IntechOpen

Bearing Technology

Edited by Pranav H. Darji

BEARING TECHNOLOGY

Edited by **Pranav H. Darji**

Bearing Technology
http://dx.doi.org/10.5772/63262
Edited by Pranav H. Darji

Contributors

Viacheslav Vavilov, Flur Ismagilov, Denis Gusakov, Valéria Cristina Maria Nasscimento Leite, Jonas Guedes Borges Da Silva, Germano Lambert-Torres, Luiz Eduardo Borges Da Silva, Giscard Giscard Francimeire Cintra Veloso, Levy Ely Oliveira, Erik Leandro Bonaldi, Fabrizio Stefani, Andrea Perrone, Luca Ratto, Ramon Francesconi, Yonmook Park, Tian Ran Lin, Yu Kun, Jiwen Tan

Notice

First published in Croatia, 2017 by INTECH d.o.o.
eBook (PDF) Published by IN TECH d.o.o.
Place and year of publication of eBook (PDF): Rijeka, 2019.
IntechOpen is the global imprint of IN TECH d.o.o.
Printed in Croatia

Legal deposit, Croatia: National and University Library in Zagreb

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

Bearing Technology
Edited by Pranav H. Darji
p. cm.
Print ISBN 978-953-51-3183-0
Online ISBN 978-953-51-3184-7
eBook (PDF) ISBN 978-953-51-4821-0

Meet the editor

Prof. (Dr.) Pranav H. Darji graduated in Mechanical Engineering from North Gujarat University, Gujarat (India), in 1998. He received his postgraduate degree from Gujarat University, Gujarat (India), in Mechanical Engineering specialization with CAD/CAM in 2001. He received his PhD degree from S. V. National Institute of Technology, Gujarat (India), in Mechanical Engineering in the specific field of Rolling-Element Bearing in 2015. He is working as a professor and head in the Department of Mechanical Engineering at C. U. Shah College of Engineering and Technology, C. U. Shah University, Wadhwan City, Gujarat (India), since 2010. He has worked as an assistant professor from 2004 to 2010 in the same organization. He has previously worked as a lecturer at N. M. Gopani Polytechnic Institute, Gujarat (India), from 1998 to 2002 and the head of the Department of Mechanical Engineering from 2002 to 2004. Since 2015, he is also working as a director of P. G. Studies and Research at Research, Development and Innovation Centre of C. U. Shah University as an additional responsibility and associated with various research activities. His research interest focuses on areas of bearing technology, tribology, CAD, FEA and advanced machine design. On the same area, he has published more than 57 research papers in various reputed national and international journals and conferences. He has edited an international book *Advances in Tribology* in the year 2016. He has organized many conferences and workshops for the benefit of the researchers, academicians and industrialists. He has guided more than 22 dissertations of postgraduate students. He has given more than 12 invited presentations.

He serves as an editorial board member of reputed international journals. He is a life member of professional bodies like Institution of Engineers (India), Tribology Society of India and Indian Society for Technical Education. He is a registered chartered engineer (Mechanical Engineering).

Contents

Preface

In the twenty-first century, bearings are expected to perform better in the form of various operating conditions, that is from low speed to extremely high speed and from low load to huge load applications. The expectations from the field of bearing technology are great. During the recent years, we have been witnessing the development of a new generation of mechanical systems that are highly miniaturized and very sophisticated, yet extremely robust. Technological progress creates increasingly arduous conditions for rolling mechanisms. Advances in many fields including aeronautics, space and atomic power involve extreme operating speeds, temperatures and environments. Moreover, rolling bearings are used in diverse precision machinery operations, for example, the high-load, high-temperature, dusty environment of steel making; the dirty environments of earthmoving and farming; the life-critical applications in aircraft power transmissions and the extreme low to high temperature and vacuum environments of deep space. They perform well in all of these applications.

Most information and data pertaining to the performance of rolling bearings are presented in manufacturer's catalogues. These data are almost entirely empirical in nature or information contained in International Organization for Standards (ISO) or similar publications. These data pertain only to applications involving slow speed, simple loading and nominal operating temperatures. If an engineer wishes to evaluate the performance of bearing applications operating beyond these bounds, it is necessary to return to the basics of rolling and sliding motions. Particularly since 1960, much research has been conducted for rolling bearings and rolling contact phenomena. In this modern age of deep-space exploration and cyberspace, many different kinds of bearings have come into use. Each of these bearing types excels in some specialized field of application.

Chapter 1 deals with the new concept for the design of a micro-GT support system. Instead of using a single type of bearing as usual, the new system includes different types, in order to take advantage of the best characteristics of each one and, simultaneously, to minimize the effects of the relevant flaws. The second chapter presents an electromagnetic levitation system for active magnetic bearing wheels. A meaningful electromagnetic force by using the singular value decomposition is also derived, and numerical simulation and experimental results on the control of the electromagnetic levitation system are discussed. Chapter 3 presents a general overview of various condition monitoring and fault diagnosis techniques for rolling-element bearings in the current practice and shows the pros and cons of each technique. The techniques introduced in this chapter include data acquisition techniques, major parameters used for bearing condition monitoring, signal analysis techniques and bearing fault diagnosis techniques using either statistical features or artificial intelligent tools. Chapter 4 examines different types of bearing supports. Technical parameters of different types of bearing supports are presented. The effectiveness of some types of bearings is

determined. General approach for the calculation of bearing overall dimensions is considered. In Chapter 5, motor current signature analysis based on squared envelope spectrum is applied in order to identify and to estimate the severity of outer race bearing faults in induction machine. Bearing fault characteristic components are extracted combining summation of phase currents, pre-whitening, spectral kurtosis and squared envelope spectrum analysis.

I have endeavoured to maintain the material presented in an up-to-date and useful format. I hope that the readers especially young researchers and engineers in this field will find this book as useful and informative. I feel amazing pleasure to edit this book. I would like to express my sincere gratitude to all authors for their outstanding chapters. I also wish to acknowledge the InTech editorial staff, in particular, Mr. Edi Lipovic, Publishing Process Manager, for indispensable technical assistance in book preparation and publishing. And of course, I would like to thank my parents, wife Veera and loving son Fagun for the support they always gave to me—thanks for being close to me during all my life.

Prof. (Dr.) Pranav H. Darji
Professor and Head, Department of Mechanical Engineering
C. U. Shah College of Engineering and Technology

Director, P. G. Studies and Research
Research, Development and Innovation Centre
C. U. Shah University
Wadhwan City, Gujarat (India)

Chapter 1

Comparative Analysis of Bearings for Micro-GT: An Innovative Arrangement

Fabrizio Stefani, Andrea Perrone, Luca Ratto and Ramon Francesconi

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/67147

Abstract

Microgas turbines are a widespread technology in cogenerative and propulsion applications. Bearings are a key factor in their design and development. The aim of the present research work is the development of the support system for a typical microturbine intended for power generation. To this goal, the present chapter defines the typical requirements of the machine and, afterward, describes the different technologies available to develop the support system of a reliable microturbine. Conventional (rolling element and oil-film) supports and cutting-edge (magnetic, aerodynamic, and aerostatic) bearings are reviewed. Particularly, their suitability to the operating conditions is compared by means of a literature review and elaboration of the relevant data. By analyzing all this information, a new concept for the design of a micro-GT support system is devised. Instead of using a single type of bearing as usual, the new system includes different types in order to take advantage of the best characteristics of each one and, simultaneously, to minimize the effects of the relevant flaws. The innovative support system requires a suitable bearing arrangement, which is compared with the conventional ones. The conceptual design of the innovation is completed by a discussion of its advantages, drawbacks, and prospective improvements.

Keywords: microturbine, gas-turbine, bearing arrangement, bearing performance, foil bearings

1. Introduction

Microturbine technology owes its origins to the military and aerospace industry, where the need of compact and high power density engines justifies significant production and development costs. Later, micro-gas turbine (micro-GT) units have been used in small-scale power

generation as well as cogeneration, and they are involved in the air compression as well as conditioning market. Recently, manufacturers are addressing their efforts to new market areas, for example, in powering hybrid electric vehicles as well as autonomous robots in the case of small-size machines.

In the following, such microturbines are intended as autonomous power generators. Despite their name, excluding portable devices and MEMS, typical shaft diameters of commercial microturbines are roughly 10 mm, and their electric power is in the order of 100 kW.

In the design of a microturbine, the bearing choice is not a trivial issue due to the high rotation speed and working temperature. In these operating conditions, the "classical" engine design criteria for choosing the most adequate bearing type (rolling or sliding bearings), which are based on dimensions and/or nominal power rate, cannot be adopted. According to such traditional design criteria, the choice for machines of small and large dimensions is unavoidable, i.e., rolling and sliding bearings are employed, respectively. For intermediate power machines, on the contrary, specific choices have to be carried out: on the one hand, rolling bearings have smaller overall dimensions and purchasing costs, and on the other hand, sliding bearings are more reliable.

Therefore, for microturbines, a comparative analysis of the available support systems, particularly focused on studying the influence of operating conditions (speed, temperature, and loads), is required. Accordingly, the present chapter provides a detailed comparison of bearing technologies by means of literature review and analysis of published data. Cutting-edge (e.g., magnetic, air, and ceramic bearings) and well-established solutions (e.g., steel rolling element and oil-lubricated slide bearings) are both considered in order to provide a large perspective.

On the basis of such a comparative study, an innovative design of a support system for micro-GTs capable of overcoming the limits of modern units is proposed. It employs different types of bearings and requires a proper design of their arrangement and coupling in order to take the maximum advantage of the peculiarity of each bearing. After a brief description of the existing bearing arrangements, the present chapter deals with the conceptual design of the new support system.

In synthesis, the present chapter defines the bearing requirements on the basis of design specifications of a typical micro-GT unit. Afterward, it studies the suitability of different bearing types for the operating condition requirements by means of a literature review. Finally, the innovative support system is proposed and compared with the current technical solutions.

2. Micro-GT design specifications

In a nutshell, the microturbines currently on the market have the design features described in the following references [1, 2]. They work according to a Brayton open cycle, which very often takes advantage of exhaust gases heat recuperation for air compressor discharge. The structure of these single-shaft microturbines includes an annular, or silo combustor, single-stage radial flow compressor as well as expander, and an optional recuperator. The electric power

ranges between 30 and 330 kW. They work with low pressure ratios (typically in the range 3–5), high nominal rotation speeds (ranging between 43,000 and 116,000 rpm), and relatively low efficiency (17–20% for simple cycle machines and around 30% for recuperated machines). The design life is usually between 60,000 and 80,000 hours with overhaul [3]. Turbine entry temperatures range between 700 and 1000°C, while exhaust temperatures are in the range of 260–310°C (for recuperated machines).

Accordingly, **Table 1** reports the main data of a micro-GT unit designed in a previous work [4], which can be assumed as the typical machine and reference for subsequent calculations.

Design variable (unit)	Value
Rotational speed N (rpm)	70,000
Shaft diameter D (mm)	15
Electric power P (kW)	110
Pressure ratio	4
Inner radius of compressor blades (mm)	7.5
Outer radius of compressor blades (mm)	59
Inner radius of turbine blades (mm)	7.5
Outer radius of turbine blades (mm)	72.5
Turbine torque M_t (N m)	30
Compressor torque M_c (N m)	−15

Table 1. Main data for the design of the reference micro-GT unit.

2.1. Bearing external loads

Table 2 gathers the bearing loads computed for the reference unit. The total radial load is due to the shaft weight, and it is directed downward. The positive direction of the axial forces listed in **Table 2** is from compressor to turbine, as also shown by the direction of *z*-axis in **Figure 1**.

The total thrust load is computed as the algebraic sum of compressor and turbine axial forces. Both of the forces are the resultant of the axial thrusts exerted on the blades and on the backside of the impellers by the working fluid.

Consequently, together with the area of the impeller backside, the pressure on the clearances between casing cover and impeller back shroud (backside pressure) plays an important role in determining the nominal axial thrusts, which are plotted in **Figure 1** as a function of the backside pressure. By means of CFD simulations, a backside pressure equal to 0.25 MPa (Case A) has been calculated so that both the impeller thrusts are directed toward the external side of the unit. Nevertheless, for different impeller geometries, which may yield different pressure drops between compressor delivery and clearances, compressor, turbine, and total thrust directions can reverse, as shown by **Figure 1**, for backside pressures lower than 0.15 MPa, and as a conse-

quence, such cases (both thrusts directed toward the inner side of the machine) must be taken into account. To this purpose, for the same reference value of the total axial load (500 N) Case B is identified in **Figure 1**. The axial loads in both Case A and B are summarized in **Table 2**. An axial force reversal can also occur during transient operation, e.g., the start/stop phase of the unit.

Design variable (unit)		Value
Radial load (rotor weight) W (N)		40
Total thrust load (absolute value) T_{ref} (N)		500
Axial thrust on compressor blades (N)		2100
Axial thrust on turbine blades (N)		−3010
Case A	Pressure on the back side of the impellers (gauge) (Pa)	250,000
	Compressor thrust T_c (N)	−600
	Turbine thrust T_t (N)	1100
Case B	Pressure on the back side of the impellers (gauge) (Pa)	80,000
	Compressor thrust T_c (N)	1200
	Turbine thrust T_t (N)	−1700

Table 2. Loads acting on the bearings of the reference micro-GT unit.

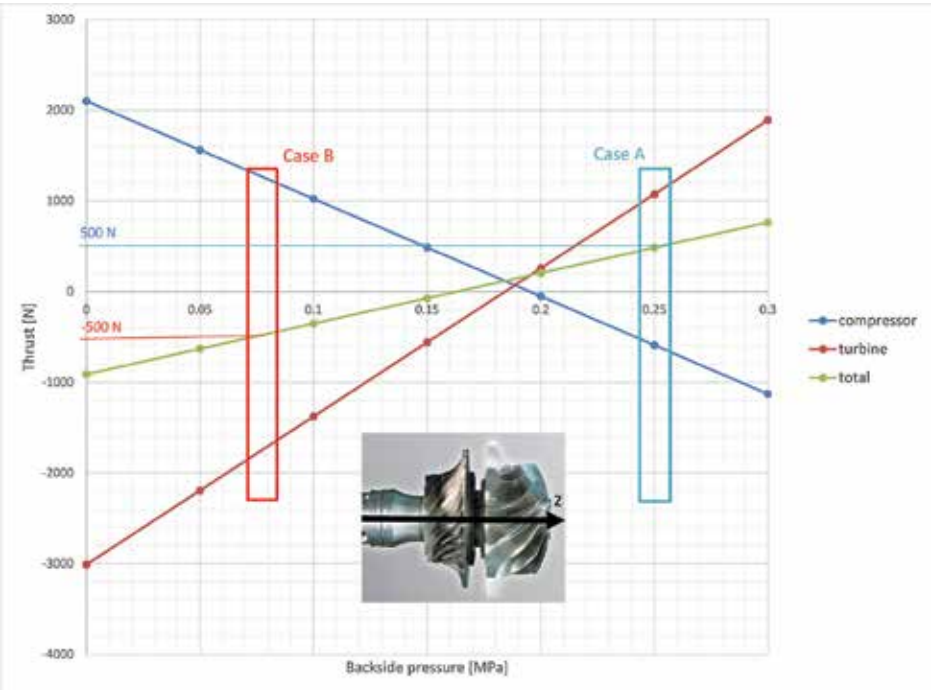


Figure 1. Compressor, turbine, and total thrust acting on the shaft in nominal conditions for the reference micro-GT as a function of the backside pressure.

http://dx.doi.org/10.5772/67147

2.2. Bearing requirements

Requirements are related to the operating conditions of the units (operative speeds and temperatures) and bearing performance (duration, load-carrying capacity, and power loss).

The high rotation speeds of micro-GT shafts yield different problems (e.g., ball spinning, centrifugal loading, and skidding), which have greater effect on larger bearings. Therefore, an important requirement is that bearings have a high peripheral speed limit.

In the following, the effect of peripheral velocity is quantified in terms of the DN speed factor (where D is the diameter of the shaft housing expressed in mm, and N is the rotation speed expressed in rpm).

Micro-GT units must usually operate in high-speed conditions, which for rotors are usually characterized by speed factors ranging between 1×10^6 and 2×10^6 mm rpm. In ultra-high-speed rotor applications DN overcomes 2×10^6 mm rpm, as in the case of mobile power systems [5], which are beyond the scope of the present paper.

Indeed, in the case of the reference microturbine, the operative speed factor is roughly 1×10^6 mm rpm (see **Table 1**) so that bearings with DN limit lower than such operative value must be discarded.

In addition, the bearings must operate throughout the domain of possible temperature conditions of the microturbine. Such temperatures range between 100 and 1000°C in nominal operation, while during start-up, they extend down to room temperature. Expected working temperature of the bearings depends on their location and relevant thermal management, e.g., type of cooling or vent.

Bearing duration should preferably overcome 70,000 hours (the average duration of the units) and should not be limited by the number of start/stop cycles so that maintenance and bearing replacements could be minimized.

Reasonably, predicted axial load is a magnitude order higher than radial load (**Table 1**). The former load is significant as far as metal fatigue and wear are concerned, while the latter is so light that it may cause stability problem at high speed in self-acting radial bearings.

Since the state-of-the-art efficiency of microturbines is quite low, bearings must not reduce further this value. To this goal, they should be designed for the minimum power loss (e.g., friction) and their possible power input should not be significant.

Finally, depending on the shaft layout, resonance eigenfrequencies usually occur at a speed lower than the operating one; therefore, the vibration amplitude should be limited by the bearing damping.

3. Bearing comparison

The adequacy of the most reliable types of bearings is studied. To this purpose, the behavior of the following types of bearings is analyzed:

(1) steel rolling element bearings, lubricated by grease or oil;

(2) sliding bearings, lubricated by oil;

(3) air (film) bearings;

(4) magnetic bearings;

(5) ceramic bearings.

The first two categories include conventional bearings, taken as a reference for performance comparison. The last item includes rolling element bearings entirely manufactured in ceramic material and hybrid bearings, made up of steel rings and ceramic balls, with or without film coatings on the races.

Only supports based on well-established technology are analyzed, while research solutions (still in development), such as squeeze film bearings, hydroinertia gas bearings, ferrofluid bearings, and metal mesh foil bearings, are not considered.

The assessment of compliance of the different bearing types with micro-GT operating conditions, including speed (DN factor), temperature, and loads, requires a proper comparison of literature and technical data.

3.1. Operating speed

Operative and absolute speed limits of the different bearing solutions are compared in **Figure 2**.

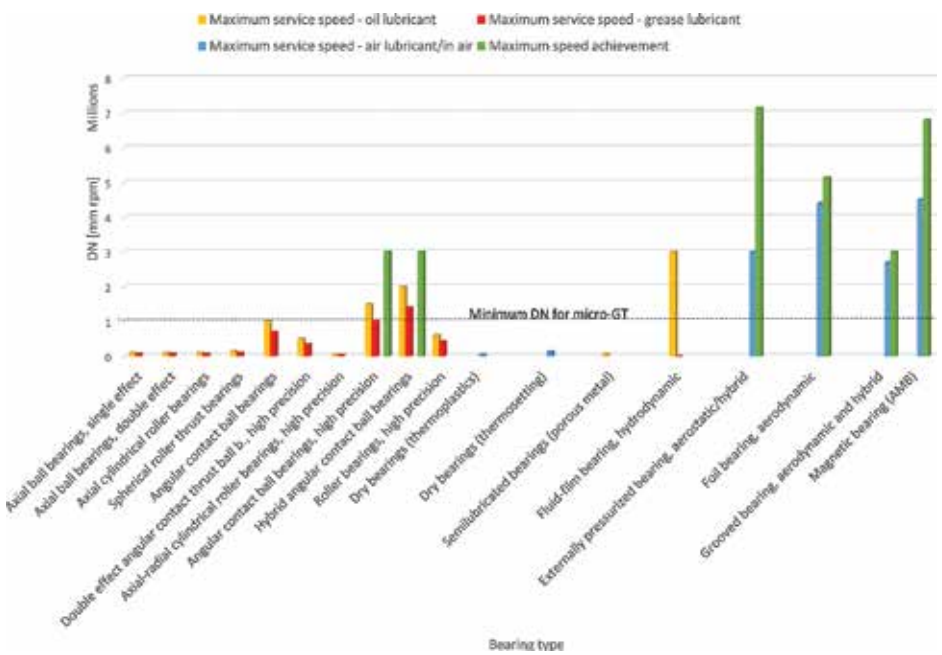


Figure 2. Maximum speed factors for different bearings.

The speed limit of rolling element bearings is mainly due to the skidding of the elements on the rings [5]. Among such bearings, high-precision angular contact ball bearings by means of suitable tolerances, osculation, rolling element size, and number can reach the high speed required by micro-GT (DN > 1×10^6 mm rpm). The maximum DN reached by high-precision ball bearings is about 3 million mm rpm [6, 7]. Higher DN values can be reached by angular contact bearings by means of specific lubrication systems [8].

Manufacturer's catalogs (SKF, Schaeffler) show that the use of ceramic balls in place of steel ones can yield an operating speed increase of roughly 20–30%. Such a result is confirmed by the data reported in reference [9], i.e., speed of hybrid bearings can be increased by 20–30% compared with conventional ones. Accordingly, in comparison with steel bearings, a 60% decrease of centrifugal load matched with the 30% load capacity reduction predicted by Hertz theory finally provides a maximum allowable increase in rotation speed of just 32%. It decreases to 10% for all ceramic bearings.

The DN operative limit reported for oil-lubricated slide bearings regards bearings of turbines for power generation plants [10] and for induction motors [11]. In such cases, operative speed is only restricted by strength limits and by the allowable temperature. In microturbine applications, the operative speed of slide bearings is also limited by their stability and suitable bearing geometries are required, e.g., elliptical and pocket bearings, multilobe bearings, tilting pad bearings (listed from least to most stable).

Waumans et al. [12, 13] report that the highest achieved DN-number for a self-acting bearing operated in air is 7.2×10^6 mm rpm. It is reached by an aerodynamic journal bearing stabilized by means of a grooved bush with a wave-shaped geometry as well as a flexible and damped support structure.

As far as foil bearings are concerned, maximum speed is relevant to a Ø8 mm bearing for microturbines operating up to 642,000 rpm (DN = 5,136,000 mm rpm) from results in reference [14].

Among the grooved bearings, the maximum operative DN of a Herringbone grooved journal bearing (HGJB) with enhanced grooved geometry [15] is 2.7×10^6 mm rpm, and it is lower than foil bearing one. In addition, experimental results confirm that grooved hybrid bearings (GHBs) can run satisfactory at speeds in excess of 3.0×10^6 rpm mm. Nevertheless, such DN is assumed as upper speed limit for grooved bearings, as they are prone to destructive whirl instability at ultrahigh speed [16].

Hybrid aerostatic bearings are suitable due to both the aerostatic stabilizing effect at high speeds and the low air consumption, which has extremely small effect on the global efficiency [17]. For a proper operation, they require air supply at high speed, when it is actually available in micro-GT systems. Data relevant to the maximum speed reached by aerostatic bearings refer to reference [18]. The relevant maximum operative speed is also documented in reference [19].

In today's industrial applications, active magnetic bearing (AMB) rotational speeds are in the range of about 180,000 rpm for a grinding spindle, or about 300,000 rpm for small

turbo-machinery [20]. The latter value, by assuming D = 15 mm as suggested by **Table 1**, corresponds to DN = 4.5 × 10^6 mm rpm, which is confirmed by the data suggested by SKF. Anyway, by means of carbon fiber bandages in the rotor, 6.8 × 10^6 mm rpm can be reached [19]. Such a value is a documented maximum speed for actual applications rather than a maximum theoretical limit, which is unknown as in the case of air foil bearings [21].

3.2. Operating temperature

Table 3 lists the maximum operational temperatures of the bearings.

Bearing type	Max operating temperature (°C)
Rolling element	125–315
Ceramic hybrid	350
Ceramic integral	800
Oil sleeve (journal)	125–150
Oil tilting pad	125–150
Hydrostatic	125–150
Magnetic (AMB)	500–600
Magnetic (PMB)	150–300
Aerodynamic, air foil	650
Aerodynamic, rigid/grooved	500
Aerostatic	900

Table 3. Maximum operating temperatures for different bearing types.

As far as steel rolling bearings are concerned, in addition to the lubrication system and the lubricant characteristic, steel reaction to heat and dimensional stability influences their endurance at high temperature. Generally, hardness of steel starts to decrease as temperature rises over 200°C. In addition, as steel heats up, phase transformation occurs and the bearing parts expand. The maximum dimensionally stable temperature ranges between 120 and 250°C, depending on steel type (source: NSK). Accordingly, a limit temperature of 180–260°C for rolling element bearings is indicated in reference [21]. Similarly, a 125–150°C operative limit for AISI 52100 bearings that can be specially stabilized up to 200°C and up to 315°C by using tool steel bearing materials is reported in reference [10]. The authors of reference [11] confirm that operative temperature is usually kept below 150°C except for heat-stabilized bearings.

Ceramic bearings behave better than steel rolling bearings at high temperature. Indeed, hardness and strength of silicon nitride do not deteriorate at high temperatures when compared with those of bearing steel. Particularly, the advantage of all ceramic bearings over hybrid

http://dx.doi.org/10.5772/67147

bearings consists especially on the major capability of working in dry or underlubricated conditions, at high temperature (<800°C) and corrosive environment [9]. Such a suitability to high temperatures is confirmed in reference [10], which indicates a limit of 650°C for ceramic bearings with vapor phase lubrication or solid lubricants.

Hydrodynamic bearings are the most limited in operating temperature. Indeed, such a limit comes from bearing surface and oil endurance. Soft metal and Babbitt limit upper temperatures are usually in the 125–150°C range [10]. Limit operating temperature of hydrocarbon oils is 93°C [22], while high temperature oils can also reach 150–200°C, e.g., silicon oils. Therefore, oil lubricated slide bearings are ultimately limited in temperature by bearing surface resistance.

Gases can be employed as lubricants over an extremely wide range of temperatures. For gas bearings, operating temperature limits come from shortcomings of solid components (journal and bearing material), not of the lubricant. Electric motors with ceramic windings supported by gas bearings can work for long periods at temperature up to 500°C [23], which is assumed as the limit temperature for aerodynamic bearings.

Foil bearings require the use of a solid lubrication to prevent wear and reduce friction during instances of contact, i.e., at low-speed conditions at start-up and shutdown. Since common lubricants, e.g., graphite and moly-disulfide (MoS_2), are limited to 150°C, solid lubrication is often obtained on the shaft and top foil layer by means of thin, soft polymeric film and sacrificial coatings. Innovative coatings, e.g., nanocomposite for journals and CuAl alloy for top foils, can reach temperatures as high as 650°C, which not by chance is also the limit operating temperature indicated for foil bearings in reference [21]. By using well-established tribo-solutions like polymer coatings, air bearing operation is roughly limited below 300°C [24].

Ceramic aerostatic bearings can reach a temperature as high as 800°C [23] and, in general, temperatures up to 900°C and speeds up to 65,000 rpm are feasible for externally pressurized gas bearings [25]. Indeed, aerostatic bearings have the highest temperature limit among the bearing analyzed. It is higher than temperature limit of aerodynamic bearings due to the external air supply, since air cools as it expands. In addition, the very low friction losses avoid thermal expansion due to viscous heating.

Among magnetic bearings, a great drawback of passive magnetic bearings (PMBs) comes from high temperature operation requirements of micro-GT systems, as permanent magnet stability is affected by temperature. Maximum practical operating and Curie (demagnetization) temperatures for the major classes of permanent magnet materials are in the ranges of 150–540 and 310–860°C, respectively. On the contrary, AMBs can work in extreme temperature environments (500–600°C) [21].

3.3. Load-carrying capacity and life

The maximum specific loads for the bearing types in analysis are summarized in **Table 4**.

Bearing type	Max service-specific load (MPa)
Rolling element	2
Ceramic hybrid	2
Ceramic integral	2
Oil sleeve	2.1
Oil tilting pad	4.4
Hydrostatic	6
Magnetic (AMB)	0.8
Magnetic (PMB)	0.4
Aerodynamic, air foil	0.7
Aerodynamic, rigid/grooved	0.1–0.2
Aerostatic	0.2

Table 4. Maximum service-specific loads for different bearing types.

For rolling element bearings, the "carrying capacity" is the ability of the bearing to carry a given load for a predetermined number of cycles or revolutions [26].

The maximum documented specific load for rolling element bearings operating in gas turbines is reported in reference [21]. **Table 5** reports the durations of bearings in a shaft support system designed by means of catalog high-precision rolling bearings solely. They are computed according to the adjusted basic rating life [27] by using data in **Table 1** and radial as well as thrust loads in **Table 2**. It is assumed that radial load is equally distributed between two sets of high-precision angular contact bearings and only one (locating) set carries the axial load in order to allow the thermal dilatation of the shaft. The solutions with both 2 and 5 matched bearings in the sets (in tandem arrangement) that carry the axial load are not satisfactory compared with the machine life (about 70,000 hours).

Set	L_1 (h)	L_{10} (h)	L_{50} (h)
One bearing, radial load W	267,960	1,276,000	6,380,200
Two bearings, radial load W and axial load T_{ref}	65	310	1549
Five bearings, radial load W and axial load T_{ref}	451	2150	10,747

Table 5. Expected life for sets of angular contact high-precision bearings, carrying half of the rotor weight and, if specified, the axial load.

Dynamic load ratings of steel bearings can also be used for ceramic bearings of the same dimensions [28], since from test results and predicted values service life of ceramic bearings is longer than that of steel bearings, except for heavy loads.

For the remaining bearings, fatigue is not the main concern, and the external load can be treated as static.

http://dx.doi.org/10.5772/67147

Sleeve bearings life is theoretically infinite and extremely long when they are properly maintained. Significant wear may occur only during extended start-up or coast-down periods, as mixed lubrication occurs at low speed. If the frequency of such events is high, hydrostatic jacking is recommended to minimize bearing wear [11].

