still reserved for men. War, although understood to be as important as women's power of child birthing, constituted the center of power of indigenous *realpolitik*.

But it is the elements of complementarity, parallelism, and reciprocity between the genders that many indigenous men and women and their mestizo/criollo allies want to claim as either still existent or in need of resurrection. This position encounters many criticisms. Perhaps most important is the fact that this gender regime did not survive colonialism intact. Colonialism itself involved a social pact between colonized and colonizer men based on the acceptance of the subordination of indigenous women to their men in exchange for limited access for colonized men to power inside the community. Indigenous men while emasculated in the public sphere were granted the control of women, children, and the elderly in the household and the community. These gender colonial norms have in time installed gender violence, something that was unknown to them in the era of pre-intrusion. As Argentinean anthropologist Rita Segato has maintained, the separation of the public and private spheres not only privatized and minoritized indigenous women; it had lethal consequences for them [26]. More recent experiences of genocide, such as the one in Guatemala where the state forced indigenous men to rape, kill, and mutilate indigenous women, have increased violence against indigenous women dramatically, and thereby led to some of the highest femicide rates in the world.

## 7. Conclusion

As noted earlier, this is a moment of deep and widespread transformation of social institutions, including universities and their related educational, research and publication contexts. In this respect, it resembles the beginnings of the early modern era in European history; this may well be the "other end" of Western modernity and its philosophies of science. It seems to be a moment when we educated elites in the modern West can only now begin to glimpse the fact that Liberal democracy's meritocracy is a contradiction in terms, as many less fortunate groups have already understood. It does not encourage us to collaborate with others or treat them as equals. We can have a meritocracy or a democracy, but not both [27].

Our recognition of this tension can be a productive event. As a start, we can learn to “walk together” respectfully with peoples whom Western modernity has marked as deeply different from us. Standpoint methodology and its strong objectivity standard can be useful resources for this project.

### **Author details**

Sandra Harding  
Philosophy, New York University, United States

\*Address all correspondence to:

### **IntechOpen**

---

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Snow CP. *The Two Cultures: And a Second Look*. Cambridge: Cambridge University Press; 1964
- [2] Rose H. Hand, brain, and heart: A feminist epistemology for the natural sciences. *Signs: Journal of Women in Culture and Society*. 1983;9(1):73-90
- [3] Smith D. *The Everyday World as Problematic: A Sociology for Women*. Boston: Northeastern University Press; 1987
- [4] Hartsock N. The feminist standpoint: Developing the ground for a specifically feminist historical materialism. In: Harding S, Hintikka M, editors. *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology and Philosophy of Science*. Dordrecht: Reidel; 1983
- [5] Collins PH. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. New York: Routledge; 1991
- [6] Jasanoff S. *States of Knowledge: The Co-Production of Science and Social Order*. New York: Routledge; 2004
- [7] Jacob M. *The Cultural Meaning of the Scientific Revolution*. New York: Knopf; 1988
- [8] Harding S. Latin American decolonial social studies of scientific knowledge: Alliances and tensions. *Science, Technology & Human Values*. 2016;41(6):1063-1087
- [9] Harding S. Latin American decolonial studies: Feminist Issues. *Feminist Studies*. 2017;43(3):624-636
- [10] Harding S. State of the field: Latin American decolonial philosophy of science. *Studies in the History and Philosophy of Science*. 2019;78:48-63. DOI: 10.1016/j.shpsa.218.10.001
- [11] Harding S, Mendoza B. Latin American decolonial feminist philosophy of knowledge production. In: Crasnow S, Intemann K, editors. *Routledge Handbook of Feminist Philosophy of Science*. Oxford: Routledge; 2020
- [12] Said E. *Orientalism*. New York: Pantheon; 1972
- [13] Escobar A. *Worlds and knowledges otherwise: The Latin American modernity/coloniality research program*. In: Mignolo WD, Escobar A, editors. *Globalization and the Decolonial Option*. New York: Routledge; 2010
- [14] Mignolo WD. *Local Histories/Global Designs: Coloniality, Subaltern Knowledges and Border Thinking*. Princeton, N.J: Princeton University Press; 2000
- [15] Mignolo WD. *The Darker Side of Western Modernity: Global Futures, Decolonial Options*. Durham: Duke University Press; 2011
- [16] Dussel E. In: Barber MD, editor. *The Invention of the Americas*. New York: Continuum; 1995
- [17] Brotherston G. America and the Colonizer Question: Two Formative Statements from Early Mexico. In: Morana M, Dussel E, Jauregui CA, editors. *Coloniality at Large: Latin America and the Postcolonial Debate*. Durham: Duke University Press; 2008. pp. 23-42
- [18] Todorov T. In: Howard R, editor. *The Conquest of America*. New York: Harper and Row; 1984
- [19] Canizares-Esguerra J. *Nature, Empire and Nation: Explorations of the History of Science in the Iberian World*. Stanford: Stanford University Press; 2006
- [20] Saldana JJ, editor. *Science in Latin America: A History*. Austin: University of Texas Press; 2006

[21] Denevan WM. The Pristine Myth: The landscape of the Americas in 1492. *Annals of the Association of American Geographers*. 1992;**82**(3):369-385

[22] Mann CC. 1491: New Revelations of the Americas before Columbus. New York: Knopf; 2005

[23] Mendoza B. Coloniality of gender and power: From postcoloniality to decoloniality. In: Disch L, Hawkesworth M, editors. *The Oxford Handbook of Feminist Theory*. UK: Oxford University Press; 2015

[24] Kellog S. The woman's room: Some aspects of gender relations in tenochitlan in the late pre-hispanic period. *Ethnohistory*. 1995;**42**(4): 563-576

[25] Silverblatt I. Moon, Sun and Witches. Princeton: Princeton University Press; 1987

[26] Segato RL. Patriarchy from margin to center: Discipline, territoriality, and cruelty in the apocalyptic phase of capital. *South Atlantic Quarterly*. 2016;**115**(3):615-624

[27] Sandel MJ. *The Tyranny of Merit: Why the Promise of Moving Up is Pulling America Apart*. New York: Farrar, Straus and Giroux; 2020

# Challenges in Flood Management

*Vijay P. Singh*

## Abstract

Each year floods occur in many parts of the world and cause huge damages to agriculture, homes, schools, hospitals, highways, industries, water supply systems, infrastructure, levees, dams, and environment. They also cause loss of animal and human lives. Looking at the history of floods and damages caused, it is evident that they are amongst the costliest natural disasters and impact hundreds of thousands of people each year. It is widely accepted that floods cannot be eliminated entirely. However, they can be managed to mitigate the loss of life and property. Revisiting the types and causes of floods, this presentation focuses on the challenges in flood management. The challenges are both technical, including hydrometeorologic, hydrologic, hydraulic, geotechnical, and structural; and nontechnical, including education, communication and Internet, legal, administrative, social, political, risk analysis, and skilled professionals. The challenges have a wide variety but fall under seemingly disparate disciplines so the emphasis here is on their integration. Compounding these challenges is climate change whose impact can be assessed but whose forecast in space and time is still a challenge. The presentation is concluded with a personal reflection on paradigm shift.

**Keywords:** natural disasters, climate change, floods, risk, management, paradigm shift

## 1. Introduction

Each year natural disasters strike many parts of the world. Some parts are hit by heavy rains, some by heavy snowstorms, some by floods, some by mudslides, some by windstorms, some by hurricanes/cyclones/typhoons, some by heat waves, some by cold waves, some by snow avalanche, some by droughts, some by tornadoes, some by wild-fires, some by earthquakes, some by lightening, some by volcanic eruption, some by tsunami, some by viral/bacterial outbreaks, and some by a combination of one or more of these disasters, such as heavy winds accompanied by heavy rains, drought accompanied by heat wave, cold wave accompanied by heavy winds, heat wave accompanied by viral outbreak, to name but a few. These natural disasters cause loss of life, damage to property, disruption in social and cultural fabric, environmental degradation, and imbalance in the ecosystem. To illustrate the impact of some of these disasters, **Table 1** lists yearly average global annual deaths by decade. It is seen that floods, droughts, and earthquakes cause more loss of life than other disasters. Of these three major disasters, floods and droughts are more common and can occur during the same year or at the same time but at different places in the same country. Further, it often happens that floods ravage one part of the country and droughts ravage other parts at the same time. An example is India where each year during the monsoon or rainy season, floods occur in the Northeast and North but droughts in the West at the same time.

Decade	Drought	Earthquake	Extreme temperature	Flood	Impact	Landslide	Mass movement (dry)	Storm	Volcanic activity	Wildfire
1900s	130000	17302	0	63	0	5	13	1801	4494	0
1910s	8500	6280	0	10138	0	0	12	5995	648	107
1920s	472400	54935	0	428	0	43	0	11999	514	10
1930s	0	23770	169	436147	0	103	4	9384	318	7
1940s	345000	16187	0	10103	0	1753	0	12712	213	25
1950s	0	2093	150	205830	0	215	0	3126	510	1
1960s	150865	5236	113	3239	0	504	218	13393	324	7
1970s	11908	44022	155	5078	0	738	7	35734	53	1
1980s	55727	6015	534	5155	0	623	127	4667	2400	40
1990s	311	10359	932	9549	0	833	87	21115	97	86
2000s	115	45364	9106	5401	0	772	28	17213	24	63
2010s	3339	43302	11644	5811	0	1069	13	3177	71	52

[Source: <https://ourworldindata.org/ofidacred-by-decade-international-disaster-data>].

**Table 1.**  
Yearly average global annual deaths from natural disasters.



No.	Name	Year	Date	Area affected	Fatalities	Cost of damage
1	South Carolina Sea Island hurricane	1893	Aug-27	Sea Island, South Carolina	2000	\$279 million
2	Galveston hurricane and storm surge	1900	Sep-09	Galveston, Texas	8000	\$602.3 million
3	Miami hurricane and flooding	1926	Sep-18	Florida Atlantic Coast, Florida	372	\$1.49 billion
4	South Florida hurricane and flood	1928	Sep-16	Lake Okeechobee, Florida	2500–3000	\$1.5 billion
5	Labor Day Hurricane	1935	Sep-02	Florida Keys, Florida	500	\$100.0 million
6	New England hurricane and flooding	1938	Sep-21	New England, Long Island, New York	700	\$5.44 billion
7	Pacific tsunami	1946	Apr-01	Hawaii, Alaska	165	\$334.1 million
8	Hurricane Agnes flood	1972	Jun-19	Susquehanna, Lackawanna, Pennsylvania	128	\$18.0 billion
9	<b>Hurricane Katrina flooding</b>	2005	<b>Aug-29</b>	<b>Southern Louisiana, Louisiana</b>	1833	<b>\$103.9 billion</b>
10	<b>Superstorm Sandy</b>	2012	<b>Oct-29</b>	<b>New Jersey, New York</b>	233	<b>\$88.4 billion</b>
11	<b>Hurricane Harvey</b>	2017	<b>Aug-26</b>	<b>Houston, Texas</b>	89	<b>\$126.3 billion</b>

**Table 2.**  
*Storms impacts are costly (examples from the U.S.).*

Heavy rainstorms cause huge losses, as shown in **Table 2**. For example, Hurricane Harvey that struck the Houston area in Texas, U.S., caused damages worth US\$126.3 billion and 89 deaths, not to speak of untold misery and disruption in the community. It took a long time to recover from this hurricane. Along a similar vein, floods cause even more losses, as shown in **Table 3**. For example, floods that occurred in the Eastern U.S. on November 8, 1996, caused 187 deaths and damage worth US\$ 4.79 billion.

No.	Name	Year	Date	Area affected	Fatalities	Cost of damage
1	Mill River Dam flood	1874	May-16	Western Massachusetts	139	\$1.0 million
2	Johnstown flood	1889	May-31	Johnstown, Pennsylvania	2209	\$12.6 billion
3	Brazos River flood	1899	Jun-17	Freeport, Texas	284	\$271.0 million
4	Oregon Heppner flash flood	1903	Jun-14	Heppner, Oregon	324	\$171 million
5	Statewide Ohio flood	1913	Mar-23	Cincinnati, Miami River, Ohio	467	\$82.4 billion
6	Brazos and Colorado River flood	1913	Dec-05	Freeport, Waco, Texas	177	\$88.7 million
7	San Antonio flood	1921	Sep-10	San Antonio, Texas	215	\$70.2 million
8	Great Mississippi flood	1927	Dec-25	Mississippi River region, Mississippi	246	\$41.7 billion
9	<b>St. Francis Dam failure</b>	1928	<b>Mar-12</b>	<b>Los Angeles, California</b>	<b>400–600</b>	<b>\$291.8 million</b>
10	Great Northeast flood	1936	Mar-11	Maryland to Maine	200	\$85.2 billion

No.	Name	Year	Date	Area affected	Fatalities	Cost of damage
11	The Ohio River flood	1937	Jan-30	Pennsylvania, Ohio, West Virginia, Tennessee, Indiana, Illinois	385	\$151.6 billion
12	Los Angeles flood	1938	Feb-27	Los Angeles, California	115	\$1.24 billion
13	East Coast flood	1955	Aug-11	New England, Northern Virginia	200	\$778 billion
<b>14</b>	<b>Hurricane Camille and flooding</b>	<b>1969</b>	<b>Aug-17</b>	<b>The Gulf Coast of Mississippi, Mississippi</b>	<b>256</b>	<b>\$9.70 billion</b>
15	Black Hills flood	1972	Jun-09	Rapid City, South Dakota	238	\$988.3 million
16	Buffalo Creek flood	1972	Feb-26	West Virginia	125	\$64.0 million
17	Big Thompson Canyon flood	1976	Jul-31	Big Thompson Canyon, Colorado	144	\$156.3 million
<b>18</b>	<b>Floods in eastern U.S.</b>	<b>1996</b>	<b>Nov-08</b>	<b>Appalachians, Mid- Atlantic, Northeast</b>	<b>187</b>	<b>\$4.79 billion</b>
<b>19</b>	<b>Southeast U.S. flood</b>	<b>1998</b>	<b>Oct-17</b>	<b>Tampa, Florida</b>	<b>132</b>	<b>\$2.49 billion</b>

**Table 3.**  
*Floods impacts are costly (examples from the U.S.).*

## 2. Types and causes of floods

Depending on where they occur, floods can be classified into different types as: watershed, riverine, urban, coastal, and glacial. These different types of floods have different spatial scales. For example, glacial outburst cause flooding at a local level but can be more extensive if the dam is broken. Coastal flooding are confined to coastal areas and can wipe out beaches and damage wetlands and vegetation by bringing in salt sea water. Flooding is quite common in urban areas these days, because urban areas turn pervious areas into impervious areas which do not infiltrate rainwater. The different types of floods are caused by extreme rainfall, hurricanes, tides, combined rainfall and snowmelt, improper drainage, improper watershed management, dam/levee breaching, or glacial outbursts. The ubiquitous cause is extreme rainfall, but rainfall and snowmelt together are also a common cause, especially in areas where snowfall is extensive as in the United States.

Likewise, in monsoon climate countries in Asia, destructive floods occur each year massive investments made in flood defenses notwithstanding. In China, damages caused by floods have been over US\$200 billion per decade. Floods during the monsoon season have been commonplace in the Yangtze and its tributaries.

## 3. Flood management

It is accepted that floods cannot be entirely eliminated because nature cannot be fully controlled, but they can be managed so that the damages caused by them are mitigated. Thus, flood management involves two aspects: technical and nontechnical. Technical aspects are primarily engineering, including hydrometeorologic, hydrologic, hydraulic, geotechnical, and structural; and nontechnical aspects are education, socio-economic, political, legal, communication, internet, and administrative.

### 3.1 Hydrologic and hydrometeorologic considerations

Hydrology is basic to flood management and to answer basic questions which are fundamental to designing a flood management project. The questions needed for design are: (1) What will be the flood producing rainfall? (2) What will be the return period of this rainfall? (3) What will be the flood magnitude due to a given rainfall event? (4) What will be the probability or return period of a given flood magnitude? (5) What will be the risk of occurrence of a flood of given magnitude? The first three questions are answered by deterministic hydrometeorologic and hydrologic modeling, also called rainfall-runoff modeling or watershed modeling. There are many types of watershed models, such as empirical (regression type), conceptual (unit hydrograph theory), and physically-based (kinematic, diffusion wave, and dynamic wave theories). A comprehensive account of most of the popular models around the globe is given in Singh [1], Singh and Woolhiser [2], and in Singh and Frevert [3–5].

When managing floods, the questions are: (1) When will the flood occur at a given location? (2) How much area will be impacted by a given flood? (3) How long will a flood last? These questions are answered by stochastic hydrologic modeling, including univariate frequency analysis, multivariate stochastic analysis, and stochastic watershed modeling. Frequency analysis is done in different ways. The most popular method of frequency analysis in practice is the empirical method which involves fitting a frequency distribution to empirical flood data, use of an appropriate parameter estimation technique, goodness of fit, selection of a distribution, establishing confidence bands, and risk analysis. A good account of the frequency distributions and their fitting and parameter estimation is given in Kite [6], Singh [7], and Rao and Hamed [8], Zhang and Singh [9]. Often multivariate frequency analysis may be needed for not only design but also for management. That is most appropriately done using copulas which along with their applications are comprehensively described in Zhang and Singh [9]. A treatise on risk and reliability analysis in environmental and water engineering is provided by Singh *et al.* [10].

### 3.2 Hydraulic considerations

Hydraulics deals with flow in the river. For designing hydraulic structures for flood control, subsequent to hydrologic questions, there are hydraulic questions that need answering. These questions are: (1) What will be the flood stage and flood discharge in a river? (2) When will the flood stage exceed the flood threshold at a given location? (3) How long a river reach will be impacted by a given flood? (4) How much area will be flooded by such a flood? (5) How long will a flood last? (6) What will be the return period of such a stage and discharge? (7) What will be the probability or return period of a given flood stage and discharge? (8) What will be the risk of such a flood stage? These questions are answered by hydraulic modeling. Deterministic flood routing which is either empirical (relation between upstream and downstream hydrographs), or conceptual (Muskingum method), or physically-based (diffusion wave, dynamic wave) answers the first five questions. Singh [11] has given a full account of the deterministic flood routing. The last three questions are answered by stochastic hydraulics involving frequency analysis and multivariate stochastic analysis. Stochastic methods in hydraulics are similar to those in hydrology [9].

### 3.3 Geotechnical considerations

Geotechnical engineering primarily deals with foundations of structures. It answers a set of basic questions which must be answered before any construction,

such as: (1) What is the most appropriate site for constructing a given structure, such as a dam? (2) Can the local soil withstand the pressure? (3) What are the local soil characteristics? (4) How high a structure, such as a levee, should be? (5) Does the foundation need reinforcement? (6) What is the reliability of a given foundation? These questions are answered by geotechnical engineering. There are standard textbooks available which provide comprehensive accounts of these and related issues.

### **3.4 Structural considerations**

Structural engineering deals with the structural design of a flood control structure, such as a dam and its associated appurtenances like spillway, tunnels, etc. or levee. It computes forces the structure must withstand and its dimensions. To that end, it answers questions such as how large a structure should be, what the type of a structure should be, how reliable the structure will be, what skill set will be needed for construction, who will do the dam construction and who would do the supervision, and how long will it take to complete the dam. These questions are answered by structural modeling which can be deterministic, including empirical, conceptual, or physically-based, and reliability and risk-based, including reliability analysis and risk analysis. Standard textbooks are available which provide full accounts of these and related issues.

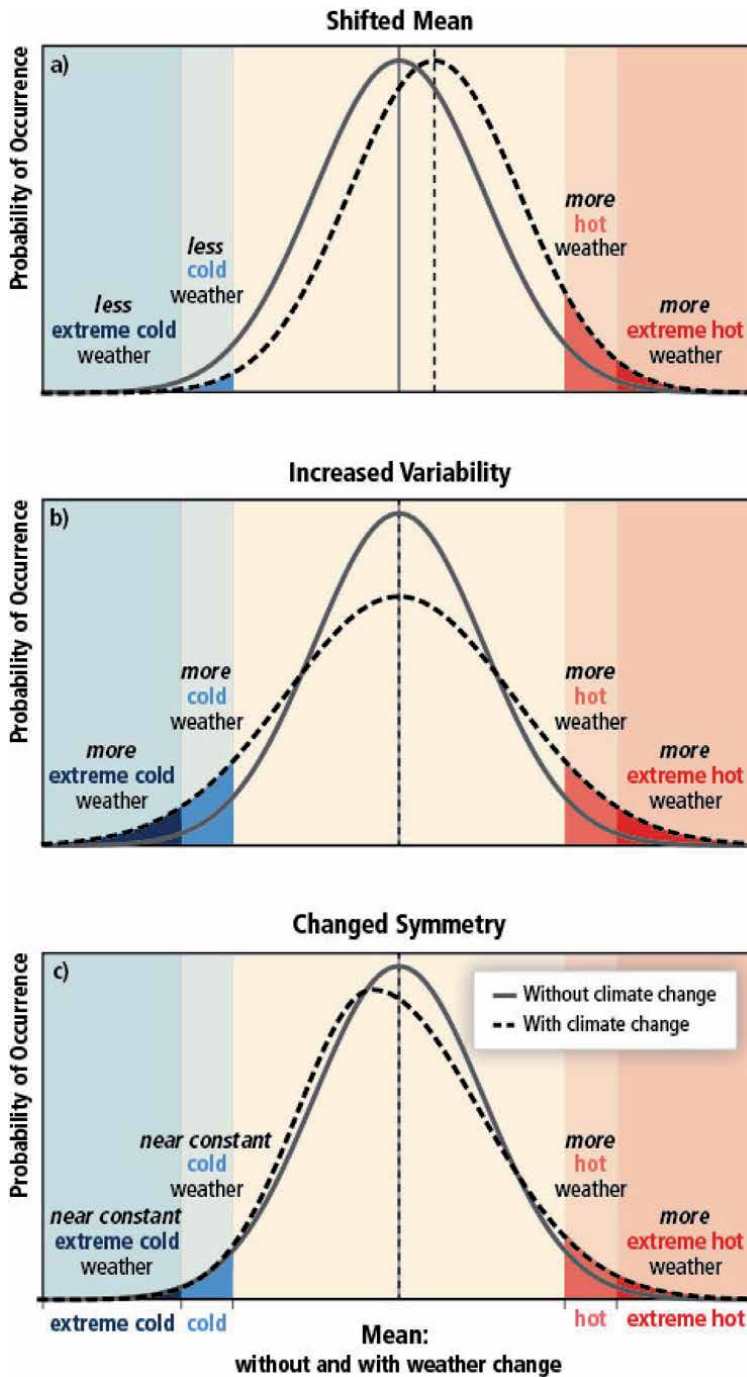
## **4. Challenges in flood management**

### **4.1 Climate change**

Climate change is a major challenge for humanity in this century. It indeed will decide the fate of our civilization. The Intergovernmental panel on Climate Change [12] notes: “A changing climate leads to changes in the frequency, intensity, spatial extent, duration, and timing of weather and climate extremes, and can result in unprecedented extremes ...” It has already started to impact the extremes of atmospheric weather and climate variables (temperature, precipitation, wind), the natural physical environment (floods, extreme sea level, waves, coastal waves, winds, and tornadoes). The questions often arise with regard to the assessment, forecasting-where, when, and how long-impact assessment-where, how much, and how serious a risk.

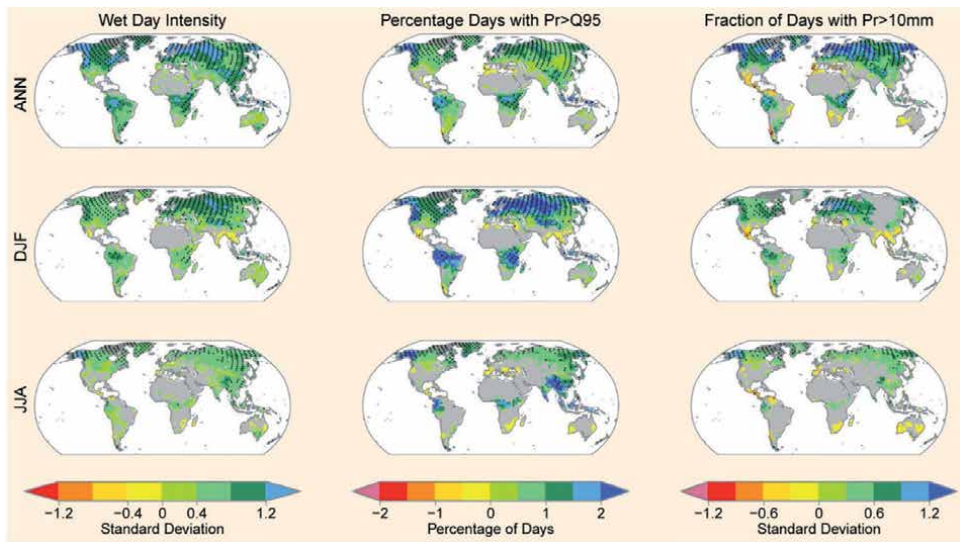
Three possible changes in weather extremes triggered by climate change are: less extreme cold but more extreme hot weather, more extreme cold and more extreme hot weather, near constant extreme cold but more extreme hot weather. It has a pronounced effect on the hydrological cycle and climate extremes as shown in **Figure 1** with and without climate change. The upper most part of **Figure 1** shows that there is a shift in the mean to the right from without climate change to with climate change, indicating less cold and less extreme cold but more hot and more extreme weather, whereas the middle part of **Figure 1** shows that there is an increase in variability with climate change, translating into more cold, more extreme cold, more hot, and more extreme hot weather, and the bottom part of **Figure 1** shows that the weather symmetry changes with climate change such that cold and extreme cold weather is nearly constant but there is more hot and more extreme hot weather.

Climate models are showing earlier occurrence of spring peak river flows in snowmelt- and glacier-fed rivers (already being observed), anthropogenic influence on changes in some components of the water cycle (precipitation, snowmelt) affecting floods, projected increases in heavy precipitation which would contribute to rain-generated local flooding in some catchments or region, and potential changes in the magnitude and frequency of floods. IPCC, SREX [12] shows the



**Figure 1.** Effect of climate change on weather extremes [Source: <https://www.ipcc.ch/report/managing-the-risks-of-extreme-events-and-disasters-to-advance-climate-change-adaptation/>].

impacts on precipitation, as shown in **Figure 2**, considering the standard deviation of wet day intensity, percentage of days with precipitation greater than Q95 (95% quantile), and standard deviation of fraction of days with precipitation greater than 10 mm for June, July and August (JJA); December, January, and February (DJF); and artificial neural network (ANN). In each case it is revealed that the standard



**Figure 2.**  
Effect of climate change on weather extremes (Source: IPCC, SREX [12]).

deviation increases over most parts of the world. IPCC, SREX [12] further shows for different parts of the world that higher 24-hour precipitation values will occur more frequently indicating their reduced return period. For example, for many parts a 20-year return period of 24-hour precipitation will reduce to 10 years or less. This means that there will be more frequent floods (Figure 3).

#### 4.2 Integration of disciplines

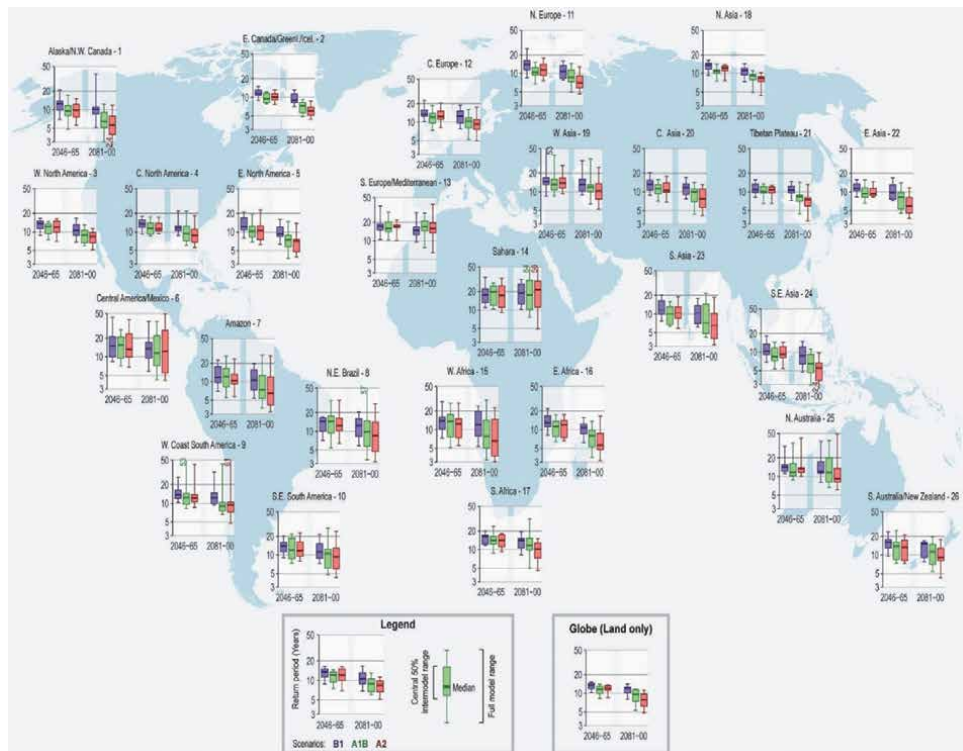
For effective management of floods, it is deemed that seemingly disparate disciplines that are associated with floods directly or indirectly should be integrated. These disciplines are: hydrometeorology, hydrology, hydraulics, agriculture, earth sciences, environmental sciences, socio-economic sciences, political and policy making, communication science, legal constraints, and administrative dimensions. These disciplines are on the flood process side. On the other hand, disciplines that provide tools for solving problems are mathematics, statistics, operations research, data science, geographical information systems, intelligent systems, and computer science. These disciplines should also be integrated with flood management.

#### 4.3 Communication

It is vitally important that agencies responsible for flood management communicate to the public as to why floods occur, likelihood of a flood in any given area, and roles and responsibilities associated with flood risk reduction and response. Needs of those people who unable to protect themselves are messages that are continually lacking to be conveyed. These messages do not “stick”, nor last, so they have to be regularly repeated even to the same audiences.

#### 4.4 Flood risk analysis

Conducting analyses of flood risks and contributors to increased flood risk are necessary to have substance in communications. That said, management of risk is unwanted, but necessary. No single organization, within the U.S. or international,



**Figure 3.** Projected return period (in years) of 20-year return values of annual maximum 24-hour precipitation rates (after IPCC, SREX [12]).

can control all aspects of population and property at risk from flooding or contributing to flooding. However, sharing risk is not desired by those who depend on or expect some other organization to provide their protection. The greater value of risk-based analyses lies in the better articulation of roles and responsibilities affiliated with flood risk reduction and response.

#### 4.5 Measurement

For developing flood control measures and flood management, spatial and temporal data from different disciplines are needed. More particularly, hydrometeorologic data, hydrometric data, watershed physiographic data, and land use and land cover data are needed to get started. Measurement technologies-remote sensing, satellite and drones- can be employed at a large scale. The remote sensing technology can provide information on rainfall fields, including storm movement, spatial variability, temporal variability, and rainfall field coverage. Also, measurements techniques are available that help describe the spatial variability of hydraulic roughness. The collected data should be subject to quality analysis/control, should be archived, and be retrievable. Then, the data needs processing and should be made accessible.

#### 4.6 Integrated hydrologic modeling

Hydrologic modeling should be integrated with remote sensing, geographical information system (GIS), data base management system, hydraulics, land use/land

cover, hydrometeorology, geomorphology, uncertainty and risk analysis. In distributed hydrologic modeling, it is important to quantify the effect of the spatial variability of watershed characteristics on runoff dynamics and hydrograph, and formation of shocks. Impacting the runoff or flood hydrograph is also the spatial variability of infiltration, hydraulic conductivity, steady infiltration, and mean infiltration. The spatial and temporal variability is directly dependent on scaling. Spatial scaling entails spatial heterogeneity in watershed characteristics, spatial variability in hydrologic processes, as well as physical spatial size involving representative elementary area, hydrologic response units, and computational grid size. On the other hand, temporal scaling involves time interval of observations, computational grid size, and temporal variability of processes. These issues play a vital role in flood model response.

An important issue in integrated modeling is calibration which involves parameter estimation algorithm, an objective function, an optimization algorithm, a termination criterion, calibration data, handling data errors, determination of data needs-quantity and information-richness, and representation of uncertainty of the calibrated model. Artificial neural networks can also be employed for modeling or model calibration.

In modern era, new tools are emerging or the existing tools are being made more accurate and versatile. These tools may include mechanistic models, data mining models, uncertainty analysis, entropy theory, risk analysis, multivariate stochastic analysis (copula theory), intelligent systems (ANN, Fuzzy logic, etc.), optimization algorithms, decision support systems, and GIS software.

With increasing demand on hydrologic models, new challenges are emerging. For flood modeling, such challenges are the need for more data at finer spatial resolutions, regional scale models, quantification of model uncertainty, long-term forecasting (ahead of time), determination of probable maximum precipitation and probable maximum flood, integration with climate models as well as with ecosystems models, and coupling with decision making models (social, political, economic, environmental, etc.).

#### **4.7 Watershed management**

Floods should be managed at the watershed scale and watershed management therefore becomes critically important. It involves land use management, drainage, soil conservation, forest management. There is growing need in the U.S. to provide increasing, and reliable, volumes of water for municipal, industrial, and agricultural needs. Reliable is a key criterion, especially during variable climatic conditions. Finding means to store flood waters in aquifers or move flood waters to areas experiencing water shortages are engineering and socio-political challenges where the U.S. will see increasing interest and pressure to address.

#### **4.8 Education**

In many cases people are unaware of the flood risk they expose themselves and families to, while in other cases people are intentionally ignorant so others can assume responsibility for their flood risk. Education is an essential long-term measure, but for education to make a difference it needs to be part of the K-12 education system. Education limited to project specific Town Halls and briefings to elected leaders is not achieving any significant change in societal behaviors.

#### **4.9 Skilled professionals**

While there is opportunity to improve hydrology and hydraulics and structural analysis tools and models, the tools available are mostly sufficient for the need.



What is lacking is experience and competence to use these tools appropriately on the most complicated projects. Identifying the right individuals and teams for unique tasks and convening multi-disciplinary teams with these special skills is a continuing issue and provides the rationale for the engineering, environmental, and social science professional fields to manage themselves and identify credentials recognizing those with advanced education and experience.

#### **4.10 Post-flood work**

Often when a flood has passed, leaving a lasting mark on people's lives and the environment, yet not as much attention is paid to post-flood work as it should be. The post-flood work involves rehabilitation, restoration, reconstruction, timely delivery of resources, and anxiety management.

#### **4.11 Environmental damage assessment**

Floods degrade water quality, damage water supply systems, cause loss of productive soil, harm the ecosystem, lead to viral and bacterial activity germane to diseases and harm to human health.

#### **4.12 Paradigm shift**

Given the socio-economic conditions prevailing these days all over the world, it is vitally important to ask as to what the development paradigm should be. Thus far, there seems to have been a more focus on concentrated development rather than distributed development. That is one reason for mounting losses due to floods. It seems that a more appropriate way to alleviate social unrest, reduce flood-caused losses, and improve the environment is to distribute the development. That will also reduce urban congestion, eliminate traffic jams, save energy, and reduce health care cost.

Another point that seems to be overlooked is the connect between decision makers and stakeholders. Policies are made for people or stakeholders but their input is not often vigorously sought. That leads to the disconnect between policy makers and the people for who the policies are being made. This seems like a contradiction but is often the case. In democracy, policies should be people-driven, not the other way round.

Further, in flood management the focus should be on a priori planning and management which is called proactive approach, but in most cases it is the reactive approach that is followed. It will require a concerted effort on the part of government agencies responsible for flood management to start adopting a proactive approach which will save lives and reduce damages.

### **5. Conclusion**

Floods are a natural disaster and cannot be eliminated entirely, but a priori planning and management can reduce their impact. Following a paradigm shift toward distributed development in place of concentrated development will go a long way in addressing the flood crisis which plagues many parts of the world each year. There is sufficient engineering technology available but society- and government-related issues still need to be fully addressed.

## **Author details**

Vijay P. Singh<sup>1,2</sup>

1 Department of Biological and Agricultural Engineering, Texas A&M University, College Station, Texas, United States

2 Zachry Department of Civil and Environmental Engineering, Texas A&M University, College Station, Texas, United States

\*Address all correspondence to: vsingh@ag.tamu.edu

## **IntechOpen**

---

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Singh VP. Hydrologic modeling: progress and future directions. *Geoscience Letters*. 2018;5(15):5-15
- [2] Singh VP, Woolhiser DA. Mathematical modeling of watershed hydrology. *Journal of Hydrologic Engineering*, ASCE. 2002;7(4):270-292
- [3] Singh VP, Frevert DK. *Mathematical Models of Large Watershed Hydrology*. Highlands Ranch, Colorado: Water Resources Publications; 2002. p. 891
- [4] Singh VP, Frevert DK. *Mathematical Models of Small Watershed Hydrology and Applications*. Highlands Ranch, Colorado: Water Resources Publications; 2002. p. 950
- [5] Singh VP, Frevert DK. *Watershed Models*. Boca Raton, Florida: CRC Press; 2006. p. 653
- [6] Kite GW. *Frequency and Risk Analyses in Hydrology*. 4th ed, Highlands Ranch, Colorado: Water Resources Publications; 2004. p. 257
- [7] Singh VP. *Entropy-based Parameter Estimation in Hydrology*. Dordrecht, The Netherlands: Springer; 1998. p. 365
- [8] Rao AR, Hamed KH. *Flood Frequency Analysis*. Boca Raton, Florida: CRC Press; 2000. p. 350
- [9] Zhang L, Singh VP. *Copulas and their Applications in Water Resources Engineering*. Cambridge, England: Cambridge University Press; 2019. p. 603
- [10] Singh VP, Jain SK, Tyagi A. *Risk and Reliability Analysis: A Handbook for Civil and Environmental Engineers*. Reston, Virginia: ASCE Press; 2007. p. 783
- [11] Singh VP. *Computer Models of Watershed Hydrology*. Highlands Ranch, Colorado: Water Resources Publications; 1995. p. 1130
- [12] IPCC, SREX. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*. New York: United Nations; 2012



# A Quest for Sustainability in the Food Enterprise

R. Paul Singh

## Abstract

The twenty-first century global food enterprise faces numerous challenges. The most critical is how to meet the food needs of the rapidly growing world's population that is expected to increase by 2 billion persons in the next 30 years. The food system is also under increasing threat from climate change. As a result, the resources required for increasing food production are becoming heavily constrained. Innovative approaches to mitigate these threats to the food system are needed. This paper's overall goal is to highlight challenges and opportunities to address the sustainability of the global food system. Various examples are drawn from the contemporary literature, including the author's research, to illustrate some of the steps needed to meet sustainability needs. Relevant issues are discussed for different food system segments from farm production to processing, distribution, storage, retail, and food preparation for consumption.

**Keywords:** sustainability, food system, climate change, food losses, food waste

## 1. Introduction

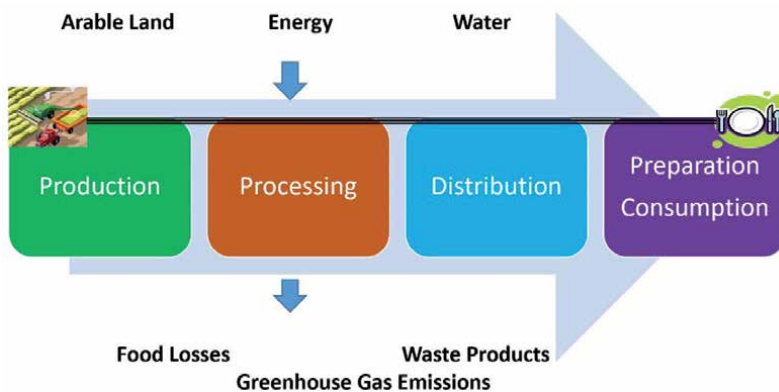
The agricultural production system's capacity to meet the increasing population demands has been questioned in the past, notably by Malthus in 1798 in "An Essay on the Principle of Population," where he theorized a specter of largescale deaths due to inadequate food production and increasing population [1]. Luckily, Malthusian prophecy did not materialize as technological advances in farming helped raise agricultural production to feed the growing population. Another alarm regarding the food system was raised by Sir William Crookes, a brilliant experimentalist known for discovering the element called thallium. Sir Crookes is also well known for his inaugural presidential speech, titled "The Wheat Problem," that he gave on September 10, 1898, to the British Association for the Advancement of Science [2]. In this talk, using data on wheat production and the increasing human population, he raised his concern about the food system's sustainability. He noted, "*we are drawing on the Earth's capital, and our drafts will not perpetually be honored. England and all... nations are in deadly peril of not having enough to eat.*" Sir Crookes' concern was based on a significant threat to the day's farming system, the potential depletion of fertilizer to grow wheat and other crops. In the late 1800s, 100% of the nitrogen used in farming was mined and shipped from Peru, Baja California, and Chile as guano. Guano is bird droppings that build up over a long period. But the mining fields were getting depleted of guano, and Sir Crookes could foresee that if the supply of guano is exhausted, then the farming will collapse, and millions will starve.

Being a chemist, he observed that the Earth's atmosphere has plenty of nitrogen. Sir Crookes challenged his fellow scientists to determine how to chemically fix nitrogen from the air to help to create what he called "chemical manure." One of the chemists, Fritz Haber, took him up on that challenge. Haber discovered the chemical reaction that allowed fixing nitrogen to make ammonia, and in 1918, he received a Nobel Prize. Working with Carl Bosch, he commercialized that research finding to create the Haber-Bosch process. The products of this process are used not only for agriculture but also for the manufacture of pharmaceuticals, plastics, textiles, and explosives.

In the 1900s, when Sir William Crookes was concerned about the food supply, the world population was less than 2 billion, now it is about 8 billion, and by the year 2050, it is predicted to increase to 10 billion. According to the current estimates, to meet the increasing population's needs and fill the food gap to 2050, an increase in agricultural production by almost 60% is required—a daunting task facing today's food and agricultural scientist [3]. Whereas the supply of guano was the main threat to the food production system in the late 1800s, today, multiple threats impact the food system. These include the increasing population, rapid urbanization rate, a dramatic ongoing depletion of natural resources, and the various impacts of climate change.

The United Nations has recognized the global scope of the problem by issuing a call for developing sustainable development goals [4]. These goals are intended to provide a blueprint to achieve a better and more sustainable future for everyone on the planet. A list of 17 sustainable development goals was identified. These goals underpin the future developmental projects supported by the United Nations. The global food system's sustainability has a significant role in several developmental goals such as zero hunger, good health, clean water, conserving marine resources, reversing land degradation, and climate action.

A simplified version of today's food system, from farm to fork, involves production to consumption, as seen in **Figure 1**. The output from agricultural production moves through processing, storage, and distribution sectors, before preparation and consumption either at home or out-of-home establishments. Primary inputs at various steps in the system include arable land, labor, energy, and water. There are food losses at each stage, and waste products, including wastewater, are generated, and greenhouse gas emissions are released into the atmosphere. Each step of the system will be next considered with a description of some of the threats it faces.



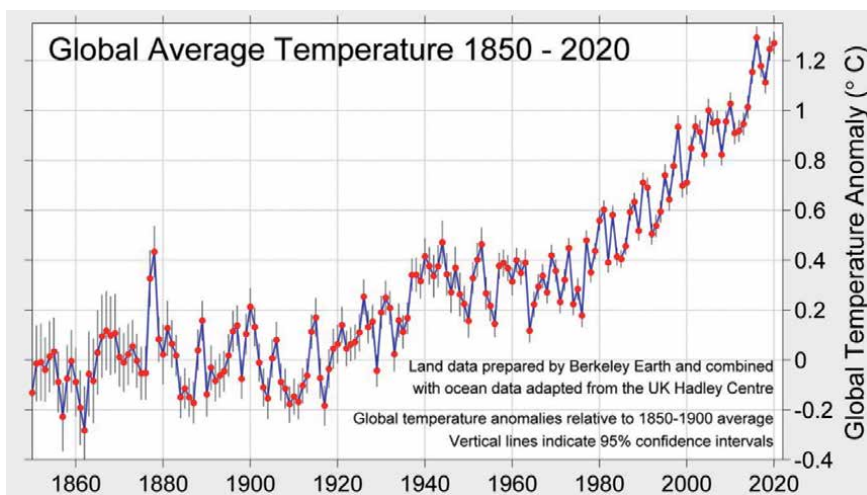
**Figure 1.**  
*A simplified version of a modern food system from farm to fork.*

## 2. Sustainability of agricultural production system

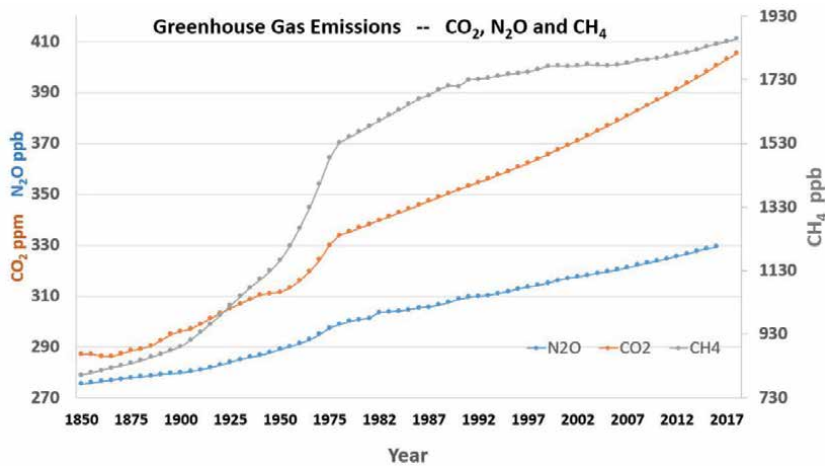
In response to the recurring famines in the early twentieth century caused by a lack of sufficient food supply and the rapidly increasing global population, several international research institutes focused on agricultural research were set up around the 1950s. Their mandate included developing science-based approaches to increase agricultural production. Among these institutes, the International Rice Research Institute (IRRI) in the Philippines is well known for developing many new rice varieties, resulting in a dramatic increase in rice production in South and South-East Asia. Similarly, the International Maize and Wheat Improvement Center (CIMMYT) in Mexico City, where Norman Borlaug and his colleagues developed dwarf varieties of wheat and new varieties of maize, significantly increased yield of these crops around the globe. Norman Borlaug received a Nobel Prize for his work at CIMMYT. World grain production has grown remarkably during the past 5 decades. Wheat production increased almost four to five times than what it was in the 1960s [3].

The agricultural production system is now under a significant threat by the many facets of climate change in meeting the impending food gap. The global average temperature has been increasing at an alarming rate, as seen in **Figure 2**. The last 6 years have been the hottest years on Earth [5]. This dramatic increase in temperature, trending upwards at a rapid rate, has severe consequences on agriculture. Along with an increase in the global average temperature, there is also a rapid increase in greenhouse gas emissions, mainly carbon dioxide, nitrous oxide, and methane (**Figure 3**). In each case, dramatic shifts have been occurring since the 1960s. A variety of economic sectors impact the global greenhouse gas emissions, such as industry, transportation, buildings, electricity and heat production, and agriculture, forestry, and land use. Up to 12% of the global greenhouse gas emission is attributed to agricultural operations (**Figure 4**) [3]. Estimates by the Intergovernmental Panel on Climate Change (IPCC) indicate that if there is no intervention within the agricultural sector, greenhouse gas emissions are likely to increase by about 30–40% by 2050 [6]. This estimated increase is mostly due to the increasing demands of the population, income growth, and dietary changes.

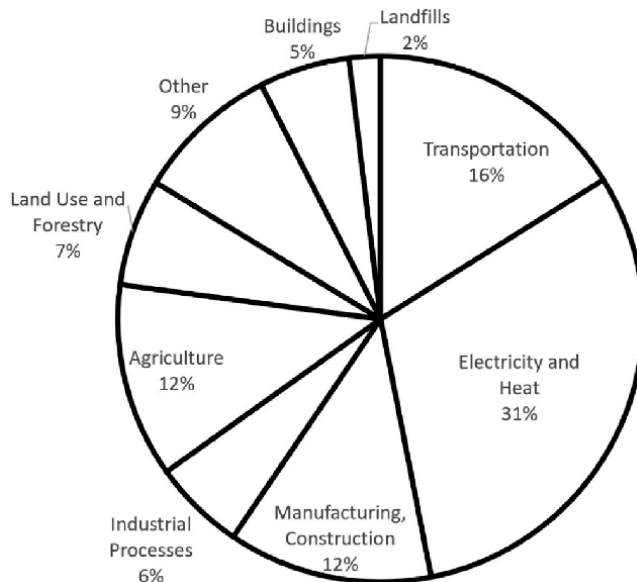
With climate change, the frequency of extreme weather events has been increasing. For example, the heat waves, melting of polar ice resulting in rising sea levels,



**Figure 2.** Global average temperatures from 1850–2020 (<http://berkeleyearth.org/global-temperature-report-for-2020/>).



**Figure 3.** Greenhouse gas emissions from 1850–2017 (based on data obtained from [5]).



**Figure 4.** World greenhouse gas emissions from various economic sectors in percent of total 49.4 Gigaton CO<sub>2</sub> equivalent in 2016 (based on data obtained from <https://www.wri.org/>).

increase in the number of heavy precipitation events causing floods, an increase in the length of drought periods, and increased incidence of wildfires as observed in California and Siberia. The strong links between agriculture and weather underscore the impact of weather on farming. In many regions with irrigated arable land, as more water is drawn from the underground aquifers for irrigation to overcome droughts' effects, the aquifers are getting depleted. For example, there has been a serious depletion of aquifers in central California in recent decades, causing land shrinkage and earthquakes [7]. Assuming the current rate of groundwater pumping for agriculture from the Ogallala Aquifer, it will be depleted by 60% by 2060 [8]. Water drawn from the Ogallala aquifer is used to meet 30% of the U.S. irrigation requirements. Similar impacts of climate change are seen in the western part of the Gulf of Mexico and the Indo-Gangetic plain, which serves as India's breadbasket.



Aquifers take a very long time to replenish. Therefore, the lowering of the water table in these heavily farmed regions is of grave concern to agricultural production sustainability.

Recent studies on the impact of climate change on agricultural production indicate that there will be a 25% reduction in maize production for most regions of the globe, a 3% reduction in wheat, and an 11% reduction in rice and potatoes [9]. These estimates indicating significant decreases in the crop yield will challenge efforts to meet the food gap predicted for the next decade. Along with reducing the yield, the increased carbon dioxide levels due to greenhouse gas emissions are also projected to lower the crops' nutritional quality. For example, when wheat is grown at high carbon dioxide levels, there is 6–12% less protein, 4–6% less zinc, and 5–7% less iron [6]. The reduction of nutrients in staple crops will have severe consequences for public health. Other climate change-driven impacts include the emergence of new pests and diseases, such as citrus greening, with growing risks and disruptions in the food system. Any shortages and subsequent increases in cereal prices will put more people at risk of hunger. Innovative farming practices are being considered to help mitigate some of the negative impacts of climate change such as increasing the soil organic matter and erosion control, improved land management, genetic improvements of crops for tolerance to heat and drought, and more diversification of the food system to implement integrated production systems. To address the needs of a sustainable agricultural production sector, many academic institutions in the United States are now focused on developing “smart” farming methods, seeking technological innovations in farming employing more efficient ways to use water and energy. For example, a multidisciplinary program referred to as SmartFarm at the University of California, Davis [10]. Similar efforts are underway at several land-grant universities in the United States.

In assessing the influence of producing foods for human consumption on the global environment, meat and dairy products rank high on the list. Meat production from livestock is responsible for using 30% of global ice-free land, 8% of global freshwater, and it generates 18% of the worldwide greenhouse gas emissions [11]. Many public and private institutions are currently engaged in research for developing cultured meats produced *in vitro* using tissue engineering techniques. Cultured meat production has the potential for substantially lowering the impact on the environment. Based on a life cycle assessment study, the environmental impact of cultured meat production in comparison to conventionally produced European meat, depending upon the product selected, shows 7–45% lower energy use, 78–96% lower greenhouse gas emissions, 99% lower land use, and 82–96% lower water [12]. Cultured meat production offers numerous opportunities for research and development for scale-up from the laboratory to the marketplace.

The increasing trend in urbanization has created numerous megacities worldwide—for example, Mexico City, with a population of 24 million, and Tokyo, with almost 40 million. Many of the cities with large populations are facing inner-city food deserts. Novel opportunities are being considered to fulfill the needs of fresh foods in the inner cities to develop urban agriculture, including vertical farming, and the production of vegetables and other crops under a controlled environment. These new farming methods in urban environments offer considerable opportunities for research and development of sustainable production, processing, and distribution systems.

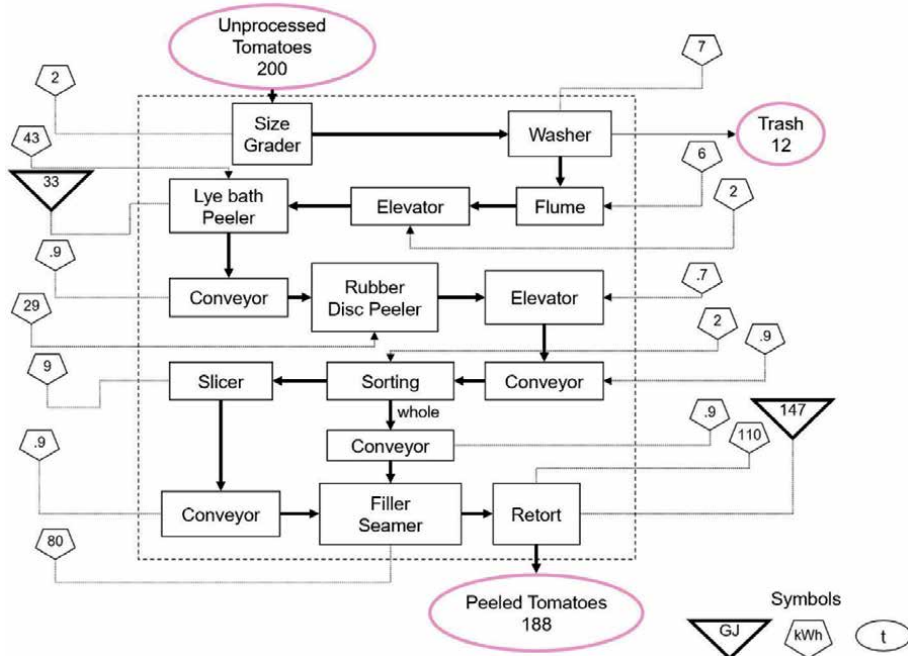
### **3. Sustainability in food processing**

In a modern food processing plant, it is not uncommon to find equipment designed and built several decades ago during the era of plentiful water and energy.

Since water use and energy use were most often not used as design constraints, there is considerable opportunity for retrofit and new design of systems to efficiently use water and energy. To identify such opportunities, industrial data of resource use in processing operations is crucial. Studies aimed at energy accounting conducted in food canning plants provide such data methodologies [13]. For example, as seen in **Figure 5**, the energy accounting diagram of canning whole-peeled tomatoes provides quantitative information on energy use in the form of electricity and natural gas and the mass flow of products. The energy use data obtained from accounting studies are helpful to identify energy-intensive operations to develop modifications and design new equipment to conserve energy.

Recent advances in sensor technology, data acquisition, and data handling offer ways to collect and retrieve data using cloud-based systems. Process data from line operations are passed on to the cloud server, stored, and made available to the equipment manufacturer for remote diagnostics and updates. Such systems offer advanced control and maintenance levels to minimize equipment breakdown and the loss of food during manufacturing operations. The development of these systems for the food industry requires skills in the computational field and electronic hardware.

A related emerging area in food manufacturing is creating digital twins of processing equipment. The digital twin technology has its origins in the aircraft industry. There is a digital twin for an airplane in flight, essentially a simulation of the plane fed with live data from the aircraft in flight to help identify any operational issues before they become severe. A similar approach is also feasible in the food processing industry. For any processing equipment, a digital twin operates in a virtual environment, providing valuable information to operators and equipment manufacturers. These systems can reduce frequent interruptions in the processing lines, thus reducing food losses during processing. While artificial intelligence and machine learning are still in their infancy, they promise to minimize human error in food processing operations.



**Figure 5.** Energy accounting diagram of canning of peeled tomatoes [13].



**Figure 6.**  
*A disc-peeler used to separate tomato peels.*

Along with energy, considerable water is used in food processing operations. Water recovery and recycling are vital for sustainability. A typical practice in a food processing plant is to discharge water streams from various processing equipment into a common floor drain. Different water streams containing multiple chemicals used in processing and cleaning equipment get mixed in the common drain, and the commingled stream is then conveyed to a water treatment facility. A potential approach to reduce water use and food waste is to recover effluent water from each piece of equipment separately to recover any food or chemicals and recycle water in the same or other operations as appropriate. For example, as shown in **Figure 6** for canning whole-peeled tomatoes, pure water is used to aid the separation of the peel from the tomato in the disc-peeling process. The effluent from the disc peeler is water with tomato solids. By separately treating the peeler's effluent using a filtration system, both the tomato solids and water are recovered. Numerous such examples exist for different food processing operations where economically valuable food and chemicals can be recovered as long as the discharge from individual operations is handled separately without mixing discharge streams into a common waste stream. Membrane-based separation systems are most suitable for such applications. A comprehensive project conducted at the University of California demonstrated this water recovery and recycling approach at over 50 food processing plants across the United States [14]. This project also reinforced the importance of industrial collaboration in academic research to reduce water use and improve the food system's sustainability.

In designing the next generation of food processing equipment, it is imperative that due consideration is given to design constraints such as low water discharge and minimal energy use. There are certain situations where these constraints become essential. For example, these constraints were at the forefront in a project to design a food processing system for a manned mission to Mars under a contract with the National Aeronautics and Space Agency (NASA) [15]. Specifically, a multipurpose fruit and vegetable processor was built for operation on the Mars surface (**Figure 7**). The design of this equipment involved a strict design constraint of zero-water discharge and the use of minimal energy. Several innovations were introduced to process fresh fruits and vegetables such as tomatoes to create multiple products. Based on the results of parallel research studies to determine optimal processing conditions, a multipurpose processor was fabricated using an ohmic heating system for rapid heating of crushed tomatoes, and membranes for separation processes. The final processed products were diced tomatoes, tomato juice, tomato sauce, and tomato paste. Water extracted from tomatoes during the concentration process was recovered and reused for cleaning equipment and other purposes. With minimal energy requirements, the



**Figure 7.**  
*A multipurpose fruit and vegetable processor built for manned mission to Mars.*

processor, although built for space applications, is equally adaptable for small-scale processing operations on Earth. Notably, the project demonstrated that it is possible to incorporate novel concepts in designing equipment that is highly conserving in its resource use. This equipment scale is particularly well suited for processing products of urban agriculture with minimal release of effluents in the inner-city setting.

Recent developments in the area of additive manufacturing offer new opportunities for precision food processing. While the 3D printing of foods is mostly in the research stage, this technique promises minimal food loss and an efficient process with low water and energy use. Additive manufacturing processes are also being considered in new food product development involving meat analogs derived from plant proteins. Meat analogs are gaining rapid growth in consumer acceptance. They offer health benefits and improved sustainability of the food system by reducing reliance on meat from livestock in the traditional diet.

#### **4. Reducing food losses and waste for a sustainable food system**

In the United States, food wastes amount to approximately \$278 billion annually, equivalent to feeding nearly 260 million people [16]. Globally, more than 1 billion metric tons of food per year never make it to the market. The market value of this lost food is almost a trillion dollars, and it has a significant negative impact on the environment. Food lost and wasted each year results in about 8% of the annual greenhouse gas emissions.

Around the globe, food losses are generally in the range of about 30% [17]. Many factors contribute to food losses, and they vary depending upon the region. In sub-Saharan Africa, a considerable amount of food loss occurs at the production stage, typically on-farm or close to a farm, during the handling and storage of harvested crops. These high losses are often due to a lack of proper infrastructure for the safe storage of cereal grains and a cold chain for perishables such as fruits and vegetables. However, in these regions, food losses during home preparation are generally small. In North America and some of the more industrialized countries, food losses during the production stage are small because of the highly developed infrastructure of the storage and transportation sector. Still, losses increase notably at the home and out-of-home preparation and consumption stage. Therefore, region-based solutions are necessary to reduce food losses for a sustainable food system.

In the food processing sector, trimming, overproduction, product and packaging damage, product graded as of low market value due to esthetic reasons, and technical malfunctions of processing equipment are often cited as fundamental causes of food losses and waste [18]. To minimize these losses in the processing sector, technological know-how and resources for operators need improvement, including

training the staff and reengineering processes to avoid product wastage during changes in product lines [17].

In most industrialized countries, packaged foods are often labeled with an expected shelf life to inform the consumer of how long the manufacturer assures safety and quality. While there is considerable merit in providing such information to the consumer, unfortunately, due to the lack of a standard shelf-life dating system, considerable confusion exists in interpreting shelf-life information. Furthermore, both elapsed time and environmental conditions, most notably temperature, affect food quality and safety. Consequently, using only a time-based shelf-life dating system, there is increased food wastage at the consumer level when acceptable food is discarded just because the label indicates that an expiration date has been reached. Since many of the food's quality characteristics change due to an integrated effect of time and temperature, there has been considerable interest in developing indicators that can be used for objective interpretation of the food's shelf life. Research in this area originated in the early 1980s [19]. Time-temperature indicators are used commercially in the distribution of vaccines and other medical drugs. They provide an objective indication of any heat abuse that a product may have received during shipment and storage and its remaining shelf life. While the early devices used biochemical or polymeric materials as indicators, with recent advances in electronic sensing and miniaturization, digital indicators are now being investigated for these applications. With cloud-based systems, data obtained from the indicators can be directly transferred to the server and used for inventory management. Such systems can be effectively used in the transportation, distribution, storage, and retail marketing of perishable foods [16, 20].

An emerging technology, blockchain, offers considerable promise to manage and share data in the food distribution systems. Blockchain allows a decentralized approach to distributing encrypted records of data securely over peer-to-peer networks. Besides information about product flow, other data relevant to food safety, quality, and resource use can be efficiently transmitted transparently. In tracking food from production to retail, this technology, when fully implemented, offers the potential to improve the safety and quality of food delivered to the consumer. Integrated systems for accessing and processing data on distribution are especially useful in time-sensitive situations involving product recalls. Innovations in food distribution such as blockchain will be necessary for the quest to improve the sustainability of the food system.

## **5. Conclusions**

The current food enterprise is under multiple threats from increasing population, depletion of resources, and the impact of climate change. Challenges in developing sustainable solutions to address these threats offer numerous research opportunities for innovations in the food processing sector, increase in agricultural production, and reduction in food wastage. Multidisciplinary efforts and cutting-edge developments will be necessary to approach many of the complex problems facing the food and agricultural enterprise. Since food is essential to sustain life, it is indeed the responsibility of everyone to ensure that the food system is sustainable not only for the current but also for future generations. Advances in science and technology are deemed to play a major role in addressing our food system's future sustainability.


### **Author details**

R. Paul Singh  
Department of Biological and Agricultural Engineering, University of California,  
Davis, CA, USA

\*Address all correspondence to: [rpsingh@ucdavis.edu](mailto:rpsingh@ucdavis.edu)

### **IntechOpen**

---

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Malthus TR. An Essay on the Principle of Population. Electronic Scholarly Publishing Project. 1798. Available from: <http://www.esp.org/books/malthus/population/malthus.pdf> [Accessed: January 5, 2021]
- [2] Giffen R. The Wheat Problem. *Nature*. 1899;**61**:169-171
- [3] World Resources Institute. Creating a sustainable food future—A menu of solutions to feed nearly 10 billion people by 2050. Available from: <https://research.wri.org/wrr-food> [Accessed: January 5, 2021]
- [4] United Nations Department of Economic and Social Affairs. Sustainable Development Goals. Available from: <https://sdgs.un.org/goals> [Accessed: January 5, 2021]
- [5] Berkeley Earth. Global Temperature Report for 2020. Available from: <http://berkeleyearth.org/global-temperature-report-for-2020/> [Accessed: January 5, 2021]
- [6] IPCC. Climate Change and Land. United Nations: The Intergovernmental Panel on Climate Change; 2019. Available from: <https://www.ipcc.ch/report/srccl/> [Accessed: January 5, 2021]
- [7] Richey AS, Thomas BF, Min-Hui L, Raeger JT, Famiglietti JS, Voss K, et al. Quantifying renewable groundwater stress with GRACE. *Water Resources Research*. 2015;**51**(7):5217-5223
- [8] Steward DR, Bruss PJ, Yang X, Staggenborg SA, Welch SM, Apley MD. Tapping unsustainable groundwater stores for agricultural production in the High Plains Aquifer of Kansas, projections to 2110. *Proceedings of the National Academy of Sciences of the United States of America*. 2013;**110**(37):E3477-E3486
- [9] Robertson R. Change in potential yields of maize, potatoes, rice, and wheat in 2050. Available from: <https://www.nationalgeographic.com/climate-change/how-to-live-with-it/crops.html> [Accessed: January 5, 2021]
- [10] University of California. Smart Farm. Available from: <https://bigideas.ucdavis.edu/sustainable-agriculture-smart-farm> [Accessed: January 5, 2021]
- [11] FAO. Livestock's Long Shadow—Environmental Issues and Options. Rome: Food and Agricultural Organization of the United Nations; 2006.
- [12] Tuomisto H, Teixeira de Mattos MJ. Environmental impacts of cultured meat production. *Environmental Science and Technology*. 2011;**45**:6117-6123
- [13] Singh RP. Energy in Food Processing. Amsterdam, The Netherlands: Elsevier Publishing; 1980
- [14] Mannapperuma J, Merson RL, Shoemaker SP. Testing and demonstration of membrane technology in the food industry. In: Abstracts of Papers of the American Chemical Society. Vol. 207. Washington, D.C.: American Chemical Society; 1994. p. 37-AGFD
- [15] Voit DC, Santos MR, Singh RP. Development of a multipurpose fruit and vegetable processor for manned mission to Mars. *Journal of Food Engineering*. 2006;**77**(2):230-238
- [16] National Academies of Sciences, Engineering, and Medicine. A National Strategy to Reduce Food Waste at the Consumer Level. Washington DC: The National Academies Press; 2020
- [17] Flanagan K, Robertson K, Hanson C. Reducing Food Loss and

Waste—Setting a Global Action Agenda. Washington, D.C: World Resources Institute; 2019

[18] CEC (Commission for Environmental Cooperation). Characterization and Management of Food Waste in North America: Foundational Report. Available from: <http://www.cec.org/islandora/en/item/11774-characterization-and-management-food-waste-in-north-americafoundational-report> [Accessed: January 5, 2021]

[19] Wells JH, Singh RP. A kinetic approach to food quality prediction using full-history time-temperature indicators. *Journal of Food Science*. 1988;53(6):1866-1871

[20] Herrero M et al. Innovation can accelerate the transition towards a sustainable food system. *Nature Foods*. 2020;1:266-272



# Evaluation of the Cytotoxic Activity of a Species of the *Buddleja* Genus in a Prostate Cancer Cell Line

*Sofía Isabel Cuevas Cianca, Luis Ricardo Hernández and Irene Vergara Bahena*

## Abstract

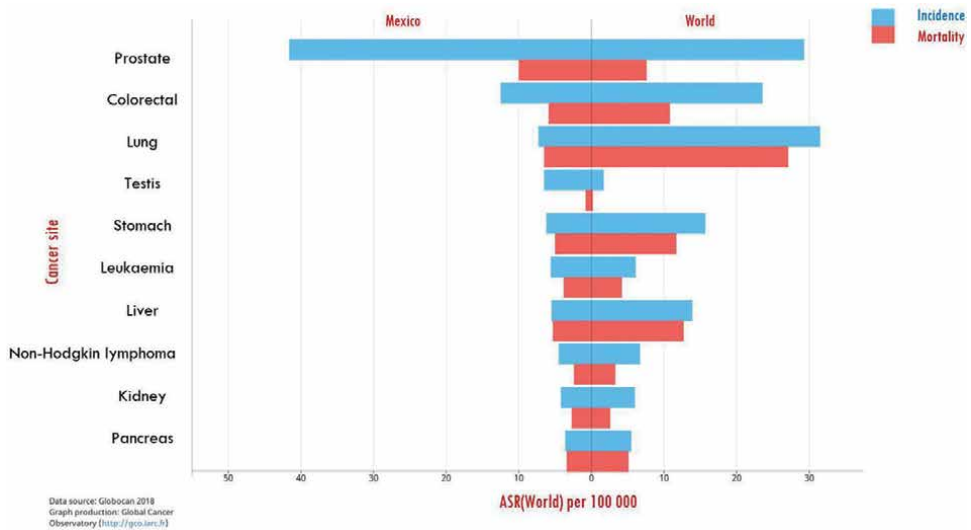
Over the centuries, humans have used medicinal plants to treat various diseases. Initially, these medications took the form of crude medications such as tinctures, teas, poultices, powders, and other herbal formulations. Almost 80% of the world population uses traditional medicines for primary health care, most of which involve the use of plant extracts. The study of plants continues, mainly, with the aim of discovering new secondary metabolites that can be used to recover health, both human and animal or vegetable. Cancer is a major public health problem worldwide and Mexico is not exempt from this problem. However, the great challenge for anticancer treatments is the specific release of the drug in the tumor tissue to avoid the adverse effects on normal cells. In this investigation, a species of the *Buddleja* genus is studied in terms of its cytotoxic activity in a prostate cancer cell line. Regarding the results found, it was obtained that the polar extract of aerial parts and the medium polarity extract of aerial parts have no cytotoxicity and high cytotoxicity, respectively against a prostate cancer cell line.