Load capacity of oil-film bearings is basically a function of speed and oil viscosity so that high temperature plays a role by reducing viscosity. Many specifications limit motor bearing specific pressures to 1.4 MPa, which is often compliant with structural strength. Nevertheless, most journal bearings safely tolerate pressures beyond 2 MPa, as reported in reference [10]. Tilting pad thrust bearings for turbines carry further increase of specific load, and pads are subject to elastic deformation.

Generally, as hydrodynamic and hydrostatic bearings distribute the load over a larger area than rolling element bearings, their load capacity can be higher. Particularly, hydrostatic bearings can support huge loads, higher than hydrodynamic supports, as their pressure distribution is more uniform.

Air (aerostatic and foil) bearings, because of the different film layer, approximately support only a fraction of the load carried by hydraulic bearings with the same dimension. Indeed, specific load capacity of air foil bearings is 0.7 MPa [21], which is roughly one-fifth of the specific load for hydrodynamic bearings (3.5 MPa, on average). Differently, specific load for aerostatic bearings is lower due to a further constraint. Indeed, aerostatic bearings are typically limited to operate at less than 10 atm of pressure (usually 0.69 MPa) for safety reasons and due to the lubricant compressibility, which yields much higher flow rates and pumping power demand for the same pressure in comparison with liquid lubricants [29]. Such value is much lower than supply pressure of hydrostatic bearings, which typically operate at 20–40 atm but can reach 200 atm, when space is not limited and large loads must be supported. Specific load capacity (load per pad area) for aerostatic and hydrostatic bearings is the efficiency multiplied by supply pressure, where the efficiency is typically 25–40%.

According to the basic rating life formula [27], for constant duration load-carrying capacity of rolling element bearings drops as speed rises. On the contrary, load capacity of foil bearings is proportional to rotational speed. Hence, foil bearings outperform rolling element bearings at high speeds [30] but require solid lubrication at low speed in order to reduce friction and wear. Particularly, conventional solid lubrication systems, i.e, thin polymer films, enable over 100,000 h of operation before requiring a major overhaul. Beyond the temperature limits of polymer coatings (300°C), innovative coatings (PS304, Korolon) have demonstrated lives in excess of 100,000 h start-stop cycles under moderate loads (0.34 MPa) and high temperatures (ranging between 178 and 650°C) where such solid lubricants become active. However, bearing operating life is cut by over half (roughly 33,000 cycles) at room temperature (25°C), where the coating does not perform as well [24]. For low temperature start-ups, usually under higher loads, life may further decrease [31].

In order to characterize the behavior of foil bearings by means of a suitable map, in reference [32], a modified Sommerfeld number S′ is correlated with the specific power loss. Foil journal bearings must be designed so that they operate in the high speed (or lightly loaded) regime

(S′ > 6) of the operating map plotted in **Figure 3** in agreement with the suitable regression proposed by the reference authors. Particularly, the nominal working point should be located in the high-speed (lightly loaded) regime, but significantly far from the shaft strength and thermal limits (specific power loss lower than 155,000 W/m^2).

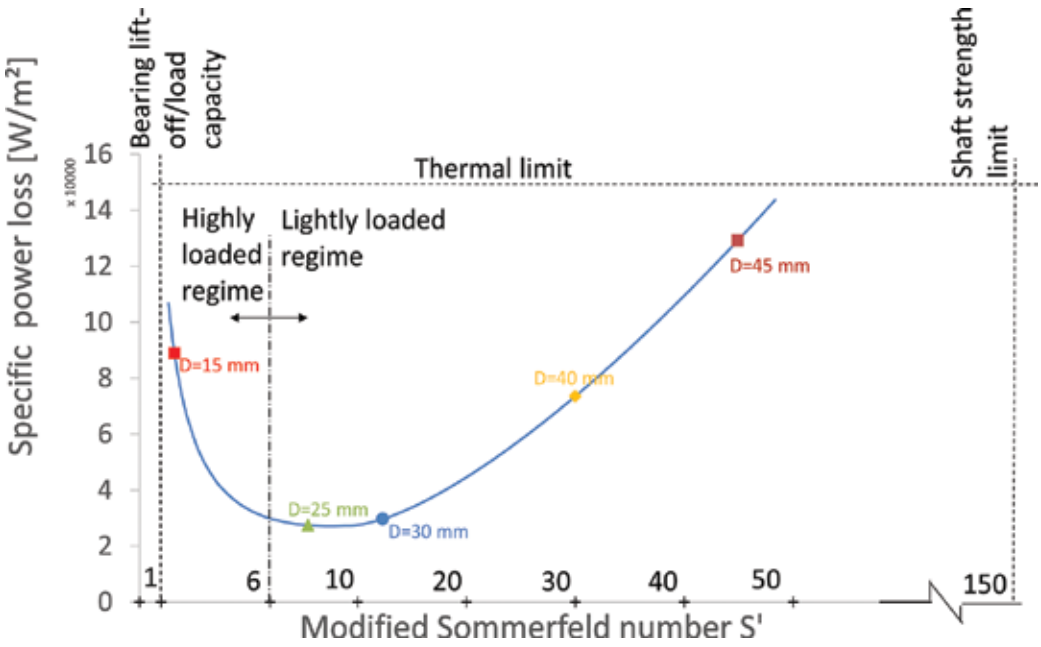


Figure 3. Performance map of foil bearings and operating points of the reference micro-GT for different journal diameters.

In order to locate in the operating map the nominal working point of a single radial foil bearing supporting the reference microturbine shaft, the data reported in **Table 1** are used together with unit (axial) length to diameter ratio (L/D = 1), load coefficient equal to 2.7×10^{-4} N/(mm^3 krpm) (third-generation bearings) and isoviscous behavior assumption. In addition, the total load is approximated by the external one (40 N) as first estimate. By means of such assumptions, the assessment of S′ according to reference [32] suggests that the working conditions are not suitable (S′ = 1.6 < 6). Therefore, the journal diameter must be increased (optimal values range between 25 and 40 mm, as suggested in **Figure 3**). Indeed, transitioning over to oil-free lubrication requires suitable design solutions in that thin shafts are not required anymore to avoid rolling element bearings from operating above their DN threshold. On the contrary, large diameter hollow shafts must be used in order to increase the peripheral speed and, as a consequence, the load capacity of air bearings [31].

As far as the remaining air film bearings are concerned, the different fluid film bearing designs (multilobe and tilting-pad geometries) achieve better stability than plain bearings at the expense of load-carrying capacity. On the contrary, gas-lubricated grooved bearings promise stability with minor reduction of lift [33].

Magnetic bearings have advantageous load-carrying characteristics, whereas load does not drop as speed decreases. Nevertheless, if the electromagnets are unable to support the applied load because they are undersized or malfunctioning, the shaft cannot be levitated and the machine shuts down [34]; therefore, backup bearings are required to protect the rotor against overloads and power loss.

Reasonable specific loads for AMBs range between 0.3 and 1 MPa. Particularly, the maximum specific load value reported in **Table 4** is suggested for AMBs in reference [21] on the basis of the data published in references [35, 36].

For a stacked structure of PMB fabricated using neodymium-iron-boron magnets with a remanence of 1.3 Tesla (42 MGOe magnetic field), the maximum specific load, referred to the axial cross section (LD), is roughly 0.4 MPa and the corresponding axial specific load is 0.6 MPa [37].

Life expectancy of passive bearings is very high, i.e., in excess of 20 years. As AMBs include more components (controllers, coils, and sensors) and a laminated rotor, their life is expected to be shorter but, anyway, AMBs can still last 20–30 years with proper substitutions of failed components.

4. Innovative support system

The conclusion of the review study is that each bearing type has different strengths and weaknesses. Therefore, by using different types of bearings in the same support system, a proper design of their arrangement should be capable of taking the maximum advantage of the peculiarity of each bearing.

To this end, the conceptual design of an innovative support system, which takes advantage of both foil and rolling element bearings, is presented.

4.1. Layout and components

A simplified scheme of the assembly of the cutting-edge support system is depicted in **Figure 4**. In the simplified scheme of the invention used, hereafter load direction is assumed as in the Case B of **Table 2**.

The angular contact ball bearing (3), at the compressor side (**Figure 4**), is capable of carrying both radial and axial loads. As axial load may reverse during start-up of the unit, the bearing must have double effect, i.e., it is made up of two (or more) matched single-row bearings in back-to-back arrangement. Although four-point contact ball bearings are normally not available for precision (high speed) applications, a single bearing is depicted in **Figure 4** for the sake of conciseness of the scheme and clarity. As shown in **Figure 5**, such a bearing is mounted by inserting the external ring inside a suitable bore on the machine frame (9) and by placing the inner ring onto the relevant seat on the shaft with proper tolerances. Particularly, the outer ring must be axially constrained in order to carry thrust loads (locating bearing).

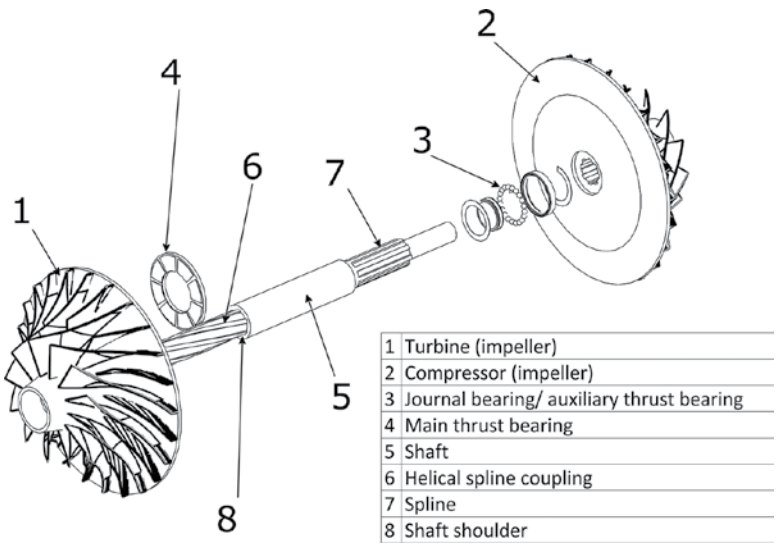


Figure 4. Simplified assembly of the innovative support system.

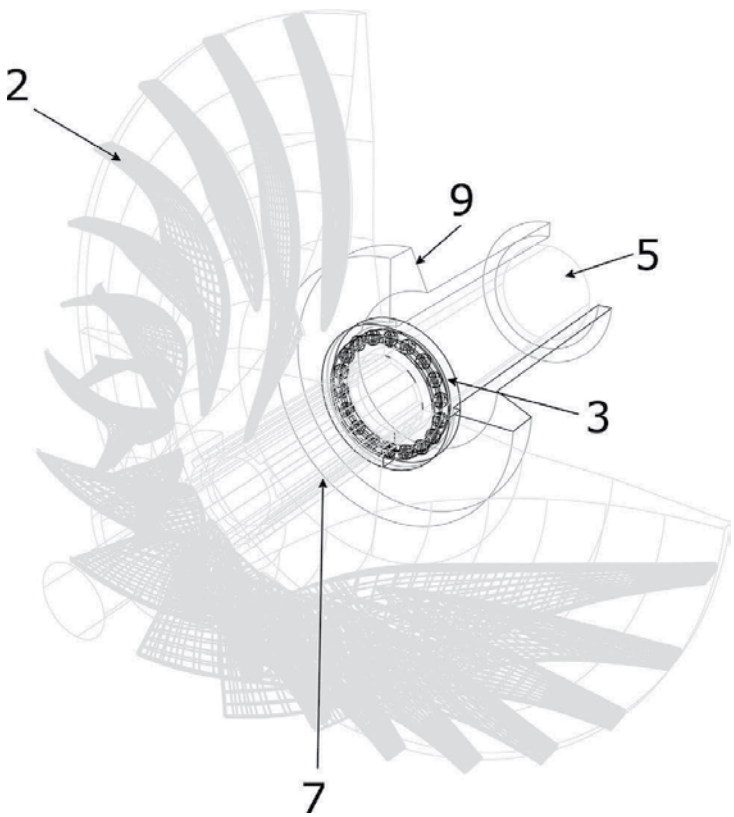


Figure 5. Ball bearing seat and spline coupling at compressor side.

http://dx.doi.org/10.5772/67147

A possible second radial support (always of rolling element type, e.g., a set of angular contact ball bearings in back-to-back arrangement) has not been included in **Figure 4** for simplicity. Nevertheless, it would be useful to avoid that the shaft is in a cantilever configuration, and it should be able to carry radial load solely. Therefore, such additional radial bearing should not be constrained axially to the housing on the frame (nonlocating bearing) so that the axial thermal expansion of the shaft would be allowed.

The plate of the foil bearing, i.e., component (4) listed in **Figure 4**, is fixed to the frame as shown in **Figure 6**, where the runner surface is located on the back of the turbine rotor (1) and the opposite sliding pairs, runner (11) and pads (12), are separated by the clearance c_z. A suitable spacer (10) must allow a proper adjustment of the clearance (or preload in static conditions) to ensure optimal operation of the foil bearing.

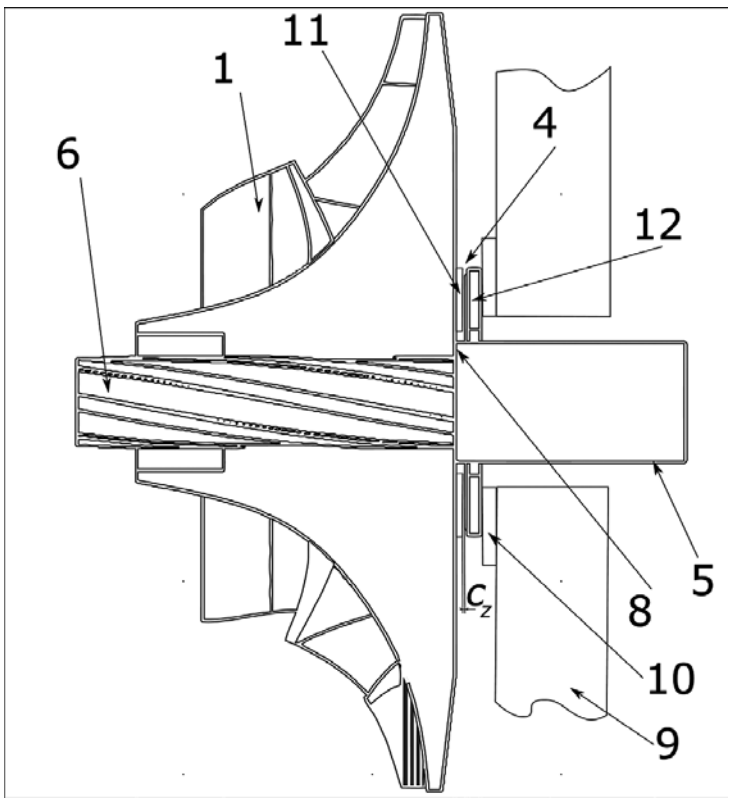


Figure 6. Foil air bearing installation and adjustment.

When the foil bearing has no aerodynamic load-carrying capacity, the turbine thrust acts on the shaft shoulder (8). Differently, after an air film is formed and the relevant aerodynamic pressure is generated, its micrometric thickness causes the runner/impeller assembly (1) to move accordingly in the axial direction so that the contact between the turbine impeller hub and the shaft shoulder does not occur anymore. Consequently, in nominal conditions and

starting from the rotation speed at which the runner is airborne, the only axial load transfer from the turbine to the shaft may occur through the helical spline coupling, as a function of the design helical angle.

Figure 4 also illustrates the coupling between the impeller hubs and the shaft (5). The compressor impeller (2) is fixed by means of a conventional spline pair (7) (made up by equally spaced straight grooves), the profile of which is depicted with parallel sides (or straight teeth). Differently, on the turbine side, a helical spline pair (6) (in which each groove forms a helix around the shaft) is machined. Such particular spline fit provides the additional function of axial load distributor. Of course, involute instead of parallel-side profiles may be chosen for both compressor and turbine wheel spline pairs. The purpose of the helical spline pair is to distribute the axial load between main (4) and auxiliary (3) thrust bearings, as explained in the following paragraphs.

4.2. Main thrust bearing relief

The innovative layout and, particularly, the direct matching of turbine impeller and main thrust bearing (4) allow for its relief during start/stop of the unit.

With reference to thrust loads and symbols given in **Table 2** (Case B), let $F_t = -T_t$ and $F_c = T_c$ be the turbine and compressor thrusts, respectively (**Figure 7**). They are caused by the pressure of the evolving fluid on the relevant impellers.

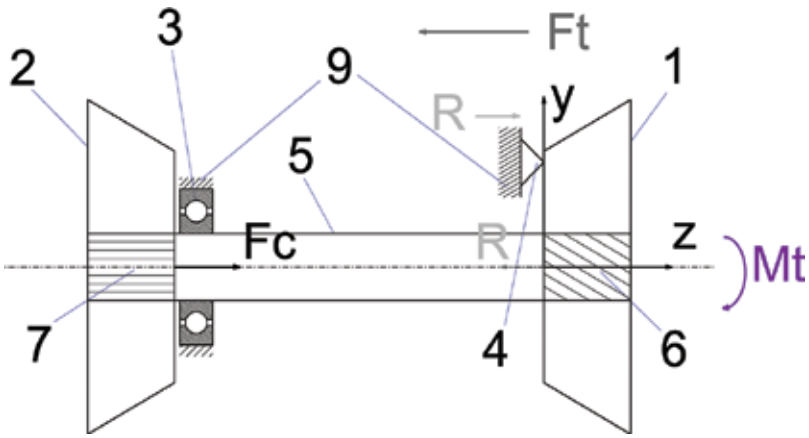


Figure 7. Static scheme (axial forces) of the innovative rotor operating in nominal conditions.

At start-up and until the onset of the aerodynamic lift, the auxiliary bearing (3) carries the whole thrust so that the main thrust bearing (4) is unloaded, and therefore, it works with minimal or no wear. Indeed, in such a condition, the turbine impeller (1) exerts the thrust F_t on the shaft shoulder (8) (visible in **Figures 4** and **6**) instead of on the foil bearing (4). Through the shoulder and the shaft, the thrust F_t is then transferred to the shaft support, i.e., the auxiliary bearing (3), like in a conventional bearing layout. Once the runner of the main bearing becomes

http://dx.doi.org/10.5772/67147

airborne at sufficient speed and a consequent translation of the turbine impeller occurs, automatically it relieves the auxiliary one, as the shaft shoulder (8) does not receive thrust anymore.

4.3. Load partition

According to the explanation of the previous paragraph, in nominal working conditions with no helical spline fit (e.g., by adopting a straight grooved spline on turbine-side too), the whole thrust of the turbine F_t would be carried by the main thrust bearing and the compressor thrust F_c would be supported by the auxiliary bearing. In a conventional shaft-bearing assembly, the reference thrust T_{ref} that loads the single axial bearing comes from the opposite thrusts exerted by turbine and compressor, i.e., $T_{ref} = F_t - F_c$. Therefore, during nominal operation in comparison with a conventional support system, the new assembly design would disadvantage the main thrust bearings, whereas $F_t > T_{ref}$. Differently, by taking advantage of the (turbine-side) helical spline pair as an actuator, a part of the turbine thrust can be transferred from the hub to the shaft. In such a way, any wanted division of the thrust load between the two (main (4) and auxiliary (3)) bearings can be obtained as a function of a single design parameter, i.e., the spline helix angle β shown in **Figure 8**.

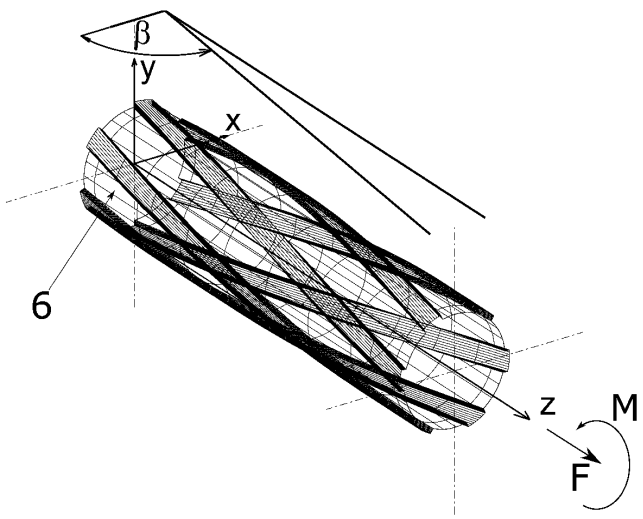


Figure 8. Helix angle and reference system of spline coupling.

Particularly, by means of a suitable choice of the helix angle in the design phase, in nominal conditions, it is also possible to subject the main axial bearing to a thrust T_{ref} as in a conventional support system, while the auxiliary bearing remains axially unloaded. In such a case, as this support is an angular contact ball bearing, it carries solely the radial load, the intensity of which is much lower than the axial loads. Obviously, this is only an example of load division. The new layout allows us the setting of the optimal load division in the design phase as a function of the load-carrying characteristics of the bearings, as well as expected duration and reliability of rolling element supports.

4.4. Law of load distribution

First, the law of load distribution followed during nominal operation by the helical spline pair, employed as a mechanical actuator besides a simple coupling system, is determined.

Figure 7 depicts the forces acting in nominal conditions on the rotor components according to the modifications resulting from the innovation. The constraint simulates the main axial bearing (4), which carries the load F_t−R. The axial forces R are the (equal) action and reaction that the turbine impeller exerts on the shaft through the helical spline. The total thrust that acts on the shaft is F_c−R and is carried by the auxiliary axial bearing (3). The torque M_t is the resisting torque of the turbine due to the pressure exerted on the relevant blades.

A campaign of FEM structural analyses has been carried out on a model of helical spline coupling (**Figure 9**) with parallel-side profiles by varying the design helix angle β from 45 to 135°. Reference system and helix angle β of the spline coupling model are shown in **Figure 8**. In agreement with the helix angle definition, the middle of the range (β = 90°) corresponds to a spline with rectilinear generatrices (straight teeth).

As shown in **Figure 9**, the spline and hub submodels are merged into the coupling model by means of contact elements. Two load cases are analyzed, where the hub section of one model end is submitted to either the axial load F_t or the torque M_t according to the values in **Tables 1** and **2**. Suitable constraints are added to the other end of the model, i.e, zero displacement components d_r, d_θ, and d_z in cylindrical coordinates.

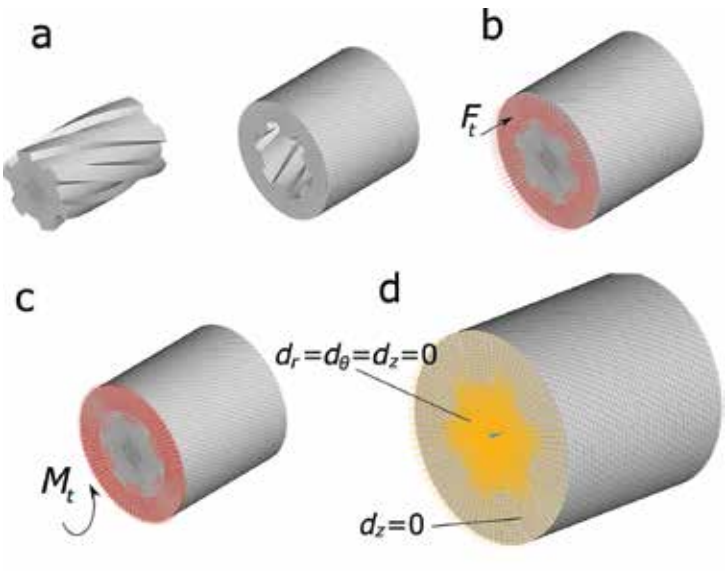


Figure 9. FEM model of helical spline coupling (β = 70°): (a) mesh of shaft and hub, (b) axial force loading, (c) torque loading, (d) boundary conditions.

Numerical results evidence that statics of spline couplings obeys two important rules. First, the load transfer through the spline due to the axial thrust F_t is small (6.7% of the thrust for β = 135° deg, i.e. the helical angle of maximum load transfer), and it is caused by deformations

http://dx.doi.org/10.5772/67147

of the kinematic pair. In other words, if the hub and the shaft were perfectly stiff, the turbine thrust would not be transmitted at all to the shaft through the spline surfaces, but it would be carried by the constraint (4), i.e., the main axial bearing. Second, the load transfer R from the hub to the shaft through the spline due to the torque M_t is actually ruled by the following relationship, valid for a spline pair with perfectly stiff members

$$R = \frac{M_t}{r_p \tan\beta} \tag{1}$$

where r_p is either the pitch radius in case of involute splines or the inner radius for parallel key splines.

Figure 10 compares the shaft and hub reactions computed by means of Eq. (1) and the FEM model with frictionless contact elements (torque load case). The shaft thrust is the reaction that the constraints exert on the grooved part of the shaft, and it represents the load transfer R from the hub to the remaining part of the shaft through the spline surfaces. The hub thrust is equal and opposite according to Newton's third law. Therefore, in case of compliant members, the rule defined by Eq. (1) is still valid with negligible error (0.6% of the transmitted load for $\beta = 135°$). Equivalently, the compliance of the kinematic pair members does not yield perceivable effects for the nominal value of torque M_t. Details of the FEM analyses will be published in the near future.

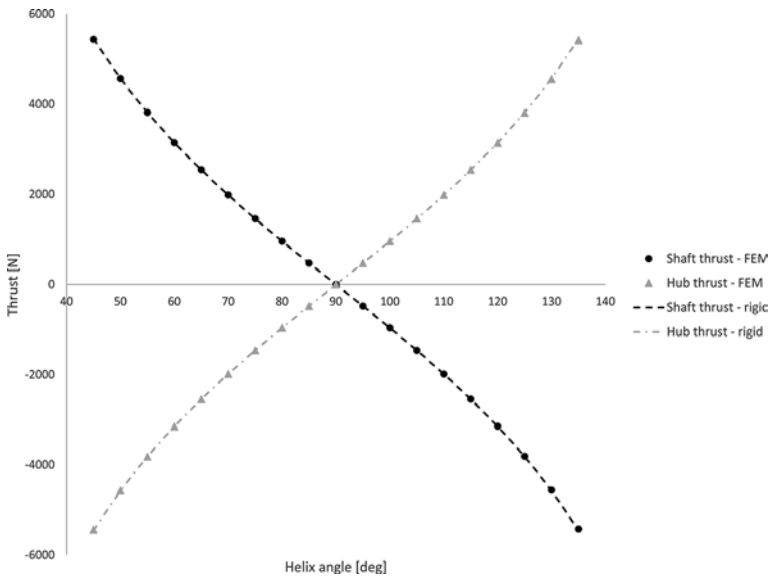


Figure 10. Reactions of shaft and hub constraints for different helix angles under torque load.

4.5. Design of the support system

In the following, the above-explained laws of load distribution are used to choose the design parameters in the assumption, adopted for the sake of simplicity, that the members of the spline pair are stiff. Such an assumption does not lead to significant errors, as proved above.

By means of the resulting design procedure, the load distribution in nominal operating conditions between the two axial bearings can be set by means of a proper choice of the helical angle. To this purpose, **Figure 11** shows for the data reported in **Tables 1** and **2** (Case B) the load transfer through the coupling together with the corresponding axial loads of the bearings as a function of the helix angle β in nominal working conditions.

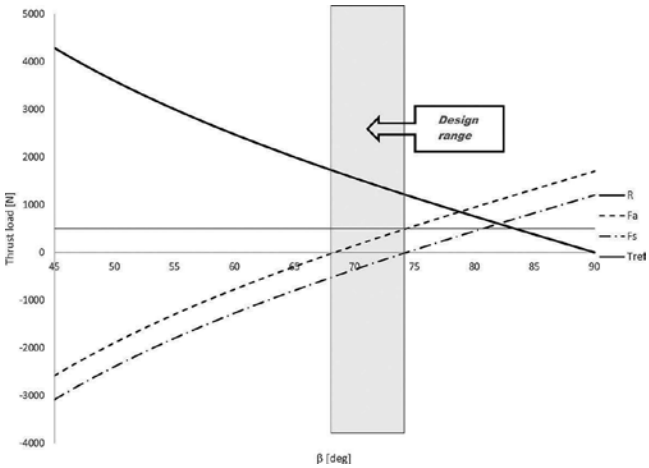


Figure 11. Trends of axial bearing thrusts and load transfer as a function of helix angle in nominal operating conditions.

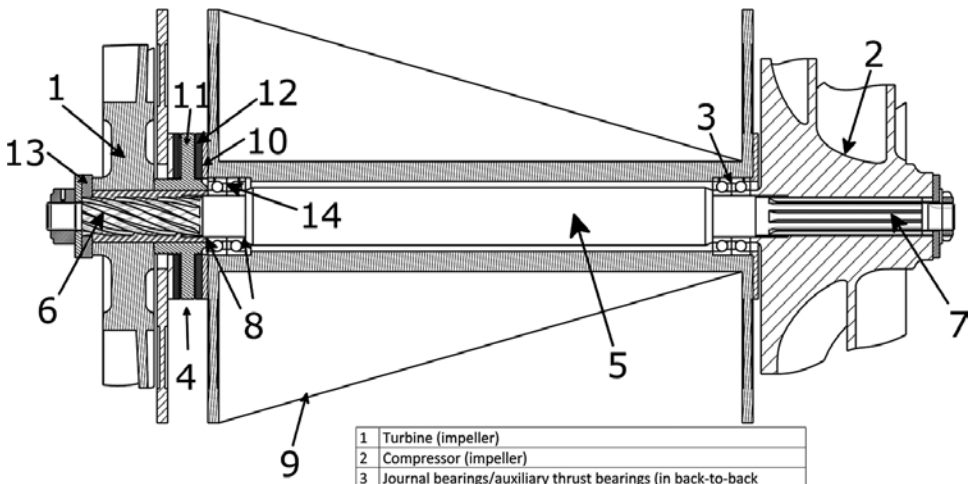
By assuming that positive torque acts on a turbine impeller ($M = M_t > 0$), handedness of the helix must be chosen so that the load transfer R is directed as in **Figure 7**. In other words, the spline coupling must exert (equal) axial forces R opposite to F_t and F_c on the hub and the shaft, respectively. According to Eq. (1), such a condition yields that the range of helix angle in the abscissa of the plot in **Figure 11** cannot exceed 90°.