**Keywords:** *Buddleja*, prostate cancer, cytotoxicity, medicinal plants

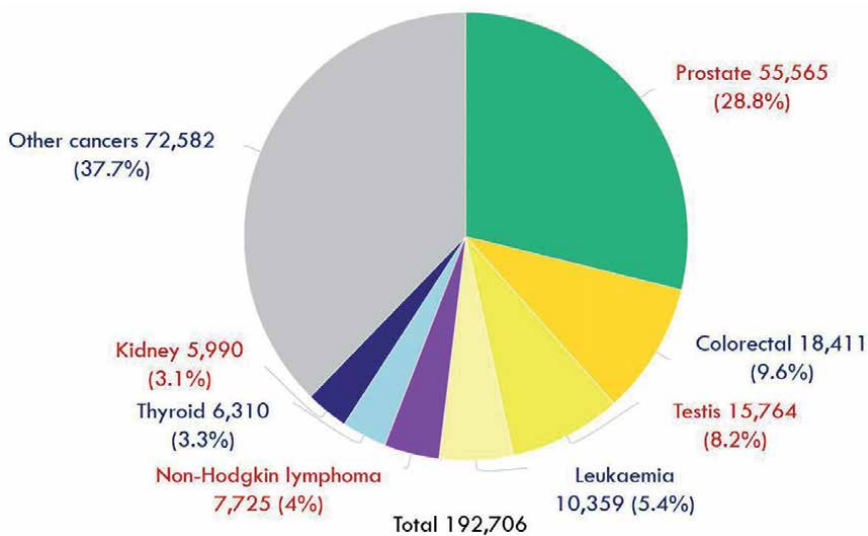
## 1. Introduction

Cancer constitutes a major public health problem worldwide and Mexico is not exempt from this problem. However, the great challenge for anticancer treatments is the specific release of the drug in the tumor tissue to avoid adverse effects on normal cells. Prostate cancer in Mexico represents the one with the highest incidence in men with 41.6 per 100,000 inhabitants in 2018. Likewise, worldwide, prostate cancer represents the second type of cancer with the highest incidence after lung cancer in 2018 (**Figure 1**). Prostate cancer in Mexico represents the highest mortality in men with 10 per 100,000 inhabitants in 2018 and the highest prevalence (5 years) in men with 55,565 cases from 2013 to 2018, respectively (**Figure 2**) [1].

Cancer occurs when healthy prostate cells change and proliferate uncontrollably, eventually forming a tumor. A tumor can be cancerous or benign. When a cancerous tumor is malignant, it means that it can grow and spread to other parts



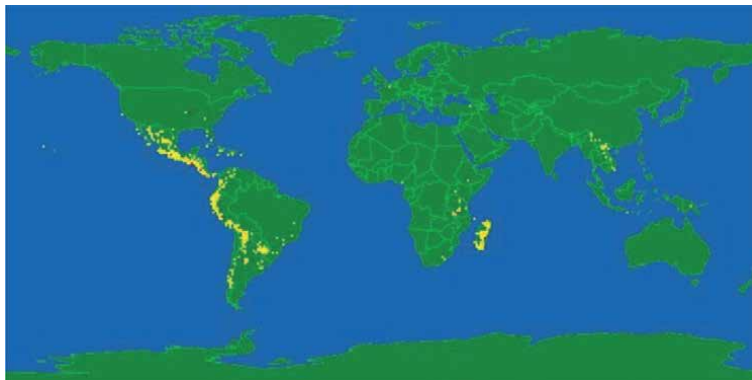
**Figure 1.** Comparison of the incidence and mortality worldwide and in Mexico of prostate cancer in men of all ages (based on [1]).



**Figure 2.** Estimated 5 year prevalence in Mexico of prostate cancer in men of all ages (based on [1]).

of the body. When a tumor is benign it means that the tumor can grow, but it will not spread [2]. Some types of prostate cancer grow very slowly and may not cause symptoms or problems for years. Even when prostate cancer has spread to other parts of the body, it can often be controlled for a long time, allowing men even with advanced prostate cancer to live in good health and quality of life for many years. However, if cancer cannot be controlled well with existing treatments, it can cause symptoms such as pain and fatigue, and can sometimes lead to death. An important part of managing prostate cancer is monitoring growth over time, to determine if it grows slowly or quickly [3].

Over the centuries medicinal plants have been used as raw medicines in the form of tinctures, teas, poultices, and powders to treat all kinds of diseases. Currently,



**Figure 3.**  
The geographical location of plants of the *Buddleja* genus (recovered from [www.Tropicos.org](http://www.Tropicos.org)).

80% of the world population uses traditional medicines, the majority involves the use of plant extracts, and 50% of all medicines for clinical use in the world come from plants, where higher plants provide no less than 25% of the total [4, 5].

The chemical study of the plant kingdom has provided a large number of potentially useful compounds, and since only a small percentage of the planet's superior plant species have been investigated for their active compounds, the chemical study of plants is considered to follow being promising for the discovery of pharmacologically useful compounds [6].

Recent phytochemical studies of plants that have or do not have an ethnobotanical history for the treatment of cancer have often resulted in the isolation of principles with antitumor activity, finding active metabolites such as flavonoids and chalcones [7], alkaloids [8, 9], sesquiterpenic lactones [10], diterpenes [11], and cardenolides [12] among others, which were shown to have activity against cancer cells.

The *Buddleja* genus (Fam. *Scrophulariaceae* Juss.) has around 300 species of shrubs, where there are both perennial and deciduous species. This group of plants is native from the southern United States to Chile and from Africa and warm parts of Asia (Figure 3). Dioecious plants are found in the southern part of the United States up to Chile, while monoecious plants are found in Africa and Asia [13].

There are some studies about species of the *Buddleja* genus, it was analyzed the chemical composition of *Buddleja polystachya* essential oil where it was found that there are monoterpenes such as bulnesol and limonene; this oil showed cytotoxic activity against carcinoma cell lines [14]. The antiproliferative and apoptotic activity of *Buddleja davidii* extracts was studied in gastric cancer and breast cancer cell lines, where it was concluded that colchicine and luteolin generate apoptosis in cells, which makes them potential drugs for the treatment of carcinoma, it was observed that they also generate apoptosis in tumor cells [15].

It should be emphasized that it is of great importance to carry out the study of extracts and fractions of a plant of the *Buddleja* genus regarding its cytotoxicity since, thanks to previous studies of plants of the same genus, highly toxic results have been obtained, which may be an indication that the plant has anticancer activity.

## 2. Methodology

To be able to determine the cytotoxic activity of the extracts, the growth of tumoral cells quantitated by the ability of living cells to reduce the yellow dye 3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) to a

purple formazan product was used. The cells were seeded and incubated in a 37°C incubator supplemented with 5% CO<sub>2</sub>. The products to be evaluated were added to the culture of the cells at different concentrations once the cells reach 80% confluence. At the end of the 24 h of incubation of the previously treated cells, 40 µL/well of the MTT solution (5 mg/mL in phosphate-buffered saline) were added, the plates were incubated for 3 h under 5% CO<sub>2</sub> and 95% air at 37°C. At the end of the 3 h of incubation, 400 µL/well of the solubilizing solution is added and gently shaken. The microplate is kept at room temperature in darkness for 24 h. The absorbance was then determined by a microplate reader at 490 nm. The percentage of growth inhibition was calculated using the following formula:

$$\% \text{ cell viability} = \left[ \frac{A_t - A_b}{A_c - A_b} \right] \times 100 \quad (1)$$

where, A<sub>t</sub> = absorbance value of test compound, A<sub>b</sub> = absorbance value of blank, and A<sub>c</sub> = absorbance value of control. The effects of extracts were expressed by LC<sub>50</sub> values (the drug concentration necessary to reduce cell viability to 50% with respect to untreated cells) [16].

### 3. Results

The results obtained in this study showed that the medium polarity extract of aerial parts of the plant of the *Buddleja* genus showed cytotoxic activity against a prostate cancer cell line, while the polar extract of aerial parts has no cytotoxic activity against a prostate cancer cell line. More studies are needed on the extract that showed activity, it is necessary to do a chromatographic separation followed by isolation of the compounds that give the desired cytotoxic activity.

### 4. Conclusions

Overall, this study evaluates that the medium polarity extract of aerial parts of the plant showed cytotoxic activity against a prostate cancer cell line, the next step is to determine which compounds are responsible for this biological activity and thus obtain potential drugs for the treatment of cancer. This study provides only basic data, further studies are necessary for isolation and identification of biologically active substances from these extracts, as well as to determine what type of death they cause through flow cytometry.

## Author details

Sofía Isabel Cuevas Cianca, Luis Ricardo Hernández\* and Irene Vergara Bahena  
Department of Chemical-Biological Sciences, Universidad de las Américas Puebla,  
Puebla, Mexico

\*Address all correspondence to: [luisr.hernandez@udlap.mx](mailto:luisr.hernandez@udlap.mx)

## IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Bray F, Ferlay J, Soerjomataram I, Siegel R, Torre L, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*. 2018;**68**(6):394-424. DOI: 10.3322/caac.21492
- [2] American Society of Clinical Oncology. 2018. Available from: <https://www.cancer.org/cancer/prostate-cancer/causes-risks-prevention.html>
- [3] Gomes I, Arinto P, Lopes C, Santos C, Maia C. STEAP1 is overexpressed in prostate cancer and prostatic intraepithelial neoplasia lesions, and it is positively associated with Gleason score. *Urologic Oncology: Seminars and Original Investigations*. 2014;**32**(1):53.e23-53.e29. DOI: 10.1016/j.urolonc.2013.08.028
- [4] Gurib-Fakim A. Medicinal plants: Traditions of yesterday and drugs of tomorrow. *Molecular Aspects of Medicine*. 2006;**27**(1):1-93. DOI: 10.1016/j.mam.2005.07.008
- [5] Savithramma N, Linga Rao M, Suhrulatha D. Screening of medicinal plants for secondary metabolites. *Middle-East Journal of Scientific Research*. 2011;**8**(3):579-584
- [6] Meléndez P, Capriles V. Antibacterial properties of tropical plants from Puerto Rico. *Phytomedicine*. 2006;**13**(4):272-276. DOI: 10.1016/j.phymed.2004.11.009
- [7] Mateeva N, Eyunni SVK, Redda KK, Ononuju U, Hansberry TD II, Aikens C, et al. Functional evaluation of synthetic flavonoids and chalcones for potential antiviral and anticancer properties. *Bioorganic & Medicinal Chemistry Letters*. 2017;**27**:2350-2356
- [8] Goietsenoven G, Andolfi A, Lallemand B, Cimmino A, Lamoral-Theys D, Gras T, et al. Amaryllidaceae alkaloids belonging to different structural subgroups display activity against apoptosis-resistant cancer cells. *Journal of Natural Products*. 2010;**73**(7):1223-1227. DOI: 10.1021/np9008255
- [9] Karla R-L, Cesar F-R, Jorge A, Julio R. Chemoprotective effect of the alkaloid extract of *Melocactus bellavistensis* against colon cancer induced in rats using 1,2-dimethylhydrazine. *Revista Peruana de Medicina Experimental y Salud Pública*. 2017;**34**:70-75
- [10] Bai N, Lai C, He K, Zhou Z, Zhang L, Quan Z, et al. Sesquiterpene lactones from *Inula britannica* and their cytotoxic and apoptotic effects on human cancer cell lines. *Journal of Natural Products*. 2006;**69**:531-535
- [11] Guerrero IC, Andrés LS, León LG, Machín RP, Padrón JM, Luis JG, et al. Abietane diterpenoids from *Salvia pachyphylla* and *S. clevelandii* with cytotoxic activity against human cancer cell lines. *Journal of Natural Products*. 2006;**69**:1803-1805
- [12] Wang T, Hojo T, Ran F, Wang R, Wang R, Chen H, et al. Cardenolides from *Saussurea stella* with cytotoxicity toward cancer cells. *Journal of Natural Products*. 2007;**70**:1429-1433
- [13] U.S. Agriculture. Genus *Buddleja* L. United States Department of Agriculture, Germplasm Resources Information Network; 2006
- [14] Ati HY, Gamal AA, Fawzy GA. Chemical composition, in vitro antimicrobial and cytotoxic activities of *Buddleja polystachya* essential oils. *Journal of Essential Oil-Bearing Plants*. 2014;**17**(6):1112-1119
- [15] Wu J, Yi W, Jin L, Hu D, Song B. Antiproliferative and cell

apoptosis-inducing activities of  
compounds from *Buddleja davidii* in  
Mgc-803 cells. *Cell Division*. 2012;7:20

[16] Cruz Paredes C, Bolívar Balbás P,  
Gómez-Velasco A, Juárez ZN, Sánchez  
Arreola E, Hernández LR, et al.  
Antimicrobial, antiparasitic, anti-  
inflammatory, and cytotoxic activities  
of *Lopezia racemosa*. *The Scientific  
World Journal*. 2013;2013. ID: 237438,  
6 pages. DOI: 10.1155/2013/237438





# Designing Magnetic Mesoporous Nanoparticles for Cancer Therapy

*Jessica Andrea Flood-Garibay, Kenneth J. Balkus Jr  
and Miguel Ángel Méndez-Rojas*

## Abstract

Cancer is the second most cause of mortality worldwide. The most common treatments are surgery, radiotherapy, and chemotherapy. Magnetic mesoporous nanoparticles (MMNPs) have attractive features such as high surface areas, large pore volumes, uniform and tunable pore sizes, high mechanical stability, and surface functionalization options for application as drug delivery systems. The latter make them a promising platform for the cancer treatment. Magnetic properties can be controlled by selecting the chemical nature and concentration of the magnetic materials to be embedded into the porous structure. These magnetic composites may be guided to allow precise targeting of a tumor using an external magnetic field. The mesoporous structure can also be loaded with different types of therapeutic agents, radiotracers, or fluorescent markers. Doping of the magnetic nanocomposite with rare earth elements may generate novel composites with physical properties useful for medical imaging or radiotherapy. The MMNPs could generate hyperthermia temperatures when exposed to an alternate-magnetic field (AMF). Many promising anticancer drugs have poor solubility, a problem that can be solved by using the MMNPs as nanocarriers, improving the bioavailability of the drugs. These MMNPs could become a promising multifunctional platform for the design of chemotherapeutic, medical imaging, drug delivery, and hyperthermia agents for cancer treatment.

**Keywords:** mesoporous, magnetic, nanoparticles, drug delivery system, theranostic

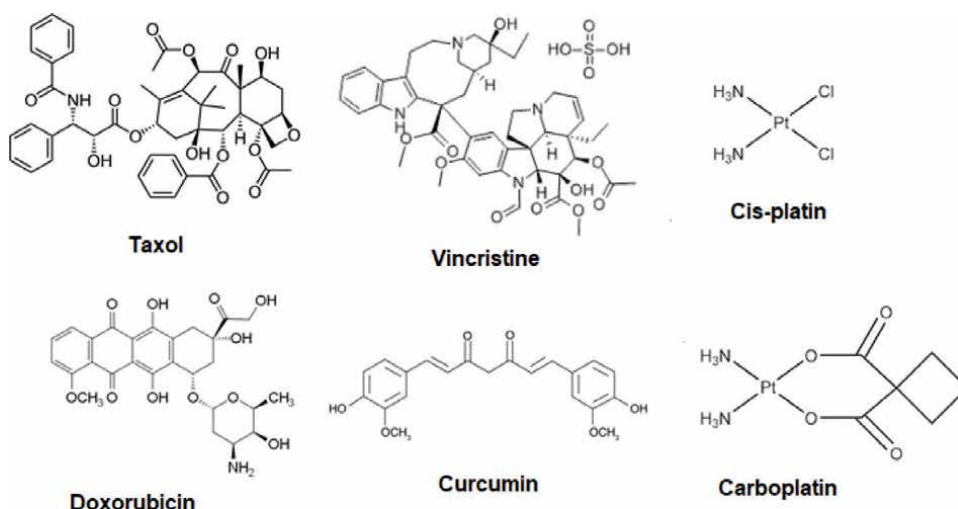
## 1. Introduction

Although great advances in the treatment and cure of several public health issues have been developed in the last decades, cancer is still a major burden worldwide. Cancer has been the second or third leading cause of death in both the United States and Mexico over the last decade [1]. Tens of millions of people are diagnosed with cancer every year, and it is considered the main cause of death globally. In the USA, there were more than 1,700,000 new cases of cancer diagnosed in 2018, with nearly 600,000 people dying from the disease, while in Mexico it was projected that nearly 1,200,000 cancer cases will be diagnosed in the next few years. Lung cancer is the leading cause of death in the US and Mexico, and this is expected to increase in coming years. Cancer therapies include surgery, chemotherapy, and/or radiotherapy. For some types of cancer, there is also the possibility of specific targeted therapy.

Chemotherapy involves the use of nonspecific cytotoxic compounds toward cancer cells, which is why they usually have multiple serious side effects in patients [2]. In order to decrease side effects, improve bioavailability and have a selective release to tumor cells, intelligent drug delivery systems (DDS) are being developed. It is important to understand that DDS should avoid high nonspecific accumulation in tissues [3]. In addition, it is important that the material of the drug carrier should be biocompatible. Furthermore, a sufficient dose of API (active pharmaceutical ingredient) should be loaded into the system and the release of the drug should be achieved without premature leakage. That way, the API could be delivered to the target site in a controlled manner; maintaining an adequate release rate in order to achieve an effective local concentration of the drug [4].

The development of new nanomaterials for biomedical applications is a rapidly growing area of research. The use of nanoparticles (NPs) as drug carriers may present different advantages, such as protecting the drug from degradation, reducing renal clearance, as well as allowing specific bioaccumulation in cancerous tumors due to improved permeation and retention effect (EPR). Magnetic nanoparticles (MNPs), in particular, iron-based ferrites are highly attractive as their magnetic properties can be easily tuned by controlling the type and ratio of metal ion substituents. Many of them have been found to be highly stable, even at physiological conditions, as well as biocompatible. Their small size may allow them to pass through several biological barriers, increasing their systemic circulation and enhancing biodistribution. Rare-earth ions can be embedded into the crystal lattice, making possible their transmutation into beta or gamma emitters by neutron activation. Also, the large surface of the mesoporous material can be used for the immobilization of different types of fluorescent dyes or biomarkers that improve both traceability and molecular recognition specificity. Localizing with precision a tumor site, either using a radiation detector or the luminescence of the nanomaterial, could be of great value for targeted delivery, helping to minimize the amount of radiation or the chemotherapeutic agent that the patient receives, thus reducing the undesirable side effects. There are several examples of nanomaterials used to deliver radionuclides *in vivo* [5]. However, controlling size to achieve an enhanced permeability and retention effect (EPR) as well as functionalization and targeted delivery remain challenges. The incorporation of radioactive isotopes into the spinel crystal structure of magnetic ferrites is a good option to achieve that goal, without compromising the size, biocompatibility, stability, or magnetic properties of the proposed nanomaterials. Another strategy could be doping the mesoporous structure around the MNPs with the radioisotope ions. That may be achieved either by adding the radioisotope-containing metal salts during the mesoporous phase synthesis or by dispersing and trapping the ions into the mesoporous structure once the material is formed. The high surface area or the mesoporous structure, depending on the choice of chemical composition and crystalline phase, may present the advantage to be easily functionalized with either radiosensitizers, fluorescent dyes, and/or to trap into the mesoporous structure different types of chemotherapeutic agents to further reduce the amount of radiation required to eliminate a tumor. In particular, the chances to improve bioavailability and aqueous dispersibility of low soluble chemotherapeutic agents make these magnetic mesoporous composites of great value for the transport and delivery of several promising anticancer drugs that have poor water solubilities, such as taxanes (paclitaxel, docetaxel), platinum-based drugs, curcumin, and many others (**Figure 1**). This is important, as poorly water-soluble drugs usually require the use of a high concentration of surfactants and co-solvents, or the administration of doses of the drug for longer periods, leading to adverse side effects [6].

Therefore, the development of new strategies for the treatment of this disease is urgently needed. The development of functionalized nanoparticles for both



**Figure 1.**  
Examples of cytotoxic agents used for cancer chemotherapy that present low solubility and, therefore, bioavailability problems.

medical imaging, diagnosis, chemo- and radio-therapeutic therapies depends in part on effective tumor targeting. Conventional approaches using tumor binding ligands have been effective in cell cultures but have been disappointing *in vivo*. Nonconventional targeting, such as magnetic nanoparticles (MNPs), are promising but in the early stages of development. The preparation of magnetic nanoparticles is a very attractive and active research field. In addition to advanced clinical treatments in modern anticancer therapies, MNPs can be used in several other practical applications such as biomarkers, magnetic storage, biomolecule separation, sensors, and medical imaging contrast agents. In particular, superparamagnetic iron oxide nanoparticles (SPIONs) offer high biocompatibility than other MNPs such as maghemite and have been widely used in several biomedical applications. Although some biocompatible, nanostructured MNPs with excellent stability, improved magnetic properties, and good biodistribution have received approval for clinical use, such as Feridex®, Resovist®, Sinerem®, Clariscan®, and Lumirem®, they are currently discontinued for biomedical use as MRI agents due to potential harmful side effects following administration [7]. However, their potential use as therapeutic agents may still make these materials clinically viable agents, as less MNPs would be required, compared to MRI use, reducing potential side effects; carefully checking of toxicity and biocompatibility is a must for these magnetic materials in order to look for real clinical applications. MNPs require them to be superparamagnetic in order to avoid spontaneous aggregation *in vivo* while they move through systemic circulation through the body. Aside from the potential use MNPs as MRI contrast agents, they can be used for drug transport and delivery, as well as for magnetic heat generation (hyperthermia). The advantages of MNPs in nanoscale delivery systems are numerous—drug delivery can be enhanced, increasing the biodistribution of the nanocarrier by avoiding clearance due to their small size and stability in physiological conditions. MNPs can be chemically modified in their surfaces by attaching functional molecules, such as proteins, antibodies, peptides, or sugars, in order to enhance bioselectivity and achieve fine-tuned drug delivery and bioaccumulation in specific targets, in particular in tumor tissues [8].

The magnetic response of MNPs can be controlled by transition metal ion substitution in the crystal lattice, a strategy highly exploited for the preparation of numerous magnetic ferrites with spinel structure [9]. Substitution using transition

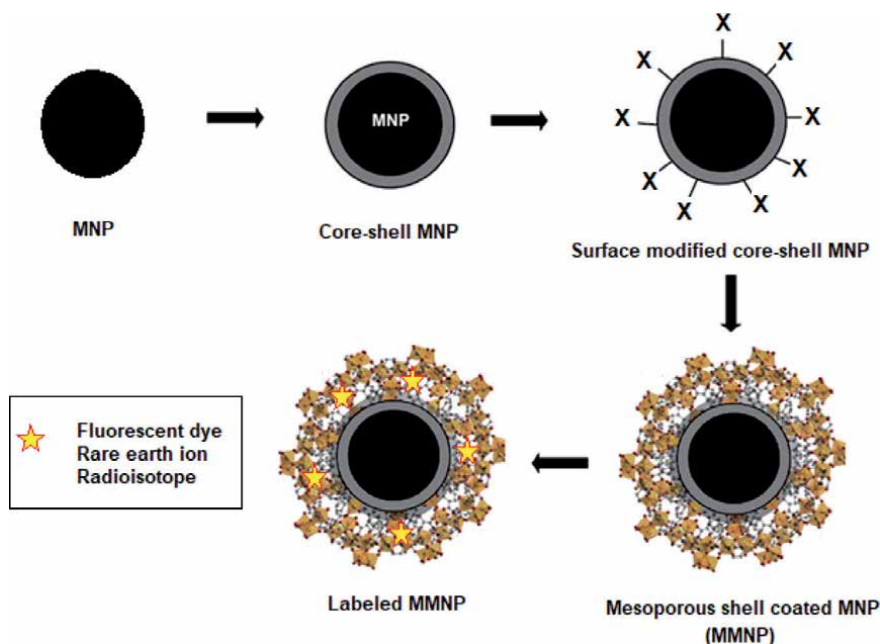
metal and rare-earth elements is an active field of research, looking to enhance saturation magnetization ( $M_s$ ), permittivity, permeability, and blocking temperature ( $T_B$ ). Several works available in the scientific literature report the design of small MNPs with controlled magnetic properties and low dispersion, with sizes less than 35 nm, by the formation of core-shell structures using the co-precipitation method [10, 11].

As a proof of concept of this idea, iron oxide nanoparticles containing Ho(III) were neutron activated and injected into athymic nude mice having tumors of non-small cell lung cancer (NSCLC) A549 cells [5]. A 12,000 Gauss magnet was placed on the tumor for 4 hours to allow the Ho-doped magnetic nanomaterial to collect in the tumor. There was a statistically significant reduction in the tumor size after 30 days and a 10-fold increase in Ho accumulation in the tumor with the magnet. While these results were promising, the Ho-doped magnetic nanomaterial presented several problems including the difficulty to functionalize the surface, as well as their relatively large size, which may lower the chances for cell internalization and efficient biodistribution. Furthermore, the low-intensity magnetic properties may not be appropriate to reach tumors below the surface. The ability to functionalize the surface of the MNP allows for the introduction of radiosensitizers and chemotherapeutic drugs as well as promote the suspension of the MNPs. The size is important to achieve the enhanced permeability and retention (EPR) effect for tumor penetration. Finally, the magnetic properties are important because treatment of certain cancers such as lung cancer may require the MNPs to be directed by a magnet several centimeters away.

## 2. Methods of preparation

Magnetic nanoparticles can be prepared easily by co-precipitation in alkaline aqueous media. Aqueous preparation is preferable to obtain products meant to be used in biomedical applications. Ferrite nanoparticles, both pristine or doped with rare-earth ions can be prepared by the addition of the corresponding Fe(III), Fe(II), and rare-earth salt precursors in the oxidation state (III) [X(III): Ce, Pr, Nd, Pm, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu], in an appropriate stoichiometry that allows to control the total amount of the rare-earth ions incorporated into the spinel structure, as previously reported [7, 10, 11]. Several rare-earth salts are available as nitrates, halides, or oxides from several commercial sources. Preparation of the rare-earth-doped magnetic nanoparticles can also follow several other modified synthetic procedures reported in the literature [9, 12]. Different types of magnetic ferrites, with a  $M^{+2}X^{+3}_xFe^{+3}_{2-x}O_4$  stoichiometry, where  $M = Zn, Co, Ni, Mn, Cu$ , can be prepared by selecting the proper amounts of the metal salt precursors. From previous works, superparamagnetic and biocompatible MNPs with strict control in size can be produced by this synthetic methodology (from 8 to 20 nm) [10, 11]. MNPs produced under these conditions are nearly monodisperse (10–15 nm), with zeta potential values higher than  $-30$  mV, low blocking temperatures ( $T_B$ ), and high magnetic saturation ( $M_s$ ), which make them small enough for internalization into tumors, stable, water-soluble and highly responsive to external magnetic fields, and suitable for biomedical applications. We have also explored recently how the incorporation of rare-earth metals induces not only structural changes but also impacts the magnetic properties, so the novel Ho-containing MNPs will have controlled magnetic and size properties [13, 14].

The surface of the magnetic nanoparticles can be easily modified by coating it with a layer of conveniently selected mesoporous materials ( $SiO_2$ , carbon,  $ZnO$ ...). Coating MNPs with a thin layer of silica could be used to produce core-shell

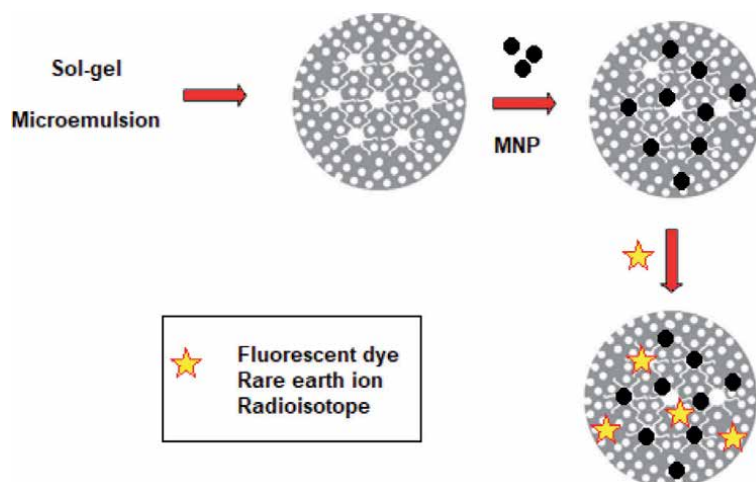


**Figure 2.**  
Schematic representation of the process for the preparation of core-shell MMNPs.

nanoparticles with an active surface that can be easily modified and derivatized (Figure 2). Once the surface of the MNPs is modified, the functional groups present in the surface could be used to grow another layer mesoporous layer, in order to increase the internal surface required for drug loading or it can be used to attach different chemical functionalities such as bioactive molecules (peptides, amino acids, antibodies, sugars), fluorescent dyes or to immobilize rare-earth ions, useful for radiotherapy or medical imaging.

A second approach for the preparation of MMNPs is to embed the magnetic nanoparticles into them by seeding them during the formation of the mesoporous structure (Figure 3). The magnetic nanoparticles could also be trapped into the voids of the mesoporous structure by sonication, stirring, or simple mixing, depending on the affinity among the materials.

After preparation and purification, the products obtained from any of these strategies can be characterized using several analytical techniques such as Fourier transform infrared (FT-IR) spectroscopy, Raman spectroscopy, fluorescence spectroscopy, dynamic light scattering (DLS), thermogravimetric analysis (TGA), powder X-ray diffraction (pXRD), magnetometry, energy dispersive spectroscopy (EDS), BET surface area analysis, and scanning and transmission electron microscopy (SEM and TEM, respectively). Once the magnetic mesoporous nanoparticles have been fully characterized, *in vitro* test of the MMNPs, can be performed using a panel of different cell lines (normal and cancer cells), in order to evaluate their biological activity. There are several methods to determine cell viability, such as the MTT viability assay, which is a quantitative colorimetric assay based on the conversion of MTT to formazan crystals by mitochondrial dehydrogenase. *In vivo* testing in small animal models may give further information on the effectiveness and performance of these MMNPs for cancer treatment, as well as on the toxicology of the nanomaterials. Morphological changes such as cell shrinkage, membrane blebbing, apoptotic body formation, cytoplasmic swelling, and cytopathic effect in cells treated with MMNPs, may be also useful to understand better the mechanisms of the biological interaction among the MMNPs



**Figure 3.** Schematic representation of the process for preparation of MMNPs where the magnetic nanoparticles were trapped into the mesoporous structure.

and the cells. Epifluorescence microscopy analysis of the cell cultures, using differentially stained wells with different types of dyes, may also be useful to understand the mechanisms of internalization and cell death.

### 3. Conclusions

The design, synthesis, and characterization of MMNPs systems, with optimal characteristics to be stable, water-soluble, biocompatible, with good size control and distribution is a promising field for the design of innovative nanoplatforms for cancer therapy. NPs with sizes lower than 100 nm, and optimal size distribution, are more easily dispersed in physiological aqueous suspensions, allowing the nanoparticles to be bioavailable and facilitating cell internalization through endocytosis or pinocytosis. Loading of the MMNPs systems with poorly soluble anticancer drugs into the mesoporous structure, and not on the surface of the nanoparticles, may be useful to improve the transport and bioavailability of these therapeutic agents, increasing their performance and lowering their side effects. Preliminary studies in our group showed that silica-based MMNPs are biocompatible, as no impact on cell viability was observed even at high concentrations of the mesoporous material. When the chemotherapeutic agent was loaded into the MMNPs, testing showed that cell viability was affected even at when low concentrations were loaded into the nanocarrier. Comparison with cell cultures exposed to the free anticancer drug showed lower antiproliferation activity with respect to that of the drug-loaded WMS nanoparticles, indicating an enhancement of bioavailability for the chemotherapeutic agent under the conditions of this study. These preliminary results are stimulating and suggest that MMNPs could become an effective alternative for the treatment of certain types of cancer.

### Acknowledgements

Financial support from ConTex-CONACYT (2019-21B) and CONACYT (JAFG, Ph.D. Scholarship) is acknowledged.

## Author details


Jessica Andrea Flood-Garibay<sup>1</sup>, Kenneth J. Balkus Jr <sup>2</sup> and Miguel Ángel Méndez-Rojas<sup>1\*</sup>

<sup>1</sup> Departamento de Ciencias Químico-Biológicas, Escuela de Ciencias, Universidad de las Américas Puebla, Puebla, Mexico

<sup>2</sup> Department of Chemistry and Biochemistry, University of Texas at Dallas, USA

\*Address all correspondence to: [miguela.mendez@udlap.mx](mailto:miguela.mendez@udlap.mx)

## IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Mohar-Betancourt A, Reynoso-Noveron N, Armas-Texta D, Gutierrez-Delgado C, Tores-Dominguez. Cancer trends in Mexico: Essential data for the creation and follow up of public policies. *Journal of Global Oncology*. 2017;**3**:740-747
- [2] American Cancer Society. *Chemotherapy Side Effects*. 2019. Available from: <https://www.cancer.org/treatment/treatments-and-side-effects/treatment-types/chemotherapy/chemotherapy-side-effects.html> [Accessed: 21 April, 2019]
- [3] Sufi SA, Pajaniradje S, Mukherjee V, Rajagopalan R. Redox nano-architectures: Perspectives and implications in diagnosis and treatment of human diseases. *Antioxidants & Redox Signaling*. 2018;**30**(5):762-785
- [4] Sun R, Wang W, Wen Y, Zhang X. Recent advances on mesoporous silica nanoparticles-based controlled release system: intelligent switches open up new horizon. *Nanomaterials (Basel)*. 2015;**5**(4):2019-2053
- [5] Munaweera I, Shi Y, Koneru B, Saez R, Aliev A, Di Pasqua AJ, et al. Chemoradiotherapeutic magnetic nanoparticles for targeted treatment of nonsmall cell lung cancer. *Molecular Pharmaceutics*. 2015;**12**:3588-3596
- [6] Narvekar M, Xue HY, Eoh JY, Wong HL. Nanocarrier for poorly water-soluble anticancer drugs-barriers of translation and solutions. *AAPS PharmSciTech*. 2014;**15**:822-833
- [7] Arteaga-Cardona F, Gutiérrez-García E, Hidalgo-Tobón S, López-Vásquez C, Brito-Barrera YA, Flores-Tochihuitl J, et al. Cell viability and MRI performance of highly efficient polyol-coated magnetic nanoparticles. *Journal of Nanoparticle Research*. 2016;**18**:345
- [8] Angulo-Molina A, Méndez-Rojas MA, Palacios-Hernández T, Contreras-López OE, Hirata-Flores GA, Flores-Alonso JC, et al. Magnetite nanoparticles functionalized with  $\alpha$ -tocopheryl succinate ( $\alpha$ -TOS) promote selective cervical cancer cell death. *Journal of Nanoparticle Research*. 2014;**16**:2528
- [9] Ali I, Islam MU, Ishaque M, Khan HM, Ashiq MN, Rana MU. Structural and magnetic properties of holmium substituted cobalt ferrites synthesized by chemical co-precipitation method. *Journal of Magnetism and Magnetic Materials*. 2012;**324**:3773-3777
- [10] Arteaga-Cardona F, Santillán-Urquiza E, de la Presa P, Hidalgo-Tobón S, Pal U, Horta-Fraijo P, et al. Enhanced magnetic properties and MRI performance of bi-magnetic core-shell nanoparticles. *RSC Advances*. 2016;**6**:77558
- [11] Arteaga-Cardona F, Santillán-Urquiza E, Pal U, Mendoza-Alvarez ME, Torres-Duarte C, Cherr GN, et al. Unusual variations of blocking temperature in bi-magnetic nanoparticles. *Journal of Magnetism and Magnetic Materials*. 2017;**441**:417-423
- [12] Bloemen M, Vandendriessche S, Goovaerts V, Brulloy W, Vanbel M, Carron S, et al. Synthesis and characterization of holmium-doped iron oxide nanoparticles. *Materials*. 2014;**7**:1155-1164
- [13] Arteaga-Cardona F, Pal U, Alonso JM, de la Presa P, Mendoza ME, Salazar-Kuri U, et al. Modifying magnetic and crystallographic properties in nanostructured  $\text{MnFe}_2\text{O}_4$  by systematically introducing different  $\text{M}^{2+}$  ions. *Journal of Magnetism and Magnetic Materials*. 2019;**489**:165496



[14] Arteaga-Cardona F, Martha-Aguilar NG, Estevez JO, Pal U, Méndez-Rojas MA, Salazar-Kuri U. Variations in magnetic properties caused by size size dispersión and particle aggregation on  $\text{CoFe}_2\text{O}_4$ . SN Applied Sciences. 2019;1:412



# Exoplanet Research Using Machine Learning and Multiresolution Analysis Techniques

*Miguel Jara-Maldonado, Vicente Alarcon-Aquino  
and Roberto Rosas-Romero*

## Abstract

The research of planets from outside our Solar System, termed exoplanets, has opened a wide range of new possibilities. Some of the current interests in exoplanet research are related to their discovery and the characterization of their atmospheres. Finding these planets is important because it may lead to answering several questions; such as the formation of planets and stellar systems, and possibly finding life outside planet Earth. There are several works that propose using artificial intelligence to ease the processes involved in exoplanet research. Many studies have focused on the detection of such celestial bodies, as well as reducing the number of false detections. Recently, the study of exoplanet atmospheres has also received considerable attention, due to its potential for finding life on these planets. In this work, we describe an artificial intelligence approach for reducing the number of spurious detections of exoplanets using the transit technique. This approach is based on using spectral multiresolution analysis techniques, which allow the artificial intelligence algorithms to better identify the exoplanet signals.

**Keywords:** artificial intelligence, deep learning, exoplanets, light curves, machine learning, multiresolution analysis, neural networks

## 1. Introduction

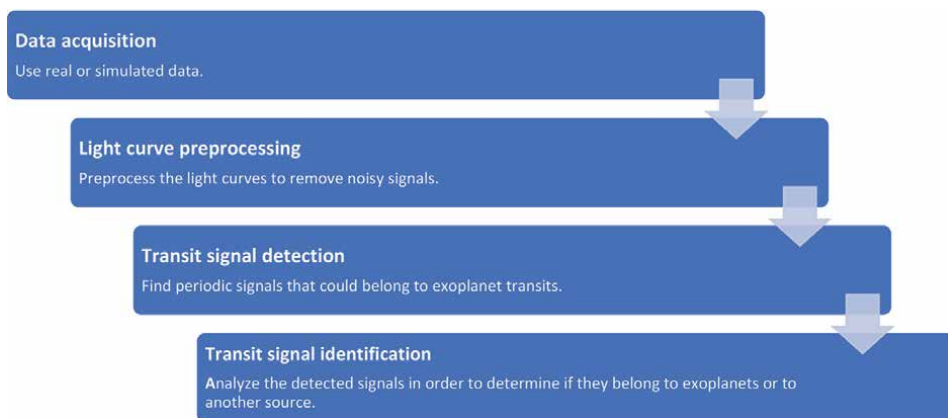
The term exoplanet is an abbreviation for extra-solar planets. Exoplanets are planets found outside our Solar System, either orbiting a star or not. Their study is important for several reasons, such as obtaining statistical information about planets, which in turn allows us to extend our understanding of how our Solar System was created. One of the reasons to study exoplanets is to look for habitable planets found outside the Solar System; which could lead to finding life outside planet Earth (although no evidence of life has yet been found in exoplanet atmospheres) [1]. In order to search for exoplanets, several missions have been launched. The Kepler [2, 3], Convection, Rotation and Planetary Transits space observatory (*CoRoT*) [4], and (*TESS*) [5] missions, are some examples of those missions.

In order to look for exoplanets, astronomers have developed different detection techniques. Among the most used are the transit method, radial velocity, gravitational microlensing, direct imaging, and others. In this work, we focus on the transit method. This method looks for transits, which occur when an exoplanet passes between the observer and its host star. To look for transits, scientists use light curves, which are records of the light flux received by the star at different moments in time. When an exoplanet transits its star, a reduction of the light flux characterized by a “U” or “V” shape is observed. This technique has provided the greatest number of exoplanet discoveries. But this technique is not infallible, and it is sensitive to noise sources that may look like transits or that hide the transit signal. In order to deal with these and other difficulties (see [1]), several artificial intelligence algorithms such as [6–12] have been developed. These approaches have the aim of ameliorating the detection and identification accuracy of exoplanet transit signals within the light curves.

In this work, we summarize the work done in [1, 13], where simulated light curves are used to test the performance of artificial intelligence and multiresolution analysis techniques for exoplanet identification.

## 2. Methodology

Automating the exoplanet discovery process requires a pipeline that describes clear instructions for the artificial intelligence algorithms to work with. We have proposed a data pipeline in [1], that establishes the whole process of exoplanet discovery with artificial intelligence. This pipeline is shown in **Figure 1**. The data acquisition step refers to the process of obtaining the light curves to work with. These light curves may be obtained by real telescopes (such as the *Kepler* satellite), or by simulating them. The light curves contain different sources that difficult their analysis. For this reason, the next step is to preprocess the light curves in order to reduce the influence of noise in the light curves. With the transit signals already enhanced, the detection step may be performed by an artificial intelligence algorithm, to search for periodic signals within the light curves that could be explained by an exoplanet. Finally, it is required to analyze the periodic signals found, to make sure that they belong to an exoplanet, and not to an event of similar geometry. In the remaining of this section, we explain how we applied this pipeline to simulated light curves generated by us, to identify exoplanet signals.



**Figure 1.**  
*Proposed pipeline for exoplanet discovery.*

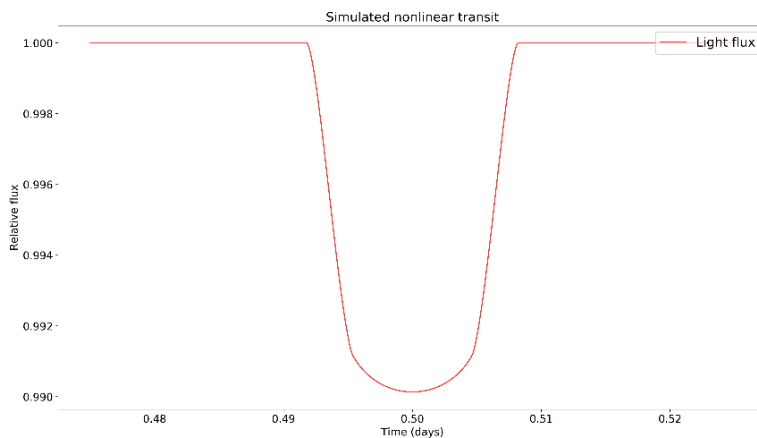
## 2.1 Light curve datasets creation

We generated two simulated datasets consisting of 10,000 light curves each. For each dataset, half of the light curves contain simulated transits and the other half does not. Each light curve contains 15,000 datapoints. These datasets can be used to train and test machine learning algorithms for exoplanet identification with controlled, though realistic, noise sources. The presented work considers four different types of transit models. Furthermore, we explain the light curve preprocessing methodology that has been used by several works such as [6, 7, 14]. The first dataset, which is called the Real-LC dataset was generated using real light curves from the Mikulski Archive for Space Telescopes (*MAST*<sup>1</sup>) with periodic events marked as non-transiting planets and then adding simulated transits to them. The second one is called the 3-median dataset, and it was created by simulating the light curves, and then adding the simulated transits. Next, it is described how the light curves were simulated.

There are several models that can be used to generate simulated transit light curves. Some examples of these models can be found in [15–19]. We used the BASIC Transit Model cAlculationN (*BATMAN*) model proposed in [15], which is a python package based on several models such as [16, 20], and others. We selected this model because it uses the model proposed by [8], and it allows one to model light curves very fast. Also, it can be parallelized with the use of OpenMP (in case it was necessary to produce a greater number of samples), and it includes a wide variety of limb darkening models including the uniform, linear, quadratic, and nonlinear models which we used. Even more, it can generate secondary eclipses which are useful for accounting for these astrophysical false positive phenomena. An example of a simulated transit is presented in **Figure 2**, which was generated using the *BATMAN* nonlinear model.

In order to add noise, we used Eqs. (1)–(4) [7]. The generated noise adds quasi-periodic systematic trends to the simulated transit data.

$$t' = t - t_{min} \quad (1)$$



**Figure 2.**  
Example of a simulated transit light curve using the *BATMAN* nonlinear model.

<sup>1</sup> Mikulski Archive for Space Telescopes (*MAST*): <https://archive.stsci.edu/>

$$A(t') = A + A \sin\left(\frac{2\pi t'}{P_A}\right) \quad (2)$$

$$\omega(t') = \omega + \omega \sin\left(\frac{2\pi t'}{P_\omega}\right) \quad (3)$$

$$F_{transit}(t) * N\left(\frac{R_p^2}{R_s^2} / \sigma_{tol}\right) * \left(1 + A(t') \sin\left(\frac{2\pi t'}{\omega(t')} + \phi\right)\right) \quad (4)$$

where  $F_{transit}(t)$  is the simulated transit signal created by using BATMAN,  $t$  is time,  $A$  is the amplitude of the stellar variability,  $\omega$  is the period of oscillation,  $\phi$  is the phase shift,  $R_p$  is the planet radius,  $R_s$  is the star radius,  $\sigma_{tol}$  is the noise parameter, and  $N$  is a Gaussian distribution to generate random numbers with a mean of 1 and standard deviation of  $\left(\frac{R_p^2}{R_s^2} / \sigma_{tol}\right)$  as explained in [7]. Each dataset contains two types of light curves, namely light curves containing a transit and light curves that do not contain a transit. Notice that for generating light curves without the transit signal, the  $F_{transit}(t)$  of Eq. (4) is omitted.

The parameters used to simulate the transits are presented in **Table 1**. These parameters were chosen from a list of 140 real exoplanets presented in the Q1-Q17 Kepler Data Release 24 [11], which were discovered using the transit method. In **Table 2**, the parameters used to simulate the noisy light curves are presented.

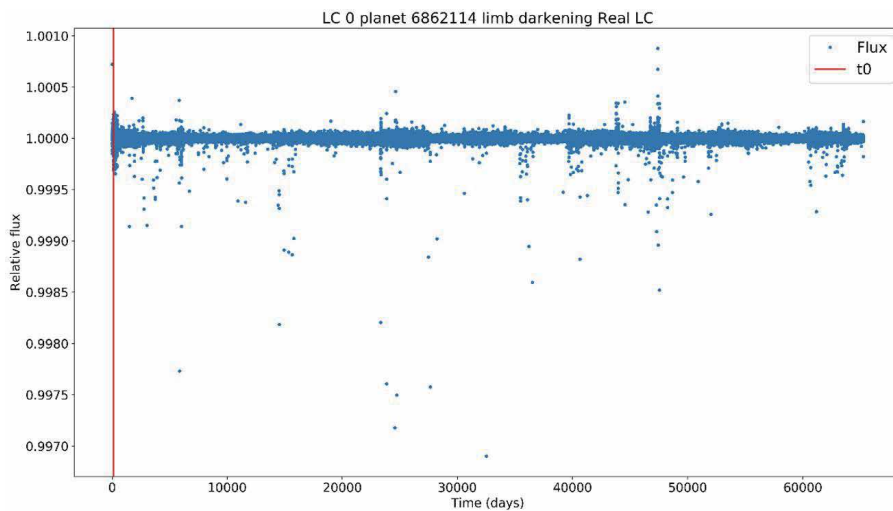
After simulating the light curves, they can be preprocessed in order to accentuate the transits and to reduce the noise sources. We used the spline fitting method proposed in [6] to preprocess the Real-LC light curves, and a 3-median filter was applied to the 3-median dataset. This process is also called flattening and it is

Fixed transit parameter	Values
Stellar radius ( $R_s$ )	0.12–2.59 Solar radii
Planet radius ( $R_p$ )	0.063–1.98 Jupiter radii
Scaled semi-major axis ( $\frac{a}{R_s}$ ), where $a$ is the semi-major axis	0.0058–0.2535 AU
Argument of periastron ( $\Omega$ )	90
Mid transit time ( $t_0$ )	75 days
Transit resolution	150 datapoints
Phase offset ( $\phi$ )	0
Amplitude variability period ( $P_A$ )	100
Wave variability period ( $P_\omega$ )	100
Light curve length	15,000 datapoints
Limb darkening model	Uniform, linear, quadratic, nonlinear
Limb darkening coefficients ( $u_1, u_2, u_3, u_4$ )	[none], [0.5], [0.5, 0.1], [0.5, 0.1, 0.1, -0.1]
Transit duration	0.253–0.4113 days
Transit Depth	0.0085–3.23%
Orbit eccentricity ( $e$ )	0–0.53
Orbit inclination ( $i$ )	78.3–96.5 deg
Orbital period ( $P$ )	0.0253–46.69 days

**Table 1.**  
Simulated transit parameters.

Varying transit parameter	Values
Noise parameter ( $\sigma_{tot}$ )	0.25, 0.75, 1.25, 1.75, 2.25, 2.75, 3, 10
Wave amplitude ( $A$ )	0.025, 0.05, 0.1, 0.2
Wave period ( $\omega$ )	6/24, 12/24, 24/24
Period offset ( $\Phi$ )	0
Amplitude variability period ( $P_A$ )	-1, 1, 100
Wave variability period ( $P_A$ )	-3, 1, 100

**Table 2.**  
 Noisy light curve simulation parameters.



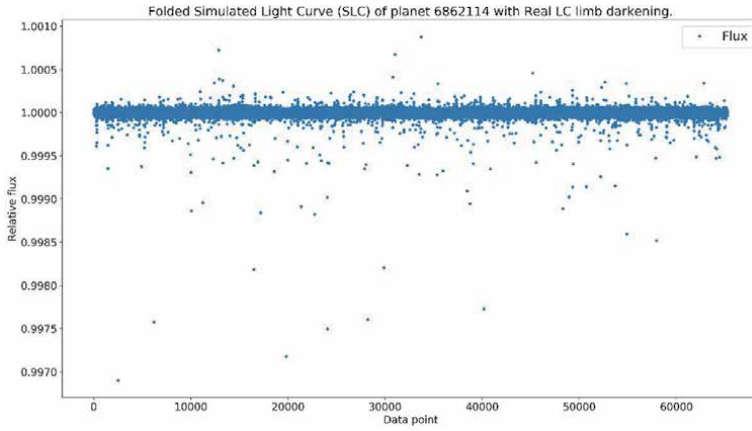
**Figure 3.**  
 Simulated light curve using synthetic noise and the BATMAN model.

performed to remove confusing data from the light curve. An example of a simulated light curve is presented in **Figure 3**. In this figure, each vertical blue line represents a transit, and the red line represents the mid-transit time from the first transit present in the light curve. This light curve consists of 15,000 simulated datapoints; which were added with a transit signal simulated using the BATMAN model.

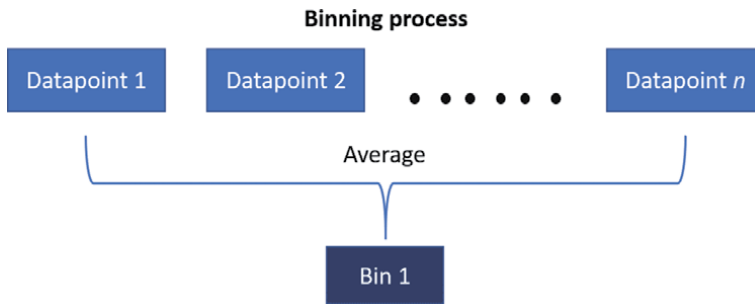
The next step is to phase fold the light curve to overlap all the points in the light using the transit event as the center. We used the *PyAstronomy* python package.<sup>2</sup> An example of a folded light curve is presented in **Figure 4**. There is a major dim in the light flux in the middle of the light curve, which corresponds to the transit. Also, there are other sources that could belong to another transit within the same light curve, although these are not centered because they do not correspond to the event that is being analyzed in this example.

Finally, the binning step allows one to reduce the dimensionality of the dataset by grouping the values in a limited number of bins. **Figure 5** explains the construction of one bin: the bins are created by calculating the mean of all the  $n$  points found inside a bin. We used 2048 bins; in other words, the length of the light curves is reduced from 15,000 to 2048 datapoints, and each bin is then represented by the

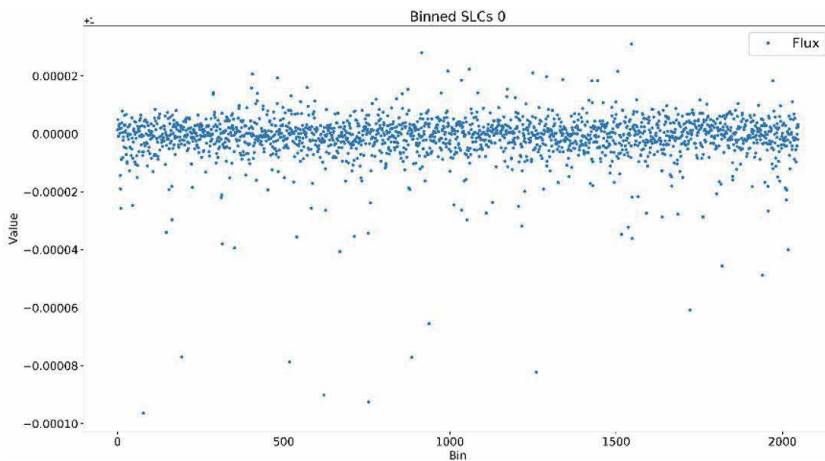
<sup>2</sup> *PyAstronomy* package: <https://www.hs.uni-hamburg.de/DE/Ins/Per/Czesla/PyA/PyA/index.html>



**Figure 4.**  
*Phase folded light curve.*



**Figure 5.**  
*Binned process.*



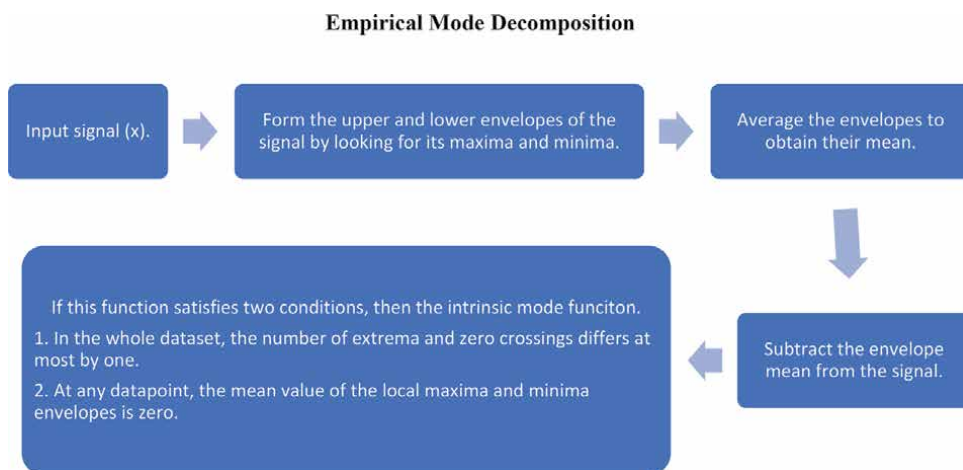
**Figure 6.**  
*Binned light curve.*

mean of all the values inside that bin. An example of a binned light curve can be seen in **Figure 6**, this is the same light curve as the one presented in **Figure 4**, with the difference that it is now binned.

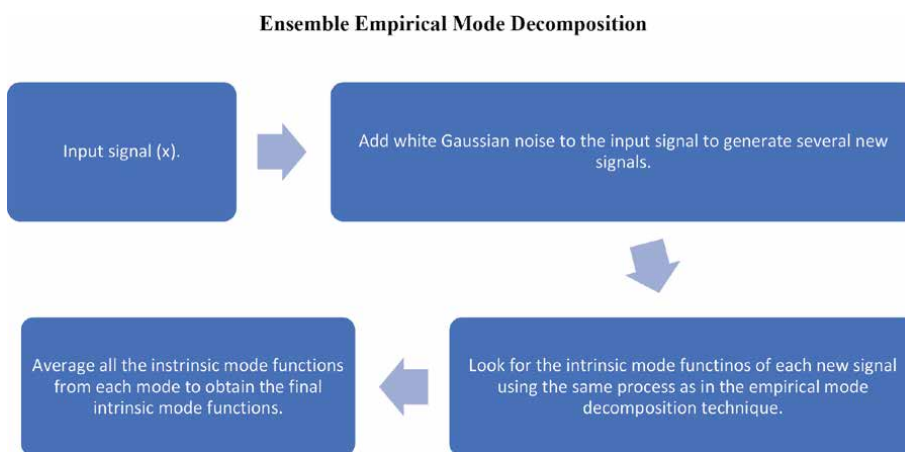


## 2.2 Transit signal identification

In our datasets, the possible transit signals have already been detected. To determine if these detections are real, we have used different machine learning models (i.e., artificial intelligence algorithms). Even more, we have occupied multiresolution analysis techniques to preprocess the light curves, and we have compared the performance of the machine learning models, using multiresolution analysis and without it. Multiresolution analysis techniques are used to obtain the different levels of resolution of a signal, in order to “look at it from different perspectives.” This process is similar to using a microscope to observe small objects, at different magnification levels different details of these objects will be visible. An example of such a technique is wavelets. Wavelets are functions that grow and decay over a finite time interval (they are short waves, hence their name, wavelet). By varying the translation and dilation parameters of the wavelet, it is possible to localize a function in both position and scale. The wavelets are convolved with the signal in order to determine how much does a section of the signal resemble the wavelet. The wavelet equation is shown in Eq. (5).



**Figure 7.**  
*Empirical mode decomposition technique.*



**Figure 8.**  
*Ensemble empirical mode decomposition technique.*

$$\psi_{\lambda,\tau}(u) = \frac{1}{\sqrt{\lambda}} \psi\left(\frac{u - \tau}{\lambda}\right) \tag{5}$$

where  $\psi(\cdot)$  is a function called mother wavelet, used to create several wavelets by varying the  $\lambda > 0$  dilation parameter and  $\tau$  translation parameter.

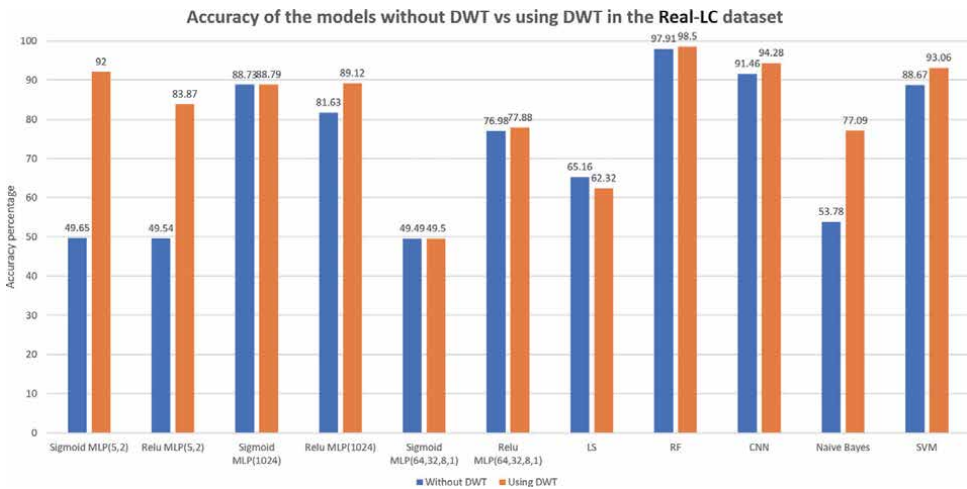
We have also used the empirical mode decomposition and ensemble empirical mode decomposition techniques. These multiresolution analysis techniques adaptively obtain intrinsic mode functions by iterating a process called *sifting*. In this process, the signal is separated into its different components. A description of these processes is shown in the diagrams from **Figures 7 and 8**. For a more detailed explanation of these techniques, refer to [13].

### 3. Results

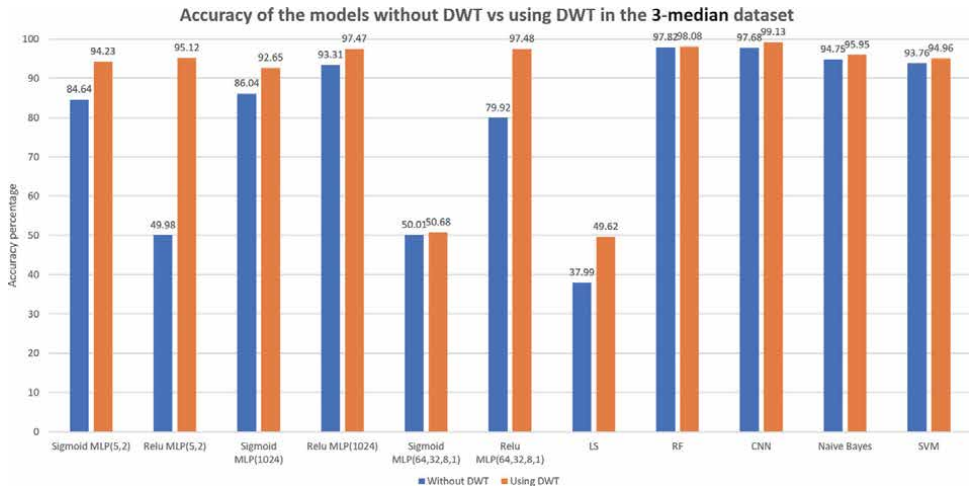
Several machine learning models were tested using these techniques to preprocess the light curves. The models tested were a convolutional neural network (CNN), different multilayer perceptron (MLP) architectures, least squares (LS), random forests (RF), Naïve Bayes, and a support vector machine (SVM) with the discrete wavelet transform. For the empirical mode decomposition and ensemble empirical mode decomposition techniques, we used a CNN, RF, K-nearest neighbors (*KNN*), and a Ridge classifier. Refer to [1, 13] for more details concerning these models and their configuration. In order to measure the performance of each model, we compared the models in terms of their accuracy and execution time. These metrics are based on the number of correctly classified exoplanets (true positives), and correctly classified nonexoplanets (false positives). The accuracy measures how many times the model was correct. The formula for this metric is presented in Eq. (6).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

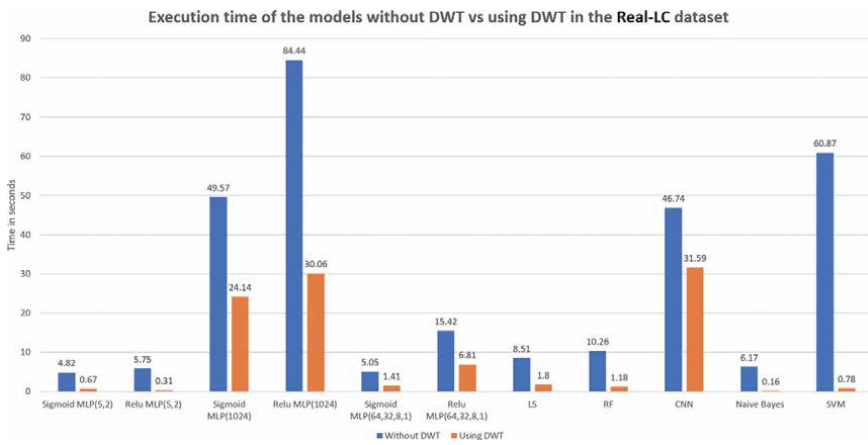
The accuracies obtained by the models that used the discrete wavelet transform with both datasets are presented in **Figures 9 and 10**, where the blue bars represent



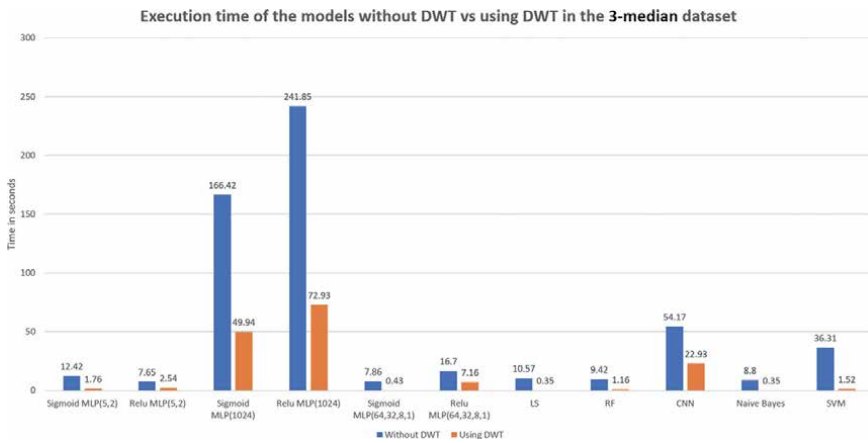
**Figure 9.** Accuracy results using the discrete wavelet transform in the Real-LC dataset.



**Figure 10.**  
 Accuracy results using the discrete wavelet transform in the 3-median dataset.



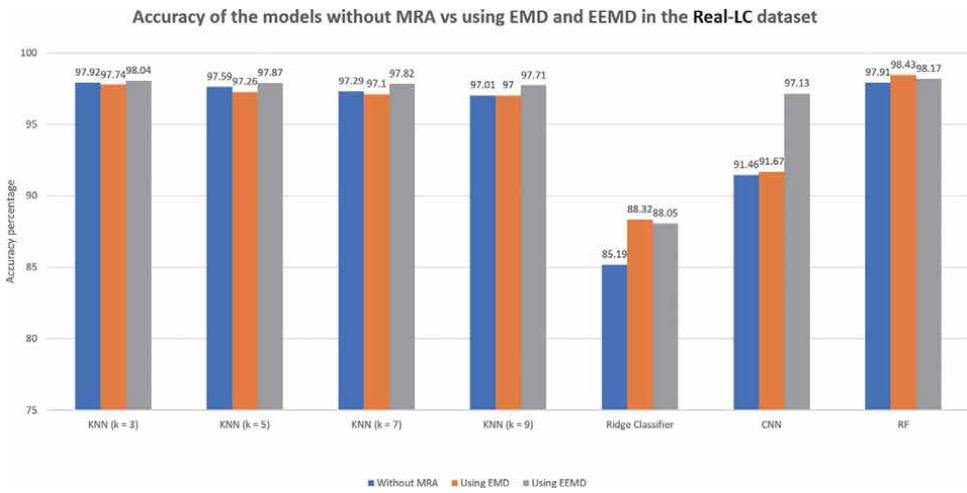
**Figure 11.**  
 Execution time results using the discrete wavelet transform in the Real-LC dataset.