The load transfer R (thick solid curve) is plotted in **Figure 11** according to Eq. (1). On the basis of its trend, the axial thrusts F_a and F_s that turbine and shaft, respectively, exert on the main bearing (4) and on the auxiliary bearing (3) are plotted as dashed and dash-dotted curves. They are evaluated by means of the relations $F_a = F_t - R$ and $F_s = F_c - R$, which can be deduced by the analysis of **Figure 7**, where turbine and compressor thrusts F_t and F_c are obviously constant in nominal operating conditions. The gray solid horizontal line represents the value of the reference load $T_{ref} = F_t - F_c$, which has to be carried by the single thrust bearing of a conventional machine. As reported in the paragraph dealing with layout and clearly visible in **Figure 11**, when a straight grooved spline (β = 90°) is picked, the load acting on the main axial bearing (4) is greater than in a conventional plant. Indeed, it is equal to the total turbine thrust F_t, while the auxiliary axial bearing (3) supports the compressor thrust F_c. By reducing the helix angle, the loads acting on both bearings begin to decrease, since the helix is oriented in such a way as to exert on the shaft and the turbine impeller a thrust R opposite to F_c and F_t, respectively (**Figure 7**). Particularly, by choosing β roughly equal to 74°, the main axial bearing (4) must carry the reference load ($F_a = T_{ref}$), while the auxiliary one is axially unloaded and, therefore, since it is only subjected to the (light) radial load, it will have an average life exceeding 6 million hours, as specified in the first row of **Table 5**. Differently, for design values of β ranging between 68 and 74°, the load of the main bearing (4) becomes lower than the

http://dx.doi.org/10.5772/67147

reference one, at the expense of the duration of the auxiliary bearing (3), on which the shaft exerts a negative thrust F_s (directed from the turbine to the compressor). For $\beta = 68°$ the load on the main axial bearing (4) is null and, consequently, the thrust exerted by the shaft on the auxiliary bearing (3) assumes its maximum value, i.e., $F_s = -T_{ref}$. In such a condition, the rolling element bearing (3) exhibits the same duration as in a conventional layout (e.g., see the basic rating life reported in the last two rows of **Table 5**). Obviously, since in the simplified layout of **Figure 4**, the main foil bearing is not of double effect, helix angles lower than 68°, which moreover would lead to an even more unsuitable life of the auxiliary bearing, are forbidden. A good design may require a helix angle slightly reduced in comparison with a value of 74°, where the amount of such reduction can be evaluated by taking into account the life and the reliability required for the bearing (3). In the design range, the highest life/reliability of auxiliary bearing (3) is obtained for $\beta = 74°$, while the lowest one, typical of a conventional layout, for $\beta = 68°$; the most severe loading case for the main axial bearing (4), equivalent to that of a conventional layout, occurs at $\beta = 74°$, the most favorable one (zero thrust) at $\beta = 68°$.

Finally, the actual assembly drawing of the invention, suited to both Cases A and B of **Table 2** as well as transient loading conditions, is reported in **Figure 12**. In this case, the total hot clearance between runner (11) and pads (12) of the double-effect air bearing must be higher than that between turbine impeller (1) and the spacer (13) used to adjust the impeller axial clearance. The second set of (nonlocating) angular contact bearings (14) cited in paragraph 4.1 is added.



1	Turbine (impeller)
2	Compressor (impeller)
3	Journal bearings/auxiliary thrust bearings (in back-to-back arrangement)
4	Main (double-effect) thrust bearing
5	Shaft
6	Helical spline coupling
7	Spline
8	Shaft shoulder
9	Frame
10	Pad spacer
11	Runner
12	Main thrust bearing pad
13	Turbine impeller spacer
14	Journal bearings (in back-to-back arrangement)

Figure 12. Section of the micro-GT support system assembly designed according to the invention.

5. Conclusions

For adoption in microturbines, the only conventional rolling element supports selected from manufacturers' catalogs (without including ad hoc solutions) compatible with the speeds at the hand are high-precision angular contact ball bearings, preferably with ceramic parts. Particularly, hybrid bearings allow for a perceivable increase in maximum speed, while all ceramic bearings can reach the high operative temperatures suitable for micro-GT units.

Air supports (hybrid aerostatic as well as aerodynamic) are the most suited oil-free solution for this application. They offer the best compromise between installation costs and performance for micro-GT systems. Aerodynamic bearings are preferred since they need no external supply of pressurized gas (e.g., tapped from the compressor) as they are self-acting. Particularly, foil bearings have good stability and compensate for shaft thermal expansion. Their main drawback is the short phase of dry friction during start up and stop, which limits their life.

Magnetic bearings are another promising oil-free solution. PMBs are suitable for speed and ease of miniaturization as well as independence from external energy input, but the high temperature limits the use of permanent magnets in microturbines since these could demagnetize. Consequently, electromagnets must be employed, although they consume a considerable amount of electrical energy. AMB remains the most suitable magnetic-driven solution because of high speed, temperature, and the control of machine dynamics.

The comprehensive comparison among the different reliable technologies carried out in the first part of this chapter proves that the design of an optimal support system might employ different type of bearings. Indeed, each of them have peculiar capabilities and limits, which make them suitable for particular tasks, e.g., carrying high load continuously or low load during frequent start/stops, working at high or low temperature, etc.

An innovative support system that relies on air as well as rolling element bearings and employs spline couplings has been proposed. The particular bearing arrangement devised for the new system is capable of taking advantage of the best characteristics of both bearing types and, simultaneously, of minimizing the effects of their flaws. Particularly, rolling element bearings behave much better than air bearings during start-stops but ensure limited duration under nominal load and speed of the most efficient micro-GT units. On the contrary, (air) foil bearings provide suitable life and load-carrying capabilities in nominal working conditions but cannot withstand a large number of start-stop cycles of the units. Therefore, the proposed system is designed to switch between the two types of bearings automatically when the unit ends the transient operation. In addition, it employs helical splines as both convenient coupling systems and actuators for the load partition between the two bearing types. Indeed, partitioning the turbine thrust is required in order to optimize the behavior of the support system.

The invention performs better than conventional support systems based on rolling element bearings, since it radically increases bearing life or load-carrying capacity as well as working temperature. In comparison with modern systems based on foil bearings, the invention is advantageous in that the number of start-stop cycles of the machine is not limited, solid

lubrication is not required, the power of the starter can be reduced and retrofitting conventional machines based on rolling element bearings is possible. In any case, replacing shrink fittings with spline couplings can simplify maintenance and inspection of the impellers. Although the proposed support system is not oil-free, the concepts used in its development might be adopted in order to match other type of bearings. For instance, combining thrust foil bearings and radial magnetic bearings may yield an oil-free design with still higher performance.

Author details

Fabrizio Stefani*, Andrea Perrone, Luca Ratto and Ramon Francesconi

*Address all correspondence to: stefani@unige.it

Department of Mechanical, Energy, Management, and Transportation Engineering, University of Genoa, Genoa, Italy

References

[1] do Nascimento M.A.R., de Oliveira Rodrigues L., dos Santos E.C., Gomes E.E.B., Dias F.L.G., Velásques E.I.G., Carrillo R.A.M. Micro Gas Turbine Engine: A Review. In: Benini E., editor. Progress in Gas Turbine Performance. Rijeka, Croatia: InTechOpen; 2013. p. 107–141. DOI: 10.5772/54444.

[2] Section 5. Technology Characterization-Microturbines. In: Catalog of CHP Technologies. U.S.: EPA & CHP; 2014.

[3] Hamilton S.L. The Handbook of Microturbine Generators. Tulsa, Oklahoma: PennWell; 2003. 205 p.

[4] Barsi D., Garbarino T., Perrone A., Ratto L., Ricci G., Stefani F., Zunino P. Micro Gas Turbine Integrated Design: Thermodynamic Cycle, Combustor, Recuperator and Bearings (Part 1). In: Fourth International Conference of Energy Water and Environmental Sciences; 15–17 December 2015; Ras Al Khaimah, United Arab Emirates.

[5] Isomura K., Togo S.I., Tanaka S. Study of Micro-High Speed Bearings and Rotor Dynamics for Micromachine Gas Turbines, Paper 7. In: Micro Gas Turbines. Educational Notes RTO-EN-AVT-131. France: Neuilly-sur-Seine; 2005.

[6] Signer H., Bamberger E.N., Zaretsky E.V. Parametric Study of the Lubrication of Thrust Loaded 120-mm Bore Ball Bearings to 3 Million DN. Journal of Lubrication Technology. 1974;**96**:515–526.

[7] Zwyssig C., Member S., Kolar J.W., Member S., Round S.D. Megaspeed Drive Systems : Pushing Beyond 1 Million r/min. IEEE/Asme Transactions on Mechatronics. 2009;**14**:564–574.

[8] Akamatsu Y., Mori M. Development of Eco-Friendly Oil Jet Lubricated. NTN Technical Review. 2004;**72**:6–10.

[9] Popp M., Sternagel R. Hybrid Ceramic and All Ceramic Anti Friction Bearings. European Space Agency-Publications-ESA SP. 1999;**438**:105–110.

[10] Khonsari M.M., Booser E.R. An Engineering Guide for Bearing Selection. Tribology and Lubrication Technology. 2004;**60**:26–32.

[11] Finley W.R., Hodowanec M.M. Sleeve versus Antifriction Bearings: Selection of the Optimal Bearing for Induction Motors. IEEE Transactions on Industry Applications. 2002;**38**:909–920.

[12] Waumans T., Peirs J., Al-Bender F., Reynaerts D. Design, Optimisation and Testing of a High-Speed Aerodynamic Journal Bearing with a Flexible, Damped Support. In: Tech Dig PowerMEMS 2009; 1–4 December 2009; Washington, DC, USA. 2009. p. 83–86.

[13] Waumans T., Peirs J., Al-Bender F., Reynaerts D. Aerodynamic Journal Bearing with a Flexible, Damped Support Operating at 7.2 Million DN. Micromechanics Microengineering. 2011;**21**:104014.

[14] Hikichi K., Isomura K., Saji N., Esashi M., Tanaka S. Ultra-High-Speed Tape-Type Radial Foil Bearing for Micro Turbomachinery. In: Proc. PowerMEMS 2009; 1–4 December 2009; Washington, DC, USA. 2009.

[15] Schiffmann J. Enhanced Groove Geometry for Herringbone Grooved Journal Bearings. Journal of Engineering Gas Turbines Power. 2013;**135**:102501.

[16] Stanev P.T., Wardle F., Corbett J. Investigation of Grooved Hybrid Air Bearing Performance. Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics. 2004;**218**:95–106.

[17] Peirs J., Waumans T., Vleugels P., Al-Bender F., Stevens T., Verstraete T., et al. Micropower Generation With Microgasturbines: A Challenge. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science. 2007;**221**:489–500.

[18] Waumans T., Peirs J., Reynaerts D., Al-Bender F. On The Dynamic Stability of High-Speed Gas Bearings: Stability Study and Experimental Validation. Sustainable Construction and Design. 2011;**2**:1–10.

[19] Görne J. Using Active Magnetic Bearings for High Speed Machining—Conditions and Benefits. In: TEHNOMUS XVII; May 2013; Suceava, Romania. 2013. p. 13–20.

[20] Schweitzer G. Active Magnetic Bearings-Chances and Limitations. In: 6th International Conference on Rotor Dynamics. Australia: Elsevier Ltd.; 2002.

[21] Clark D.J., Jansen M.J., Montague G.T. An Overview of Magnetic Bearing Technology for Gas Turbine Engines. National Aeronautics and Space Administration Technical Memorandum. 2004;NASA/TM-2004-213177:1–13. Available from: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20040110826.pdf [Accessed: 2017-01-10].

[22] Cappellino C.A., Osborne J.C. The Prediction of Bearing Lubricant Temperatures and Cooling Requirements for a Centrifugal Pump. In: Proceedings of the Second International Pump Symposium; April 1985; Houston, Texas. 1985. p. 77–84.

[23] Ghosh M., Majumdar B., Sarangi M. Theory of Lubrication. New Delhi: Tata McGraw-Hill; 2013.

[24] DellaCorte C., Lukaszewicz V., Valco M.J., Radil K.C., Heshmat H. Performance and Durability of High Temperature Foil Air Bearings for Oil-Free Turbomachinery. Tribology Transactions. 2000;**43**:774–780.

[25] Venkatesh V. Precision Engineering. New Delhi: Tata McGraw-Hill Education; 2007.

[26] Hamrock B.J., Anderson W.J. Rolling-Element Bearings. National Aeronautics and Space Administration Reference Publication. 1983;NASA RP-1105:1–57. Available from: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19830018943.pdf [Accessed: 2017-01-10].

[27] Pope J.E. Rules of Thumb for Mechanical Engineers. Houston, Texas: Gulf Publishing Company; 1997.

[28] Tanimoto K., Ikeda T. Evaluation of Extra-Small Ceramic Ball Bearings. KOYO Engineering Journal English Edition. 2000;**156**(E):23–29.

[29] Powell J. Design of Aerostatic Bearings. Brighton, U.K.: The Machinery Publishing Co.; 1970.

[30] DellaCorte C., Valco M.J. Load Capacity Estimation of Foil Air Journal Bearings for Oil-Free Turbomachinery Applications. STLE Tribology Transactions. 2000;**43**(4):795–801.

[31] Radil K.C., DellaCorte C. Foil Bearing Starting Considerations and Requirements for Rotorcraft Engine Applications. Army Research Laboratory Final Report. 2009;ARL-TR-4873:1–12. Available from: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20120012857.pdf [Accessed: 2017-01-10]

[32] DellaCorte C., Radil K. A Preliminary Foil Gas Bearing Performance Map. National Aeronautics and Space Administration Technical Memorandum. 2006;NASA/TM-2006-214343:1–15. Available from: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20060051824.pdf [Accessed: 2017-01-10]

[33] Cunningham R.E., Fleming D.P., Anderson W.J. Experiments of Stability of Herringbone-Grooved Gas Lubricated Journal Bearings to High Compressibility Numbers. National Aeronautics and Space Administration Technical Report. 1968;NASA TN-D-4440:1–20. Available from: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19680012496.pdf [Accessed: 2017-01-10]

[34] Reitsma T.W. Lifting Off Friction [Internet]. 1 September 2006.Available from: http://machinedesign.com/archive/lifting-friction.

[35] Montague G., Jansen M., Provenza A., Jansen R., Ebihara B., Palazzolo A. Room Temperature Characterization of a Magnetic Bearing for Turbomachinery. National Aeronautics and Space Administration Technical Memorandum. 2002;NASA/TM–2002-211904:1–9.

Available from: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20020090900.pdf [Accessed: 2017-01-10]

[36] Montague G., Jansen M., Provenza A., Palazzolo A., Jansen R., Ebihara B. Experimental High Temperature Characterization of a Magnetic Bearing for Turbomachinery. National Aeronautics and Space Administration Technical Memorandum. 2003;NASA/TM–2003-212183:1–12. Available from: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20030039177.pdf [Accessed: 2017-01-10]

[37] Paden B., Groom N., Antaki J.F. Design Formulas for Permanent-Magnet Bearings. ASME Journal of Mechanical Design. 2003;**125**:734–738.

Chapter 2

Electromagnetic Levitation System for Active Magnetic Bearing Wheels

Yonmook Park

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/67227

Abstract

In this chapter, the author presents an electromagnetic levitation system for active magnetic bearing wheels. This system consists of a rotor, a shaft, a cover, and a base. The author derives a meaningful electromagnetic force by using the singular value decomposition. The author develops a control system using the proportional-integral-derivative controller to control the position of the rotor and regulate the two gimbal angles of the rotor. The author gives the numerical simulation and experimental results on the control of the electromagnetic levitation system.

Keywords: active magnetic bearing, electromagnetic levitation system, motion control

1. Introduction

As a reaction wheel in spacecraft, a ball bearing wheel, a magnetic bearing wheel, and an active magnetic bearing wheel have mainly been used. First, a ball bearing wheel uses a ball bearing to maintain the separation between the bearing races. Ball bearings reduce rotational friction and support radial and axial loads by using at least two races to contain balls and transmit the loads through balls. Ball bearings tend to have a lower load capacity than other kinds of rolling element bearings mainly due to the small contact area between balls and races. Also, ball bearings should be lubricated periodically with a lubricant such as oil and grease for ball bearings to operate properly [1]. Next, a magnetic bearing is used in a magnetic bearing wheel. A magnetic bearing supports a load by the magnetic levitation principle. In magnetic bearing wheels, permanent magnets are used to carry a wheel, a control system is used to hold a wheel stable, and power is used when a levitated wheel deviates from its target position. A magnetic bearing wheel also requires a back-up bearing in case of control system or power failure and

during initial start-up conditions. A magnetic bearing has two kinds of instabilities. One is that attractive magnets provide an unstable static force that decreases at distant distances and increases at close distances. The other is that magnetism gives rise to oscillations that may cause loss of suspension if driving forces are present [1]. Finally, in an active magnetic bearing wheel, a rotating shaft is levitated by the principle of electromagnetic suspension. A wheel is supported in an active magnetic bearing wheel without physical contact. The contactless operation of the active magnetic bearing wheels eliminates the need of lubrication of the bearing components, which allows them to operate cleanly. Moreover, it can accommodate irregularities in the mass distribution automatically, which allows it to spin around its center of mass with very low vibration, and can suppress the nutation and precession of the rotor effectively. The components of an active magnetic bearing wheel are an active magnetic bearing, a wheel, a control system, an electromagnet assembly, power amplifiers, and gap sensors. This bias current is mediated by a control system that offsets the bias current by equal but opposite perturbations of current as the rotor deviates by a small amount from its center position [1].

The active magnetic bearing wheel exhibits very lower vibration than ball bearing wheels and magnetic bearing wheels. Thus, it is a desirable reaction wheel for the spacecraft attitude control since vibration is the critical factor for the high precision spacecraft attitude control. The active magnetic bearing is the very important component among components of active magnetic bearing wheels. Due to this importance, various kinds of active magnetic bearings have been developed and their control methods have been studied (e.g., [2–10]).

In this chapter, the author presents an electromagnetic levitation system for active magnetic bearing wheels. This system consists of a rotor, a shaft, a cover, and a base. Also, this system does not include a mechanism for spinning the rotor around its rotating axis. The author derives a meaningful electromagnetic force by using the singular value decomposition [11]. The proportional-integral-derivative (PID) controller is used to control the position of the rotor and regulate the two gimbal angles of the rotor. The author gives the numerical simulation and experimental results on the control of the electromagnetic levitation system.

2. Electromagnetic levitation system

In **Figure 1**, the schematic of the electromagnetic levitation system developed in this chapter is given. The cover protects the rotor, and the base supports the rotor, shaft, and cover. This system can levitate the rotor up to 0.8 mm from the ground in the z-axis, rotate the rotor, and gimbal the rotor within a small angle of ±0.2°.

The dynamic equations of motion of the electromagnetic levitation system are given as follows:

$$m\ddot{z} = F_z \tag{1}$$

$$I\ddot{\phi} = T_x \tag{2}$$

$$I\ddot{\theta} = T_y \tag{3}$$

http://dx.doi.org/10.5772/67227

In Eqs. (1)–(3), ϕ and θ are the gimbal angles of the rotor in the x- and y-axes, respectively, z is the displacement of the rotor in the z-axis, F_z is the control force in the z-axis, T_x and T_y are the control torques applied to the rotor in the x- and y-axes, respectively, m = 0.72 kg is the mass of the rotor, and $I = 877.367 \times 10^{-6}$ kg m^2 is the inertia of the rotor for the x- and y-axes.

Let us consider the four pairs of electromagnets shown in **Figure 1**. Then, the control inputs F_z in Eq. (1), T_x in Eq. (2), and T_y in Eq. (3) can be represented as follows:

$$\sum_{i=1}^{4} F_{ei} = F_z + mg \tag{4}$$

$$D_e\left(F_{e2} - F_{e4}\right) = T_x \tag{5}$$

$$D_e\left(F_{e3} - F_{e1}\right) = T_y \tag{6}$$

In Eqs. (4)–(6), F_{ei} is the electromagnetic force generated by the ith pair of electromagnets, and g = 9.8 m/s^2 is the acceleration of gravity. Then, Eqs. (4)–(6) can be written as follows:

$$A F_e = u + u_g \tag{7}$$

where

$$A \triangleq \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & D_e & 0 & -D_e \\ -D_e & 0 & D_e & 0 \end{bmatrix} \tag{8}$$

$F_e \triangleq [F_{e1}\ F_{e2}\ F_{e3}\ F_{e4}]^T$, $u \triangleq [F_z\ T_x\ T_y]^T$, and $u_g \triangleq [mg\ 0\ 0]^T$. After designing the control inputs F_z, T_x, and T_y, the four electromagnetic forces F_{ei}, i = 1,…,4 have to be determined by Eq. (7). Among solutions for Eq. (7), the minimal norm solution is derived by using the singular value decomposition [11]. Let the singular value decomposition of the matrix $A \in R^{3\times4}$ in Eq. (8) be $U\Sigma V^T$ and define

$$A^+ \triangleq V\Sigma^+ U^T \tag{9}$$

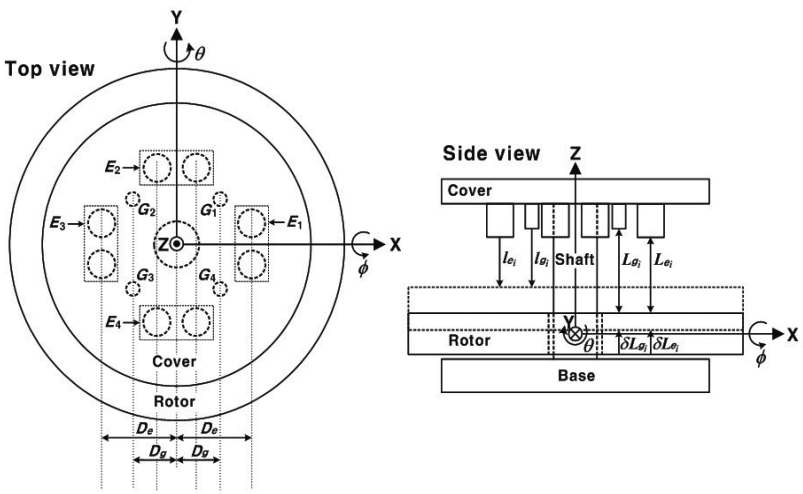


Figure 1. Schematic representation of the electromagnetic levitation system.

where $U \in R^{3\times3}$ and $V \in R^{4\times4}$ are orthogonal matrices, $\Sigma \in R^{3\times4}$, and $A^{+} \in R^{4\times3}$ and $\Sigma^{+} \in R^{4\times3}$ denote the pseudoinverse matrices of the matrices A and Σ, respectively. With some calculations, we obtain the following by the singular value decomposition of the matrix A in Eq. (8)

$$V = \frac{1}{2}\begin{bmatrix} 1 & 0 & -\sqrt{2} & -1 \\ 1 & \sqrt{2} & 0 & 1 \\ 1 & 0 & \sqrt{2} & -1 \\ 1 & -\sqrt{2} & 0 & 1 \end{bmatrix} \tag{10}$$

$$\Sigma^{+} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{D_e} & 0 \\ 0 & 0 & \frac{\sqrt{2}}{D_e} \\ 0 & 0 & 0 \end{bmatrix} \triangleq \begin{bmatrix} \Sigma_1^{-1} \\ 0_{1\times3} \end{bmatrix} \tag{11}$$

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{12}$$

In Eq. (11), $0_{1\times3}$ implies the 1 × 3 zero matrix. Then, with A^{+} of Eq. (9), we obtain

$$F_e = A^{+}\left(u + u_g\right) = \begin{bmatrix} \frac{1}{4}\left(F_z + mg\right) - \frac{1}{2}\frac{T_y}{D_e} \\ \frac{1}{4}\left(F_z + mg\right) + \frac{1}{2}\frac{T_x}{D_e} \\ \frac{1}{4}\left(F_z + mg\right) + \frac{1}{2}\frac{T_y}{D_e} \\ \frac{1}{4}\left(F_z + mg\right) - \frac{1}{2}\frac{T_x}{D_e} \end{bmatrix} \tag{13}$$

The following condition holds for any other solution $\hat{F}_e$ to Eq. (7) [12]

$$\|F_e\|_2 < \|\hat{F}_e\|_2 \tag{14}$$

where '$\|\cdot\|_2$' denotes the Euclidean norm of a vector (i.e., for a vector $x \in R^{n}$, $\|x\|_2 \triangleq \sqrt{\sum_{i=1}^{n} x_i^2}$).

Since the two gimbal angles, ϕ and θ, are very small, we can approximate $\sin(\phi) \cong \phi$ and $\sin(\theta) \cong \theta$ by the small-angle approximation. Thus, the displacement from the bottom surface of the ith gap sensor to the top surface of the rotor in the z-axis can be calculated as follows:

$$l_{g1} = L_{g1} + \delta L_{g1} = L_{g1} - z + D_g \sin(\theta) - D_g \sin(\phi) \cong L_{g1} - z + D_g\theta - D_g\phi \tag{15}$$

$$l_{g2} = L_{g2} + \delta L_{g2} = L_{g2} - z - D_g \sin(\theta) - D_g \sin(\phi) \cong L_{g2} - z - D_g\theta - D_g\phi \tag{16}$$

$$l_{g3} = L_{g3} + \delta L_{g3} = L_{g3} - z - D_g \sin(\theta) + D_g \sin(\phi) \cong L_{g2} - z - D_g\theta + D_g\phi \tag{17}$$

$$l_{g4} = L_{g4} + \delta L_{g4} = L_{g4} - z + D_g \sin(\theta) + D_g \sin(\phi) \cong L_{g4} - z + D_g\theta + D_g\phi \tag{18}$$

Then, from Eqs. (15) to (18), the system state z, ϕ, and θ can be calculated as follows:

$$z = \frac{1}{4}\left(\sum_{i=1}^{4} L_{gi} - \sum_{i=1}^{4} l_{gi}\right) \tag{19}$$

$$\phi = \frac{1}{4D_g}\left[\left(l_{g3} + l_{g4} - l_{g1} - l_{g2}\right) - \left(L_{g3} + L_{g4} - L_{g1} - L_{g2}\right)\right] \tag{20}$$

http://dx.doi.org/10.5772/67227

$$\theta = \frac{1}{4D_g}\left[\left(l_{g1} + l_{g4} - l_{g2} - l_{g3}\right) - \left(L_{g1} + L_{g4} - L_{g2} - L_{g3}\right)\right] \tag{21}$$

Similarly, the displacement from the bottom surface of the ith pair of electromagnets to the top surface of the rotor in the z-axis can be calculated as follows:

$$l_{e1} = L_{e1} + \delta L_{e1} = L_{e1} - z + D_e \sin(\theta) \cong L_{e1} - z + D_e\theta \tag{22}$$

$$l_{e2} = L_{e2} + \delta L_{e2} = L_{e2} - z - D_e \sin(\phi) \cong L_{e2} - z - D_e\phi \tag{23}$$

$$l_{e3} = L_{e3} + \delta L_{e3} = L_{e3} - z - D_e \sin(\theta) \cong L_{e3} - z - D_e\theta \tag{24}$$

$$l_{e4} = L_{e4} + \delta L_{e4} = L_{e4} - z + D_e \sin(\phi) \cong L_{e4} - z + D_e\phi \tag{25}$$

By the Maxwell's equation [13], the following equation is obtained for the control currents supplied to the coils of the four pairs of electromagnets

$$i_i = \frac{2l_{ei}}{n}\sqrt{\frac{F_{ei}}{\mu_0 G}},\ i = 1, \ldots, 4 \tag{26}$$

where $\mu_0 = 4\pi \times 10^{-7}$ N/A^2 is the permeability constant of free space, $n = 240$ is the number of coil turn, i_i is the control current of the ith pair of electromagnets, and $G = 50.265 \times 10^{-6}$ m^2 is the cross-sectional area of a pair of electromagnets. The author limits each control current in Eq. (26) by 1 A.

3. Numerical simulation and experimental results

The author gives the numerical simulation and experimental results on the control of the electromagnetic levitation system in this section.

The author uses the following discretized PID controller to control the displacement of the rotor in Eq. (19) and the two gimbal angles of the rotor in Eqs. (20) and (21).

$$u(t) = K_p x_e(t) + K_i T_s \sum_{i=0}^{t} x_e(i) + K_d \frac{\left(x_e(t) - x_e(t-1)\right)}{T_s} \tag{27}$$

where T_s is the sampling time and given by $T_s = 1$ ms, $x_e \triangleq x_t - x$ is the error between the system state $x \triangleq [z\ \phi\ \theta]^T$ and the target system state $x_t \triangleq [z_t\ \phi_t\ \theta_t]^T$ and $K_p \triangleq \text{diag}[K_{pz}, K_{p\phi}, K_{p\theta}]$, $K_i \triangleq \text{diag}[K_{iz}, K_{i\phi}, K_{i\theta}]$, and $K_d \triangleq \text{diag}[K_{dz}, K_{d\phi}, K_{d\theta}]$ denote the 3 × 3 diagonal positive definite matrices.