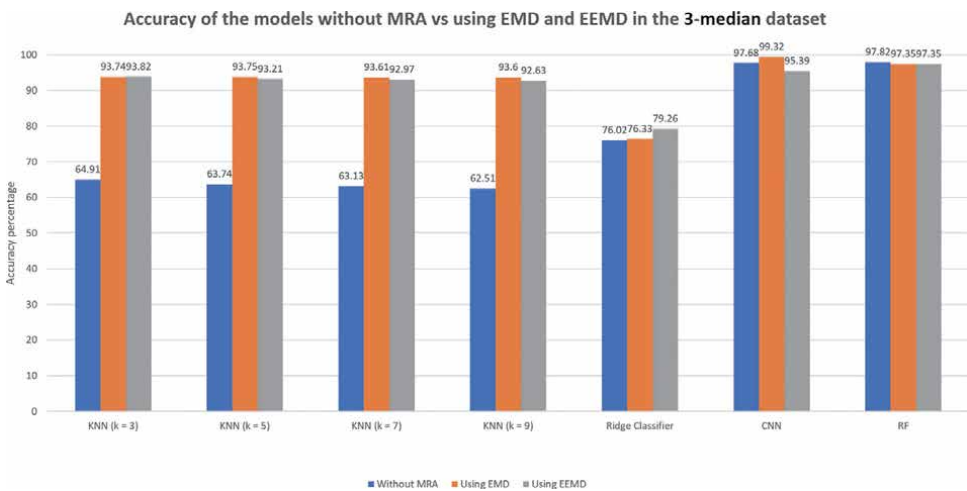
**Figure 12.**  
 Execution time results using the discrete wavelet transform in the 3-median dataset.

the results obtained without using the discrete wavelet transform, and the orange ones are the results obtained using it. It is noticeable that in most cases, the accuracy is increased, or at least it does not decrease. Then, in **Figures 11** and **12**, the execution time results are presented. As it can be seen, the execution times are always reduced, and this is due to the downsampling property of the discrete wavelet transform. At each level of resolution, the length of the signal is reduced by half.

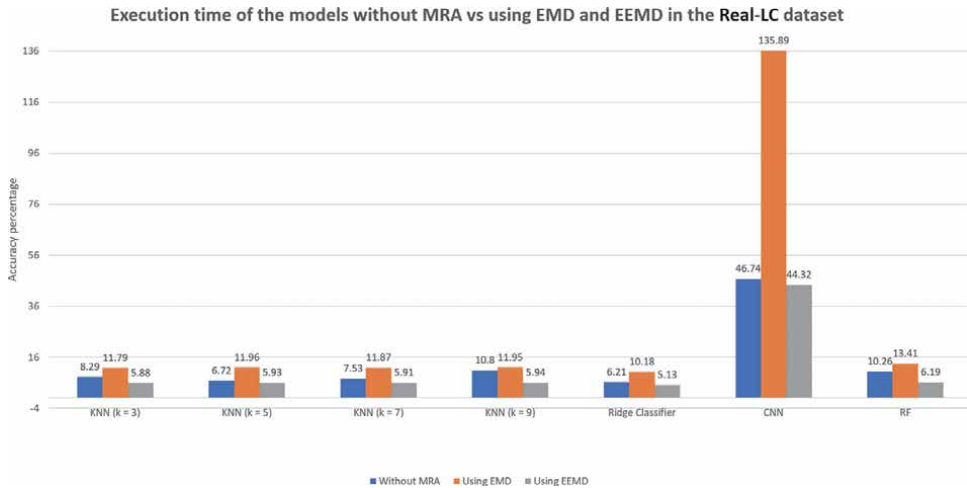
In **Figures 13** and **14**, the accuracy results of the empirical mode decomposition and its ensemble variant are presented. The blue bars, again, represent the signal without multiresolution preprocessing. The orange bars represent the results obtained using the empirical mode decomposition technique, and the gray bars represent the results obtained using the ensemble empirical mode decomposition



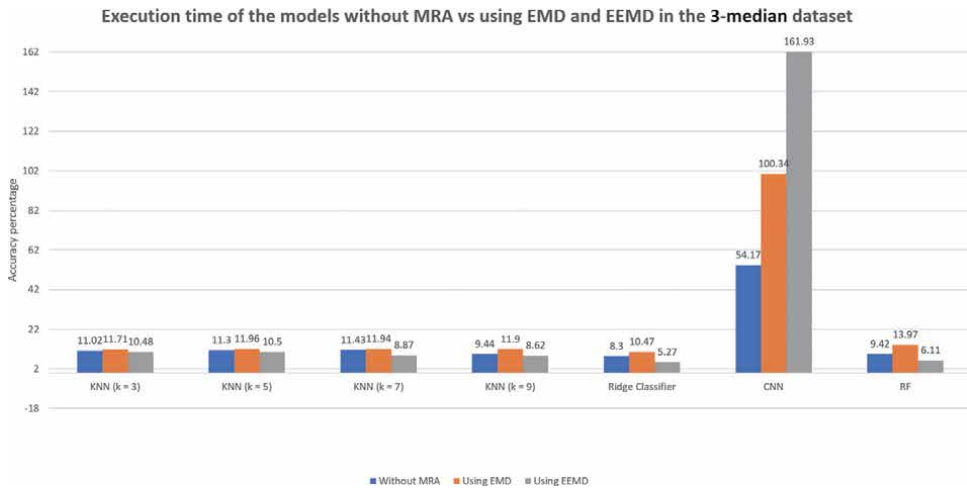
**Figure 13.** Accuracy results using the empirical mode decomposition and ensemble empirical mode decomposition techniques in the Real-LC dataset.



**Figure 14.** Accuracy results using the empirical mode decomposition and ensemble empirical mode decomposition techniques in the 3-median dataset.



**Figure 15.**  
 Execution time results using the empirical mode decomposition and ensemble empirical mode decomposition techniques in the Real-LC dataset.



**Figure 16.**  
 Execution time results using the empirical mode decomposition and ensemble empirical mode decomposition techniques in the 3-median dataset.

technique. Finally, **Figures 15** and **16** show the execution times for these techniques. These figures demonstrate that in most cases, using these techniques increase the performance of the identification models, both in time and accuracy. The only case in which the execution time is severely affected by these techniques is with the CNN model. We attribute this to the fact that the data obtained several decimal positions after the sifting processes.

#### 4. Conclusions and future work

The huge amounts of data, faced when analyzing transiting exoplanet light curves, have encouraged data scientists to develop machine learning models capable of automatically identifying exoplanets. These models can reduce the time spent

eyeballing the light curves while enhancing the identification accuracy. For such algorithms to exist, simulated light curves are necessary, because they provide a wide variety of labeled scenarios that can be used to train the models. For this reason, in this work, we presented the methodology followed to create two datasets of simulated light curves with different parameters, labeled as transit and nontransit signals. These light curves were used to train machine learning algorithms, and later test them. Once that the results obtained with the simulated data are satisfying enough, real data can be used to identify transiting exoplanets and contribute to the existing catalogs of exoplanet discoveries. Furthermore, some useful preprocessing steps were explained in this work. They can be used with simulated or real data. Our results show that using the multiresolution analysis techniques to preprocess the light curves improves the identification rates of the machine learning models. Future work will be done in proposing a new machine learning model based on multiresolution analysis techniques, instead of using them to preprocess the light curves.

## **Acknowledgements**

The authors would like to acknowledge the Mexican National Council on Science and Technology (*CONACyT*) and the Universidad de las Américas Puebla (*UDLAP*) for their support through the doctoral scholarship program. Also, the authors would like to thank Kyle A. Pearson for his feedback regarding the light curve preprocessing steps.

## **Author details**

Miguel Jara-Maldonado, Vicente Alarcon-Aquino\* and Roberto Rosas-Romero  
Department of Computing, Electronics and Mechatronics, Universidad de las Américas Puebla, Puebla, Mexico

\*Address all correspondence to: [vicente.alarcon@udlap.mx](mailto:vicente.alarcon@udlap.mx)

## **IntechOpen**

---

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Jara-Maldonado M, Alarcon-Aquino V, Rosas-Romero R, et al. Transiting exoplanet discovery using machine learning techniques: A survey. *Earth Science Informatics*. 2020;**13**:573-600. DOI: 10.1007/s12145-020-00464-7
- [2] Borucki WJ et al. Kepler planet-detection mission: Introduction and first results. *Science*. 2010;**327**(5968):977. DOI: 10.1126/science.1185402
- [3] Basri G, Borucki WJ, Koch D. The Kepler mission: A wide-field transit search for terrestrial planets [review article]. *New Astronomy Reviews*. 2005;**49**:478-485. DOI: 10.1016/j.newar.2005.08.026
- [4] Auvergne M et al. The CoRoT satellite in flight: Description and performance. *Astronomy and Astrophysics*. 2009;**506**:411-424. DOI: 10.1051/0004-6361/200810860
- [5] Ricker GR et al. The transiting exoplanet survey satellite. *Journal of Astronomical Telescopes, Instruments, and Systems*. 2014;**1**(1):014003. DOI: 10.1117/1.JATIS.1.1.014003
- [6] Shallue CJ, Vanderburg A. Identifying exoplanets with deep learning: A five planet resonant chain around Kepler-80 and an eighth planet around Kepler-90. *The Astronomical Journal*. 2018;**155**(2):94. DOI: 10.3847/1538-3881/aa9e09
- [7] Pearson KA, Palafox L, Griffith CA. Searching for exoplanets using artificial intelligence. *Monthly Notices of the Royal Astronomical Society*. 2018;**474**(1):478-491. DOI: 10.1093/mnras/stx2761
- [8] Kovács G, Zucker S, Mazeh T. A box-fitting algorithm in the search for periodic transits. *Astronomy and Astrophysics*. 2002;**391**(1):369-377. DOI: 10.1051/0004-6361:20020802
- [9] Armstrong DJ et al. Automatic vetting of planet candidates from ground based surveys: Machine learning with NGTS. *Monthly Notices of the Royal Astronomical Society*. 2018;**478**(3):4225-4237. DOI: 10.1093/mnras/sty1313
- [10] Petigura EA, Marcy GW, Howard AW. A plateau in the planet population below twice the size of earth. *Astrophysical Journal*. 2013;**770**:69. DOI: 10.1088/0004-637X/770/1/69
- [11] Coughlin JL et al. Planetary candidates observed by Kepler. VII. The first fully uniform catalog based on the entire 48 month dataset (Q1-Q17 DR24). *The Astrophysical Journal Supplement Series*. 2016;**224**(1):27. DOI: 10.3847/0067-0049/224/1/12
- [12] Catanzarite JH. Autovetter Planet Candidate Catalog for Q1-Q17 Data Release 24. p. 36
- [13] Jara-Maldonado M, Alarcon-Aquino V, Rosas-Romero R. A multiresolution machine learning technique to identify exoplanets. In: Martínez-Villaseñor L, Herrera-Alcántara O, Ponce H, Castro-Espinoza FA, editors. *Advances in Soft Computing. MICAI 2020. Lecture Notes in Computer Science*. Vol. 12468. Cham: Springer; 2020. pp. 50-64. DOI: 10.1007/978-3-030-60884-2\_4
- [14] Thompson SE et al. A machine learning technique to identify transit shaped signals. *The Astrophysical Journal*. 2015;**812**(1):46. DOI: 10.1088/0004-637x/812/1/46
- [15] Kreidberg L. batman: BAsic Transit Model cAlculatioN in Python. *Publications of the Astronomical Society of the Pacific*. 2015;**127**(957):1161-1165. DOI: 10.1086/683602
- [16] Mandel K, Agol E. Analytic light curves for planetary transit searches.

The Astrophysical Journal. 2002;**580**(2):  
L171-L175. DOI: 10.1086/345520

[17] Pearson KA, Griffith CA,  
Zellem RT, Koskinen TT, Roudier GM.  
Ground-based spectroscopy of the  
exoplanet XO-2b using a systematic  
wavelength calibration. The  
Astronomical Journal. 2018;**157**(1):21.  
DOI: 10.3847/1538-3881/aaf1ae

[18] Parviainen H. PyTransit: Fast and  
easy exoplanet transit modelling in  
Python. Monthly Notices of the Royal  
Astronomical Society. 2015;**450**(3):  
3233-3238. DOI: 10.1093/mnras/stv894

[19] McHardy IM, Emmanoulopoulos D,  
Papadakis IE. Generating artificial light  
curves: Revisited and updated. Monthly  
Notices of the Royal Astronomical  
Society. 2013;**433**(2):907-927.  
DOI: 10.1093/mnras/stt764

[20] Eastman J, Gaudi BS, Agol E.  
EXOFAST: A fast exoplanetary fitting  
suite in IDL. Publications of the  
Astronomical Society of the Pacific.  
2013;**125**(923):83-112. DOI: 10.1086/  
669497



# Network Intrusion Detection Using Dendritic Cells and Danger Theory

*David Limon-Cantu and Vicente Alarcon-Aquino*

## Abstract

The Dendritic Cell Algorithm (DCA) is a bioinspired, population-based, supervised binary classifier, designed for anomaly detection in communication networks. The proposed model is inspired by the behavior of Dendritic Cells and Danger Theory. The main contribution of this research addresses two contemporary challenges of Network-based Intrusion Detection Systems, namely feature selection and generalization capabilities to improve classification performance. Feature selection improvement is achieved by using information gain and mutual information. A Decision Tree model is incorporated as a classification mechanism in order to improve accuracy, as a substitute to the classification threshold of the DCA. The proposed model is assessed using two publicly available datasets, namely UNSW-NB15 and NSL-KDD. Experimental results are compared against state of the art bioinspired and machine learning approaches for binary classification. The proposed approach provides competitive results when compared to other state of the art approaches, such as Support Vector Machines, and Artificial Neural Networks, achieving a 97.25 and 93.28% accuracy for the UNSW-NB15 and NSL-KDD datasets, respectively. Future challenges include multi-class classification, further performance improvements, and online detection.

**Keywords:** Anomaly detection, Dendritic Cell Algorithm, Decision Tree, binary classifier, Danger Theory, Artificial Immune System

## 1. Introduction

Anomaly detection refers to the problem of finding unexpected behavior. These are often known as anomalies, outliers, or discordant observations [1], and are usually patterns not conforming with a notion of normal behavior. The detection of anomalous patterns consists of defining a region represented as normal behavior, and any element distant from such a region is determined as anomalous; this distinction is achieved through several methods, including searching, signature-based, anomaly-based, feature learning, and feature reduction.

Intrusion Detection Systems (IDS) aim to prevent undesired usage of computer networks. This is performed using tools such as machine learning algorithms and signature-based detection, to generate alerts based on the status of the protected resources. This helps system administrators to make decisions that can affect the network systems, depending on important factors, such as response time and

accuracy of the status. IDS can be classified into two broad groups, namely Network Intrusion Detection Systems (NIDS) and Host-Based Intrusion Detection Systems (HIDS). NIDS are IDS whose main purpose is to analyze network communications, find anomalies and predict incoming attacks; whereas HIDS are specific purpose IDS whose objective is to protect a specific computer system.

Machine learning NIDS have generated relevant results [2, 3]. Alternative approaches aim to solve relevant NIDS anomaly detection challenges, namely high computational complexity and online detection. Artificial Immune Systems (AIS) are a type of evolutionary computing algorithms and models, inspired by the behavior of the Human Immune System (HIS). Their aim is to imitate the favorable qualities of their biological counterpart. Although there exist other evolutionary computing algorithms, such as Genetic Algorithms (GA), the immune system is solely focused on the protection of its host system.

The Dendritic Cell Algorithm (DCA) is a computational model developed around the immune Danger Theory (DT) and is a population-based binary classifier designed for anomaly detection, where Dendritic Cells are represented as agents known as artificial Dendritic Cells (DCs). The algorithm is able to assess whether a group of observations are anomalous or normal through temporal correlation of preprocessed features and linear equations to simulate part of the observed behavior of biological DCs. The DCA algorithm evolution has been marked by three different contributions, starting with the “prototype” DCA [4], followed by a more elaborated version using stochastic elements, known as the “stochastic” DCA [5], and further developed as the “deterministic” DCA [6–9].

### **1.1 Related work**

Several machine learning, bioinspired, and meta-heuristic methods have been developed for anomaly detection in communication networks. Machine learning algorithms used for intrusion detection can be divided into two broad groups. Deep learning models have achieved remarkable results and can automatically learn feature representations, such as Convolutional Neural Network (CNN) [10], and Deep Neural Network (DNN) [3]. Traditional machine learning techniques, conversely, are characterized for their lack of “depth” in the analysis, such as Support Vector Machine (SVM) [11] K-Nearest Neighbor (KNN) [12], Decision Forest [2], Random Forest [3] and Naive Bayes classifier (NB) [13].

Artificial Immune Systems (AIS) are classified into two major categories, namely network-based and population-based. Network-based algorithms make use of the Immune Network Theory and are based on Artificial Immune Networks [14]. Population-based algorithms, on the other hand, imitate immune cell behavior through artificial agent interactions and are based on Negative Selection [15], Clonal Selection [16, 17], or Danger Theory [8, 18–23]. AIS models have focused on imitating some characteristics of the HIS, such as multiple-level detection mechanisms based on DT [20], and modifications to the DCA. Said modifications include incorporating probability theory [19], fuzzy inference systems [21], feature selection [22], and detection improvements in a semi-supervised context [23].

### **1.2 Contribution**

The main contribution of this research is a biologically inspired NIDS approach based on the deterministic DCA [6]. This model aims to tackle two challenges (and contemporary issues) of NIDS, namely feature selection, and generalization capabilities to improve classification accuracy. A comparison with different bioinspired and machine learning techniques using two publicly available benchmark datasets

(NSL-KDD and UNSW-NB15) is presented. The rest of this paper is organized as follows. Section 2 details the related methodology, as well as the proposed model. Section 3 presents datasets definition, model parameters, and numerical results, as well as a comparison of efficiency metrics with state of the art approaches for binary classification. Section 4 presents conclusions, challenges, and future work.

## 2. Methodology

Binary classification is the task of classifying elements of a given set into two groups, on the basis of a classification rule [18]. The objective of the proposed model consists of achieving anomaly classification based on the provided observations. The first process consists of performing feature selection and data categorization, to provide the proposed algorithm with input data. The DCA performs context assessment and finally, a classifier is used to produce a concrete assessment. Each observation is then classified as normal or anomalous and performance metrics are generated. The objective of this section is to introduce mathematical and algorithmic background. The proposed methodology contains four phases, namely dataset preprocessing, algorithm initialization, detection, and classification.

The Danger Theory model [24] was proposed by French immunologist Polly Matzinger and is mainly centered on the interactions of signals emitted by cells and antigens. These signals denote when a cell or a tissue is experiencing regular or abnormal behavior, such as programmed or unexpected cell death (known as apoptosis and necrosis respectively) or stress caused by antigens (pathogen or harmful organism signatures). The signals are categorized into three groups, namely Pathogen Associated Molecular Patterns (PAMP), Safe Signals (SS) and Danger Signals (DS). Biological Dendritic Cells are Human Immune System cells, constantly sensing the environment for such signals. These are collected (ingested) in order to assess whether the present alterations are due to an attacking organism or as a result of a normal process, for which an immune response is not necessary (known as a regulatory or tolerance process).

### 2.1 Feature selection

The DCA requires input data to be represented as three input signals, namely PAMP, SS, and DS, as well as antigen representation (such as data ID's or attack type). Each input signal used by the algorithm denotes part of the context for the observations analyzed. As antigens in the immune system are organisms associated with disease, this signal category is related to the presence of attacks. Safe Signals are associated with the normal behavior of a biological cell life cycle. This signal category is related to normal behavior in the observed network communications. Danger Signals are emitted by cells and tissues that are stressed or damaged. This signal category indicates suspicious behavior in the network.

The preprocessing phase assigns a set of features from the original dataset to each of the signal categories (PAMP, SS, DS). This is commonly done by using expert knowledge or feature reduction methods such as PCA, Fuzzy Set Theory [18], or K-Nearest Neighbors [25]. In order to determine the features with the most influence [21, 26], the proposed approach relies on the information gain method, along with maximizing feature-class mutual information for signal categorization, followed by an average feature aggregation and normalization for each category. The information gain of an attribute  $F$  and a given dataset  $S$  is evaluated as shown in Eq. (1),

$$G(S, F) = H(S) - \sum_{v \in \text{values}(F)} \frac{|S_v|}{|S|} * H(S_v) \quad (1)$$

where  $\text{values}(F)$  represents all the possible values of a given feature  $F$  in the set  $S$ ,  $S_v \subset S$  where  $v$  is a potential value that attribute  $F$  may take,  $G$  is the information gain function and  $H$  represents the entropy of a system, as shown in Eq. (2),

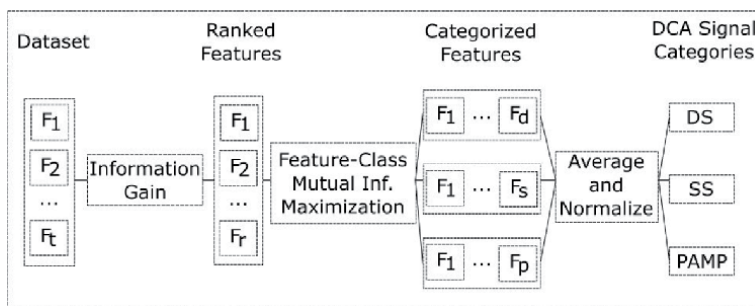
$$H(S) = \sum_{i=1}^2 -p_i * \log_2 p_i \quad (2)$$

where  $p_i$  represents the probability of a given class  $i$  in the dataset  $S$ , based on the values of attribute  $F$ . High entropy implies the attribute provides a high amount of information about a feature in the dataset, high ranking attributes are preserved such as to have at least one feature per signal category. Each selected feature is assigned into one of the three signal categories, namely PAMP, DS, and SS. This is performed by performing feature-class mutual information maximization. Given two random features  $F$  and  $C$ , the mutual information among them  $I(F;C)$  is the amount of information that the feature  $C$  gives about  $F$ , as shown in Eq. (3),

$$I(F; C) = \sum_{f \in \text{values}(F), c \in \text{values}(C)} p(f, c) * \log \left( \frac{p(f, c)}{p(f)p(c)} \right) \quad (3)$$

where  $p(f,c)$  represents the joint probability of attribute values  $f$  and  $c$ ,  $p(f)$  and  $p(c)$  are the marginal probabilities. In order to categorize the selected features, the feature-class mutual information between each attribute and class is calculated. If a given attribute has higher mutual information with the normal class than it has with the anomalous class, it is categorized as SS. Conversely, if the attribute has higher mutual information with the anomalous class than with the normal class, it is categorized as PAMP. The remaining features are classified as DS.

The DCA contains a population of artificial Dendritic Cells, to simulate the behavior of biological cell context assessment capabilities in a human body. Each cell in the population has a predefined migration threshold (or lifespan). After which the cell does not sense signals or antigens. Its state is aggregated to the antigen repository used to classify after all observations have been processed. Algorithm initialization is performed in order to provide the detection phase with the required parameters, namely migration threshold and DC population size. The preprocessing phase is summarized in **Figure 1**. Dataset features are defined as  $Dataset = \{F_1, F_2, \dots, F_t\}$ ,  $t$  is the total number of dataset features. The information gain selected features  $Ranked = \{F_1, F_2, \dots, F_r\} \subseteq Dataset$ ,  $r$  is the total number of



**Figure 1.**  
Dataset preprocessing.

ranked features, are then compared against normal and anomalous data in order to generate three subsets of categorized features, namely danger signals  $\{F_1, F_2, \dots, F_d\} \subset Ranked$ , safe signals  $\{F_1, F_2, \dots, F_s\} \subset Ranked$ , and PAMP signals  $\{F_1, F_2, \dots, F_p\} \subset Ranked$ ,  $d, s, p$ , are the total number of features for each signal category (DS, SS, and PAMP). Categorized features are averaged and normalized in the closed range of  $[0, 1]$ , in order to generate the processed dataset, where only four predictors are present namely DS, SS, PAMP, and antigen representation.

## 2.2 Detection phase

The detection phase aims to generate an antigen repository. This process is achieved after a population of artificial DCs (or agents) is created. The agent population performs signal ( $PAMP_i, DS_i, SS_i, i = 1, 2, \dots, n$ ,  $n$  is the dataset size) and antigen ( $\alpha$ ) collection until a threshold is met. Antigen types that are collected by each cell are counted and stored as cell state signals  $\alpha_g$ , where  $g$  represents antigen categories. For each observation fed into the algorithm, the entirety of the DC population samples signals and antigens. The proposed approach incorporates cumulative signals known as Costimulatory Molecule Signal (CSM), Semi-mature Signal (smDC) and Mature Signal (mDC) [4]. These are defined in Eq. (4),

$$C_{CSM,smDC,mDC} = (W_P * C_P) + (W_S * C_S) + (W_D * C_D) \quad (4)$$

where  $C_{CSM,smDC,mDC}$  represents the signal concentration for CSM, smDC, and mDC respectively,  $W_{P,S,D}$  are the weights used for PAMP, SS, and DS [5, 27].  $C_{P,S,D}$  are the signal concentration values for each antigen sampled by the artificial DC. The role of CSM is to limit the time an artificial DC spends on antigen sampling by imitating the cell's lifespan (or signal collection limit). The smDC and mDC signals determine the cell context for the antigens collected in the DC population and are the basis used to generate the  $\hat{k}$  anomaly context. When a DC has exceeded the DC maturation threshold (set in algorithm initialization), it migrates to a separate DC pool where it no longer samples antigens. A new DC is created in the original DC population pool to always preserve the initial number of DCs. The deterministic DCA employs  $\hat{k} \in R$  to reflect the anomaly characteristic (or signature) of a migrated cell, this is shown in Eq. (5), where  $s$  represents the signals received by each artificial DC,  $C_{mDC}$  and  $C_{smDC}$  are the intermediary mature and semi-mature signals respectively.

$$\hat{k} = \sum_1^s C_{mDC} - C_{smDC} \quad (5)$$

After all data instances in the dataset have been processed, all migrated cells anomaly context and observed antigen count are summarized using  $k_\alpha$ , defined as the sum of all  $\hat{k}^\alpha$  presented by each DC for antigen category  $\alpha$ , in proportion to the amount of antigens presented in all migrated DCs, as defined in Eq. (6), where  $m$  represents the index of a DC in the migrated population.

$$K_\alpha = \frac{\sum_m \hat{k}_m^\alpha}{\sum_m \alpha_m} \quad (6)$$

## 2.3 Classification

The classification phase generates a distinction criterion for all obtained  $k_\alpha$  anomaly signatures in the antigen repository. The DCA classification was based on a

constant classification threshold [5, 6, 8]. This threshold was commonly set as a user-defined parameter, or derived from observations obtained in the detection phase. This approach is known to have issues [28], as the assigned threshold may not properly separate normal  $k_\alpha$ . The proposed model removes the use of such anomaly threshold, in favor of including a Decision Tree Classifier.

A Decision Tree (DT) is a supervised learning model commonly used for classification and regression tasks. The main objective of a DT is to build a model based on (simple) decision rules that are derived from data predictors. Decision Trees are commonly easy to understand, as they can be visualized. Some favorable characteristics of Decision Trees are low computational complexity for prediction, not requiring large amounts of observation to generate a model, and transparency (as generated rules can be visualized and understood). Decision Trees are also known to overfit. In order to solve this, several constraints and optimization features have been developed, such as pruning, sample number minimum for each leaf node, and maximum tree depth [29].

A Decision Tree is built in a sequential manner, where a set of simple tests are combined logically. For example, comparing a numeric value against a threshold or a specific range, or comparing a categorical value against a set of possible categorical values. As observation is compared against the set of rules generated by a DT, it is determined as belonging to the most frequent class present in that “region”. A Decision Tree can be constructed using graphs, and can be expressed as shown in Eq. (7),

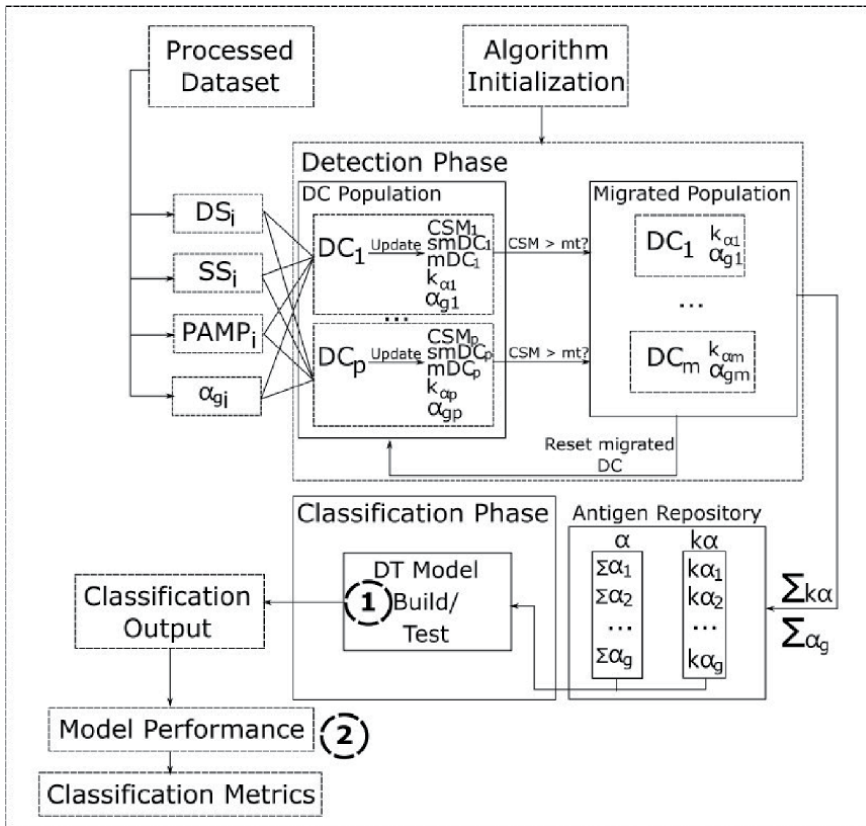
$$G = (V, E) \tag{7}$$

where  $E \in V^2$ ,  $V$  is a set of nodes, and  $E$  is a set of edges. The set of nodes  $V$  can be further described as the joint of three sets, namely  $D, U, T$ , where  $D$  are decision nodes,  $U$  are chance nodes, and  $T$  are terminal nodes, this set is expressed in Eq. (8). Decision nodes execute decision making, in which an action is selected. A chance node randomly selects a related edge. Terminal nodes are the end of action and chance nodes. Each edge contains a parent node association, as well as a child node. Decision Trees have further functions and conditions [30].

$$V = D \cup C \cup T \tag{8}$$

## 2.4 Proposed model

The proposed model is summarized in **Figure 2**. Similar to the deterministic DCA approach, feature ranking is obtained by using Information Gain. Selected features are sorted into one of the three signal categories, namely SS, DS, and PAMP. Each feature set selected for each category is aggregated and normalized. Segment size, migration threshold, and DC population size are set as the algorithm initializes. Data from the processed dataset is fed to the algorithm sequentially, where a set of  $\{DS_i, SS_i, PAMP_i, \alpha_{gi}\}$ ,  $i = 1, 2, \dots, n$ ,  $n$  is the dataset size, and  $g$  is the antigen category for observation  $i$ . Each cell  $DC_1, \dots, DC_p$  in the DC population, where  $p$  is the amount of DC in the population, receives the same set of signals and antigen. An update process is performed to  $CSM_p, smDC_p, mDC_p, k_{ap}, \alpha_{gp}$ . After signal collection in the current iteration, the  $CSM$  status signal is compared against the migration threshold for all  $DC_p$ . If the said threshold is surpassed, the  $DC_p$  is migrated and does no longer perform signal and antigen collection. The accumulated status signals  $k_{am}, \alpha_{gm}$ , where  $m$  is the migrated population size, are accumulated into the antigen repository.



**Figure 2.**  
 DCA with Decision Trees.

Finally, all migrated DCs in the current iterations are reset. Classification is performed after all data elements are processed and by using a Decision Tree (DT). Stage (1) denotes Decision Tree model building. After the model has been built, testing can be performed by providing the testing dataset and starting the algorithm again. Stage (2) achieves classification by using the previously trained DT model after all data elements have been processed. Classification metrics are finally obtained to analyze the model performance.

### 3. Experimental work

The proposed model was tested using the NSL-KDD and the University of New South Wales (UNSW-NB15) datasets. The dataset preprocessing and algorithm was developed using the MATLAB R2020 environment and executed in a computer running the Linux operating system with an Intel Core i7 8700 CPU and 16.0 GB of RAM. A confusion matrix is used to describe performance. For a binary classifier, the confusion matrix consists of positive and negative classes. The positive class refers to any anomaly (attack) present in the dataset. The negative class refers to normal behavior. In order to generate a confusion matrix, the classified records are compared against the dataset actual classes (i.e., ground truth). The anomalous records correctly classified are called True Positives (TP). When TP records are wrongfully classified, they are False Negatives (FN). In the case of normal behavior, correctly classified records are known as True Negatives (TN). Wrongfully

classified normal records are known as False Positives (FP). The resulting performance metrics are then used to generate statistic measures for further analysis and comparison, namely precision, sensitivity, specificity, and accuracy. Precision reflects the proportion of correct classifications and is given by Eq. (9). Sensitivity (also known as TP rate) refers to the proportion of correctly classified anomalies and is given by Eq. (10). In contrast, specificity (or TN rate) is the proportion of correctly classified normal behavior, given by Eq. (11). Finally, accuracy reflects the proportion of true results, either of anomaly or normal behavior, and is given by Eq. (12).

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (10)$$

$$Specificity = \frac{TN}{TN + FP} \quad (11)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

### 3.1 Dataset description

The UNSW-NB15 is a publicly available dataset [31]. It contains nine different attack types, namely Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms, as well as normal traffic. The dataset is divided into train and test sets. The training set contains 175,341 records (119,341 anomalous and 56,000 normal). The testing set, conversely, contains 82,332 records (45,332 anomalous and 37,000 normal). Two tools (Argus and Bro-IDS) along with 12 developed algorithms were used to generate 49 different features, which are categorized into flow features, content features, time features, basic features, and additionally generated features. Statistical analysis, feature correlation, and complexity were evaluated and showed the train and test sets to be of similar distributions [13].

The KDD-99 dataset was developed for the Third International Knowledge Discovery and Data Mining Tools Competition and is publicly available. It was generated to support NIDS development by simulating several intrusions in a military network environment. This dataset contains four attack types, namely Denial of Service (DoS), Probe, User to Root Attack (U2R), and Remote to Local Attack (R2L), and normal traffic. The dataset is divided into two subsets, namely train, and test. The train set contains 494,021 records (97,278 normal and 396,743 anomalous). The test set consists of 311,029 records (60,593 normal and 250,436 anomalous). In total, 41 features for each connection were generated. This dataset was widely used in IDS research. However, it has been the subject of wide criticism due to the probability distribution of the records in the testing set, as well as inconsistencies in the values of the training and testing sets. This has led to an unbalance in normal and anomalous observations, as well as several duplicate data instances [31, 32].

The NSL-KDD [32] is a publicly available dataset developed by the Canadian Institute for Cybersecurity. It was created to solve two main problems of the KDD-99 dataset, namely the distribution of the attacks in the train and test sets, and the over-inclusion of Denial of Service (DoS) attack types (*neptune* and *smurf*) in the test dataset. This dataset also provides the following improvements. The omission of



redundant or duplicate records in the train and test sets, the balancing of records for the train and tests sets, in order to avoid dataset sub-sampling, to reduce computational time in model testing. Given this dataset is an improved version of the KDD-99, it has the same features and attack types. The complete training dataset contains 125,973 features (58,630 anomalous and 67,343 normal). There is a reduced version of the train set (KDD + Train 20%) that contains a 20% subset of the training set. The full testing dataset contains 22,544 records (12,833 anomalous and 9711 normal). Additionally, there exists a testing dataset that does not include records that were not validated by all 21 classifiers used to match the KDD-99 ground truth labels in the dataset creation [32]. The attack types for the presented datasets are detailed in **Table 1**.

### 3.2 Dataset preprocessing

As part of the proposed model phases, dataset preprocessing was performed by ranking the most relevant features to be used for each signal category (PAMP, Safe, and Danger) required for the DCA. The feature ranking, selection, and categorization were based on information gain and feature-class mutual information maximization [21]. As a result, 10 and 17 features were selected for the NSL-KDD and UNSW-NB15 datasets respectively, as shown in **Tables 2** and **3**. Anomalous records of any category are labeled as one, whereas normal records are labeled as zero, to fit binary classification constraints. The selected features were combined by performing normalization in the range from zero to one. Each signal category is equal to the average of its corresponding features, similar to the approach in [21]. Antigen representation was achieved by using several dataset categorical features to

Type	Description	Dataset
Normal	Normal transaction data.	NSL-KDD, UNSW-NB15
Fuzzers	Attempting to cause a program or network to suspend by feeding it with randomly generated data.	UNSW-NB15
Analysis	A series of port scans, spam, and HTML file attacks.	UNSW-NB15
Backdoors	Technique to bypass security mechanisms stealthily.	UNSW-NB15
DoS	Malicious attempt to make a network resource unavailable by overwhelming its capacity to serve requests.	NSL-KDD, UNSW-NB15
Exploits	Leverage the knowledge of a system vulnerability by exploiting it to achieve unauthorized access to a system.	UNSW-NB15
Generic	A technique that works against all block ciphers (encryption method) without consideration of its structure.	UNSW-NB15
Reconnaissance	Attacks that aim to gather information about the network.	NSL-KDD, UNSW-NB15
Shellcode	A small piece of code is used to exploit a software vulnerability.	UNSW-NB15
Worms	A piece of code that replicates itself in order to spread over the network, relying on exploits to gain access.	UNSW-NB15
User to Root Attack (U2R)	The gains access to a regular account on the system and exploits vulnerabilities to gain root access.	NSL-KDD
Remote to Local Attack (R2L)	An attacker without an account sends packets to a system to gain access as a user by exploiting vulnerabilities.	NSL-KDD

**Table 1.**  
*Attack types and descriptions for NSL-KDD and UNSW-NB15 datasets.*

Feature name	Description	Signal category
count	Number of connections to the same host as a current connection.	Danger
srv count	Number of connections to the same service as the current connection in the past 2 seconds.	Danger
logged in	Indicates if a user is logged in.	Safe
srv diff host rate	Percentage of connections to different hosts.	Safe
dst host count	Count of connections having the same destination host.	Safe
error rate	Percentage of connections that have "SYN" errors.	PAMP
srv error rate	Percentage of connections that have "SYN" errors.	PAMP
same srv rate	Percentage of connections to the same service.	PAMP
dst host error rate	Percentage of connections to current host that has an SO error.	PAMP
dst host error rate	Percentage of connections to current host that has an RST error.	PAMP

**Table 2.**  
Feature descriptions for signal categorization, NSL-KDD dataset.

Feature name	Description	Signal category
sbytes	Source to destination bytes.	Danger
dbytes	Destination to source bytes.	Danger
dload	Destination bits per second.	Danger
dmean	Mean of the flow packet size transmitted by the destination.	Danger
dpkts	Destination to source packet count.	Safe
sttl	Source to destination time to live.	Safe
smean	Mean of the flow packet size transmitted by the source.	Safe
ct state ttl	No. for each state according to source/destination time to live.	Safe
ct dst sport ltm	No. of connections of the same destination address and source port in 100 connections.	Safe
ct srv dst	No. of connections that contain the same service and destination address in 100 connections.	Safe
dur	Record total duration.	PAMP
rate	Transfer rate.	PAMP
dttl	Destination to source time to live.	PAMP
sload	Source bits per second.	PAMP
ct srv src	No. of connections that contain the same service and source address in 100 connections.	PAMP
ct src dport ltm	No. of connections of the same source address and destination port in 100 connections.	PAMP
ct dst src ltm	No. of connections of the same source and destination address in 100 connections.	PAMP

**Table 3.**  
Feature descriptions for signal categorization, UNSW-NB15 dataset.

generate antigen categories. Attack categories can be compared to biological antigens invading the body, as they tend to have similar patterns and can also attack recurrently [33].

### 3.3 Model parameters

The proposed model has two configurable parameters, namely migration threshold and DC population size. The DC population size was set to 10 artificial DCs. The impact of migration threshold selection for the DC population is still an open research question. As noted in [9], a high migration threshold results in degraded performance for the DCA. Migration threshold selection was performed analyzing input signals for the datasets tested. Currently, this process needs to be adjusted depending on the dataset and selected features. The migration threshold was set as a uniform distribution generated in the closed real interval  $[0, .001]$ . This was chosen to have at least one migrated cell per iteration, in order to avoid oversampling in the antigen signature generation in the detection phase. The classification phase was performed by building a Decision Tree with pruning using the *fitctree* MATLAB model builder.

The parameters used to build the DT are detailed in **Table 4**. The Decision Tree contains only two predictor categories, namely “Normal” and “Anomalous”. Antigen categories are defined as a combination of categorical features from the dataset, namely *flag* and *attack category* for the NSL-KDD dataset, and *protocol*, *service*, and *attack category* for the UNSW-NB15 dataset. Each predictor is determined as Anomalous if it is an attack of any kind. This process aims to increase antigen signature diversity, as only providing two antigen categories for the DCA detection phase will produce a two-observation classification task. This also reduces the performance penalties for miss-classification. One predictor is used as input for the classifier, namely  $K_{\alpha}$ , as detailed in Eq. (6). An exact search is used as a predictor split. The cost for miss-classification is set as one. No maximum depth is set for the training process. For each split node, a maximum of 10 category levels is set, as to not increase computational complexity considerably. Leaf merging is also performed, where all leaves coming from the same parent are merged, if the risk value total is considered greater or equal to the associated parent risk value. The minimum amount of branch nodes is set to 10. Prior probabilities are set as empirical, as class probabilities are obtained from class frequencies in the class label.

Parameter	Value
Predictor categories	Normal, anomalous
Predictors	$k_{\alpha}$
Predictor split	Exact search
Miss-classification cost	1
Max. categories	10
Leaf merging	Yes
Min. branch nodes	10
Prior probabilities	Empirical

**Table 4.**  
 DT model parameters.

Dataset	Stage	Precision (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)	Computation time in seconds
UNSW-NB15	Train	99.98	99.98	99.96	99.97	183.95
<b>UNSW-NB15</b>	<b>Test</b>	<b>95.01</b>	<b>100</b>	<b>94.24</b>	<b>97.25</b>	<b>66.95</b>
NSL-KDD	Train	95.31	94.85	95.90	95.41	126.42
<b>NSL-KDD</b>	<b>Test</b>	<b>88.91</b>	<b>99.20</b>	<b>87.11</b>	<b>93.28</b>	<b>19.78</b>

**Table 5.**  
*Experimental results.*

### 3.4 Numerical results

The tested model performance is summarized in **Table 5**, the testing performance for each dataset is highlighted in bold. The model was tested using the full train sets for each tested dataset, namely the UNSW-NB15 training set, and KDDTrain+. Testing performance is analyzed and used for comparisons. Classification performance for each dataset was tested using the UNSW-NB15 testing set and KDDTest+. Precision indicates the correctly classified anomaly proportion. The precision for the UNSW-NB15 dataset achieved 95.01%, NSK-KDD dataset was 88.91%. Specificity (or true negative rate) showed 94.24% for the UNSW-NB15 dataset, whereas the NSL-KDD dataset achieved 87.11%.

Additionally, the UNSW-NB15 and the NSL-KDD showed 99.98 and 94.85% in sensitivity. Higher sensitivity indicates the algorithm excels at identifying anomalies, whereas higher specificity denotes normal behavior correctly identified. Accuracy indicates the overall number of corrected assessments, and the UNSW-NB15 dataset achieved 97.25%, whereas the NSL-KDD achieved a 93.28%. Computation time in seconds was calculated. Training and testing time are aggregated to show the total algorithm runtime. The UNSW-NB15 dataset training time was 183.95 seconds, whereas the testing time was 66.95 seconds. Conversely, the NSL-KDD dataset training time was 126.42 seconds, and the testing time was 19.78 seconds.

Contemporary models based on the DCA are compared and presented in **Table 6**, the proposed method results are highlighted in bold. The proposed model was able to surpass other approaches and achieved a 97.25% in accuracy. The stochastic DCA [5] was tested using the UNSW-NB15 in [21], two proposals are included in the comparison and achieved results between 60.4 and 78.04% accuracy. The deterministic DCA [8] achieved a 90.14% accuracy. The fuzzy inference

Dataset	Model	Accuracy (%)
UNSW-NB15	<b>Deterministic DCA with Decision Trees</b>	<b>97.25</b>
	DCA Without Signal Categorization [22]	90.23
	Deterministic DCA [6]	90.14
	Takagi-Sugeno-Kang and Fuzzy Inference DCA [21]	89.30
	Improved Stochastic DCA [23]	84.2
	Stochastic DCA [21, 23]	60.4–78.04
NSL-KDD	Deterministic DCA with Antigen Multiplication [34]	98.6
	Deterministic DCA without Antigen Multiplication [34]	96.1
	<b>Deterministic DCA with Decision Trees</b>	<b>93.28</b>

**Table 6.**  
*DCA accuracy comparison.*

DCA [21] achieved 89.30% accuracy. The deterministic DCA without signal categorization achieved the second best result with a 90.23% accuracy for the UNSW-NB15. The NSL-KDD dataset model accuracy was compared with two other models. The deterministic DCA with the multiplication of antigens [34] achieved the best results with a 98.6% accuracy, whereas the same model without implementing antigen multiplication achieved a 96.1% accuracy. The proposed approach achieved a 93.28% accuracy.

The accuracy of contemporary methods for binary classification is presented in **Table 7**. Accuracy for the NSL-KDD and UNSW-NB15 datasets were compared, the proposed model results are highlighted in bold. A comparison with state of the art machine learning-based models was performed. The best accuracy result for the NSL-KDD dataset was obtained by K-Nearest Neighbors classifier [35] with a 94.92% accuracy. The second-best result was achieved by the proposed model with 93.28% accuracy, followed by a Deep-Learning Long-Short Term Memory model [36] with an 86.99%. Other methods compared include Random Forest classifier [36] with 85.44% accuracy and Artificial Neural Network with 85.31% accuracy.

### 3.5 Discussion

The deterministic DCA performs context assessment by using a population of artificial DCs. Each element in the dataset is sequentially processed. All cells in the population receive the same signals and antigens for the current iteration. When a cell migration threshold is met, a cell does not receive any new signals or antigens and its antigen context values, namely the accumulated antigen signature of all cells that migrated in the current iteration, for each antigen type  $\alpha$  ( $k_\alpha$ ) and the sum of antigens  $\alpha$  received by cell in its lifetime  $\hat{s}_\alpha$ . Said outputs are accumulated in an antigen repository. All cells in the population are able to determine a spatial correlation between signals and antigen types  $\alpha$  by using coefficient  $\hat{k}_\alpha$ , as the accumulated difference of two linear functions, namely  $smDC$  and  $mDC$ . Antigen type  $\alpha$  is determined in the dataset preprocessing phase and can be a distinctive categorical feature that represents similar observations (i.e., attack type, protocol, source port, etc.). Once all signals and antigens in the dataset have been processed, the anomaly metric coefficient  $k_\alpha$  is obtained, and is given as the relation between the sum of all  $k$  for each antigen type  $\alpha$  and the amount of times antigen category  $\alpha$  was sensed by any migrated DC. For classification, the DCA proposed a constant classification

Dataset	Model	Accuracy (%)
UNSW-NB15	Deep Feed-Forward Neural Network [3]	99.19
	Random Forest [3]	98.86
	Gradient Boosted Tree [3]	97.72
	<b>Deterministic DCA with Decision Trees</b>	<b>97.25</b>
	Locally Deep Support Vector Machine [2]	93.30
NSL-KDD	K-Nearest Neighbors [35]	94.92
	<b>Deterministic DCA with Decision Trees</b>	<b>93.28</b>
	Deep Long-Short Term Memory [36]	86.99
	Random Forest [36]	85.44
	Artificial Neural Network [36]	85.31

**Table 7.**  
 Proposed model comparison with machine learning models.

threshold, based on the collected data [6]. Any antigen category  $\alpha$  above said threshold is considered an anomaly. As the threshold calculated using the proposed equation for the deterministic DCA is a constant, it may be prone to large classification penalties when any antigen category is miss-classified, as all instances in the dataset that present this antigen category are affected. This issue may increase when antigen category count is low, or as a large dataset is processed and  $k_\alpha$  tends to have low variance. When the count of signal instances is large enough, the classification threshold tends to zero, and even though the normal antigen category (or categories) may be linearly separable, the classification threshold may not be adequate. This is further worsened if the mean of safe signals is greater or equal to the mean of danger signals, as Equation  $k_\alpha \in R$  can produce negative values. To solve this, the proposed model builds a DT classifier after the detection phase. The decision rules derived from this model are used to classify the antigen repository, generated from all migrated DCs. The proposed model aims to avoid the dependability on a linear classification threshold, as DT can perform classification using a non-linear approach.

The presented computation time results are related to the computational complexity, where the deterministic DCA presents a big O notation of  $O(n^2)$  for a worst-case scenario. Computational complexity increased with the incorporation of a DT classifier in the classification phase. As  $N$  (DC population size) changes, the DT construction does not present an increment or reduction in computation time, since all antigen signatures are summarized in the antigen repository of size  $m$ . Conversely, increasing the amount of antigen types  $m$  presents an increment in computation time. The main drawback of this model resides on the dependence on the DC migration threshold, dataset size, and antigen categories. It is necessary to provide a migration threshold that does not cause cells to migrate prematurely or late, as the over and under-sampling of signals in a migrated cell tends to cause classification errors or reduce antigen signature separability. This affects the DT classifier as it may not be able to assess several signatures of similar magnitudes, and all observations presenting this antigen category are thus incorrectly classified. To decrease this likelihood, it is necessary to provide dataset selection and signal categorization that can produce a relatively low average migration rate. The classification threshold proposed in the deterministic DCA is also highly dependent on the amount of observations and attack distribution in the observed data for training. The proposed model introduced an increase in computational cost. One final issue is, as Decision Trees receive a large number of observations for training, it is known to over-fit, as well as when dealing with high dimensional problems. Further DT optimization procedures in relation to dataset features may need to be implemented to solve such issues.

#### **4. Conclusions**

Anomaly detection in computer networks is a complex task that requires the distinction of normality and anomaly. Artificial Immune Systems are biologically inspired computational models designed for the development of Intrusion Detection Systems. The Dendritic Cell Algorithm (DCA) is a population-based binary classifier, initially designed for network anomaly detection. The proposed model was inspired by the behavior of Dendritic Cells and immune Danger Theory. This research proposed solutions to two relevant anomaly detection challenges, namely feature selection and generalization capabilities to improve classification performance. The proposed model was based on the DCA and incorporated Decision Trees for the classification phase. Two publicly available datasets, namely UNSW-

NB15 and NSL-KDD, were used. The model was trained using each training set provided. A comparison to assess the accuracy of other DCA models, along with state of the art approaches for network anomaly detection was performed. The proposed approach achieved a 97.25% accuracy, with the contemporary UNSW-NB15 dataset, and provided competitive results when compared to other state of the art machine learning approaches. The results using the NSL-KDD dataset achieved a 93.28% accuracy and surpassed machine learning methods, such as Artificial Neural Network and Random Forest. The proposed model was able to surpass other contemporary proposals using the DCA. Relevant challenges derived from the results obtained are the following. The potential of large miss classification due to the low amount of antigen categories; model dependence on migration threshold and their relationship with dataset features; lack of online detection; dependence on a large amount of observations to perform classification; as well as the lack of multi-class classification. There have been several proposals to address some of the presented issues, such as a variable functional migration threshold function [23], and signal categorization optimizations [22]. Said approaches need to be analyzed to further improve the proposed model. Multi-resolution analysis may provide insight to solve some of the mentioned challenges, such as reducing dependence on feature selection and multi-class classification. The proposal of a segmented version of the DCA [7] may provide a framework to implement online classification, reduction of computational complexity, and further increase the model learning capabilities. Although other proposals have included the use of machine learning techniques to perform classification in the DCA [34], the proposed method provides a starting point to incorporate a robust feature selection and classification mechanism to the ongoing research and development challenges of the DCA.

## Acknowledgements

The authors of this paper would like to thank the Mexican National Council of Science and Technology (CONACYT), as well as the Universidad de las Americas Puebla, Mexico, for providing funding for this research.

## Author details

David Limon-Cantu and Vicente Alarcon-Aquino\*  
Department of Computing, Electronics and Mechatronics, Universidad de las Americas Puebla, San Andres Cholula, Puebla, Mexico

\*Address all correspondence to: [vicente.alarcon@udlap.mx](mailto:vicente.alarcon@udlap.mx)

## IntechOpen

---

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Chandola V, Banerjee A, Kumar V. Anomaly detection: A survey. *ACM Computing Surveys*. 2009;**41**(3):1-58
- [2] Rajagopal S, Siddaramappa Hareesha K, Panduranga Kundapur P. Performance analysis of binary and multiclass models using azure machine learning. *International Journal of Electrical and Computer Engineering*. 2020;**10**(1):978. DOI: 10.11591/ijece.v10i1.pp978-986
- [3] Faker O, Dogdu E. Intrusion detection using big data and deep learning techniques. In: *Proceedings of the 2019 ACM Southeast Conference on ZZZ—ACM SE '19*. Kennesaw, GA, USA: ACM Press; 2019. pp. 86-93. DOI: 10.1145/3299815.3314439
- [4] Greensmith J, Aickelin U, Cayzer S. Introducing dendritic cells as a novel immune-inspired algorithm for anomaly detection. In: Jacob C, Pilat ML, Bentley PJ, Timmis JL, editors. *Artificial Immune Systems*. Berlin, Heidelberg: Springer; 2005. pp. 153-167
- [5] Greensmith J, Twycross J, Aickelin U. Dendritic cells for anomaly detection. In: *IEEE International Conference on Evolutionary Computation, CEC 2006, Part of WCCI 2006*. 16–21 July 2006. BC, Canada: Vancouver; 2006. pp. 664-671. DOI: 10.1109/CEC.2006.1688374
- [6] Greensmith J, Aickelin U. The deterministic dendritic cell algorithm. In: Bentley PJ, Lee D, Jung S, editors. *Artificial Immune Systems*. Berlin, Heidelberg: Springer; 2008. pp. 291-302
- [7] Gu F, Greensmith J, Aickelin U. Theoretical formulation and analysis of the deterministic dendritic cell algorithm. *Biosystems*. 2013;**111**(2):127-135
- [8] Greensmith J, Gale MB. The functional dendritic cell algorithm: A formal specification with Haskell. In: 2017 IEEE Congress on Evolutionary Computation (CEC). Donostia, Spain: IEEE; 2017. pp. 1787-1794. DOI: 10.1109/CEC.2017.7969518
- [9] Greensmith J. Migration threshold tuning in the deterministic dendritic cell algorithm. In: Martín-Vide C, Pond G, Vega-Rodríguez MA, editors. *Theory and Practice of Natural Computing*. Cham: Springer International Publishing; 2019. pp. 122-133
- [10] Potluri S, Ahmed S, Diedrich C. Convolutional neural networks for multi-class intrusion detection system. In: Groza A, Prasath R, editors. *Mining Intelligence and Knowledge Exploration*. Cham: Springer International Publishing; 2018. pp. 225-238
- [11] Jing D, Chen H. SVM based network intrusion detection for the UNSW-NB15 dataset. In: *2019 IEEE 13th International Conference on ASIC (ASICON)*. Chongqing, China: IEEE; 2019. pp. 1-4. DOI: 10.1109/ASICON47005.2019.8983598
- [12] Ravipati RD, Abualkibash M. Intrusion detection system classification using different machine learning algorithms on KDD-99 and NSL-KDD datasets—A review paper. In: *SSRN Scholarly Paper ID 3428211*. Rochester, NY: Social Science Research Network; 2019. DOI: 10.2139/ssrn.3428211
- [13] Moustafa N, Slay J. The evaluation of network anomaly detection systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set. *Information Security Journal: A Global Perspective*. 2016;**25**(1–3):18-31. DOI: 10.1080/19393555.2015.1125974
- [14] Shi Y, Peng X, Li R, Zhang Y. Unsupervised anomaly detection for network flow using immune network based K-means clustering. In: Zou B, Li



M, Wang H, Song X, Xie W, Lu Z, editors. Data Science. Singapore, Singapore: Springer; 2017. pp. 386-399

[15] Belhadj Aissa N, Guerroumi M, Derhab A. NSNAD: Negative selection-based network anomaly detection approach with relevant feature subset. *Neural Computing and Applications*. 2019;**32**(8):3475-3501. DOI: 10.1007/s00521-019-04396-2

[16] Lysenko S, Bobrovnikova K, Savenko O. A botnet detection approach based on the clonal selection algorithm. In: 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT). Kyiv, Ukraine: IEEE; 2018. pp. 424-428. DOI: 10.1109/DESSERT.2018.8409171

[17] Elshafie HM, Mahmoud TM, Ali AA. Improving the performance of the snort intrusion detection using clonal selection. In: 2019 International Conference on Innovative Trends in Computer Engineering (ITCE). Aswan, Egypt: IEEE; 2019. pp. 104-110. DOI: 10.1109/ITCE.2019.8646601

[18] Chelly Z, Elouedi Z. A survey of the dendritic cell algorithm. *Knowledge and Information Systems*. 2016;**48**(3): 505-535

[19] Sharma VP, Tiwari R. Immunity based intrusion detection system using probabilistic dendritic cell algorithm. *International Journal of Recent Research Aspects*. 2018;**5**(1):90-99

[20] Alaparthy VT, Morgera SD. A multi-level intrusion detection system for wireless sensor networks based on immune theory. *IEEE Access*. 2018;**6**: 47364-47373. DOI: 10.1109/ACCESS.2018.2866962

[21] Elisa N, Yang L, Fu X, Naik N. Dendritic cell algorithm enhancement using fuzzy inference system for network intrusion detection. In: 2019 IEEE International Conference on Fuzzy

Systems (FUZZ-IEEE). New Orleans, LA, USA: IEEE; 2019. pp. 1-6. DOI: 10.1109/FUZZ-IEEE.2019.8859006

[22] Elisa N, Chao F, Yang L. A study of the necessity of signal categorization in dendritic cell algorithm. In: Ju Z, Yang L, Yang C, Gegov A, Zhou D, editors. *Advances in Computational Intelligence Systems*. Cham: Springer International Publishing; 2020. pp. 210-222

[23] Farzadnia E, Shirazi H, Nowroozi A. A new intrusion detection system using the improved dendritic cell algorithm. *The Computer Journal*. 2020;**64**(8): 1193-1214. DOI: 10.1093/comjnl/bxaa140

[24] Matzinger P. Tolerance, danger, and the extended family. *Annual Review of Immunology*. 1994;**12**(1):991-1045

[25] Kuttranont P, Boonprakob K, Phaudphut C, Permpol S, Aimtongkhamand P, KoKaew U, et al. Parallel KNN and neighborhood classification implementations on GPU for network intrusion detection. *Journal of Telecommunication, Electronic and Computer Engineering*. 2017;**9**(2-2):29-33

[26] Witten IH, Frank E, Hall MA, Pal CJ. *Data Mining: Practical Machine Learning Tools and Techniques*. Cambridge, MA, United States: Morgan Kaufmann; 2017

[27] Greensmith J, Aickelin U, Twycross J. *Articulation and Clarification of the Dendritic Cell Algorithm*. Berlin, Germany: Springer; 2006. pp. 404-417

[28] Gu F, Feyereisl J, Oates R, Reys J, Greensmith J, Aickelin U. Quiet in class: Classification, noise and the dendritic cell algorithm. In: Liò P, Nicosia G, Stibor T, editors. *Artificial Immune Systems*. Berlin, Heidelberg: Springer; 2011. pp. 173-186

[29] Kotsiantis SB. Decision trees: A recent overview. *Artificial Intelligence*

Review. 2013;**39**(4):261-283. DOI: 10.1007/s10462-011-9272-4

IEEE; 2020. pp. 195-199. DOI: 10.1109/ICIoT48696.2020.9089659

[30] Kamiński B, Jakubczyk M, Szufel P. A framework for sensitivity analysis of decision trees. *Central European Journal of Operations Research*. 2018;**26**(1): 135-159. DOI: 10.1007/s10100-017-0479-6

[36] Kasongo SM, Sun Y. A deep long short-term memory based classifier for wireless intrusion detection system. *ICT Express*. 2020;**6**(2):98-103. DOI: 10.1016/j.icte.2019.08.004

[31] Moustafa N, Slay J. UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In: 2015 Military Communications and Information Systems Conference (MilCIS). Canberra, ACT, Australia: IEEE; 2015. pp. 1-6. DOI: 10.1109/MilCIS.2015.7348942

[32] Tavallae M, Bagheri E, Lu W, Ghorbani AA. A detailed analysis of the KDD CUP 99 data set. In: 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications. Ottawa, ON, Canada: IEEE; 2009. pp. 1-6. DOI: 10.1109/CISDA.2009.5356528

[33] Gu F, Greensmith J, Aickelin U. Further exploration of the dendritic cell algorithm: Antigen multiplier and time windows. In: Bentley PJ, Lee D, Jung S, editors. *Artificial Immune Systems*. Berlin, Heidelberg: Springer; 2008. pp. 142-153

[34] Igbe O, Ajayi O, Saadawi T. Denial of service attack detection using dendritic cell algorithm. In: 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON). New York, NY, USA: IEEE; 2017. pp. 294-299. DOI: 10.1109/UEMCON.2017.8249054

[35] Alsaadi HS, Hedjam R, Touzene A, Abdesslem A. Fast binary network intrusion detection based on matched filter optimization. In: 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIOT). Doha, Qatar:

# Automatic Terrain Perception in Off-Road Environments

*Ethery Ramírez-Robles and Oleg Starostenko*

## Abstract.

Autonomous driving is a growing research area; however, there are no fully autonomous vehicle (AV) in the world. Existing AVs have different capabilities and can drive by themselves only in specific scenarios with several constraints. This paper discusses several studies from the point of view of a modular system approach. This approach perceives autonomous driving as separate tasks to solve. Studies are classified in object/pedestrian detection, road detection, obstacle avoidance, terrain perception, mapping of the environment, and path planning. Furthermore, various perception sensors are reviewed and compared. Important datasets and metrics found in the literature are presented. Finally, one of our experiments obtained a weighted IoU of 83.88% in the segmentation of five classes. Since this is a work in progress, more research needs to be done, but our proposal shows promising results in terrain perception in off-road environments.

**Keywords:** autonomous driving, terrain perception, semantic segmentation

## 1. Introduction

Autonomous driving is a growing research area; recently, it has received a lot of attention due to its many advantages. According to a study by Morgan Stanley Research, autonomous vehicles (AVs) can save money from reduced labor costs, improved productivity, lower fuel consumption, and fewer accidents [1]. There are two types of scenes in autonomous driving, on-road and off-road. In the first type, we can find pavement roads, lane markings, defined cues, etc. In the second, there are uneven surfaces, not clear delimiters, vegetation, and different terrains. Several projects have brought significant advances and had a meaningful impact on the state of the art of autonomous driving. The DARPA Challenge held in 2004 was one of the first important competitions initially; it was oriented for military applications, but then the focus change to civilian purposes in urban scenarios. None of the contestants finished; in the second edition, five teams complete the challenge without human intervention.

Autonomous vehicles (AVs) are complex systems. The Society of Automotive Engineers (SAE) [2] defines six autonomy levels in cars, starting from 0 to 5. In level 0, there is no driving automation; the human driver performs all driving tasks. In level 1, some tasks are performed by the car, like adaptive cruise speed control, stability control, and anti-lock braking systems. Partial driving automation is level 2. In this level, there are combined automated functions like acceleration and deceleration in defined situations. Level 3, known as conditional automation, is when the vehicle can control some functions under limited conditions for a certain period. In level 4, the vehicle is capable of fulfilling all driving tasks under certain

conditions. Level 5 is full automation; there is no need for a human driver, the car can drive under all conditions.

There are two main classifications for system architecture in AVs, based on their connectivity and their algorithmic design [3]. In the first, we find ego-only systems and connected systems; for the second, there are modular and end-to-end systems. The majority of the research focuses on modular systems since it is an easier form to implement an AV. In this work, we present several proposals found in the literature from the point of view of modular systems. In addition, there is a short review of the most common sensors used to perceive on AVs. Finally, we implement an existing model for segmenting different terrain types using a lightweight and fast network that can be used in mobile devices.

## **2. Related work**

As mentioned before, there exist different classifications for system architecture in AVs; based on their connectivity, we find ego-only systems and connected systems. Ego-only systems are when a single self-sufficient vehicle carries all the necessary automated driving operations at all times. In contrast, connected systems depend on other vehicles and infrastructure to make decisions. This last approach is still in an initial phase, but in the future, with the growing area of the Internet of Things (IoT), this will be possible. It is expected to have vehicle-to-vehicle (V2V) communication, vehicle to infrastructure (V2I) communication, and vehicle to everything (V2X) communication. A large amount of data will be available to vehicles, so more informed decisions will be taken; however, new challenges will exist, and AVs could become even more complex.

A second classification is based on the algorithmic design; we find modular systems and end-to-end systems. In modular systems, AVs are seen as separate tasks to accomplish. Every module represents one task that can be solved separately, and then the results of each are integrated to form a complete system. However, this approach is prone to error propagation. On the other hand, in end-to-end systems, all the modules are seen as a black box. In general, the system receives data from the sensors, and the output is the directions for the actuators of the vehicle.

In this work, we will discuss different proposals found in the literature based on modular systems. Some tasks of this classification are object/pedestrian detection, road detection, obstacle avoidance, terrain perception, mapping of the environment, and path planning.

### **2.1 Object/pedestrian detection**

Object detection identifies and locates instances of objects in an image. In this task, when an object is detected, it is marked with a rectangular bounding box. Some general steps in object detection are preprocessing, Region of Interest (ROI) extraction, object classification, and localization. In the preprocessing steps, some subtasks are performed, such as exposure and gain adjustment, camera calibration, and image rectification. Extract regions of interest can also be implemented as a preprocessing step. Approaches that use ROI extraction usually have more computational cost since the system becomes more complex; however, the results are better. Another disadvantage of ROI is processing time; in modular systems, time is an essential consideration because other modules need to be executed so the vehicle can take and implement the decisions in real-time.

The most common approach is Deep Convolutional Neural Networks (DCNN). One of the most known DCNN is YOLO (You Only Look Once) [4]. YOLO works

with a single neural network that predicts bounding boxes, confidence for those boxes, and a class probability map. This network process 45 FPS, and there is a modified version that is faster, but the accuracy is lower. A different method is proposed by [5]; this method consists of a multi-scale CNN. First, they use a proposal sub-network and then a detection sub-network. The proposal network could work as a detector, but it is not strong since its sliding windows do not cover objects well. Thus, the reason to include a detection network was to increase detection accuracy.

Alternatively, Tabor et al. [6] not only apply their own initial implementation of a CNN but also considered aggregated channel features (ACF) and deformable parts model (DPM). ACF uses a sliding window approach where candidate bounding boxes are considered at regular intervals throughout images. DPM uses a two-stage classification process to model parts of an object to move relative to each other and the object centroid. Another approach is Region Proposal Networks followed (RPN) by boosted forests [7] which is a more simple but effective method. RPN generates the candidate boxes as well as convolutional feature maps, while Boosted Forest classifies the proposal using convolutional features.

## 2.2 Road detection

There is no general definition of the road detection problem. Mei et al. [8] defines the problem as “detecting the region in front of the robot that is mechanically traversable by the robot that is apt to be chosen by a human to drive.” This definition can be applied to off-road environments where there are no defined roads like in cities. Approaches in the literature usually consider the scenarios since some methods are more reliable in urban scenarios than off-road.

Lane and road boundary detectors are proposed despite the lack of boundaries in some unstructured scenarios. Jiménez et al. [9] presents a new algorithm based on this using a laser scanner and a digital map when available. They applied two methods in parallel to increase the robustness of their results. Their first method is about the study of variations in the detection of each layer of the laser scanner. They detect boundaries when there are sidewalks higher than the road. The second method is for the study of the separation between intersecting sections of consecutive laser scanner layers. The solution proposed in the method is to try to identify areas with constant radius differences within a predefined tolerance, which allows determining the roadway area.

Cameras are a common form of perception; some works consider the color model of the terrain surfaces and illumination conditions to extract and segment roads [9]. The problem, in that case, is formulated as a joint classification. Moreover, Procházka [10] uses the Monte-Carlo algorithm to segment road regions. They estimate the probability density function (PDF) of road pixels from a sequence of observations. The sequential Monte-Carlo estimation is the one that approximates the PDF. In contrast, Li et al. [11] combines camera information with laser data. They apply a preprocessing step to detect roads, and then they analyze texture features in grayscale images. The laser sensors provide a traversable region near the front of the vehicle.