The target position and target gimbal angles of the rotor are set to be $z_t = 0.3$ mm and $\phi_t = 0°$ and $\theta_t = 0°$, respectively. The author initially decides the feedback gains of the PID controller in Eq. (27) that can achieve the control objective by adopting the well-known Ziegler-Nichols method [14] and then finely tunes the feedback gains of the PID controller in Eq. (27) by an experiment. Also, the antiwindup compensator is used to make the overshot as small as possible. According to the Ziegler-Nichols method [14], first, $K_i = K_d = \text{diag}[0, 0, 0]$ are set, and the proportional gain K_p is then increased until the system just oscillates. The proportional gain is then multiplied by 0.6, and the integral and derivative gains are calculated as $K_p = 0.6\ K_m$, $K_i = K_p(\omega_m/\pi)$, and $K_d = K_p(0.25\ \pi/\omega_m)$ where $K_m \triangleq \text{diag}[K_{mz}, K_{m\phi}, K_{m\theta}]$ denotes the 3 × 3 diagonal positive definite matrix with the gain elements at which the proportional system oscillates,

and $\omega_m \triangleq \text{diag}[\omega_{mz}, \omega_{m\phi}, \omega_{m\theta}]$ denotes the 3 × 3 diagonal positive definite matrix with the oscillation frequency elements. As a result, the feedback gains of the PID controller in Eq. (27) are chosen as follows: $K_{pz} = 1000$, $K_{p\phi} = K_{p\theta} = 20$, $K_{iz} = 5000$, $K_{i\phi} = K_{i\theta} = 8.33$, $K_{dz} = 100$, and $K_{d\phi} = K_{d\theta} = 0.013$. Consequently, we see that the dominant feedback gains in this PID controller are K_p and K_i, and thus, one can obtain $K_{mz} \cong 1666.67$, $K_{m\phi} = K_{m\theta} \cong 33.33$, $\omega_m \cong 15.71$ rad/s and $\omega_{m\phi} = \omega_{m\theta} \cong 1.31$ rad/s.

The control flow diagram of the system is shown in **Figure 2**. After we measure the displacements from the bottom surfaces of the four gap sensors to the top surface of the rotor, we calculate the displacement of the rotor and the two gimbal angles of the rotor by Eqs. (19)–(21), respectively. The control input $u \triangleq [F_z \; T_x \; T_y]^T$ is made by the PID controller in Eq. (27). Then, we calculate the four electromagnetic forces and the displacements from the bottom surfaces of the four pairs of electromagnets to the top surface of the rotor by Eq. (13) and the Eq. (22)–(25), respectively. After we calculate the control currents by Eq. (26), they pass through the current limiters and are supplied to the coils of the four pairs of electromagnets by the power electronics.

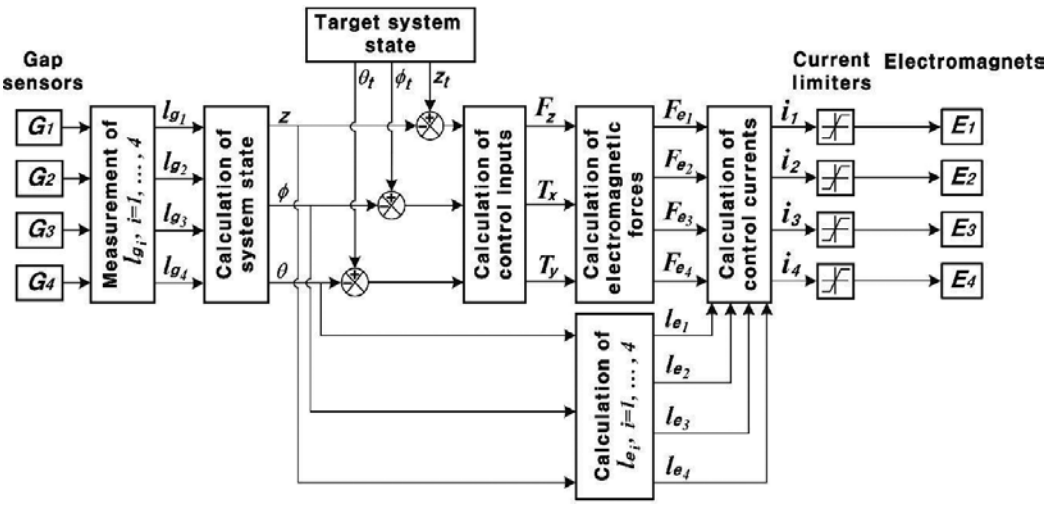


Figure 2. Control flow diagram of the electromagnetic levitation system.

In the numerical simulation, it will be demonstrated that the electromagnetic force F_e of Eq. (13) satisfies the condition of Eq. (14) with respect to the following electromagnetic force $\hat{F}_e$, which is another solution to Eq. (7)

$$\hat{F}_e = \begin{bmatrix} mg \\ \frac{1}{2}F_z + \frac{1}{2}\frac{(T_x - T_y)}{D_e} - \frac{1}{2}mg \\ \frac{T_y}{D_e} + mg \\ \frac{1}{2}F_z - \frac{1}{2}\frac{(T_x + T_y)}{D_e} - \frac{1}{2}mg \end{bmatrix} \triangleq \begin{bmatrix} \hat{F}_{e1} & \hat{F}_{e2} & \hat{F}_{e3} & \hat{F}_{e4} \end{bmatrix}^T \tag{28}$$

With the initial system state given by $z_{ini} = 0$ mm, $\dot{z}_{ini} = 0$ mm/s, the values are $\phi_{ini} = \theta_{ini} = 0°$ and $\dot{\phi}_{ini} = \dot{\theta}_{ini} = 0°/s$. In **Figures 3–5**, the numerical simulation results on the control of the electromagnetic levitation system using F_e of Eq. (13) and $\hat{F}_e$ of Eq. (28) are shown. As shown in **Figure 3**, the rotor reaches the target position $z_t = 0.3$ mm, and the control forces using F_e of Eq. (13) and $\hat{F}_e$ of Eq. (28) show the same behaviors, respectively. In **Figure 4**, we see that each control current using F_e of Eq. (13) reaches the same value, the control currents of i_1 and i_3 using $\hat{F}_{e1}$ and $\hat{F}_{e3}$ of Eq. (28), respectively, reach the same value, and the control currents of i_2 and i_4 using $\hat{F}_{e2}$ and $\hat{F}_{e4}$ of Eq. (28), respectively, reach the same value. In **Figure 5**, we see that the Euclidean norm of F_e of Eq. (13) is about 3.162 times smaller than that of $\hat{F}_e$ of Eq. (28) in the steady-state region. Therefore, the numerical simulation results shown in **Figure 5** illustrate that F_e of Eq. (13) satisfies the condition of Eq. (14) with respect to $\hat{F}_e$ of Eq. (28).

The experimental results on the control of the system using F_{ei}, $i = 1,\ldots,4$ of Eq. (13) are shown in **Figures 6–9**. As shown in **Figures 6** and **7**, the PID controller successfully levitates the rotor at the target position $z_t = 0.3$ mm with well regulating the two gimbal angles. The trajectories of control currents are shown in **Figure 8**. And the Euclidean norm of F_e of Eq. (13) is shown in **Figure 9**. As shown in **Figures 4**, **5**, **8**, and **9**, the trajectories of control currents and Euclidean norm of F_e of Eq. (13) obtained by the experiment move around the values obtained by the numerical simulation.

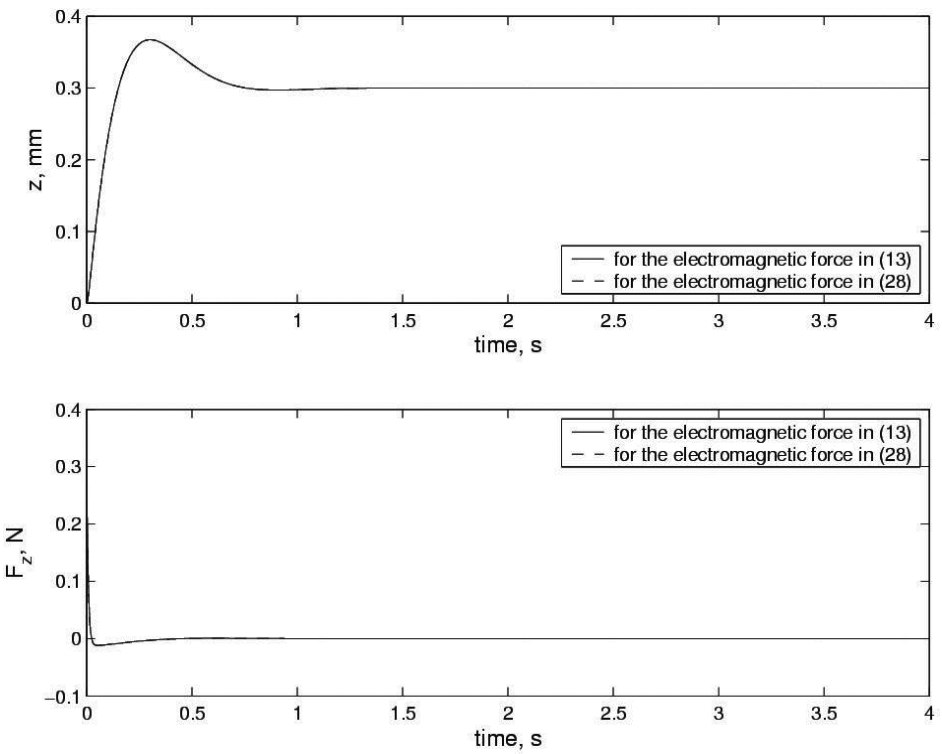


Figure 3. Time histories of the position of the rotor and the control force, which are obtained by the numerical simulation.

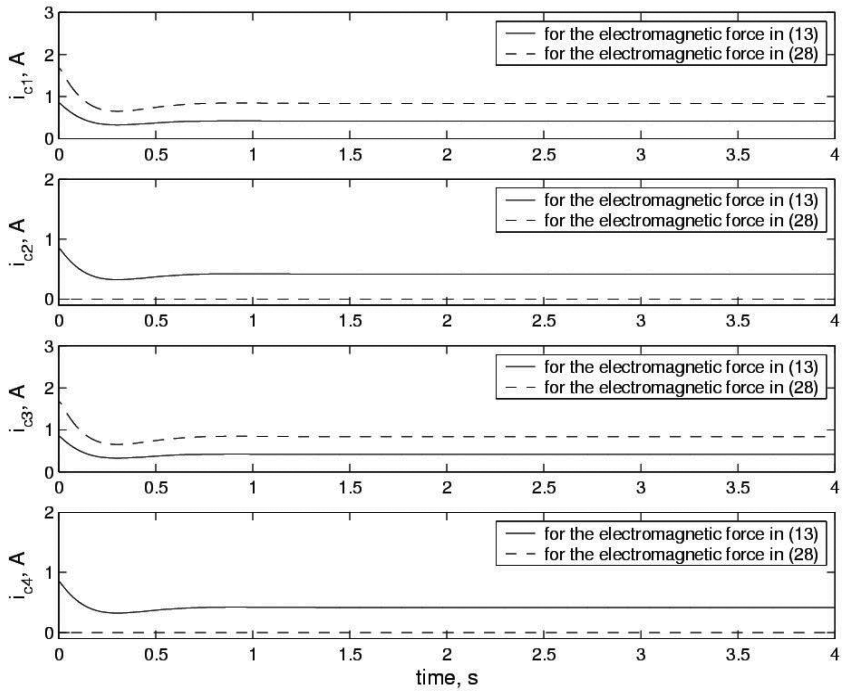


Figure 4. Time histories of the control currents, which are obtained by the numerical simulation.

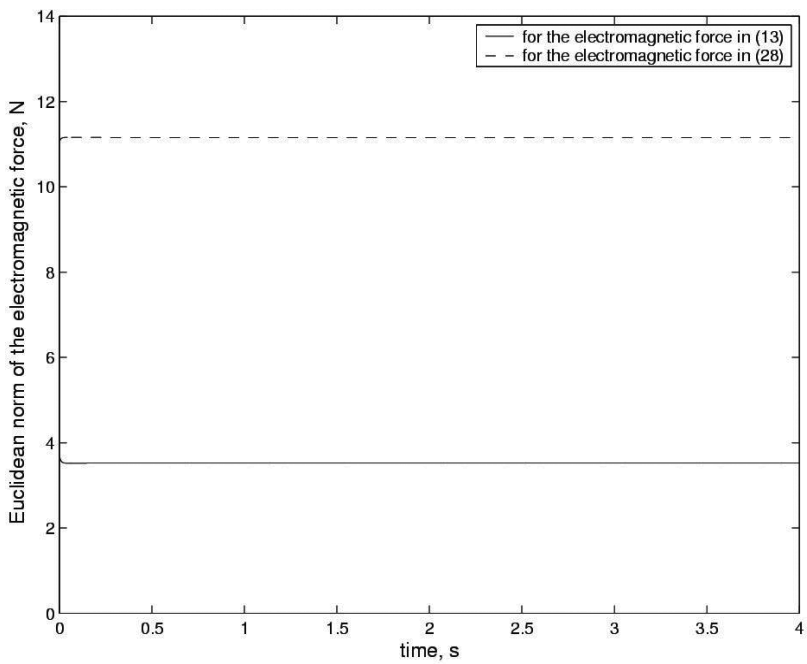


Figure 5. Time histories of the Euclidean norm of the electromagnetic force, which are obtained by the numerical simulation.

http://dx.doi.org/10.5772/67227

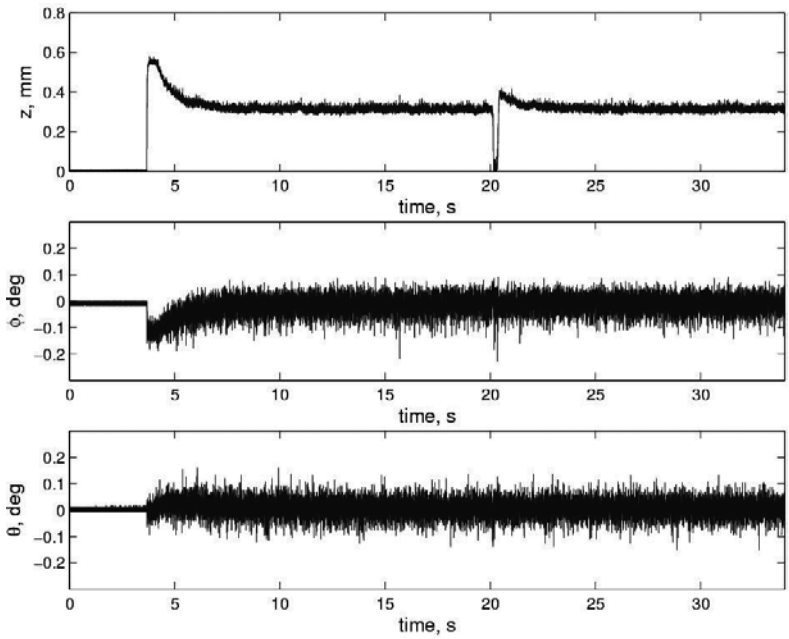


Figure 6. Time histories of the system state of the electromagnetic levitation system using Fe of Eq. (13), which are obtained by the experiment.

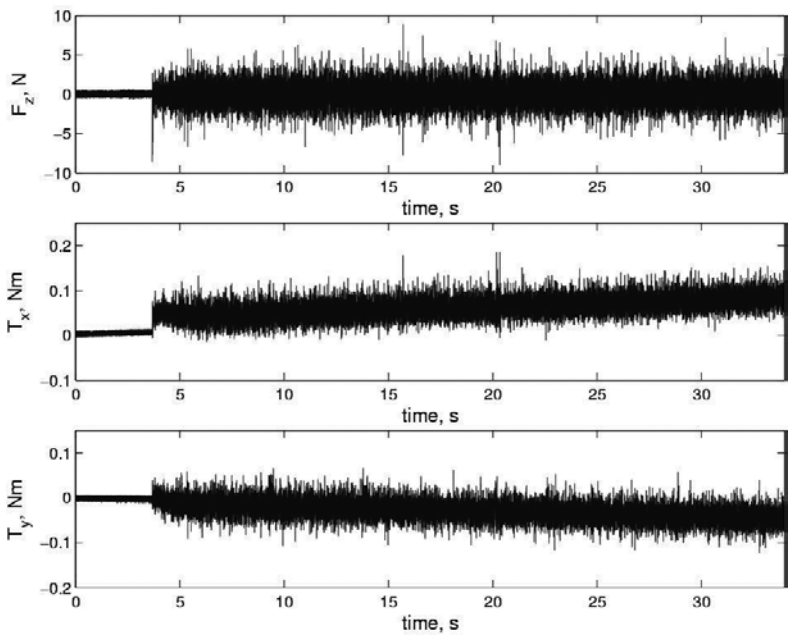


Figure 7. Time histories of the control inputs of the electromagnetic levitation system using Fe of Eq. (13), which are obtained by the experiment.

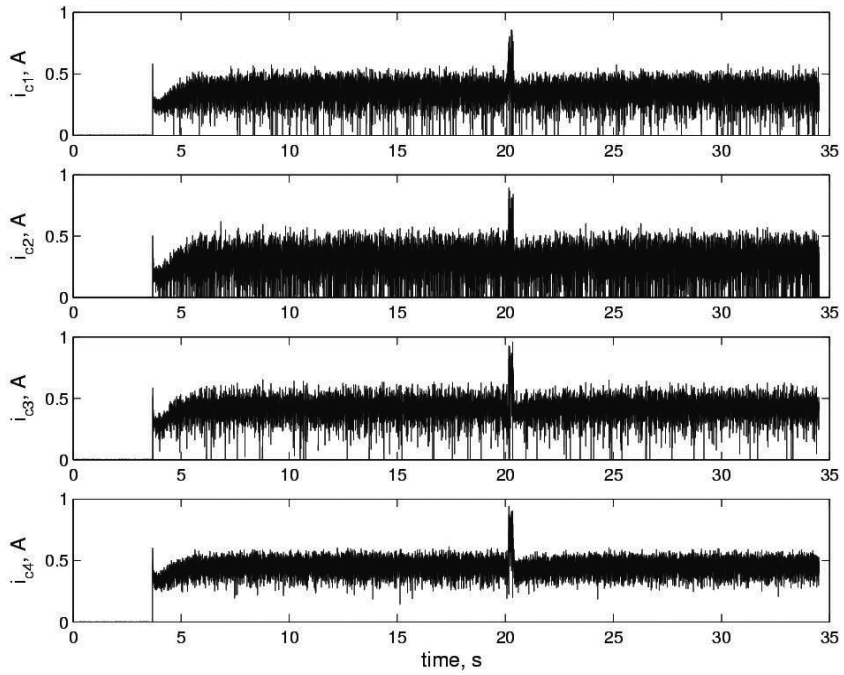


Figure 8. Time histories of the control currents of the electromagnetic levitation system using Fe of Eq. (13), which are obtained by the experiment.

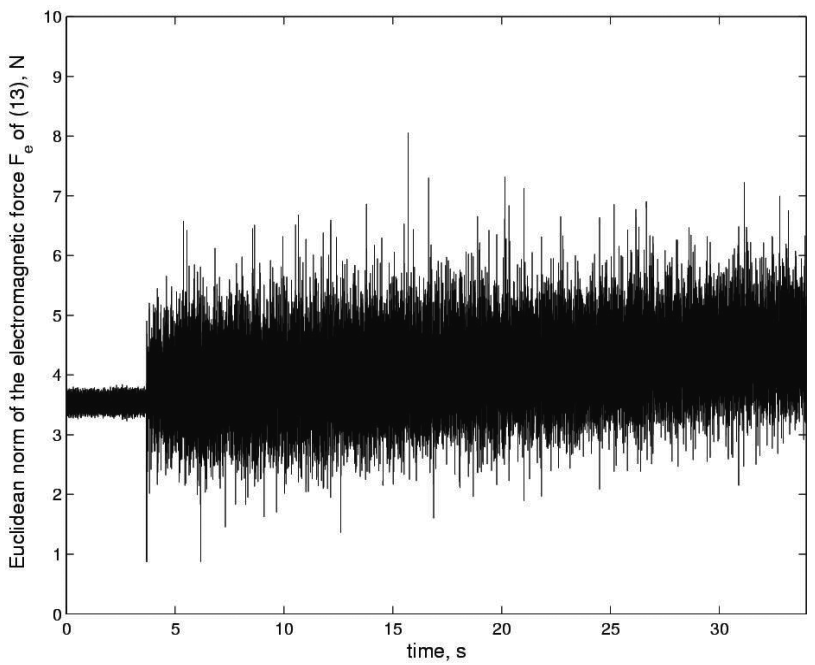


Figure 9. Time histories of the Euclidean norm of Fe of Eq. (13), which are obtained by the experiment.

In order to simulate an external disturbance, a human hand presses down hard on the rotor to the ground at about 20 s, and it is removed from the rotor momentarily. In **Figure 6**, we see that, after we remove the external disturbance from the rotor, the system state becomes a steady state within 2 s. Also, in **Figure 8**, we see that, as the external disturbance applied to the rotor increases, each control current increases to resist the external disturbance. It should be remarked that the operating parameters like an applied load and a speed of the rotor may influence on the design of electromagnetic levitation system because these parameters make an impact on the system dynamics.

4. Conclusion

In this chapter, the electromagnetic levitation system was developed as a prototype for developing active magnetic bearing wheels. A control system was developed to control the position and two gimbal angles of the rotor. The experimental results demonstrated that the control system can control the position of the rotor and regulate the two gimbal angles. The refinement of the electromagnetic levitation system for the development of active magnetic bearing wheels is the further research topic.

Acknowledgements

The author would like to thank the Satellite Technology Research Center at the Korea Advanced Institute of Science and Technology for its support to develop the electromagnetic levitation system.

Author details

Yonmook Park

Address all correspondence to: ym-park@kaist.ac.kr

Smart Machine Center, MINDs Lab. Co. Ltd., Bundang-gu, Seongnam-si, Gyeonggi-do, Republic of Korea

References

[1] Schweitzer G., Maslen E.H., editors. Magnetic bearings. Berlin, Germany: Springer; 2009. pp. 1–535.

[2] Trumper D.L., Olson S.M., Subrahmanyan P.K. Linearizing control of magnetic suspension systems. IEEE Trans. Control Syst. Technol. 1997;**5**(4):427–438.

[3] Hong S.-K., Langari R. Robust fuzzy control of a magnetic bearing system subject to harmonic disturbances. IEEE Trans. Control Syst. Technol. 2000;**8**(2):366–371.

[4] Lindlau J.D., Knospe C.R. Feedback linearization of an active magnetic bearing with voltage control. IEEE Trans. Control Syst. Technol. 2002;**10**(1):21–31.

[5] Mazenc F., de Queiroz M.S., Malisoff M., Gao F. Further results on active magnetic bearing control with input saturation. IEEE Trans. Control Syst. Technol. 2006;**14**(5):914–919.

[6] Sivrioglu S. Adaptive backstepping for switching control active magnetic bearing system with vibrating base. IET Control Theory Appl. 2007;**1**(4):1054–1059.

[7] Jugo J., Lizarraga I., Arredondo I. Nonlinear modelling and analysis of active magnetic bearing systems in the harmonic domain: a case study. IET Control Theory Appl. 2008;**2**(1):61–71.

[8] Tang J., Fang J., Wen W. Superconducting magnetic bearings and active magnetic bearings in attitude control and energy storage flywhee for spacecraft. IEEE Trans. Appl. Supercond. 2012;**22**(6):5702109.

[9] Schuhmann T., Hofmann W., Werner R. Improving operational performance of active magnetic bearings using kalman filter and state feedback control. IEEE Trans. Ind. Electron. 2012;**59**(2):821–829.

[10] Fang J., Zheng S., Han B. Attitude sensing and dynamic decoupling based on active magnetic bearing of MSDGCMG. IEEE Trans. Instrum. Meas.. 2012;**61**(2):338–348.

[11] Strang G. Linear Algebra and Its Applications. 3rd ed. Florida, USA: Harcourt Brace Jovanovich Publishers; 1988. pp. 1–505.

[12] Golub G.H., Van Loan C.F. Matrix computations. 3rd ed. Maryland, USA: The Johns Hopkins University Press; 1996. pp. 1–728.

[13] Maxwell J.C. A dynamical theory of the electromagnetic field. Phil. Trans. R. Soc. Lond. 1865;**155**:459–512.

[14] Ogata K. Modern control engineering. 5th ed. New Jersey, USA: Prentice-Hall; 1996. pp. 1–894.

Chapter 3

Condition Monitoring and Fault Diagnosis of Roller Element Bearing

Tian Ran Lin, Kun Yu and Jiwen Tan

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/67143

Abstract

Rolling element bearings play a crucial role in determining the overall health condition of a rotating machine. An effective condition-monitoring program on bearing operation can improve a machine's operation efficiency, reduce the maintenance/replacement cost, and prolong the useful lifespan of a machine. This chapter presents a general overview of various condition-monitoring and fault diagnosis techniques for rolling element bearings in the current practice and discusses the pros and cons of each technique. The techniques introduced in the chapter include data acquisition techniques, major parameters used for bearing condition monitoring, signal analysis techniques, and bearing fault diagnosis techniques using either statistical features or artificial intelligent tools. Several case studies are also presented in the chapter to exemplify the application of these techniques in the data analysis as well as bearing fault diagnosis and pattern recognition.

Keywords: rolling element bearings, condition monitoring, fault diagnosis

1. Introduction

Rolling element bearings are the most critical but vulnerable mechanical components in a rotating machine. A bearing failure can lead to a complete machine breakdown causing unintended interruption to a production process and financial losses. It is important to have an effective bearing condition monitoring (CM) and fault diagnosis system in place so that incipient bearing faults can be detected and correctly diagnosed on time to prevent them from deteriorating further to cause damage to a machine. For instance, an early detection of incipient defect of a rolling element bearing in a high speed train or a wind turbine can lead to a timely maintenance/replacement to prevent potential disastrous consequence and human loss caused by unexpected failure of critical mechanical components.

Many condition-monitoring and fault diagnosis techniques have been developed in the last few decades to improve the reliability of rolling element bearings. This chapter provides an overview on the most commonly employed condition-monitoring, signal analysis, and fault diagnosis techniques for rolling element bearings and discusses some of the pros and cons of these techniques.

A starting but most fundamental information in bearing condition monitoring is the characteristic bearing defect frequencies. The characteristic defect frequency components in a signal are generated by flaws or faults presented in a bearing when the bearing is operated at a specific machine rotating speed under certain loading conditions. Alternatively, defect signals can also be produced accompanying the normal wear process during a bearing's operational life.

Figure 1 shows the graphical and the cross-sectional representations of a rolling element bearing. The bearing comprises four mechanical components: an outer race, an inner race, rollers (balls), and a cage that hold the rollers (balls) in place. Correspondingly, there are four possible characteristic defect frequencies for a rolling element bearing: ball (roller) pass frequency at the outer race (BPFO), ball (roller) pass frequency at the inner race (BPFI), ball (roller) spin frequency (BSF), and fundamental train frequency (cage frequency) (FTF). The formulae for these four characteristic bearing defect frequencies are listed in **Table 1**.

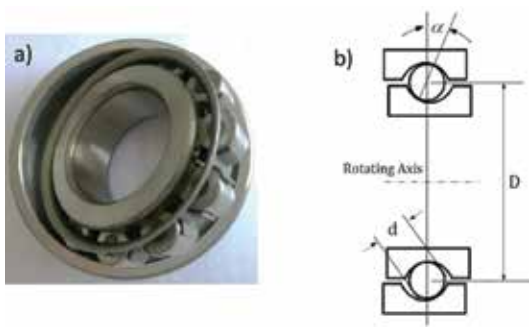


Figure 1. (a) A graphical illustration of a roller element bearing and (b) a cross-sectional view of the roller element bearing.

Ball-pass frequency at outer race (BPFO)	$\mathrm{BPFO} = \frac{n}{2}\frac{N}{60}\left(1-\frac{d}{D}\cos\alpha\right)$
Ball-pass frequency at inner race (BPFI)	$\mathrm{BPFI} = \frac{n}{2}\frac{N}{60}\left(1+\frac{d}{D}\cos\alpha\right)$
Ball-spin frequency (BSF)	$\mathrm{BSF} = \frac{D}{2d}\left[1-\left(\frac{d}{D}\cos\alpha\right)^2\right]$
Fundamental train frequency (FTF)	$\mathrm{FTF} = \frac{1}{2}\frac{N}{60}\left(1-\frac{d}{D}\cos\alpha\right)$

Note: N is the shaft speed in revolutions per minute (RPM), n is the number of roller elements in a bearing, α is the contact angle of the bearing due to the load from the radial plane, d is the diameter of the roller, and D is the mean diameter of the bearing as shown in **Figure 1**.

Table 1. Formulae of the bearing defect frequencies.

http://dx.doi.org/10.5772/67143

A bearing defect signal can be simulated using the following equations [1]:

$$s(t,n) = Qe^{-\gamma(t-\frac{n}{\mathrm{BDF}})}\sin\left[2\pi f_r\left(t-\frac{n}{\mathrm{BDF}}\right)\right] + \frac{O(t)}{r_{\mathrm{sn}}},\ t < \frac{n+1}{\mathrm{BDF}},\ n = 0,1,2,\ldots. \tag{1a}$$

and

$$s(t,n) = Qe^{-\gamma(t-\frac{n+1}{\mathrm{BDF}})}\sin\left[2\pi f_r\left(t-\frac{n+1}{\mathrm{BDF}}\right)\right] + \frac{O(t)}{r_{\mathrm{sn}}},\ t \geq \frac{n+1}{\mathrm{BDF}},\ n = 0,1,2,\ldots. \tag{1b}$$

where Q is the assumed maximum loading intensity for a bearing defect and t is the time variable, BDF represents a bearing defect frequency, f_r is the assumed bearing resonance frequency and α is the energy decay constant of the bearing race. The first part in Eqs. (1a) and (1b) is the signal produced by a bearing defect, and the second part of the equations is the superimposed white Gaussian noise representing the machine background noise. n is the pulse index of the bearing defect frequency, $O(t)$ is the white Gaussian noise and r_{sn} is the assumed signal-to-noise ratio (SNR). A typical bearing defect signal is shown in **Figure 2**. The parameters used in the simulation of the signal are listed in **Table 2**.