## 2.3 Obstacle avoidance

This task satisfies the objective of non-intersection or non-collision with objects, and it is very related to path planning. Obstacle avoidance is a crucial aspect of autonomous driving; however, some researchers took more emphasis on optimizing the avoidance of crashes while others only comply with satisfying this task but not in the best form. Some important aspects to consider are the vehicle's

characteristics, like the turning radius and the velocity. Similar to some path planning proposals, some authors use cubic splines [12] to generate several paths considering obstacles, and the best path is selected using optimization techniques. Other approaches use fuzzy algorithms [13] to control AVs, considering vehicle dynamics and the geometry of the obstacles.

## **2.4 Terrain perception**

This task is a vital component of AVs in an off-road environment. In contrast with cities, off-road scenes are more unstructured, and surfaces are not expected to be flat. AVs must be able to decide whether the terrain ahead is passable easily, passable with caution, or whether it is better to avoid. Usually, the information processed in this task comes from images. Another widely used sensor is a laser; information obtained with this kind of sensor helps build a 3D map of the scene to understand terrains with different altitudes.

Cameras are usually mounted on top of AVs, but in some cases, like in small robots, cameras view only the ground, so only one type of terrain is perceived at the time. In automobiles, the perspective is different; bigger pictures are obtained containing also information from the sky. Some researchers use a more classical approach; it is common to see feature extractors and classifiers. There are different forms of feature extraction; for example, Filitchkin and Byl [14] uses a bag of visual words from speeded up robust features (SURF). Other works use local binary patterns (LBP) and local ternary patterns (LTP) [15]. Besides that, some approaches create a combination of features, for instance, color and edge directivity (CEDD) and fuzzy color and texture histogram (FCTH) [16].

A common classifier used not only in computer vision tasks but in other areas is Support Vector Machine (SVM), which is one of the most robust prediction methods in the literature; however, this classifier uses supervised learning. Random forests had been found useful to classify asphalt, tiles, and grass with information from cameras and lasers [15]. A different approach is the use of CNN [17, 18]; usually, no preprocessing steps are performed; only RGB images are the inputs to the network. One disadvantage is the need for large amounts of data needed to train this kind of network.

## **2.5 Mapping of the environment**

Mapping presents a digital representation of the environment; it helps to decide a safer path to follow. Usually, 2 and 3 Dimensional (2D and 3D) information is used. Create 3D maps can be computationally expensive and can increase processing time. Some approaches use a priori maps; the system compares real-time readings with previous data. The main disadvantage in a priori map is the changes in the environment; specifically, in off-road scenes, it is difficult to have the same characteristic all the time, that is, vegetation growth.

Some representations in this task are superpixels, stixels, and 3D primitives [19]. In pixel-based representation, each pixel is a separate entity; due to this, in high-resolution images, the complexity is more. Superpixels are groups of pixels used to solve the problem of complexity. These groups are obtained by segmenting the image into small regions; these should be similar in color and texture. Stixels are presented as a medium-level representation of 3D traffic scenes with the goal to bridge the gap between pixels and objects. These are represented by a set of rectangular sticks standing vertically on the ground to approximate surfaces. 3D primitives are blocks of 3D basic geometric forms such as cubes, pyramids, cones, among others.

With the help of 2D and 3D information obtained from LIDAR and other sensors, the systems can have a sense of the geometric structure around the vehicle. A way to map the environment is by using semantic segmentation combined with CNN [20]. The combination of neural networks with other approaches creates more robust methods than the approaches that used a single algorithm. Nevertheless, sometimes there are difficulties in estimating the pose of lasers, which is required for the proper registration of the range measurements. As a result, Parra-Tsunekawa et al. [21] proposed the use of the extended Kalman filter to estimate in real-time the instantaneous pose of the vehicle and the laser rangefinders by considering various measurements acquired by different sensors.

## 2.6 Path planning

In this task, the main goal is to find a geometric path from an initial point to an endpoint. Sometimes, the vehicle dynamics can be considered in the problem even though this will mean the work can only be applied to vehicles with the same characteristics [22, 23]. A better approach is to work in path planning considering a general solution that is not tied to any specific vehicle [24]. There are two main approaches global route planning and local path planning. Global planners search routes from origin to the final destination, some proposals focus on efficiency in real-time traffic. In contrast, others can compute directions in milliseconds, and others consider space requirements. Local planners find trajectories in real-time considering obstacles, and their objective is to complete the global route. Despite the different approaches, there is an existing controversy among some researchers that if an AV should drive like a human or should look for the optimum path.

Some proposals use SVM [22] and Genetic Algorithms (GA) [25]; with these algorithms usually, other methods are applied in the first step, for instance, A\* algorithm which is a typical graph search algorithms in pathfinding. Another method extensively used is the use of Artificial Neural Networks (ANN). There are different variants, the more used in autonomous driving include CNN [23] and Fully Convolutional Networks (FCN) [24].

## 3. Sensing hardware

In this type of system, sensor redundancy is commonly used. Sensors are a necessary part of AVs; there are different devices to perceive the environment. Some of the most used are presented next.

- **Monocular vision:** RGB images are usually obtained with cameras mounted in front of the upper part of the vehicle. One of the most significant advantages is their cost, compared with other devices, is cheaper, and the results are overall good. Nevertheless, this type of sensor is affected by weather and illumination. Several studies use cameras to perceive color, which is important to know in some tasks.
- **Light detection and ranging (LIDAR):** This type of sensor is commonly seen in AVs; the data obtained can be helpful to achieve better success rates. LIDARs work emitting light waves and measuring the reflection to have distances with objects. Until years ago, the main disadvantage was their size and cost; however, recently, the size started to decrease, and now we can find these sensors even in mobile devices like the iPhone 12. Nonetheless, these new smaller

sensors do not have the same range of detection as the bigger ones. LIDARs are generally used for mapping the environment and object detection.

- Stereo imaging: This modality provides similar information to LIDARs. 3D data is obtained through two cameras and can be used in some basic tasks that a LIDAR can perform, but the accuracy and reliability are not the same; however, stereo imaging is cheaper.
- Radio detection and ranging (RADAR): This kind of sensor work in the same way as LIDARs, with the difference that it uses radio waves instead of a laser, and its resolution is lower. One of the main differences between these two sensors is that RADARs can detect at longer distances than LIDARs, while both are not affected by illumination conditions.
- Global positioning systems (GPS): GPS devices are commonly used in several systems, not only AVs. They communicate with several satellites to provide geographical information about where the sensor is located in the world. These devices obtain precise information; however, there are scenarios where the signal can be lost, for example, in tunnels, tree-lined streets, or underpasses. In these scenarios, the Inertial Measurement Unit (IMU) is very important; they can improve the accuracy and help to estimate the position of the vehicle.
- Vehicle dynamics: In this section, we find all the sensors typically installed in vehicles. They perceive speed, yaw rate, and acceleration. These sensors are useful in the implementation of control navigation. Nevertheless, sometimes automobiles do not provide an easy form to obtain this information from the communication bus.

#### 4. Datasets and evaluation metrics

Datasets are important to train and evaluate algorithms. In the literature, there are several known datasets to use on autonomous driving projects. Some of the most popular are PASCAL VOC [26], KITTI Vision Benchmark [27], MS-COCO [28], ImageNet [29], Berkeley DeepDrive [30], nuScenes [31], Oxford RobotCar [32], Waymo Open [33], and Cityscapes [34]. There are other small datasets like Freiburg Forest Dataset (FFD) [18], Hand-Labeled DARPA LAGR Datasets [35], and NREC Agricultural Person-Detection Dataset [36]. Every dataset contains a different structure, but in general, all have a set for training and others for evaluation. In the area of autonomous driving, the majority of datasets contain images, but there are others containing information from LIDAR sensors. Only a few contain other kinds of data like depth, near-infrared, radar, GPS, vegetation indexes, etc.

One of the most common metrics to evaluate classification algorithms is accuracy [Eq. (1)]. It is defined as the number of correct predictions over the total number of predictions made. In binary classification, it can also be calculated in terms of positives and negatives [Eq. (2)]. In the case of imbalanced data, accuracy does not present an accurate representation. In those cases, precision [Eq. (3)] and recall [Eq. (4)] are better metrics to use. The first metric attempts to answer the question what proportion of positive identifications was actually correct? And the second answer the question what proportion of actual positives was identified correctly?

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions made}} \quad (1)$$



$$Accuracy = \frac{True\ Positives + True\ Negatives}{True\ Positives + True\ Negatives + False\ Positives + False\ Negatives} \quad (2)$$

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (3)$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (4)$$

A common metric used in object detection, road detection, and terrain perception is the Jaccard index, also known as Intersection over Union (IoU) [Eq. (5)]. This metric measures the similarity between two finite sets, in this case, the ground truth and the prediction. Depending on the task, the ground truth can be a bounding box or a mask. IoU is defined as the area of overlap between the ground truth and the prediction divided by the area of the union of both. The metric range goes from 0 to 1 (0–100%), where 0 means no overlap at all and 1 is a perfect overlap of masks.

$$IoU = \frac{Area\ of\ union}{Area\ of\ overlap} \quad (5)$$

There are other forms of evaluation proposed by several researchers that are not standardized metrics. In some papers for obstacle avoidance, it is not only evaluated if the vehicle hits or not an obstacle but also the distance of the AV with objects. The evaluation can be very subjective for path planning since there is not an exact and unique path to follow. An important consideration is the mechanical characteristics of the AV. Not all vehicles can traverse through the same roads, that is, an all-terrain vehicle compared to a commercial automobile or a military vehicle. In Mei et al. [8], they proposed a metric called mechanical traversability; they defined it as the percentage of extracted road pixels that are mechanically traversable. Another form of evaluating an AV is proposed by Bojarski et al. [23], which measure the percentage of autonomy of the vehicle [Eq. (6)]:

$$autonomy = \left( 1 - \frac{(number\ of\ interventions) \cdot 6\ seconds}{elapsed\ time} \right) \cdot 100 \quad (6)$$

They assumed that human intervention in an AV would require 6 seconds to take control of the vehicle, re-center it, and restart the autonomous mode. The elapsed time is the total time in seconds of the simulated test.

## 5. Our proposal

Our proposal is focused on the terrain perception stage for off-road environments. Our approach uses an existing public dataset named Freiburg Forest Dataset (FFD) [18] to train a convolutional neural network to segment five different classes in daylight considering good weather conditions.

FFD is an open dataset that contains multi-modal/spectral images. There are 230 training images and 136 validation images. It also contains manually annotated pixel-wise ground truth segmentation masks. Besides RGB images, the other modalities included are two vegetation indexes: the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI); also Near-infrared (NIR)



**Figure 1.**  
*Sample image Freiburg forest dataset with its ground truth mask.*

and depth data. The five classes in this dataset are Obstacle, Trail, Sky, Grass, and Vegetation. All the data was captured at 20 Hz with a camera resolution of  $1024 \times 768$  pixels (**Figure 1**).

The semantic segmentation was achieved using convolutional neural networks. For this step, we select DeepLab [35], a model created by Google to perform semantic segmentation. This model is commonly used to segment objects like persons, vehicles, animals, etc. There are different versions of DeepLab, but the latest v3+ implements a novel encoder-decoder structure and a spatial pyramid pooling module (ASPP). DeepLab supports different network backbones like Xception [36], MobileNet [37], ResNet [38], and PNASNet [39]. Besides, there are pretrained checkpoints to retrain with different data.

For this work, we select MobileNet as the backbone due to its fast and light-weight structure. We select two checkpoints, the first is pretrained on ADE20K and the second in MS-COCO. The main reason to select these two checkpoints is based on the content of that datasets. Both datasets contain labels related to off-road environments, that is, tree, sand, ground, among others. Since the data is in a different format all the RGB images and the PNG ground truth images were transformed into TFRecord format.

All the training and evaluation were run in Python on a Laptop with a Core i7-8750H and an NVIDIA GTX 1050Ti GPU.

## 6. Results

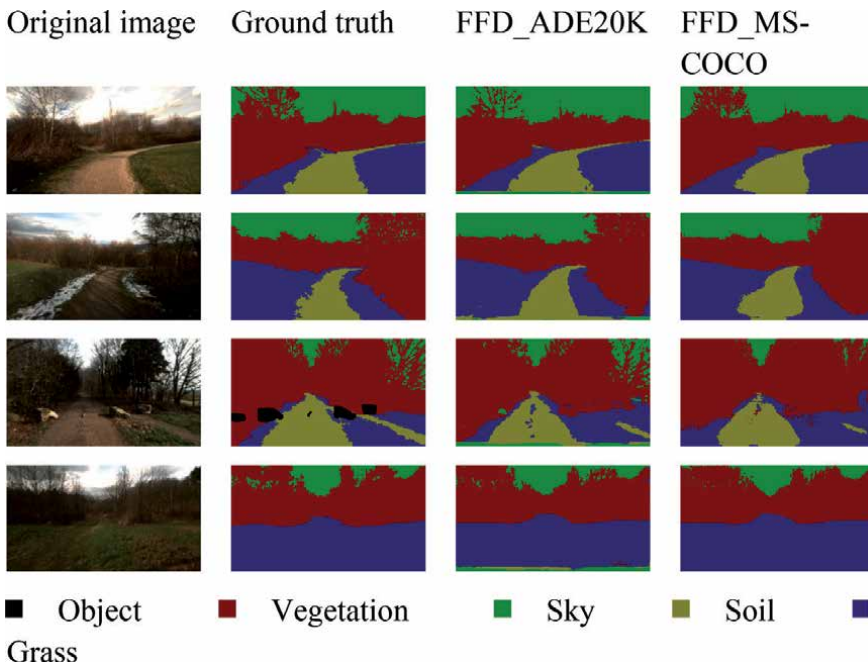
In this work, we study several perspectives on how to attack the problem of autonomous driving. As mentioned before, the two main approaches are modular and end-to-end. Our approach is based on the task of terrain perception. Our approach was to apply transfer learning to retrain an existing model to segment different terrains. Two checkpoints were selected, and five classes were segmented: (0) Object, (1) Vegetation, (2) Sky, (3) Soil, and (4) Grass.

We use Intersection over Union as a metric to evaluate the performance of our approach. Also, the mean IoU (mIoU) is presented, but it is not a good form to evaluate since it does not consider the number of times that a class is presented in the data. mIoU can be skewed by imbalanced datasets giving more importance to classes with more presence. In order to present a more accurate metric, we obtained weighted IoU that gives us an average IoU of each class, weighted by the number of pixels in that class.

**Table 1** presents results for each class, as it can be seen the best results were obtained by the checkpoint pre-trained in MS-COCO with 83.88% of IoU. We believe the reason is the bigger number of images in this set of databases that contained off-road scenes. Nevertheless, class 0, which is objects, was not detected at all in FFD\_MS-COCO. For that specific class, the best results were obtained with FFD\_ADE20K with an IoU of 20.45%. It is important to mention that this class is not found in every image, so the network does not have enough examples to learn and give better predictions.

	ADE20K (%)	MS-COCO (%)
Class 0	20.45	0
Class 1	85.00	85.57
Class 2	78.83	89.46
Class 3	69.77	75.07
Class 4	75.70	80.10
mIoU	65.95	66.04
Weighted IoU	79.67	83.88

**Table 1.**  
IoU comparison results.



**Figure 2.**  
Qualitative results.

**Figure 2** shows some qualitative results; the first column shows the original images. The ground truth mask is presented in the second column, and the rest of the columns are the results of the two different experiments. As shown in some of the results with FFD\_ADE20K, the bottom part of the image was detected as the sky and in some results as soil. This problem was persistent in most of the images, in the same way for FFD\_MS-COCO; it was existent but only in a few images.

## 7. Conclusions

This research presents different approaches found in the literature of modular systems in autonomous vehicles. We focused on modular systems since it is an easier form to solve the problem of autonomous driving. One of the main advantages is redundancy. This type of system needs to be redundant and reliable since it can be dangerous consequences in case of error like human fatalities. Alternatively,

end-to-end systems have become more studied in the latest years. In the future it is expected to have more proposals using this approach, however, at this moment there is more limitation with this type of system like the lack of hardcoded safety measures.

Our approach is based on the terrain perception stage of modular systems. We select an existing model that obtains similar results to the one found in the literature. Our model is lightweight to be run on mobile devices but still robust enough. Two checkpoints were compared, obtaining 83.88% of weighted IoU for the best result.

We expect to improve the semantic segmentation results by augmenting the dataset for future experiments so the network has more data to learn. Further research could explore different parameters and hyperparameters of the model and their influence on the results. Since the architecture selected is oriented to run in mobile devices, we look for implementing and testing the video segmentation in smartphones.


## **Author details**

Ethery Ramírez-Robles\* and Oleg Starostenko  
Department of Computing, Electronics and Mechatronics, Universidad de las Américas Puebla, Puebla, México

\*Address all correspondence to: [ethery.ramirezrs@udlap.mx](mailto:ethery.ramirezrs@udlap.mx)

## **IntechOpen**

---

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Hermann A, Stadler R, Brenner W, Brenner W, Stadler R. *Autonomous Driving: How the Driverless Revolution will Change the World*. England: Emerald Publishing Limited; 2018
- [2] SAE. Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. In: SAE J3016. Tech. Rep. 2016
- [3] Yurtsever E, Lambert J, Carballo A, Takeda K. A survey of autonomous driving: Common practices and emerging technologies. *IEEE Access*. 2020;**8**:58443-58469
- [4] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE; 2016. pp. 779-788
- [5] Cai Z, Fan Q, Feris R, Vasconcelos N. A unified multi-scale deep convolutional neural network for fast object detection. In: *Computer Vision – ECCV 2016*. ECCV 2016. Lecture Notes in Computer Science. Vol. 9908. Springer, Cham; 2016. pp. 354-370
- [6] Tabor T, Pezzementi Z, Vallespi C, Wellington C. People in the weeds: Pedestrian detection goes off-road. In: 2015 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR). West Lafayette, IN, USA: IEEE; 2015
- [7] Zhang L, Lin L, Liang X, He K. Is faster R-CNN doing well for pedestrian detection? In: *Computer Vision—ECCV 2016*. Amsterdam, The Netherlands: Springer, Cham; 2016. pp. 443-457
- [8] Mei J, Yu Y, Zhao H, Zha H. Scene-adaptive off-road detection using a monocular camera. *IEEE Transactions on Intelligent Transportation Systems*. 2018;**19**(1):242-253
- [9] Jiménez F, Clavijo M, Castellanos F, Álvarez C. Accurate and detailed transversal road section characteristics extraction using laser scanner. *Applied Sciences*. 2018;**8**(5):724
- [10] Procházka Z. Road tracking method suitable for both unstructured and structured roads. *International Journal of Advanced Robotic Systems*. 2013;**10**:158
- [11] Li L, Lian J, Huang H, Yao B, Guo L. 1804. Back-stepping variable structure controller design for off-road intelligent vehicle. *Journal of Vibroengineering*. 2015;**17**(7):3876-3886
- [12] Chu K, Lee M, Sunwoo M. Local path planning for off-road autonomous driving with avoidance of static obstacles. *IEEE Transactions on Intelligent Transportation Systems*. 2012;**13**(4):1599-1616
- [13] Stoian V. A control algorithm for autonomous electric vehicles by fuzzy logic. *Advanced Engineering Forum*. 2018;**27**:103-110
- [14] Filitchkin P, Byl K. Feature-based terrain classification for LittleDog. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. Vilamoura-Algarve, Portugal: IEEE; 2012
- [15] Laible S, Khan YN, Bohlmann K, Zell A. 3D LIDAR- and camera-based terrain classification under different lighting conditions. *Informatik Aktuell: Autonomous Mobile Systems*. 2012; **2012**:21-29
- [16] Zou Y, Chen W, Xie L, Wu X. Comparison of different approaches to visual terrain classification for outdoor mobile robots. *Pattern Recognition Letters*. 2014;**38**:54-62
- [17] Zhang W, Chen Q, Zhang W, He X. Long-range terrain perception using

- convolutional neural networks. *Neurocomputing*. 2018;**275**:781-787
- [18] Valada A, Oliveira GL, Brox T, Burgard W. Deep multispectral semantic scene understanding of forested environments using multimodal fusion. In: *Proceedings in Advanced Robotics 2016 International Symposium on Experimental Robotics*. Nagasaki, Japan: Springer, Cham; 2017. pp. 465-477
- [19] Janai J, Güney F, Behl A, Geiger A. *Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art*. Foundations and Trends in Computer Graphics and Vision. 2020;**12**:1-308
- [20] Maturana D, Chou P, Uenoyama M, Scherer S. Real-time semantic mapping for autonomous off-road navigation. In: *Field and Service Robotics*. Zurich, Switzerland: Springer, Cham; 2018. pp. 335-350
- [21] Parra-Tsunekawa I, Ruiz-Del-Solar J, Vallejos P. A Kalman-filtering-based approach for improving terrain mapping in off-road autonomous vehicles. *Journal of Intelligent and Robotic Systems*. 2014;**78**(3-4):577-591
- [22] Chen J, Jiang W, Zhao P, Hu J. A path planning method of anti-jamming ability improvement for autonomous vehicle navigating in off-road environments. *Industrial Robot: An International Journal*. 2017;**44**(4): 406-415
- [23] Bojarski M et al. End to end learning for self-driving cars. *CoRR*. 2016:1604.07316. Available from: <https://arxiv.org/abs/1604.07316>
- [24] Holder C, Breckon T. Learning to drive: using visual odometry to bootstrap deep learning for off-road path prediction. In: *The 29th Intelligent Vehicles Symposium (IEEE IV 2018)*. Changshu, China: IEEE; 2018
- [25] Al-Jarrah R, Al-Jarrah M, Roth H. A novel edge detection algorithm for mobile robot path planning. *Journal of Robotics*. 2018;**2018**:1-12
- [26] Everingham M, Eslami S, Van Gool L, Williams C, Winn J, Zisserman A. The pascal visual object classes challenge: A retrospective. *International Journal of Computer Vision*. 2015;**111**(1):98-136
- [27] Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? The KITTI vision benchmark suite. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Providence, RI, USA: IEEE; 2012. pp. 3354-3361
- [28] Lin TY et al. Microsoft COCO: Common objects in context. In: *Computer Vision – ECCV 2014*. ECCV 2014. Lecture Notes in Computer Science. Vol. 8693. Springer, Cham; 2016. pp. 740-755
- [29] Russakovsky O et al. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*. 2015;**115**(3):211-252
- [30] Yu F, Chen H, Wang X, Xian W, Chen Y, Liu F, Madhavan V, Darrell T. BDD100K: A diverse driving dataset for heterogeneous multitask learning. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA: IEEE; 2020. pp. 2633-2642
- [31] Caesar H, Bankiti V, Lang A, Vora S, Liong V, Xu Q, Krishnan A, Pan Y, Baldan G, Beijbom O. nuScenes: A multimodal dataset for autonomous driving. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle WA, USA; IEEE; 2020. pp. 11618-11628
- [32] Maddern W, Pascoe G, Linegar C, Newman P. 1 year, 1000 km: The oxford robotcar dataset. *The International*

Journal of Robotics Research. 2017;  
**36**(1):3-15

[33] Sun P et al. Scalability in perception for autonomous driving: waymo open dataset. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA: IEEE; 2020. pp. 2443-2451

[34] Cordts M, Omran M, Ramos S, Rehfeld T, Enzweiler M, Benenson R, Franke U, Roth S, Schiele B. The cityscapes dataset for semantic urban scene understanding. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE; 2016. pp. 3213-3223

[35] Procopio M. Hand-labeled DARPA LAGR datasets. 2007. Available from: <https://mikeprocopio.com/labeledlagrdata.html>

[36] Pezzementi Z et al. Comparing apples and oranges: Off-road pedestrian detection on the National Robotics Engineering Center agricultural person-detection dataset. Journal of Field Robotics. 2018;**35**(4):545-563

[37] Chen L, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Computer Vision – ECCV 2018. ECCV 2018. Lecture Notes in Computer Science. Vol. 11211. Springer, Cham; 2018. pp. 833-851

[38] Chollet F. Xception: Deep learning with depthwise separable convolutions. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE; 2017. pp. 1800-1807

[39] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L. MobileNetV2: Inverted residuals and linear bottlenecks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, UT, USA: IEEE; 2018





# Analysis of Voice and Magnetic Resonance Images to Assist Diagnosis of Parkinson's Disease with Machine Learning

*Gabriel Solana-Lavalle and Roberto Rosas-Romero*

## Abstract

Parkinson's disease (PD) is a chronic neurodegenerative disease that affects 1% of the population and whose diagnosis is considered one of the most challenges in the area of neurology. The goal of our work is to assist physicians with the correct diagnosis and early detection of PD. This chapter provides a review of previous work on PD detection under two perspectives, voice analysis and Magnetic Resonance Imaging analysis, by comparing our work with those from other authors. For the case of voice-based PD detection, accuracy reaches 95.9% in female patients and 94.36% in male patients on the largest available dataset. Another contribution in this area is the analysis of voice features to assist the clinical interpretation of the binary result of voice-based detection. For the case of structural Magnetic Resonance Imaging (sMRI)-based PD detection, detection accuracy reaches 96.97% for female patients and 99.01% for male patients using the Parkinson's Progression Marker Initiative dataset. We provide a discussion about the finding of new regions of interest to assist in the detection of PD on sMRI. There is also a comparison between voice-based and MRI-based PD detection methods. Finally, a perspective on future work for PD detection is discussed.

**Keywords:** Parkinson's disease, machine learning, biomedical engineering, magnetic resonance imaging, voice analysis, diagnostic tool

## 1. Introduction

Parkinson's disease (PD) is a chronic neurodegenerative disease that affects over six million people worldwide. Because PD is most common in people over the age of 50, the number of PD patients is expected to double by 2040 due to the increase in life expectancy [1]. The loss of dopaminergic cells in the substantia nigra region of the brain reduces the amount of dopamine in PD patients, causing dyscontrol in several areas of the brain. Some of the main symptoms of PD are motor symptoms, such as tremors, rigidity, and slow movement (bradykinesia). These symptoms, however, become apparent at an intermediate-advanced stage of the disease when the patient may have had the disease for over ten years [2].

The diagnosis of PD is considered one of the most challenging in the area of neurology. The autopsy of the brain of PD patients has shown that 35% of the

cases clinically diagnosed with PD were incorrect [3]. Usually, the diagnosis is done by a physician who looks for cardinal symptoms of the disease and starts dopaminergic therapy as a differential diagnosis. However, these symptoms appear at a late stage, and the patient may have lived with PD for years. Added to the similarity to other parkinsonian disorders that in some cases have the same motor symptoms as PD, may cost the patient crucial time and money, as inadequate treatments could be given by physicians. On the other hand, if detected in time, PD patients can improve their quality of life by taking the correct medication and therapy [2].

Great efforts are being made to find biomarkers that share some light into the causes and development of PD. Advances in technology provide alternatives to help physicians correctly diagnose PD patients at an early stage, and at the same time, obtain relevant information for understanding the disease. As shown in this article, non-invasive techniques such as medical images and voice recordings combined with machine learning and signal processing have proven to be adequate tools for solving the problem of PD detection with great accuracy.

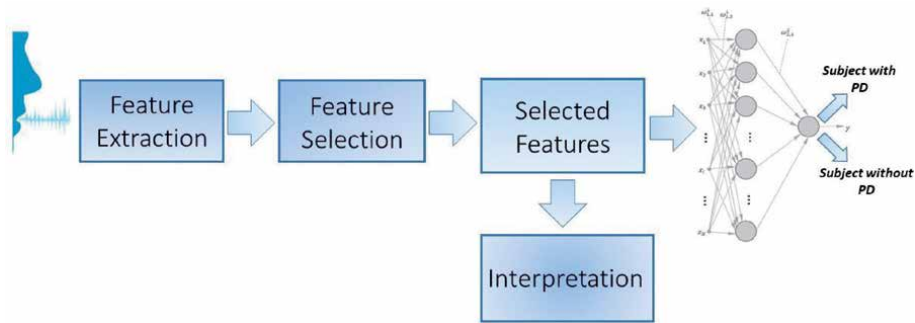
The interest in using voice recordings for PD detection comes from the knowledge that voice disorders are prodromal symptoms present in over 90% of PD patients at an early stage. Some of the alterations include dysphonia (defective use of the voice), hypophonia (reduced vocal loudness), and imprecise hypokinetic articulation [4]. The advantages of this method are that voice recordings can be obtained without going to a hospital, and the economic cost is low, among others.

On the other hand, medical imaging techniques are important tools for understanding and helping diagnose PD. The images give information about neuro-anatomical and pathophysiological processes related to the disease [5]. Some of the most used imaging techniques for neurological disease detection are DaTscan, *Magnetic Resonance Imaging* (MRI), and *Diffusion Tensor Imaging* (DTI). DaTscan images detect the concentration levels of dopamine in different regions of the brain, but the availability and cost of the studies may be prohibitive for patients. Structural Magnetic Resonance Imaging (sMRI) is a technique that provides structural information of the tissues and connectivity of the brain. It is available in most countries and is economically viable compared to other studies.

Work on voice-based and sMRI-based detection of PD and their clinical interpretation is reviewed in this article. The rest of the article is structured as follows: Section 2 presents voice-based analysis, including the database characteristics, classification results, and clinical interpretation of the extracted features. Section 3 introduces the analysis of sMRI of PD patients, the detection performance, and regions of interest for the diagnosis of both female and male patients. Section 4 gives conclusions and future work on the area of PD detection.

## **2. Voice-based analysis applied to the diagnosis of Parkinson's disease**

In general, PD detection based on voice analysis consists of two stages: feature extraction and classification; however, to train classifiers, two additional stages are used: feature selection and performance assessment. The main reasons to analyze voice for PD detection are: (1) voice-based analysis is a low-cost and non-invasive technique, (2) speech problems start at early stages of the disease so that voice-analysis is appropriate for early detection, (3) we have conducted research so that the detection of PD is extended to provide clinicians with quantitative information to help in the understanding of a binary result [6]. In the following, a description



**Figure 1.**  
The most relevant ROIs for PD detection in 1.5 T and 3 T MRI of female patients are highlighted with colors red yellow and green.

of voice-based PD detection is given, the advantages of conducting separate tests for men and women are highlighted. Another issue to explain is the importance of the used dataset size when conducting separate PD detection experiments. It is also explained that the most contributing features to a high detection performance are those obtained with extraction processes that resemble the way the auditory system works.

The first step in the analysis of voice recordings consists of the extraction of features. Different groups of features have been used by researchers, from which *baseline features* are the most common. Baseline features include *jitter*, *shimmer*, *detrended fluctuation analysis* (DFA) among others, and are the most traditional set of features. Other commonly used features are *Mel Frequency Cepstral Coefficients* (MFCC), *Wavelet transform*, and *Tunable-Q Wavelet Transform* (TQWT). These features are obtained by using banks of filters, which extract information over multiple frequency bands at different bandwidths so that the higher the frequency content, the higher the bandwidth. Among all the features, a reduced set of relevant features is obtained through a selection process where correlated features are eliminated. An observation, within the used dataset, is characterized by 754 features extracted from voice recordings. These features belong to six groups of features; however, in our work, there were two groups of features that were not relevant for the classification of voice recordings.

The classifiers are trained with the sets of selected relevant features. In our work, we have used four different classifiers. The classification result is binary since the subject, under analysis, is identified as a *patient with PD* or as not having PD. The different stages (feature extraction, feature selection, classification, interpretation), within the methodology, are shown in **Figure 1**.

### 3. Dataset

Most of the previous works have conducted PD detection by using a population of patients and controls without separate studies for female and male subjects. One reason for not conducting separate studies has to do with the dataset size, which is not large enough to conduct such separate analyses. However, the work from Sakar et al. [7] has provided the research community with the largest voice-based dataset publicly available so far. This dataset was built from 756 voice recordings, where 754 voice features were extracted from each recording by using different signal processing techniques. The recordings were obtained from 107 male patients, 81

female patients, and 64 controls. This number of observations, within the dataset, is high enough to obtain statistically relevant results after partitioning it into two datasets according to gender. A total of 252 individuals were involved during the generation of this dataset. The involvement of each individual consisted in pronouncing vowel /a/ in front of a 44.1-kHz microphone three times. Each recording duration is 220 seconds (9,702,000 samples per recording). Each recording was divided into frames of 25 ms to conduct stationary signal processing for feature extraction. Feature vectors, from different frames, were averaged. Six signal processing techniques were applied for feature extraction. This dataset is found in the Machine Learning Repository of the University of California Irvine. This dataset was generated by the Cerrahpsa Faculty of Medicine at the Department of Neurology, Istanbul University, from 188 PD patients (107 men and 81 women) with an age range between 33 and 87 years old, and from 64 controls (23 men and 41 women) with an age range between 41 and 82 years old.

#### 4. Classification results

*Feature selection* was applied to reduce the dimensionality of feature vectors. Feature selection was conducted by running *Wrappers feature subset selection*, which results in an optimal subset of features for a specific classifier. Feature subset selection was accomplished for each classifier. The most relevant groups of features, selected by Wrappers, were the *Mel Frequency Cepstral Coefficients* (MFCC) and the *Tunable Q-factor Wavelet Transform* (TQWT) features. MFCCs are based on the way the human auditory system works. The computation of MFCCs involves the use of multiple band-pass filters where the filter bandwidth is increased as the central frequency is higher. In the work by Sakar et al. [7], the two most relevant groups of features were TQWT and MFCCs, and the work by Solana-Lavalle et al. [8] is also based on these features.

Multiple classifiers have been applied to the problem of voice-based PD detection such as the *k Nearest Neighbors* (kNN), *Multi-Layer Perceptron* (MLP), *Random Forest* (RF), and the *Support Vector Machine* (SVM). However, after conducting separate studies for male and female populations, it was found that the classifiers, with the highest detection performance, were (1) the *Support Vector Machines* (SVM) with a *radial basis function kernel* (RBF), and (2) the *Multi-Layer perceptron* (MLP) [6]. The highest accuracy reported is 94%, which is a considerable improvement over the previous works that used the same dataset [7, 8]. In addition, the complexity of the last reported model has been reduced from 50 to only 20 used features.

#### 5. Classification results in male and female populations

In the work by Tsanas et al. [9] it was claimed that the problem of PD detection would be more adequate if the problem were addressed by conducting separate analyses on male and female subjects. At that time (2012), such experiments were not possible due to the reduced size of publicly available datasets. However, an adequate number of recordings, for separate statistical studies, has been currently available since the introduction of the dataset by Sakar et al. [7].

In the first work by Solana-Lavalle et al. [8], datasets for PD detection on male and female subjects, were unbalanced, i.e., the number of PD patients is greater than the number of controls. Experiments, with balanced sets of PD patients and controls, were later conducted by Solana-Lavalle et al. [6] with interesting results.

Author, year	Dataset	Results
Peker, 2016 [10]	195 sound measurements from 8 healthy people and 23 PD patients	Accuracy = 0.99, sensitivity = 0.96, specificity = 1
Guruler, 2017 [11]	195 sound measurements from 8 healthy subjects and 23 with PD	Accuracy = 0.99, sensitivity = 1, specificity = 0.99, $F_1$ score = 0.99
Sakar et al., 2017 [12]	42 patients with PD and 8 healthy controls	Accuracy = 0.96, MCC = 0.77
Braga et al., 2019 [13]	22 speakers with PD and 30 healthy speakers	Accuracy = 0.99 for RF classifier
Sakar et al., 2013 [4]	20 patients with PD and 20 healthy individuals	Accuracy = 0.85, sensitivity = 0.85, specificity = 0.9
Raza et al., 2020 [14]	195 voice samples from 8 healthy people and 23 PD patients	Accuracy = 0.97
Vital et al., 2021 [15]	1200 voice samples from 51 healthy people and 62 PD patients	Accuracy = 1
Peker et al., 2015 [16]	195 sound measurements from 23 PD patients and 8 healthy people	Accuracy = 1, sensitivity = 1, specificity = 1
Tsanas et al., 2011 [17]	10 healthy controls and 33 patients with PD	Accuracy = 0.977 and accuracy = 99.03
Montaña et al. 2018 [18]	27 healthy controls and 27 patients with PD	Accuracy = 0.944
Sakar et al., 2019 [7]	756 voice recordings from 64 healthy individuals and 188 patients with PD	Accuracy = 0.86, MCC = 0.59
Proposed approach	756 voice recordings from 64 healthy individuals and 188 patients with PD	Accuracy = 0.947, sensitivity = 0.984, specificity = 0.9268, precision = 0.9722, false alarm rate = 0.0277, MCC = 0.8686

**Table 1.**  
*Voice-based Parkinson detection.*

It was found that detection performance is increased if balanced datasets are used to train and test classifiers.

A comparison of the different methods of voice-based PD detection, proposed by the research community, is shown in **Table 1**. It is observed that different datasets have been analyzed; however, the largest one is the dataset introduced by Sakar et al. [4]. The method that achieves the highest detection performance with the largest dataset is the one proposed by Solana-Lavalle et al. [6]. In addition, to reach the highest detection performance, this method is characterized by the lowest feature vector size.

## 6. Clinical interpretation

The binary output from a classifier implies that a clinician will need further tests to gather strong evidence at the time of taking a diagnosing decision if the patient presents PD or not. Thus, a deeper quantitative analysis of the results must be carried out if the binary results from multiple voice-based tests are contradictory.

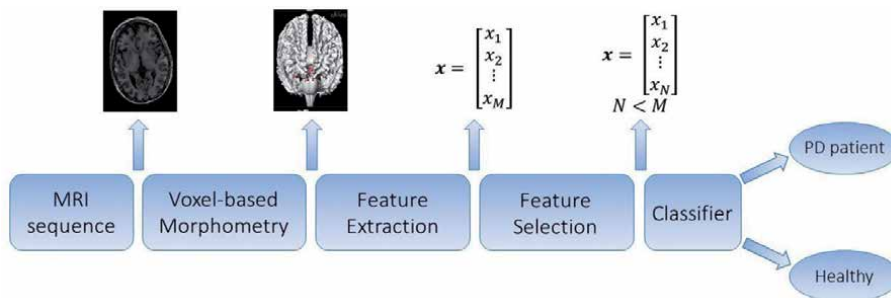
For this reason, the work from Solana-Lavalle et al. [19], provides an analysis of the most important features used to classify a subject as a PD patient or control. By using *Principal Component Analysis* (PCA), the features, with the highest contribution to the detection of PD, were obtained and analyzed. It was found that the features which explained better the diagnosis result, for the case of female subjects, are related to higher frequencies, such as the 32nd and 33rd TQWT coefficients. On the other hand, for the case of male subjects, it is found that features, with the highest contribution to PD detection, are related to lower frequencies such as the fifth TQWT coefficient. The mean and the standard deviation of the most important features were computed for the PD.

Patients and controls and a comparison (PD patients vs. control) is done by using box-plots. According to the box plots, it is shown that there is a clear separation between both groups in most cases. This analysis could help the physician during the interpretation of a binary result to understand how much affected voice is, and the likeness that a patient belongs to one group or the other.

## 7. Analysis of MRI to assist the diagnosis of Parkinson’s disease

Medical images are an important tool to assist the detection and track the progression of neurodegenerative diseases. For PD detection, *structural Magnetic Resonance Imaging* (sMRI) provides relevant information on the thickness and structure of brain tissues. A quantitative analysis is recommended to assist the visual interpretation of the physician [20, 21]. When working with sMRI, some parameters should be taken into account including the strength of the magnetic field (measured in teslas), contrast, noise, relaxation times (T1 and T2), among others. These factors may vary depending on the characteristics of the equipment. The work by Solana et al. [6] aims to identify the regions of the brain that are affected by the disease. It shows how different regions of the brain contribute to the classification, depending on the gender of the patient, and the strength of the magnetic field (1.5 T or 3 T).

*Voxel-based morphometry* (VBM) is a technique to determine the differences in local concentrations of gray matter by comparing MRI voxels between two templates or atlases, where an atlas or template represents a group of subjects. For the cases of PD detection, one group corresponds to PD patients and the other to controls. To apply VBM, images are extracted from multiple individuals, then these images are registered and integrated to generate a brain atlas that represents that particular group of individuals. This study is useful since PD patients are



**Figure 2.**  
Main stages of VBM-based PD detection from MRI.

characterized by a decrease in gray matter volume when compared with controls. The motivation for applying VBM to PD detection is to identify regions of interest for subsequent classification.

According to reported research efforts, VBM-based PD detection from MRI consists of the following stages: (1) VBM to identify regions of interest, (2) feature extraction from regions of interest, (3) selection of the most relevant features for subsequent classification of regions of interest, (4) classification, and (5) performance assessment. The different stages for VBM-based PD detection are shown in **Figure 2**.

## 8. Dataset description

To conduct VBM-based PD detection on MRI on separate datasets, one for men and another for women, the largest publicly available collection of MR images is the *Parkinson's Progression Makers Initiative* (PPMI) dataset. This dataset is the result of collecting clinical data (including images) for PD research around the world. Clinical data includes genomics, patient data, and imaging data. The PPMI dataset is publicly shared to accelerate research discoveries to assist the treatment and diagnosis of PD. PPMI's T1-weighted MR images have been applied to VBM-based PD detection. T1-weighted MR images were generated by using a 1.5–3 T scanner with (1) a scanning time between 20 and 30 min, (2) a slice thickness of 1.5 mm or less, and (3) at three different views, axial, sagittal and coronal. The MR images were obtained from 226 men with PD, 86 male controls, 104 women with PD, and 64 female controls.

## 9. Classifiers

For classification over the regions of interest, detected with VBM, texture information is very useful since the measurement of texture requires statistical analysis to determine how voxel intensity values are distributed. Texture measurement involves the computation of the first-order and second-order statistics of the regions of interest.

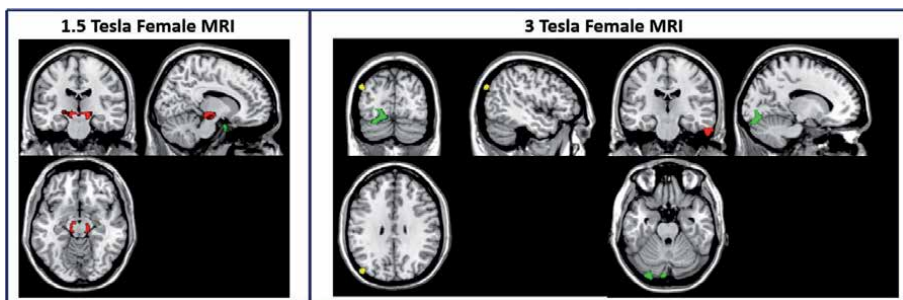
The number of features extracted from one atlas is very large because of the number of regions of interest, the number of directions for second-order statistics features (co-occurrence matrix), number of views (sagittal, axial, coronal), number of different first-order statistics features, and number of second-order statistics features. Thus, *Principal Component Analysis* and *Wrappers* were applied to detect the most relevant features for the classification of regions of interest.

## 10. Regions of interest

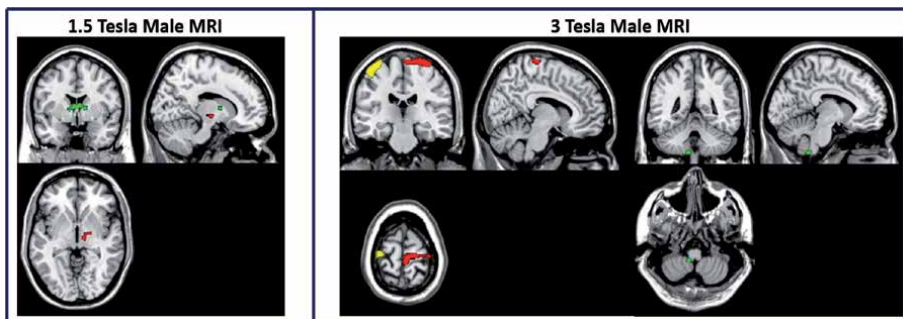
From the results of applying VBM to brain MRI, it has been found that regions of interest for PD detection in men are the *basal ganglia*, *brainstem*, *fourth ventricle*, *lateral ventricle*, *cerebellum*, *frontal lobe*, *temporal lobe*, *putamen*, and *thalamus*. The generation of signals for involuntary movement and instincts is generated within the putamen and thalamus. Other regions of interest lie in the upper cortex, which is related to brain functions such as reasoning, decision making. On the other hand, the application of VBM to female brain MRI shows that regions of interest, for PD detection in women, are *occipital lobe*, *basal ganglia*, a small part of the *cerebellum*, *frontal lobe*, *thalamus*, *brainstem*, and *temporal gyrus*. The last three regions are

associated with visual stimuli processing and spatial awareness. Regions of interest, within the cortex area, are not as large as those in men. These results are significant since most works, for automated PD detection, have been focused on the *striatum region* of the brain to detect damage. Another finding is that regions of interest in men are bigger than those in women, which agrees with medical findings that state that men are more prone to PD than women by almost twice. It is also found that the number of regions of interest are more in women than in men and that regions of interest, in men and women, are generally scattered over the same brain zones. Regions of interest, in men, are found within areas where more and smaller regions of interest for women occur. The number of selected features for PD detection in women is more reduced than in men.

Another finding from Solana et al. [6] is that the regions of interest from which most features are selected for PD detection, vary if the image is acquired with a different magnetic field. When the scanner uses 1.5-T for obtaining the MRI images, the features from the striatum region of the brain were chosen for the classification algorithm. On the other hand, when 3-T MRI are analyzed, features from regions like the primary *somatosensory cortex*, the *cerebellum*, and *temporal lobe* are selected as it is shown in **Figures 3** and **4**. Detection of PD with MRI achieves good performance with both genders and magnetic fields. When classifying female patients' MRI, accuracies of 96.77% and 93.28% for 1.5 and 3 T respectively. For male patients' MRI, excellent results were obtained, with 99.01% and 95.56% accuracy for 1.5 and 3 T respectively. **Table 2**, shows the results obtained by different methods in recent years, and how they compare to the proposed work.



**Figure 3.** The most relevant regions of interest for PD detection in 1.5 and 3 T MRI of female patients, are highlighted with colors red, yellow, and green.



**Figure 4.** The most relevant regions of interest for PD detection in 1.5 and 3 T MRI of male patients, are highlighted with colors red, yellow, and green.



Author, year	Dataset	Performance
Long et al., 2012 [22]	MRI from 19 PD patients and 27 healthy subjects	Accuracy 86.96%, sensitivity 78.95%, specificity 92.59%
Lei et al., 2018 [23]	PPMI MRI dataset	Accuracy 86.48%
Sivaranjini et al., 2020 [24]	CNN	Accuracy 88.9%
Esmailzadeh et al., 2018 [25]	PPMI MRI dataset and personal information (age, gender)	Accuracy 100%
Shah et al., 2018 [26]	PPMI MRI dataset	Accuracy 93%
Salvatore et al., 2014 [27]	MRI from 28 PD patients and 28 healthy controls	Accuracy, sensitivity and specificity above 90%
Shinde et al., 2019 [28]	Neuromelanine MRI	85% of accuracy
Amoroso et al., 2018 [29]	PPMI MRI dataset	Accuracy 93%, sensitivity 93%, specificity 92%
Proposed method	PPMI MRI dataset	Accuracy 99.01% (men) and 96.97% (women), sensitivity 99.35% (men) and 100% (women), specificity 100% (men) and 96.15% (women)

**Table 2.**  
 A comparison between different works on PD detection based on MRI.

## 11. Conclusions

Parkinson's disease (PD) detection is an active area of research. These efforts are oriented to assist the provision of a better quality of life for PD patients. Vocal-based detection of PD is a non-invasive and inexpensive alternative for the early detection of the disease. According to neurology studies, the female brain and the male brain are functionally different and this is the motivation to conduct separate studies, according to gender. Fortunately, the availability of large datasets allows such research efforts. The work by Solana-Lavalle et al. [8, 19] is based on the largest publicly available dataset to train and test different classifier so that separate studies for male and female patients are carried out. Experiment results show that the most relevant features for accurate classification are highly dependent on gender. In the case of male patients, low-frequency voice content is the most significant, while for female patients, high-frequencies give better results. Most features selected in the feature selection process are extracted by using the *Tunable Q-factor Wavelet Transform* (TQWT) and the *Mel Frequency Cepstral Coefficients* (MFCC). Both groups of features are obtained through the use of banks of filters, where these extraction mechanisms operate in a similar way the human auditory system does. The accuracy obtained by the classifying algorithms reaches up to 95.9%, showing the best results with the male population. Also, a statistical analysis of the variability of the most significant features, from each gender, is done to assist the clinical interpretation of the classification result (PD positive and PD negative).

Another method to detect neurological alterations is through medical images such as *Magnetic Resonance Imaging*, *DaTscan*, and *Diffusion Tensor Imaging*. Physicians have used these images modalities to help diagnose PD. However, they rely on visual inspection, which is prone to misdiagnosis due to human error. For this reason, a quantitative analysis of these images is suggested. Solana et al. [6] proposed a method for using structural MRI combined with signal processing and machine learning classifiers to assist the diagnosis of PD. This method achieves

competitive results and insights. The classification results deliver an accuracy of 99.01% in male patients and 96.97% in female patients.

*Voxel-Based Morphometry* is a statistical study that has been used to identify brain regions that show differences between PD patients and controls. Features, based on first-order (histogram) and second-order statistics (co-occurrence matrix), have been extracted from the regions of interest identified by VBM. Since the number of features, extracted from multiple regions of interest, is very large, feature selection techniques have to be used such as *wrapper for feature selection*. The aim of using *feature subset selection* is to identify the most important features for discrimination and to reduce computational complexity. Regions of interest for PD detection usually include the striatum. However, by using feature subset selection it has been possible to identify several regions, outside the striatum, suggesting an affection in those areas of the brain. These regions include the *somatosensory cortex*, *temporal gyrus*, and *cerebellum*.

Future work on the detection of PD could make use of other imaging techniques such as *Functional Magnetic Resonance Imaging* and *Diffusion Tensor Imaging*. These imaging techniques provide information about the activity within the brain, and about the connectivity of the brain respectively. Thus, these modalities are good candidates to provide new information about PD and an alternative to assist the physicians with early detection of the disease. On the other hand, some of the best classification results in voice recordings are obtained using *deep learning* techniques which demand the availability of a larger dataset. To the best of our knowledge, deep learning has not been applied to the largest dataset from Sakar et al. [7] and could be an opportunity to compare these new learning techniques with classical approaches.

## Acknowledgements


The authors would like to acknowledge the support of the National Council for Research and Technology (CONACYT) in Mexico (Scholarship 934454 and stimulus 68150).

## Author details

Gabriel Solana-Lavalle\* and Roberto Rosas-Romero\*  
Departamento de Computación, Electrónica y Mecatrónica, Universidad de las Américas Puebla, San Andrés Cholula, Puebla, Mexico

\*Address all correspondence to: gabriel.solanale@udlap.mx  
and roberto.rosas@udlap.mx

## IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Dorsey ER, Constantinescu R, Thompson JP, Biglan KM, Holloway RG, Kieburtz K, et al. Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology*. 2007;**68**(5):384-386. DOI: 10.1212/01.wnl.0000247740.47667.03
- [2] Factor SA, Weiner W. *Parkinson's Disease: Diagnosis & Clinical Management*. 2nd ed. New York City: Demos Medical Publishing; 2007
- [3] Schrag A, Ben-Shlomo Y, Quinn N. How valid is the clinical diagnosis of parkinson's disease in the community? *Journal of Neurology, Neuro-surgery & Psychiatry*. 2002;**73**(5):529-534
- [4] Sakar BE, Isenkul ME, Sakar CO, Sertbas A, Gurgen F, Delil S, et al. Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. *IEEE Journal of Biomedical and Health Informatics*. 2013;**17**(4):828-834. DOI: 10.1109/JBHI.2013.2245674
- [5] Saeed U, Compagnone J, Aviv RI, Strafella AP, Black SE, Lang AE, et al. Imaging biomarkers in parkinson's disease and parkinsonian syndromes: Current and emerging concepts. *Translational Neurodegeneration*. 2017;**6**(1):1-25
- [6] Solana-Lavalle G, Rosas-Romero R. Classification of PPMI MRI scans with voxel-based morphometry and machine learning to assist in the diagnosis of Parkinson's disease. *Computer Methods and Programs in Biomedicine*. 2021;**198**:105793. DOI: 10.1016/j.cmpb.2020.105793
- [7] Sakar CO, Serbes G, Gunduz A, Tunc HC, Nizam H, Sakar BE, et al. A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable q-factor wavelet transform. *Applied Soft Computing*. 2019;**74**:255-263. DOI: 10.1016/j.asoc.2018.10.022
- [8] Solana-Lavalle G, Galán-Hernández J-C, Rosas-Romero R. Automatic Parkinson disease detection at early stages as a pre-diagnosis tool by using classifiers and a small set of vocal features. *Biocybernetics and Biomedical Engineering*. 2020;**40**(1):505-516. DOI: 10.1016/j.bbe.2020.01.003
- [9] Tsanas A, Little MA, McSharry PE, Spielman J, Ramig LO. Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. *IEEE Transactions on Biomedical Engineering*. 2012;**59**(5):1264-1271. DOI: 10.1109/TBME.2012.2183367
- [10] Peker M. A decision support system to improve medical diagnosis using a combination of k-medoids clustering based attribute weighting and svm. *Journal of Medical Systems*. 2016;**40**(116):1-16. DOI: 10.1007/s10916-016-0477-6
- [11] Guruler H. A novel diagnosis system for Parkinson's disease using complex-valued artificial neural network with k-means clustering feature weighting method. *Neural Computing and Applications*. 2017;**28**:1657-1666. DOI: 10.1007/s00521-015-2142-2
- [12] Sakar BE, Serbes G, Sakar CO. Analyzing the effectiveness of vocal features in early telediagnosis of Parkinson's disease. *PLoS One*. 2017;**12**:1-18. DOI: 10.1371/journal.pone.0182428
- [13] Braga D, Madureira AM, Cuelho L, Ajith R. Automatic detection of Parkinson's disease based on acoustic analysis of speech. *Engineering Applications of Artificial Intelligence*.

2019;77:148-158. DOI: 10.1016/j.engappai.2018.09.018

[14] Rizvi DR, Nissar I, Masood S, Ahmed M, Ahmad F. An lstm based deep learning model for voice-based detection of Parkinson's disease. *International Journal of Advanced Science and Technology*. 2020;29(5):8

[15] Vital TPR, Nayak J, Naik B, Jayaram D. Probabilistic neural network-based model for identification of Parkinson's disease by using voice profile and personal data. *Arabian Journal for Science and Engineering*. 2021;46:1-25

[16] Peker M, Sen B, Delen D. Computer-aided diagnosis of Parkinson's disease using complex-valued neural networks and mrmr feature selection algorithm. *Journal Healthcare Engineering*. 2015;6(3):281-302. DOI: 10.1260/2040-2295.6.3.281

[17] Tsanas A, Little MA, McSharry PE, Ramig LO. Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity. *Journal of the Royal Society Interface*. 2011;8(59):842-855. DOI: 10.1098/rsif.2010.0456

[18] Montaña D, Campos-Roca Y, Pérez CJ. A diadochokinesis-based expert system considering articulatory features of plosive consonants for early detection of Parkinson's disease. *Computer Methods and Programs in Biomedicine*. 2018;154:89-97. DOI: 10.1016/j.cmpb.2017.11.010. Available from: <https://www.sciencedirect.com/science/article/pii/S0169260717305151>

[19] Solana-lavalle G, Rosas-romero R. Analysis of voice as an assisting tool for detection of Parkinson's disease and its subsequent clinical interpretation. *Biomedical Signal Processing and Control*. 2021;66:102415. DOI: 10.1016/j.bspc.2021.102415

[20] Hotter A, Esterhammer R, Schocke MF, Seppi K. Potential of advanced mr imaging techniques in the differential diagnosis of parkinsonism. *Movement Disorders*. 2009;24(S2):S711-S720

[21] Heim B, Krismer F, De Marzi R, Seppi K. Magnetic resonance imaging for the diagnosis of Parkinson's disease. *Journal of Neural Transmission*. 2017;124(8):915-964

[22] Long D, Wang J, Xuan M, Gu Q, Xu X, Kong D, et al. Automatic classification of early Parkinson's disease with multi-modal MR Imaging. *PLoS One*. 2012;7:1-9. DOI: /10.1371/journal.pone.0047714

[23] Lei H, Zhao Y, Wen Y, Luo Q, Cai Y, Liu G, et al. Sparse feature learning for multi-class Parkinson's disease classification. *Technology and Health Care : Official Journal of the European Society for Engineering and Medicine*. 2018;26(1):193-203. DOI: 10.3233/thc-174548

[24] Sivaranjini S, Sujatha CM. Deep learning based diagnosis of Parkinson's disease using convolutional neural network. *Multimedia Tools and Applications*. 2020;79:15467-15479. DOI: 10.1007/s11042-019-7469-8

[25] Esmaeilzadeh S, Yang Y, Adeli E. End-to-end Parkinson disease diagnosis using brain MR-images by 3D-CNN. *Computer Vision and Pattern Recognition*. 2018:1-7

[26] Shah PM, Zeb A, Shafi U, Zaidi SFA, Shah MA. Detection of Parkinson disease in brain MRI using convolutional neural network. In: 2018 24th International Conference on Automation and Computing (ICAC). Manhattan, New York, US: IEEE; 2018. pp. 1-6

[27] Salvatore C, Cerasa A, Castiglioni I, Gallivanone F, Augimeri A, Lopez M,

et al. Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and progressive supranuclear palsy. *Journal of Neuroscience Methods*. 2014;**222**:230-237. DOI: 10.1016/j.jneumeth.2013.11.016

[28] Shinde S, Prasad S, Saboo Y, Kaushick R, Saini J, Pal PK, et al. Predictive markers for Parkinson's disease using deep neural nets on neuromelanin sensitive MRI. *Neuroimage: Clinical*. 2019;**22**:101748. DOI: 10.1016/j.nicl.2019.101748

[29] Amoroso N, La Rocca M, Monaco A, Bellotti R, Tangaro S. Complex networks reveal early MRI markers of Parkinson's disease. *Medical Image Analysis*. 2018;**48**:12-24. DOI: 10.1016/j.media.2018.05.004



# A Systematic Review of Sensitivity Analysis of Activated Sludge Modeling

*Rafael Andrés Borobio-Castillo,*

*José Manuel Cabrera-Miranda, Alberto Vargas-Hidalgo*

*and Benito Corona-Vásquez*

## Abstract

There are a series of sensitivity analysis performed around activated sludge models for wastewater treatment. Comparison is presented both for local and global approaches, and the most used methods are reported. It is observed that sensitivities depend on the modeling objectives. Furthermore, local methods are applicable only for linear models, thus, the global ones are often preferred. Due to the current wastewater resource recovery trend, more sensitivity analysis regarding phosphorus removal and model refinement will be required. Finally, knowledge gaps are identified in association with uncertainty in the influent fractions, and variance-based methods for factor interaction. The sensitivity analyses are quality assurance tools that, if applied properly, it is expected to improve complex phenomena understanding as well as decision making.

**Keywords:** activated sludge models (ASM), benchmark simulation model (BSM), membrane bioreactor (MBR), uncertainty, sensitivity analysis, local sensitivity analysis (LSA), global sensitivity analysis (GSA)

## 1. Introduction

Disposal of urban wastewater (WW) with adequate treatment is a major concern in developing countries. In most of them, a considerable amount of WW is discharged into the environment (rivers, lakes, and oceans) as raw WW or poorly treated WW. Consequently, surface water and groundwater get polluted, affecting human health, aquatic ecosystems, food production, and drinking water availability [1]. Thereupon, it is vital to treat wastewater to mitigate the environmental impact.

Wastewater treatment plants (WWTP) are infrastructure dedicated to water sanitation. The most commonly applied process is activated sludge, a biological treatment consisting of a bioreactor coupled with a secondary settler. Within the bioreactor, biomass (heterotrophic, autotrophic, and/or phosphorus accumulating) is synthesized for biodegradation of the pollutants as well as for the removal of nutrients suchlike Nitrogen and Phosphorus [2]. Then, the secondary settler concentrates the biomass for its removal and further solid treatment. Finally, a clarified effluent result after the treatment.

However, in developing countries, most WWTPs only aim for primary (physical) and secondary (biological) treatment, without tertiary treatment nor sludge treatment (anaerobic digestion) [3]. Hence, the lack of advanced treatment techniques as well as inefficient operation/control of the WWTPs results in increased water pollution.

Lately, it has surged a trend to conceive wastewater treatment plants as water resource recovery facilities (WRRF). This is because it is possible to recover organic matter, nutrient-rich by-products, energy, and water itself, representing an economic revenue for the WRRF [4]. Consequently, there is a need for designing new infrastructure and for process optimization to meet stringent water quality standards together with resource recovery.

Either for design or diagnosis, it is vital to consider the processes influencing the WWTP performance. In the AS process, it is governed by the interaction of raw wastewater fluctuations (quality and quantity), the biokinetics, the mixing conditions, the aeration system, together with the operational conditions [5]. Due to process complexity, mathematical models have arisen as an ideal tool for assessment of the AS performance, allowing to provide continuous feedback in an understandable, faster, and cheaper manner.

Over the past few decades, process models have been established for designing, upgrading, and optimizing wastewater treatment plants [4]. In the wastewater industry, specifically for the biological processes area, the activated sludge models (ASM) were introduced for the latter, given its capability to proximately simulate the process kinetics taking place in the bioreactor in a simpler fashion. Mind that the ASMs, are deemed as core models, i.e., that can be modified according to modelers' needs. While benchmarking frameworks-also known as BSMs-have been proposed to assess environmental and economic aspects in an activated sludge plant-wide context [6]. Moreover, recently membrane bioreactor (MBR) models-activated sludge process plus membrane filtration have surged as an alternative for meeting stringent water regulations and for resource recovery of water given its high-quality effluent post-membrane treatment [7, 8]. Mind that the above-mentioned activated sludge modeling frameworks have been developed by various modeling task groups of the International Water Association (IWA). Thus, the group of them will be referred to as the IWA models.

According to Saltelli et al. [9], building any kind of model requires specifying model archetype, parameters, resolutions, and calibration data including its acceptance criteria, and so on. Nevertheless, sometimes information and data are missing or are not well-known, resulting in uncertainty in each of the previous requirements. Hence, model implementation highly depends on the understanding of the AS process, as lack of it results in augmentation of model uncertainty.

Mind that model applicability relies on how proximate model inputs and outputs are to the real-plant data. Therefore, appropriate modeling practices based on high-quality data collection and model calibration are essential. According to Rieger et al. [10], a good modeling practice (GMP) of the activated sludge process consists of 5 phases: project definition, data collection and reconciliation, plant model set-up, calibration, and validation, as well as scenario simulation. Inherently, uncertainty is present in the input data required in each phase (e.g., influent flow rate, pollutant fluxes, seasonal conditions, model parameters), that if not heed and reduced, will spoil model applicability.

Consequently, Belia et al. [11] stated the importance of identifying the sources of uncertainty in WWTP modeling for project risk reduction and model validation. Thereby, identification and classification of the sources of uncertainty as input data (influent, operational settings, etc.), model data (e.g., structure and process interaction), model parameters (hydraulic, biokinetic, settling), and technical aspects



(solver setting and computational thresholds) within each of the GMP phases is strongly recommended [11]. After sources identification, an uncertainty analysis is to be conducted. It consists of propagating the model input-also called input factor-uncertainty in the desired model output(s) via Monte Carlo simulations or by probabilistic methods [9, 12]. Hence, determining probabilistic distributions of the model output given uncertain input factors.

Nevertheless, the IWA models (including model refinements) are often over parameterized. Hence, a vast number of input factors may be uncertain, troubling calibration principally of complex models. Therefore, after an uncertainty analysis, a sensitivity analysis is conducted for quantifying how much uncertainty is related to an individual input factor or a group of them [13]. So, a sensibility analysis (SA) is a method used to characterize and prioritize uncertainty. According to Al et al. [14], and SA can be used as a quality assurance technique for modelers as it improves promotes a better understanding of the activated sludge model behavior.

There are a series of sensitivity analysis performed around activated sludge models for wastewater treatment. These can be classified according to their nature, is to say, local or global. Local approaches-also called OAT approaches or LSA-assess parameters sensitivity in the function of partial derivatives of the outputs given small perturbations of an input factor for control/identification of problems [15]. Nevertheless, local approaches are fiercely criticized due to method limitations such as linearity, normality assumptions, and local variations of the input space [16]. Still, there is a significant amount of local sensitivity analysis (LSA) in the activated sludge modeling field.

To overcome the limitations of LSAs, global approaches have been established for the assessment of the entire domain of the input space of parameters variation. Global sensitivity analysis (GSAs) methods can be deemed as an analysis of variance (ANOVA), thus, it fractions variance among the uncertain input factors to elucidate its influence in model output [9, 15]. Therefore, unlike local approaches, the GSAs allow studying mathematical models as a whole, even, some methods account for the effect of factor interaction [14, 16]. Fortunately, over the past years, most activated sludge modelers have taken previous considerations and conducted several GSA methods to reduce model uncertainty in predicting system performance.

However, either for local or global approaches, in the activated sludge modeling area, there has been published a wide range of sensibility analysis methods, along with a different focus, i.e., for SA method introduction or application of it. The latter under different modeling goals and their respective scenarios demonstrating the applicability of the sensibility analysis in the field. Consequently, due to the advantages of the SAs, together with the need for complex models for improving process understanding, it is expected that more AS stakeholders will rely on sensitivity analysis results for process design, control, and upgrade.

Yet, up to the author's knowledge, there is not a systematic review concerning the sensibility analysis around IWA models. Hence, according to the statements above mentioned, the objective of this review is to (I) compare the sensitivity analysis performed in the IWA models, distinguishing them from local and global approaches, (II) report the used method, (III) look for similarities and misinterpretations found in the reviewed papers, (IV) determine if the purpose was developing a methodology or the sensitivity analysis were used for an application, (V) catalog the papers according to the aim of the papers (e.g., control, operation, etc.), and (VI) demonstrate lacunae in knowledge concerning the sensitivity analysis in the IWA models.

According to these objectives, this paper presents the collective effort of the authors to collect the up-to-date most relevant works in the activated sludge modeling area. We summarize what we consider the most relevant features that

current and future AS modeling practitioners must heed. For improving readers' understanding the paper is divided into six major sections. First, an overview of the IWA models is presented for the readers to become au fait of the activated sludge models, the benchmark simulations models, as well as the membrane bioreactor models. Then, some of the most applied sensitivity analysis methods in the area are presented, principally distinguishing them as local or global approaches. In the third section, we outline the systematic selection of the papers across the activated sludge modeling area. The results of the systematic review are presented in section four. Finally, the results discussion and our main conclusions are reported in sections five and six, respectively.

## 2. The IWA models

Mathematical modeling of activated sludge systems is an optimal technique for WWTP design and operation, human resource training, and research [10]. Therefore, the International Waster Association (IWA) has developed activated sludge models together with benchmark models for assessing control strategies of the AS process even for a plant-wide context including a primary treatment and sludge digestion [5, 6]. Moreover, due to the advantages to treat and reuse wastewater, membrane bioreactor (MBR) models have gained attention [17]. Therefore, the activated sludge models (ASM), benchmark simulation models (BSM) together with those for membrane bioreactor modeling are briefly discussed below.

### 2.1 Activated sludge models (ASM)

The activated sludge models were introduced in the 1980s with the core model known as ASM1 [10]. Its main purpose was to assess the activated sludge process utilizing simple relationships to mimic the biokinetics occurring within the bioreactor. It consists of a set of biokinetic rates for biological WW treatment based on Monod-like equations (Eq. (1)) for particulate and soluble compounds or state variables (denoted by  $S$  and  $X$ , respectively) [5].

$$\frac{S}{K_s + S} \text{ or } \frac{X}{K_x + X} \quad (1)$$

Since the introduction of the ASM1, many attempts were made to improve the model's capability for reproducing biological nutrient removal. **Table 1** presents a brief overview of the most applied ASM developed by the IWA. Mind other models have been developed, however, only the most applied and those strictly related to biological wastewater treatment were considered. Notice the models have different scopes together with a variation in the number of model parameters. Moreover, it is important to notice the similarity in the overall process among the ASM1 and ASM3 as well as for ASM2d and ASM3 BioP. However, these differ from the state variables to be modeled and their parameters.

For example, the ASM3 was developed to deal with the ASM1 limitations concerning the kinetics for nitrogen and alkalinity of heterotrophic microorganisms [5]. While the ASM2d consists of a modified version of the ASM2 (not included in **Table 1**), as ASM2 does not consider denitrification due to phosphorus accumulating organisms (PAOs), together with the glycogen storage as carbon storage for PAOs [5]. Finally, the ASM3 BioP adds biological phosphorus removal to the ASM3. It differs from ASM2d as it does not include P chemical precipitation (easily implemented), the use of endogenous respiration rates, lower rates for anoxic rates

Model	Overall process	State variables	# of parameters	# of processes	Reference
ASM1	C/N	13	23	8	[5]
ASM2d <sup>a</sup>	C/N/P	19	74	21	[5]
ASM3	C/N	13	46	12	[5]
ASM3 BioP	C/N/P	17	83	23	[18]

*C, carbon removal; N, nitrogen removal; P, phosphorus removal.*  
<sup>a</sup>Fermentable COD fractions state variables.