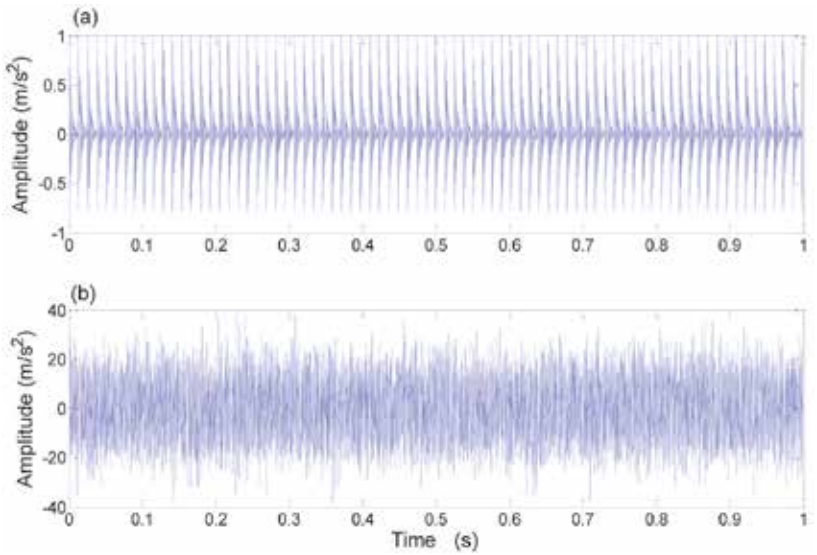


Figure 2. A simulated defect signal of a roller element bearing due to a fault at the outer race: (a) pure defect signal; (b) noise-added signal.

Fault type	Defect frequency (Hz)	Resonance frequency, f_r (Hz)	Decay constant, α	Loading intensity, Q	SNR, r_{sn}
BPFO	78.18	4000	200	1	−20 dB

Table 2. Parameters used in the simulation.

2. Condition-monitoring techniques

Condition-monitoring (CM) techniques are required to acquire the operation condition data of a rotating machine. The CM data can then be processed and analyzed using appropriate signal analysis techniques to obtain the most relevant characteristic parameters before being used in a diagnosis or prognosis algorithm to evaluate/predict the health state of the machine. The data employed in a condition-monitoring program can be vibration, noise, electric current, oil and grease, temperature or a combination of these data. The CM data can be analyzed using either time-domain techniques, frequency-domain techniques, or time-frequency techniques according to the data properties such as linearity or nonlinearity, stationary or nonstationary, to extract the most useful and effective features for bearing fault diagnosis. This section summarizes some of the most commonly employed techniques in condition-monitoring applications of rolling element bearings.

2.1. Vibration technique

Vibration technique is the most frequently employed technique for machine condition monitoring. A change of the vibration signal in a machine without changing the operation condition can imply a change of health state of the machine. Vibration signals are typically generated by defects in the moving components of the machine such as defects in a bearing, gearboxes, reciprocating components, and so on. For example, when a rolling element bearing operates under a faulty condition, an impulse signal will be generated as other bearing components passing through the faulty position. This will lead to an increase in overall vibration amplitude of the machine. The defect component of the bearing and its severity can be determined by the characteristic defect frequency component contained in the condition monitoring signal (see **Figure 2**). The frequency range of vibration measurement can be as low as in the infrasonic region (below 20 Hz) and span across to the upper limit of the audible frequency (20 kHz). Though the vibration technique can be problematic in acquiring useful signals when it is deployed for condition monitoring of large low-speed rotating machine where the energy of an incipient defective signal is usually weak and often submerges under the background noise. High-frequency techniques such as acoustic emission (AE) technique can be employed to overcome this limitation.

2.2. Acoustic emission technique

Acoustic emission is a transient elastic wave generated by the sudden release or redistribution of stress in a material. For example, in bearing condition-monitoring applications, AE is generated by the sudden release of energy caused by the material deformation as other bearing components passing through a defect part. The signal then propagates to the bearing house to be detected by a monitoring AE sensor. Compared to vibration signals, the AE signal is less likely to be affected by the dominating noise and vibration generated by the moving mechanical components of the monitored system due to the high-frequency nature of the signal (typically above 100 kHz). Care should be taken in choosing the mounting locations of AE sensors to minimize the energy loss along the AE propagation path for better signal clarity. Furthermore, AE also comes with inherited problems such as calibration, nonlinearity, data

http://dx.doi.org/10.5772/67143

storage and transfer, data processing and interpretation. Recently, Lin et al. [1–3] developed a number of signal-processing algorithms to overcome the problems of nonlinearity and large AE data. Nowadays, a large industry deployment of AE technique in bearing condition monitoring is still restricted by the expensive highly specialized AE data acquisition devices.

2.3. Current signal technique

Current signal technique is mainly used to monitor the bearing condition of electric motors. The technique is based on the principle that the vibration signal generated in a motor is closely related to the change of magnetic flux density in the motor. Stator current technique is the frequently employed current signal technique. When a motor bearing operates on a faulty condition, the radial motion of the motor axis would lead to a small shift of the rotor causing a change in the magnetic flux density between the stator and the rotor. The induced voltage will then cause the variation of the stator current. The stator current technique employs noninvasive sensors to monitor the variation of the stator current; therefore, it is easy to implement and simple to operate.

2.4. Oil and debris-monitoring technique

The application of oil and debris-monitoring technique is a frequently employed tribology approach for machine condition monitoring. The health condition of a roller element bearing can be monitored by analysis of the properties and particle contents of the lubrication oil. The application of oil and debris-monitoring technique is rather straightforward. Though such tribology analysis is normally undertaken at laboratories using spectrometers and scanning electron microscopes rather than *in situ*. The technique is also limited to condition-monitoring applications of lubrication-related or wear-related problems.

2.5. Thermography technique

Thermography technique detects faults in a bearing by measuring the emission of infrared energy using thermo-infrared devices during bearing operation process. Laser thermography devices or thermo-infrared cameras are the most common instruments used for such measurement. This technique can be applied to monitor the heat variation caused by the change of bearing lubrication, load, and operation speed. It is particularly sensitive in monitoring the overheating phenomenon caused by improper lubrication but less sensitive in detecting incipient fault developed in a bearing such as early pitting, slight wear, or peeling.

3. Data analysis and fault diagnosis techniques

Many signal analysis and machine fault diagnosis techniques have been developed in the last few decades in order to improve the reliability, efficiency, and lifespan of machines, as well as to reduce the maintenance and operation cost. In this section, the most frequently employed signal-processing and fault diagnosis techniques for rolling element bearings are briefly discussed and summarized.

3.1. Time-domain data analysis techniques

3.1.1. Time series analysis

Time series analysis is a mathematical method, which handles an observed data series in a statistic manner. Time series analysis is based on the assumption that the variation trend of a historical observed data can be extended to provide an indication/prediction of future data variation of a same monitored system. A typical time series analysis approach is to establish a predictive model based on the observed data series. The three widely adopted univariate time series models in machine fault diagnosis are autoregressive (AR) model, moving average (MA) model, and autoregressive moving average (ARMA) model [4].

A general linear ARMA (p, q) model can be expressed as

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{2}$$

where y_t is the time series needed to be modeled, c is a constant, p is the number of autoregressive orders, q is the number of moving average orders, ϕ_i is the autogressive coefficients, θ_j is the moving average coefficients, and ε_t is the independent and identically distributed terms, which are commonly assumed to have zero mean and a constant variance, or (i) $E(\varepsilon_t) = 0$; (ii) $E(\varepsilon_t \varepsilon_T) = 0$ for t ≠ T ; and (iii) $E(\varepsilon_t^2) = \sigma^2$.

An AR model and an MA model can be viewed as the special case of an ARMA model. For instance, when all autoregressive coefficients ϕ_i equal to 0 ($\phi_i = 0,\ 1 \le i \le p$), an ARMA model degrades to an MA model. On the other hand, when all moving average coefficients θ_j equal to 0 ($\theta_j = 0,\ 1 \le j \le q$), an ARMA model degrades to an AR model.

After selecting the most suitable time series model for the data analysis, the next step is to determine the orders of AR or MA models. The commonly adopted order parameter determination criteria for these time series models are the final prediction error (FPE) criterion, Akaike information criterion (AIC), and Bayesian information criterion (BIC).

On the other hand, the common methods used to determine the autoregressive coefficients and moving average coefficients are least squares estimation, maximum likelihood estimation, Yuler-Walker estimation, and so on.

It is worth mentioning that ARMA models are developed based on the assumption that the signal is stationary. If the signal is not stationary, some data preprocessing steps need to be adopted. For example, (1) performing a differential operation of the data term by term until the signal satisfies the stationary criterion and (2) decomposing a nonstationary signal by empirical mode decomposition (EMD) to obtain the stationary intrinsic mode functions (IMFs).

3.1.2. Minimum entropy deconvolution

Minimum entropy deconvolution (MED) technique is proposed originally by Wiggins, which has been successfully employed in dealing with the seismic response signal [5]. The basic idea of MED is to find an inverse filter that counteracts the effect of the transmission path [6]. It is designed to reduce the spread of impulse response functions (IRFs) and then obtain signals,

http://dx.doi.org/10.5772/67143

which are close to the original impulses giving rise to them. The MED technique has been successfully employed in bearing fault diagnosis such as in [6].

Figure 3 illustrates the basic idea of the MED technique. In this process, an unknown impact signal $x(n)$, which can be as highly impulsive as possible, passes through a structural filter **h** and then mixes with a noise $e(n)$ to produce an intermediate output $o(n)$. The signal $o(n)$ then passes through an inverse (MED) filter **f** to produce a final output $y(n)$. The process eliminates the structure resonance and the final output $y(n)$ after the inverse filter needs to be as close as possible to the original input $x(n)$.

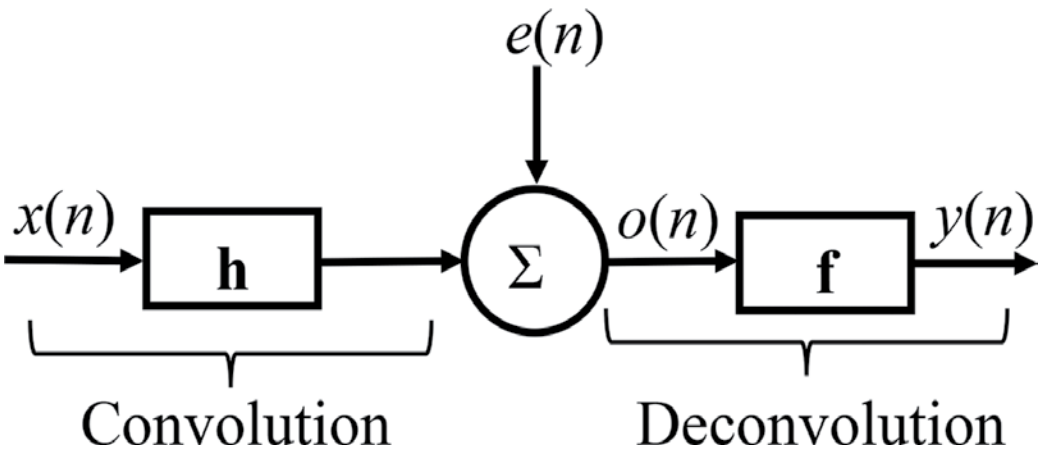


Figure 3. Inverse filtering (deconvolution) process for MED.

The inverse filter *f(l)* can be modeled as a finite impulse response (FIR) filter with L coefficients:

$$y(n) = \sum_{l=1}^{L} f(l)x(n-l), \tag{3}$$

and

$$f(l) * h(l) = \delta(n - l_m). \tag{4}$$

where δ is a Dirac delta function and l_m is the time delay after the MED process, which displaces the entire signal by l_m but keeps the impulse spacing of the signal.

The inverse filter *f(l)* is implemented for the MED technique by the objective function method (OFM). The OFM is an optimization process designed to maximize the kurtosis of the output signal, $y(t)$. OFM achieves this by changing the coefficients of the filter *f(l)*. The kurtosis is taken as the normalized fourth-order moment given by

$$K\Big(f(l)\Big) = \sum_{n=1}^{N} y^4(n) / \left[\sum\nolimits_{n=1}^{N} y^2(n)\right]^2, \tag{5}$$

and the maximum kurtosis of *y(t)* can be obtained according to *f(l)* for which the derivative of the objective function is zero:

$$\partial K(f(l))/\partial f(l) = 0 \tag{6}$$

where the filter coefficients of *f(l)* can converge to a given tolerance by the iterative process [7].

3.1.3. Spectral kurtosis

Spectral kurtosis (SK) was first proposed in the 1980s for detecting impulsive events in sonar signals. SK was first applied in bearing fault diagnosis by Antoni [8]. The method basically computes a kurtosis at "each frequency line" in order to discover the presence of hidden nonstationarities and to indicate in which frequency band it takes place. The method has been proved to be relatively robust against strong additive noise. An SK of nonstationary signals is defined based on the Wold-Cramer decomposition, which describes any stochastic random process *x(t)* as the output of a causal, linear, and time-varying system [9]:

$$x(t) = \int_{-1/2}^{+1/2} H(t,f) e^{j2\pi f n} dZ_x(f), \tag{7}$$

where $dZ_x(f)$ is an orthonormal spectral process of unit variance and $H(t, f)$ is the time-varying transfer function interpreted as the complex envelope of $x(t)$ at frequency f. The SK of a signal $x(t)$ is defined as a normalized fourth-order spectral accumulation given by [8, 9]:

$$K_x(f) = \frac{\langle |H(t,f)|^4 \rangle}{\langle |H(t,f)|^2 \rangle^2} - 2, \tag{8}$$

in which $K_x(f)$ is the spectral kurtosis of signal $x(t)$ around frequency f and $\langle \cdot \rangle$ denotes the averaging over time.

Antoni and Randall [9] proposed two techniques to calculate the spectral kurtosis, one is based on short-time Fourier transform (STFT) (the so-called kurtogram for finding the optimal filter) and the other is based on one-third binary filter banks (fast kurtogram for on-line condition monitoring and fault diagnosis). Kurtogram is a powerful tool for the analysis of nonstationary signals in bearing fault diagnosis, though it has been reported that the technique fails to detect a bearing fault when the defect signal has low signal-to-noise ratio and contains non-Gaussian noise with high peaks [10, 11].

3.1.4. Singular-value decomposition

Singular-value decomposition (SVD) is a numerical method which states that a matrix $\mathbf{A}$ of rank L can be decomposed into the product of three matrices: $\mathbf{U}$ (an orthogonal matrix), $\mathbf{\Lambda}$ (a diagonal matrix) and $\mathbf{V}^{\mathrm{T}}$ (the transpose of an orthogonal matrix $\mathbf{V}$) [12]. This method is usually presented as

$$\mathbf{A}_{m\times n} = \mathbf{U}_{m\times m} \cdot \mathbf{\Lambda}_{m\times n} \cdot \mathbf{V}^{\mathrm{T}}_{n\times n}, \tag{9}$$

where $\mathbf{U^T U} = 1$ and $\mathbf{V^T V} = 1$; $\mathbf{\Lambda}_{m\times n}$ is a diagonal matrix containing the square roots of eigenvalues of $\mathbf{A^T A}$ which can be expressed as $\mathbf{\Lambda} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_L)$, where $L = \min(m, n)$.

The non-zero diagonal terms $\sigma_i (i = 1, 2, \cdots, l)$ together with the zero terms $\sigma_{l+1} = \ldots = \sigma_L = 0$ form the singular values of the matrix **A**.

SVD can be an effective tool for data reduction, for example, a reduction of the fault feature dimensions. It is also a useful tool for signal de-noising. The application of SVD in signal de-noising is mainly operated as follows:

Suppose a bearing CM signal $x(k)$ can be modeled as

$$x(k) = y(k) + n(k), \tag{10}$$

where $y(k)$ and $n(k)$ are respectively, the uncontaminated signal and noise. A conversion method such as phase space reconstruction can be used to transform the signal from a one-dimensional vector into a two-dimensional matrix as follows:

$$\mathbf{A} = \overline{\mathbf{A}} + \mathbf{N}, \tag{11}$$

where the matrix $\overline{\mathbf{A}}$ represents the uncontaminated data $y(k)$, which contains characteristic fault information and the matrix **N** signifies the unwanted noise part $n(k)$. From Eqs. (9) and (11), we have

$$\mathbf{A} = \overline{\mathbf{A}} + \mathbf{N} = [\,\mathbf{U}_1 \quad \mathbf{U}_0\,] \begin{bmatrix} \mathbf{\Lambda}_1 & 0 \\ 0 & \mathbf{\Lambda}_0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^{\mathrm{T}} \\ \mathbf{V}_0^{\mathrm{T}} \end{bmatrix}. \tag{12}$$

$\mathbf{\Lambda}_1$ in Eq. (12) contains the significant singular values $\sigma_i (i = 1, 2, \cdots, l)$ which are used to construct the uncontaminated data and $\mathbf{\Lambda}_0$ contains small singular values σ_i $(i = l + 1, \ldots, L)$, which can be viewed as noise.

From Eq. (12), the de-noising signal matrix can be written as

$$\overline{\mathbf{A}} = \mathbf{U}_1 \mathbf{\Lambda}_1 \mathbf{V}_1^{\mathrm{T}} \tag{13}$$

$\overline{\mathbf{A}}$ is a reconstructed matrix using only the largest l number of singular values whose values are greater than a pre-set threshold value ε. The rest of the singular values are replaced by zero such that

$$\sigma_i = 0 \text{ when } \sigma_i \leq \varepsilon,\ t = l + 1, \ldots, L. \tag{14}$$

3.2. Frequency-domain data analysis techniques

3.2.1. Power spectrum

Spectrum analysis is the most popular frequency-domain analysis techniques, which transforms a time-domain data into discrete frequency components by Fourier transform. The power spectrum of a time-domain signal is defined as the square of the magnitude of the Fourier transform of a signal. It can be written as

$$P(\omega) = \left|\int_{-\infty}^{+\infty} x(t)\mathrm{e}^{-\mathrm{j}\omega t}dt\right|^2 = X(\omega)X^*(\omega), \tag{15}$$

where $X(\omega) = \int_{-\infty}^{+\infty} x(t)e^{-\mathrm{j}\omega t}dt$ is the Fourier transform of a signal, $X^*(\omega)$ is its complex conjugate and ω is the angular frequency in radian/s. The power spectrum is frequently employed to extract useful characteristic defect frequency components of a stationary CM signal.

3.2.2. Cepstrum

A cepstrum is defined as the power spectrum of the logarithm of the power spectrum of a signal [13]. The name of cepstrum was derived by reversing the first four letters of spectrum. There are four types of cepstra: a real cepstrum, a complex cepstrum, a power cepstrum and a phase cepstrum.

The real cepstrum of a signal $x(t)$ is given by

$$\mathrm{c}(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log|X(\omega)|\mathrm{e}^{\mathrm{j}\omega t}d\omega. \tag{16}$$

The complex cepstrum of a signal $x(t)$ is given by

$$\vec{c}\,(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\Big(X(\omega)\Big)\mathrm{e}^{\mathrm{j}\omega t}d\omega. \tag{17}$$

The power cepstrum of a signal $x(t)$ is given by

$$c(t)^2 = \frac{1}{2\pi}\left|\int_{-\pi}^{\pi} \log|X(\omega)|\mathrm{e}^{\mathrm{j}\omega t}d\omega\right|^2. \tag{18}$$

The most commonly used cepstrum in machine condition monitoring is power cepstrum. The following steps can be taken to calculate the power cepstrum of a signal: A signal → power spectrum of the signal → logarithm of the power spectrum → power spectrum of the data from the previous step → inverse Fourier transform of the log power spectrum from the previous step (power cepstrum). The main application of cepstrum analysis in machine condition monitoring is for signals containing families of harmonics and sidebands where it is the whole family rather than individual frequency component characterizing the fault [13] (typical for bearing CM signals). The technical terms used in a cepstrum are quefrency, gamnitude and rahmonics corresponding to frequency, amplitude and harmonics in a spectrum analysis. The cepstrum can be used for the harmonics generated by bearing faults, but only if they are well separated. In contrast, envelope analysis to be introduced in the next section is not limited by such restriction.

3.2.3. Envelope spectrum

It has been pointed out by Randall [13] that the benchmark method for rolling element bearing diagnosis is envelope analysis as the spectrum of raw bearing CM signals often contains little information about bearing faults. In envelope analysis, a time waveform is bandpass filtered in a high-frequency band and the fault signal is amplified by the structural resonances and

http://dx.doi.org/10.5772/67143

amplitude modulated to form the envelope signal for bearing diagnosis [13]. The procedures for envelope analysis are briefly described below:

Given a real signal $x(t)$, its Hilbert transform, $h(t) = H\{x(t)\}$ is defined as

$$h(t) = H\{x(t)\} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau = \frac{1}{\pi\tau} x(t), \tag{19}$$

The real signal $x(t)$ and its Hilbert transform $h(t)$ together form a new complex analytic signal $z(t)$,

$$z(t) = x(t) + jh(t). \tag{20}$$

The envelope signal $E(t)$ is simply the absolute value of the analytic signal $z(t)$,

$$E(t) = |z(t)| = |x(t) + jh(t)| = \sqrt{x^2(t) + h^2(t)}. \tag{21}$$

After taking a fast Fourier transform on the envelope signal $E(t)$, an envelope spectrum can be obtained. An envelope spectrum can reveal the repetition characteristic defect frequencies caused by bearing faults even in the presence of a small random fluctuation. The envelope spectrum of the noise-added bearing defect signal shown in **Figure 2(b)** is given in **Figure 4**. It is noted that due to a high noise level in the signal (SNR = 0.05, representing an initial bearing defect), the envelope spectrum is compromised by many artificial frequency components and produces a subtle bearing defect information. Also due to such interference, the calculated defect frequency (the modulated frequency) of the outer race fault and its higher harmonics (1 × BPFO, 2 × BPFO and so on) shown in the envelope spectrum are also slightly lower than that of the simulated defect frequency (1 × BPFO = 78.18 Hz).

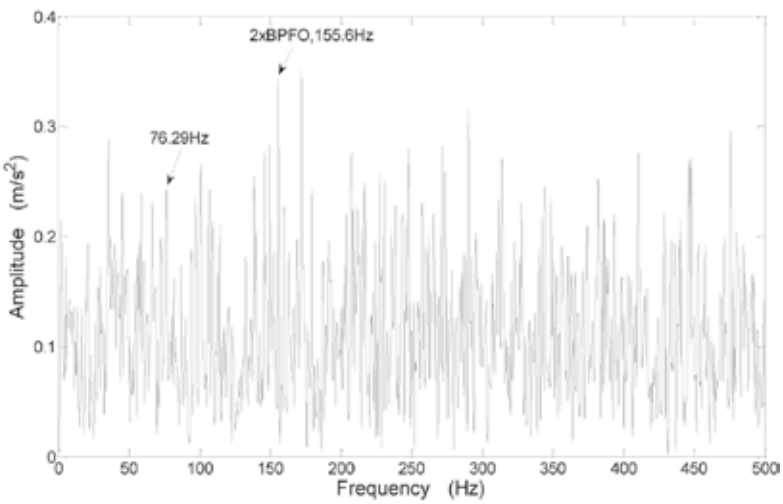


Figure 4. Envelope spectrum of the noise-added bearing defect signal shown in **Figure 2(b)**.

3.2.4. Higher-order spectra

Higher-order spectra (HOS) (also known as polyspectra) consist of higher-order moment of spectra, which are able to detect nonlinear interactions between frequency components [14]. Higher-order spectra are defined as the Fourier transform of the corresponding cumulant sequences of a signal. For instance, the first-order cumulant of a stationary process is the mean of the signal:

$$C_1 x = E\{x(n)\}. \tag{22}$$

The second- and third-order cumulants of a stationary process are defined as

$$C_2 x(k) = E\{x^*(n)x(n+k)\}, \tag{23}$$

and

$$C_3 x(k,l) = E\{x^*(n)x(n+k)x(n+l)\} - C_2 x(k)C_2 x(l-m) - C_2 x(l)C_2 x(k-m) \tag{24}$$

From Eqs. (23) and (24), it is clear that the power spectrum is, in fact, the FT of the second-order cumulant of a signal and the third-order spectrum also termed as bispectrum is the FT of the third-order cumulant. Note that the bispectrum $S_{3x}(\omega_1, \omega_2)$ is a function of two frequencies. Therefore, it can detect phase coupling between two frequency components which appears as a third frequency component at the sum or difference of the first two (frequencies) with a phase that is also the sum or difference of the first two.

Traditionally, power spectrum is used to break down a time waveform signal into a series of frequency components. However, power spectrum cannot determine whether peaks at harmonically related positions are phase coupling since power spectrum uses only the magnitude of Fourier components and the phase information is neglected. Higher-order spectra such as bispectrum use the phase information of the signal and are capable to detect phase coupling of frequency components in the spectra. Therefore, a bispectrum can provide additional phase information than a power spectrum analysis.

The motivation behind the use of higher-order spectrum analysis is summarized as follows. Firstly, the technique can suppress Gaussian noise in the data processing of unknown spectral characteristics for fault detection, parameter estimation and classification problems. If Gaussian noise is embedded in a non-Gaussian signal, a HOS transform can eliminate the noise. On the other hand, periodic, quasi-periodic signals and self-emitting signals from complex machinery in practical applications are typical non-Gaussian signals which will be preserved in the transform. Secondly, a HOS analysis can preserve the phase information. For example, there are situations in practice in which the interaction between two harmonic components creates a third component at their sum and/or difference frequencies. Thirdly, HOS can play a key role in detecting and characterizing the type of nonlinearity in a system from its output data.

For a better illustration of the signal-processing techniques discussed above in data analysis, a number of case studies are given in the following section to exemplify the usage of the above algorithms in bearing fault detection.

3.2.5. Case studies

Case 1: (AR & MED de-noising): In this case study, the AR model described in Section 3.1.1 and the MED method described in Section 3.1.2 are employed in the analysis to filter the noise-added bearing defect signal shown in **Figure 2(b)** prior to an envelope analysis to enhance the bearing

http://dx.doi.org/10.5772/67143

defect frequency components in the envelope spectrum for a more reliable bearing fault diagnosis. The results are presented in **Figures 5** and **6**. Compared to the result in **Figure 4**, it is shown that the two-step de-noising by the AR and MED models has successfully suppressed the artificial components and enhanced the defect signal representation in the spectrum.

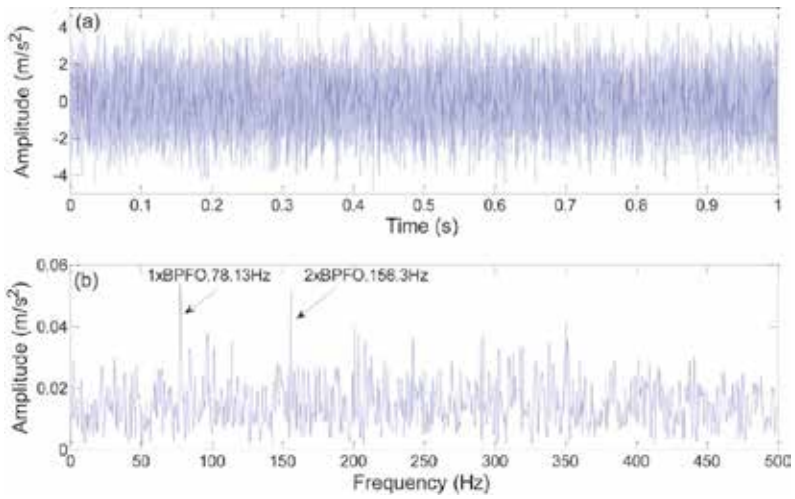


Figure 5. The time waveform and the envelope spectrum of the simulated bearing defect signal after preprocessed by an AR model: (a) the time waveform after filtered by an AR model; (b) envelope spectrum of the signal shown in **Figure 5(a)**.

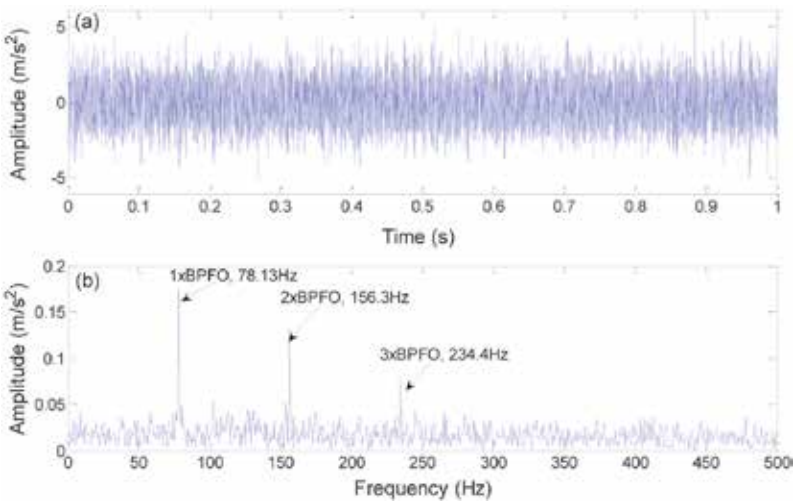


Figure 6. The time waveform and the envelope spectrum of the simulated bearing defect signal after preprocessed by an AR model and filtered further by an MED model: (a) the time waveform after de-noising; (b) envelope spectrum of the signal shown in **Figure 6(a)**.

Case 2 (spectrum kurtogram): The signal used in this case study is shown in **Figure 7**. The signal is generated by adding 0-dB white noise to the simulated bearing defect signal presented in **Figure 2(a)**. In the analysis of the signal, the fast kurtogram algorithm [9] described in Section 3.1.3 is first employed to determine the bearing resonance band (to obtain the center

frequency and the bandwidth) having the highest band energy (corresponding to the highest kurtosis value) in the signal. A five-level fast kurtogram based on a filter band and fast Fourier transform is shown in **Figure 8** and the highest band energy is found to occur in the band encircled by the white ellipse in the figure. The band is found to be centered at 3958.33 Hz (close to the simulated bearing resonance frequency of 4000 Hz as listed in **Table 2**) and has the bandwidth of 416.67 Hz. The band has the highest kurtosis value of 0.1 which occurs at level 4.5 in the decomposition. The next step after the decomposition is to take an envelope analysis based on the band-filtered optimum frequency band signal obtained from the fast kurtogram and the result is shown in **Figure 9(a)** and **(b)**. It is shown that the spectrum kurtosis technique can detect the characteristic defect frequency from weak-bearing defect signals.

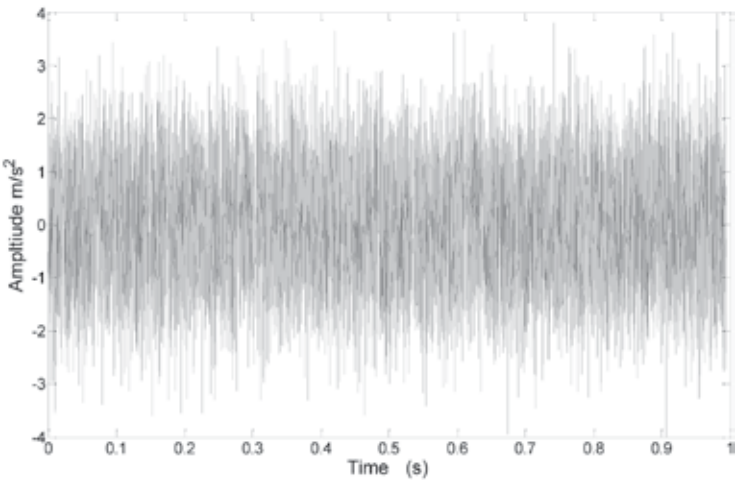


Figure 7. Simulated bearing defect signal with 0-dB white noise added.

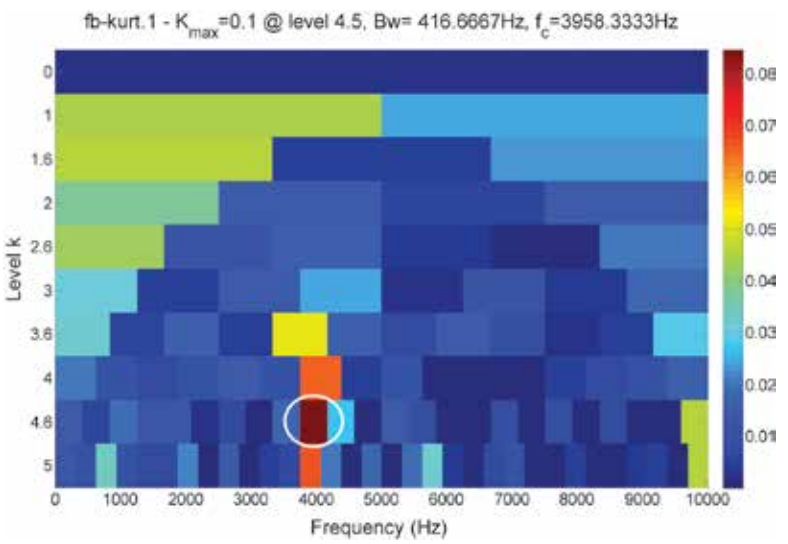


Figure 8. Fast kurtogram calculated based on the defect signal shown in **Figure 7**.

http://dx.doi.org/10.5772/67143

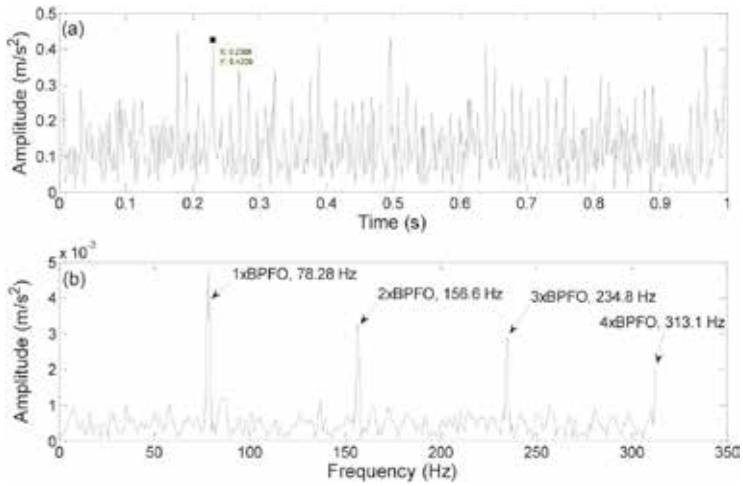


Figure 9. Spectrum kurtosis analysis of the bearing defect signal: (a) envelope of the band-filtered signal; (b) Fourier transform of the envelope signal.

3.3. Time-frequency analysis

A major limitation of frequency-domain analysis is that it is only useful in dealing with stationary signals. The presence of transient or nonstationary signals in the data would not be captured in a traditional frequency-domain analysis. To overcome this limitation, time-frequency analysis techniques are then developed for a better understanding of how spectrum properties change with time. In time-frequency analysis, waveform signals are analyzed in both time and frequency domains to capture the progressive change of spectrum components. The most commonly employed time-frequency analysis techniques are short-time Fourier transform (STFT), wavelet transform (WT), Wigner-Ville distribution (WVD) and adaptive signal analysis techniques such as empirical mode decomposition (EMD) technique.

3.3.1. Short-time Fourier transform

Short-time Fourier transform adds a time variable to the traditional Fourier spectrum, thus allowing it to investigate the time-varying nature of a signal. In a SFFT analysis, a continue time-domain waveform is multiplied by a sliding narrow time window and a Fourier transform is computed on the windowed signal at each time step of the sliding time window. The STFT analysis is based on the consideration that if a signal can be considered stationary over the length of the chosen sliding time window, a Fourier transform can be performed on the windowed signal segment for each new position of the sliding time window to obtain a satisfactory time-frequency analysis of a nonstationary signal.

A short-time Fourier transform of a continuous signal $x(t)$ can be computed by

$$\mathrm{STFT}_x(\tau,\omega) = \int_{-\infty}^{+\infty} x(\tau) w^*(\tau - t) \mathrm{e}^{-\mathrm{j}\omega\tau} d\tau, \tag{25}$$

where $w(\tau - t)$ is a finite time window function centered at time t. The asterisk sign (*) in the time window indicates a complex conjugate and the analysis window can be regarded as the impulse response of a low-pass filter. A spectrogram, which is the squared amplitude of an

STFT transform, is often used in the display of the transformation result for signal analysis and interpretation. A major limitation of STFT is that the analysis has a uniform resolution in both time and frequency planes implying that a small time window will have a good time resolution but poor frequency resolution and vice versa. This drawback limits the technique for the analysis of signals with slow-changing dynamic only. Another time-frequency analysis technique, wavelet transform, has then been developed to overcome this limitation.

3.3.2. Wavelet transform

Wavelet transform (WT) is a time scale representation of a signal. It is the inner product of a signal with the translated and scaled family of a mother wavelet function $\psi(t)$. In general, WT analysis can be categorized into three forms: a continuous wavelet transform (CWT), a discrete wavelet transform (DWT), and a more general wavelet packet decomposition (WPD).

In CWT, the wavelet transform of a continuous signal $x(t)$ is calculated using

$$W_f(u,s) = \frac{1}{\sqrt{s}} \int x(t)\psi^* \left(\frac{t-u}{s}\right) dt. \tag{26}$$

where s is the scale parameter and u is the time translation. If one considers the wavelet function $\psi(t)$ as a bandpass impulse response, then the wavelet transform is simply a bandpass analysis. Care should be taken to ensure that the decomposed signal can be perfectly reconstructed from its wavelet representation when using a wavelet transform. Thus, a WT has to meet the criterion which is also known as the admissibility condition.

A CWT is mainly used in data analysis of scientific research. In practical applications, the discrete version, a discrete wavelet transform (DWT), is more popular due to the small computation cost and excellent signal compaction properties. A DWT decomposes a signal into different frequency bands by passing it through a series of filters. In each step of decomposition, a signal is passed through a pair of low- and high-pass filters simultaneously accompanying by down sampling. A more general form of wavelet analysis is the so-called wavelet packet decomposition (WPD). In WPD, a signal is split into two parts, one contains a vector of approximation coefficients and the other contains a vector of detail coefficients in each stage of decomposition. Both the details and the approximations can be split further in the next level of decomposition which offer a great range of possibilities to decode a signal than ordinary wavelet analysis. Wavelet transform has been widely employed in signal processing for condition monitoring and fault diagnosis of rotating machine [15].

3.3.3. Wigner-Ville distribution

Another popular time-frequency analysis technique is Wigner-Ville distribution (WVD). WVD is the core distribution for the quadratic class of quadratic time-frequency distributions. It yields an ideal resolution for mono-component, linearly frequency-modulated signals, but produces undesired cross-terms for multicomponent and nonlinearly frequency-modulated signals. Wigner-Ville distribution can be viewed as a particular case of the Cohen class

http://dx.doi.org/10.5772/67143

distributions which yields a time-frequency energy density computed by correlating the signal with a time and frequency translation of itself.

The WVD of a signal $x(t)$ is defined by

$$W_x(t,\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} x\left(t+\frac{\tau}{2}\right)\cdot x^*\left(t-\frac{\tau}{2}\right)\cdot e^{-j\omega\tau}d\tau, \tag{27}$$

where x^* denotes the conjugate of x. Thus, the Wigner-Ville integral is in fact a Fourier transform of the inner product of a signal and its conjugate with a time delay variable τ. The bilinear nature of this procedure therefore avoids the loss of time-frequency resolution in the transform such as the resolution problem encountered when performing the finite sliding time windowing in STFT.

Compared with other distributions, the WVD has the desirable property of fulfilling the marginal condition, thus the total signal energy can be calculated in time or in frequency using the Plancherel formula:

$$||x^2|| = \int_{-\infty}^{+\infty}|x(t)|^2 dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty}|X(\omega)|^2 d\omega \tag{28}$$

The values $||x(t)||^2$ and $||X(\omega)||^2$ can be interpreted as the energy densities in time and frequency domain, respectively. This enables a direct computation of the energy present at a given time-frequency box from the WVD output.

Another important feature of WVD is that its first conditional moment for a given time t_c equals the instantaneous frequency:

$$\frac{d\varphi}{dt} = \langle\omega\rangle_{t_c} = \frac{1}{|s(t_c)|^2}\int_{-\infty}^{+\infty}\omega W(t_c,\omega)d\omega. \tag{29}$$

Therefore, it can be computed as the average of all frequencies ω present in the time-frequency plane at a time t_c.

In discrete form, the WVD is defined as

$$\mathrm{WVD}(n,k) = \sum_{p=-N}^{N-1} R[n,q]\cdot e^{-j2\pi kq/N}, \tag{30}$$

where $R[n, q]$ is the instantaneous correlation given by

$$R[n,q] = x\left[n+\frac{q}{2}\right]\cdot x^*\left[n-\frac{q}{2}\right], \tag{31}$$

in which n is the number of samples of the analytical or interpolated form of the discrete signal $x[n]$ and q is an odd integer.

Since the instantaneous correlation is centered on a value, the delay q is distributed between the delayed sample $x[n-q/2]$ and the corresponding advanced sample, $[n+q/2]$. It is thus necessary to calculate the value of $x[n]$ at the two half integer positions using an interpolation. In addition, positions at the extremes of $x[n]$ are padded with zeros in order to compute the Fourier transform. A major drawback of WVD analysis is that it can induce artifacts and negative values which need to be properly compensated in the signal analysis [16].

3.3.4. Adaptive signal decomposition

Empirical mode decomposition (EMD) is an adaptive time-frequency analysis technique originally proposed by Huang [17] in 1998. It is based on the local characteristic time scales of a signal and can decompose the signal into a set of complete and almost orthogonal components termed as intrinsic mode functions (IMF). Lei et al. [18] provided a review on the successful application examples of EMD technique in fault diagnosis of rotating machines.

In EMD analysis, the decomposed signal can be represented by

$$\mathrm{x}(t) = \sum_{i=1}^{N} C_i(t) + r_N(t), \tag{32}$$

where $C_i(t)$ represents the ith IMF component and $r_N(t)$ is the residual after the EMD decomposition.

The IMFs represent the natural oscillatory modes imbedded in the signal and can serve as the basis functions, which are determined by the signal itself, rather than predetermined kernels. Thus, the decomposition is a self-adaptive signal process suitable for nonlinear and nonstationary data analysis.

Although the EMD technique has been successfully employed in the analysis of nonlinear and nonstationary signals in various applications, the algorithm itself also has a number of weaknesses, for instance, a lack of a solid theoretical foundation, end effects, a sifting stop criterion and extremum interpolation. To overcome some of these deficiencies, an improved EMD algorithm or a so-called ensemble empirical mode decomposition (EEMD) technique has been developed [19]. EEMD is a noise-assisted data analysis technique which imposes a white Gaussian noise into a signal and then decomposes the mixed signal by using the EMD algorithm. A major advantage of the EEMD technique is that no missing scales will be presented in the decomposition and the IMF components in different scales of the signal are automatically projected into proper reference scales established by the white noise in the background [20].

3.3.5. Case studies

Case 3 (EMD de-noising): In this analysis, the noise-added signal shown in **Figure 7** is adaptively decomposed into 14 IMF components and a residual component using the EMD technique described in the previous section. The decomposition result is shown in **Figure 10**. The correlation coefficient of each IMF component with the original signal can be calculated using

http://dx.doi.org/10.5772/67143

$$\rho_{xy} = \frac{n\sum x_i y_{j,i} - \sum x_i \sum y_{j,i}}{\sqrt{n\sum x_i^2 - \left(\sum x_i\right)^2}\sqrt{n\sum y_{j,i}^2 - \left(\sum y_{j,i}\right)^2}}, \tag{33}$$

where x_i $(i = 1, 2, \ldots, n)$ is the original data, $y_{j,i}$ $(i = 1, 2, \ldots, n)$ is the data of one of the jth IMF components and n is the data record length. The calculated correlation coefficients for the 14 IMF components and the residual are listed in **Table 3**.

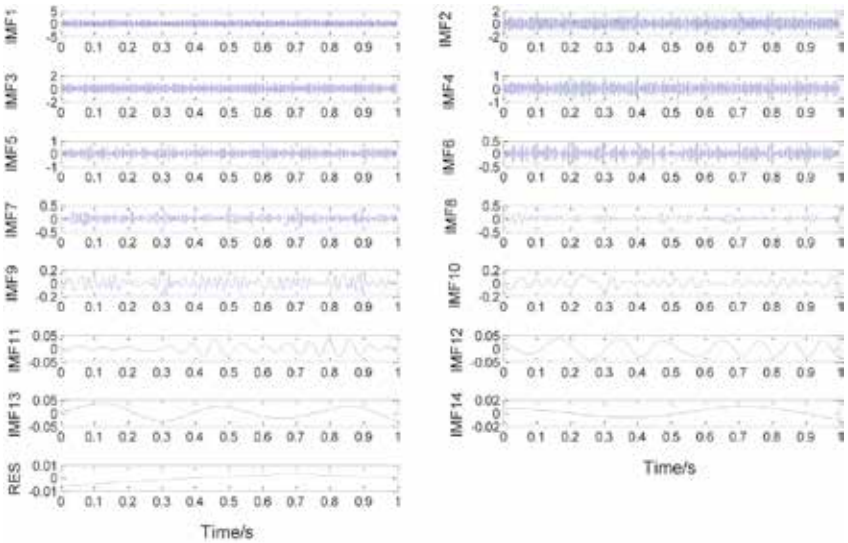


Figure 10. The IMF components and residual of the bearing defect signal from the EMD decomposition.

Component	Correlation coefficient	Component	Correlation coefficient	Component	Correlation coefficient
IMF1	**0.7084**	IMF6	0.1348	IMF11	0.0152
IMF2	0.4767	IMF7	0.0913	IMF12	0.0152
IMF3	0.3220	IMF8	0.0535	IMF13	0.0169
IMF4	0.2353	IMF9	0.0485	IMF14	0.0087
IMF5	0.1747	IMF10	0.0340	RES	−0.0048

Table 3. The correlation coefficients of the IMF components and the residual.

It is shown in the figure that the first IMF component (IMF1) has the highest correlation coefficient with the original signal implying that it is most closely related to the defect signal. Therefore, an envelope analysis is undertaken on the IMF1 component in the next step of analysis. The result is shown in **Figure 11** where the bearing defect frequency and its second-order harmonic can be clearly discriminated from the envelope spectrum.

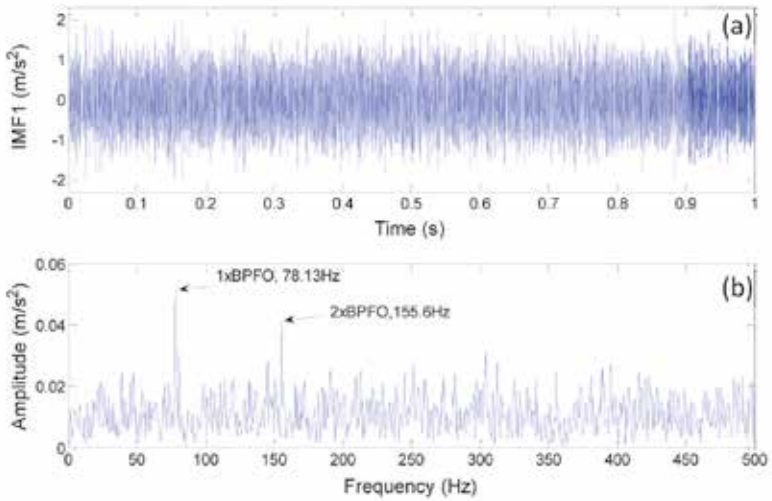


Figure 11. The envelope spectrum of the first IMF component of the bearing defect signal.

Case 4 (Signal de-noising using SVD decomposition): The singular-value decomposition described in Section 3.1.4 can be employed to filter out the noise from bearing condition-monitoring signals to enhance the impulses produced by a bearing defect. In this approach, the one-dimension time waveform of the bearing condition-monitoring signal $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ (as shown in **Figure 12(a)**) is rearranged by an incrementing process to form a Hankel matrix $\mathbf{A}(p, g)$ which is then decomposed into three matrices using Eq. (9) in Section 3.1.4 to obtain the singular values $\sigma_i,\ i = 1, 2, \ldots L,$ whose values are shown in **Figure 13(a)**.

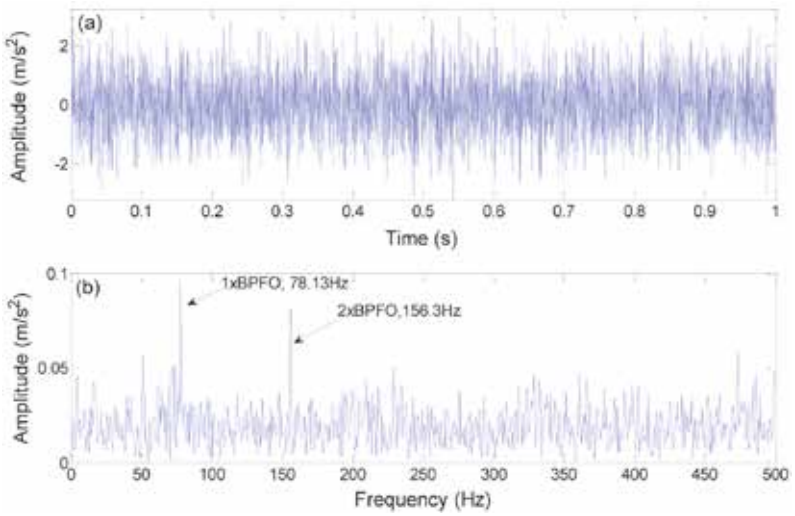


Figure 12. Noise-added bearing defect signal and its envelope spectrum: (a) time waveform (note: the sampling frequency in this simulation is reduced to 5 kHz to avoid the requirement of large computer memory in SVD decomposition; (b) envelope spectrum.

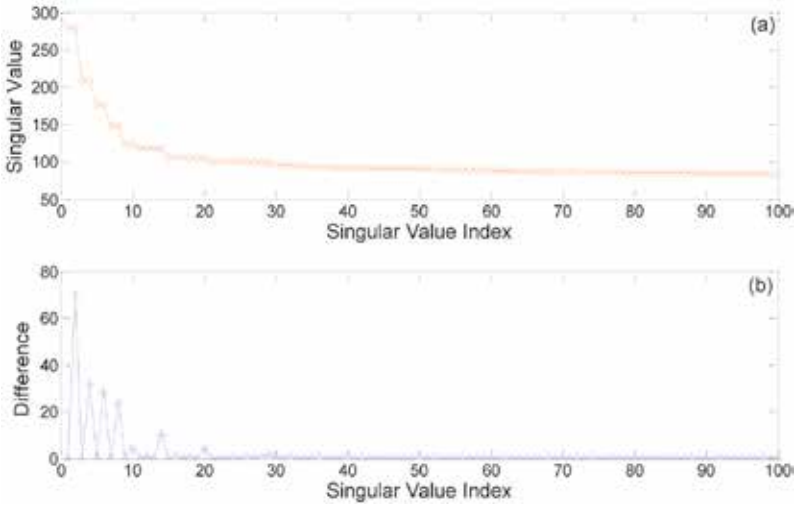


Figure 13. (a) Singular values and (b) the series of the difference between two sequential singular values.

The next step is to calculate the difference between the subsequent singular values $b_i = \sigma_i - \sigma_{i+1},\ (i = 1, 2, \cdots, L-1)$ to form a series vector $\mathbf{B} = (b_1, b_2, \ldots, b_{L-1})$ which reflects the variation of the two neighboring singular values. Vector **B** is plotted in **Figure 13(b)** for illustration. When the difference between two neighboring singular values is large, they will form a larger peak in the difference vector **B** (see **Figure 13(b)**) indicating that there is a small correlation between the defect and noise signals at the corresponding singular values. Keeping the singular values prior to as well as the largest difference peak and letting the singular values after this peak to zero and then substituting the modified singular-value vector to Eq. (9) to obtain a new Hankel matrix (note, as the first difference peak happens to be the largest peak in our case, the singular values associated with the sequential three peaks are also used in the modified singular value vector). We can then reverse the process of the first step to obtain the de-noised bearing defect signal as shown in **Figure 14(a)**. The envelope spectrum of the de-noised signal (**Figure 14(a)**) is shown in **Figure 14(b)**. It is shown that the SVD de-noise can yield a much clean spectrum mainly containing the bearing defect frequency component and its higher harmonics.

3.4. Statistical-based bearing fault diagnosis

3.4.1. Statistical features in the time domain

Some useful statistical features obtained directly from a time waveform signal can be used to evaluate the health condition of a rolling element bearing. Such features can be grouped into two categories: (a) dimensional features and (b) nondimensional features. The group of dimensional statistical features includes peak value, root-mean-square (RMS) value, absolute mean value and variance. This feature group is closely related to bearing fault severity, for instance, their values increase as the bearing fault severity increases. Though, care needs to be taken as their values are also influenced by the working conditions such as load or rotate speed of a machine.

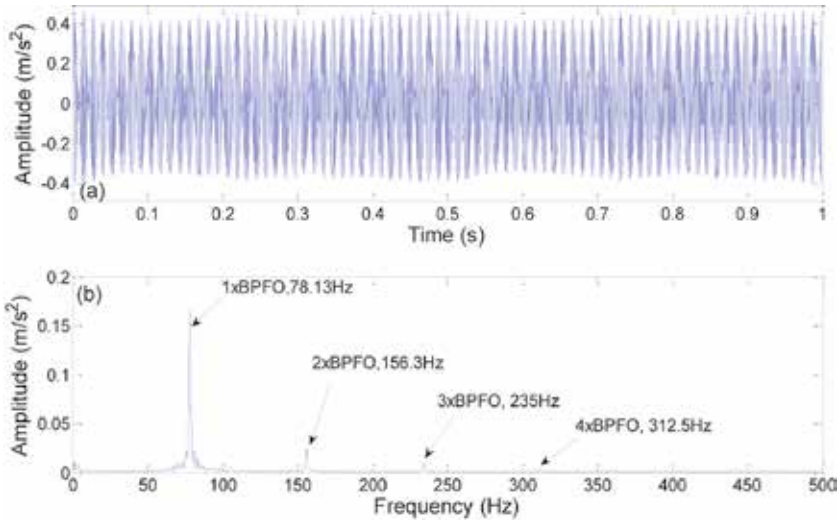


Figure 14. (a) The de-noised bearing defect signal after SVD decomposition; (b) the envelope spectrum.

For a smooth and ergodic continuous time-domain signal $x(t)$, its peak value can be calculated as

$$x_p = \text{Max}[|x(t)|]. \tag{34}$$

Its RMS value which reflects the power level of the signal is calculated by

$$x_{\text{rms}} = \sqrt{\int_{-\infty}^{+\infty} x(t)^2 p(x) dx}, \tag{35}$$

where $p(x)$ is the probability density function of the signal $x(t)$, which represents the probability level that the signal $x(t)$ fall into a certain interval.

Alternatively, an approximate formula can be used to calculate the RMS value as

$$x_{\text{rms}} = \sqrt{\frac{1}{T}\int_0^T x^2(t) dt}. \tag{36}$$

The absolute mean value is defined as

$$x_{\text{av}} = \int_{-\infty}^{+\infty} |x| p(x) dx. \tag{37}$$

Or it can be calculated using the approximate formula:

$$x_{\text{av}} = \frac{1}{T}\int_{-\infty}^{+\infty} |x(t)| dt. \tag{38}$$

Variance is used to depict the fluctuation of a signal that deviated from the center, which can be viewed as the dynamic feature of a signal. The variance of a signal $x(t)$ is

http://dx.doi.org/10.5772/67143

$$D_x = \sigma_x^2 = \int_{-\infty}^{+\infty} (x - \mu_x)^2 p(x) dx, \tag{39}$$

where μ_x is the mean value and σ_x is the standard deviation of the signal.

Variance can also be calculated approximately using the following formula:

$$D_x = \sigma_x^2 = \frac{1}{T} \int_0^T \left(x(t) - \mu_x \right)^2 dt. \tag{40}$$

Table 4 lists the mathematical formula for calculating the same features for a corresponding discrete time waveform, $\{x(n) | n = 1, 2, \cdots, N\}$.

Statistical features	Formula
RMS	$\mu_x = \frac{1}{N} \sum_{i=1}^{N} x_i$
Mean	$x_{\text{rms}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} x_i^2}$
Absolute mean	$x_{\text{av}} = \frac{1}{N} \sum_{i=1}^{N} \lvert x_i \rvert$
Variance	$D_x = \sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - x_{\text{av}})^2$

Table 4. The dimensional statistical features of a discrete time series.

The group of nondimensional statistical features includes crest factor, shape factor, impulsion factor, clearance factor, skewness, and kurtosis factors which are the ratio of two-dimensional statistical features. This type of features is insensitive to the change amplitude or frequency in a signal and thus is not influenced by the working condition of a machine and can accurately reflect a fault condition of a bearing. The formulas for these features are listed in **Table 5**:

Statistical features	Formula
Crest factor	$C = \frac{x_p}{x_{\text{rms}}}$
Shape factor	$S = \frac{x_{\text{rms}}}{x_{\text{av}}}$
Impulsion factor	$I = \frac{x_p}{x_{\text{av}}}$
Clearance factor	$Cl = \frac{x_p}{x_r}$
Skewness	$Sk = \frac{\int_{-\infty}^{+\infty} \left(x(t) - \mu_x \right)^3 p(x) dx}{\sigma_x^3}$ (continuous) $Sk = \frac{\sum_{i=1}^{N} (x_i - \overline{X})^3}{(N-1)S^3}$ (discrete)
Kurtosis	$K = \frac{\int_{-\infty}^{+\infty} \left(x(t) - \mu_x \right)^4 p(x) dx}{\sigma_x^4}$ (continuous) $K = \frac{\sum_{i=1}^{N} (x_i - \overline{X})^4}{(N-1)S^4}$ (discrete)

Table 5. The nondimensional features of a time waveform.

In **Table 5**, x_r is the root amplitude of a continuous time waveform, which is given by

$$x_r = \left(\int_{-\infty}^{+\infty} |x(t)|^{1/2} p(x) dx\right)^2. \tag{41}$$

Its discrete form is given by

$$x_r = \left(\frac{1}{N}\sum_{i=1}^{N} \sqrt{|x_i|}\right)^2. \tag{42}$$

Shape factor and impulsion factor are often used to examine whether there exists pulse shocks in a signal. Clearance factor is sometimes used to examine the wear condition of a machine. Yet, the most frequently employed nondimensional features in data analysis are the skewness and kurtosis factors. The skewness is the third statistic moment that measures the degree of asymmetric distribution of a time waveform. The kurtosis is the fourth statistic moment that measures the "Peakness" of the data distribution.

3.4.2. Statistical features in the frequency domain

The power spectrum depicts the power amplitude level of the frequency components in a signal. If the power amplitude levels for some of the frequency components change, the weight-averaged center frequency of a power spectrum will also change. For example, if the number of frequency components with dominant amplitude in a spectrum increases, the energy distribution will be more dispersed. On the contrary, the energy distribution will be concentrated more around the dominant frequency components if there are only a few of such components in the spectrum. Hence, the fluctuation of a signal in the frequency domain can reflect the change of machine condition by observing the change in the weight-averaged center frequency of power spectrum or the disperse degree of power amplitude level distribution.

The typical statistical featured used to depict the weight-averaged center of a power spectrum are "Frequency center" and "Mean-square frequency," which are defined as follows:

$$\text{Frequency center : FC} = \frac{\int_{-\infty}^{+\infty} f S(f) df}{\int_{0}^{+\infty} S(f) df}, \tag{43}$$

and

$$\text{Mean-square frequency : MSF} = \frac{\int_{0}^{+\infty} f^2 S(f) df}{\int_{0}^{+\infty} S(f) df}, \tag{44}$$

where $S(f)$ represents the power spectrum of a continuous waveform $x(t)$.