**Table 1.**  
 Overview of the activated sludge models.

(compared to aerobic ones), as well as neglecting fermentation [18], so, influent fractionation becomes simpler.

As the activated sludge models are core models, i.e., this can be subjected to refinements to meet modelers' needs, usually for representing the AS process more accurately. For example, to surpass the limitation of the ASM3 of modeling nitrification in a single-step process ( $S_{NH_4} \rightarrow S_{NO_3}$ ), Iacopozzi et al. [19] developed a two-step nitrification process ( $S_{NH_4} \rightarrow S_{NO_2} \rightarrow S_{NO_3}$ ). Hence, they were able to represent the separation of autotrophic biomass into ammonia and nitrite oxidizers within their model. Other modifications have been made for portraying microbial processes in detail [20–22], including modeling the AS process in an MBR scheme, discussed later.

However, the ASMs (including refinements) were developed for assessing the efficiency of the bioreactor. Consequently, the need for a modeling framework that couples the bioreactor, the secondary clarifier, as well as sludge treatment together with key performance indicators, among other features, result in the development of the benchmarking simulation models.

## 2.2 Benchmark simulation models (BSM)

The benchmark simulation model No 1 (BSM1) is a model framework for evaluation using performance indexes of an AS process based on simulations of the WWTP [23]. The BSM1 models the bioreactor following the ASM kinetics and dividing the reactor into anaerobic, anoxic, and oxic (aerobic) phases according to the AS model being reproduced. The secondary clarifier is modeled using the Takács settling model [24]. Also, the BSM1 considers recycling flowrates as well as the wastage and return of the activated sludge to the system.

Moreover, the BSM1 framework allows the modeler to assess plant performance by measuring the effluent quality index (EQI, in kg pollution units  $d^{-1}$ ) and an operational cost index (OCI) [6]. The EQI is a measure of the water quality being discharged to the environment. It sums the main effluent pollutant fluxes (BOD<sub>5</sub>, COD, TKN, NO<sub>x</sub>-N, and TSS) by employing weighting factors. Like the EQI, the OCI weights the sum of different costs within the system suchlike energy requirements (aeration, pumping, mixing), sludge disposal, external carbon sources, and methane production (income) if available. Nevertheless, it does not provide an operational cost but could be easily calculated.

The BSM1 limits to assess local control strategies for the AS process, without accounting the interactions with the primary and sludge treatment. Consequently, the benchmark simulation model No. 2 (BSM2) was developed as a plant-wide assessment model [6, 25]. Its framework couples the features of the BSM1 with a primary clarifier model [26], and a sludge treatment including anaerobic digester following the ADM1 kinetics [27], a sludge thickener as a dewatering unit.

Consequently, the BSM2 allows for the evaluation of unit process interaction in a wider context than BSM1.

### **2.3 Membrane bioreactor (MBR) models**

According to Judd [7], membrane bioreactors (MBR)-a promising biological-physical technology for WW treatment that couples an AS process with microfiltration (MF) or ultrafiltration (UF)-considerably reduce WWTP footprint and achieve higher effluent quality and reduced sludge yield. Mind that models are fully capable to be simulated within membranes bioreactors (MBR) since both systems are alike from the biochemical engineering aspect [17]. Moreover, it is also possible to modify the BSM frameworks to include membrane bioreactor processes.

However, for modeling the system, the ASMs are compatible with or without modifications. The refinement of the ASMs in MBRs, are extended versions mainly for incorporating the release and degradation of soluble microbial products (SMP) and extracellular polymer substances (EPS) [28–30]. The EPSs are mixtures of organics (proteins, lipids, DNA residuals, etc.) that support bacterial growth in high-density biomass communities (as in MBRs), while SMPs are soluble excreta produced during biomass growth and decay that serve as indicators of substrate consumption and biomass decay rates [17]. According to Hai et al. [17], EPS and SMPs play a major role in membrane fouling as these can adhere to the membrane surface, thus, limiting its permeability.

Moreover, for modeling more precisely the MBR operation, physical sub-models can be coupled. An example of these can be found in Mannina et al. [8], who proposed a physical model for modeling cake deposition, deep-bed filtration, and membrane resistances (to simulate transmembrane pressure and resistance variations, e.g., pore fouling, sludge cake, among others). Mind that the model includes mathematical representations of particles' drag and buoyant forces, particle deposition in membrane probability, biomass and sludge attachment and detachment (including backwashing effect) rates, together with cake deposition, deep-bed filtration, and membrane resistances themselves [8].

## **3. Sensibility analysis**

Sensibility analysis is a tool for modelers for the appreciation of the dependency between input factors and model outputs, allowing them to investigate the relevance of each factor around the outputs [16]. Hence, elucidating which model inputs provoke most of the uncertainty in model outputs according to the studied scenario, usually done via Monte Carlo simulations [15]. Thereby, the SA's scope is factor prioritization together with factor fixing (non-influential), and in some cases, to ascertain factor interactions, for potentially reducing model uncertainty [9].

There are two classifications, local sensitivity analysis (LSA) and global sensitivity analysis (GSA) [13]. However, according to Saltelli et al. [9], local approaches, i.e., varying one factor at a time (OAT), are not recommended when dealing with non-linear models, as this approach does not explore a multi-dimensional space, thus missing important effects such as factor interaction (FI). While the GSA methods do vary all the factors together like in an analysis of variance (ANOVA), thus, informing the modeler about factors' global influence in model output variance [9]. Therefore, local approaches (LSA) and global ones (GSA) are to be briefly explained as follows.

### 3.1 Local sensitivity analysis

A local sensitivity analysis (LSA) is a simple analysis where only one factor (OAT approach) changes value between consecutive simulations [13]. An advantage of this method is that the modeler can determine the influence of the perturbed parameter under a local range in a rapid manner. According to the author's knowledge, the most common LSA method applied for AS modeling is the normalized sensitivity index (NSI) described in Eq. (2).

$$NSI = \frac{\theta \Delta Y}{Y \Delta \theta} \quad (2)$$

where  $\Delta\theta$  and  $\Delta Y$  are the selected are the observed differences of model input and output, respectively. Nevertheless, there seems to be an issue with the NSI method term as sometimes it is called by different names, however, those terms seem to be the same according to Eq. (2). Moreover, some authors have reported other LSA techniques [31–33].

### 3.2 Global sensitivity analysis

Unlike the LSA, global sensitivity analysis does consider the entire probability distributions of the input factors, thus assessing the entire domain of the input space [16]. Global sensitivity analysis methods most widely used can be classified as elementary effects methods, linear regression models, variance-based, as well as derivative-based sensitivity analysis. Consequently, Morris Screening, Standardized Regression Coefficients, Sobol Sensitivity Indices, Extended-FAST, together with derivative-based global sensitivity measures are discussed below, as these were the GSA methods applied in the sensitivity analysis-activated sludge modeling literature of this review.

#### 3.2.1 Morris screening method

The Morris screening method measures the factor's sensitivity by adding up Elementary Effects (EEs), i.e., averaging local measures. An EE (see Eq. (3)) indicates the variation between model output ( $y$ ) predisposed to a factor ( $x_n$ ) perturbation being replicated [29, 30]. Where  $A$  is the model output after perturbation,  $B$  represents no perturbation, while  $\Delta$  is a factor depending on the number of levels of the  $n$ -dimensional  $p$ -level grid  $\{1/(p-1), \dots, 1-1/(p-1)\}$ , comparable with the uncertainty range.

$$EE_i(x_1, \dots, x_n, \Delta) = \frac{y(x_1, \dots, x_{i-1}, x_i + \Delta, x_{i+1}, \dots, x_n) - y(x_1, \dots, x_n)}{\Delta} = \frac{A - B}{\Delta} \quad (3)$$

Morris screening standardizes the model inputs and outputs ( $y$  and  $x_n$ ) according to its mean and standard deviation for measuring the sensitivity indices. The mean ( $\mu$ ) measures the influence of the factor in model output uncertainty, whereas the standard deviation ( $\sigma$ ) determines the factor's influence. For example, high values of  $\sigma$  indicate the output variance is related to non-linearity or interactions. To avoid the effect of opposite signs the EEs are referred to as the absolute mean ( $\mu^*$ ) [34]. Whenever  $\mu^* >$  mean threshold the factors are considered as influential and vice versa. The mean standard error ( $\sigma_i \cdot r^{(-0.5)}$ ) provides information about the factor effect. Whenever the factor lies above or below the threshold line ( $\mu^*_i = 2 \sigma_i \cdot r^{(-0.5)}$ ), its effect is involved to model linearity and interactions, respectively [35]. According to Morris [35], the number of simulations (or replicas) is equal to  $r \cdot (n + 1)$ .

### 3.2.2 Standardized regression coefficients method

The standardized regression coefficients (SRC) are sensitivity measures that fit a first-order linear multivariate model to a scalar output ( $b_0 \dots b_i$ ) of the MCS and correlate the model inputs and outputs ( $y$  and  $\theta_i$ , see Eq. (4)). The quantification of the SRCs ( $\beta_i$ ) is done by scaling the regression coefficients ( $b_i$ ) according to the standard deviation of model input and outputs (Eq. (5)).

$$y = b_0 + \sum_i b_i \cdot \theta_i \quad (4)$$

$$\beta_i = \frac{\sigma_{\theta_i}}{\sigma_y} \cdot b_i \quad (5)$$

According to Saltelli et al. [13],  $\beta_i^2$  are deemed as output variance contributors or as a first-order sensitivity index ( $S_i$ ) as long as the coefficient of determination is high enough to imply model linearity ( $R^2 \geq 0.7$ ). They also mentioned  $\beta_i$  values range between  $-1$  and  $+1$ , where a high absolute value indicates a large effect in output variance (sign indicates positive and negative effects, and close-to-zero values indicate negligible effects). However, the SRC method does not measure the effect of factor interactions in output variance diminishing the reliability of the results. Still, it is a rapid method useful for first approximations on sensitivity measures. Sampling is usually done by Latin hypercube sampling (LHS).

### 3.2.3 Sobol sensitivity indices method

Sobol indices is a variance decomposition method for quantifying the input factor individual effects together with the effect due to factor interaction (FI) [13]. It decomposes the output variance in first-order sensitivity indices ( $S_i$ ) and the total sensitivity indices ( $S_{Ti}$ ).  $S_i$  is the amount of variance extracted from total output variance. It is measured as the conditional variance over the unconditional variance according to the factor uncertainty range (see Eq. (6)). The  $S_i$  represents the factor's contribution to the model output variance. The model will be deemed as linear whenever  $\sum S_i = 1$ , a sum different to 1 indicates model non-linearity. This indicates there is a contribution in output variance due to FI, which can be highlighted by the difference between  $1 - \sum S_i$ .

$$S_i = \frac{V[E(Y|X_i)]}{V(Y)} \quad (6)$$

Total sensitivity indices ( $S_{Ti}$ ) quantify the total effect of the input factor on model output variance (including FI, see Eq. (7)) [13]. Therefore, the strength of the interactions can be assumed by the difference between  $S_{Ti}$  and  $S_i$ , as both indices follow the same linearity or non-linearity principles.

$$S_{Ti} = 1 - \frac{V[E(Y|X_i)]}{V(Y)} \quad (7)$$

Estimating  $S_i$  and  $S_{Ti}$  requires approximately 2000 Monte Carlo simulations per factor [14]. The MCS requires a design of experiments, thus, space-filling sampling methods are recommended to improve estimators' accuracy (e.g., Latin Hypercube sampling or Sobol sequences sampling).

### 3.2.4 Extended FAST

Like the Sobol method, the Extended Fourier Amplitude Sensitivity Test (E-FAST) is a variance decomposition method. It decomposes the variance in first sensitivity indices ( $S_i$ ) and total sensitivity indices ( $S_{Ti}$ ) and determines the influence of FI by the difference between these indices (see Eqs. (6) and (7)). However, E-FAST differs from Sobol as the numerical simulations (around 500–1000 per factor) for computation of the indices are based on a spectral method rather than MCS [36]. It is also, less computationally expensive than the Sobol method given the lesser number of simulations per factor.

### 3.2.5 Derivative-based global sensitivity analysis (1241)

Derivative-based global sensitivity measures (DGSM) is a method that exhibits strong similarities between the Morris screening method and Sobol sensitivity indices, with the advantage of its ease for implementation and numerical evaluation [16]. Previous reasons have not gone unnoticed from practitioners, thus, recently becoming a popular GSA method.

DGSM method combines Morris screening and Sobol capabilities, but attending its major drawbacks [16]: (1) the EEs are added by using a random sampling of the n-p grid and measuring the finite differences when incrementing  $\Delta$  (see Morris method above), thus, EEs cannot be accounted in a range lesser than delta, limiting is accuracy, and (2) Sobol indices are computationally expensive to measure. While DGSM shows a higher convergence rate and more accuracy than Morris, and a lower computational cost in various magnitude orders compared to Sobol [37].

Essentially, DGSM is based on local derivatives suchlike NSI (see Eq. (2)), but as an average of the sensitivity measures evaluated by Quasi-Monte Carlo sampling methods, rather than by point from a fixed grid [16, 37]. For detailed information about the derivative-based global sensitivity measures please refer to Ghanem et al. [16].

## 4. IWA models sensitivity analysis review

To ascertain the sensitivity analysis across the activated sludge modeling field, an extensive literature review was carried out. The review was based on sensitivity analysis in the IWA models including refinements of these. Model modifications possessing increased complexity were discarded. The logic behind the previous statement is that some of these refinements are still not understood in deep to put into a model for simulating scenarios. Also, research outputs that were not cited, excluding those published during the last year (2020) were not included, except for one article that concerned a project continuation. Consequently, guaranteeing only the inclusion of IWA models as well as cited papers.

### 4.1 Selection procedure

The literature search was conducted on the Web of Science (WoS) database-also known as Web of Knowledge (WoK). As this review focus is to evaluate the sensitivity analysis conducted up to date around the IWA models the search query presented in Box 1 was used. The timeline chosen for the search was 2008-present, being the last search was run until 30 December 2020. This period was selected given the discussion taken place in the WWTmod2008 workshop about the importance of dealing with uncertainty for certifying model accuracy [11]. A summary of the conference proceedings surrounding WWTP modeling together with a

structure to identify sources of uncertainty within the facilities can be found at Belia et al. [11].

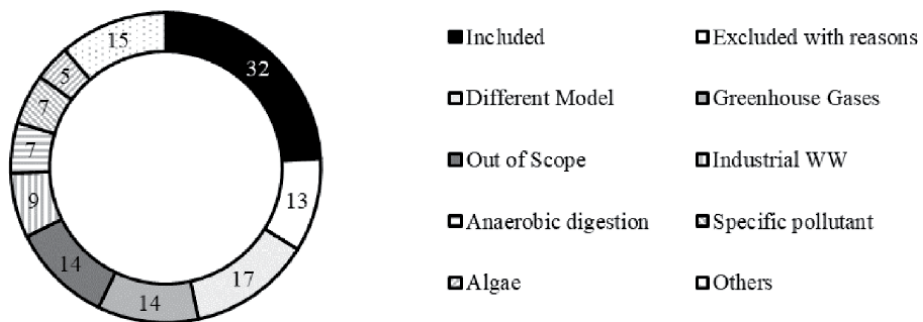
*Sensitivity analysis AND (Activated sludge model OR Benchmark simulation model OR MBR model OR BSM) AND (Wastewater OR membrane)* Covers all articles dealing with sensibility analysis applied to WWTP modeling according to the IWA models, including variations for MBR modeling and microbial process upgrades.

**Box 1.**  
Search query (2008–2020).

A total of 133 research outputs were found in the WoK database matching our search query (see **Box 1**) either in the abstract, title, or keywords. Hence, for determining the studies included within this review, a screening protocol was established for determining the articles' eligibility. Which was conducted by the main author. A graphical summary of the screening process is presented in **Figure 1**. Notice that it provides a classification of the included articles, those excluded by reasons as well as the subject areas and the total of research outcomes for those excluded due to the screening process.

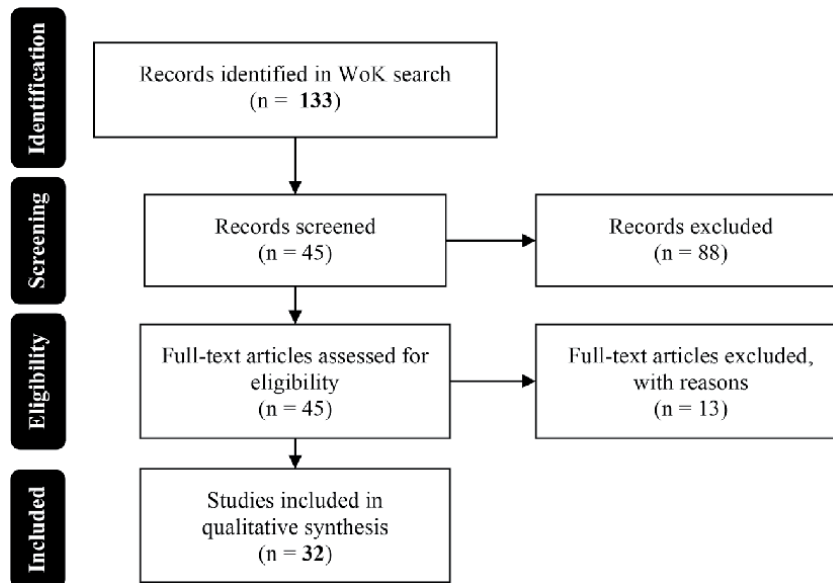
**Figure 2** presents the flow diagram of the screening process followed during this review. It is noticeable that from the total research outcome (133), 88 were discarded due to mismatching the eligibility criteria. According to the information published in **Figure 1**, a total of 17 research groups studied different models such as aeration models MBRs, sludge dewatering units, among others. Even some of these publications conduct a SA, the researchers assessed the model biokinetics via Monod equations rather than using the ASMs. The articles concerning greenhouse gases (GHGs) emissions (14) were discarded as the formation of GHGs is a complex process not clearly enough understood to be modeled [4]. Moreover, now the major concern in developing countries is improving WWTP effluent quality rather than carbon footprint reduction.

As for the anaerobic digestion field (7) together with the one for industrial WW (9) were excluded due to the following. First, the anaerobic digestion itself is not an activated sludge process due to the lack of dissolved oxygen. Even so, plant-wide models as BSM2 consider anaerobic digestion processes but the main purpose is to assess the whole plant performance and the interaction between unitary processes (e.g., primary clarifier, bioreactor, etc.). Nevertheless, sensitivity analysis for these models have been published and might be useful for the interested [38–40]. Second, industrial WW usually possess high COD loads that must be treated anaerobically, and usually recalcitrant pollutants (7) are present, which are also beyond the scope



**Figure 1.**  
Graphical summary of the screening process.





**Figure 2.**  
*Research outcomes screening process flow diagram.*

of this review. While algae-bacterial processes (5) were discarded as still an emerging technology [41]. Finally, the rest were out of the scope (17) deemed as casualties of matching the search query or dealing with sensors, some were unclear, among other features.

As shown in **Figure 2**, From the 45 full-text articles assessed for eligibility, 13 were excluded due to the following reasons. Three (3) did not comply with the desired quality. Mostly because there seem not to be a justification about the scenarios being studied, there was no clarity in the results or provide bare conclusions, and the application of the ASMs was confusing. Additionally, these were not cited. Mind that only the articles published in 2020, rather for one exception, could have no citations given its recent publication. Another three (3) publications use the BioWin General Model (EnviroSim Associates Ltd., Canada)-also called BioWin AS/AD model, an extended version of the IWA Baker & Dold model, coupled with an AD process. Given the complexity of the model, the articles were excluded. Yet, if it is of the readers' curiosity, please refer to Liwarska-Bizukojs and Biernacki [42]. While the rest studied Monod kinetics or Metcalf and Eddy guidelines (4), the AS model was not clear (2), or its purpose was to include specific pollutants for assessing its degradation within an AS process (1). Consequently, a total of 32 articles were included within this review.

#### 4.2 Review criteria

Each paper was review against the following set of criteria:

1. Was an IWA activated sludge model performed? If so, was there any modification of it?
2. Was a benchmark simulation model or an MBR model performed?
3. Was the type of sensibility analysis local or global (i.e., LSA or GSA)?

4. Which was the method studied?
5. Does the focus of the sensibility analysis was the introduction of a method or an application of it?
6. Which were the input factors and the model outputs assessed?
7. Was scenario analysis performed? If so, these are to be reported.
8. What was the scope of the paper? For example, if these aims to develop control and operation strategies, among others.

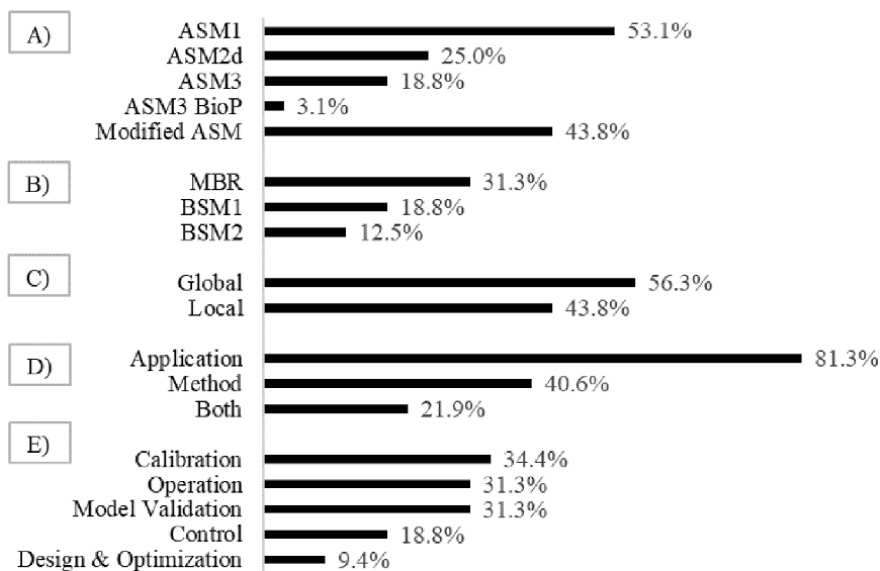
Additionally, notes on each paper were taken for improving the discussion.

## 5. Synthesis of research outcomes

### 5.1 Frequency across review criteria

**Figure 3** overviews the frequency of the results matching the review criteria. For ease of understanding, review features were cataloged according to the ASM studied, modifications for MBR and AS process improvements were grouped, whether a benchmark framework or a membrane bioreactor was applied, the type of sensibility analysis, namely local or global approaches, together with the focus of the sensitivity analysis as a method introduction or an application of it. Finally, the scope of the article and the use of the sensibility analysis.

The ASM1 is the most common model whose parameters have been assessed via SA, followed by the ASM2d, ASM3, and the ASM3 BioP. However, up to 44% of the articles introduce modifications to ASMs, mainly to include soluble microbial products, extracellular polymeric substances, or two-step nitrification processes. Only 32% outputs were related to the benchmark simulation models, as well as for the



**Figure 3.** Features of the included research outputs. (A) Activated sludge model, (B) model framework, (C) analysis type, (D) paper focus, and (E) scope of the article.

Model	ASM refinement			Modeling framework				Type of analysis				Paper focus			Total
	SMP	SMP-EPS	Upgrade	BSM1	BSM1	Mod. MBR	MBR	Global	Local	U	A	M	Both		
ASM1	2	1	2	4	4	2	1	9	7	1	11	3	3	17	
ASM2d	4	-	-	1	-	4	1	6	3	-	4	2	2	8	
ASM3	—	3	2	1	—	2	—	3	3	—	4	—	2	6	
ASM3-BioP	—	—	—	—	—	—	—	—	1	—	1	—	—	1	

U, unclear; A, applications; M, method.

**Table 2.**  
 Summary of study characteristics.

ASM	BSM	MBR	Analysis Method	Method	Focus	Input FAC	Output	Scenarios	Scope	Reference
ASM1	BSM1	○	G	SRC	M	Biokyn-Stoyk-F <sub>Influent</sub>	EQI-OCI-TKN	Comb. of controllers (DO, S <sub>NO3</sub> , S <sub>NH4</sub> ) & ctrl. Strategies (k <sub>L,a</sub> , Q <sub>Innr</sub> , C <sub>Dese</sub> , DO setpoint)	Control	[43]
ASM3-BioP	○	○	L	SI	A	Biokyn-Stoyk-SET <sub>PAR</sub>	EQI-EQSSK	EQI weights (steady-state & dynamic)	Calibration	[44]
ASM1	○	○	L	Spearman Rank	A	Biokyn-Stoyk	TN	Ctrl. operations: • DO-SRT • HRT-R-BCOD:TKN	Control-optimization	[33]
ASM2	○	○	L	NSI	M-A	Biokyn-Stoyk-F <sub>Influent</sub>	COD-S <sub>NH4</sub> -X <sub>TSS</sub> -S <sub>PO4</sub>	Unique	Calibration	[45]
ASM1	BSM1	○	G	SRC	M-A	Biokyn-Stoyk-F <sub>Influent</sub> -HYD & M <sub>trans</sub>	S <sub>NH4</sub> -S <sub>NO3</sub> -TN-SL <sub>PROD</sub> -X <sub>TSS</sub> -AE	Uncertain: • F <sub>Influent</sub> , Biokyn, Stoyk • HYD & M <sub>trans</sub> • All the above	Design	[15]
ASM1 <sup>1</sup>	○	●	G	SRC	A	Biokyn-Stoyk-F <sub>Influent</sub> -MEM <sub>PAR</sub>	MLSS-COD-S <sub>NH4</sub> -TMP-MEM <sub>RES</sub>	Unique	Model validation	[8]
ASM2d	○	○	L	CRS	M	Biokyn-Stoyk	S <sub>NH4</sub>	AM & data frequency: • AM1: k <sub>L,a</sub> • AM2: SCADA-k <sub>L,a</sub> • AM3: αSOTE	Calibration	[31]
ASM1	BSM2	○	G	SRC & PCC	M-A	Biokyn-Stoyk-HYD & M <sub>trans</sub> -SET <sub>PAR</sub>	EQI-OCI-S <sub>NH4</sub>	Solver setting: • Runge-Kutta • CVDOE	Design and control	[46]
ASM3 <sup>1</sup>	○	●	G	E-FAST	M-A	Biokyn-Stoyk	COD-SMP-S <sub>NH4</sub> -S <sub>NO3</sub> -MLSS	SRT: • 10d • 30d • 50d	Model validation-operation	[28]
ASM1	○	○	L	U	A	Biokyn-Stoyk	X <sub>TSS</sub> -COD-TKN	AS & DNR	Calibration	[47]

ASM	BSM	MBR	Analysis Method	Focus	Input FAC	Output	Scenarios	Scope	Reference
ASM1 <sup>1</sup>	○	●	G <sup>2</sup>	A	Biokyn-Stoyk-F <sub>Influent</sub> -MEM <sub>PAR</sub>	MLSS-COD-S <sub>NH4</sub> -TMP-MEM <sub>RES</sub>	Start-up strategy: <ul style="list-style-type: none"> <li>AS inoculum</li> <li>No inoculum</li> </ul>	Calibration-operation	[48]
ASM2d	○	●	L/G	A	HYD & M <sub>trans</sub>	(S <sub>PO4</sub> -TMP-S <sub>NO3</sub> -S <sub>NH4</sub> )/(S <sub>NH4</sub> -S <sub>PO4</sub> -TP-TMP-EQJ-OCI)	Ctrl. strategies: <ul style="list-style-type: none"> <li>Q<sub>WAS</sub></li> <li>C<sub>Dose</sub></li> <li>Relaxation time</li> <li>Flux</li> <li>AER recirculation</li> <li>ANOX recirculation</li> <li>k<sub>La</sub> membrane</li> <li>DO aerobic tank</li> </ul>	Operation	[49]
ASM3 <sup>1</sup>	BSM1	●	L	A	MEM <sub>PAR</sub>	Energy requirement	Energy requirements: <ul style="list-style-type: none"> <li>SAD<sub>filtering</sub></li> <li>SAD<sub>cleaning</sub></li> <li>t<sub>c</sub> (TMP above-30 kPa)</li> </ul>	Control	[50]
ASM2d <sup>1</sup>	○	●	G	M-A	Biokyn-Stoyk-F <sub>Influent</sub> -M <sub>trans</sub> -MEM <sub>PAR</sub>	COD-S <sub>NO3</sub> -S <sub>NH4</sub> -S <sub>PO4</sub> -MLSS-TN	Unique	Calibration	[29, 30]
ASM1	BSM1	○	L	M-A	HYD & M <sub>trans</sub>	OPEX	Disturbances due to weather profiles in: COD, Q, X <sub>TSS</sub> , TN, Temp (°C)	Control-Operation	[32]
ASM2d <sup>1</sup>	○	●	G	M	Biokyn-Stoyk-F <sub>Influent</sub> -MEM <sub>PAR</sub>	COD-S <sub>NH4</sub> -S <sub>NO3</sub> -S <sub>PO4</sub> -MLSS	Unique for each GSA	Calibration	[30]
ASM2d <sup>1</sup>	○	●	G	A	Biokyn-Stoyk-F <sub>Influent</sub> -MEM <sub>PAR</sub>	COD-S <sub>NH4</sub> -S <sub>NO3</sub> -S <sub>PO4</sub> -MLSS	Unique	Model validation-operation	[51]
ASM1	○	○	L	A	Biokyn-Stoyk	ASM state VAR	Noise in data: <ul style="list-style-type: none"> <li>Original</li> <li>X4</li> </ul>	Calibration	[52]

ASM	BSM	MBR	Analysis Method	Focus	Input FAC	Output	Scenarios	Scope	Reference
ASM1	BSM2	O	G	Morris & SRC	A	Biokyn-Stoyk-SET <sub>PAR</sub> (1 & 2 order)	CH <sub>4</sub> PROD-SL <sub>PROD</sub> -OCI-EQI- X <sub>TSS</sub> -S <sub>NH4</sub>	Calibration	[53, 54]
ASM1	BSM1	O	G	SRC	A	Biokyn-Stoyk-F <sub>Influent</sub> - SET <sub>PAR</sub> (1st & 2nd order)	X <sub>TSS</sub> -SRT-SBH-S <sub>NH4</sub> - S <sub>NO3</sub> -TN-SO2-OTR-AE	Calibration	[54]
ASM3 <sup>1</sup>	O	O	G	DGSM	M-A	Biokyn-Stoyk	ASM state VAR	Calibration	[37]
ASM2d <sup>1</sup>	O	●	G	SRC	A	Biokyn-Stoyk-F <sub>Influent</sub> - M <sub>trans</sub> -MEM <sub>PAR</sub>	COD-S <sub>NO3</sub> -S <sub>NH4</sub> - S <sub>PO4</sub> -MLSS-TN	Operation	[55]
ASM1 <sup>1</sup>	O	O	U	SSF	M	Biokyn-Stoyk-M <sub>trans</sub>	COD-TKN-TN-VSS-S <sub>NH3</sub> - MeOH	Model Validation	[20]
ASM3 <sup>1</sup>	O	O	L	RSF	A	Biokyn-Stoyk	COD-S <sub>NH4</sub> -SMP-EPS-OUR- X <sub>STO</sub>	Model Validation	[21]
ASM2d	BSM1	O	G	Morris	A	HYD & M <sub>trans</sub>	EQI-OCI-OQI-TVOP-S <sub>NH3</sub> - DO	Control- Operation	[56]
ASM1 <sup>1</sup>	O	O	L	RMSSA	A	Biokyn-Stoyk	COD-TN-EPS-SMP	Model validation	[57]
ASM3	O	O	L	SC	A	Biokyn-Stoyk-M <sub>trans</sub>	OUR-AUR-NUR	Operation	[58]

ASM	BSM	MBR	Analysis Method	Focus	Input FAC	Output	Scenarios	Scope	Reference
ASM1 <sup>1</sup>	○	○	L	A	Biokyn-Stoyk-Logistic	COD-S <sub>NH4</sub> -S <sub>NO2</sub> -S <sub>NO3</sub>	COD to N: <ul style="list-style-type: none"> <li>• COD:N = 6</li> <li>• COD:N = 9.8</li> </ul>	Model validation-operation	[22]
ASM1	BSM2	○	G	M	Biokyn-Stoyk-F <sub>Influent</sub> -HYD & M <sub>Trans</sub>	CH <sub>4</sub> PROD-SL <sub>PROD</sub> -EQI-S <sub>NH4</sub> -S <sub>NO3</sub> -AE	Uncertainty: <ul style="list-style-type: none"> <li>• F<sub>Influent</sub></li> <li>• Biokyn &amp; Stoyk</li> <li>• HYD &amp; M<sub>Trans</sub></li> <li>• All the above</li> </ul>	Model validation-operation	[14]
ASM3 <sup>1</sup>	○	○	G	M	Biokyn-Stoyk	ASM state VAR	Unique	Model validation	[59]
ASM1	BSM2	○	G <sup>2</sup>	A	Biokyn-Stoyk	BOD-X <sub>TSS</sub> -TN	Flood	Calibration	[12]
ASM1	○	●	L	A	F <sub>Influent</sub> -HYD	ASM state VAR	SRT: <ul style="list-style-type: none"> <li>• 10d</li> <li>• 30d</li> <li>• 50d</li> </ul>	Model validation	[60]

1, modified model; 2, absent; A, applications; AE, aeration energy; AER, aerobic tank; AM, aeration model; ANOX, anoxic tank; ASM, activated sludge model; AUR, ammonia uptake rate; BCOD, biodegradable COD; Biokyn, biokinetics; BOD, bio-chemical oxygen demand; BSM, benchmark simulation model; C<sub>DOSSES</sub>, carbon dosage; CH<sub>4</sub>, methane; COD, chemical oxygen demand; CRS, central relative sensitivities; CSTR, completely stirred tank reactor; DO, dissolved oxygen; DNR, denovo nutrient removal; E-FAST, extended Fourier amplitude sensitivity test; EPS, extracellular polymeric substances; EQ, effluent quality index; EQSSK, effluent quality index South Korea; FAC, factors; F<sub>Influent</sub>, influent fractions; G, global; HRT, hydraulic retention time; HYD, hydraulics; k<sub>L</sub>a, aeration coefficient; L, local; M, method; MBR, membrane bioreactor; MEM, membrane; MeOH, methanol; MLSS, mixed liquor suspended solids; M<sub>Trans</sub>, mass transfer; NSI, normalized sensitivity index; NUR, nitrate uptake rate; OAT, one at the time; OCI, operational cost index; OPEX, operational expenditure; OQJ, overflow quality index; OTR, oxygen transfer rate; PAR, parameter; PCC, partial correlation coefficients; PROD, production; Q<sub>WAS</sub>, internal recycle flow; Q<sub>WAS</sub>, waste activated sludge flow; R, recycle ratio; RES, resistances; RMSSA, root mean square sensitivity analysis; RSF, relative sensitivity function; SAD, coarse bubble aeration intensity; SBH, sludge blanket height; SBR, sequential batch reactor; SC, sensitivity coefficient; SCADA, system control and data acquisition; SET, settler; SI, sensitivity index; SL, sludge; SMP, soluble microbial products; S<sub>NH3</sub>, ammonia; S<sub>NO2</sub>, nitrite; S<sub>NO3</sub>, nitrate; SOTE, standard oxygen transfer efficiency; S<sub>PO4</sub>, phosphate; SRC, standardized regression coefficient; SRT, solid retention time; S<sub>s</sub>, readily biodegradable substrate; Stoyk, stoichiometric; SSF, scaled sensitivity factor; t<sub>c</sub>, cleaning time; TKN, total Kjeldahl Nitrogen; TMP, transmembrane pressure; TN, total nitrogen; TP, total phosphorus; TVOF, total overflow volume; U, unclear; VAR, variables; VSS, volatile suspended solids; X<sub>S70</sub>, COD storage material; X<sub>TSS</sub>, total suspended solids.

**Table 3.**  
 Synthesis of the results for each of the included scientific articles.

membrane bioreactor scheme. Moreover, local, and global approaches seem to be balanced with a slighter preference for and global sensitivity analysis, even of the disadvantages of local approaches. The good news is that 81% of the papers highlighted the applicability of the sensibility analysis for activated sludge modeling. Finally, the general article scopes are also balanced in terms of model validation and calibration, as well as studying operation strategies, while control strategies and design scopes were less studied. Mind that model validation refers to the introduction of a model or method and usually includes calibration of it.

## **5.2 Study characteristics and individual results**

**Table 2** reports in more detail the occurrence of the activated sludge models within the sensitivity analysis field. It is evident that modelers aim for nitrogen removal either with the ASM1 or ASM3. As these models are simpler in all criteria (see **Table 1**), and the fact that most treatment facilities' objective is to remove carbon and nitrogen might be the reason for the imbalance between phosphorus removal models.

An interesting result was that 14 outcomes deal with ASM refinements, hence, demonstrating the versatility of the models to materialize modeler and stakeholders' needs. Also, it is important to notice that most MBR applications were subjected to ASM modifications, as only 2 articles applied MBRs without modeling SMP/EPSSs. Moreover, 14 articles used local approaches besides their disadvantages.

For all the outcomes considered, a detailed summary of the review criteria is presented in **Table 3**. Here the sensitivity analysis method, the input factors, the model outputs, and the scenarios evaluated in each study are presented. The most common method for the LSA and the GSAs were the normalized sensitivity index (NSI, 5) and the standardized regression coefficients (SRC, 12). As for the input factors, the biokinetic (Biokyn) and stoichiometric (Stoyk) parameters were assessed together in 27 papers. Moreover, the influent fractions were studied by 12 researchers even of normally occurring fluctuations that increase uncertainty for modeling of the activated sludge process.

All the outputs were related to plant performance, either for discharged effluent quality, operational costs, or process evaluation (e.g., MLSS in the reactor, membrane and settler parameters, nutrient uptake, etc.). It is important to highlight the vast scenarios studied, remarking the applicability of the SAs for activated sludge modeling. Mind that because of the vast number of model parameters under study, discussion about the sensitivity indices results is beyond the scope of this review.

## **6. Discussion**

### **6.1 Activated sludge modeling refinements (No MBR)**

To sum up the modifications for modeling organic colloids and biopolymers, Alikhani et al. [20] improved the ASM1 for assessing the anoxic growth and decay of methylotrophs together with the growth of heterotrophs on methanol (MeOH). Gao et al. [57] introduced an ASM1-SMP-EPS for assessing the interaction between biomass and SMP and EPS kinetics. While Gao et al. [21] assay an ASM3-SMP-EPS. The main findings were the following. The methylotrophic parameters were significantly sensitive (according to the scaled sums of sensitivities), mostly those related to the growth and decay of the latter. For the ASM1-SMP-EPS, the original ASM parameters were more sensitive than the newly introduced ones. While from ASM3-SMP-EPS, 27 model parameters were proven sensitive with positive or



negative effects around the outputs. Nevertheless, all the latter used local approaches (except for one whose method remains unclear) for model validation as well as calibration, Thus, suggesting only a first approach for practitioners about the sensitive parameters, however, these must be ascertained via global analysis.

As for the two-step nitrification (SND) models, Yan et al. [22] introduced an ASM1-SND, while Zhu et al. [37] and Fortela et al. [59] employed the ASM3 model developed by Iacopozzi et al. [19]. The first used NSI for determining that the newly introduced saturation coefficients (introduced) did were proved sensitive for the model outputs. However, only the effect (positive or negative) was reported. Special attention was taken for the rest of the refinements. For example, Zhu et al. [37], were the first (up to authors' knowledge) to use the DGSM method for assessing an IWA model. The coupled the DGSM a pseudo-global covariance matrix for assessing parameter correlation. Their method was effective to prove parameter sensitivity. Also, they demonstrate a significant difference between sensitive parameters for a CSTR and an SBR. Nevertheless, comparison versus other global methods could improve the capabilities of the method. While Fortela et al. [59] developed a method that couples a Morris screening and a principal component analysis for transforming the original data into low-dimensional variables. The method allowed to provide a ranking matrix of the sensitive parameters for all the ASM state variables. Yet, it is important to mind the drawbacks of the Morris method.

## 6.2 Membrane bioreactor

Mannina et al. [8] introduced the ASM1-SMP coupled with a physical model (cake deposition, deep-bed filtration, membrane resistances) to assess the effect of SMPs in the reduction of membrane permeability together with their foulant influencing features (e.g., hydrophobicity, floc morphology). They used the SRC method as a calibration protocol and found that 25 from the 45 model parameters were sensitive ( $SRC > 0.2$ ) on the evaluated model outputs, thus, fixing 20 parameters to their original value. These results were useful for assessing start-up strategies of an MBR and found that fouling rate was lower when adding AS inoculum serves as a prefilter for dissolved and colloidal components (SMPs), thus reducing the membrane foulants [48].

All the publications that modified the ASM2d improved the model for including soluble microbial products in an MBR scheme. Like Mannina et al. [8], Cosenza et al. [29, 30] introduced an ASM2d-SMP coupled to a physical model and conducted an SRC for ease of the calibration, resulting in 24 sensitive parameters ( $SRC > 0.2$ ), hence fixing 55 model parameters. Later, Cosenza et al. [30] compared the modified model between the SRC, Morris screening, and E-FAST methods. The results showed that the SRC method is not recommended as its application is considered out of range ( $SRC < 0.7$ ). The sensitivity outcomes from Morris screening and E-FAST were similar for COD,  $S_{NH_4}$ , and  $S_{NO_3}$  but for  $S_{PO_4}$  and MLSS differ substantially in most sections of the WWTP layout. Additionally, influencing factors were very similar for the SRC and E-FAST. Nevertheless, due to convergence issues, the results between Morris screening and E-FAST protocols presented low similarity either for influencing or non-influencing factors. Finally, it was stated that SRCs are useful as a rapid method for factor prioritization due to the lower computational capacity, while the E-FAST comes with high-quality results as the method measures factor interaction (FI). Then, Cosenza et al. [51] used the E-FAST for the evaluation of sensitive parameters using real wastewater in a pilot plant and divided the influential factors into groups (e.g., COD,  $S_{NH_4}$ , etc.). Finally, Mannina et al. [55] proposed a phase protocol for assessing the uncertainty of the model for ease of the practitioners.

As for MBR biokinetics with the ASM3, Chen et al. [28] introduced the ASM3-SMP to investigate the model application in the operation of aerobic MBR. The use of the E-FAST allowed determining that 10 parameters were more sensitive within the three SRT scenarios. Mind that only the uncertain parameters were evaluated (32). Also, results indicate a strong FI in all the output, but significantly bigger for COD and MLSS. Suh et al. [50] studied a BSM1-MBR framework for rating membrane fouling using control strategies using the CES-ASM3, a model that couples EPS and SMPs. The use of an LSA highlighted that 7 membrane parameters (from 13 assessed) were more sensitive to affect the overall energy requirements or the system.

Is it important to notice that most practitioners conducted global sensitivity analysis, thus, certified as valid according to Saltelli et al. [9] perception. Even three of them did assess first-order and total sensitivity indices using the E-FAST method, thus, accounting for factor interaction. Mind that most MBR applications are concerned with ASMs refinement. Hence, only two outputs did not modify models. Besides the application of an LSA, special attention must be accounted for the “good MBR operational strategies decision tree” from Dalmau et al. [49].

### **6.3 Benchmark simulation models**

According to Gernaey et al. [6], the purpose of benchmark simulation models (BSM) is to provide a measure of reference relative to the activated sludge performance. These aim to achieve minimum costs, optimum effluent quality, and minimal sludge production by the combined assessment of control and monitoring strategies for identifying faults and optimization opportunities.

To sum up, the most relevant studies using BSMs, Flores-Alsina et al. [43] studied the sensitivity of the most uncertain ASM1 parameters within the BSM1 under different controller scenarios (combinations of oxygen, nitrate, ammonium). Sin et al. [15] use the BSM1 and the SRC for studying the design of WWTP under uncertainty scenarios concerning influent fractionation, model parameters (biokinetics, stoichiometry), plant hydraulics, and mass transfers phenomena, together with combinations of the latter. Benedetti et al. [46] evaluated the sensitivity (SRC) of the BSM2 model parameters, diving then into three categories, operation and design, water line, and sludge-line parameters. While Al et al. [14], up to the authors' knowledge were the first to use the Sobol sensitivity indices, for assessing the BSM2 framework over the most relevant plant-wide performance indicators. Notice that these articles present valid sensitivity indices as the approach was global.

### **6.4 General findings**

After conducting the model simulations and the sensitivity analysis, the previously mentioned authors (as well as the non-mentioned) found clear differences between output variance and that factor prioritization can significantly vary from one alternative to another. This is a compelling fact that practitioners must account for factor sensitivities in the function of the modeling goals to comply with [15].

Once the thorough assessment of the research outcomes concluded, the following aspects capture the authors' attention. (I) NSI, SI, SC, and RSF, seem to follow the same fundamental principles. Thus, the authors recommend using the NSI term for ease of method comparison. (II) Up to the authors' consideration, some researchers conflate local and global sensitivity analysis. Usually, because they provide a sum of the LSAs sensitivity indices, lacks an appropriate justification to be classified as global according to Saltelli et al. [9], or because the SA method is not

clear enough, as in some cases the type of analysis or the method used was unclear. (III) Most of the articles studied the sensitivity of biokinetic and stoichiometric parameters. However, sometimes calibration values can fall under non-plausible ranges. Therefore is it important to conduct appropriate assays like the ones conducted by De Arana-Sarabia et al. [58], who used respirometric systems to assess the oxygen, ammonia, and nitrate uptake rate of the activated sludge of an operating WWTP for model calibration. (IV) Only 38% of the include influent fractions as uncertain parameters. Nevertheless, these as well as temperature, are some of the most significant parameters affecting uncertainty in WWTP [15]. Especially in developing countries where WW composition is usually stronger and exhibits abrupt temporal and spatial fluctuations. Finally, (V) there are barely four articles studying total sensitivity indices ( $S_{Ti}$ ) for factor interaction, and in one specific results of the GSA were absent. Consequently, there still knowledge gaps concerning factor interaction.

## 6.5 Local versus global

Despite the capabilities of the sensibility analysis, modelers tend to rely on local approaches. In accordance with Saltelli et al. [9], local sensitivity analysis have two major drawbacks. First, these are not efficient when dealing with non-linear models, given factor interaction is not accounted for. Second OAT methods leave most of the input space unexplored.

Whenever more and more factor combinations are to be studied, the dimensionality of the input space increases becoming a hypercube. Thereby, LSAs only comprise a small fraction of the input space leaving out important phenomena that could improve systems understanding [9].

It is just enough to look at the biological processes of the ASMs or the behavior of its state variable to prove models' non-linearity. Therefore, LSAs can be classified as a non-valid sensitivity measure unless model linearity is justified. Bear in mind that it is not the authors' intention to disparage previous works concerning LSAs, rather it is to encourage practitioners to consider global approaches due to their advantages over the latter. Consequently, to overcome the latter, global sensitivity analysis are preferred.

## 6.6 Future trends

Fortunately, most papers focus on the applicability of the sensibility analysis, to investigate the uncertainty of their parameters despite the scenario studied. Consequently, application-focused methods are considered to have a broader impact on modelers and project stakeholders for decision-making [9]. Still, there is room for improvement as most WWTP have their plant configurations, objectives, as well as issues to attend to. It has been proven that sensibility analysis could be used for a wide range of purposes, mostly for calibration, model validation (of ASM refinements), control and operation strategies. Moreover, SAs can be used for project design and reengineering purposes.

An example of it is that future plant development will likely emphasize meeting stringent water quality regulations and resource recovery (water, nutrients, organics, energy) [4]. According to Regmi et al. [4], water resource recovery facilities (WRRF) will need models capable of accounting for stringent water-product quality, process performance stability, and operating costs, either for design, operation, or control. Hence, resource recovery will incur on integrating broader frameworks such as watershed models or similar, improved settler models, more phosphorus removal applications, as well as ASM improvements like the

above mentioned. For example, Saagi et al. [56] coupled an urban water system to a BSM1 framework for assessing the influence WWTP and sewer control handles on river quality under different rainfall scenarios. Using a Morris screening (due to the low computational capacity) they found that sewer control handles were more influential for TVOF and OQI. For EQI and  $S_{NH_3}$  and DO exceedance both controllers seem to be sensitive, and for OCI the WWTP ones were more influential.

Another significant model improvement was introduced by Ramin et al. [53, 54]. They compared the performance of the Takács 1D settler model and second-order model that include the effect of hydrodynamic features like convection-dispersion phenomena. After conducting the GSA, it was demonstrated that settling parameters are as influential as biokinetics on the assessed outputs. However, the second-order model seemed to provide more realistic measures, thus, suggesting the model will lead to significantly less variance in model outputs.

Regardless of the modeling application or the use of the GSA, data quality and its abundance result in better experimental designs as it provides information to support decision making [4]. Consequently, due to data abundance, the current and forthcoming increase of computer capacity, and the advancements in data-driven models, more complex models including a large number of uncertain input factors will surge [9]. Therefore, if used effectively and responsibly the sensitivity analysis could improve complex phenomena understanding concerning the activated sludge process, together with decision making.

## **7. Conclusions**

It is important to highlight that if carefully performed, sensibility analysis serve as a tool for quality assurance for either of the activated sludge modeling frameworks stated in this review, including refinement of them. The main conclusions of this research are summed up as follows.

The sensitivity indices are in function of the project modeling objectives and the scenarios being evaluated. Consequently, data quality is essential as the model will reproduce it, so if non-valid data are reproduced, will spoil sensitivity analysis applicability.

There are still knowledge gaps due to uncertainty in the influent fractions given the inherent variations in wastewater composition, as well as for the application of global sensitivity analysis that consider factor interaction as almost half of the research outputs conducted a local analysis. The latter is only applicable for assessing linear models, thus, within the activated sludge modeling field, global approaches provide more accurate measures, especially variance-based methods.

Due to the current move beyond water sanitation to resource recovery, it is expected that number of research outcomes concerning sensibility analysis in phosphorus removal models, benchmarking frameworks, membrane bioreactor models, and modification of these will increase. Finally, the sensitivity analysis are capable of propagation uncertainty among the activated sludge modeling framework parameters. Hence, its capability to improve process understanding leading to innovative solutions.


## **Author details**

Rafael Andrés Borobio-Castillo, José Manuel Cabrera-Miranda,  
Alberto Vargas-Hidalgo and Benito Corona-Vásquez\*  
Civil and Environmental Engineering Department, Universidad de las Américas  
Puebla, San Andrés Cholula, Puebla, Mexico

\*Address all correspondence to: [benito.corona@udlap.mx](mailto:benito.corona@udlap.mx)

## **IntechOpen**

---

© 2022 The Author(s). Licensee IntechOpen. Distributed under the terms of the Creative Commons Attribution - NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited. 

## References

- [1] UNESCO. Agua y Empleo. UN: United Nations World Water Assessment Programme; 2016
- [2] Grady L, Daigger GT, Love NG, Filipe CDM. Biological Wastewater Treatment. 3rd ed. London, UK: IWA Publishing; 2011
- [3] Awad H, Gar Alalm M, El-Etriby HK. Environmental and cost life cycle assessment of different alternatives for improvement of wastewater treatment plants in developing countries. *Science of the Total Environment*. 2019; **660**:57-68
- [4] Regmi P, Stewart H, Amerlinck Y, Arnell M, García PJ, Johnson B, et al. The future of WRRF modelling—Outlook and challenges. *Water Science and Technology*. 2019; **79**(1):3-14
- [5] Henze M, Gujer W, Mino T, van Loosdrecht M. Activated Sludge Models ASM1, ASM2, ASM2d and ASM3. Henze M, Gujer W, Mino M, Loosdrecht Van. Scientific and Technical Report Series (First, Vol. 5). London, UK: IWA Publishing; 2000
- [6] Gernaey KV, Jeppsson U, Vanrolleghem PA, Copp JB. Benchmarking of Control Strategies for Wastewater Treatment Plants. Scientific and Technical Report No. 23. London, UK: IWA Publishing; 2014
- [7] Judd S. In: Judd S, Judd C, editors. *The MBR Book Principles and applications of Membrane Bioreactors for Water and Wastewater Treatment*. 2nd ed. Oxford, UK: Elsevier; 2011
- [8] Mannina G, Di Bella G, Viviani G. An integrated model for biological and physical process simulation in membrane bioreactors (MBRs). *Journal of Membrane Science*. 2011b; **376**(1–2):56-69
- [9] Saltelli A, Aleksankina K, Becker W, Fennell P, Ferretti F, Holst N, et al. Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling and Software*. 2019; **114**(January):29-39
- [10] Rieger L, Gillot S, Langergraber G, Ohtsuki T, Shaw A, Takács I, et al. Guidelines for Using Activated Sludge Models. Scientific and Technical Report No. 22. London, UK: IWA Publishing; 2012
- [11] Belia E, Amerlinck Y, Benedetti L, Johnson B, Sin G, Vanrolleghem PA, et al. Wastewater treatment modelling: Dealing with uncertainties. *Water Science and Technology*. 2009; **60**(8): 1929-1941
- [12] Olyaei MA, Karamouz M. Bayesian approach for estimating biological treatment parameters under flooding condition. *Journal of Environmental Engineering*. (Reston, Virginia, United States of America). 2020; **146**(8):1-14
- [13] Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, et al. *Global Sensitivity Analysis. The Primer*. West Sussex, England: John Wiley & Sons; 2008
- [14] Al R, Behera CR, Zubov A, Gernaey KV, Sin G. Meta-modeling based efficient global sensitivity analysis for wastewater treatment plants—An application to the BSM2 model. *Computers and Chemical Engineering*. 2019; **127**:233-246
- [15] Sin G, Gernaey KV, Neumann MB, van Loosdrecht MCM, Gujer W. Global sensitivity analysis in wastewater treatment plant model applications: Prioritizing sources of uncertainty. *Water Research*. 2011; **45**(2):639-651
- [16] Ghanem R, Higdon D, Owhadi H. *Handbook of Uncertainty*

Quantification. Cham, Switzerland: Springer; 2017

[17] Hai FI, Yamamoto K, Lee C-H. *Membrane Biological Reactors: Theory, Modeling, Design, Management and Applications for Wastewater Reuse*. London, UK: IWA Publishing; 2013

[18] Rieger L, Koch G, Kühni M, Gujer W, Siegrist H. The EAWAG Bio-P module for activated sludge model No. 3. *Water Research*. 2001;**35**(16):3887-3903

[19] Iacopozzi I, Innocenti V, Marsili-Libelli S, Giusti E. A modified Activated Sludge Model No. 3 (ASM3) with two-step nitrification-denitrification. *Environmental Modelling and Software*. 2006;**22**(6):847-861

[20] Alikhani J, Takács I, Al-Omari A, Murthy S, Massoudieh A. Evaluation of the information content of long-term wastewater characteristics data in relation to activated sludge model parameters. *Water Science and Technology*. 2017;**75**(6):1370-1389

[21] Gao F, Nan J, Zhang X. Simulating a cyclic activated sludge system by employing a modified ASM3 model for wastewater treatment. *Bioprocess and Biosystems Engineering*. 2017;**40**(6): 877-890

[22] Yan D, He J, Zuo X, Li Z. Simultaneous effect of organic carbon and ammonium on two-step nitrification within sequential batch reactor (SBR). *International Journal of Environmental Science and Technology*. 2019;**16**(5):2239-2248. DOI: 10.1007/s13762-018-1769-9

[23] Copp JB. *The COST Simulation Benchmark: Description and Simulator Manual*. Luxembourg: Office for Official Publications of the European Community; 2002

[24] Takács I, Patry GG, Nolasco D. *A dynamic model of the clarification*

*thickening process*. *Water Research*. 1991;**25**(10):1263-1271

[25] Jeppsson U, Rosen C, Alex J, Copp J, Gernaey KV, Pons MN, et al. Towards a benchmark simulation model for plant-wide control strategy performance evaluation of WWTPs. *Water Science and Technology*. 2006;**53**(1):287-295

[26] Otterpohl R. *Dynamische Simulation zur Unterstützung der Planung und des Betriebes von kommunalen Kläranlagen*. Schriftenreihe Gewässerschutz Wasser Abwasser. 1995;**151**:28-36

[27] Batstone DJ, Keller J, Angelidaki I, Kalyuzhnyi SV, Pavlostathis SG, Rozzi A. The IWA Anaerobic Digestion Model No 1 (ADM1). *Water Science and Technology*. 2002;**45**(10):65-73

[28] Chen L, Tian Y, Cao C, Zhang S, Zhang S. Sensitivity and uncertainty analyses of an extended ASM3-SMP model describing membrane bioreactor operation. *Journal of Membrane Science*. 2012;**389**:99-109

[29] Cosenza A, Mannina G, Neumann MB, Viviani G, Vanrolleghem PA. Biological nitrogen and phosphorus removal in membrane bioreactors: Model development and parameter estimation. *Bioprocess and Biosystems Engineering*. 2013a;**36**(4): 499-514

[30] Cosenza A, Mannina G, Vanrolleghem PA, Neumann MB. Global sensitivity analysis in wastewater applications: A comprehensive comparison of different methods. *Environmental Modelling and Software*. 2013b;**49**:40-52

[31] Cierkens K, Plano S, Benedetti L, Weijers S, De Jonge J, Nopens I. Impact of influent data frequency and model structure on the quality of WWTP model calibration and uncertainty.

Water Science and Technology. 2012;  
**65**(2):233-242

[32] De Araujo ACB, Gallani S, Mulas M, Skogestad S. Sensitivity analysis of optimal operation of an activated sludge process model for economic controlled variable selection. *Industrial and Engineering Chemistry Research*. 2013; **52**(29):9908-9921

[33] Dey A. Modeling simultaneous nitrification-denitrification process in an activated sludge continuous flow stirred-tank reactor: System optimization and sensitivity analysis. *Environmental Engineering Science*. 2010;**27**(9):757-765

[34] Campolongo F, Cariboni J, Saltelli A. An effective screening design for sensitivity analysis of large models. *Environmental Modelling and Software*. 2007;**22**:1509-1518

[35] Morris MD. Factorial sampling plans for preliminary computational experiments. *Technometrics*. 1991;**33**: 161-174

[36] Saltelli A, Tarantola S, Chan KPS. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*. 1999; **41**(1):39-56

[37] Zhu A, Guo J, Ni BJ, Wang S, Yang Q, Peng Y. A novel protocol for model calibration in biological wastewater treatment. *Scientific Reports*. 2015;**5**:1-10

[38] Arnell M, Astals S, Åmand L, Batstone DJ, Jensen PD, Jeppsson U. Modelling anaerobic co-digestion in Benchmark Simulation Model No. 2: Parameter estimation, substrate characterisation and plant-wide integration. *Water Research*. 2016;**98**: 138-146

[39] Mannina G, Cosenza A. Quantifying sensitivity and uncertainty analysis of a new mathematical model for the

evaluation of greenhouse gas emissions from membrane bioreactors. *Journal of Membrane Science*. 2015;**475**:80-90

[40] Sweetapple C, Fu G, Butler D. Identifying sensitive sources and key control handles for the reduction of greenhouse gas emissions from wastewater treatment. *Water Research*. 2014;**62**:249-259

[41] Ramanan R, Kim BH, Cho DH, Oh HM, Kim HS. Algae-bacteria interactions: Evolution, ecology and emerging applications. *Biotechnology Advances*. 2016;**34**(1):14-29

[42] Liwarska-Bizukojc E, Biernacki R. Identification of the most sensitive parameters in the activated sludge model implemented in BioWin software. *Bioresource Technology*. 2010;**101**(19):7278-7285

[43] Flores-Alsina X, Rodriguez-Roda I, Sin G, Gernaey KV. Uncertainty and sensitivity analysis of control strategies using the Benchmark Simulation Model No1 (BSM1). *Proc. IEMSs 4th Biennial Meeting—Int. Congress on Environmental Modelling and Software: Integrating Sciences and Information Technology for Environmental Assessment and Decision Making, IEMSs*. 2008;**2**:1331-1338

[44] Kim J, Kim Y, Lee S, Moon T, Kim C. Dynamic parameter estimation to calibrate the activated sludge model for an enhanced biological phosphate removal process. *Desalination and Water Treatment*. 2009;**4**(1–3):262-268

[45] Mannina G, Cosenza A, Vanrolleghem PA, Viviani G. A practical protocol for calibration of nutrient removal wastewater treatment models. *Journal of Hydroinformatics*. 2011; **13**(4):575-595

[46] Benedetti L, Batstone DJ, De Baets B, Nopens I, Vanrolleghem PA. Uncertainty analysis of WWTP control



- strategies made feasible. *Water Quality Research Journal of Canada*. 2012;**47**(1): 14-29
- [47] Lim JJ, Kim MH, Kim MJ, Oh TS, Kang OY, Min B, et al. A systematic model calibration methodology based on multiple errors minimization method for the optimal parameter estimation of ASM1. *Korean Journal of Chemical Engineering*. 2012;**29**(3):291-303
- [48] Mannina G, Di Bella G. Comparing two start-up strategies for MBRs: Experimental study and mathematical modelling. *Biochemical Engineering Journal*. 2012;**68**:91-103
- [49] Dalmau M, Rodriguez-Roda I, Ayesa E, Odriozola J, Sancho L, Comas J. Development of a decision tree for the integrated operation of nutrient removal MBRs based on simulation studies and expert knowledge. *Chemical Engineering Journal*. 2013;**217**:174-184
- [50] Suh C, Lee S, Cho J. Investigation of the effects of membrane fouling control strategies with the integrated membrane bioreactor model. *Journal of Membrane Science*. 2013;**429**:268-281
- [51] Cosenza A, Mannina G, Vanrolleghem PA, Neumann MB. Variance-based sensitivity analysis for wastewater treatment plant modelling. *Science of the Total Environment*. 2014; **470–471**:1068-1077
- [52] Sharifi S, Murthy S, Takács I, Massoudieh A. Probabilistic parameter estimation of activated sludge processes using Markov Chain Monte Carlo. *Water Research*. 2014;**50**:254-266
- [53] Ramin E, Flores-Alsina X, Sin G, Gernaey KV, Jeppsson U, Mikkelsen PS, et al. Influence of selecting secondary settling tank sub-models on the calibration of WWTP models—A global sensitivity analysis using BSM2. *Chemical Engineering Journal*. 2014a; **241**:28-34
- [54] Ramin E, Sin G, Mikkelsen PS, Plósz BG. Significance of settling model structures and parameter subsets in modelling WWTPs under wet-weather flow and filamentous bulking conditions. *Water Research*. 2014b;**63**: 209-221
- [55] Mannina G, Cosenza A, Viviani G. Sensitivity and uncertainty analysis of an integrated membrane bioreactor model. *Desalination and Water Treatment*. 2016;**57**(21):9531-9548
- [56] Saagi R, Kroll S, Flores-Alsina X, Gernaey KV, Jeppsson U. Key control handles in integrated urban wastewater systems for improving receiving water quality. *Urban Water Journal*. 2018; **15**(8):790-800
- [57] Gao F, Nan J, Li S, Wang Y. Modeling and simulation of a biological process for treating different COD:N ratio wastewater using an extended ASM1 model. *Chemical Engineering Journal*. 2018;**332**:671-681
- [58] De Arana-Sarabia ME, Vasiliadou IA, Vitanza R, Cortesi A, Gallo V. Mathematical Simulation and Validation of a Wastewater Treatment Plant in Northern Italy. *Environmental Engineering Science*. 2018;**35**(10): 1136-1147
- [59] Fortela DLB, Farmer K, Zappi A, Sharp WW, Revellame E, Gang D, et al. A methodology for global sensitivity analysis of activated sludge models: Case study with activated sludge model no. 3 (ASM3). *Water Environment Research*. 2019;**91**(9):865-876
- [60] Lahdhiri A, Lesage G, Hannachi A, Heran M. Steady-state methodology for activated sludge model 1 (ASM1) state variable calculation in MBR. *Water*. 2020;**12**(11):1-13



# Microbial Photobioelectrochemical Systems: A Scoping Review

*Luis Erick Coy-Aceves, José Luis Sánchez-Salas,  
Mónica Cerro-López, Miguel Ángel Méndez-Rojas  
and Benito Corona-Vázquez*

## Abstract

The combination of characteristics belonging to bioelectrochemistry and photoelectrochemistry produces a relatively new area that can be called photobioelectrochemistry. The main idea consists of an electrochemical device that can make use of light, microorganisms, biotic materials, and/or abiotic materials for multiple applications such as energy generation, hydrogen production, CO reduction into compounds like methane, heavy metals reduction, and green material synthesis. Light can be harvested by semiconductors, phototrophic microorganisms, and biotic substances. This area of research is recent; there is no classification system to identify different combinations of materials or configurations of these devices. This heterogeneity makes it difficult for scientists to search for an application or a specific material integration to know the state of the art of the field, as well as slow down the identification of unexplored configurations of these systems. Consequently, this makes the direct comparison of contributions and advances in this area an enormous challenge. This work proposes sets for all the possible permutations of photobioelectrochemical systems, as well as a classification system for them. Additionally, it presents a scoping review of investigations regarding combinations of experimental arrangements for microbial photobioelectrochemical systems, as well as their applications to identify areas of opportunity for topics that remain unexplored.

**Keywords:** wastewater treatment, bioelectrochemical systems, photoelectrochemical systems, photobioelectrochemical systems, standardization, microbial photobioelectrochemical systems

## 1. Introduction

The available data on worldwide wastewater treatment indicate that, on average, high-income countries treat 70% of their generated wastewater, while upper-middle-income countries treat 38%, lower-middle-income-countries treat 28% [1]; in contrast, 80 to 90% of the wastewater in low-income-countries is neither collected nor treated [2]. Developing countries do not perform enough wastewater treatment due to their high operating costs, high energy consumption, and low economic return [3]. It is for this reason that polluted water is constantly accumulating.

The energy contained in wastewater pollutants is 6 to 13 times greater than that required for its treatment [4, 5], thus, much of new research and development initiatives in this area revolves around the recovery of such energy to save or profit from wastewater treatment. Some of the technological developments that are applied to recover energy from wastewater are the combustion of organic solid waste to generate electricity and anaerobic digestion to generate biogas from pollutants [6].

Much of the technology that can harness energy from water pollutants are currently under research and development. Within this set, bioelectrochemical systems are particularly interesting because they can perform wastewater treatment and generate electricity, hydrogen, or high-value chemical compounds simultaneously; furthermore, they can be designed for other purposes such as desalinating water or removing nitrogen, producing electricity at the same time [7].

The number of studies about bioelectrochemical systems has increased by orders of magnitude in the last couple of decades [8]; however, there is no standardized parameter to report their performance. Many publications report results using different indicators, which, in turn, complicates direct comparison among results [7]. This is arguably one of the main factors that slow the optimization and advance of these devices.

There are bioelectrochemical systems that require the application of voltage to work, which reduces the net profits in their operation. It has been proven that the use of semiconductors or phototropic bacteria can generate the required potential difference using sunlight [9]. This new set of devices that use bacteria and semiconductors in electrochemical cells can be called photobioelectrochemical systems.

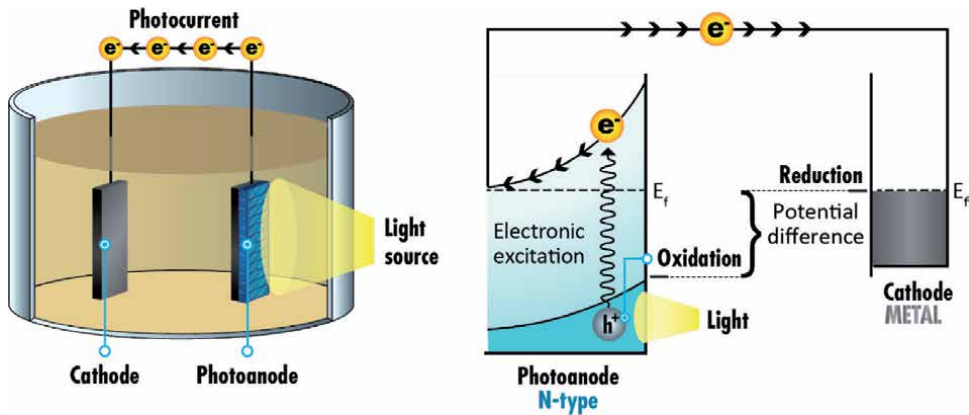
On the other side, there is little information available in the literature about this mix of photoelectrochemical and bioelectrochemical systems. However, the variety of combinations of materials, as well as subproducts and functions that these devices can perform is very large. The scoping review approach can open a way to explore this new research area to identify areas of opportunity that can help research teams work towards its evolution.

The main objective of the present work is to identify and map the existing literature on photobioelectrochemical systems, as well as to identify gaps in knowledge that can be useful to research and advance their development. Also, this paper proposes a way to organize the knowledge surrounding these devices, by systematically reviewing in the future when more literature will be available.

## **1.1 Photoelectrochemistry**

The term “photoelectrochemistry” refers to the area of electrochemistry that studies photoactive electrodes, also called photoelectrodes, exposed to light. Photoactivity is usually achieved by using semiconductor materials that, when irradiated by light, generate an electrochemical reaction that produces an electrical current, also known as photocurrent. This process represents the conversion of light energy into electric and chemical energy. Photoelectrochemical systems are widely studied because they can have potential applications for renewable energy generation and storage [10], as well as environmental applications. After all, their performance can be useful for the design of advanced oxidation processes [11].

The simplest form of a photoelectrochemical cell, illustrated in **Figure 1**, consists of a photoelectrode and a metallic electrode. A photoelectrode can behave as anode or cathode depending on the semiconductor nature. It is more common for photoanodes, made with n-type semiconductors, to be used because photocathodes, made with p-type semiconductors, tend to corrode in solution. In photoanodes, the voltage generated by the production of electron–hole pairs drives the photocurrent from the anode to the cathode. The most studied application



**Figure 1.**  
*Photoelectrochemical cell with photoanode.*

of photoanode devices is water splitting because they can store solar energy by generating hydrogen [10].

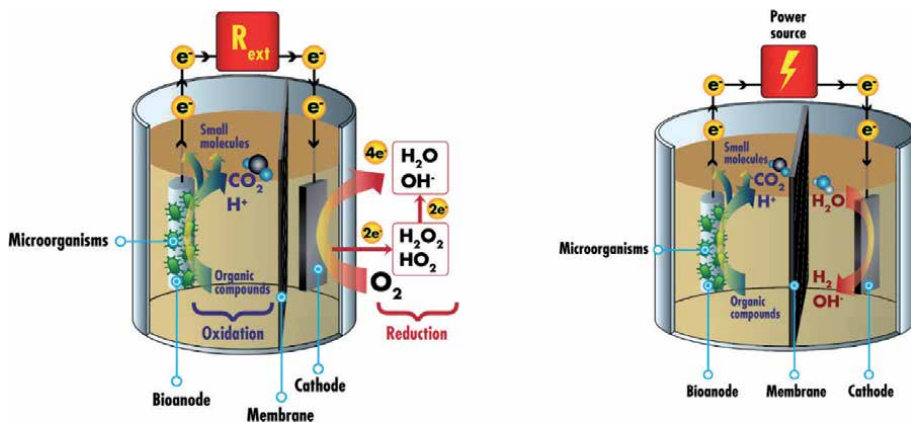
Another configuration of photoelectrochemical cells consists of two photoelectrodes, allowing the full use of sunlight. As the Fermi level of the photoanode, due to its n-type semiconductor properties, is higher than that of the photocathode, which tends to have a lower Fermi level due to its p-type semiconductor properties, the generated photocurrent between the electrodes is enhanced and stronger redox abilities of the electrons and holes on each photoelectrode can be achieved. Used as an advanced oxidation process, these systems are also known as photocatalytic fuel cells, and they are also widely studied as an environmentally friendly technique to oxidize nonbiodegradable compounds in polluted water [11].

## 1.2 Bioelectrochemistry

Many microorganisms are exoelectrogenic. This means that they can generate an electrical current by connecting their electron transport system outside of their cell membranes either directly to an electrode or indirectly through added mediators or electron shuttles. Some microorganisms are electrotrophic, which means that they can harvest electrons to grow [12]. Bioelectrochemistry studies show how these microorganisms can be used in electrochemical cells to remove organic matter from wastewater and generate electricity.

There is a wide variety of bioelectrochemical systems. Some of the most studied, are the microbial fuel cells, illustrated in **Figure 2** (left). These devices generate electricity by using an exoelectrogenic biofilm-coated anode, also called bioanode, to oxidize organic matter and spontaneously generate a current that reduces oxygen at the cathode. [8]. On some occasions, electrotrophic biofilm-coated cathodes, also called biocathodes, can be implemented instead of regular counter electrodes to drive other reactions such as ammonia oxidation and nitrate reduction to nitrogen [13]. Another widely studied form of bioelectrochemical system is the microbial electrolysis cells. These devices have the same components as microbial fuel cells, with the difference that there is no oxygen in the cathodic side of the cell, which allows for other chemical species such as hydrogen ions to be reduced. The main drawback of these systems is that they need an applied voltage to drive the reduction reaction [8].

Exoelectrogenic and electrotrophic microorganisms are so diverse that they can be used to build bioelectrochemical systems with many applications for wastewater and



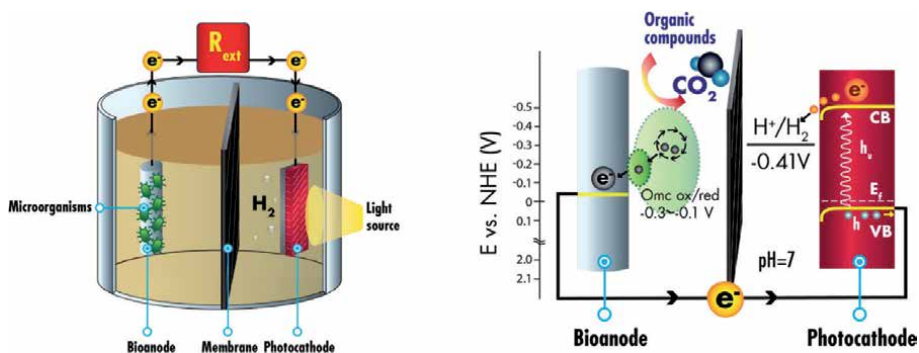
**Figure 2.** Microbial fuel cell (left) and microbial electrolysis cell (right).

saltwater treatment such as removal of organic compounds, water desalination, denitrification, and sulfate removal while generating electricity or storing chemical energy in the form of hydrogen or other high-value compounds such as methane [8, 12, 13].

### 1.3 Photobioelectrochemistry

The net energy generated by microbial electrolysis cells is significantly less than that of the microbial fuel cells due to the applied voltage that they require to work, which can represent up to 95% of the required operation energy of these devices [7]. To overcome this limitation, as well as improve energy generation, these systems have been modified either by adding photoelectrodes [14–16] or phototrophic microorganisms, which use light to drive their metabolism, [17] to harness the power of solar light. These new systems, for which an example is illustrated in **Figure 3**, have designs that integrate characteristics of photoelectrochemical and bioelectrochemical systems, generating a new field of study that will be called photobioelectrochemistry in this review.

Two other reviews address photobioelectrochemical systems. Both of them acknowledge that there is a new trend that integrates bioelectrochemical systems with photoelectrochemistry; however, their views do not clarify the potential scope that this field can have in terms of electrode materials and cell configurations [9, 18]. The present review aims to provide means to identify the whole set of



**Figure 3.** A microbial photoelectrochemical cell with bioanode and photocathode.

photobioelectrochemistry elements and analyze one of its subsets. To start defining photoelectrochemistry as its field, it is necessary to identify the electrode materials that it uses and find the proper terminology to refer to them.

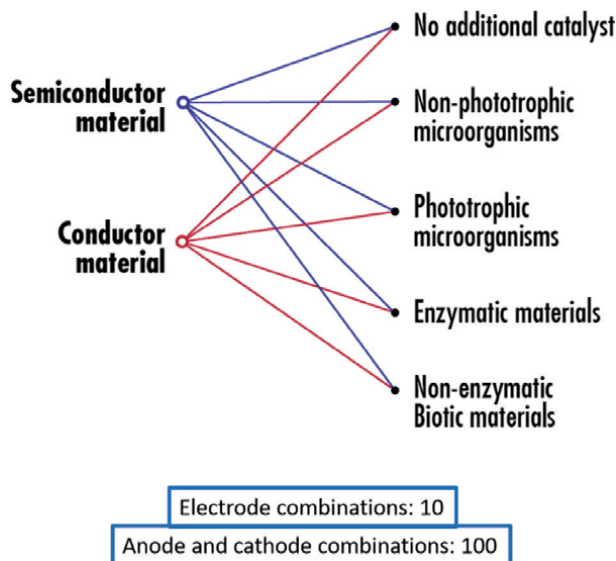
### 1.3.1 Electrode materials

Photoelectrochemical systems use a wide range of electrode materials. Conductor or semiconductor electrodes can be used directly or as a supporting matrix for catalysts such as chemotrophic microorganisms [19], phototrophic microorganisms [20], enzymes [21] or any other biotic material like cell membranes [22] or chlorophyll [23]. As seen in **Figure 4**, there are at least 10 combinations of electrodes that can be formed assuming that there are no combinations between catalysts and all the biotic materials that are not enzymes are grouped. Considering that an electrochemical cell requires at least two electrodes to work, there can be near 100 possible configurations of photobioelectrochemical systems, not counting systems with more than two electrodes or electrode pairs that are only bioelectrochemical or photoelectrochemical.

This extensive variety of electrodes and cell configurations makes the field of photobioelectrochemistry so large that doing a complete review that identifies all the systems that have been tested to date is a big challenge. As different conditions are required to preserve an enzyme, to maintain an active biofilm, or to keep other biotic materials that are not greatly affected by the environment, different cell designs and configurations are required according to the biological catalysts that they contain, which increases the difficulty for classifying all the possible photobioelectrochemical devices.

### 1.3.2 The field of a thousand names

It is quite difficult to find specific information about this new area of study. One of the reasons is that this field has not been properly established. Another



**Figure 4.** Electrode materials and catalysts for photobioelectrochemical systems.

one is that the terminology used to address photobioelectrochemical devices is not standardized. Photobioelectrochemical systems are referred to with several different terms in the articles available in the scientific literature. **Table 1** contains some examples where different prefixes and suffixes are used based on physical characteristics or functionalities that these devices present.

This mix of prefixes and suffixes are also used to name each combination of electrode materials, as well as functional capabilities. For instance, **Table 2** enlists 16 different names that were used to refer to the same set of microbial bioanode and photocathode electrodes in 26 articles, out of 76 that reported any type of photobioelectrochemical system. While some researchers use dashes to build the terms, others make a brief description of the devices. For the biological components only the prefix bio and the word microbial are implemented, which is easy to remember;

Field name	Reference
Photobioelectrochemical Systems	[24]
Bioelectrochemical solar Cells	[22]
Bio-photoelectrochemical Cells	[25]
Photo-bioelectrochemical Systems	[26]
Photoelectrochemical Biofuel Cells	[21]
Bio-photo-electro-catalytic reactors	[27]
Biophotovoltaic cell	[28]
Semiconductor biohybrid system	[29]

**Table 1.**  
*Current terms used for referring to Photobioelectrochemical systems.*

No.	System name	Reference
1	Biophotoelectrocatalytic reactor with bioanode and photo-electrocatalytic cathode	[27]
2	Biophotoelectrochemical cell	[30]
3	Biophotoelectrochemical system with photocathode	[31]
4	Biophotoelectrochemical cell with photocathode	[25]
5	Microbial coupled photoelectrochemical fuel cell	[32]
6	Microbial fuel cell equipped with a photocatalytic cathode	[33]
7	Microbial fuel cell with photocathode	[34]
8	Microbial photoelectrochemical cell	[35]
9	Microbial photoelectrochemical cell with photocathode	[36]
10	Photo-assisted microbial electrolysis cell using a photocathode	[37]
11	Photocatalytic microbial electrolysis cell	[38]
12	Photo-driven bioelectrochemical photocathode	[39]
13	Photoelectrocatalytic microbial fuel cell	[40]
14	Semiconductor cathode coupling with bioanode	[41]
15	Semiconductor cathode in microbial fuel cells	[42]
16	Microbial photoelectrosynthesis system	[43]

**Table 2.**  
*The terminology used to define a photobioelectrochemical device with bioanode and photocathode materials.*



however, for the photoelectrode components, the prefixes photo and/or electro are coupled with verbs such as assistance, catalysis, driving, and synthesis. These prefixes are usually substituted by the word semiconductor to refer to these electrodes. Also, some papers describe the devices, while others create a unique name for them.

The wide variety of terms used to define photobioelectrochemical systems further increases the difficulty of making a literature review. This work aimed to compile as much information on this field as possible, but there are probably so many combinations of keywords that some of them may be missed even with a systematic search process like the one used here.

## 2. Methodology

Given the considerable number of permutations that photobioelectrochemical systems can have, it is better to approach the analysis of this field in terms of subsets of devices that share the same cell designs and operating conditions for a selection of electrode materials and catalysts. It is convenient to group devices that share similarities in terms of the biological materials used. For instance, microbial photobioelectrochemical cells were reviewed in this study. These systems use chemotrophic and phototrophic microorganisms as the biological component, which means that their designs must have the conditions necessary to maintain microbial growth and to harness light.

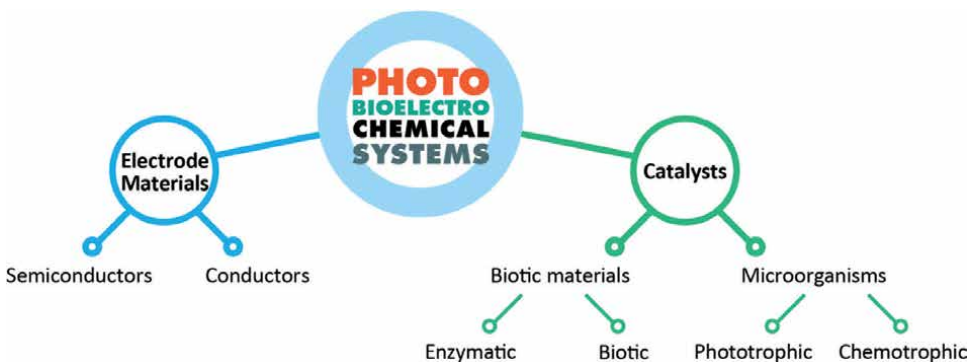
### 2.1 Nomenclature system

To avoid confusion and present this review in the clearest way possible, a nomenclature system was developed. The main purpose of this tool is to identify any photobioelectrochemical device, as well as the nature of its components, just by reading the constructed name. The prefixes, suffixes, and descriptive terms have been based on literature published, as well as the Compendium of Chemical Terminology of the International Union of Pure and Applied Chemistry (IUPAC) [44]. Although this nomenclature was developed for this review, it was designed to name any photobioelectrochemical system.

#### 2.1.1 *Electrode name building*

As the electrodes can be made from a considerable variety of materials, their nomenclature should be able to describe their composition as accurately as possible. To achieve this, two categories for supporting electrode materials and four for catalysts, as illustrated in **Figure 5**, will be considered. It is assumed that conductors are a component of all electrodes since they are connected to an electrochemical circuit, so no nomenclature will be added for these materials. The prefix “photo” will be used to refer to semiconductors and the prefix “bio” will be used for any biological material. To distinguish the four categories of biomaterials, the prefixed words “enzymatic” and “biotic” will be used to distinguish enzymatic catalysis from any other material with a biological origin that cannot be classified as microorganisms or enzymes, such as chlorophyll. The words “phototrophic” and “chemotrophic” will be used to indicate the type of microorganisms.

The prefixes of the compound name for an electrode are ordered to indicate the material that would be exposed to light and the words “cathode” or “anode” indicate the electrode where reduction or oxidation, respectively, takes place. The order is determined by looking at the electrode material layers from the outside into the center of the electrochemical cell, as illustrated in **Figure 6**, where a chemotrophic

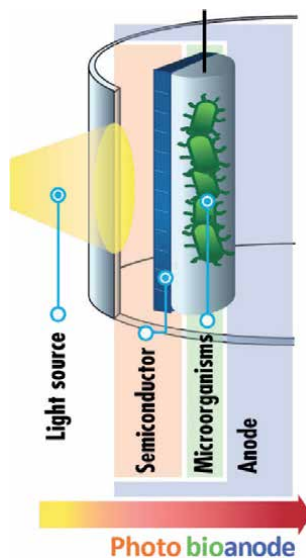


**Figure 5.**  
*Categories for electrode materials.*

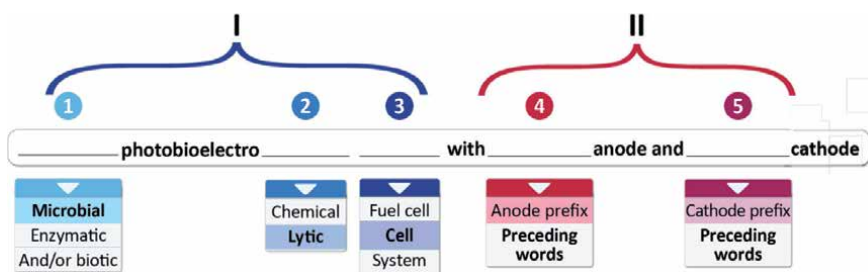
photo bioanode would indicate an anodic electrode that has a semiconductor side exposed to light and chemotrophic microorganisms on the other side.

### 2.1.2 System name building

The name for a photobioelectrochemical system will follow the structure shown in **Figure 7** where, in the first section (**Figure 7, I**) the first word(s) indicate if the biological component of the device is enzymatic, microbiological, or any other biotic material (**Figure 7, I, 1**). The second part of the name (**Figure 7, I, 2**) has the prefixes “photobioelectro” to indicate that a system is an object of study from the field of photobioelectrochemistry and this word will end with the suffix “chemical” if the electrical current generated by the system is spontaneous, or the suffix “lytic” if it is non-spontaneous [44]. The third part of the name (**Figure 7, I, 3**) ends with



**Figure 6.**  
*Example of nomenclature for an electrode composed of a semiconductor exposed to light and chemotrophic microorganisms.*



**Figure 7.**  
*Photobioelectrochemical device nomenclature.*

the words “fuel cell” to indicate electricity generation, “cell” for electrolysis-related reactions, and “system” for any other functionality that the device may have.

The second section (**Figure 7, II**) mentions the electrode and material combinations, following the corresponding nomenclature for electrodes. The complete name built with this nomenclature should describe all the characteristics of a device. For instance, a microbial and biotic photobioelectrolytic cell with chemotrophic bioanode and biophotocathode could refer to a two-electrode electrochemical chamber with a biofilm coated anode that does not need to be exposed to light and a chlorophyll-sensitized semiconductor cathode with the capacity to generate hydrogen from water using solar light. All the specific details would be described later here for any paper that reports the studies made on such a system.

## 2.2 Determination of electrode variety and cell permutations

The subset of microbial photobioelectrochemical cells contains conductors and semiconductors, as well as chemotrophic and phototrophic microorganisms as electrode materials. **Table 3** shows the possible electrode variations found after estimating combinations of these materials. As the order in which these materials are added to the surface of the electrode can modify their performance, terms like photobioelectrode and biophotoelectrode would mean that the former has a surface with a semiconductor covered in a biofilm, while the latter has a biofilm coated with a semiconductor film.

Of the 16 possible electrode combinations shown in **Table 3**, the 9 that are in italics are exclusively photobioelectrodes because they either have phototrophic microorganisms or combine semiconductors with chemotrophic microorganisms. These electrodes would need to be exposed to light and be in an environment that meets the appropriate conditions to maintain cellular growth. Bioelectrodes and Photoelectrodes are already widely studied in their respective fields; however, photobioelectrochemical systems can also be built by combining them [39], so this review covers reported studies that either use photobioelectrodes or combines bioelectrodes and photoelectrodes.

## 2.3 Search strategy

The articles were selected according to the following inclusion criteria:

- a. Use of microorganisms and light: All references that state in their title or abstract that the studied device uses either phototrophic microorganisms and/or semiconductors in addition to chemotrophic microorganisms would need

Component	Conductor	Semiconductor	Chemotrophic biofilm	Phototrophic biofilm
Conductor	Electrode	Photoelectrode	Chemotrophic Bioelectrode	Phototrophic bioelectrode
Semiconductor	Photoelectrode	Photoelectrode	Chemotrophic photobioelectrode	Phototrophic photobioelectrode
Chemotrophic biofilm	Chemotrophic Bioelectrode	Chemotrophic biophotoelectrode	Chemotrophic Bioelectrode	Phototrophic and chemotrophic bioelectrode
Phototrophic biofilm	Phototrophic bioelectrode	Phototrophic biophotoelectrode	Phototrophic bioelectrode	Phototrophic bioelectrode

**Table 3.** Microbial photobioelectrochemical electrodes. Each material on the first column is the surface of the electrode.

a similar cell design and operating conditions, which is within the scope of this review.

- b. Use of electrochemical configurations: Many studies use phototrophic microorganisms, semiconductors, and combinations of chemotrophic microorganisms with semiconductors. Only the ones that used these materials on electrochemical cells were included.
- c. Use of bioelectrodes: The references must use exoelectrogenic biofilms grown on conductive surfaces.
- d. Publication status: All included references have a digital object identifier to demonstrate that they are legitimately published.

Likewise, articles were discarded according to the following exclusion criteria:

- i. Use of enzymatic materials: The use of enzymes in combination with microorganisms and/or semiconductors requires cell designs and operating conditions that prevent the denaturation of these macromolecules. These characteristics can be significantly different from the ones belonging to the subset of systems studied.
- ii. Use of biotic materials and not microorganisms: Photoelectrochemical devices with bioelectrodes that contain solely biotic materials such as chlorophyll can resist a broader variety of operating conditions in comparison to microorganisms, and this characteristic sets them out of the subset of microbial photobioelectrochemical systems.

Other data such as year of publication and applications or subproducts of the studied devices were not limited as exclusion criteria because information about microbial photobioelectrochemical systems is very scarce and this was considered as an area of opportunity for this manuscript to harness all the available information.

Publications were collected from Web of Science (WoS), covering a period of 27 years from 1993 to 2020. The most recent search was executed on December 15th, 2020.

A preliminary search with the keywords “Microbial photoelectrochemical cells” was realized. After obtaining a set of terms, some of which were mentioned in **Tables 1** and **2**, the sampling was carried out by using 12 simultaneous WoS search fields, each referring to combinations of prefixes to name photobioelectrochemical devices or combinations of electrodes. All the search fields are linked by OR operators and are explained in the following points:

1. Photoelectro\* AND (Microbial OR Bioelectro\*) NOT Photoelectron\$: This covers all references that mention photoelectrochemistry combined with microbial materials, as well as any bioelectrochemical phenomenon.
2. Photobioelectro\* OR Bio-photoelectro\* OR Photo-bioelectro\* OR Bio-photo-electro\* OR Photo-bio-electro\*: This covers all the references that combine the photo and bio prefixes commonly used to denote photobioelectrochemical systems.
3. Bioelectro\* AND Solar\* AND Light\* AND Photo\*: This covers reports that mention bioelectrochemical devices that use light to operate.

4. Biophoto\* NOT (Biophoton\$ OR Biophotonic\$): This covers any publications that mention biological components, semiconductor components, and/or exposure to light. The words biophoton(s) and biophotonic(s) were excluded because they refer to light with a biological origin, which is out of the scope of this review.
5. Semiconductor\$ AND Biohybrid\$: This covers any articles that mention semiconductor biohybrid systems, some of which can be photobioelectrochemical.
6. (“Microbial fuel cell\$” OR “Microbial electrolysis cell\$”) AND (Semiconductor\* OR Photocathode\$ OR Solar\* OR Photo\*): This covers any reference that specifically mentions microbial fuel or electrolysis cells that have any relationship to the use of light harnessed by a semiconductor or phototrophic microorganisms.
7. Bioanode\$ AND (Semiconductor\* OR Photocathode\$ OR Photobiocathode\$ OR Biophotocathode\$ OR Solar\* OR Light\* OR Photo\*): This covers all photobioelectrochemical devices that include a bioanode, which can contain phototrophic microorganisms, and any other photoelectrode or photobioelectrochemical electrode.
8. Photoanode\$ AND (Biocathode\$ OR Photobiocathode\$ OR Biophotocathode\$): This covers all the combinations of photoanodes with any photobioelectrochemical electrode, including biocathodes.
9. Biocathode\$ AND (Semiconductor\* OR Photoanode\$ OR Photobioanode\$ OR Biophotoanode\$ OR Solar\* OR Light\* OR Photo\*): This covers all biocathodes, which can have phototrophic microorganisms, combined with photobioelectrochemical electrodes.
10. Photocathode\$ AND (Bioanode\$ OR Photobioanode\$ OR Biophotoanode\$): This covers all the combinations of photocathodes with bioanodes and photobioelectrochemical anodes.
11. (Photobioanode\$ OR Biophotoanode\$) AND (Photobiocathode\$ OR Biophotocathode\$ OR Solar\* OR Light\* OR Phototrophic): This covers combinations of photobioelectrochemical electrodes in which the anodes contain phototrophic microorganisms.
12. (Photobioanode\$ OR Biophotoanode\$ OR Solar\* OR Light\* OR Phototrophic) AND (Photobiocathode\$ OR Biophotocathode\$): This covers combinations of photobioelectrochemical electrodes in which the cathodes contain phototrophic microorganisms.

The search fields were set to find the defined queries in all fields; however, as the number of references initially found was 2009, the survey was refined by restricting the query to the title of the publications. The remaining reports were analyzed by their relevance to the topic of interest by screening their titles, abstracts, and the full texts, selecting them according to the eligibility criteria.

The data was charted by using a form developed on an Excel file. The spreadsheet was continuously updated in an iterative process when new information was found.

The data extracted from the references include general characteristics (such as digital object identifier, year, country, and keywords), electrodes that were used, and subproducts obtained (such as electricity, hydrogen, or other synthetic materials).

Data from research done using microbial consortiums or pure cultures, citing their names if they were reported in the publications was extracted too.

From the results reported in these studies, only the highest values regarding current density, power density, and efficiency of contaminant removal or product synthesis were included in this review. Also, the units in which values are reported were adjusted using the data available in the publications to report them in the most homogeneous way possible in order to make direct comparison feasible. In the cases where the cell compartments have different volumes, the unit standardization was done using the anolyte's volume. When a result is reported in a quantity per unit of area, it is based on the anode's projected area.

The studies were grouped according to the combination of electrodes and electrode materials used. Subgroups were made by classifying the subproducts generated such as electricity, hydrogen, or other synthetic compounds. If any review articles about photobioelectrochemical systems were found, an analysis of their references was performed to identify studies missed with the search strategy reported above in this manuscript.

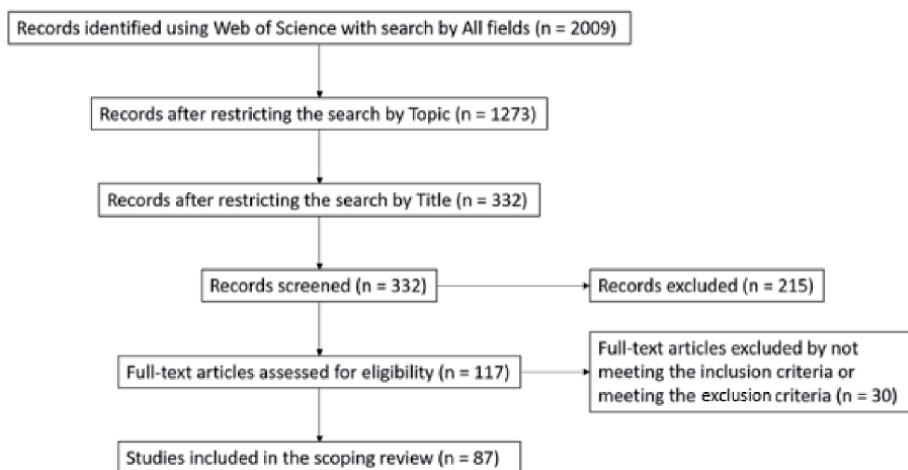
### 3. Results

The initial search, which included all the fields, resulted in 2009 publications. Refining the search to look by topic, 1273 results were obtained. Of these, 332 included the keywords in their titles. After a full-text analysis, 87 articles met the inclusion criteria. **Figure 8** illustrates the process of selection that was carried out. The selected studies were published after 2009.

It was found that the combinations of materials and electrodes that have been studied are chemotrophic bioanode and photocathode, photoanode and biocathode, phototrophic bioelectrodes, chemotrophic photobioelectrodes, and systems with three or more electrodes.

#### 3.1 Chemotrophic bioanode and photocathode systems

Microbial photobioelectrochemical cells with chemotrophic bioanodes and photocathodes are the most studied devices because their main function is to use light



**Figure 8.**  
*Selection of sources of evidence.*

as the source of energy to drive a hydrogen reduction reaction at the photocathode. Most of them can simultaneously degrade organic compounds at the bioanode, which can increase the energy efficiency and lower operating costs compared to conventional microbial electrolysis cells [45]. Aside from hydrogen production, these systems have been used to generate electricity or synthesize semiconductors, as well as to reduce heavy metals.

### *3.1.1 Hydrogen production*

A list of microbial photobioelectrochemical cells that focus on hydrogen production is summarized in **Table 4**. The semiconductor materials that are more frequently used in photocathodes are titanium dioxide [30, 38, 41, 46] and copper(I) oxide [14–16, 45, 47]. These tests have only been done with known substances such as synthetic nutrient mediums with acetate or trypticase soy broth. This limits the extrapolation of the results obtained because there are no tests done with wastewater that can prove the effectiveness of these devices for wastewater treatment.

There were studies in which the power density obtained by illuminating the photocathode was so low that an additional voltage had to be applied to obtain measurable amounts of electricity and hydrogen [14–16, 35, 37, 43, 47]. These references did not provide enough information to determine if the low current generation could be attributed to the semiconductor used, the electrode dimensions, or the concentration of organic matter. It should be noted that applying an additional voltage to a photobioelectrochemical system would defeat the purpose of using it as an alternative for a bioelectrochemical system.

The biofilms were mostly inoculated from wastewater and activated sludge, except for one that used a pure culture of *Shewanella oneidensis* MR-1 [45]. As the results reported on the reference do not include any information on energy efficiency, the amount of electrical power or hydrogen produced cannot be compared between different studies, as more information is needed to standardize the results and make a correct assessment.

### *3.1.2 Electricity generation*

A list of microbial photobioelectrochemical cells that focus on electricity generation is summarized in **Table 5**. These studies also use titanium dioxide [27, 32–34] and copper(I) oxide [32, 48, 51, 52] as the most common photocathode materials. There is a study that used n-type  $\text{Cu}_2\text{O}$ , which sets it apart from other studies because photocathodes are usually made with p-type semiconductors and copper(I) oxide is normally a p-type material [52].

There was one study in which the device used anaerobic sludge as substrate. The increased generation of electricity in comparison to a regular microbial fuel cell demonstrates the potential for microbial photobioelectrochemical cells with chemotrophic bioanode and photocathode to be used in wastewater treatment plants [34].

The most used pure culture for these systems is *Shewanella Oneidensis* MR-1 [27, 32, 53], while microbial consortiums are mostly inoculated from wastewater or anaerobic sediments from large bodies of water [33, 51]. As the energy efficiency was not reported in any of these references, their results in terms of power generation cannot be properly compared.

### *3.1.3 Synthetic materials production*

Reduction of chromium (VI) and simultaneous electricity production was achieved using a photocathode built with graphite coated with titanium dioxide.



Bioanode inoculum	Anolyte	Photocathode	Catholyte	Light source	Internal resistance	External resistance and/or bias	Power / current density	Hydrogen produced (efficiency)	Reference
<i>Shewanella oneidensis</i> MR-1	Trypticase soy broth	Cu <sub>2</sub> O nanowires	Potassium phosphate buffer	Solar simulator	—	—	0.4 mA/cm <sup>2</sup>	—	[45]
Fruit wastewater	Acetate and nutrients	TiO <sub>2</sub> nanorod arrays	Na <sub>2</sub> SO <sub>4</sub>	Xenon lamp	10,000 Ω	10,000 Ω	6 mW/m <sup>2</sup>	4.2 L/m <sup>3</sup> d	[41, 46]
Domestic wastewater	Acetate and nutrients	MoS <sub>3</sub> modified p-type Si nanowires	Phosphate buffer solution	Solar simulator	2000 Ω	500 Ω	71 mW/m <sup>2</sup>	0.1886 L/m <sup>3</sup> d	[25]
Fruit wastewater	Acetate and nutrients	TiO <sub>2</sub>	Na <sub>2</sub> SO <sub>4</sub>	Xenon lamp	—	1000 Ω	474 mW/m <sup>2</sup>	14.1 L/m <sup>3</sup> d	[30]
Activated sludge	Acetate	Cu <sub>2</sub> O coated with NiO <sub>x</sub>	N/A*	Visible light**	—	0.2 V	—	5.09 μL/cm <sup>2</sup> h	[14–16]
N/A	Acetate	p-type polyaniline nanofibers	N/A*	Fluorescent light	—	0.8 V	—	1780 L/m <sup>3</sup> d (66.2%)	[37]
Domestic wastewater	Acetate, nutrients, buffer, and methyl orange	TiO <sub>2</sub> -coated Ni foam	N/A*	UV lamp	8.8 Ω	10 Ω	—	1555.86 L/m <sup>3</sup> d	[38]
Domestic wastewater	Acetate, nutrients, buffer, and methyl orange	g-C <sub>3</sub> N <sub>4</sub> /BiOBr heterojunction	N/A*	Xenon lamp	—	10 Ω	—	143.8 L/m <sup>3</sup> d	[31]
N/A	17.5 mL Acetate, nutrients, minerals, and vitamins	CaFe <sub>2</sub> O <sub>4</sub>	28 mL Na <sub>2</sub> SO <sub>4</sub>	Xenon lamp	—	100 Ω	0.09 mA/cm <sup>2</sup>	36.75 L/m <sup>3</sup> d	[36]

Bioanode inoculum	Anolyte	Photocathode	Catholyte	Light source	Internal resistance	External resistance and/or bias	Power / current density	Hydrogen produced (efficiency)	Reference
Activated sludge	Acetate, nutrients, minerals, and vitamins	GaInP <sub>2</sub> -TiO <sub>2</sub> -MoS <sub>x</sub>	N/A*	Solar simulator	—	0–0.8 V	0.59–12.1 mA/cm <sup>2</sup>	(97% faradaic efficiency)	[43]
N/A	Acetate	Cu <sub>2</sub> O coated with MoS <sub>2</sub>	N/A*	Solar simulator	—	0.8 V	—	2720 L/m <sup>3</sup> d (83%)	[47]
N/A	25 mL Acetate	WO <sub>3</sub> and WO <sub>3</sub> /NiFe <sub>2</sub> O <sub>4</sub>	13 mL Buffer solution (not specified)	Solar simulator	—	0.3 V	0.074 mA/cm <sup>2</sup>	1401.6 L/m <sup>3</sup> d	[35]

\*The device is a single-chamber cell.

\*\*\*The light source was not specified, but it did mention that it was visible light.

**Table 4.** Microbial photobioelectrochemical cells with chemotrophic bioanode and photocathode that produce hydrogen.

Bioanode inoculum	Anolyte	Photocathode	Catholyte	Light source	Internal resistance	External resistance and/or bias	Power / current density	Reference
Domestic wastewater	Anaerobic sludge (10% v/v)	Graphite coated with rutile TiO <sub>2</sub>	Methyl Orange and KCl (1 M)	Xenon lamp	-	500 Ω	0.16 A/m <sup>2</sup>	[34]
N/A	Artificial wastewater with acetate and nutrients	Pd nanoparticle-modified p-type Si nanowires	Methyl orange (25 mg/L)	Xenon lamp	2 kΩ (dark) 670 Ω (light)	1000 Ω	0.119 W/m <sup>2</sup>	[40]
Anaerobic sediments from a lake (10% v/v)	Acetate and nutrients	Graphite coated with rutile TiO <sub>2</sub>	KCl (1 M)	Solar simulator	85 Ω (dark) 65 Ω (light)	1000 Ω	12.03 W/m <sup>3</sup>	[33]
Inoculum from another MFC chamber	Acetate, minerals and nutrients	CuO nanowires	Na <sub>2</sub> SO <sub>4</sub> (0.5 M)	Xenon lamp	354 Ω (dark) 333 Ω (light)	5000 Ω	0.04644 W/m <sup>2</sup>	[48]
Shewanella oneidensis MR-1	Tryptic soy broth	p-n TiO <sub>2</sub> /Cu <sub>2</sub> O/ITO composite junction	Tryptic soy broth	60 W fluorescent light bulb	—	100 Ω	—	[32]
S. oneidensis MR-1	Trypticase soy broth, lactate, and fumarate	Flower-like CuInS <sub>2</sub>	Phosphate buffer solution (0.1 M)	Solar simulator	—	—	1.08 W/m <sup>2</sup>	[49]
N/A	—	3D nitrogen-doped graphene	—	—	—	—	2.607 W/m <sup>2</sup>	[50]
Sea sediment	Sea water	Carbon-coated Cu <sub>2</sub> O	N/A*	white light LED lamp	21.4 Ω (dark) 16.4 Ω (light)	—	0.249 W/m <sup>2</sup>	[51]
S. oneidensis	Glucose and nutrients	TiO <sub>2</sub> nanosheets air cathode	N/A**	365 nm UV lamp	54 Ω	1000 Ω	928 W/m <sup>2</sup>	[27]

Bioanode inoculum	Anolyte	Photocathode	Catholyte	Light source	Internal resistance	External resistance and/or bias	Power / current density	Reference
Domestic wastewater	Wastewater with acetate and nutrients	n-type Cu <sub>2</sub> O doped activated carbon air cathode	N/A**	—	—	1000 Ω	1.39 W/m <sup>2</sup>	[52]

\*The device is a single-chamber cell.  
 \*\*Air cathodes do not have catholyte.

**Table 5.** Microbial photobiorelectrochemical cells with chemotrophic bioanode and photocathode that generate electricity.

The device was able to reduce 97% of an initial concentration of 25 mg/L of Cr(VI) while it was illuminated; moreover, it could still maintain the reduction reaction, albeit slower, while the light was absent [42].

Two devices used polydopamine coated titanium dioxide nanotubes as photocathode material. Both were used to catalyze MoS<sub>2</sub> nanoparticle synthesis. In this case, the focus of the devices was to be used as catalyzers to produce the semiconductor nanoparticles, instead of oxidizing organic matter or generate added value such as electricity or hydrogen. It was demonstrated that MoS<sub>2</sub> obtained through these systems can be used for hydrogen generation under visible light illumination [39, 54].

### **3.2 Photoanode and biocathode systems**

Although the most common photobioelectrochemical systems that have been studied consist of bioanodes and photocathodes, there are systems in which semiconductors are used as anodes instead of cathodes, and microorganisms are used on the cathodes instead of the anodes. This allows the device to use an advanced oxidation process with the photoelectrochemically generated holes in the photoanode and, at the same time, perform reduction processes with the microorganisms at the biocathode.

At the time of this writing, only four publications that use this electrode configuration were found. The most recent one, published in 2017, demonstrated the successful oxidation of rhodamine B with more than 90% efficiency using a photoanode composed of TiO<sub>2</sub> nanotubes with Ag nanoparticles and an oxygen-reducing biocathode, achieving a maximum power density of 0.318 W/m<sup>2</sup>. This research highlighted that the biocathode eliminated the persistent kinetic limitations of abiotic cathodes, which ensured the stable operation of the photoanode. This was also the only study in which the cathodic microbial community was characterized [55].

Another research team developed a nitrogen removal strategy in a single reactor by combining photoelectrocatalytic oxidation of ammonium using an anode composed of AgI/TiO<sub>2</sub> nanotubes and denitrification of nitrates using a cathode with an autotrophic biofilm (Q. [56]).

The two remaining publications were made at the same laboratory and focused on organic pollutant oxidation at the photoanode using TiO<sub>2</sub> nanotubes with coupled nitrification of ammonia nitrogen at the biocathode. One of the publications compared this system with a conventional photoelectrochemical cell with a TiO<sub>2</sub> photoanode and a Pt/C cathode, demonstrating that the photobioelectrochemical device does more pollutant removal and electricity generation [57], while the other focused on the parameter that influences the performance of this device such as pollutant types, electrolyte concentration and gas atmosphere of the photoanode [58].

### **3.3 Phototrophic bioelectrodes**

Cyanobacteria, algae, and similar phototrophic microorganisms can be considered as part of photobioelectrochemical systems because they need light to grow and catalyze reactions. Phototrophic bioelectrodes can be anodic or cathodic and can be used for a wide variety of applications.

#### *3.3.1 Phototrophic bioanodes*

Most of the studies related to phototrophic bioanodes have been aimed only at their potential for electricity generation. This is mainly because these microorganisms are photosynthetic and do not oxidize organic matter. Also, while

chemotrophic bioanodes are usually built with microbial consortiums, phototrophic bioanodes are assembled with pure culture biofilms. **Table 6** summarizes some of the publications found about these electrodes.

Unlike most of the works published about these electrodes, one published work used a phototrophic microbial consortium obtained from sediment and seawater, therefore, it can function as a self-assembled and self-repairing device to generate energy from sunlight. The measurements made on the cell showed that it can produce up to  $0.017 \text{ W/m}^2$  [64].

Also, a research team from Italy developed a photobioelectrochemical cell with phototrophic bioanode and chemotrophic biocathode that generate hydrogen by coupling the photosynthetic capabilities of cyanobacteria in the anode and the dark fermentation process of heterotrophic bacteria degrading an organic substrate. This research shows that cathodic biofilms can consume organic substrates similar to anodic biofilms [65].

Phototrophic bioanodes can be integrated into electronics to supply them with power, as demonstrated by a study that built an invasive ultramicroelectrode array and a microfluidic chamber using silicon microfabrication techniques to immobilize photosynthetic microorganisms and use them in a microtip array as a solar power generator to integrate into solar cells and sensors. The prototype registered a current of 250 pA and 45 mV [66].

### 3.3.2 Phototrophic biocathodes

It is common to find microbial fuel cells with phototrophic biocathodes to be named microbial solar cells, as oxygenic microorganisms can produce the oxygen required for the device to generate electricity without aeration. This research shows that phototrophic cathodic biofilms made of mixed cultures have microorganisms that are either photosynthetically active or catalyzes the reduction of oxygen in the electrode [67]. An oxygenic phototrophic biocathode with a mixed culture of the cyanobacteria *Synechococcus leopoliensis*, *Anabaena cylindrica*, and the algae *Chlorella pyrenoidosa* has been used in a microbial fuel cell to generate two times more electricity in comparison to a regular carbon fiber veil cathode [68].

Phototrophic biocathodes can be inoculated without using pure cultures. One example is an experiment in which the biofilms were grown on a carbon fiber veil that was submerged in pond water for 2 months in a well-illuminated room. Unfortunately, the consortium was not characterized, but the resulting electrode increased the power generation of the microbial fuel cell by 42% in comparison to a regular carbon fiber veil electrode [69]. In other work, a biofilm grown with microorganisms obtained from surface soil from the base of a drainage ditch mixed with distilled water and exposed to sunlight for 1 month was reported. The resulting phototrophic biocathode allowed a sediment-type microbial fuel cell, in which the anode is a bed of sediment, to generate a maximum power density of  $11 \text{ mW/m}^2$  over 6 months without feeding [70].

Some electrodes have been investigated for many other applications. A microbial fuel cell in which both the anode and cathode were inoculated with anaerobic sludge from a wastewater treatment plant achieved a microbial fuel cell with phototrophic bioanode and phototrophic biocathode. The function of the bioanode was to dechlorinate 4-chlorophenol so it can be mineralized at the anode [71]. Another application is using the biomass generated on the cathodic compartment of a microbial fuel cell with chemotrophic bioanode and phototrophic biocathode to feed the anodic compartment, thus creating a device with self-sustainable electricity production when exposed to sunlight [20]. Another way to use these phototrophic biocathodes is to take the light/dark cycles of the microorganisms. The biocathode

Microorganism	Anolyte (mediator)	Cathode catalyst	Catholyte	Light source	Internal resistance	External resistance and/or bias	Power/ current density	Reference
<i>Synechococcus</i> sp. PCC7942	1,4-benzoquinone or diaminodurene	Bilirubin oxidase	2,2'-azinobis(3-ethylbenzothiazolin-6-sulfonate)	15 W fluorescent lamp	130 Ω	50 to 100 kΩ	0.35 W/m <sup>2</sup>	[59]
Anabaena	Alga-Gro medium and 0.01 M methylene blue	Cr/Au	0.02 M potassium ferricyanide in 0.1 M sodium phosphate buffer	—	—	10 Ω	61 μW/L	[60]
<i>Rhodobacter sphaeroides</i>	Sistrom's minimal medium	Pt-coated Toray carbon paper	N/A*	10 W incandescent light	510 Ω	10,000 Ω	0.79 W/m <sup>2</sup>	[61]
<i>Spirulina platensis</i>	Zarrouk medium	graphite carbon cloth	N/A*	White fluorescent lamp	—	1000 Ω	0.01 W/m <sup>2</sup>	[62]
<i>Nostoc</i> sp. ATCC 27893	Phosphate buffer 0.1 M	laccase/ carbon nanotubes	N/A*	Dolan-Jenner Industries Fiber-Lite lamp (model 190)	—	—	0.1 W/m <sup>2</sup>	[26]
<i>Paulschulzia pseudovolvox</i>	Phosphate buffer and NaCl 0.01 M, MgCl <sub>2</sub> 0.005 M	Pt foil	N/A*	FOI-150-220 (150 W and 220 V) with an FOI-5 light guide	—	0.35 V	11.5 μA/cm <sup>2</sup>	[63]

\*The device is a single-chamber cell.

**Table 6.** Microbial photobioelectrochemical cells with phototrophic bioanodes.

can generate oxygen to be used as the electron acceptor when exposed to light and use nitrate as the electron acceptor (and reducing it to N<sub>2</sub> in the process) in dark conditions [17].

### 3.4 Chemotrophic Photobioelectrodes

Chemotrophic biofilms can be grown on semiconductors in a way that the bias generated by the photoelectric effect can enhance the current generated by the microorganisms. These electrodes are referred to as photobioelectrodes in the literature, which agrees with the nomenclature used in this document, except that the word chemotrophic is added here to clarify that the photoelectric effect is produced at the semiconductor and the electrical current is provided by the chemotrophic microorganisms. So far, only chemotrophic photobioanodes have been reported in the literature.

Microorganisms could serve as a protective layer for corrosive semiconductors while providing catalytic functionality. One example of these electrodes is a hematite nanowire electrode with a biofilm of *Shewanella oneidensis* strain MR-1, which shows a synergistic effect in which the power generation and substrate consumption is greater with live cells instead of dead cells or the hematite photoanode alone [53].

Titanium dioxide can be used to build chemotrophic photobioanodes. A TiO<sub>2</sub>/Ti electrode was operated with a biofilm inoculated from non-chlorinated Dutch tap water to reduce phenol with a removal efficiency of 62% after 4 hours of light irradiation [72].

It has been demonstrated that the use of a semiconductor such as  $\alpha$ -Fe<sub>2</sub>O<sub>3</sub> as chemotrophic biofilm support can reduce the resistance of the electrode, accelerate biofilm formation, enrich exoelectrogens, shorten the startup time for the microbial fuel cell and significantly increase the current produced [73]. Scientific studies show that a chemotrophic photobioelectrodes can consist of a stainless steel sheet substrate in which the semiconductor material is deposited on one side and the chemotrophic biofilm is grown on the other side, achieving the same benefits previously reported [16].

The synergistic effect of a chemotrophic photobioelectrode to produce electrical current was used in a desalination cell with anion-exchange membranes and cation-exchange membranes. The metallic anode of the system was modified on one side with nanostructured  $\alpha$ -Fe<sub>2</sub>O<sub>3</sub> and the other side was inoculated with the fresh anodic effluent of an anaerobic granulate sludge blanket reactor. The device achieved a maximum current of 8.8 A/m<sup>2</sup> and a salt removal performance of at least 96%, which demonstrated the capacity for this system to generate electricity and desalinate water [15].

Although phototrophic bioelectrodes and photobioelectrodes have been studied, there are no available parameters that can measure its performance in terms of solar energy conversion. An appropriate place to begin building a diagnostic formula for phototrophic biofilms may be the solar-to-hydrogen efficiency formula used in photoelectrochemistry for water splitting, which can be written as the Eq. (1):

$$\varepsilon_{SH} = \left[ \frac{j_{CC} \left( \frac{mA}{cm^2} \right) \times 1.23 V \times \eta_F}{P_{total} \left( \frac{mW}{cm^2} \right)} \right]_{AM1.5G} \quad (1)$$

In Eq. (1),  $j_{CC}$  is the photocurrent density per electrode area, 1.23 V is the required voltage for water splitting,  $\eta_F$  is the faradaic efficiency for hydrogen production and  $P_{total}$  is the light energy per electrode area when the electrode is irradiated [46]. Results evaluated with this formula are standardized by doing



experiments with a solar simulator that can produce light resembling the AM 1.5G spectrum. It may be possible that the same relationship that is shown in Eq. (1) can be used to evaluate the energetic efficiency of phototrophic bioelectrodes, replacing the voltage for water splitting with the required bias for the reduction reaction that the microorganisms are performing.

#### **4. Conclusions**

Photobioelectrochemical systems have been studied for more than a decade. However, the heterogeneity of materials selected to build electrodes, as well as the many applications that they can achieve has led this research area to be in a relative state of chaos. The lack of standardization of terminology makes it difficult for researchers to learn from the work of others, which in turn slows down the development and improvement of cells that have the possibility of solving many environmental problems while adding value through electricity generation. The nomenclature system proposed in this work aims to contribute to solve this problem, so future reviews could be done more easily and gather more information about these devices.

Moreover, there are still several topics that remain unexplored to this day. There are no publications that report chemotrophic photobiocathodes which, based on the information available on their anodic versions, should present a synergic effect that could increase the performance of reduction reactions made by chemotrophic biofilms.

There are no reports of phototrophic photobioelectrodes. The notion of covering a semiconductor with phototrophic microorganisms can open the possibility of using biofilms as sensitizers. These hypothetical electrodes could function as tandem solar cells, in which phototrophic microorganisms would absorb one part of the solar spectrum and generate current while the semiconductor absorbs another part, generating more current and, possibly, a synergistic effect that remains to be observed.

Lastly, there are no comparisons between microbial photobioelectrochemical systems and microbial bioelectrochemical systems assisted with solar panels to generate their required bias by harnessing sunlight. It is important to know which systems are more energetically and economically efficient so further research can be focused on the technology that has more potential to generate benefit and thus be applied at an industrial scale.

#### **Acknowledgements**

Funding for scholarship was provided partially by both the Universidad de las Américas Puebla (UDLAP), and the Consejo Nacional de Ciencia y Tecnología (CONACYT). LECA is thankful to CONACYT for granting a Ph.D. Scholarship.

#### **Conflicts of interest**

The authors declare no conflicts of interest.

## **Author details**

Luis Erick Coy-Aceves<sup>1</sup>, José Luis Sánchez-Salas<sup>2</sup>, Mónica Cerro-López<sup>2</sup>, Miguel Ángel Méndez-Rojas<sup>2</sup> and Benito Corona-Vázquez<sup>2\*</sup>

1 Departamento de Ingeniería Civil y Ambiental, Universidad de las Américas Puebla, Puebla, México

2 Departamento de Ciencias Químico-Biológicas, Universidad de las Américas Puebla, Puebla, México

\*Address all correspondence to: [benito.corona@udlap.mx](mailto:benito.corona@udlap.mx)

## **IntechOpen**

---

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Sato T, Qadir M, Yamamoto S, Endo T, Zahoor A. Global, regional, and country level need for data on wastewater generation, treatment, and use. *Agricultural Water Management*. 2013;**130**:1-13. DOI: 10.1016/j.agwat.2013.08.007
- [2] UN-Water. WWAP (UNESCO World Water Assessment Programme), 2019, United Nations World Water Development Report 2020. Paris, France: Water and Climate Change; 2020
- [3] Buchauer K. East Asia and Pacific, Wastewater to Energy Processes: A Technical Note for Utility Managers in EAP countries. 2015. 1-253
- [4] Barroso Soares R. Comparative Analysis of the Energy Consumption of Different Wastewater Treatment Plants. *International Journal of Architecture, Arts and Applications*. 2017;**3**(6):79. DOI: 10.11648/j.ijaaa.20170306.11
- [5] Beegle JR, Borole AP. Energy production from waste: Evaluation of anaerobic digestion and bioelectrochemical systems based on energy efficiency and economic factors. *Renewable and Sustainable Energy Reviews*. 2018;**96**(July):343-351. DOI: 10.1016/j.rser.2018.07.057
- [6] Maktabifard M, Zaborowska E, Makinia J. Achieving energy neutrality in wastewater treatment plants through energy savings and enhancing renewable energy production. In *Reviews in Environmental Science and Biotechnology*. 2018;**17**:655-689. DOI: 10.1007/s11157-018-9478-x
- [7] Zou S, He Z. Efficiently “pumping out” value-added resources from wastewater by bioelectrochemical systems: A review from energy perspectives. *Water Research*. 2018;**131**:62-73. DOI: 10.1016/j.watres.2017.12.026
- [8] Santoro C, Arbizzani C, Erable B, Ieropoulos I. Microbial fuel cells: From fundamentals to applications. A review. *Journal of Power Sources*. 2017;**356**:225-244. DOI: 10.1016/j.jpowsour.2017.03.109
- [9] Fischer F. Photoelectrode, photovoltaic and photosynthetic microbial fuel cells. *Renewable and Sustainable Energy Reviews*. 2018;**90**(March):16-27. DOI: 10.1016/j.rser.2018.03.053
- [10] Jiang C, Moniz SJA, Wang A, Zhang T, Tang J. Photoelectrochemical devices for solar water splitting-materials and challenges. *Chemical Society Reviews*. 2017;**46**(15):4645-4660. DOI: 10.1039/c6cs00306k
- [11] Li M, Liu Y, Dong L, Shen C, Li F, Huang M, et al. Recent advances on photocatalytic fuel cell for environmental applications—The marriage of photocatalysis and fuel cells. *Science of the Total Environment*. 2019;**668**:966-978. DOI: 10.1016/j.scitotenv.2019.03.071
- [12] Logan BE, Rossi R, Ragab A, Saikaly PE. Electroactive microorganisms in bioelectrochemical systems. *Nature Reviews Microbiology*. 2019;**17**(5):307-319. DOI: 10.1038/s41579-019-0173-x
- [13] Slate AJ, Whitehead KA, Brownson DAC, Banks CE. Microbial fuel cells: An overview of current technology. *Renewable and Sustainable Energy Reviews*. 2019;**101**(March 2018): 60-81. DOI: 10.1016/j.rser.2018.09.044
- [14] Liang D, Han G, Zhang Y, Rao S, Lu S, Wang H, et al. Efficient H<sub>2</sub> production in a microbial photoelectrochemical cell with a composite Cu<sub>2</sub>O/NiO<sub>x</sub> photocathode under visible light. *Applied Energy*. 2016c;**168**:544-549. DOI: 10.1016/j.apenergy.2016.01.118

- [15] Liang Y, Feng H, Shen D, Li N, Long Y, Zhou Y, et al. A high-performance photo-microbial desalination cell. *Electrochimica Acta*. 2016a;**202**:197-202. DOI: 10.1016/j.electacta.2016.03.177
- [16] Liang Y, Feng H, Shen D, Long Y, Li N, Zhou Y, et al. Metal-based anode for high performance bioelectrochemical systems through photo-electrochemical interaction. *Journal of Power Sources*. 2016b;**324**:26-32. DOI: 10.1016/j.jpowsour.2016.05.059
- [17] Sun J, Xu W, Yang P, Li N, Yuan Y, Zhang H, et al. Enhanced oxytetracycline removal coupling with increased power generation using a self-sustained photo-bioelectrochemical fuel cell. *Chemosphere*. 2019;**221**:21-29. DOI: 10.1016/j.chemosphere.2018.12.152
- [18] Wang AJ, Liang B, Li ZL, Cheng HY. *Bioelectrochemistry Stimulated Environmental Remediation: From Bioelectrorespiration to Bioelectrodegradation*. Singapore: Springer Nature Singapore Pte Ltd.; 2018a. pp. 1-300. DOI: 10.1007/978-981-10-8542-0
- [19] Ali Ansari S, Mansoob Khan M, Omaish Ansari M, Hwan Cho M. Improved electrode performance in microbial fuel cells and the enhanced visible light-induced photoelectrochemical behaviour of PtO<sub>x</sub>@M-TiO<sub>2</sub> nanocomposites. *Ceramics International*. 2015;**41**(7): 9131-9139. DOI: 10.1016/j.ceramint.2015.03.321
- [20] Gajda I, Greenman J, Melhuish C, Ieropoulos I. Self-sustainable electricity production from algae grown in a microbial fuel cell system. *Biomass and Bioenergy*. 2015;**82**:87-93. DOI: 10.1016/j.biombioe.2015.05.017
- [21] Amao Y, Sakai Y, Teshima Y. Photoelectrochemical starch-O<sub>2</sub> biofuel cell consisting of chlorophyll derivative-sensitized TiO<sub>2</sub> anode and enzyme-based cathode. Research on Chemical Intermediates. 2016;**42**(11):7761-7770. DOI: 10.1007/s11164-016-2661-1
- [22] Pankratova G, Pankratov D, Di Bari C, Goñi-Urtiaga A, Toscano MD, Chi Q, et al. Three-dimensional graphene matrix-supported and thylakoid membrane-based high-performance bioelectrochemical solar cell. *ACS Applied Energy Materials*. 2018;**1**(2):319-323. DOI: 10.1021/acsaem.7b00249
- [23] Duan S, Dall'Agnese C, Chen G, Wang XF, Tamiaki H, Yamamoto Y, et al. Bilayer chlorophyll-based biosolar cells inspired from the z-scheme process of oxygenic photosynthesis. *ACS Energy Letters*. 2018;**3**(7):1708-1712. DOI: 10.1021/acsenerylett.8b00797
- [24] Grattieri M, Beaver K, Gaffney EM, Minteer SD. Tuning purple bacteria salt-tolerance for photobioelectrochemical systems in saline environments. *Faraday Discussions*. 2019;**215**:15-25. DOI: 10.1039/c8fd00160j
- [25] Zang GL, Sheng GP, Shi C, Wang YK, Li WW, Yu HQ. A bio-photoelectrochemical cell with a MoS<sub>3</sub>-modified silicon nanowire photocathode for hydrogen and electricity production. *Energy and Environmental Science*. 2014;**7**(9): 3033-3039. DOI: 10.1039/c4ee00654b
- [26] Sekar N, Umasankar Y, Ramasamy RP. Photocurrent generation by immobilized cyanobacteria via direct electron transport in photo-bioelectrochemical cells. *Physical Chemistry Chemical Physics*. 2014;**16**(17):7862-7871. DOI: 10.1039/c4cp00494a
- [27] Wang L, Liu L, Yang F. Efficient gas phase VOC removal and electricity generation in an integrated bio-photo-electro-catalytic reactor with bio-anode and TiO<sub>2</sub> photo-electro-catalytic air

- cathode. *Bioresource Technology*. 2018b; **270**:554-561. DOI: 10.1016/j.biortech.2018.09.041
- [28] Kim MJ, Bai SJ, Youn JR, Song YS. Anomalous power enhancement of biophotovoltaic cell. *Journal of Power Sources*. 2019; **412**(March 2018):301-310. DOI: 10.1016/j.jpowsour.2018.11.056
- [29] Cestellos-Blanco S, Zhang H, Kim JM, Shen Y, xiao, & Yang, P. Photosynthetic semiconductor biohybrids for solar-driven biocatalysis. *Nature Catalysis*. 2020; **3**(3):245-255. DOI: 10.1038/s41929-020-0428-y
- [30] Chen Q-Y, Liu J-S, Liu Y, Liu X-H, Wang Y-H. Influence of operating parameters on the bio-photoelectrochemical cell for hydrogen and electricity production. *Environmental Progress*. 2015; **34**(6):1796-1800. DOI: 10.1002/ep.12156
- [31] Hou Y, Gan Y, Yu Z, Chen X, Qian L, Zhang B, et al. Solar promoted azo dye degradation and energy production in the bio-photoelectrochemical system with a g-C<sub>3</sub>N<sub>4</sub> /BiOBr heterojunction photocathode. *Journal of Power Sources*. 2017a; **371**(August):26-34. DOI: 10.1016/j.jpowsour.2017.10.033
- [32] Gorrepati K. The effect of a p-n TiO<sub>2</sub>/Cu<sub>2</sub>O/ ITO composite junction created using electron beam evaporation on a *Shewanella oneidensis* MR-1 powered microbial coupled photoelectrochemical fuel cell. *Journal of the South Carolina Academy of Science*. 2016; **14**(2):7
- [33] Lu A, Li Y, Jin S, Ding H, Zeng C, Wang X, et al. Microbial fuel cell equipped with a photocatalytic rutile-coated cathode. *Energy and Fuels*. 2010; **24**(2):1184-1190. DOI: 10.1021/ef901053j
- [34] Ding H, Li Y, Lu A, Jin S, Quan C, Wang C, et al. Photocatalytically improved azo dye reduction in a microbial fuel cell with rutile-cathode. *Bioresource Technology*. 2010; **101**(10):3500-3505. DOI: 10.1016/j.biortech.2009.11.107
- [35] Tahir MB. Microbial photoelectrochemical cell for improved hydrogen evolution using nickel ferrite incorporated WO<sub>3</sub> under visible light irradiation. *International Journal of Hydrogen Energy*. 2019; **44**(32):17316-17322. DOI: 10.1016/j.ijhydene.2019.01.067
- [36] Chen Q-Y, Zhang K, Liu J-S, Wang Y-H. Hydrogen and electricity production in a light-assisted microbial photoelectrochemical cell with CaFe<sub>2</sub>O<sub>4</sub> photocathode. *Journal of Photonics for Energy*. 2017; **7**(2):026501. DOI: 10.1117/1.jpe.7.026501
- [37] Jeon Y, Kim S. Persistent hydrogen production by the photo-assisted microbial electrolysis cell using a p-Type polyaniline nanofiber cathode. *ChemSusChem*. 2016; **9**(23):3276-3279. DOI: 10.1002/cssc.201600906
- [38] Hou Y, Zhang R, Yu Z, Huang L, Liu Y, Zhou Z. Accelerated azo dye degradation and concurrent hydrogen production in the single-chamber photocatalytic microbial electrolysis cell. *Bioresource Technology*. 2017b; **224**:63-68. DOI: 10.1016/j.biortech.2016.10.069
- [39] Zeng L, Li X, Fan S, Zhang M, Yin Z, Tadé M, et al. Photo-driven bioelectrochemical photocathode with polydopamine-coated TiO<sub>2</sub> nanotubes for self-sustaining MoS<sub>2</sub> synthesis to facilitate hydrogen evolution. *Journal of Power Sources*. 2019; **413**(May 2018):310-317. DOI: 10.1016/j.jpowsour.2018.12.054
- [40] Han HX, Shi C, Yuan L, Sheng GP. Enhancement of methyl orange degradation and power generation in a photoelectrocatalytic microbial fuel cell. *Applied Energy*. 2017; **204**:382-389. DOI: 10.1016/j.apenergy.2017.07.032

- [41] Chen QY, Liu JS, Liu Y, Wang YH. Hydrogen production on TiO<sub>2</sub> nanorod arrays cathode coupling with bio-anode with additional electricity generation. *Journal of Power Sources*. 2013b;238:345-349. DOI: 10.1016/j.jpowsour.2013.04.066
- [42] Li Y, Lu A, Ding H, Jin S, Yan Y, Wang C, et al. Cr(VI) reduction at rutile-catalyzed cathode in microbial fuel cells. *Electrochemistry Communications*. 2009;11(7):1496-1499. DOI: 10.1016/j.elecom.2009.05.039
- [43] Lu L, Williams NB, Turner JA, Maness PC, Gu J, Ren ZJ. Microbial Photoelectrosynthesis for Self-Sustaining Hydrogen Generation. *Environmental Science and Technology*. 2017;51(22):13494-13501. DOI: 10.1021/acs.est.7b03644
- [44] Book G. IUPAC Compendium of Chemical Terminology. In: Nič M, Jiráč J, Košata B, Jenkins A, McNaught A, editors. USA: International Union of Pure and Applied Chemistry (IUPAC); 2014. Available from: 10.1351/goldbook
- [45] Qian F, Wang G, Li Y. Solar-driven microbial photoelectrochemical cells with a nanowire photocathode. *Nano Letters*. 2010;10(11):4686-4691. DOI: 10.1021/nl102977n
- [46] Chen Z, Dinh HN, Miller E. Photoelectrochemical water splitting: Standards, experimental methods, and protocols. 1st ed. New York, Heidelberg, Dordrecht, London: Springer; 2013a. DOI: 10.1007/978-1-4614-8298-7
- [47] Jeon Y, Kim JH, Koo K, Kim S. A photo-assisted microbial electrolysis cell for the exclusive biohydrogen production using a MoS<sub>2</sub>-coated p-type copper oxide. *Journal of Power Sources*. 2018;373(November 2017):79-84. DOI: 10.1016/j.jpowsour.2017.11.003
- [48] Sun Z, Cao R, Huang M, Chen D, Zheng W, Lin L. Effect of light irradiation on the photoelectricity performance of microbial fuel cell with a copper oxide nanowire photocathode. *Journal of Photochemistry and Photobiology A: Chemistry*. 2015;300:38-43. DOI: 10.1016/j.jphotochem.2014.12.003
- [49] Wang S, Yang X, Zhu Y, Su Y, Li C. Solar-assisted dual chamber microbial fuel cell with a CuInS<sub>2</sub> photocathode. *RSC Advances*. 2014;4(45):23790-23796. DOI: 10.1039/c4ra02488e
- [50] Guo D, Song RB, Shao HH, Zhang JR, Zhu JJ. Visible-light-enhanced power generation in microbial fuel cells coupling with 3D nitrogen-doped graphene. *Chemical Communications*. 2017;53(72):9967-9970. DOI: 10.1039/c7cc04666a
- [51] Jia Y, Zhang D, You H, Li W, Jiang K. Benthic microbial fuel cell equipped with a photocatalytic Cu<sub>2</sub>O-coated cathode. *Journal of Nanoparticle Research*. 2019;21(1):1-10. DOI: 10.1007/s11051-018-4444-7
- [52] Zhang X, Li K, Yan P, Liu Z, Pu L. N-type Cu<sub>2</sub>O doped activated carbon as catalyst for improving power generation of air cathode microbial fuel cells. *Bioresource Technology*. 2015;187:299-304. DOI: 10.1016/j.biortech.2015.03.131
- [53] Qian F, Wang H, Ling Y, Wang G, Thelen MP, Li Y. Photoenhanced electrochemical interaction between shewanella and a hematite nanowire photoanode. *Nano Letters*. 2014;14(6):3688-3693. DOI: 10.1021/nl501664n
- [54] Zeng L, Li X, Fan S, Zhang M, Yin Z, Tade M, et al. Insight into MoS<sub>2</sub> synthesis with biophotoelectrochemical engineering and applications in levofloxacin elimination. In: *ACS Applied Energy Materials*. Vol. 1. Washington, DC: American Chemical Society; 2018. DOI: 10.1021/acsaem.8b00524
- [55] Sui M, Dong Y, Wang Z, Wang F, You H. A biocathode-driven

- photocatalytic fuel cell using an Ag-doped TiO<sub>2</sub>/Ti mesh photoanode for electricity generation and pollutant degradation. *Journal of Photochemistry and Photobiology A: Chemistry*. 2017;**348**:238-245. DOI: 10.1016/j.jphotochem.2017.08.047
- [56] Wang Q, Xu J, Ge Y, Zhang Y, Feng H, Cong Y. Efficient nitrogen removal by simultaneous photoelectrocatalytic oxidation and electrochemically active biofilm denitrification. *Electrochimica Acta*. 2016;**198**:165-173. DOI: 10.1016/j.electacta.2016.03.011
- [57] Du Y, Feng Y, Qu Y, Liu J, Ren N, Liu H. Electricity generation and pollutant degradation using a novel biocathode coupled photoelectrochemical cell. *Environmental Science and Technology*. 2014;**48**(13):7634-7641. DOI: 10.1021/es5011994
- [58] Du Y, Qu Y, Zhou X, Feng Y. Electricity generation by biocathode coupled photoelectrochemical cells. *RSC Advances*. 2015;**5**(32):25325-25328. DOI: 10.1039/C4RA15965A
- [59] Tsujimura S, Wadano A, Kano K, Ikeda T. Photosynthetic bioelectrochemical cell utilizing cyanobacteria and water-generating oxidase. *Enzyme and Microbial Technology*. 2001;**29**(4-5):225-231. DOI: 10.1016/S0141-0229(01)00374-X
- [60] Lam KB, Chiao M, Lin L. A micro photosynthetic electrochemical cell. *IEEE Xplore*. 2003;**2003**:391-394. DOI: 10.1109/memsys.2003.1189768
- [61] Cho YK, Donohue TJ, Tejedor I, Anderson MA, McMahan KD, Noguera DR. Development of a solar-powered microbial fuel cell. *Journal of Applied Microbiology*. 2008;**104**(3):640-650. DOI: 10.1111/j.1365-2672.2007.03580.x
- [62] Lin CC, Wei CH, Chen CI, Shieh CJ, Liu YC. Characteristics of the photosynthesis microbial fuel cell with a *Spirulina platensis* biofilm. *Bioresource Technology*. 2013;**135**:640-643. DOI: 10.1016/j.biortech.2012.09.138
- [63] Hasan K, Çevik E, Sperling E, Packer MA, Leech D, Gorton L. Photoelectrochemical wiring of *pauschulzia pseudovolvox* (Algae) to osmium polymer modified electrodes for harnessing solar energy. *Advanced Energy Materials*. 2015;**5**(22):1-11. DOI: 10.1002/aenm.201501100
- [64] Malik S, Drott E, Grisdela P, Lee J, Lee C, Lowy DA, et al. A self-assembling self-repairing microbial photoelectrochemical solar cell. *Energy and Environmental Science*. 2009;**2**(3):292-298. DOI: 10.1039/b816417g
- [65] Bensaïd S, Ruggeri B, Saracco G. Development of a photosynthetic microbial electrochemical cell (PMEC) reactor coupled with dark fermentation of organic wastes: Medium term perspectives. *Energies*. 2015;**8**(1):399-429. DOI: 10.3390/en8010399
- [66] Ha JG, Song YS, Jung S, Jang S, Kim YK, Bai SJ, et al. Novel microbial photobioelectrochemical cell using an invasive ultramicroelectrode array and a microfluidic chamber. *Biotechnology Letters*. 2017;**39**(6):849-855. DOI: 10.1007/s10529-017-2307-4
- [67] Strycharz-Glaven SM, Glaven RH, Wang Z, Zhou J, Vora GJ, Tender LM. Electrochemical investigation of a microbial solar cell reveals a nonphotosynthetic biocathode catalyst. *Applied and Environmental Microbiology*. 2013;**79**(13):3933-3942. DOI: 10.1128/AEM.00431-13
- [68] Walter XA, Greenman J, Ieropoulos IA. Oxygenic phototrophic biofilms for improved cathode performance in microbial fuel cells. *Algal Research*. 2013;**2**(3):183-187. DOI: 10.1016/j.algal.2013.02.002

[69] Gajda I, Greenman J, Melhuish C, Ieropoulos I. Photosynthetic cathodes for microbial fuel cells. *International Journal of Hydrogen Energy*. 2013; **38**(26):11559-11564. DOI: 10.1016/j.ijhydene.2013.02.111

[70] Commault AS, Lear G, Novis P, Weld RJ. Photosynthetic biocathode enhances the power output of a sediment-type microbial fuel cell. *New Zealand Journal of Botany*. 2014;**52**(1): 48-59. DOI: 10.1080/0028825X.2013.870217

[71] Kong F, Wang A, Ren HY, Huang L, Xu M, Tao H. Improved dechlorination and mineralization of 4-chlorophenol in a sequential biocathode-bioanode bioelectrochemical system with mixed photosynthetic bacteria. *Bioresource Technology*. 2014;**158**:32-38. DOI: 10.1016/j.biortech.2014.01.142

[72] Bennani Y, Peters MCFM, Appel PW, Rietveld LC. Electrochemically active biofilm and photoelectrocatalytic regeneration of the titanium dioxide composite electrode for advanced oxidation in water treatment. *Electrochimica Acta*. 2015;**182**:604-612. DOI: 10.1016/j.electacta.2015.09.101

[73] Feng H, Liang Y, Guo K, Li N, Shen D, Cong Y, et al. Hybridization of photoanode and bioanode to enhance the current production of bioelectrochemical systems. *Water Research*. 2016;**102**:428-435. DOI: 10.1016/j.watres.2016.06.061



# Methods for Persistent Organic Pollutants Removal in Wastewater: A Review

*Valérie Pihen and Jose Luis Sanchez-Salas*

## Abstract

Various persistent organic pollutants (POPs) are increasingly being detected in numerous environmental matrices, including water. Even though there are currently some technologies for the elimination of these pollutants, it is necessary to evaluate their advantages, disadvantages, process time, and cost to find the optimal treatment depending on the characteristics of the pollutants and the matrix to remediate. This work was carried out to compare phase change technologies, advanced oxidation processes, and biological treatments for the elimination of POPs. In this chapter, a recent literature review of the aforementioned methods was performed. Studies are still being carried out to find the best way to eliminate POPs, as this depends on the treatment conditions, the type of water and the policies of each country, but biological treatments seem to be the best option so far.