The statistical feature used to depict the disperse degree of energy distribution in the frequency domain is the "Variance of frequency," which is defined as

http://dx.doi.org/10.5772/67143

$$\text{Variance of frequency : VF} = \frac{\int_0^{+\infty}(f-FC)^2 S(f)df}{\int_0^{+\infty} S(f)df} = \text{MSF}-\text{FC}^2. \tag{45}$$

Accordingly, for a discrete time series $\{x(n)|n = 1, 2, \cdots, N\}$, the three frequency-domain statistical features are given by

$$\text{FC} = \frac{1}{2\pi\Delta}\frac{\int_0^{\pi}\omega S(\omega)d\omega}{\int_0^{\pi} S(\omega)d\omega}, \tag{46}$$

$$\text{MSF} = \frac{1}{4\pi^2\Delta^2}\frac{\int_0^{\pi}\omega^2 S(\omega)d\omega}{\int_0^{\pi} S(\omega)d\omega}, \tag{47}$$

and

$$\text{VF} = \frac{1}{4\pi^2\Delta^2}\frac{\int_0^{\pi}(\omega-2\pi\Delta FC)^2 S(\omega)d\omega}{\int_0^{\pi} S(\omega)d\omega} = \text{MSF}-\text{FC}^2, \tag{48}$$

where Δ is the sampling frequency, $S(\omega)$ is the power spectrum of a discrete time series $x(n)$, which can be obtained using the following formula:

$$S(\omega) = X(\omega)\overline{X(\omega)}, \tag{49}$$

and

$$X(\omega) = \sum_{n=1}^{N-1} x(n)e^{-j\pi\omega}, \tag{50}$$

where ω is the angular frequency.

3.4.3. Data complexity index

The complexity of a signal can be described by two measures: entropy and Lempel-Ziv complexity.

Entropy is a measure of randomness, suggested by Shannon in 1948 [21]. Entropy is used to depict the randomness (or "uncertainty") existed in a signal or the amount of information carried by the signal.

For a discrete random variable x with probability density function $\{x_i|i = 1, 2, \cdots, N\}$, the Shannon entropy is given by

$$H(x_i) = -\sum_{i=1}^{n} p(i)\log_b p(i), \tag{51}$$

where $p(i)$ is the probability of the ith event of the random variable x, b is the base of the logarithm which takes the common logarithm base values of either 2, e, or 10 depending on the

application. For events with a probability around 0 or 1, the term $p(i)\log_b p(i)$ converges to zero and the entropy is zero. For random signals with uniform probability density function such as pure noise, the entropy is maximum.

The entropy-based techniques have been widely used in bearing fault diagnosis in the last decade. It has been shown that the entropy value is closely related to the working condition of a machine and its value decreases monotonously with aggravation of faults or conditions [22]. Entropy is often combined with other techniques to capture the detail changes of the nonlinearity and nonstationary properties of a signal in machine fault diagnosis [23, 24]. Typical entropies used in machine fault diagnosis are approximate entropy, sample entropy, fuzzy entropy and permutation entropy. Yet, some of these features are still not ideal and problematic in CM applications. For example, approximate entropy is heavily dependent on the data record length which could yield lower estimation value. Sample entropy uses Heaviside step function which is mutational and discontinue at the boundary. Fuzzy entropy is calculated based on the membership function which is difficult to determine accurately. Permutation entropy requires the reconstruction of phase space though the embedding dimension and time lag of the reconstructed matrix need to be selected manually which has so far limits its application.

Ziv and Lempel [25] presented a specific complexity algorithm termed as Lempel-Ziv complexity (LZC) to calculate the complexity of a finite length time series. LZC can reflect the rate for generating the new condition pattern feature as the nonlinearity of a time series grows [26]. Its value represents the degree of random variation of a time series. In their algorithm, the complexity values of a time series is evaluated based on a ''coarse-graining'' operation by which the data sequence is transformed into a pattern of only a few symbols, for example, 0 and 1 and involves data sequence comparison and number counting in one dimension only. A flow chart of the LZC algorithm is shown in **Figure 15** and the process is described below:

1. Coarse-grain process to a finite binary sequence. A discrete time series $A = \{a_1, a_2, \cdots, a_n\}$ is converted into a binary sequence $S_N = \{s_1, s_2, \cdots, s_n\}$ in the initiation and preprocessing phase.

2. Copy and Insert. A binary sequence up to $s_r (1 < r < N)$ of complexity c_N can be reconstructed by simply copying and inserting some of the existing vocabulary of $SQv_r = \{s_1, s_2, \cdots, s_v\}$ $(v < r)$. To check the rest string $S_{N-r} = \{s_{r+1}, \cdots, s_N\}$ can be reproduced by the same approach, the process is executed by the following steps:

   Step 1: Take $Q_r = \{s_r\}$ and check whether this string belongs to the vocabulary of SQv_r. If so, string $Q_r = \{s_r\}$ is a simple repetition of an existing substring of SQv_r (i.e., a simple "copy" of the existing vocabulary can restore it) and the complexity remains unchanged or $c_N(r) = c_N(r-1)$.

   Step2: Read the next string and take $Q_{r+1} = \{s_r, s_{r+1}\}$. Check if $Q_{r+1} = \{s_r, s_{r+1}\}$ belongs to SQv_{r+1}.

   Step 3: If string Q_{r+1} does not belongs to SQv_{r+1}, increase the complexity by one, i.e., $c_N(r+1) = c_N(r) + 1$, nullify $Q_{r+1} = \{\}$, read the next string and take $Q_{r+3} = \{s_{r+3}\}$.

http://dx.doi.org/10.5772/67143

Step 4: Repeat the above process until $S_N = \{s_1, s_2, \cdots, s_n\}$ is covered. The resulting $c_N(n)$ is the complexity of a given string.

3. Normalization of the complexity value. The complexity obtained above equals the number of nullification of Q. It indicates that the complexity is affected by the length of the string, or the number of data sample N. To find a robust complexity measure, Lempel and Ziv [27] suggested a normalized measure $c_{LZ} = \frac{c_N(n)}{b_N(n)}$, which is termed as LZC after their names and defined by

$$0 \leq c_{LZ} = \frac{c_N(n)}{b_N(n)} \leq 1 \tag{52}$$

where $b_N(n) = \lim_{n\to\infty} c_N(n) \approx \frac{n}{\log_k n}$.

It is shown in **Figure 15** that the coarse-grained process of a finite time sequence serves as the basis for the LZC algorithm. Commonly employed technique in coarse-grained process is a single-scale process which converts a discrete time series $A = \{a_1, a_2, \cdots, a_n\}$ into a symbolic binary sequence $S_N = \{s_1, s_2, \cdots, s_n\}$ using the following formula:

$$s_i = \begin{cases} 1, & a_i \geq \overline{a} \\ 0, & a_i < \overline{a} \end{cases} \tag{53}$$

where $\overline{a} = (a_1 + a_2 + \cdots + a_n)/n$ is the mean value of the discrete time series which is used as a threshold in the coarse-graining preprocess.

In the single-scale coarse-grain process, all segments of a discrete time series larger than or equal to the mean value are set to 1 while all segments smaller than the mean value are set to 0. The process neglects the fluctuation between the data intervals and loses many detailed information of the data series during the process. In order to capture the details contained in a discrete time series, the division interval should be reduced and a multidivision scale should be adopted so that the variation in the data can be reflected in the binary sequence at a multiscale process. The preprocess of a revised multiscale coarse grain is outlined below:

1. Divide the discrete time series into various scales. A two-scale division process is used here as an example. In this process, a discrete time series is divided into two regions, with the mean value of each region as the boundary. The same approach can be applied to a multiscale division where a discrete time series can be divided into several regions.

2. When the first number in the discrete time series is larger than the mean value of the entire discrete time series, this point is set to 1 or 0 vice versa. Starting from the second data point of the discrete time series, the binary value is determined by comparing its value with the value of the previous point in the discrete time series. If the two points are in the same division interval, the binary value of the latter point will be the same as the previous one. When the value of the latter point increases to another division interval, the point will

be assigned to the binary value of 1. When the value of the latter point decreases into another division interval, the point will be assigned to the binary value of 0.

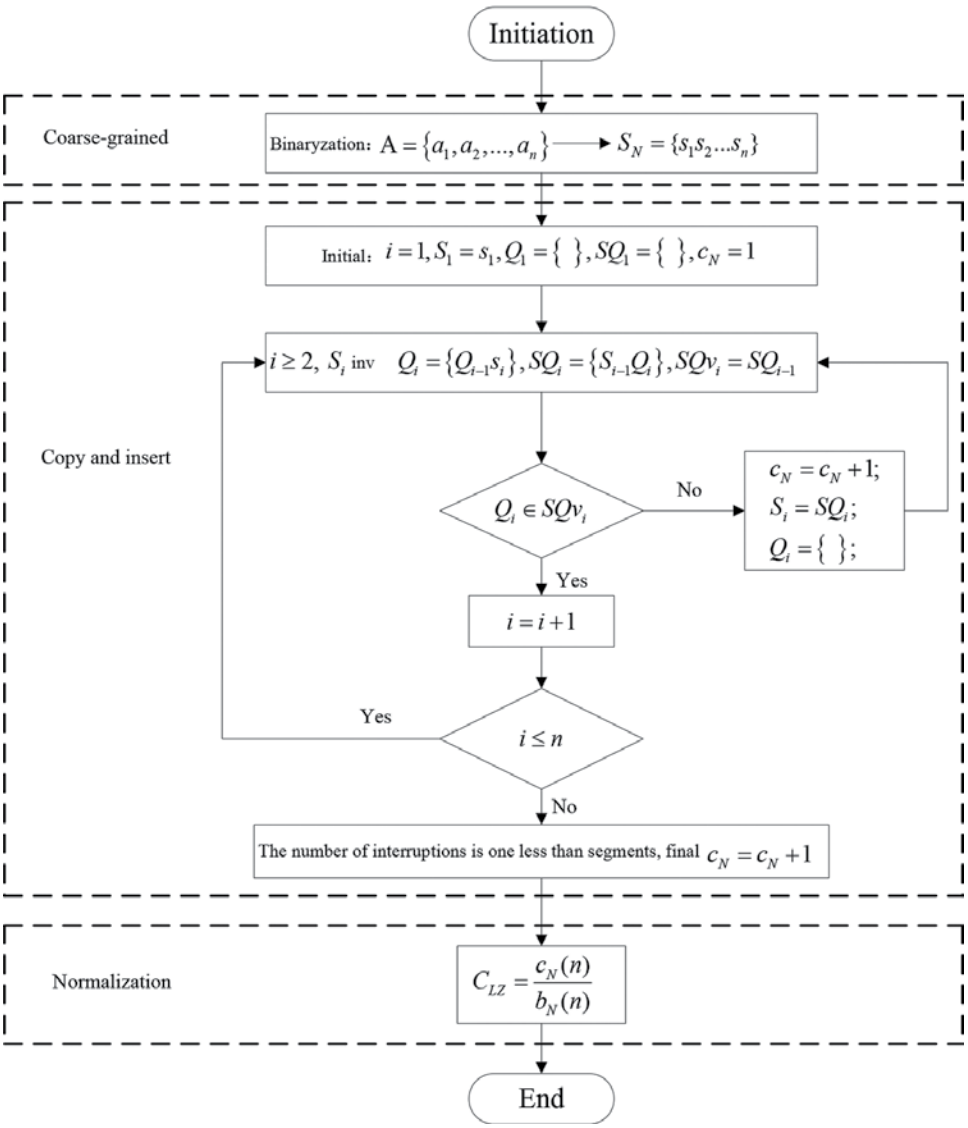


Figure 15. Flow chart of the LZC algorithm.

A data sample given by **Figure 16** is used here as an example to illustrate the difference between the single-scale and multiscale coarse-grain process. For single-scale coarse-grain process, the binary sequence of the data sample is $S_N = (1,1,1,1,1,1,0,0,0,0,0,0)$ which does not reflect the fluctuation between the data interval of the time series. When a two-scale coarse-grain process is employed to process the same data, the binary sequence becomes S_N=(1,1,0,1,0,1,0,1,0,1,0,1). Therefore, a multiscale coarse-grain process can better capture the detailed fluctuation in a discrete time series.

http://dx.doi.org/10.5772/67143

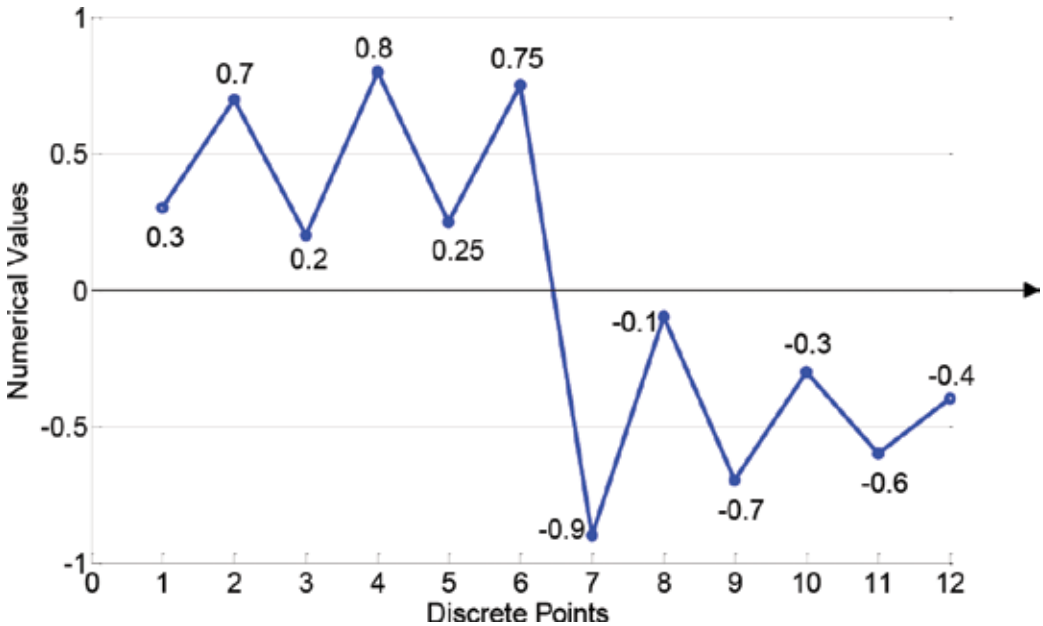


Figure 16. An example of a discrete time series.

3.4.4. Manifold learning

Classical dimensionality reduction techniques, such as multidimensional scaling (MDS) and independent component analysis (ICA), are only applicable to linearly structured data sets but not suitable for high-dimensional, nonlinear data sets such as bearing CM data. To overcome this problem, a nonlinear dimensionality reduction technique, manifold learning, has recently developed for machine fault diagnosis [28]. Manifold learning technique projects the original high-dimensional data onto a lower-dimension feature space while preserving the local neighborhood structure to detect the intrinsic structure of nonlinear high-dimensional data. Manifold learning can be realized through several algorithms including locally linear embedding (LLE), isometric feature mapping (IsoMap), local tangent space alignment (LTSA) and local preserving projection (LPP).

The application of manifold learning in mechanical fault diagnosis can be in twofolds. Firstly, fault features with a large dimension can have many redundant components, which can increase the complexity and operation time of a fault diagnosis process. Manifold learning can be used to eliminate the redundant components and extract the nonlinear features for fault classification in this case. Secondly, manifold learning can discard the noise components and extract the intrinsic manifold features related to nonlinear dynamic of a CM signal; therefore, it can also be used as a de-noise technique. It should be noted that manifold learning only operates on a matrix, so a preprocessing such as a reconstruction of phase space converting an original one-dimensional signal into a two-dimensional data is required. For a detailed information of the manifold learning algorithm and its application in fault diagnosis, interested readers are referred to [28, 29].

3.5. Bearing fault diagnosis based on artificial intelligent

3.5.1. Shallow architecture machine learning

The most commonly employed machine-learning techniques in fault diagnosis of rotating machines are hidden Markov models (HMMs), support vector machines (SVMs) and artificial neural networks (ANNs), which exploit shallow architectures either contain a single hidden layer or without a hidden layer. Such shallow architecture-based models have achieved great success both in theory and in practical applications. Though, shortcomings of these algorithms such as poor universality, lacking of theory basis in parameter selection, easier to fall into a local optimum value have limited the application of the algorithms in machine fault diagnosis.

3.5.2. Deep neural network

A consensus criterion for an effective bearing fault diagnosis technique is that it should not only be able to identify various bearing fault conditions but also be able to discriminate different fault severities in each fault condition [30]. This leads to a stricter requirement on identification procedures where a classifier must have a greater capability in discriminating different fault classes. When fault data contain more than one level of fault severities in each fault condition, the accuracy of fault diagnostic result using shallow architecture classifiers described in the previous section will reduce dramatically. Hinton and Salakhutdinov [31] proposed a deep learning technique to overcome the deficiencies of single-layer architecture classifiers for a better pattern recognition capability. The concept is further extended to become a so-called deep belief networks (DBNs) [32] which relieves the training difficulties of deep network structures by adopting a layer-by-layer unsupervised forward pretraining learning and then back fine-tuning mechanism. A description of the training process of a DBN is given in **Figure 17**.

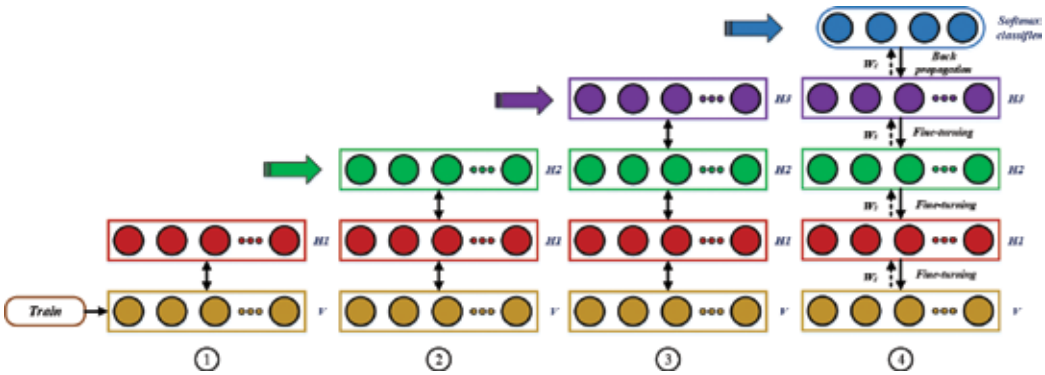


Figure 17. A description of the training process of a DBN.

Bengio et al. [33] proposed a deep stack auto-encoder (SAE) network by using a network structure similar to DBNs which is stacked with a number of auto-encoder networks. Furthermore, Le Cun et al. [34] proposed a convolutional neural network (CNN), a multilayer network where each layer is composed of several two-dimensional planes, to reduce the number of parameters in the learning process using a unique weight-sharing mechanism. The above cited deep-learning-based neural network algorithms have been widely employed in machine fault diagnosis nowadays [35, 36].

http://dx.doi.org/10.5772/67143

3.5.3. A case study on bearing fault diagnosis

Case 5: In this case study, a combination of a four-level wavelet packet decomposition (WVD), a locality preserving projection (LPP), a particle swarm optimization (PSO) and support vector machine (SVM) algorithms are employed for an intelligent bearing fault diagnosis and recognition. Bearing condition-monitoring data on various operation conditions such as healthy bearing, outer race fault, inner race fault and ball fault acquired from a bearing fault simulation test-rig as shown in **Figure 18** are used in this analysis.

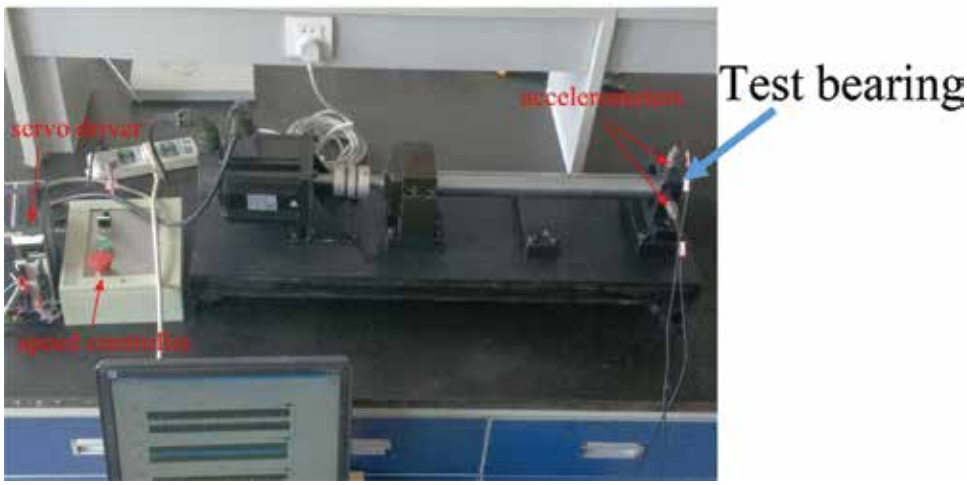


Figure 18. A graphical illustration of the bearing fault simulation test rig.

Three types of bearing defects are simulated in this experiment: an outer race fault, an inner race fault and a ball fault as shown in **Figure 19**. The measured vibration signals from an accelerometer mounted on the test bearing house for the four bearing operation conditions (the three simulated faults and a healthy bearing) are shown in **Figure 20**. Hundred data sets are acquired in the experiment which are divided into two groups: one contains 70 sets of data and is used as the training data set and the other contains 30 sets of data which is used as the test data set.

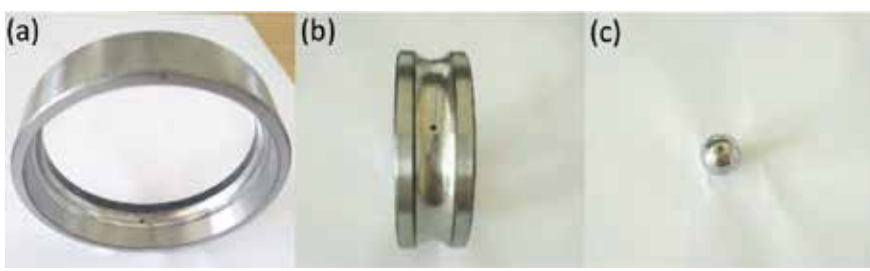


Figure 19. Simulated bearing faults: (a) outer race fault; (b) inner race fault; and (c) ball fault.

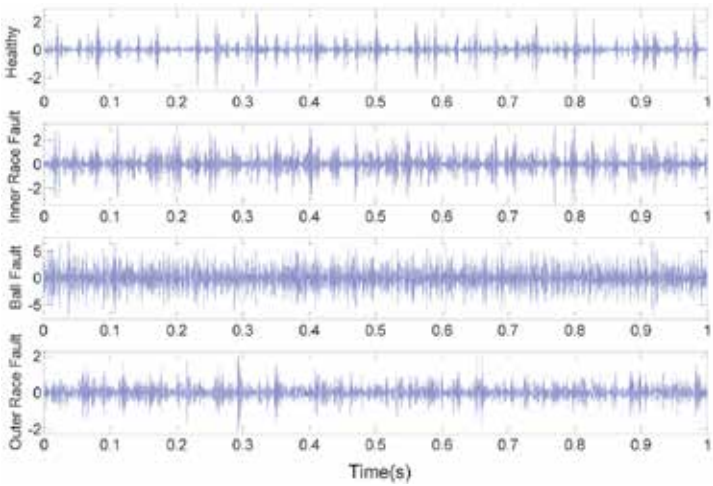


Figure 20. The measured vibration signals for the four bearing operation conditions. Note, sampling frequency, 8192, data length, 8192.

Five major steps are taken in the analysis of the bearing condition-monitoring data:

1. Each data is decomposed by a four-level wavelet packet decomposition leading to 16 components of different frequency contents (mutual orthogonal subspaces). The wavelet components of the condition-monitoring data for the outer race fault are shown in **Figure 21** for illustration.

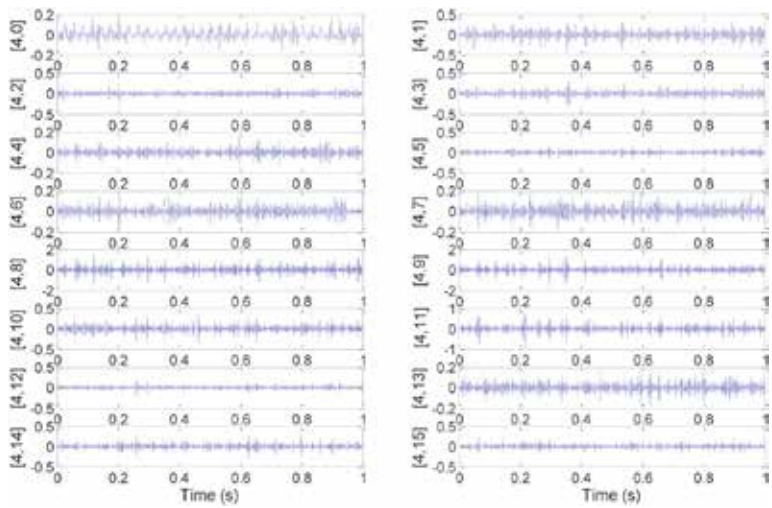


Figure 21. The wavelet components of the outer race bearing defect signal.

2. The energy value for each wavelet component is calculated and is used as the fault feature since the energy of the component can discriminate different classes and contains the fault information and its fluctuation in the particular component corresponding to the occurrence of the fault. A fault feature vector containing 16 energy features can be obtained for

http://dx.doi.org/10.5772/67143

each data. The energy distribution of the 16 wavelet components for the four bearing operation conditions is shown in **Figure 22**.

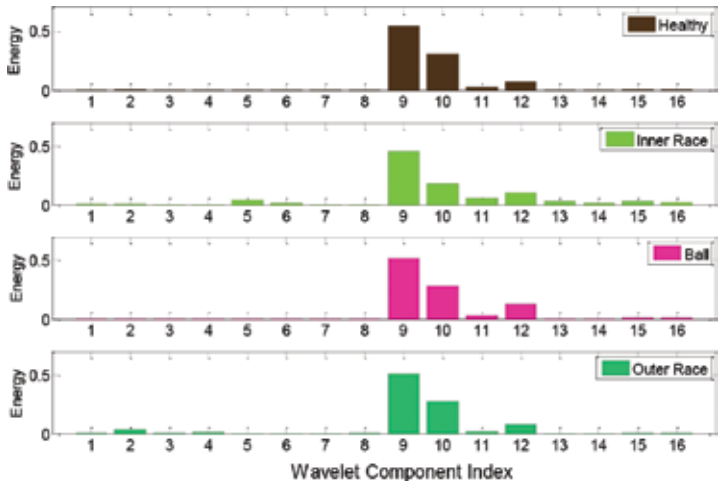


Figure 22. The energy values of the 16 wavelet components for the four bearing operation conditions.

3. For the 100 data sets (of four bearing operation conditions), the process in Step (2) will generate a 400×16 feature set. The dimension of the large feature set can cause a problem for the following algorithm leading to misclassification of the bearing fault classes in the diagnosis step using SVM algorithm. The locality preserving projection (LPP) [37], a linear dimensionality reduction algorithm, which can effectively reduce the dimension while preserving the neighborhood structure of a data set, is utilized to reduce the feature set dimension to 400×3. The three-dimensional feature distribution for the four bearing operation conditions after the dimensional reduction by LPP is shown in **Figure 23**.

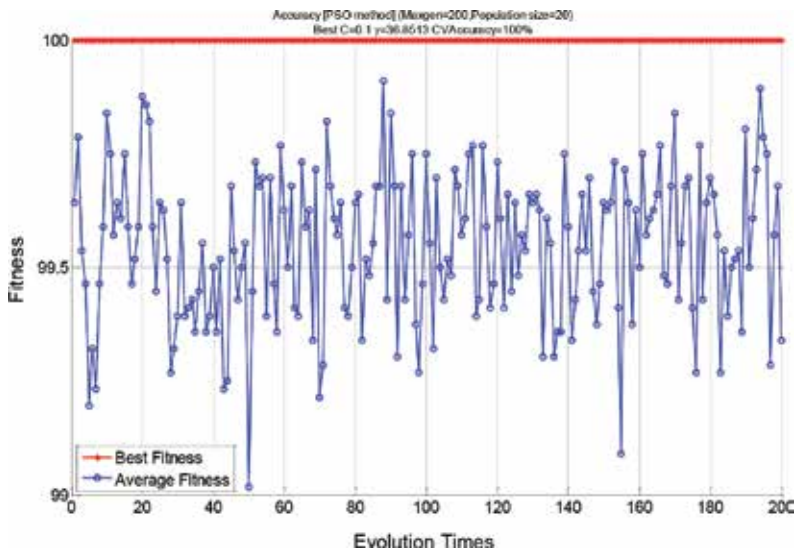


Figure 23. The optimization result of the PSO algorithm.

4. Particle swarm optimization (PSO) is adopted in this step to find the optimal kernel parameters and penalty factor used in SVM algorithm from the training data set to enhance the performance of the SVM algorithm. The optimization result of the PSO is shown in **Figure 23** where the fitness function of the optimization is above 99% (close to zero misclassification rate) throughout the time evolution progress.

5. Training and prediction (recognition) of the bearing experimental data using the SVM algorithm. The prediction result by the SVM is shown in **Figure 24**.

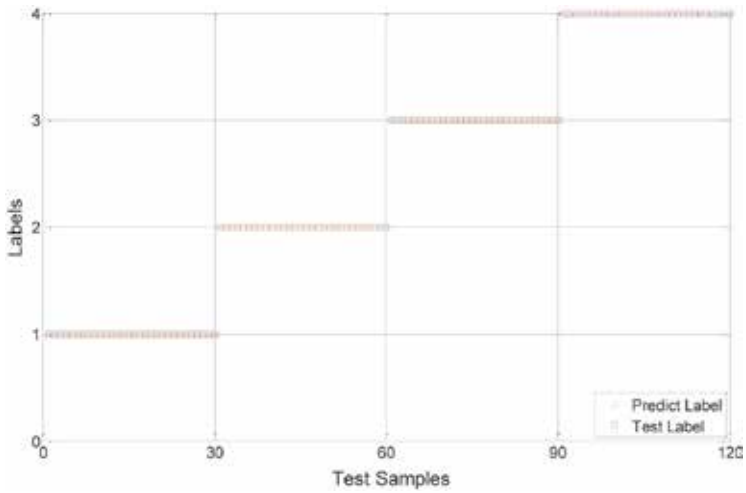


Figure 24. Prediction result of the SVM algorithm. Note, Levels 1–4 correspond to healthy bearing, inner race fault, ball fault and outer race fault, respectively.