**Keywords:** persistent organic pollutants, wastewater, treatment processes, phase change technologies, advanced oxidation processes, biological processes

## 1. Introduction

In recent decades, the world population has grown exponentially, generating many negative ecological impacts, derived from residues such as drugs, pesticides, dyes, hormones, and personal care products. These wastes have persistent organic pollutants (POPs) in their composition [1–3].

POPs are usually halogenated and mostly chlorinated compounds that normally have nitro, sulfo, halogen, and/or aromatic residues that are responsible for their recalcitrance [4]. Its carbon–chlorine bonds are very stable against hydrolysis and at a higher number of these bonds, the higher is the resistance to degradation by biological or photolytic action. These POPs usually have ring structures with a single or branched chain. Due to their low solubility in water and high in lipids, they can pass through biological membranes and accumulate in the fat deposits of organisms [5]. Therefore, they can negatively affect human health and several other living organisms, causing mutagenicity, carcinogenicity, reproductive instability, as well as acute and chronic toxicity [3, 6].

The amounts of POPs used around the world have been considered a major concern, which is why 127 countries signed, on May 22, 2001, in Sweden, the Stockholm Convention, as part of the United Nations Environment Program (UNEP). This agreement initially generated a list with 12 of the most used priority toxic substances to prohibit and minimize their use [5, 7]. The number of chemical groups found on

this list continues to grow as new compounds that are part of this classification are identified, such as pentachlorophenol, nonylphenol, octylphenol, and dicofol [8].

The toxic properties of these substances persist for a long time in the environment and can travel long distances before being stored in fatty tissues [5, 7]. POPs are distinguished by being semi-volatile, which allows them to appear in the form of vapor and be present in the atmosphere. These POPs are transported long distances by air, soil, and water, affecting particularly fish and marine mammals. Due to the bioaccumulation properties of such pollutants, they can be transmitted through trophic chains [5, 9].

POPs can be found in large numbers of water bodies in concentrations ranging from  $\text{ngL}^{-1}$  to  $\mu\text{gL}^{-1}$ , even in drinking water, but mainly in industrial, domestic, agricultural, and hospital discharges. These POPs are difficult to degrade with conventional wastewater treatments due to its physicochemical characteristics [1, 9].

Although there are current technologies for advanced wastewater treatment, they have limitations such as their high cost, the formation of toxic by-products and damage to the environment. Therefore, it is important to find, which method is optimal depending on the water characteristics. This chapter reviews the literature of existing processes and emerging technologies for the elimination of persistent organic pollutants in wastewater, analyzing their advantages, disadvantages, comparing time, cost, and degradation conditions of the treatment.

## **2. Treatment processes**

Unconventional wastewater treatment technologies have changed over time, developing a wide range of approaches to pollutant removal. Treatment processes are broadly divided into phase change technologies, advanced oxidation processes, and biological treatments.

Persistent organic pollutants comprise a wide range of various compounds and transformation products, but only those found mostly in the literature will be mentioned. It should be noted that the removal efficiencies were taken directly from the bibliography without any additional modification.

### **2.1 Phase change technologies**

These are processes capable of moving pollutants from one phase to another. They are commonly used in POPs removal and are divided into adsorption processes and membrane technologies filtration.

#### *2.1.1 Adsorption processes*

Different adsorption processes have been studied for the removal of different pollutants, but the use of activated carbon (AC) is the most frequent due to its high porosity and specific surface [10, 11]. In general, good results of POPs removal have been obtained with AC, up to 99% depending on the contaminant and the application time [12, 13]. This type of treatment lasts approximately 90–150 minutes depending on the amount of the adsorbent and the pollutant [14, 15].

Carbon nanotube adsorption (CNT) is another technology, consisting of an allotrope of carbon that has different adsorption characteristics depending on the degree of waviness, diameter, internal geometry, physicochemical properties, and the treatment process used to its synthesis. CNTs are defined as single-walled nanotubes (SWNT), which have an internal diameter of approximately 1.0 nm, and multi-walled nanotubes (MWNT), which consist of several concentric tubes or layers of

laminated graphene. Currently, there are only limited studies of this technology, but single-walled nanotubes are known to be more effective in removing POPs [16, 17].

There are countless studies of the use of clay minerals in adsorption processes, of which it has been observed that, depending on the type of clay, the amount of nitrogen, iron or other present minerals, different removal efficiencies can be produced. Although these approaches have shown very promising results, more research is required as the fate of the contaminants and the removal mechanisms involved remain largely unknown [18, 19].

Other adsorbent materials for POPs removal include zeolites, mesoporous, microporous materials, resins, and metal oxides. In which the nature of the pollutant significantly influences. One limitation of the use of these materials is the sustainability of their production since the use of soils, clays or other natural materials can be unsustainable in the long term. On the other hand, as in the adsorption methods mentioned above, there has been a limited application of these materials for the removal of POPs [20, 21].

**Table 1** shows a comparison of some of the adsorption processes, in terms of removal percentages, degradation conditions and an approximation of the application costs. There is a wide range between the removal percentages depending on the type of contaminant; this may be due to the carboxyl and hydroxyl groups they contain. On the other hand, these processes worked with acidic to neutral pH and with temperatures ranging between 22–30°C. Operating costs vary depending on the type of material used and can range from 0.01 to 7 USD per m<sup>3</sup>.

Process type	POPs	Removal efficiency (%)	Treatment conditions	Approximate cost (USD/m <sup>3</sup> )	Reference
Activated carbon (AC)	Trimethoprim, Tetracycline, Ciprofloxacin, Ibuprofen, Paracetamol, Cephalixin, Diclofenac, Perfluorooctane sulfonate, Perfluorooctanoic acid	60–79	pH = 4-5 T = 25-30°C	0.01	[13, 17]
	Tetracycline, Amoxicillin, Penicillin, Ketoprofen, Naproxen, Diclofenac, Phenol, Polychlorinated biphenyls	88–100	pH = 5-6 T = 25°C		
Carbon nanotubes (CNTs)	Amoxicillin, Ciprofloxacin, Ibuprofen, Triclosan, Tetracycline, Ofloxacin, Norfloxacin	5–99	pH = 4-5 T = 27 °C	7	[17]
Clay minerals	Sulfadimethoxine, Sulfamethoxazole, Tetracycline hydrochloride	13–60	pH = 5-7 T = 22°C		[17]
	Ciprofloxacin, Oxytetracycline, Ampicillin, Tetracycline	90–100			
Other adsorbents	Ciprofloxacin, Chlortetracycline, Oxytetracycline, Carbamazepine	25–75	pH = 5-7 T = 22°C		[17]
	Tetracycline, Diclofenac, Norfloxacin	90–99			

**Table 1.** POPs removal characteristics by adsorption processes.

In general, for adsorption processes, the characteristics of the adsorbent material dictate the efficiency of the removal process, mainly because this determines other properties such as pore size, metallic or non-metallic nature, and the ability to couple with a second treatment [22]. It should be noted that a significant limitation is that most research studies discuss laboratory-scale tests and do not provide information for the expansion or large-scale viability of the processes [17].

### *2.1.2 Membrane technologies*

Membrane processes are another type of phase change technology with great POPs removal capacities; they are based on the use of hydrostatic pressure to eliminate suspended solids and high or low molecular weight solutes, allowing the water passage through. The duration of these processes is approximately 2–8 hours [15]. As in adsorption processes, membranes are produced from different materials with specific filtering characteristics such as pore size, surface charge, and hydrophobicity, which will determine the type of contaminant that can be retained [17, 23].

Ultrafiltration (UF) has been used for the removal of a significant variety of POPs since this process has a pore size in the range of 0.001–0.1  $\mu\text{m}$ . The removal efficiency may vary according to the type of membrane and the contaminant. Generally, highly water-soluble polar pollutants are efficiently removed by ultrafiltration compared to non-polar compounds, poorly soluble in water [24–26].

Nanofiltration (NF) can be used for the removal of some POPs due to its small pore size (10 to 100  $\text{\AA}$ ). This process could be considered more efficient than ultrafiltration for contaminants removal. Another advantage is the lower cost since it operates at low water pressure [24, 25]. However, the half-life of the membranes and their cost should be considered.

Microfiltration (MF) is a technique that has many advantages, but unfortunately, it is not useful for the removal of POPs as it cannot remove contaminants smaller than 1.0  $\mu\text{m}$  [27, 28].

Reverse osmosis (RO) and forward osmosis (FO) use a semi-permeable membrane to separate water from dissolved solids, especially ions, by osmotic pressure. In comparison, reverse osmosis is more effective for the removal of persistent organic pollutants, since it can remove particles up to 10  $\text{\AA}$ , as well as colloidal particles [29, 30].

Depending on the type of process and the type of contaminants, removal efficiencies are highly variable with values between 11 and 99%. This is associated with the molecular weight of the contaminants, the lower the better removal and vice versa. On the other hand, these technologies are of average cost, ranging from 0.3 to 0.5 USD per  $\text{m}^3$ , as shown in **Table 2**.

According to the membrane technologies discussed in this section, as the size of the pores decreases, the efficiency of the POPs removal process improves significantly. However, membrane plugging can occur due to particles and colloids present in the feed streams. These processes are still being updated to achieve greater elimination in terms of quantity and quality of contaminants.

In summary, phase change processes can be effective in removing some persistent organic pollutants. However, the final disposal of the contaminants is challenging, because they pass into the solid phase after treatment, in the case of adsorption, or flow with the rejected effluent, in the case of membrane processes. Therefore, POPs will continue to be a problem for the environment [17].

## **2.2 Advanced oxidation processes**

In recent years there has been great interest in advanced oxidation processes (AOP), due to their great capacity to decompose pollutants, which is associated with

Process type	POPs	Removal efficiency (%)	Treatment conditions	Approximate cost (USD/m <sup>3</sup> )	Reference
Ultrafiltration (UF)	Acetaminophen, Metoprolol, Antipyrine, Sulfamethoxazole, Ketorolac, Atrazine, Hydroxybiphenyl, Diclofenac, Nonylphenol	11-60	In pure water	0.3	[17]
	17 $\alpha$ -ethynilestradiol, Naproxen, Gemfibrozil, Ketoprofen, Bis-2-ethylhexylphthalate, Butylbenzylphthalate, Bisphenol-A, Triclosan, Estrone	70-98			
Nanofiltration (NF)	Acetaminophen, Sulfamethoxazole, Ibuprofen, Naproxen, Atrazine, 17 $\alpha$ -ethynilestradiol, Estrone, Nonylphenol, Bisphenol A, Perfluorohexanoic acid	81-99	In pure water NF-90 and NF-200	0.5	[17]
Reverse osmosis (RO)	Carbamazepine, 17 $\alpha$ -ethynilestradiol, Ibuprofen, Naproxen, Fenoprofen, Gemfibrozil, Ketoprofen, Polycyclic aromatic hydrocarbons	65-99	Aromatic polyamide membrane	0.3	[17, 31]
Forward osmosis (FO)	1,4-dioxane, Acetaminophen, Metronidazole, Phenazone, Bisphenol A	40-99	Hydration innovation	0.3	[17]

**Table 2.**  
 Removal properties of POPs by membrane technologies.

the production of hydroxyl radicals in situ (oxidation potential, 2.8 V) that mineralizes pollutants [17].

AOPs are processes with different routes of free radical production, with specific working conditions and they can involve different materials. These processes have been applied for the elimination of different POPs and their effectiveness depends on the concentration of the oxidizing agent, the pH of the reaction mixture, the chemical structure, the initial concentration of the target contaminant, the wavelength, the intensity of the source, and presence of other organic matter. Therefore, there is no single AOP capable of eliminating all persistent organic pollutants [17].

Some of the most used and best-performing AOPs for POPs removal are UV/H<sub>2</sub>O<sub>2</sub> treatment, Fenton, wet air oxidation, and ozonation. The combined UV/H<sub>2</sub>O<sub>2</sub> process is more effective in degrading POPs in water than UV irradiation or H<sub>2</sub>O<sub>2</sub> oxidation alone, due to photolysis of hydroxyl radicals that generate H<sub>2</sub>O<sub>2</sub> [32].

UV/H<sub>2</sub>O<sub>2</sub> wastewater treatment complemented with microwave irradiation is very effective because of its short reaction time, reduction in activation energy, smaller equipment size, ease of operation and high product performance [15].

The Fenton process is capable of oxidizing aromatic contaminants. Iron (II) reacts with hydrogen peroxide to form iron (III) and hydroxyl radicals. Iron (III) is regenerated back to Fe (II) by hydrogen peroxide in an acidic environment. This technology can be used as a pretreatment method to reduce the toxicity of contaminants. In addition, the Fenton process has variants, such as the Fenton-like and photo-Fenton processes. The Fenton-like process uses iron (III) as a catalyst to convert the reaction from homogeneous to heterogeneous process and is more economical and efficient compared to the classical Fenton process, while having a similar mechanism. On the other hand, the photo-Fenton process is a more efficient and less pH-dependent treatment method, in which hydrogen peroxide can generate hydroxyl radicals under ultraviolet light and iron (III) can accept an ultraviolet photon to regenerate iron (II) [33, 34].

Wet Air Oxidation (WAO) can be used to treat toxic organic wastewater, with high temperature and pressure alone or with catalysts. In this process, the microorganisms mix with gaseous oxygen at temperatures ranging from 150 to 400°C and with pressures of 2 to 40 MPa [35].

Ozonization is a good alternative to degrade POPs at a very fast rate. The method involves a direct reaction between molecular ozone and dissolved compounds or by transformation to oxidants, hydroperoxyl radicals and other species that react with the target compounds [36].

As shown in **Table 3**, AOPs have a high degradation performance for all studies, for which pH is important, with values between 3.5 and 7. Another great advantage is the short time of operation that goes from 0.5 to 8 hours. The main difference between these processes is the source of energy provided, which raises the cost

Process type	POPs	Removal efficiency (%)	Treatment conditions	Approximate time (h)	Reference
UV/H <sub>2</sub> O <sub>2</sub>	Aldrin, Diazinon, Malathion, Steroid estrogens, 17β-estradiol, Estriol, 17α-ethynilestradiol, 4-nonylphenol, Bisphenol A	90-99	Concentration 0.5-10 mgL <sup>-1</sup>	1.5	[15, 37]
Fenton	Polyphenols	82-90	pH = 3.5	8	[15]
Fenton-like	4-nonylphenol	99	T = 60°C pH = 7.4	6.5	[15]
Photo-Fenton	Phenol, Atrazine, Triclosan, Bisphenol A, Ibuprofen, Diclofenac, Ofloxacin, Trimethoprim	90-100	Concentration 0.5 ppm	0.5-2.5	[15, 17, 38]
Wet air oxidation (WAO)	Phenol, 2-chlorophenol	42-74	T = 150-400°C pH = 5	5	[15]
Ozonation	Phenol, <i>p</i> -nitrophenol, Bisphenol A, Naproxen, Ibuprofen, Diclofenac	85-98	pH = 7	1-4	[15, 17]

**Table 3.**  
*Advanced oxidation processes for POPs removal.*

between 7.5 and 9 USD per m<sup>3</sup>, subsequently, solar radiation is a potential alternative to reduce it [15, 17].

### 2.3 Biological processes

In biological processes, microorganisms such as bacteria, fungi, yeasts, even enzymes are used to remove POPs from water. In these processes, microorganisms use the contaminant as a substrate, generating the production of enzymes, transforming the contaminants into smaller molecules that are generally less toxic [39, 40].

The most widely used biological process is activated sludge, so the search for new biological treatments continues and some interesting ones have emerged, such as immobilized cells using a membrane bioreactor. Microalgae are also a developing biological technology capable of removing pesticides from water. Another interesting alternative is fungal biosorption coupled to a membrane bioreactor [39, 41].

The use of yeasts such as *Candida tropicalis* is a new treatment process and is another alternative for the elimination of persistent organic pollutants in water, since these microorganisms have a high degradation potential, a high tolerance to the toxicity of the pollutant, a nearly constant reaction rate, they degrade the contaminant in approximately 66 hours and they can serve as single-cell protein as food. In these processes, the concentration of the contaminant is an important factor, because with a large amount the yeast can die [42–44].

Enzyme treatments are another of the emerging technologies, in which a biocatalyst (an enzyme) is used to transform pollutants. This technique can be applied as a primary treatment or in combination with a biological unit. Different enzymes are used, but the most common ones to degrade POPs are laccases and peroxidases [2, 15]. Although the use of enzymes is an effective technology with proven bioremediation potential, its application is currently a long way off, as it has not yet been incorporated into large-scale water treatment systems. The main reason for the lack of applications of this type of bioremediation is the cost of obtaining pure enzymes, their stability, the number of enzymes required, as well as its lack of reuse [2].

**Table 4** shows a comparison of biological processes, where the wide diversity of actions in these technologies for the elimination of POPs can be observed since to obtain the desired results using them it is necessary to have the correct temperature, pH and ionic strength. Depending on the compound and the treatment conditions, the removal efficiency values vary in a range from 40 to 100%, using a pH of 3–7, with temperatures between 25–37°C and with a removal time of 3–200 hours.

In these technologies, the application of aerobic or anaerobic conditions is related to the acceptance of terminal electrons. Some research indicates that easily biodegradable POPs can be eliminated by this technology, while those with low biodegradability may not go away completely. Then it is recommended to couple them sequentially with other tertiary treatment processes, carry out a pretreatment in the case of some very toxic compounds or find a strain that can endure those [46]. Some proposed methodologies are bioelectrochemical systems (BES), the combination of phase change, biological and electrochemical processes, as well as the use of electrochemical membrane bioreactors (EMBR) [47].

Biological treatment processes have many advantages compared to phase change technologies and advanced oxidation processes, as they are safer, less disruptive, less expensive, require less energy use, are considered green catalysis processes, generate biomass and can be used with contaminants that have very low concentrations, which cannot be achieved by physicochemical techniques. One disadvantage of this method is the high variation in each treatment that depends on a load of organic matter, the concentration of toxic compounds, changes in pH and

Process type	POPs	Removal efficiency (%)	Treatment conditions	Approximate time (h)	Reference
Activated sludge	Phenol, 17 $\beta$ -estradiol, 4-nonylphenol, Naproxen, Trimethoprim, Diclofenac, Ibuprofen, Ketoprofen, Triclosan	40–97	In WWTP	80–200	[15, 17]
Microalgae	Atrazine, Diazinon	78–89	Pilot plant scale, WWTP		[45]
Laccase	Triclosan, Diclofenac, Bisphenol A, Nonylphenol, Triclosan, Diclofenac, Naproxen, Estrone	95–100	T = 25–37°C pH = 5–6	8–24	[2]
Peroxidase	Nonylphenol, Octylphenol, Triclosan, Estrone, Phenol, Paracetamol, Triclosan, Bisphenol A, Naproxen, Diclofenac	80–100	T = 25–37°C pH = 3–7	3–24	[2]
Yeasts ( <i>Candida tropicalis</i> )	Phenol Cresol	90	T = 30°C pH = 6	8–66	[42–44]

**Table 4.**  
*Biological treatment for POPs removal.*

temperature. However, a major disadvantage of biological treatments is the time required, as shown in **Table 4**, so it is possible that microorganisms cannot survive and grow in hard and adverse environmental conditions [48, 49].

### 3. Conclusions

POPs and their by-products, being present in small amounts, have proven to be a challenge for a wide range of water treatment technologies. Conventional treatments such as adsorption in all its variants and membrane technologies have a moderate cost, they are carried out in an acceptable period, but their removal efficiency depends on the pollutant characteristics. On the other hand, advanced oxidation processes have very good removal percentages, their process time is short, but there are methods with a very high cost, and they are not friendly to the environment due to the large amount of energy they consume. However, biological treatments take a little more time compared to other processes, but they have very good removal efficiency, are environmentally friendly, and save energy. It is important to consider that each study carried out works with different treatment conditions, such as pH and temperature, toxic compounds concentrations, organic matter load, and its composition, which influences the success of the tests. While studies are still being conducted to find the best way to remove POPs from water, everything indicates that biological treatments are so far the best option.



## Author details

Valérie Pihen<sup>1\*</sup> and Jose Luis Sanchez-Salas<sup>2\*</sup>

1 Departamento de Ingeniería Civil y Ambiental, Universidad de las Américas Puebla, Puebla, Mexico

2 Departamento de Ciencias Químico-Biológicas, Universidad de las Américas Puebla, Puebla, Mexico

\*Address all correspondence to: [valerie.pihenmz@udlap.mx](mailto:valerie.pihenmz@udlap.mx)  
and [jluis.sanchez@udlap.mx](mailto:jluis.sanchez@udlap.mx)

## IntechOpen

---

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Deblonde T, Cossu-Leguille C, Hartemann P. Emerging pollutants in wastewater: A review of the literature. *International Journal of Hygiene and Environmental Health*. 2011;**214**(6):442-448
- [2] Morsi R, Bilal M, Iqbal HM, Ashraf SS. Laccases and peroxidases: The smart, greener and futuristic biocatalytic tools to mitigate recalcitrant emerging pollutants. *Science of the Total Environment*. 2020;**136572**
- [3] Sauvé S, Desrosiers M. A review of what is an emerging contaminant. *Chemistry Central Journal*, 8, 15. Deegan, A. M., Shaik, B., Nolan, K., Urell, K., Oelgemöller, M., Tobin, J., & Morrissey, A. (2011). Treatment options for wastewater effluents from pharmaceutical companies. *International Journal of Environmental Science & Technology*. 2014;**8**(3):649-666
- [4] Jeon J, Murugesan K, Baldrian P, Schmidt S, Chang Y. Aerobic bacterial catabolism of persistent organic pollutants-potential impact of biotic and abiotic interaction. *Current Opinion in Biotechnology*. 2016;**38**:71-78
- [5] Yarto M, Gavilán A, Barrera J. El Convenio de Estocolmo sobre contaminantes orgánicos persistentes y sus implicaciones para México. *Secretaría de Medio Ambiente y Recursos Naturales*. 2003;**69**:7-28
- [6] Bilal M, Adeel M, Rasheed T, Zhao Y, Iqbal HM. Emerging contaminants of high concern and their enzyme-assisted biodegradation—a review. *Environment International*. 2019;**124**:336-353
- [7] Hung H, Katsoyiannis AA, Bronstrom-Lunden E, Olafsdottir K, Aas W, Breivik K, et al. Temporal trends of persistent organic pollutants (POPs) in arctic air: 20 years of monitoring under the Arctic monitoring and assessment programme (AMAP). *Environmental Pollution*. 2016;**217**:52-61
- [8] Wang Y, Sun Y. The causes of the scientific and regulatory gap in the listing of new persistent organic pollutants in the Stockholm convention. *Environmental Science & Technology*. 2016;**50**:6117-6118
- [9] Bakir A, Rowland SJ, Thompson RC. Transport of persistent organic pollutants by microplastics in estuarine conditions. *Estuarine Coast Shelf Sci*. 2014;**140**:14-21
- [10] Rivera-Utrilla J, Sánchez-Polo M, Ferro-García MÁ, Prados-Joya G, Ocampo-Pérez R. Pharmaceuticals as emerging contaminants and their removal from water. A review. *Chemosphere*. 2013;**93**(7):1268-1287
- [11] Sotelo JL, Rodríguez AR, Mateos MM, Hernández SD, Torrellas SA, Rodríguez JG. Adsorption of pharmaceutical compounds and an endocrine disruptor from aqueous solutions by carbon materials. *Journal of Environmental Science and Health, Part B*. 2012;**47**(7):640-652
- [12] Wong KT, Yoon Y, Snyder SA, Jang M. Phenyl- functionalized magnetic palm-based powdered activated carbon for the effective removal of selected pharmaceutical and endocrine-disruptive compounds. *Chemosphere*. 2016;**152**:71-80
- [13] Zhang D, Luo Q, Gao B, Chiang S, Woodward D, Huang Q. Sorption of perfluorooctanoic acid, perfluorooctane sulfonate and perfluoroheptanoic acid on granular activated carbon. *Chemosphere*. 2016;**144**:2336-2342
- [14] Kulkarni SJ, Kaware JP. Review on research for removal of phenol from

wastewater. *International Journal of Scientific and Research Publications*. 2013;**3**(4):1-5

[15] Villegas LGC, Mashhadi N, Chen M, Mukherjee D, Taylor KE, Biswas N. A short review of techniques for phenol removal from wastewater. *Current Pollution Reports*. 2016;**2**(3):157-167

[16] Ren X, Chen C, Nagatsu M, Wang X. Carbon nanotubes as adsorbents in environmental pollution management: a review. *Chemical Engineering Journal*. 2011;**170**(2-3):395-410

[17] Rodriguez-Narvaez OM, Peralta-Hernandez JM, Goonetilleke A, Bandala ER. Treatment technologies for emerging contaminants in water: A review. *Chemical Engineering Journal*. 2017;**323**:361-380

[18] Li Z, Chang PH, Jean JS, Jiang WT, Wang CJ. Interaction between tetracycline and smectite in aqueous solution. *Journal of Colloid and Interface Science*. 2010;**341**(2):311-319

[19] Rahardjo AK, Susanto MJJ, Kurniawan A, Indraswati N, Ismadji S. Modified Ponorogo bentonite for the removal of ampicillin from wastewater. *Journal of Hazardous Materials*. 2011;**190**(1-3):1001-1008

[20] Ahmed MB, Zhou JL, Ngo HH, Guo W. Adsorptive removal of antibiotics from water and wastewater: Progress and challenges. *Science of the Total Environment*. 2015;**532**:112-126

[21] Mabrouki H, Akretche DE. Diclofenac potassium removal from water by adsorption on natural and pillared clay. *Desalination and Water Treatment*. 2016;**57**(13):6033-6043

[22] Acero JL, Benitez FJ, Real FJ, Teva F. Coupling of adsorption, coagulation, and ultrafiltration processes for the removal of emerging contaminants in a secondary effluent. *Chemical Engineering Journal*. 2012;**210**:1-8

[23] Schäfer AI, Akanyeti I, Semião AJ. Micropollutant sorption to membrane polymers: a review of mechanisms for estrogens. *Advances in Colloid and Interface Science*. 2011;**164**(1-2):100-117

[24] Cissé M, Vaillant F, Pallet D, Dornier M. Selecting ultrafiltration and nanofiltration membranes to concentrate anthocyanins from roselle extract (*Hibiscus sabdariffa* L.). *Food Research International*. 2011;**44**(9):2607-2614

[25] Lidén A, Persson KM. Comparison between ultrafiltration and nanofiltration hollow-fiber membranes for removal of natural organic matter: A pilot study. *Journal of Water Supply: Research and Technology AQUA*. 2016;**65**(1):43-53

[26] Mohammadi S, Kargari A, Sanaeepur H, Abbassian K, Najafi A, Mofarrah E. Phenol removal from industrial wastewaters: A short review. *Desalination and Water Treatment*. 2015;**53**(8):2215-2234

[27] Lu NC, Liu JC. Removal of phosphate and fluoride from wastewater by a hybrid precipitation–microfiltration process. *Separation and Purification Technology*. 2010;**74**(3):329-335

[28] Zhou JE, Chang Q, Wang Y, Wang J, Meng G. Separation of stable oil–water emulsion by the hydrophilic nano-sized ZrO<sub>2</sub> modified Al<sub>2</sub>O<sub>3</sub> microfiltration membrane. *Separation and Purification Technology*. 2010;**75**(3):243-248

[29] Lee KP, Arnot TC, Mattia D. A review of reverse osmosis membrane materials for desalination development to date and future potential. *Journal of Membrane Science*. 2011;**370**(1-2):1-22

[30] Malaeb L, Ayoub GM. Reverse osmosis technology for water treatment: State of the art review. *Desalination*. 2011;**267**(1):1-8

- [31] Smol M, Wlodarczyk M. The effectiveness in the removal of PAHs from aqueous solutions in physical and chemical processes: A review. *Polycyclic Aromatic Compounds*. 2017;**37**(4):292-313
- [32] Zhang A, Li Y. Removal of phenolic endocrine disrupting compounds from waste activated sludge using UV, H<sub>2</sub>O<sub>2</sub>, and UV/H<sub>2</sub>O<sub>2</sub> oxidation processes: effects of reaction conditions and sludge matrix. *Science of the Total Environment*. 2014;**493**:307-323
- [33] Mofrad MR, Nezhad ME, Akbari H, Atharizade M, Miranzadeh MB. Evaluation of efficacy of advanced oxidation processes fenton, fenton-like and photo-fenton for removal of phenol from aqueous solutions. *Journal of the Chemical Society of Pakistan*. 2015;**37**(02):266
- [34] Pariente MI, Molina R, Melero JA, Botas JÁ, Martínez F. Intensified-Fenton process for the treatment of phenol aqueous solutions. *Water Science and Technology*. 2015;**71**(3):359-365
- [35] Chen C. Wet air oxidation and catalytic wet air oxidation for refinery spent caustics degradation. *Journal of the Chemical Society of Pakistan*. 2013;**35**(2):244-250
- [36] Kuosa M, Kallas J, Häkkinen A. Ozonation of p-nitrophenol at different pH values of water and the influence of radicals at acidic conditions. *Journal of Environmental Chemical Engineering*. 2015;**3**(1):325-332
- [37] Belgiorino V, Rizzo L, Fatta D, De la Roca C, Lofrano G, Nikolaou A, et al. Review on endocrine disrupting-emerging compounds in urban wastewater: Occurrence and removal by photocatalysis and ultrasonic irradiation from wastewater reuse. *Desalination*. 2007;**215**:166-176
- [38] Klammerth N, Malato S, Agüera A, Fernandez-Alba A, Mailhot G. Treatment of municipal wastewater treatment plant effluents with modified photo-Fenton as a tertiary treatment for the degradation of micro pollutants and disinfection. *Environmental Science & Technology*. 2012;**46**:2885-2892
- [39] Ahmed MB, Zhou JL, Ngo HH, Guo W, Thomaidis NS, Xu J. Progress in the biological and chemical treatment technologies for emerging contaminant removal from wastewater: A critical review. *Journal of Hazardous Materials*. 2017;**323**:274-298
- [40] Tran NH, Urase T, Ngo HH, Hu J, Ong SL. Insight into metabolic and cometabolic activities of autotrophic and heterotrophic microorganisms in the biodegradation of emerging trace organic contaminants. *Bioresource Technology*. 2013;**146**:721-731
- [41] Grandclement C, Syssiecq I, Piram A, Wong-Wah-Chung P, Vanot G, Tiliacos N, et al. From conventional biological wastewater treatment to hybrid processes, the evaluation of organic micropollutant removal: A review. *Water Research*. 2017;**111**:297-317
- [42] Adav SS, Chen MY, Lee DJ, Ren NQ. Degradation of phenol by aerobic granules and isolated yeast *Candida tropicalis*. *Biotechnology and Bioengineering*. 2007;**96**(5):844-852
- [43] Jiang Y, Wen J, Li H, Yang S, Hu Z. The biodegradation of phenol at high initial concentration by the yeast *Candida tropicalis*. *Biochemical Engineering Journal*. 2005;**24**(3):243-247
- [44] Jiang Y, Wen J, Bai J, Wang D, Hu Z. Phenol biodegradation by the yeast *Candida tropicalis* in the presence of m-cresol. *Biochemical Engineering Journal*. 2006;**29**(3):227-234

[45] Matamoros V, Gutierrez R, Ferrer I, Garcia J, Bayona JM. Capability of microalga-based wastewater treatment systems to remove emerging organic contaminants: A pilot-scale study. *Journal of Hazardous Materials*. 2015;288:34-42

[46] Sui Q, Huang J, Deng S, Chen W, Yu G. Seasonal variation in the occurrence and removal of pharmaceuticals and personal care products in different biological wastewater treatment processes. *Environmental Science & Technology*. 2011;45(8):3341-3348

[47] Yang X, Flowers RC, Weinberg HS, Singer PC. Occurrence and removal of pharmaceuticals and personal care products (PPCPs) in an advanced wastewater reclamation plant. *Water Research*. 2011;45(16):5218-5228

[48] Al-Maqdi KA, Hisaindee SM, Rauf MA, Ashraf SS. Comparative degradation of a thiazole pollutant by an advanced oxidation process and an enzymatic approach. *Biomolecules*. 2017;7(3):64-76

[49] Rauf MA, Salman Ashraf S. Survey of recent trends in biochemically assisted degradation of dyes. *Chemical Engineering Journal*. 2012;209:520-530



# A Critical Review on Algal-Bacterial Granular Sludge Process as Potential Economical Alternative to AOPs for Textile Wastewater Treatment

*Celina Sanchez-Sanchez, Guillermo Baquerizo  
and Ernestina Moreno-Rodríguez*

## Abstract

Textile wastewaters are complex effluents that may cause severe effects on environmental and human health when not treated properly. Traditionally, textile wastewater treatment has been carried out through conventional processes using individual or combined advanced oxidation processes (AOPs) with acceptable removal efficiencies. Although AOPs may remove textile pollutants, the operational costs are expensive. The ABGS process is a technology based on symbiotic microbial interactions between algal and bacterial consortia. This novel biological process could be a promising alternative versus conventional technologies for textile wastewater treatment due to having a higher capability for effectively removing nutrients and refractory organic pollutants. Therefore, this chapter presents a critical review of Algal-Bacterial Granular Sludge processes as an effective and economical alternative treatment for textile wastewater to comparison to AOPs. Furthermore, the symbiotic characteristics of ABGS processes allow to decrease the environmental effects and operating costs and improve biomass production containing high value-added compounds. Therefore, the ABGS-based processes could be considered a feasible technology for textile wastewater treatment with lower energy requirements than conventional treatments. Although to guarantee ABGS optimal performance, algae and bacteria strains must be selected according to their adaptive capacity to textile wastewater characteristics.

**Keywords:** ABGS process, textile wastewater, AOPs, cost estimation

## 1. Introduction

Textile wastewaters are defined as complex industrial effluents due to their wide variability of chemical compounds. Most of the compounds present in the textile effluents can be classified as refractory and priority pollutants, and thus showing low biodegradability. Uncontrolled disposal may provoke severe environmental effects, on receiving the water bodies and thus in human health. This low

biodegradability of textile wastewater is generally attributable to the presence of recalcitrant organic compounds, including dyes, phenolic pollutants, and dyeing aids, where conventional treatment methods are ineffective in degrading them successfully [1, 2]. The latter has led experts to explore diverse practices technologies for removing this type of pollutants. These practices have comprised physical, biological, and chemical processes, in which the conventional biotechnologies and advanced oxidation processes (AOPs) are the commonly selected treatments.

Regarding AOPs, the ozonation ( $O_3$ ), hydrogen peroxide ( $H_2O_2$ ), UV-irradiation, Fenton-based process, and combinations among them, are well known for their capacity to reach efficient pollutant removals. However, these processes may generate high loads of toxic wastes during their application. Also, AOPs usually require significant processing costs due to the overall cost of AOPs being typically represented by their capital, operating, and maintenance costs [3].

On the other hand, conventional biotechnologies such as activated sludge processes, facultative lagoons, wetlands, etc., are considered reliable technologies able to treat wastewater from economic and environmental aspects. However, the conventional biotechnologies exhibit not only high operating costs derived from aeration requirements but also a large footprint [4]. Although the activated sludge processes have been widely used for industrial wastewater treatment, most of the dyes are poorly removed by using this technology. Besides, when textile wastewaters are mixed with sewage, the conventional biological processes have been shown to be inefficient in decolorizing textile effluents. Moreover, direct biological oxidation is not possible due to the presence of recalcitrant organic molecules of dyes [5]. Therefore, textile industry effluents after biological treatment are not inconsistent with the wastewater quality standards established in the legislation applicable [6]. On the other hand, biological and chemical processes have been combined to achieve better removal efficiencies and even to accomplish pollutants mineralization with more competitive treatment costs compared to AOPs, although operating costs are still considerable.

Therefore, in the last decade, great attention has been paid to novel biological technologies for wastewater treatment based on microalgae in combination with bacteria, forming the called microalgae-bacteria symbiotic granular sludge. The algal-bacterial granular sludge (ABGS) process has been positioned as a cost-effective technology capable of simultaneously removing nutrients and mineralizing recalcitrant organic pollutants from municipal and industrial wastewater [4]. The presence of algae in ABGS systems not only provides oxygen for the bacterial metabolism but also maintains the structure of the original biochemical pool while further reducing the pollutants. Besides, the low concentration of organic matter in the resulting effluent may reduce the cost of the subsequent chemical treatment [7].

In this sense, the goal of this chapter is to present an overview of ABGS-based processes as an economic and friendly treatment alternative for textile wastewater with potential for resource recovery. In the following sections of this chapter, the environmental effects caused by textiles effluents, AOPs characteristics, and ABGSs potential are discussed in detail. Section 2 briefly presents the environmental, human health, and marine biota effects caused by untreated textile wastewater effluents or treated by conventional processes. In Section 3, the advantages, disadvantages, and operating costs of the main AOPs used for textile wastewater treatment are discussed. The selection of the most suitable physicochemical treatment process is determined in each particular case, thus the operating cost of AOPs for textile wastewater treatment varies depending on the energy/reagent amount used. Section 4 provides a complete description of the ABGS system's potential for the remediation of textile wastewater treatment from environmental, economic, and resource recovery aspects. Finally, Section 5 presents the main conclusions of the conducted research.



## 2. Environmental impacts of untreated textile wastewater effluents

The textile and clothing industry is considered a vital industry, with a global market of over U\$450 billion, in terms of nominal sales, contributing with 7% of the total world exports and employing around 35 million workers around the world. Additionally, it is one of the most polluting industrial sectors [8, 9]. The major activities from textile industry that may cause severe environmental impacts are associated with—(1) energy consumption throughout the production of man-made fibers, in yarn manufacturing, in finishing processes, and in washing and drying clothes; (2) solid waste generated from textile and clothing manufacturing and, mostly, from the disposal of products at the end of their cycle life; (3) direct carbon dioxide (CO<sub>2</sub>) emissions, particularly related to transportation processes within globally-dispersed supply chains; and (4) water large volumes consumption and chemical products requirement associated with fiber growth, wet pretreatment, dyeing and finishing activities, and laundry [9, 10].

The United Nations Framework Convention on climate change indicated that textile and fashion industries produced about 20% of global wastewater in 2018 [11]. Textile mills are also responsible for being a 20% contributor to the world's industrial water pollution, using thousands of toxic chemicals during production, some of them cataloged as carcinogenic compounds [12]. A broad type of chemical product is used in the textile industry including inorganic compounds, polymers, and organic products with very complex compositions. The effluents from textile processes are characterized by alkaline reaction, significant salinity, intensive color, and toxicity since they contain dyes, heavy metals (like chromium, cobalt, and copper), pentachlorophenol, chlorine bleaching, halogen carriers, carcinogenic amines, free formaldehyde, biocides, salts, surfactants, disinfectants, solvents, and softeners [13, 14]. **Table 1** shows the average values of main wastewater parameters according to different types of textile processing, variability in pollutant concentrations, and operating conditions present in the discharged effluent.

From **Table 1**, process type (raw wool scouring) may reach high values of chemical oxygen demand (COD) and biochemical oxygen demand (BOD), which are difficult to degrade by using conventional biological treatments. By contrast, type 3 (woven fabric finishing) with the highest BOD/COD ratio suggest that biological-based treatment can be a suitable option. On the other hand, process type 7 (stock and yarn dyeing and finishing) is of major concern since selected dyes are mainly water-soluble and nonbiodegradable with an important content of recalcitrant compounds. In this context, according to the processing type carried out will be the complexity of pollutant loads generated, and thus, the type of wastewater treatment required by textile effluent discharged.

More than 100,000 commercial dyes are currently available worldwide, with over 1 million tons of dye-stuff produced annually [16]. However, 10% of produced dyes are released to the environment in the form of toxic effluents [13, 17]. Many dyes are difficult to decolorize due to their complex structure and synthetic origin. Brightly colored, water-soluble reactive azo and acid dyes are the main concerns. Indeed, they normally pass through conventional biological treatments without suffering metabolic degradation [18], textile dyes may cause allergic reactions, carcinogenicity, mutagenicity, and cytotoxicity effects on various plants, rats, fishes, mollusks, microbes, and mammalian cells [19]. Even, low concentrations of dyes in effluents are highly visible and undesirable.

The environmental effects caused by textile pollutants include the detriment of the visual aspect of superficial water bodies, which in turn interferes with aquatic biological processes, prevents penetration of light and, causes eutrophication in

Parameters	Production type						
	1	2	3	4	5	6	7
BOD/COD (mg/L)	0.2	0.29	0.35	0.54	0.35	0.3	0.31
BOD (mg/L)	6000	300	350	650	350	300	250
TSS (mg/L)	8000	130	200	300	300	120	75
COD (mg/L)	30,000	1040	1000	1200	1000	1000	800
Oil and grease (mg/L)	5500	—	—	14	53	—	—
Total chromium (mg/L)	0.05	4	0.014	0.04	0.05	0.42	0.27
Phenol (mg/L)	1.5	0.5	—	0.04	0.24	0.13	0.12
Color (ADMI)	2000	1000	—	325	400	600	600
pH	8.0	7.0	10	10	8.0	8.0	11
Temperature (°C)	28	62	21	37	39	20	38

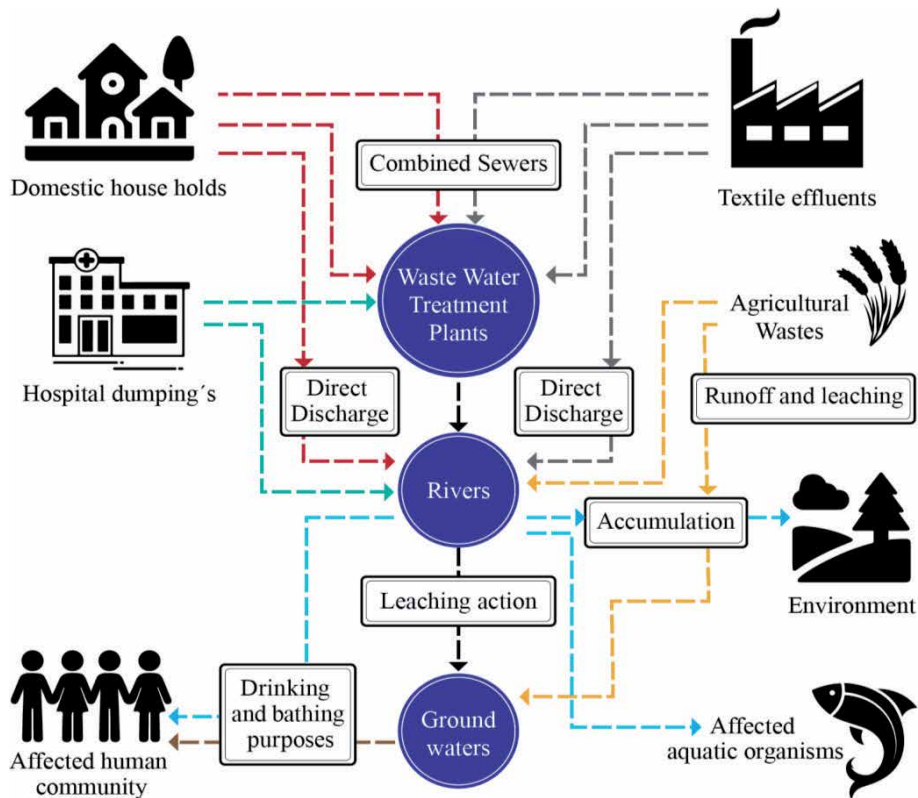
Categories description—(1) raw wool scouring; (2) yarn and fabric manufacturing; (3) wool finishing; (4) woven fabric finishing; (5) knitted fabric finishing; (6) carpet finishing; (7) stock and yarn dyeing and finishing. TSS, total suspended solids. Adapted from [15].

**Table 1.**  
Effluent characteristics from different production processes.

water bodies [20]. It is known that effluent with a high concentration of COD (from 800 to 30,000 mg/L) indicates the presence of recalcitrant toxic compounds that can lead to depletion of dissolved oxygen in the receiving water bodies [21]. The strong development of synthetic dyes has caused several detrimental effects on the environment and human health as an improper discharge of strongly colored effluents and the associated metabolites in aqueous ecosystems to reduce sunlight penetration, causing inhibitory effects on photosynthesis, and the presence of aromatic amines generated when dyes are broken anaerobically, which are toxic, carcinogenic and mutagenic [22, 23].

On the other hand, textile wastewaters also contain many persistent organic pollutants, such as phenols, aromatic amines, dioxazine, anthraquinone, consider toxic chemicals that can be transported by wind and water. Moreover, these pollutants may persist for long periods in the environment and can be accumulated, passing from one species to the next through the food chain [15]. **Figure 1** shows the sources, transport pathways, and fates of persistent organic pollutants of textile wastewater. The impacts on human health and the environment are also shown.

From **Figure 1**, a portion volume of textile wastewater is usually mixed with municipal wastewater, to be treated in wastewater treatment plants through conventional biological processes, while the remaining volume is directly discharged into rivers. Also, another portion volume of untreated hospital and agricultural wastewater is discharged to the rivers contributing to severe contamination to receiving water bodies. The discharge of a wide pollutants variability from all these sources generates wastewater very complex which is transported and bioaccumulated by rivers, which severely impacts to marine biota presents in the superficial waters, to the exposed population and reached environment. It is estimated that 10% of textile chemicals are potentially toxic to human health (i.e., carcinogenic compounds), and about 5% of these substances are highly toxic to the environment [19, 25]. Therefore, the minimization of the discharges of untreated textile wastewater into the environment together with improving the processes for textile wastewater treatment through more efficient and sustainable technologies arises as an obligatory task.



**Figure 1.**  
*Source, transport pathways, and the fate of persistent organic pollutants in the environment. Adapted from [24].*

### 3. Operating cost of main AOPs used for textile wastewater treatment

AOPs are processes based on the production and utilization of hydroxyl ( $\text{OH}\cdot$ ) radicals [18]. These processes can be broadly classified into four groups—photocatalytic process ( $\text{H}_2\text{O}_2/\text{UV}$ ,  $\text{O}_3/\text{H}_2\text{O}_2/\text{UV}$ ,  $\text{UV}/\text{TiO}_2$ ,  $\text{H}_2\text{O}_2/\text{TiO}_2/\text{UV}$ ,  $\text{O}_3/\text{TiO}_2/\text{UV}$ ); the Fenton reaction-based processes ( $\text{Fe}_2^+/\text{H}_2\text{O}_2$ ,  $\text{Fe}_2^+/\text{H}_2\text{O}_2/\text{UV}$ , Electro-Fenton), ozone-based processes ( $\text{O}_3/\text{UV}$ ,  $\text{O}_3/\text{H}_2\text{O}_2$ ,  $\text{O}_3/\text{Fe}^{++}$ ,  $\text{O}_3/\text{metal oxide catalyst}$ ,  $\text{O}_3/\text{activated carbon}$ ,  $\text{O}_3/\text{ultrasound}$ ); and other processes which may include activated persulfate, ionizing radiation or electron beam technology [2], the AOPs can be applied either individually or in combination among them. Thus, the operating costs of AOPs applied to wastewater treatment may widely vary according to the type and the amount of energy/reagent used.

Although the mechanisms of AOPs rely on the formation of  $\text{OH}\cdot$  radicals, the formation pathways might be different under different operating conditions [26], showing a strong impact on the estimated cost for the selected treatment process. Moreover, AOPs for complete pollutant mineralization are generally expensive because the intermediates formed during treatment tend to be even more resistant to complete chemical degradation; the intermediates treatment also represents a substantial part of energy and chemicals, which increase with the treatment duration [27].

In this sense, the cost-effectiveness of each technology is one of the main concerns of decision-makers. It must be considered the different costs involved and the efficiencies achieved in the proposed treatment system. The total costs estimation

AOP	Energy cost (U\$-m <sup>3</sup> )	Reagent cost (U\$-m <sup>3</sup> )	Observations
O <sub>3</sub>	5.28	—	High cost. Extremely short half-life (20 minutes at most). Incomplete degradation. Possible generation of intermediates. Unstable method.
UV-based process	1.1	—	Pretreatment needed to remove suspended solids. High cost. Treatment length limited.
Fenton process	—	0.519	Unable to remove disperse dyes. Generation of High iron-content sludge. Long reaction time. Works only at low pH. Production of toxic by-products
Photo-Fenton process	1.1	0.519	High cost. Better performance at low pH (i.e., 3.0).
UV/persulfate	0.608	3.99	Pretreatment needed to remove suspended solids, radical scavengers, and competing ions
UV/H <sub>2</sub> O <sub>2</sub>	1.14	0.304	Expansive. Formation of the high amount of by-products
O <sub>3</sub> /UV	8.5	—	High cost of energy and equipment
O <sub>3</sub> /H <sub>2</sub> O <sub>2</sub>	5.35	0.5	Generation of oxidation-refractory compounds
O <sub>3</sub> /UV/H <sub>2</sub> O <sub>2</sub>	6.2–11	0.29	High cost. Treatment length limited.

*Adapted from [3, 18, 27–31].*  
*O<sub>3</sub>, ozone-based process.*

**Table 2.**  
*Operating cost in terms of energy and chemical consumption for textile wastewater treatments.*

comprises installation, operation, maintenance and additional requirements of the AOP used and that could arise during the process. **Table 2** presents the operating cost related to energy and chemical consumption for AOPs commonly used for the treatment of textile wastewaters.

### 3.1 Ozone-based processes

Ozone (O<sub>3</sub>) is a strong oxidant, characterized by its extremely high oxidative potential (E<sup>0</sup> 2.07 V), which can decompose many hardly degradable pollutants [32, 33]. During ozonation, the organic compounds can be oxidized in two ways. In the first oxidation way, the generated ozone, which is a highly selective oxidant, can react directly with dissolved organics at variable rates. In the second way, ozone is involved in a chain reaction mechanism to form hydroxyl radicals, which are responsible for pollutant decomposition/oxidation [3]. The two pathways can lead to several final products, with different transformation kinetics and represent different treatment costs.

The increasing popularity in recent years of ozone applications is mainly explained by two factors—(1) costs associated with ozone production have considerably decreased in the last decade, and (2) ozone presents some environmental advantages over chlorine. The benefits of ozonation in wastewater treatment plants include sludge reduction and removal of recalcitrant organic contaminants from hazardous wastewater. However, the ozone is an unstable gas that must be generated *in situ* and the associated generation process is still considered an expensive technology [2].

In this sense, ozone treatment costs for textile wastewater involve installation and maintenance costs in the site. Ozonation technology cost is defined by the cost of an ozone generator and its cooling system. The process is also affected by the cost of a pretreatment unit for drying the oxygen (or air) that fed the ozonator; and a post-treatment system for treating the residual ozone in the off-gas, that is, a catalytic ozone destruction unit [3]. Depending on the water quality requirements and treatment objectives, the estimation of operating cost is impacted by different design variables, including flow rate, site constraints, type of manufacturer, among others, which determine the applied ozone doses [31]. Operation and maintenance costs are based on the energy consumption and replacement part costs. In the case of ozone-based treatment, its high costs are mainly related to energy consumption, and the cost of equipment for oxygen or air generation [31, 34].

The removal efficiencies using an ozone-based process in the treatment of textile wastewater may attain values of 97% for color removal and 60% for phenol removal [35, 36]. For water contaminated with phenol, treatment costs are \$0.03 and \$0.51/L using  $O_3$  and  $O_3/UV$ , respectively. This same tendency is observed in the case of water contaminated with reactive azo dye, with treatment costs of \$0.04 and \$0.24/L for  $O_3$  and  $O_3/UV$  processes, respectively [3]. From **Table 2**; the ozonation process is the most expensive AOP with operating costs up to 10 and five times higher than Fenton and UV-irradiation processes, respectively. Even, it has higher costing than UV/ $H_2O_2$  and UV/persulfate combined processes up to 265.7 and 14.8%, respectively. Although the ozonation process is an effective method to degrade several toxic recalcitrant pollutants, it is still viewed as an expensive technology in application aimed to complete substance mineralization [37], compared with other AOPs.

### 3.2 Fenton and Photo-Fenton processes

The Fenton process is based on the enhanced oxidative potential of hydrogen peroxide ( $H_2O_2$ ) when iron (Fe) is used as a catalyst under acidic conditions. The Fenton reaction mechanism is well-known where—the  $Fe^{+3}$  ion, dissolved in water, form different complexes. For instance, at pH close to 3, the pentaqua-iron (III) hydroxide ( $[Fe(H_2O)_5(OH)]^{2+}$ ) becomes the predominant stable species [38]. On the other hand, the combination of  $H_2O_2$  and UV radiation with a  $Fe^{2+}$  or  $Fe^{3+}$  ion,  $Fe(OH)^{2+}$ , Fe-radical, etc., produces more hydroxyl radicals and in turn, increases both the degradation rate of persistent organic pollutants, as the applicability of the process. The latter process is known as the Photo-Fenton process [39].

The Photo-Fenton application in textile wastewater is able to remove a wider range of pollutants than the Fenton-based process since is considered the most effective treatment for the decolorization of wastewater. It also provides high energy efficiencies compared to other AOPs [25]. Several studies have shown that the treatment of textile wastewater using Fenton and Photo-Fenton processes resulted in 74 and 87% of COD removal, respectively, while color removal through Fenton's oxidation process for direct blue 71, and acid orange 24, reach efficiencies of 94 and 92.7%, respectively [38, 40, 41].

Additional advantages of Fenton's technologies include a simple application procedure, low investment cost, lack of residues, ability to treat complex compounds, and low environmental impacts [5]. However, the main drawbacks of the Fenton and Photo-Fenton process are the sludge production and the discarded unused ferrous ions, especially in the case of the homogenous processes. Additionally, textile wastewater is usually generated at alkaline conditions (see **Table 1**), while the Fenton processes usually require a pH of around 3. Furthermore, the addition

of strong acids may even prove counterproductive for ensuring optimal treatment conditions since precipitation phenomena can appear at low pHs [30].

Commonly the iron species involved in the oxidation is normally found at concentrations between 10 and 150 mg/L in textile wastewater, which may cause a hindrance in the effluent and thus prevent water reuse in the textile industry. High iron concentrations can cause fabric stains and wear during bleaching and dyeing [42]. In this sense, the Fenton processes are better recommended for the pretreatment of textile wastewater.

Moreover, the conventional Fenton process is mainly influenced by the cost of the required chemical reagents as hydrogen peroxide, ferrous iron, and those aimed at pH adjustment. Many studies have addressed the significant impact of hydrogen peroxide on process costs for Fenton-based technologies [3]. For this process, the costs of Fenton reagents like  $H_2O_2$  and  $FeSO_4$ ; and acidification/neutralization chemicals ( $H_2SO_4$  and  $NaOH$ ) are commonly used for operating cost calculation, while energy consumption is considered negligible. However, in the Photo-Fenton processes, energy plays an important role, which has a considerable impact on the treatment costs of the process [27].

Photo-Fenton processes are considered more energy-efficient than other AOPs [25, 39]. However, Fenton-based processes are considerably less expensive than Photo-Fenton technology in terms of total costs. From **Table 2**, the Fenton process is the most economical AOP among all reported processes with an annual operating cost of about 92% lower compared to the ozonation process. However, the operating cost of the Photo-Fenton process rises over the UV-based and UV/ $H_2O_2$  combined processes in 47.2 and 12.12%, respectively. The above is explained by the electrical energy requirement for treatment application. In this sense, the Photo-Fenton is ranked as the second most expensive AOP when it is applied as a single treatment process.

### 3.3 UV-based processes

Ultraviolet (UV) light (200–400 nm) is usually utilized for degrading organic compounds by direct photolysis [33]. The compounds absorb UV light and undergo degradation from their excited state. The types of lamps commonly applied in photolysis reactions are low-pressure UV and medium-pressure UV lamps [3], which define the cost of process installation and operation.

Advantages of UV-irradiation processes in the treatment of textile wastewater treatment include no sludge production, no use of hazardous chemicals, and no generation of unpleasant odors [29]. However, UV-based systems are widely known as energy-intensive processes with important maintenance costs. Electricity consumption of UV irradiation is the major contributor to the total operating cost. Moreover, it has been reported that routine replacement of key parts in a UV system may equate to about 45% of the annual electrical power consumption costs [26].

This technology has been applied as a single treatment for textile wastewater [33, 43], although it is commonly combined with other techniques such as UV/ $H_2O_2$ , UV/ $O_3$ , or UV/ $H_2O_2$ / $O_3$  and thus the operating costs may vary depending on contact time, light intensity, dose, among other operating conditions [44]. It has been reported that treatment cost wastewater contaminated with phenol and reactive azo dyes using  $H_2O_2$ /UV might reach values of \$1.64 and \$0.32/L, respectively [3].

The combination of these advanced oxidation processes tends to substantially improve the removal efficiencies of textile pollutants since possess the ability to provide high color removals in the range of 80–100% after 45–120 minutes of reaction, and different behaviors in terms of COD removal [42]. However, Paździor

*et al.* [30] reported that investment costs, particularly in UV lamp equipment, may significantly increase the total costs of wastewater treatment (approximately 47% compared to ozonation). The authors demonstrated that operational costs of AOPs based on UV radiation may increase 44% in comparison to other oxidation technologies, including the Photo-Fenton process, and UV-process/ozonation.

From **Table 2**, the operating cost of the UV-irradiation process is competitive compared to other AOPs. However, the combination of UV-irradiation with other processes (i.e., O<sub>3</sub>, H<sub>2</sub>O<sub>2</sub>, or persulfate) increases up to eight times the final operational cost, which becomes them the process most expensive among all available AOPs. Just the operating cost of the combined UV/O<sub>3</sub> process may cost up to \$8.5 m<sup>-3</sup>, which is approximately equal to the total operating cost of all the individual oxidation processes. Therefore, the application of each treatment process either single or combined will depend on the type of textile wastewater treated, target pollutants, desired level of mineralization of the pollutants, and applicable legislation for the particular case.

The efficiency of these technologies mainly depends on the volume of wastewaters to be treated, the concentration and nature of the specific pollutants, and the co-occurrence of other substances [45]. In this sense, the operating cost in each AOP for wastewater treatment varying according to application strategy and energy/reagent amount used.

#### **4. Potential of the ABGS systems in the remediation of textile wastewater treatment**

Algal-bacterial granular sludge (ABGS) process is considered a cost-effective and environmental friendly alternative to conventional technologies for the treatment of wastewater. In addition, this technology is viewed as an attractive alternative for resource recovery due to the presence of the algae consortium [4, 46]. In the last decade, ABGS processes have been intensively studied due to the inherent operational advantages, such as lower energy demands, higher nutrient removal ratio, and potential for resource recovery potential. These operational advantages have positioned the ABGS-based process as a promising technology to improve textile wastewater treatment.

##### **4.1 Environmental advantages**

The ABGS processes are considered as a highly efficient technology for nutrient removal and degradation of both inorganic and organic pollutants in wastewaters, compared to classical biological treatments. The use of microalgal-bacteria consortia for wastewater treatment processes provides advantages at several levels. For instance, through oxygenic photosynthesis, microalgae generate the O<sub>2</sub> required for the aerobic degradation of organic molecules by heterotrophic bacteria, while uptake the CO<sub>2</sub> released during the bacterial aerobic mineralization of the organic substrates. This symbiotic interaction contributes to the prevention of greenhouse gas emissions during the operation of wastewater treatment plants [45].

Moreover, the developed granules are composed of different layers-phases (aerobic, anoxic, and anaerobic) that combined with the microalgae are able to remove toxic pollutants which conventional biological processes usually cannot remove. Other advantages offered by microalgae-bacteria granulation include—(1) the capacity of degrading priority pollutants, (2) enhanced settlement rate, (3) improvement of microalgae separation, and (4) lower operation and maintenance cost than conventional processes [47, 48].

On the other hand, as microalgae are photosynthetic in nature, pollutants are used for producing food and release oxygen into the system, facilitating aerobic pollutant degradation. In toxic environments, microalgae are able to acclimatize themselves to these extreme conditions, with an increasing tolerance to the pollutant toxicity, which allows its removal from water bodies [49]. On the other hand, microalgae-bacteria symbiotic associations contribute toward species microbiological growth by playing an integral part in environmental ecosystems [50].

Several studies have shown the potential of symbiotic algal-bacterial consortia to decolorize dyes and metabolize the aromatic amines typically released during the physicochemical oxidation of dyes. These compounds are considered even more hazardous than their predecessors' dyes [4, 51]. For instance, it was shown that ABGS systems used for the treatment of synthetic textile wastewater may attain efficiencies in dyes decolorization of  $99 \pm 1\%$  and  $96 \pm 3\%$  for dispersing orange-3 and dispersing blue-1, respectively [51].

These efficiencies could be explained by the fact that microalgae present three different mechanisms for decolorization or assimilation of the colored compounds. The chromophores are utilized—(1) for the production of algal biomass, carbon dioxide, and water; (2) for the transformation of the colored compounds to uncolored ones; and (3) for the adsorption of the dye on the algal biomass. It has been reported that *Chlorella* and *Oscillatoria* are able to degrade azo dyes to aromatic amines to simple compounds and subsequently to  $\text{CO}_2$  [1]. Besides, bacteria also have contributed to reaching the highest decolorization and mineralization rates of dyes present in the textile wastewaters. Species belonging to the genera *Pseudomonas*, *Bacillus*, *Aeromonas*, and *Proteus* are some of the most studied bacteria for the degradation of dyes and other toxic effluents [22, 52]. However, despite the high potential of algal-bacterial processes for wastewater treatment, only a few studies have dealt with the bioremediation of textile wastewaters [4].

#### 4.2 Economic benefits

The microalgal-bacterial symbiotic interactions based on the mutualistic exchange of  $\text{O}_2$  and  $\text{CO}_2$  between microalgae and bacteria also allow the system to operate without an additional oxygen source, representing an economic advantage during the treatment of xenobiotic pollutants. Aeration from photosynthetic metabolism is, therefore, especially interesting to reduce operation costs. In addition, many recalcitrant and toxic compounds are much easier to degrade under aerobic conditions than anaerobically [47, 53].

Several economic advantages linked to the resource recovery from microalgal can be highlighted. For instance, since microalgae are naturally found in the environment, at the proper temperature and light intensity microalgae development is possible with minimal costs, allowing the recovery of high-value products with minimum costs [49].

Consequently, ABGS-based processes arise as an alternative technology to AOPs for the treatment of textile wastewater. Tolerances to high pollutant loads, large removal efficiencies, as well as effective operating costs have been highlighted as the main advantages of ABGS technology in the treating of textile wastewater. Pollutant adsorption in the wall of microalgae cells facilitates the mineralization of dyes present in textile wastewater. Moreover, bacterial communities associated with microalgal cultures also can simulate microalgae growth by releasing growth-promoting factors. For instance, bacterial consortia can provide vitamins for improving microalgal growth which may result in lower cost for microalgal biomass production and, therefore in greater production efficiency [54].



An essential factor that must be considered for the development of the stable microbial consortia is related to the compatibility between the selected species/strains. In the particular case of microalgae-bacteria consortia, a balanced exchange of CO<sub>2</sub> and O<sub>2</sub> is essential for attaining optimal performance. Under high CO<sub>2</sub> concentrations, a decreased pH is occurred, causing inhibitory episodes in some microalgal strains. The amount of CO<sub>2</sub> required for microalgae growth varies according to the selected species, and it also depends on the specific configuration of the cultivation system.

On the other hand, the O<sub>2</sub> accumulation produced by photosynthesis must be avoided, since high levels of dissolved oxygen may induce photooxidative damage in microalgae. Therefore, maintaining both CO<sub>2</sub> and O<sub>2</sub> under optimal concentration ranges is essential to guarantee stable removal efficiencies and low-operating costs throughout the process [45].

### 4.3 Resource recovery potential

ABGS-based processes can be considered as potential technology where microalgae can generate valuable resources that can be recovered during the textile wastewater treatment. In the last decade, many efforts have been put to obtain axenic algal monocultures aimed at developing biomass production processes. However, the interactions between microalgae and microorganisms are currently recognized for the potential to improve algal biomass production and to enrich this biomass with valuable chemical and energy compounds with industrial interest such as lipids and carbohydrates [55]. In this respect, the general attributes of bacterial consortia are viewed with a higher interest due to the interactions with microalgae, which may affect algae growth, including motility, chemotaxis, type IV secretion systems, quorum sensing systems, and synthesis of growth promoters [54].

Microalgae are capable of synthesizing several biofuels as lipids and carbohydrates which represent the major energy storage molecules in the microalgae. In contrast, proteins in microalgae are generally not considered as substrates for biofuel production but rather for both food and feed use in human and animal nutrition [55]. In particular, the use of microalgae for the production of biodiesel has focused considerable attention in the past decades, since some species are able to accumulate hydrocarbons up to 30–70% of their dry weight [45].

**Table 3** presents several studies on microalgae and microalgae-bacteria cultures used for the treatment of different wastewater. As it can be noted, the obtained resources depend on the culture systems and type of treated wastewater. These studies showed that microalgae tend to produce a large number of proteins independently from the cultivation method used, as well as other high-value resources with large-energy content such as lipids compounds.

Algal lipids are used to generate biodiesel, a sustainable alternative to fossil fuels. Through photosynthesis, algae can convert CO<sub>2</sub> and water into the organic matter like carbohydrates and lipids. Under ideal conditions, algae can produce carbohydrates, while external stress by limitation of nutrients arises, the algae tend to accumulate lipids. These lipids can be converted into fatty acyl methyl esters through transesterification reactions and can be used as fuel because of their excellent energy density [70].

On the other hand, recent studies have shown that under adequate operating conditions, microalgae and bacteria can form aggregates showing good settleability [71, 72]. The latter facilitates biomass harvesting for its use as feedstock for other energy-producing bioprocesses. Therefore, ABGS systems show a promising future as zero or even negative-energy systems [73]. In addition, the suitable control of the biological interactions between microalgae and bacteria could help to improve

Wastewater	Microbial species	Cultivation method	Obtained products	References
Swine	<i>Chlorella sorokiniana</i>	PBR	Protein-rich microalgal biomass	[56]
Pharmaceutical	<i>Scenedesmus abundans</i>	PBR and PMFC	Production of biodiesel and electricity.	[57]
Textile	<i>Tetradesmus obliquus/proteobacteria</i>	PBR	Protein-rich microalgal biomass	[51]
Poultry	<i>Scenedesmus obliquus</i>	SAR'CENA	Biodiesel and proteins	[58]
Synthetic	<i>Chlorella vulgaris</i>	Batch PBR	—	[59]
Swine		Open PBR	Methane production	[60]
Synthetic	<i>Chlorella pyrenoidosa</i>	Airlift circulation PBR	Biomass and lipid productivity	[61]
Agro-alimentary		Batch PBR	High lipid productivity	[62]
Textile	<i>Chlorella sp./ Pseudomonas sp.</i>	Open Pond	Protein microalgal biomass	[63]
Pharmaceutical	<i>Microcystis aeruginosa</i>	Batch PBR	—	[64]
Synthetic	<i>Chlorophyta sp.</i>	Batch PBR	High aromatic proteins productivity	[65]
Paper Mill	<i>Scenedesmus sp.</i>	Open circular ponds	Protein-rich microalgal biomass and significant presence of a-linolenic acid	[66]
Synthetic	<i>Microalgae consortia</i>	Batch PBR	Biofuel production	[67]
Textile	<i>Green algae/cyanobacteria</i>	Batch PBR	Protein microalgal biomass	[45]
Municipal	<i>Scenedesmus acutus</i>	Batch PBR	High content of lipids, convenient for biodiesel production	[68]
Urban	<i>Galdieria sulphuraria</i>	Enclosed PBR	Higher biomass yield for energy recovery	[69]

*PBR, photobioreactor; PMFC, photosynthetic microbial fuel cell; SAR'CENA, synergistic algal refinery for circular economy using nutrient analogs.*

**Table 3.**  
Microalgae growth and resource recovery potential according to the cultivation method.

microalgae-based biomass and biofuel production in the future. Finally, all these economic, environmental, and resource recovery advantages allow considering the ABGS-based process as a sustainable technology.

## 5. Conclusions

Although many studies have proposed different treatment strategies using chemical, and biological methods for treating textile effluents, no technique is suitable or all-around appropriate for treating the wide range of pollutants present in the textile and clothing industry effluents, causing severe impacts on the environment and human health.

In order to reduce the impact caused on the environment, these substances need to be oxidized by AOPs to become them in biodegradable components, and finally

can be degraded through biological treatments. Although the combination of typical biological techniques with AOPs can provide good efficiencies for the removal of complex high-strength textile wastewaters, the overall process is characterized by high-operating costs. ABGS-based processes have is able to effectively remove nutrients and refractory organic pollutants, with several advantages over any typical individual treatment process. ABGS advantages are attributed to microalgae-bacteria interactions taking place during the process which result in high removal efficiencies under relatively low-operating costs. The treatment efficiencies reached by these processes are also achieved with lower footprint requirements concerning to AOPs. These operating and efficiency advantages allow considering the ABGS technology as an alternative to AOPs concerning the economic and environmental aspects.

ABGS-based processes also promote algal biomass production, and thus, the generation of high-value products. Although ABGS processes have shown a great potential for resource recovery, the number of works evaluating the algal-bacterial symbiosis potential for textile wastewater treatment is still limited. Therefore, more studies on the performance mechanisms, removal efficiency, and cost-effectiveness as an alternative for textile wastewaters treatment are required.

## Author details


Celina Sanchez-Sanchez<sup>1\*</sup>, Guillermo Baquerizo<sup>2</sup> and Ernestina Moreno-Rodríguez<sup>1</sup>

1 Departamento de Ingeniería Civil and Ambiental, Universidad de las Américas Puebla, Puebla, México

2 Instituto de Investigación en Medio Ambiente Xabier Gorostiaga S.J., Universidad Iberoamericana Puebla, Puebla, Mexico

\*Address all correspondence to: [celina.sanchezsz@udlap.mx](mailto:celina.sanchezsz@udlap.mx)

## IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Mullai P, Yogeswari MK, Vishali S, Tejas Namboodiri MM, Gebrewold BD, Rene ER, et al. Aerobic treatment of effluents from textile industry. In: *Current Developments in Biotechnology and Bioengineering: Biological Treatment of Industrial Effluents*. Amsterdam, Netherlands: Elsevier B.V.; 2017. DOI: 10.1016/B978-0-444-63665-2.00001-1
- [2] Rekhate CV, Srivastava JK. Recent advances in ozone-based advanced oxidation processes for treatment of wastewater—A review. *Chemical Engineering Journal Advances*. 2020;**3**:100031. DOI: 10.1016/j.cej.2020.100031
- [3] Huddersman K, Ekpruke A, Asuelimen L. Application of AOPs in the treatment of OSPAR chemicals and a comparative cost analysis. *Critical Reviews in Environmental Science and Technology*. 2019;**49**(4):277-317. DOI: 10.1080/10643389.2018.1537770
- [4] Dhaouefi Z, Toledo-Cervantes A, García D, Bedoui A, Ghedira K, Chekir-Ghedira L, et al. Assessing textile wastewater treatment in an anoxic-aerobic photobioreactor and the potential of the treated water for irrigation. *Algal Research*. 2018;**29**:170-178. DOI: 10.1016/j.algal.2017.11.032
- [5] Sharma S, Kapoor S, Christian RA. Effect of Fenton process on treatment of simulated textile wastewater: Optimization using response surface methodology. *International journal of Environmental Science and Technology*. 2017;**14**(8):1665-1678. DOI: 10.1007/s13762-017-1253-y
- [6] Khan R, Patel V, Khan Z. Bioremediation of dyes from textile and dye manufacturing industry effluent. In: *Abatement of Environmental Pollutants*. Cambridge, MA, United States: Elsevier Inc.; 2020. pp. 107-125. DOI: 10.1016/B978-0-12-818095-2.00005-9
- [7] Lin C, Cao P, Xu X, Ye B. Algal-bacterial symbiosis system treating high-load printing and dyeing wastewater in continuous-flow reactors under natural light. *Water (Switzerland)*. 2019;**11**(3):469. DOI: 10.3390/w11030469
- [8] Lellis B, Fávaro-Polonio CZ, Pamphile JA, Polonio JC. Effects of textile dyes on health and the environment and bioremediation potential of living organisms. *Biotechnology Research and Innovation*. 2019;**3**(2):275-290. DOI: 10.1016/j.biori.2019.09.001
- [9] Resta B, Gaiardelli P, Pinto R, Dotti S. Enhancing environmental management in the textile sector: An organisational-life cycle assessment approach. *Journal of Cleaner Production*. 2016;**135**:620-632. DOI: 10.1016/j.jclepro.2016.06.135
- [10] Buscio V, López-Grimau V, Álvarez MD, Gutiérrez-Bouzán C. Reducing the environmental impact of textile industry by reusing residual salts and water: ECUVal system. *Chemical Engineering Journal*. 2019;**373**:161-170. DOI: 10.1016/j.cej.2019.04.146
- [11] United Nations. UN Helps Fashion Industry Shift to Low Carbon | UNFCCC. 2018. Available from: <https://unfccc.int/news/un-helps-fashion-industry-shift-to-low-carbon>
- [12] Elmogahzy YE. Engineering design for sustainability in the textile and garment industry. In: *Engineering Textiles*. Duxford, United Kingdom: Elsevier Ltd.; 2020. DOI: 10.1016/B978-0-08-102488-1.00006-x
- [13] Jadhav SB, Chougule AS, Shah DP, Pereira CS, Jadhav JP. Application of response surface methodology for the

optimization of textile effluent biodecolorization and its toxicity perspectives using plant toxicity, plasmid nicking assays. *Clean Technologies and Environmental Policy*. 2015;17(3):709-720. DOI: 10.1007/s10098-014-0827-3

[14] Paździor K, Wrębiak J, Klepacz-Smółka A, Gmurek M, Bilińska L, Kos L, et al. Influence of ozonation and biodegradation on toxicity of industrial textile wastewater. *Journal of Environmental Management*. 2017;195:166-173. DOI: 10.1016/j.jenvman.2016.06.055

[15] Mustereț CP, Teodosiu C. Removal of persistent organic pollutants from textile wastewater by membrane processes. *Environmental Engineering and Management Journal*. 2007;6(3):175-187. DOI: 10.30638/eemj.2007.022

[16] Sahu O, Singh N. Significance of bioadsorption process on textile industry wastewater. In: *The Impact and Prospects of Green Chemistry for Textile Technology*. London, United Kingdom: Elsevier Ltd.; 2018. pp. 367-416. DOI: 10.1016/B978-0-08-102491-1.00013-7

[17] Muda K, Aris A, Razman M, Ibrahim Z. Sequential anaerobic-aerobic phase strategy using microbial granular sludge for textile wastewater treatment. In: *Biomass Now—Sustainable Growth and Use*. London, Greater London, United Kingdom: IntechOpen; 2013. DOI: 10.5772/54458

[18] Mahamuni NN, Adewuyi YG. Advanced oxidation processes (AOPs) involving ultrasound for waste water treatment: A review with emphasis on cost estimation. *Ultrasonics Sonochemistry*. 2010;17(6):990-1003. DOI: 10.1016/j.ultsonch.2009.09.005

[19] Saxena G, Kishor R, Zainith S, Bharagava RN. Chapter 17: Environmental contamination, toxicity

profile and bioremediation technologies for treatment and detoxification of textile effluent. In: *Bioremediation for Environmental Sustainability*. Cambridge, MA, United States: Elsevier Inc; 2021. pp. 415-434. DOI: 10.1016/B978-0-12-820524-2.00017-1

[20] Bahadur N, Das P, Bhargava N. Improving energy efficiency and economic feasibility of photocatalytic treatment of synthetic and real textile wastewater using bagasse fly ash modified TiO<sub>2</sub>. *Chemical Engineering Journal Advances*. 2020;2:100012. DOI: 10.1016/j.cej.2020.100012

[21] Ibrahim Z, Amin MFM, Yahya A, Aris A, Muda K. Characteristics of developed granules containing selected decolourising bacteria for the degradation of textile wastewater. *Water Science and Technology*. 2010;61(5):1279-1288. DOI: 10.2166/wst.2010.021

[22] Paz A, Carballo J, Pérez MJ, Domínguez JM. Biological treatment of model dyes and textile wastewaters. *Chemosphere*. 2017;181:168-177. DOI: 10.1016/j.chemosphere.2017.04.046

[23] Solís M, Solís A, Pérez HI, Manjarrez N, Flores M. Microbial decolouration of azo dyes: A review. *Process Biochemistry*. 2012;47(12):1723-1748. DOI: 10.1016/j.procbio.2012.08.014

[24] Singh R, Behera M, Kumar S, Rani A. Current state of knowledge on algae-mediated remediation of endocrine-disrupting chemicals (EDCs) from wastewater. In: *Application of Microalgae in Wastewater Treatment*. Cham, Switzerland: Springer; 2019. pp. 101-120. DOI: 10.1007/978-3-030-13913-1\_6

[25] Starling MCVM, Castro LAS, Marcelino RBP, Leão MMD, Amorim CC. Optimized treatment conditions for textile wastewater reuse using photocatalytic processes under UV and visible light sources. *Environmental Science and Pollution*

Research. 2017;**24**:6222-6232. DOI: 10.1007/s11356-016-6157-8

[26] Chong MN, Sharma AK, Burn S, Saint CP. Feasibility study on the application of advanced oxidation technologies for decentralised wastewater treatment. *Journal of Cleaner Production*. 2012;**35**:230-238. DOI: 10.1016/j.jclepro.2012.06.003

[27] Cai QQ, Wu MY, Li R, Deng SH, Lee BCY, Ong SL, et al. Potential of combined advanced oxidation—Biological process for cost-effective organic matters removal in reverse osmosis concentrate produced from industrial wastewater reclamation: Screening of AOP pre-treatment technologies. *Chemical Engineering Journal*. 2020;**389**:123419. DOI: 10.1016/j.cej.2019.123419

[28] Bustillo-Lecompte CF, Mehrvar M. Treatment of an actual slaughterhouse wastewater by integration of biological and advanced oxidation processes: Modeling, optimization, and cost-effectiveness analysis. *Journal of Environmental Management*. 2016;**182**:651-666. DOI: 10.1016/j.jenvman.2016.07.044

[29] Katheresan V, Kansedo J, Lau SY. Efficiency of various recent wastewater dye removal methods: A review. *Journal of Environmental Chemical Engineering*. 2018;**6**(4):4676-4697. DOI: 10.1016/j.jece.2018.06.060

[30] Paździor K, Bilińska L, Ledakowicz S. A review of the existing and emerging technologies in the combination of AOPs and biological processes in industrial textile wastewater treatment. *Chemical Engineering Journal*. 2019;**376**:120597. DOI: 10.1016/j.cej.2018.12.057

[31] Plumlee MH, Stanford BD, Debroux J-F, Hopkins DC, Snyder SA. Costs of advanced treatment in water reclamation. *Ozone Science and*

*Engineering*. 2014;**36**(5):485-495. DOI: 10.1080/01919512.2014.921565

[32] Bilińska L, Blus K, Foszpańczyk M, Gmurek M, Ledakowicz S. Catalytic ozonation of textile wastewater as a polishing step after industrial scale electrocoagulation. *Journal of Environmental Management*. 2020;**265**:110502. DOI: 10.1016/j.jenvman.2020.110502

[33] Suzuki H, Araki S, Yamamoto H. Evaluation of advanced oxidation processes (AOP) using O<sub>3</sub>, UV, and TiO<sub>2</sub> for the degradation of phenol in water. *Journal of Water Process Engineering*. 2015;**7**:54-60. DOI: 10.1016/j.jwpe.2015.04.011

[34] McGivney W, Kawamura S. Operation and maintenance cost impacts. In: *Cost Estimating Manual for Water Treatment Facilities*. Hoboken, NJ, New Jersey: John Wiley & Sons, Inc.; 2008. pp. 95-103. DOI: 10.1002/9780470260036.ch6

[35] Asaithambi P, Sajjadi B, Abdul Aziz AR, Daud WMABW. Ozone (O<sub>3</sub>) and sono (US) based advanced oxidation processes for the removal of color, COD and determination of electrical energy from landfill leachate. *Separation and Purification Technology*. 2017;**172**:442-449. DOI: 10.1016/j.seppur.2016.08.041

[36] Wang Y, Yang W, Yin X, Liu Y. The role of Mn-doping for catalytic ozonation of phenol using Mn/ $\gamma$ -Al<sub>2</sub>O<sub>3</sub> nanocatalyst: Performance and mechanism. *Journal of Environmental Chemical Engineering*. 2016;**4**(3):3415-3425. DOI: 10.1016/j.jece.2016.07.016

[37] Krichevskaya M, Klauson D, Portjanskaja E, Preis S. The cost evaluation of advanced oxidation processes in laboratory and pilot-scale experiments. *Ozone Science and Engineering*. 2011;**33**(3):211-223. DOI: 10.1080/01919512.2011.554141

- [38] GilPavas E, Dobrosz-Gómez I, Gómez-García MÁ. Coagulation-flocculation sequential with Fenton or Photo-Fenton processes as an alternative for the industrial textile wastewater treatment. *Journal of Environmental Management*. 2017;**191**:189-197. DOI: 10.1016/j.jenvman.2017.01.015
- [39] Ameta RK, Chohadia A, Jain A, Punjabi PB. Fenton and photo-Fenton processes. In: *Advanced Oxidation Processes for Waste Water Treatment*. San Diego, CA, United States: Elsevier; 2018. pp. 49-87. DOI: 10.1016/B978-0-12-810499-6.00003-6
- [40] Ertugay N, Acar FN. Removal of COD and color from Direct Blue 71 azo dye wastewater by Fenton's oxidation: Kinetic study. *Arabian Journal of Chemistry*. 2017;**10**:S1158-S1163. DOI: 10.1016/j.arabjc.2013.02.009
- [41] Sohrabi MR, Khavaran A, Shariati S, Shariati S. Removal of Carmoisine edible dye by Fenton and photo Fenton processes using Taguchi orthogonal array design. *Arabian Journal of Chemistry*. 2017;**10**:S3523-S3531. DOI: 10.1016/j.arabjc.2014.02.019
- [42] Silva LGM, Moreira FC, Cechinel MAP, Mazur LP, de Souza AAU, Souza SMAGU, et al. Integration of Fenton's reaction based processes and cation exchange processes in textile wastewater treatment as a strategy for water reuse. *Journal of Environmental Management*. 2020;**272**:111082. DOI: 10.1016/j.jenvman.2020.111082
- [43] Samsami S, Mohamadi M, Sarrafzadeh MH, Rene ER, Firoozbahr M. Recent advances in the treatment of dye-containing wastewater from textile industries: Overview and perspectives. *Process Safety and Environmental Protection*. 2020;**143**:138-163. DOI: 10.1016/j.psep.2020.05.034
- [44] Rosa JM, Tambourgi EB, Vanalle RM, Carbajal Gamarra FM, Curvelo Santana JC, Araújo MC. Application of continuous H<sub>2</sub>O<sub>2</sub>/UV advanced oxidative process as an option to reduce the consumption of inputs, costs and environmental impacts of textile effluents. *Journal of Cleaner Production*. 2020;**246**:119012. DOI: 10.1016/j.jclepro.2019.119012
- [45] Maza-Márquez P, González-Martínez A, Juárez-Jiménez B, Rodelas B, González-López J. Microalgae-bacteria consortia for the removal of phenolic compounds from industrial wastewaters. In: *Approaches in Bioremediation, Nanotechnology in the Life Sciences*. Cham, Switzerland: Springer; 2018. pp. 135-184. DOI: 10.1007/978-3-030-02369-0\_8
- [46] Zhang B, Lens PNL, Shi W, Zhang R, Zhang Z, Guo Y, et al. Enhancement of aerobic granulation and nutrient removal by an algal-bacterial consortium in a lab-scale photobioreactor. *Chemical Engineering Journal*. 2018;**334**:2373-2382. DOI: 10.1016/j.cej.2017.11.151
- [47] Liu L, Fan H, Liu Y, Liu C, Huang X. Development of algae-bacteria granular consortia in photo-sequencing batch reactor. *Bioresource Technology*. 2017;**232**:64-71. DOI: 10.1016/j.biortech.2017.02.025
- [48] Kumar R, Venugopalan VP. Development of self-sustaining phototrophic granular biomass for bioremediation applications. *Current Science*. 2015;**108**(9):1653-1661. Available from: <http://www.jstor.org/stable/24905531>
- [49] Gupta SK, Bux F. Application of microalgae in wastewater treatment. In: Gupta SK, Bux F, editors. *Application of Microalgae in Wastewater Treatment*. Vol. 1. Gewerbestrasse, Cham, Switzerland: Springer International Publishing; 2019. DOI: 10.1007/978-3-030-13913-1

- [50] Tandon P, Jin Q. Microalgae culture enhancement through key microbial approaches. *Renewable and Sustainable Energy Reviews*. 2017;**80**:1089-1099. DOI: 10.1016/j.rser.2017.05.260
- [51] Dhaouefi Z, Toledo-Cervantes A, Ghedira K, Chekir-Ghedira L, Muñoz R. Decolorization and phytotoxicity reduction in an innovative anaerobic/aerobic photobioreactor treating textile wastewater. *Chemosphere*. 2019;**234**:356-364. DOI: 10.1016/j.chemosphere.2019.06.106
- [52] Zhu Y, Wang W, Ni J, Hu B. Cultivation of granules containing anaerobic decolorization and aerobic degradation cultures for the complete mineralization of azo dyes in wastewater. *Chemosphere*. 2020;**246**:125753. DOI: 10.1016/j.chemosphere.2019.125753
- [53] Muñoz R, Guieysse B. Algal-bacterial processes for the treatment of hazardous contaminants: A review. *Water Research*. 2006;**40**(15):2799-2815. DOI: 10.1016/j.watres.2006.06.011
- [54] Fuentes JL, Garbayo I, Cuaresma M, Montero Z, González-Del-Valle M, Vilchez C. Impact of microalgae-bacteria interactions on the production of algal biomass and associated compounds. *Marine Drugs*. 2016;**14**(5):100. DOI: 10.3390/md14050100
- [55] Yao S, Lyu S, An Y, Lu J, Gjermansen C, Schramm A. Microalgae-bacteria symbiosis in microalgal growth and biofuel production: A review. *Journal of Applied Microbiology*. 2019;**126**(2):359-368. DOI: 10.1111/jam.14095
- [56] Chen CY, Kuo EW, Nagarajan D, Ho SH, Dong CD, Lee DJ, et al. Cultivating *Chlorella sorokiniana* AK-1 with swine wastewater for simultaneous wastewater treatment and algal biomass production. *Bioresource Technology*. 2020;**302**:122814. DOI: 10.1016/j.biortech.2020.122814
- [57] Nayak JK, Ghosh UK. Post treatment of microalgae treated pharmaceutical wastewater in photosynthetic microbial fuel cell (PMFC) and biodiesel production. *Biomass and Bioenergy*. 2019;**131**:105415. DOI: 10.1016/j.biombioe.2019.105415
- [58] Agarwal A, Mhatre A, Pandit R, Lali AM. Synergistic biorefinery of *Scenedesmus obliquus* and *Ulva lactuca* in poultry manure towards sustainable bioproduct generation. *Bioresource Technology*. 2020;**297**:122462. DOI: 10.1016/j.biortech.2019.122462
- [59] de-Bashan L. Removal of ammonium and phosphorus ions from synthetic wastewater by the microalgae *Chlorella vulgaris* coimmobilized in alginate beads with the microalgae growth-promoting bacterium *Azospirillum brasilense*. *Water Research*. 2002;**36**(12):2941-2948. DOI: 10.1016/S0043-1354(01)00522-X
- [60] González-Fernández C, Mahdy A, Ballesteros I, Ballesteros M. Impact of temperature and photoperiod on anaerobic biodegradability of microalgae grown in urban wastewater. *International Biodeterioration and Biodegradation*. 2016;**106**:16-23. DOI: 10.1016/j.ibiod.2015.09.016
- [61] Chu HQ, Tan XB, Zhang YL, Yang LB, Zhao FC, Guo J. Continuous cultivation of *Chlorella pyrenoidosa* using anaerobic digested starch processing wastewater in the outdoors. *Bioresource Technology*. 2015;**185**:40-48. DOI: 10.1016/j.biortech.2015.02.030
- [62] Hongyang S, Yalei Z, Chunmin Z, Xuefei Z, Jinpeng L. Cultivation of *Chlorella pyrenoidosa* in soybean processing wastewater. *Bioresource Technology*. 2011;**102**(21):9884-9890. DOI: 10.1016/j.biortech.2011.08.016



- [63] Tait K, White DA, Kimmance SA, Tarran G, Rooks P, Jones M, et al. Characterisation of bacteria from the cultures of a *Chlorella* strain isolated from textile wastewater and their growth enhancing effects on the axenic cultures of *Chlorella vulgaris* in low nutrient media. *Algal Research*. 2019;**44**:101666. DOI: 10.1016/j.algal.2019.101666
- [64] Liu Y, Zhang J, Gao B, Feng S. Combined effects of two antibiotic contaminants on *Microcystis aeruginosa*. *Journal of Hazardous Materials*. 2014;**279**:148-155. DOI: 10.1016/j.jhazmat.2014.07.002
- [65] Zhou D, Zhang C, Fu L, Xu L, Cui X, Li Q, et al. Responses of the microalga *Chlorophyta* sp. to bacterial quorum sensing molecules (N-acylhomoserine lactones): Aromatic protein-induced self-aggregation. *Environmental Science and Technology*. 2017;**51**(6):3490-3498. DOI: 10.1021/acs.est.7b00355
- [66] Usha MT, Sarat Chandra T, Sarada R, Chauhan VS. Removal of nutrients and organic pollution load from pulp and paper mill effluent by microalgae in outdoor open pond. *Bioresource Technology*. 2016;**214**:856-860. DOI: 10.1016/j.biortech.2016.04.060
- [67] Hu Y, Hao X, van Loosdrecht M, Chen H. Enrichment of highly settleable microalgal consortia in mixed cultures for effluent polishing and low-cost biomass production. *Water Research*. 2017;**125**:11-22. DOI: 10.1016/j.watres.2017.08.034
- [68] Sacristán de Alva M, Luna-Pabello VM, Cadena E, Ortíz E. Green microalga *Scenedesmus acutus* grown on municipal wastewater to couple nutrient removal with lipid accumulation for biodiesel production. *Bioresource Technology*. 2013;**146**:744-748. DOI: 10.1016/j.biortech.2013.07.061
- [69] Henkanatte-Gedera SM, Selvaratnam T, Karbakhshravari M, Myint M, Nirmalakhandan N, van Voorhies W, et al. Removal of dissolved organic carbon and nutrients from urban wastewaters by *Galdieria sulphuraria*: Laboratory to field scale demonstration. *Algal Research*. 2017;**24**:450-456. DOI: 10.1016/j.algal.2016.08.001
- [70] Johnson I, Girijan S, Tripathy BK, Ali MAS, Kumar M. Algal-bacterial symbiosis and its application in wastewater treatment. In: *Emerging Technologies in Environmental Bioremediation*. Cambridge, MA, United States: Elsevier; 2020. pp. 341-372. DOI: 10.1016/B978-0-12-819860-5.00015-8
- [71] Arcila JS, Buitrón G. Microalgae-bacteria aggregates: Effect of the hydraulic retention time on the municipal wastewater treatment, biomass settleability and methane potential. *Journal of Chemical Technology and Biotechnology*. 2016;**91**(11):2862-2870. DOI: 10.1002/jctb.4901
- [72] Medina M, Neis U. Symbiotic algal bacterial wastewater treatment: Effect of food to microorganism ratio and hydraulic retention time on the process performance. *Water Science and Technology*. 2007;**55**(11):165-171. DOI: 10.2166/wst.2007.351
- [73] Vargas A, Escobar Alonso S, Arcila JS, Buitrón G. A dynamic model for microalgae-bacteria aggregates used for wastewater treatment. In: Mannina G, editor. *Frontiers in Wastewater Treatment and Modelling*. Vol. 4. Gewerbestrasse, Cham, Switzerland: Springer International Publishing; 2017. pp. 602-606. DOI: 10.1007/978-3-319-58421-8\_94



# Isolation and Identification of Molds in Selected Dried Fruits and Seeds Sold in Bulk in México

David González-Albarrán, Aurelio López-Malo  
and Enrique Palou

## Abstract

In Mexico, dried fruits and seeds are commonly purchased by consumers in bulk at loosely regulated markets. Lack of oversight in these points of sale translates also to a lack of knowledge of the sanitary conditions of the product being sold and of potential risks to the health of consumers. The objective of this work was to assess the incidence of molds in three different products (peanuts, pecan nuts, and squash seeds) sold in bulk at local traditional produce markets. The isolated molds were analyzed via optical microscopy and colony morphology in selective growth media, and species were identified *via* dichotomic keys. Results of the assessment indicate a high incidence of contamination with toxigenic mold species, such as *Aspergillus flavus*, *Fusarium* spp., and *Penicillium* spp., as well as deteriorative molds such as *Aspergillus niger* and *Rhizopus oryzae*. However, although the incidence of contamination was high, the degree of contamination of most studied samples did not exceed counts permitted by Mexican regulation (300 CFU/g). This would indicate that, in spite of the lack of oversight, storage conditions for the sampled products were, for the most part, adequate, and that the risk to the consumer associated with these kinds of products is marginal.

**Keywords:** peanuts, pecan nuts, squash seeds, *Aspergillus*, *Fusarium*, *Penicillium*

## 1. Introduction

Main causes of deterioration of dried fruits and seeds are fungi (yeasts and molds); thanks to their low water activity preventing bacterial growth for the most part. The edible part of the seeds is also usually protected from spoilage by thick, durable shells [1]. However, dried fruits are often sold to consumers with the shell removed, which exposes the edible portion to microbial contamination and spoilage. Some species of molds are capable of producing toxic substances, such as aflatoxins, ochratoxin, and patulin.

In Mexico, purchasing fresh produce loose or in bulk at traditional markets is a common practice, even by small (i.e., non-corporate) consumers, such as families and individuals. These points of sale tend to be loosely regulated, which translates also to a lack of knowledge of the sanitary conditions of the product being sold and of potential risks to the health of consumers. More knowledge about the mycobiota

of these food products can help make better, more informed policy decisions, as well as help in the development of solutions to potential problems caused by fungi.

Peanuts are an important food crop all around the world, valued for their edible seed as well as a source of edible oils. As a crop that grows underground, peanuts are highly susceptible to microbial contamination. Mycobiota of peanuts mostly consists of *Aspergillus* and *Penicillium* species, including *A. flavus* and *A. parasiticus*, both known to produce highly toxic aflatoxins [2, 3]. Due to their low water activity, spoilage of peanuts is often caused by molds; however, it is only likely to happen when storage conditions are inadequate and water activity of the kernels is allowed to rise above the threshold for mold growth.

Tree nuts are similar to peanuts in composition and spoilage characteristics; they have a high lipid content and low water activity. Contamination of tree nuts with molds occurs mostly when they are handled improperly after being dehulled since their thick shells are usually able to protect them from microbial contamination. After being dehulled, tree nuts are susceptible to spoilage by molds when their water activity is allowed to increase above the threshold at which molds are capable of growth. Tree nuts are important food crops all around the world, and their nutritional and flavor properties, as well as their relatively high market price, make them of economic importance to the countries in which they are produced.

The seeds from different species of squash or gourds are consumed either on their own or as an ingredient on several dishes all around the world. The species *Cucurbita pepo* (winter squash) and *Cucurbita argyrosperma* (silver-seed gourd) are of particular importance for their production of edible seeds [4]. Amongst edible seeds, squash seeds are unique in that they have a relatively high moisture content, which makes them more susceptible to microbial spoilage [5]. Despite this, little information exists concerning the mycobiota and spoilage of edible squash seeds.

In recent years, the identification of molds is more commonly carried out via molecular identification techniques. However, morphological analysis remains an accessible and effective, if labor-intensive, technique for the identification of fungal isolates. Although often dismissed by some specialists as difficult and inconsistent, it has certain advantages over molecular techniques. In particular, many mold species produce polymerase-inhibiting substances, which hinder identification via techniques based on the polymerase chain reaction or PCR [6]. Lack of knowledge about microbiota of foods is a common problem in developing nations, more accessible methods for isolation and identification of microorganisms can help researchers in these areas to investigate such matters more effectively. For these purposes, further study and development of morphological techniques for mold identification are not only useful but also essential.

## 2. Methodology

Peanut (*Arachis hypogaea*), pecan (*Carya illinoensis*), walnut (*Juglans regia*), and squash seed (*C. argyrosperma*) samples were purchased in municipal produce markets of cities and towns across different regions of Mexico, where dried fruits can be purchased in bulk. Samples consisted of at least 250 g of the selected dried fruit dehulled, unsalted, and unroasted. A total of 30 samples were obtained and transferred to the Laboratory of Food Microbiology of the Universidad de las Americas Puebla for further analysis. None of the samples showed visible signs of fungal growth/deterioration.

Total yeast and mold content were determined on dried fruit samples. 10 g of each dried fruit were weighed and placed in sterile sampling bags. Samples were diluted with 90 mL of 0.1% peptone (Becton-Dickinson, Mexico) solution, and homogenized using a Stomacher® laboratory blender (Seward, United Kingdom.).

Serial dilutions of the sample were prepared using the same 0.1% peptone solution and plated in potato dextrose agar (PDA, Becton-Dickinson, Mexico) plates acidified with 1.6 mL/100 mL of a 10% tartaric acid solution to a final pH of 3.5. Colonies in inoculated plates were counted after being incubated at 25°C over a period of 5 days. For samples in which *Rhizopus* spp. colonies were readily apparent, Dicloran Chloramphenicol Rose Bengal agar (DRBC, Becton-Dickinson, USA) was used instead of PDA to limit colony growth and allow for more accurate quantification of yeasts and molds. Assays were carried out by triplicate; results of these trials showed generally low counts and are not presented in this article.

Individual mold colonies were taken from counting plates and inoculated in PDA plates, without tartaric acid, for isolation. Plates were incubated at 25°C for 7 days, or 3 days when a *Rhizopus* spp. the colony was apparent and inspected afterward for contamination. Resampling and inoculation of the mold colonies were repeatedly carried out until only a single mold species was apparent in the agar plate. Afterwards, molds were point-inoculated in malt extract agar (MEA, Becton-Dickinson, Mexico), and Czapek Yeast Extract agar (CYA, Becton-Dickinson, Mexico) or Czapek Dox agar (CZD, Becton-Dickinson, Mexico). For identification, molds were incubated at 25°C for 7 days, or 3 days when a *Rhizopus* spp. the colony was readily apparent. Colony diameter, front and back colony color, colony texture and shape, presence and color of exudates, as well as presence and color of soluble pigments in the different media were the main criteria used for mold identification.

After incubation of the fungal isolates, colony macro and microscopic morphology was observed and identification was carried out following dichotomous keys as outlined by Pitt and Hocking [1], and Samson *et al.* [7]. For microscopic observation and measurement, a small (2 mm<sup>2</sup>, approx.) and shallow (<1 mm deep) sample were cut from the outer edges of a colony using a sterile dissection needle, including a small portion of the culture media along with the mold sample. The colony sample was then placed on a glass microscope slide and a drop of aniline blue solution (0.1% aniline blue, Química Meyer, Mexico) in 85% lactic acid (Química Meyer, Mexico) was added, as mounting fluid. Afterwards, a drop of 70% ethanol was added to help disperse conidia and aid in the visualization of fruiting structures. A slide cover was then placed on top of the sample, and the sample was carefully heated with a Meker-Fisher burner to melt the culture medium. The sample was then observed under 100-x magnification in an optical microscope (American Optical, USA) equipped with an Axiocam ERc 5S camera (Zeiss Microscopy, Germany) and associated software for microphotography and measurements. Shape, size, texture, and color of conidia and fruiting structures were the main characteristics used for discrimination. Data is only presented for media in which the sampling for microscopic observation yielded measurable, observable structures and images of satisfactory quality.

For identification of *Penicillium* subgenus *Penicillium* species, an additional culture was prepared using creatine sucrose neutral agar (CSN), as described by Pitt and Hocking [1]. Briefly, 10 g of sucrose, 5 g of creatine, 1 g of monopotassium phosphate, 0.05 g of bromocresol purple, 10 mL of a solution containing potassium chloride, magnesium sulfate heptahydrate, ferrous sulfate, zinc sulfate, and copper (II) pentahydrate in trace amounts, and 15 g of bacteriological agar were dissolved in 1000 mL distilled water and heated to a boil before being autoclaved at 121°C for 15 min. The medium adds an additional criterion for discrimination of the otherwise similar species of *Penicillium* subgenus *Penicillium* by producing an acid (yellow), alkaline (violet), or neutral (gray) reaction based on the capacity and extent to which the species metabolizes sucrose and/or creatine. CSN reaction, along with culture morphology, is part of the criteria used in the dichotomous keys presented by Pitt and Hocking [1].

### 3. Results

Results of the assessment showed the presence of some toxigenic mold species, such as *Aspergillus flavus*, *Fusarium* spp., and *Penicillium* spp., as well as deteriorative molds such as *Aspergillus niger* and *Rhizopus oryzae*, as well as innocuous mold species such as *Phoma* spp. and *Cladosporium* spp. In general, mycobiota of the sampled seeds is consistent with reports for dried fruits from a tropical region, with a predominance of molds adapted to warm, humid climates such as *R. oryzae*, *A. flavus*, and *A. niger* [7].

A comprehensive list of the identified molds can be seen in **Tables 1–3**, along with the main characteristics that allowed their identification. A greater diversity of mold species was found in peanut samples, whereas squash seed mycobiota was largely dominated by *R. oryzae* and *A. niger*. Mold species identified in peanuts have been previously reported by other authors: *A. flavus*, *A. niger*, *Fusarium* spp., *Alternaria alternata*, *Rhizopus* spp., *Cladosporium* spp., and *Mucor* spp. which are ubiquitous in shelled peanut samples, both raw and roasted [2, 3, 8–10].

Valle-Garcia *et al.* [11] reported *A. niger*, *Fusarium* spp., *Rhizopus* spp., and a variety of *Penicillium* species in pecans from Rio Grande do Sul State in Brazil. The particular species of *Penicillium* found in the aforementioned study were different from those encountered in this study; however, *Penicillium* is an extremely diverse genus, and species are bound to be different in separate geographic areas.


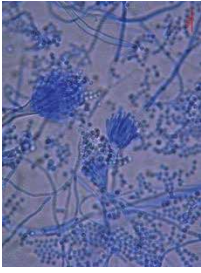
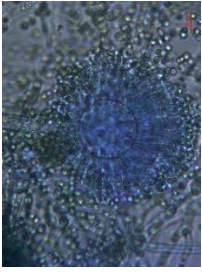

The mycobiota of the seeds of *C. argyrosperma* remains poorly studied. Phytopathogenic molds such as *Phytophthora capsici*, *Rhizoctonia solani*, and *Sclerotium rolfsii* have been reported to cause fruit rot in silver-seed gourd [12]; however, no such studies exist about the mycology of the seeds themselves, which is, presumably, different from that of the rest of the fruit. It can be assumed that the mycobiota of *C. argyrosperma* seeds is similar to that of the seeds of *C. pepo*. In that regard, *R. oryzae* and *A. niger* have also been reported in the seeds of *C. pepo* [13]. Furthermore, *Penicillium* species have also been reported in *C. pepo* seeds [14], although data concerning squash seeds in that study is deficient.

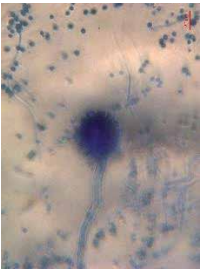
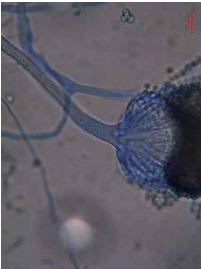
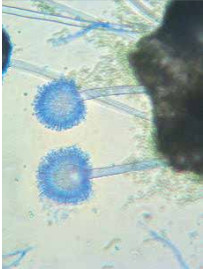
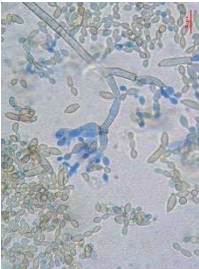
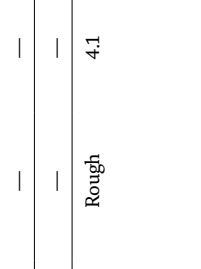
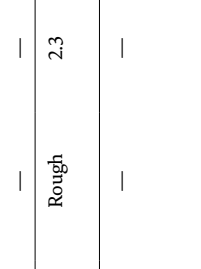
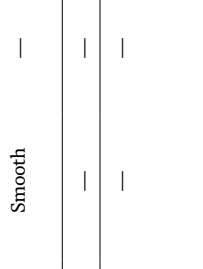
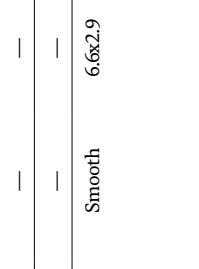
Morphological characteristics of *A. alternata* are in accordance with those described by Armitage *et al.* [15], except for colony diameter in PDA culture, which was reported as a maximum of 68 mm by these authors. Other authors [16] have also reported the *Aspergillus candidus* morphology encountered in this study. Its distinctive features are the characteristic white color of its colonies and the smooth texture of its conidia. Conidium diameter was smaller than reported in other studies, at 2.0  $\mu\text{m}$ .

*A. flavus* was differentiated from similar species such as *A. oryzae* and *A. parasiticus* by the texture of its stipe (rough) and its conidia (slightly roughened), as reported by Diba *et al.* [17] and Samson *et al.* [7]. *A. niger* was identified by its characteristic black, extensive colonies, as well as by its small conidia (<6  $\mu\text{m}$  in diameter), borne in conidiophores bearing both phialides and metulae, as outlined by Pitt & Hocking [1]. *A. tamarii* was identified by its characteristic brown colonies across different culture media, its large (>3  $\mu\text{m}$  diameter), rough-walled conidia, and its radiate conidiophores [1, 18].


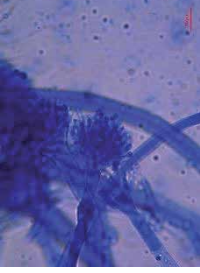
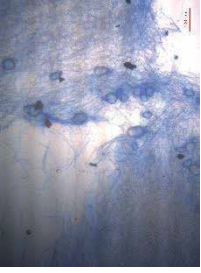
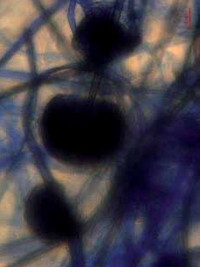
Two unidentified *Aspergillus* species were encountered in peanut samples. Morphological characteristics were insufficient for identification with the means available, and further testing would be necessary.

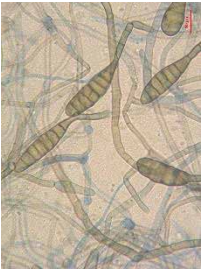
For the differentiation of some *Penicillium* species, creatine sucrose neutral agar proved useful, as macroscopic colony characteristics and microscopic morphology of *Penicillium* species are often too similar to allow for speciation. A wide variety of *Penicillium* species were recovered from peanut kernels and pecan nuts, which is in accordance with other studies of the microbiota of both dried fruits [11, 19, 20]; however, the particular blend of species identified in this study was different than

Mold species	Growth media*	Macroscopic characteristics				Microscopic characteristics						Optical micrographs
		Colony diameter (mm)	Colony color			Stipe			Conidia			
			Front	Back	Color	Texture	Shape	Texture	Size (µm)			
<i>Alternaria alternata</i>	PDA	80.0	Black/brown	Brown	Hyaline	Smooth	Clavate	Rough	19.5x8.9			
	MEA	60.0	Black/brown	Brown	—	—	—	—	—			
	CZD	58.0	Black/brown	Brown	—	—	—	—	—			
<i>Aspergillus candidus</i>	PDA	30.0	White	Brown	Hyaline	Smooth	Spherical	Smooth	2.0			
	MEA	25.0	White	Brown	—	—	—	—	—			
	CZD	33.0	White	Brown	—	—	—	—	—			
<i>Aspergillus flavus</i>	PDA	60.8	Green	White	Hyaline	Rough	Spherical	Slightly roughened	2.6			
	MEA	60.0	Green	White	—	—	—	—	—			
	CZD	45.0	White	White	—	—	—	—	—			
<i>Aspergillus niger</i>	PDA	85.0	Black	White	Hyaline	Slightly roughened	Spherical	Slightly roughened	2.8			
	MEA	36.2	Black	White	—	—	—	—	—			
	CZD	48.7	White	White	—	—	—	—	—			

Mold species	Growth media*	Macroscopic characteristics				Microscopic characteristics				Optical micrographs
		Colony diameter (mm)	Colony color		Stipe	Conidia				
			Front	Back		Color	Texture	Shape	Texture	
<i>Aspergillus tamarii</i>	PDA	71.2	Brown	White	—	—	—	—	—	
	MEA	59.0	Brown	Cream	—	—	—	—	—	
	CYA	47.5	Golden brown	Cream	Hyaline	Smooth	Spherical	Rough	4.1	
<i>Aspergillus</i> spp.	PDA	70.0	Olive brown	White	—	—	—	—	—	
	MEA	78.0	White	White	Hyaline	Finely roughened	Spherical	Rough	2.3	
	CZD	80.0	Brown	Yellow	—	—	—	—	—	
<i>Aspergillus</i> spp.	PDA	55.0	Olive brown	White	Hyaline	Finely roughened	Spherical	Smooth	—	
	MEA	80.0	White	White	—	—	—	—	—	
	CZD	65.0	Brown	White	—	—	—	—	—	
<i>Cladosporium cladosporioides</i>	PDA	22.0	Dark green	Black	—	—	—	—	—	
	MEA	24.0	Dark green	Black	—	—	—	—	—	
	CZD	22.0	Dark green	Black	Hyaline	Smooth	Ellipsoidal	Smooth	6.6x2.9	

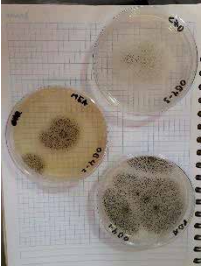

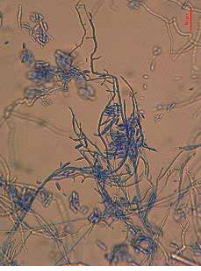



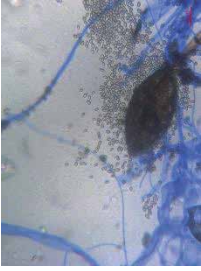
Mold species	Growth media*	Macroscopic characteristics			Microscopic characteristics					Optical micrographs
		Colony diameter (mm)	Colony color		Stipe	Conidia				
			Front	Back		Color	Texture	Shape	Texture	
<i>Epicoccum nigrum</i>	PDA	76.0	Black	Black	Brown	Smooth	Spherical	Rough	8.1	
	MEA	29.0	White	White	—	—	—	—	—	
	CYA	62.0	Gray	White	—	—	—	—	—	
<i>Neosartorya fisheri</i>	PDA	59.1	White	White	—	—	—	—	—	
	MEA	52.6	White	White	—	—	—	—	—	
	CYA	74.8	White	Yellow	Hyaline	Smooth	Spherical	Spheroidal	2.2	
<i>Phoma spp.</i>	PDA	13.9	Golden brown	Brown	—	—	—	—	—	
	MEA	12.7	Golden brown	Brown	—	—	—	—	—	
	CYA	16.5	Golden brown	Yellow	Hyaline	Smooth	Cylindrical	Smooth	1.9	
<i>Rhizopus oryzae</i>	PDA	87.4	White	Gray	Brown	Smooth	N/P	N/P	N/P	
	MEA	87.4	White	Gray	—	—	—	—	—	
	CYA	87.4	White	Gray	—	—	—	—	—	

Mold species	Growth media*	Macroscopic characteristics			Microscopic characteristics					Optical micrographs	
		Colony diameter (mm)	Colony color		Stipe	Conidia			Size (µm)		
			Front	Back		Color	Texture	Shape			Texture
<i>Ulocladium</i> spp.	PDA	53.0	Black	Brown	Brown	Brown	Smooth	Apiculate	Smooth	29.0x7.9	
	MEA	40.0	Black	Brown	—	—	—	—	—	—	
	CZD	45.0	Black/ brown	Brown	—	—	—	—	—	—	

\*PDA: potato dextrose agar; MEA: malt extract agar; CZD: Czapek Dox agar; CYA: Czapek yeast extract agar.

**Table 1.** Identification of isolated non-Penicillium molds in peanut samples.

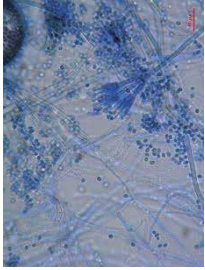
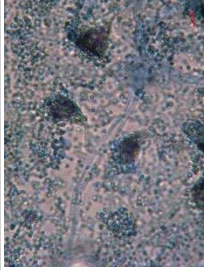
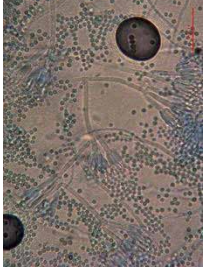

Mold species	Growth media*	Macroscopic characteristics			Microscopic characteristics					Optical micrographs / Photos
		Colony diameter (mm)			Stipe					
		Front	Back	Colony color	Color	Texture	Shape	Texture	Size (µm)	
<i>Aspergillus niger</i>	PDA	85.0	Black	Yellow	—	—	—	—	—	
	MEA	32.8	Black	Yellow	—	—	—	—	—	
	CZD	49.2	Gray	Gray	—	—	—	—	—	
<i>Euotium</i> spp.	PDA	57.3	White	White	—	—	—	—	—	
	MEA	52.6	White	White	—	—	—	—	—	
	CZD	42.0	White	White	—	—	—	—	—	
<i>Fusarium</i> spp.	PDA	37.2	Violet	Violet	—	—	—	—	—	
	MEA	30.7	Brown	Brown	—	—	—	—	—	
	CYA	60.8	Light pink	Yellow	Hyaline	Rough	Fusiform	Smooth	28.6 x 3.6	
<i>Mucor hiemalis</i>	PDA	87.4	Gray	Gray	Hyaline	Smooth	Reniform	Smooth	5.0	
	MEA	87.4	Gray	Gray	—	—	—	—	—	
	CZD	87.4	Gray	Gray	—	—	—	—	—	


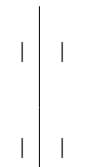
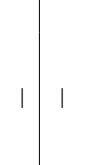
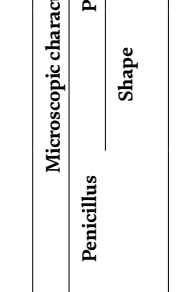
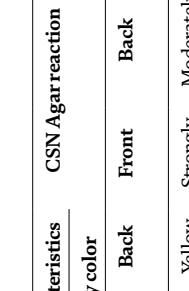
Mold species	Growth media*	Macroscopic characteristics			Microscopic characteristics					Optical micrographs / Photos
		Colony diameter (mm)	Colony color		Stipe	Conidia				
			Front	Back		Color	Texture	Shape	Texture	
<i>Rhizopus oryzae</i>	PDA	87.4	White	Gray	Brown	Smooth	Fusiform	Striate	5.2 x 2.9	
	MEA	87.4	White	Gray	—	—	—	—	—	
	CZD	79.7	White	Gray	—	—	—	—	—	

\*PDA: potato dextrose agar; MEA: malt extract agar; CZD: Czapek Dox agar; CYA: Czapek yeast extract agar.

**Table 2.**  
Identification of isolated non-Penicillium molds in pecan samples.

Mold species	Dried fruit	Growth media*	Macroscopic characteristics			CSN Agar reaction			Microscopic characteristics			Optical micrographs
			Colony diameter (mm)	Colony color		Front	Back	Penicillus	Phialide			
				Front	Back				Shape	Length (µm)	Width (µm)	
<i>Paecilomyces variotii</i>	Walnut	PDA	50.0	Brown	Cream	N/A	N/A	—	—	—		
		MEA	62.7	Brown	Cream			Biverticillate	Ampulliform with long collula	11.2		2.7
		CZD	53.2	Brown	Cream			—	—	—		—
<i>Penicillium citrinum</i>	Peanut	PDA	15.0	Pale yellow/green	Yellow	N/A	N/A	Biverticillate	Ampulliform	4.4	1.2	
		MEA	15.0	Green	Brown			—	—	—	—	
		CZD	10.0	Pale brown/white	Pale brown			—	—	—	—	
<i>Penicillium crustosum</i>	Peanut	PDA	36.6	Olive green	Pale yellow	Strongly acidic	Moderately alkaline	Terverticillate	Ampulliform	9.0	2.1	
		MEA	29.4	Green	Yellow			—	—	—	—	
		CZD	31.8	White	White			—	—	—	—	
<i>Penicillium expansum</i>	Peanut	PDA	30.6	Gray/green	Cream	Strongly acidic	Strongly acidic, brown	Terverticillate	Cylindrical	9.4	2.1	
		MEA	27.1	Gray/green	Cream			—	—	—	—	
		CYA	37.2	White/green	Cream			—	—	—	—	

Mold species	Dried fruit	Growth media*	Macroscopic characteristics			CSN Agar reaction			Microscopic characteristics				Optical micrographs
			Colony diameter (mm)	Colony color		Front	Back	Penicillus	Phialide				
				Front	Back				Shape	Length (µm)	Width (µm)		
<i>Penicillium funiculosum</i>	Pecan	PDA	34.6	Green	White	N/A	N/A	Terverticillate	Acerose	10.0	2.0		
		MEA	38.2	Pale green	White								
		CZD	30.0	White/green	White								
<i>Penicillium glabrum</i>	Peanut	PDA	45.0	Gray/green	White	N/A	N/A	Biverticillate	Ampulliform	8.6	3.6		
		MEA	40.0	Green	Brown								
		CZD	35.0	Green/white	Yellow								
<i>Penicillium griseofulvum</i>	Peanut	PDA	25.1	Gray/green	Cream	Neutral	Neutral	Biverticillate	Ampulliform	6.9	1.5		
		MEA	10.0	Green	White								
		CZD	20.0	Gray	White								
<i>Penicillium nidigovense</i>	Peanut	PDA	23.8	White	Brown	Neutral	Neutral	Terverticillate	Ampulliform	9.5	2.3		
		MEA	22.5	Gray/green	Brown								
		CZD	23.7	White	Gray								

Mold species	Dried fruit	Growth media*	Macroscopic characteristics			CSN Agar reaction		Microscopic characteristics			Optical micrographs			
			Colony diameter (mm)	Colony color		Front	Back	Penicillus	Shape	Length (µm)	Width (µm)			
				Front	Back									
<i>Penicillium naligovense</i>	Pecan	PDA	23.8	White	Yellow	Strongly acidic	Moderately alkaline	—	—	—	—			
		MEA	20.0	Gray/green	Yellow			—	—	—	—			
		CZD	30.6	White	Yellow			—	—	—	—	—		
<i>Penicillium rugulosum</i>	Walnut	PDA	14.4	Green	Brown	N/A	N/A	Tertverticillate	Acerose	7.6	1.2			
		MEA	12.1	Green	Brown			—	—	—	—			
		CZD	15.2	Brown	Cream			—	—	—	—	—		
<i>Penicillium</i> spp.	Peanut	PDA	20.0	Gray/Green	Yellow	N/A	N/A	Biverticillate	Ampulliform	6.6	1.6			
		MEA	10.0	Green	White			—	—	—	—			
		CZD	25.0	White/Brown	White			—	—	—	—	—		

\*PDA: potato dextrose agar, MEA: malt extract agar, CZD: Czapek Dox agar, CYA: Czapek yeast extract aga.

**Table 3.** Identification of isolated *Penicillium* and related species in dried fruit samples.

those reported elsewhere. *P. citrinum* has been reported in peanut kernels [19] as well as *Penicillium crustosum* [14] and *P. expansum* [21], whereas *P. variotii* is a mold associated with oily foods [1] that has previously been isolated from walnuts as well as from peanuts [22].

It was possible to identify isolates of *P. variotii*, *P. citrinum*, *P. funiculosum*, *P. glabrum*, and *P. rugulosum* without the need for a CSN culture thanks to the characteristics of their penicilli, colony colors in the front and back, colony diameter, and phialide dimensions. It is also known that CSN often lacks discriminatory power against species outside of the subgenus *Penicillium* [23].

*P. variotii* presented the characteristic phialides of the species: flask-shaped (ampulliform) with an elongated neck (collula) that curves away from the center, as well as the ellipsoid conidia that distinguish it from other members of the *Paecilomyces* genus [1]. These same characteristics were reported by Houbraken *et al.* [24].

*P. citrinum* presented white mycelium with dull green conidia, 15 mm diameter colonies, brown colony backs, and ampulliform phialides borne in biverticillate penicilli, as outlined by Pitt & Hocking [1], whose criterion include colonies 12–18 mm in diameter in MEA, and ampulliform phialides 7–12  $\mu\text{m}$  in length, and as reported by Houbraken *et al.* [23], who did not report phialide length, but reported conidia colored blue/gray/green with a brownish-yellow colony back in MEA culture. Differences in phialide length between studies could be attributed to the media of choice.

*P. funiculosum* presented white mycelium with dull green conidia, pale green colony backs, 38.2 mm colony diameter, and acerose phialides borne in terverticillate penicilli, measuring 10.0  $\mu\text{m}$  in length, as outlined by Pitt & Hocking [1], whose criterion include colonies between 25 and 45 mm diameter in MEA, dull green conidia with white to reddish mycelium and acerose phialides between 9 and 11  $\mu\text{m}$  in length. Bandh *et al.* [6] reported colonies colored green on the front and brown in the back, measuring 39 mm in diameter in MEA, as well as cylindrical phialides borne in terverticillate penicilli.

*P. glabrum* presented production of green conidia and brown colony reverse, 40 mm diameter colonies, ampulliform phialides 8.6  $\mu\text{m}$  in length borne in biverticillate penicilli, and smooth conidia. Pitt & Hocking [1] described the species as presenting green conidia with white mycelium and green, yellow, or brown colony backs, as well as ampulliform phialides 8–12  $\mu\text{m}$  in length borne in mono or biverticillate penicilli. Barreto *et al.* [25] reported similar macroscopic colony morphology, as well as the species' main distinctive feature: smooth to slightly roughened conidia.

Lastly, *P. rugulosum* presented green colonies (due to conidia production) with brown colony backs, colony diameters of 12.1, and acerose phialides 7.6  $\mu\text{m}$  in length borne from terverticillate penicilli. These characteristics are mostly in line with those reported by Pitt & Hocking [1], who reported green/yellow colony fronts with brown backs and 10–20 mm diameter colonies. These authors also report phialides between 8 and 11  $\mu\text{m}$ , which are slightly longer than those encountered in this work. Similarly, Sidrim *et al.* [26] reported *P. rugulosum* as green colonies of approximately 12 mm in diameter, presenting biverticillate penicilli and ellipsoidal conidia.

#### 4. Conclusions

The mycobiota of the studied products is similar to that reported by other studies, mostly consistent with mold species known to appear in peanuts, pecans, walnuts, and squash seeds. Morphological techniques were used to successfully



identify both genus and species in a variety of molds, proving the value of traditional techniques even as PCR-based speciation becomes the norm. *Penicillium* subgenus *Penicillium* species, in particular, tend to be difficult to speciate thanks to their very similar morphology; however, it was possible to differentiate them with the use of the creatine-sucrose neutral medium. Of the identified species, several are known to produce toxins, while others are known as spoilage microorganisms of dried fruits and grains. Samples were obtained shelled and non-roasted; therefore, contamination is likely to have occurred during transportation, handling, and exhibition at the point of sale of the samples. This would indicate that more measures that are stringent are required for the sale of loose produce as to prevent potential health hazards to consumers. Further, more studies regarding the mycobiota of dried fruits, and in particular the mycobiota of squash seeds, are necessary to better understand the potential for disease vectoring and spoilage of this commodity.

## Acknowledgements


Author González-Albarrán gratefully acknowledges Universidad de las Américas Puebla (UDLAP) and the National Council for Science and Technology (CONACyT) of Mexico financial support for his PhD studies. This work was supported by CONACyT (grant number CB-2016-2101-283636) and UDLAP (grant number 3555).

## Author details

David González-Albarrán, Aurelio López-Malo and Enrique Palou\*  
Department of Chemical, Food, and Environmental Engineering, Universidad de las Américas Puebla Cholula, Puebla, Mexico

\*Address all correspondence to: [enrique.palou@udlap.mx](mailto:enrique.palou@udlap.mx)

## IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Pitt JI, Hocking AD. Fungi and Food Spoilage. 3rd ed. London: Springer; 2009. DOI: 10.1007/978-0-387-92207-2
- [2] Khandaker M, Ahmed R, Rahim M, Hassan T, Begum M. Determination of mycoflora and mycotoxins in raw and roasted peanuts in Bangladesh. International Journal of Plant and Environment. 2019a;5(4):259-264
- [3] Khandaker MM, Rahman T, Rahim MM, Hassan MT, Begum M, Bhuiyan MAB. Mycoflora and mycotoxin detection in roasted peanut kernels of Bangladesh. Bangladesh Phytopathological Society. 2019b;35(1&2):53-58
- [4] Sánchez-de la Vega G, Castellanos-Morales G, Gámez N, Hernández-Rosales HS, Vázquez-Lobo A, Aguirre-Plante E, et al. "Genetic resources in the 'Calabaza Pipiana' Squash (*Cucurbita agryrosperma*) in Mexico: Genetic diversity, genetic differentiation and distribution models". Frontiers in Plant Science. 2018;9:400. DOI: 10.3389/fpls.2018.00400
- [5] Chayjan RA, Salari K, Abedi Q, Sabziparvar AA. Modeling moisture diffusivity, activation energy and specific energy consumption of squash seeds in a semi fluidized and fluidized bed drying. Journal of Food Science and Technology. 2013;50:667-677
- [6] Bandh SA, Kamili AN, Ganai BA. Identification of some *Penicillium* species by traditional approach of morphological observation and culture. African Journal of Microbiology Research. 2011;5(21):3493-3496
- [7] Samson RA, Hoekstra ES, Frisvad JC, Filtenberg O. Introduction to Food-Borne Fungi. 4th ed. Baarn and Delft, The Netherlands: Centraalbureau voor schimmelcultures; 1995
- [8] Adetunji MC, Ezeokoli OT, Ngoma L, Mwanza M. Phylogenetic diversity and prevalence of mycoflora in ready-to-eat supermarket and roadside-vended peanuts. Mycologia. 2020;113(1):1-11
- [9] Adjou ES, Dahouenon-Ahoussi E, Soumanou MM. Investigations on the mycoflora and processing effects on the nutritional quality of peanut (*Arachis hypogea* L. var. TS 32-1). Journal of Microbiology, Biotechnology and Food Sciences. 2012;2(3):1025-1039
- [10] Gürses M. Mycoflora and aflatoxin content of hazelnuts, walnuts, peanuts, almonds, and roasted chickpeas (leblebi) sold in Turkey. International Journal of Food Properties. 2006;9:395-399
- [11] Valle Garcia M, Machado Moraes V, Olivier Bernardi A, Schneider Oliveira M, Mallmann CA, Boscardin J, et al. Mycological quality of pecan nuts from Brazil: absence of aflatoxigenic fungi and aflatoxins. Ciência Rural. 2019;49(96):1-8
- [12] Díaz-Nájera JF, Alvarado-Gómez OG, Leyva-Mir SG, Ayvar-Serna S, Michel-Aceves AC, Vargas-Hernández M. Identification and control of fungi causing fruits rot in pipiana pumpkin (*Cucurbita argyrosperma* Huber). African Journal of Agricultural Research. 2015;10(11):1150-1157
- [13] Rahim S, Dawar S, Tariq M, Zaki MJ. Microflora associated with the seed samples of *Cucurbita pepo* L. collected from Pakistan. Pakistan Journal of Botany. 2013;45:2173-2179
- [14] Njobeh PB, Dutton MF, Koch SH, Chuturgoon A, Stoev S, Seifert K. Contamination with storage fungi of human food from Cameroon. International Journal of Food Microbiology. 2009;135(3):193-198

- [15] Armitage AD, Barbara DJ, Harrison RJ, Lane CR, Sreenivasaprasad S, Woodhall JW, et al. Discrete lineages within *Alternaria alternata* species group: Identification using new highly variable loci and support from morphological characters. *Fungal Biology*. 2015;**119**(11):994-1006
- [16] Varga J, Frisvad JC, Samson RA. Polyphasic taxonomy of *Aspergillus* section *Candidi* based on molecular, morphological and physiological data. *Studies in Mycology*. 2007;**59**:75-88
- [17] Diba K, Kordbacheh P, Mirhendi SH, Rezaie S, Mahmoudi M. Identification of *Aspergillus* species using morphological characteristics. *Pakistan Journal of Medical Sciences*. 2007;**23**(6):867-872
- [18] Balajee SA. *Aspergillus terreus* complex. *Medical Mycology*. 2009;**47**:S42-S46
- [19] Chein SH, Sadiq MB, Datta A, Anal AK. Prevalence and identification of *Aspergillus* and *Penicillium* species isolated from peanut kernels in central Myanmar. *Journal of Food Safety*. 2019;**39**:e12686. DOI: 10.1111/jfs.12686
- [20] Gonçalves E, Nogueira JHC, Fonseca H, Felicio JD, Pino FA, Corrêa B. Mycobiota and mycotoxins in Brazilian peanut kernels from sowing to harvest. *International Journal of Food Microbiology*. 2008;**123**(3):184-190
- [21] Aslam MF, Irshad G, Khan HM, Ghuffar S, Azam F. Identification of seed-borne mycoflora associated with peanut (*Arachis hypogaea* L.) in Pothwar, Pakistan. *Plant Protection*. 2017;**1**, 2:91-95
- [22] Rahouti M, Seigle-Murandi F, Steiman R, Eriksson K-E. Metabolism of ferulic acid by *Paecilomyces variotii* and *Pestalotia palmarum*. *Applied and Environmental Microbiology*. 1989;**55**(9):2391-2398
- [23] Houbraken JAMP, Frisvad JC, Samson RA. Taxonomy of *Penicillium citrinum* and related species. *Fungal Diversity*. 2010b;**44**:117-133
- [24] Houbraken J, Verweij PE, Rijs AJMM, Borman AM, Samson R, A. Identification of *Paecilomyces variotii* in clinical samples and settings. *Journal of Clinical Microbiology*. 2010a;**48**(8): 2754-2761
- [25] Barreto MC, Houbraken J, Samson RA, Frisvad JC, San-Romão MV. Taxonomic studies of the *Penicillium glabrum* complex and the description of a new species *P. subericola*. *Fungal Diversity*. 2011;**49**:23-33
- [26] Sidrim JJC, Moreira Filho RE, Cordeiro RA, Rocha MFG, Caetano EP, Monteiro AJ, et al. Fungal microbiota dynamics as a postmortem investigation tool: focus on *Aspergillus*, *Penicillium*, and *Candida* species. *Journal of Applied Microbiology*. 2010;**108**:1751-1756

*Edited by Luis Ricardo Hernández  
and Martín Alejandro Serrano Meneses*

From the success of the first and second volume of this series, we are enthusiastic to continue our discussions on research topics related to the fields of Food Science, Intelligent Systems, Molecular Biomedicine, Water Science, and Creation and Theories of Culture. Our aims are to discuss the newest topics, theories, and research methods in each of the mentioned fields, to promote debates among top researchers and graduate students and to generate collaborative works among them.

Published in London, UK

© 2022 IntechOpen

© andromeda / Dollarphotoclub

**IntechOpen**