4. Conclusion

This chapter presented a comprehensive, step-by-step approach on condition monitoring of roller element bearings aiming to provide an introductory and referencing material for engineers and researchers new to the field. It summarized the most frequently employed data acquisition, signal analysis, feature and parameter extraction and fault diagnosis techniques in the current practice. Pros and cons of each technique are briefly discussed in the text. The formulation and discussion are also supported by ample tables, graphs and figures throughout the text for a better illustration and understanding of how to utilize the techniques presented in the chapter for real-life problems.

Author details

Tian Ran Lin*, Kun Yu and Jiwen Tan

*Address all correspondence to: trlin888@163.com

School of Mechanical Engineering, Qingdao University of Technology, Huangdao District, Qingdao, PR China

http://dx.doi.org/10.5772/67143

References

[1] T. R. Lin, E. Kim and A. C. C. Tan, "A simple signal processing approach for condition monitoring of low speed machinery using Peak-Hold-Down-Sample algorithm", Mechanical Systems and Signal Processing 36(2), 256–270, 2013.

[2] T. R. Lin, W. Wu and A. C. C. Tan, "A signal processing approach to solve the non-linearity problem of acoustic emission sensors", Proceedings of the 9th World Congress on Engineering Asset Management, 28–31 Oct 2014, Pretoria, South Africa.

[3] W. Wu, T. R. Lin, A. C. C. Tan, "Normalization and source separation of acoustic emission signals for condition monitoring and fault detection of a multi-cylinder diesel engine", Mechanical Systems and Signal Processing 64–65, 479–497, 2015.

[4] F. Kadri, F. Harrou, S. Chaabane, Y. Sun, C. Tahon, "Seasonal ARMA-based SPC charts for anomaly detection: Application to emergency department systems", Neurocomputing 173, 2102–2114, 2016.

[5] R. A. Wiggins, "Minimum entropy deconvolution", Geoexploration, 16, 21–35, 1978.

[6] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial", Mechanical Systems and Signal Processing 25, 485–520, 2011.

[7] J. Y. Lee and A. K. Nandi, "Extraction of impacting signals using blind deconvolution", Journal of Sound and Vibration **232** (5), 945–962, 1999.

[8] J. Antoni, "The spectral kurtosis: a useful tool for characterizing non-stationary signals", Mechanical Systems and Signal Processing 20 (2), 282–307, 2006.

[9] J. Antoni, R. B. Randall, "The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines", Mechanical Systems and Signal Processing 20 (2), 308–331, 2006.

[10] Y. Wang, J. Xiang, R. Markert, M. Liang, "Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications", Mechanical Systems and Signal Processing 66–67, 679–698, 2016.

[11] Y. G. Lei, J. Lin, Z. J. He and Y. Zi, "Application of an improved kurtogram method for fault diagnosis of rolling element bearings", Mechanical Systems and Signal Processing 25, 1738–1749, 2011.

[12] X. Zhao, B. Ye, "Similarity of signal processing effect between Hankel matrix-based SVD and wavelet transform and its mechanism analysis", Mechanical Systems and Signal Processing 23, 1062–1075, 2009.

[13] R. B. Randall, "Vibration-based condition monitoring", John Wiley & Sons, West Sussex, 2011.

[14] K. C. Chua, V. Chandran, U. R. Acharya, C. M. Lim, "Application of higher order statistics/spectra in biomedical signals—A review", Medical Engineering & Physics 32, 679–689, 2010.

[15] Z. K. Peng and F. L Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnosis: A review with bibliography", Mechanical Systems and Signal Processing 18, 199–221, 2004.

[16] Y. Yin, X. Hu and T. R. Lin, "A practical approach to analyze the non-stationary signals of a quayside container crane motor using a combined empirical mode decomposition and wavelet packet quantization", Noise Control Engineering Journal 64(2), 126–133, 2016.

[17] N. E. Huang, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", Proceedings of the Royal Society, Series A: Mathematical, Physical and Engineering Sciences 454, 903–995, 1998.

[18] Y. Lei, J. Lin, Z. J. He, M. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery", Mechanical Systems and Signal Processing 35, 108–126, 2013.

[19] Z. H. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method", Royal Society of London Proceedings: Mathematical, Physical and Engineering Sciences 460(2046), 1597–1611, 2004.

[20] Z. H. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method", Advances in Adaptive Data Analysis 1(1), 1–41, 2011.

[21] C. E. Shannon, "A mathematical theory of communication", Bell System Technical Journal 27, 379–423, 1948.

[22] H. Cui, L. Zhang, R. Kang and X. Lan, "Research on fault diagnosis for reciprocating compressor valve using information entropy and SVM method", Journal of Loss Prevention in the Process Industries 22, 864–867, 2009.

[23] S. Pan, T. Han, A. C. C. Tan and T. R. Lin, "Fault diagnosis system of induction motors based on multiscale entropy and support vector machine with mutual information algorithm", Shock and Vibration 2016, Article ID 5836717, 2016.

[24] Y. Tian, Z. Wang, C. Lu, "Self-adaptive bearing fault diagnosis based on permutation entropy and manifold-based dynamic time warping" (In Press), Mechanical System and Signal Processing.

[25] J. Ziv and A. Lempel, "Compression of individual sequences via variable-rate coding", IEEE Transactions on Information Theory 24(5), 530–536, 1978.

[26] H. Hong and M. Liang, "Fault severity assessment for rolling element bearings using the Lempel-Ziv complexity and continuous wavelet transform", Journal of Sound and Vibration 320, 425–468, 2009.

[27] A. Lempel and J. Ziv, "On the complexity of finite sequences", IEEE Transactions on Information Theory 22(1), 75–81, 1976.

[28] B. Tang, T. Song, F. Li, L. Deng, "Fault diagnosis for a wind turbine transmission system based on manifold learning and Shannon wavelet support vector machine", Renewable Energy 62, 1–9, 2014.

[29] H. S. Seung and D. D. Lee, "The manifold ways of perception", Science 5500(290), 2268–2269, 2000.

[30] M. Gan, C. Wang and C. Zhu, "Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings", Mechanical Systems and Signal Processing 72–73, 92–104, 2016.

[31] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks", Science 313(5786), 504–507, 2006.

[32] D. Yu and L. Deng, "Deep learning and its applications to signal and information processing", IEEE Signal Processing Magazine 28(1), 145–154, 2011.

[33] Y. Bengio, P. Lamblin, D. Popovici and H. Larochelle, "Greedy layer-wise training of deep networks", Advances in Neural Information Processing Systems 19, 153–160, 2007.

[34] Y. Le Cun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and L. D. Jackel, "Handwritten digit recognition with a back-propagation network", Advances in Neural Information Processing Systems 2, 396–404, 1990.

[35] Z. Chen, C. Li and R. Sanchez, "Gearbox fault identification and classification with convolutional neural networks", Shock and Vibration 2015, Article ID 390134, 2015.

[36] W. Sun, S. Shao, R. Zhao, R. Yan, X. Zhang and X. Chen, "A sparse auto-encoder-based deep neural network approach for induction motor faults classification", Measurement 89, 171–178, 2016.

[37] X. He and P. Niyogi, "Locality preserving projections", Advances in Neural Information Processing Systems 16 (NIPS 2003), Vancouver, Canada, 2003.

Chapter 4

Design Aspects of the Bearing Supports

Ismagilov Flur Rashitovich,

Vavilov Vyacheslav Evgenievich and D.V. Gusakov

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/67144

Abstract

This chapter examines different types of bearing supports. Technical parameters of different types of bearing supports are presented. The effectiveness of some types of bearings is determined. General approach for the calculation of bearing overall dimensions is considered.

Keywords: bearing supports, high-speed electrical machines, electromagnetic processes

1. Introduction

One of the main problems in the design of high-speed electrical machines (HS EM) is the task of selecting the bearing assemblies that forms design and determines its application area, allowable load, and efficiency.

The complexity of this task lies in the fact that the high-speed electromagnetic bearing assemblies must meet various criteria, which often contradict each other. So, bearing supports of HS EM should ensure minimum friction losses and maximum resource, wide operating temperature range (which is typical for non-contact bearings, and almost impossible to achieve on the mechanical bearings), but they must have a minimum ductility (maximum stiffness) for sub-critical rotor speeds and rotor dynamics requirements, providing significant mechanical stress and have a minimum weight and overall dimensions (it is ensured well enough mechanical bearing assemblies and is difficult to achieve on a contactless bearing supports).

Therefore, mechanical (ball and roller), hydrostatic and gas bearings, as well as various types of magnetic bearings applied in modern HS EM. The choice of bearings depends on the specific tasks and function of HS EM.

It is advisable to consider the selection criteria in more detail before analyzing the advantages and disadvantages of various types of bearings.

Bearing stiffness is a value that is characterized by an elastic deformation of the bearing under load. It is expressed as the ratio of the load to the elastic deformation, depending on the type, design and size of the bearing. In simplified form, the bearing stiffness can be defined as follows:

$$k_x = -\frac{F}{\delta}, \tag{1}$$

where F—load acting on the bearing; δ—change in the bearing gap under the load; k_x—bearing stiffness.

Typically, the stiffness is defined in the technical catalogs for the bearing supports.

The so-called stiffness background is used more often rather than bearing stiffness when calculating bearing supports in EM:

$$k_{x1} = \frac{k_x}{LD}, \tag{2}$$

where L—bearing length; D—bearing diameter.

Also used the damping coefficient attributable to the area of the bearing support:

$$c_{x1} = \frac{c_x}{LD}. \tag{3}$$

The static load is load acting on the bearing when the rotor is stationary and dynamic load is the load exerted on the bearing with a rotating rotor.

Bearing speed is a technical parameter that determines the maximum speed of the bearing. Bearing speed is measured in mm × rpm/min and defined as follows:

$$DN = D \cdot n, \tag{4}$$

where n—rotor rotational speed; D—bearing diameter.

The main producers of high-speed mechanical bearings are FAG, SKF, GMN and NTR companies.

High-speed bearings of SKF are made in accordance with ISO 683 ⟨⟨Heat-treated steels, alloy steel sand-free-cutting steels—Part 17: Ball and roller bearing steels⟩⟩ and presented in the N10 series. Under the conditions of the liquid lubrication of the bearings, rotation speed can be achieved 40,000 rpm and can be used at temperatures from −40 to + 150°C.

Rotor rotational speed of HS EM on the FAG bearings with oil lubrication can reach 170,000 rpm. In this mode, the bearings temperature is within the range from −40 to +150°C [1].

GMN Company produces mechanical bearings with speed limit of 75,000 rpm and its temperature limit corresponds to the analogues presented above [2].

Undoubtedly, mechanical bearings have reached significant technical heights. However, they have inherent weaknesses such as limited speed, considerable noise emission and low operating temperature.

http://dx.doi.org/10.5772/67144

HS EM on the mechanical rolling bearings is characterized by thermal deformation; trajectory instability is caused by a change in the rotation angle of the separator with a set of rolling elements and a manufacturing error of the mechanical support rings, as well as limited service life determined by mechanical friction between the dynamic rotating parts. Therefore, for a more promising use in high-speed and high-temperature, EM have a contactless bearing supports: magnetic [active magnetic bearings (AMB), hybrid magnetic bearings (HMB)] or gas [aerodynamic bearings (ADB) or air bearings].

2. Bearings supports types

2.1. Air bearings (AB)

AB is a slide bearings (according to Standard ISO 4378-1-2001) in which the lubricating membrane pressure is created by the gas supply system. AB operating principle is based on the injection of air through a system of holes under pressure into the gap between the pin and the bearing. At the same time, the pin is separated by a layer of pressurized air from bearing. They are not used in HS EM due to the fact that the air bearings require additional pressurization system compressor.

2.2. Aerodynamic bearings (ADB)

Aerodynamic bearings (ADB) is the sliding bearings (according to ISO 4378-1-2001), in which the lubricant membrane pressure, and hence load bearing capacity is created by the surface movement. The operating principle of the ADB is that in the absence of rotation of the pin rests on the inner surface of the bearing, while rotating air or other gas is sucked from the environment, creating an air cushion with increased pressure, thus, lifting the pin and separating it from the bearing (**Figure 1**).

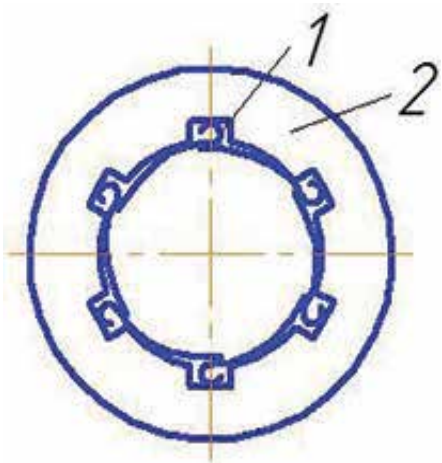


Figure 1. Aerodynamic bearing: 1—trunnion; 2—foil.

Table 1 shows the characteristics of radial ADP made by Russian production (produced by National Research University "Moscow Power Engineering Institute"). **Table 2** shows the axial ADB made by Russian production.

The advantages of ADB is the absence of necessity for a control system (as compared with the electromagnetic bearings), as well as their noncontact (compared to mechanical). The

Bearing type	The nominal diameter of the pin (mm)	The axial length (mm)	The recommended maximum speed (rpm)	The static load-bearing capacity (*N*)		The frequency of surfacing (rpm)
				Usual scheme	Enhanced scheme	
FGB11	10.5	13	364,000	2	–	19,000
FGB14	13.5	16	283,000	3	–	14,800
FGB16	15.5	18	247,000	4	–	13,000
FGB20	19.5	24	196,000	7	–	10,000
FGB30	30	34	127,000	15	–	6700
FGB35	35	31	109,000	16	27	5700
FGB40	39	44	98,000	25	42	5100
FGB61	61	70	63,000	63	105	3300
FGB67	67	70	57,000	69	115	3000
FGB74	74	70	52,000	76	127	2700
FGB80	80	70	48,000	82	137	2500
FGB84	84	85	46,000	105	175	2400
FGB103	103	70	37,000	–	177	1900
FGB103l	103	120	37,000	–	303	1900

Table 1. Radial ADB made by Russian production.

Type	The diameter of the heel (mm)	The outer diameter of the bearing (mm)	The inner diameter of the bearing (mm)	Rated speed (rpm)	Bearing capacity at rated speed (N)
TFGB37	37	43	19	207,000	95
TFGB44	44	49	22	174,000	137
TFGB64	64	74	34	119,000	277
TFGB72	72	82	42	106,000	322
TFGB85	85	95	52	90,000	426
TFGB105	95	116	93	75,000	702
TFGB120	120	132	70	64,000	895

Table 2. Axial ADB made by Russian production.

disadvantage is that ADB provides noncontact rotation of the rotor only with a certain speed (rate surfacing), and up to this frequency ADB acts as a mechanical bearing of a high friction (for example, from **Table 1** it is seen that the bearing frequency surfacing is 2400 rpm with a load capacity of 105 N). Furthermore, using ADB is an increased requirement for the surface treatment of the shaft. Also ADB cannot be operated in the absence of a gas environment, such as a vacuum, which limits their use in cosmic space.

2.3. Active magnetic bearings (AMB)

AMB (according to ISO 14839-1-2011) is a rotor maintenance device without mechanical contact by the magnetic attraction forces and uses feedback servo, in which the circuit typically contains sensors, solenoids, power amplifiers, power supplies and the controller (**Figure 2**).

AMBs are widely used in Russian and foreign industry [in Russia are engaged in the development of the ⟨⟨VNIIEM Corporation⟩⟩ JSC and ⟨⟨Pskov engineering company⟩⟩, among foreign manufacturers can mark SKF, CalnetixTechnologies (USA), the Synchrony (USA) and others.].

The advantages of the AMB are their features such as controllability, contactless operation, providing rotor levitation when power is supplied to the control electromagnets (unlike ADB), the ability to work at high temperatures and in corrosive environments, bearing stiffness control possibility (due to pulse changes the electromagnetic force) and the bearing damping ability and the ability to work in vacuum.

The AMB disadvantages include the complexity of their design, the complexity of their control systems, significant product price and their high weight and overall dimensions. The stiffness of the AMB under normal operating conditions is comparable or slightly higher than the stiffness of the ADB.

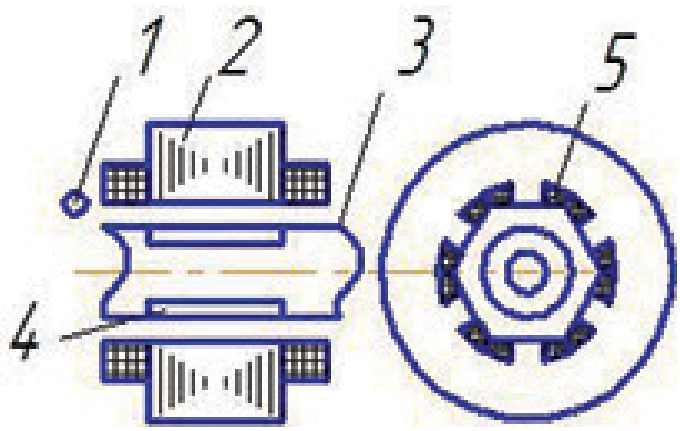


Figure 2. AMB: 1—rotor position sensor; 2—AMB magnetic core; 3—shaft; 4—ferromagnetic sleeve; 5—AMB winding.

Despite these disadvantages, the AMB are widely used in HS EM. Moreover, the use of AMB in the Russian Federation is rationed by the technical documentation (ISO 14839-1-2011, ISO 14839-2-2011, ISO 14839-3-2013, ISO 14839-4-2014).

Importantly, the AMB are not only electromagnets, in which the shaft is concentrically located but an intellectual complex system consisting of microscopic sensors, signal amplifiers, etc. A more complete design of AMB control systems, as well as their control algorithms is described in Refs. [3, 4].

Table 3 shows the geometric dimensions of the AMB, produced by ⟨⟨Pskov Engineering Company⟩⟩.

To evaluate the effectiveness of the AMB and ADB energy characteristics, it is advisable to make a comparison on the specific speed and static load, which is accepted in the form:

$$F_{sp} = \frac{F}{DL} \tag{5}$$

ADB and AMB of Russian production are considered when comparing.

From **Table 3**, it is seen that when the static load-bearing capacity is 180 N, the specific speed of AMB is 3,750,000 rpm and the specific static load is 100,000 N/m^2. At the same time at the static load of 175 N, specific speed is 3,864,400 rpm and the specific static load is 24,509 N/m^2. That is, the AMB of Russian production exceeds ADB by the specific static load, and the specific speed of both variants is about the same (AMB's specific speed at 2.4% less than ADB).

d (mm)	*D* (mm)	*L* (mm)	*n* (10^3 rpm)	*F* (N)	*m* (kg)
15	44	14	252	20	0.07
20	52	16	190	30	0.12
25	58	20	150	50	0.18
30	66	24	125	70	0.3
35	72	27	110	90	0.4
40	80	30	95	120	0.52
50	94	36	75	180	0.84
60	110	42	63	250	1.32
70	130	46	54	360	2
80	148	50	47	450	2.7

Notes: d—diameter of the shaft; D—external diameter; m—mass of the AMB; L—active length; F—static load-bearing capacity; n—permissible speed.

Table 3. Standards of ⟨⟨Pskov engineering company⟩⟩ for radial AMB.

To eliminate the AMB and ADB disadvantages, hybrid magnetic bearings (HMB) are applied in HS EM. HMB is the bearing that combines AMB design and magnetic bearing at permanent magnets (MB PM) in accordance with ISO 14839-1-2011.

At the same time, as shown in Ref. [5], the concept of the HMB goes far beyond the definition of an ISO and represents a combination of different bearing types in a single product that allows them to combine their design merits and ADB, and AMB, reaching thus the minimum weight and overall dimensions, controllability and stability of the entire HS EM.

There are three main types of structural HMB: gas-magnetic bearing, magnetomechanical and various combinations of MB PM with AMB.

2.4. HMB, as a combination of MB PM and AMB

This type of HMB is the most common and used in practice. Moreover, it is considered the most promising design of HMB. This area has two main ways of development: the permanent magnets are installed in the magnetic AMB (**Figure 3**) to increase the magnetic flux. Separation of the AMB and MB PM, for example, two radial MB PM placed on one shaft, and the rotor axial fixation is provided by axial AMB (**Figure 4**).

A significant pulling force of the electromagnet is required when using the second option, so the first design most widely used in industry. At the same time some technical branches of second design has broad application prospects.

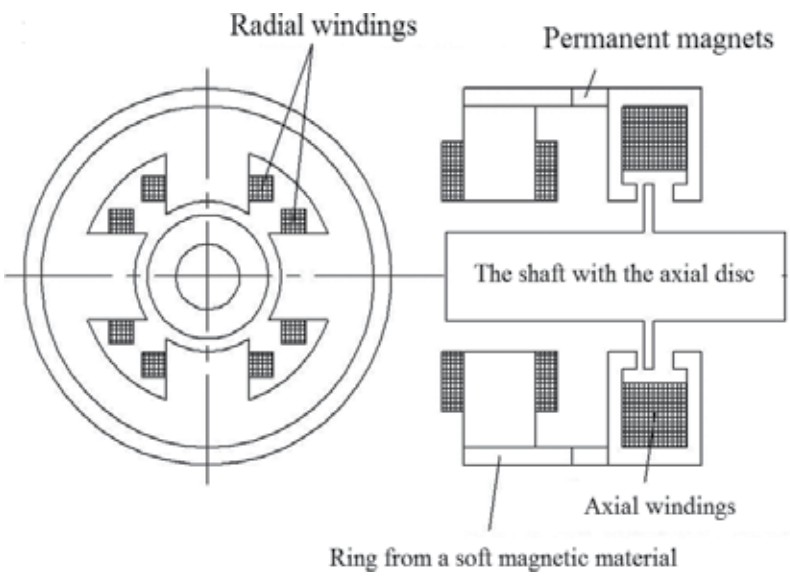


Figure 3. Radial-axial HMB, in which PM are used to amplify the magnetic flux.

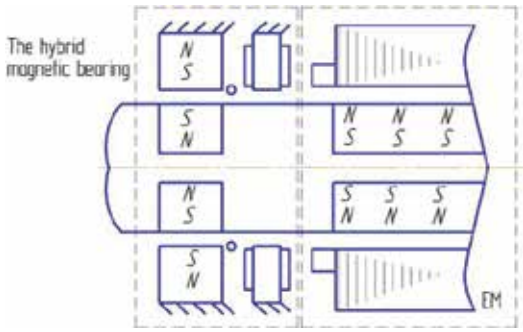


Figure 4. HMB where AMB and MB PM used separately.

2.5. Magnetomechanical HMB

This HMB class is a combination of mechanical bearings, which serve as the main shaft support and MB PM, which are intended for unloading mechanical bearings. The advantages of this HMB type is the lack of a control system and simplicity of design and the disadvantages is the presence of mechanical bearings friction, and consequently also their low reliability.

For example, it is known that magnetomechanical bearing (MMB) design [6] for the electromechanical battery consists of flywheel and a high-speed electric generator with a vertical shaft. A feature of this design is the use of a ball in HMB made of sapphire, which provides the axial support system. **Table 4** shows the effectiveness of different ball materials and plate in the MMB.

To improve the efficiency of MMB in rotary system of the HS EM the passive vibration damper also entered besides mechanical bearings and MB PM, which is needed for damping vibration energy. A passive vibration damper is an electrically conductive plate installed with a gap relative to the PM. Eddy currents are induced in the copper sleeve with displacements of the PM, which provide damping of vibration energy.

MMB are actively developing due to their simple application design. The major trends in the development of this type of HMB are reduction of friction in the mechanical bearings by the use of coatings and materials, as well as by the maximum discharge of mechanical bearings

Ball material	Material plate	Friction coefficient	Friction losses at 50,000 rpm (MW)
Sapphire	Sapphire	0.1	152
Steel	Steel	0.42	628
Cast iron	Cast iron	0.15	230
Teflon	Steel	0.04	63

Table 4. The effectiveness of different ball materials and plate in the MMB.

and levitating shaft vibration reduction. It is obvious that in a number of industries, especially in high-speed systems with short life cycles, the HMB type have broad prospects.

2.6. Gas-magnetic HMB

This HMB is a combination of ADB and AMB. **Figure 5** shows a design of this HMB type [7].

The advantages of this type of HMB include high stiffness and handling, but they have considerable design complexity of execution, so they are not widely used in the industry. Gas-magnetic HMB are considered in Refs. [8–10] in more detail.

2.7. Electrostatic bearings

At low mass of the rotor, as well as to the possibility of providing vacuum in the cavity of the EM, it seems appropriate to use electrostatic poles. Electrostatic support is a noncontact bearing assembly, in which efforts are created by attractive forces between two surfaces having different potentials (**Figure 6**). The created ascensional power in electrostatic supports is insignificant and is accepted in the form:

$$f = \frac{\varepsilon E^2}{2} \tag{6}$$

where ε—the dielectric constant of the suspended body; E—the electric field strength.

The advantage of the electrostatic poles relates primarily to no energy losses due to eddy currents. The electrostatic bearings application allows creating ultra-high-speed, contactless, vacuumed, miniature EM with low noise and heat generation. Electrostatic supports are controlled.

In the Russian Federation industry, the electrostatic support is most widely used as gyro bearings. Basic electrostatic bearing theory is presented in Refs. [11–16].

Additionally, certain industrial application perspectives have bearings, which are based on the Lorentz force, which are defined as follows:

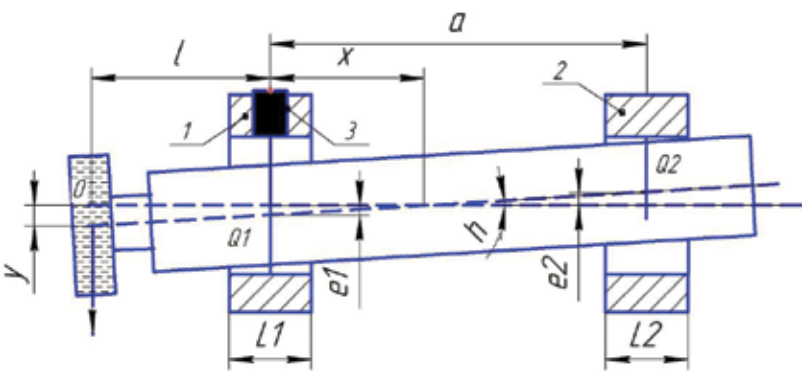


Figure 5. Hybrid gas-magnetic shaft suspension of high-speed spindle: 1—front gas-magnetic bearing; 2—rear gas-static bearing; 3—electromagnet.

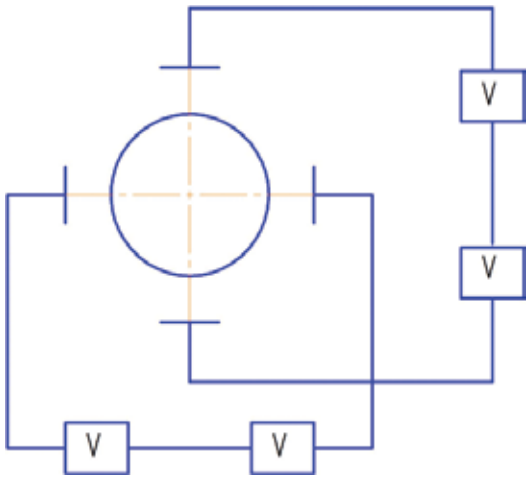


Figure 6. The electrostatic support.

$$f = q(E + [v \times B]). \tag{7}$$

This type of bearings has broad prospects for use in HS EM. For example, the Swiss company Seleroton has developed ultra-high-speed vacuumed motor CM-AMB-400 using this type of bearings (power of 250 W, the rotor speed of 400 000 rpm).

Using the suspension based on the Lorentz forces in the electric motor in conjunction with vacuum allowed to almost completely solve the problems of the rotor friction of the air and the friction in the bearing supports. Overall efficiency of the EM reaches 91–92%.

3. A generalized approach to the calculation of the basic AMB and HMB overall dimensions

In view of prospects for using of AMB and HMB, it is useful to consider the approach for the calculation of their overall dimensions in more detail.

In view of the design similarity of AMB and HMB (HMB, in which PM are used to create an additional magnetic flux), the development of a generalized approach for the AMB and HMB calculation is proposed.

To solve this problem, consider HMB design with radial or axial magnetic inserts. The fundamental difference of these designs is the arrangement of the PM for reinforcing the magnetic flux in the way of the magnetic field line. Thus, these design differences have no significant influence on the mathematical description of the HMB. Moreover, one can get AMB, equating the energy characteristics of a PM to zero that allow making a conclusion about the generalization of considered designs for AMB and HMB.

The following assumptions are used in solving the problems:

1. Permeability of air is equal to the magnetic permeability of vacuum;

2. Temperature and HMB electromagnetic mode are established;

3. AMB active materials are isotropic.

Based on the terms of the problem, the developed generalized approach should take into account both the thermal and electromagnetic processes in HMB. Therefore, the equivalent circuit method (equivalent circuits) has been selected for HMB research that is widely used in the electromagnetic and thermal processes calculations. **Figure 5** shows the equivalent circuit of the magnetic (a) and thermal circuit (b) of HMB.

The strength of the HMB determined as:

$$F = \frac{pl\tau B_\delta^2}{8\mu_0}, \tag{8}$$

where p—number of poles; l—active length of HMB; $\tau = \frac{\pi D}{2p}$—pole pitch; B_δ—flux density in the HMB air gap.

According to the equivalent circuit from the total current law, it should be:

$$F_m + 2Iw = 2F_\delta + 2F_z + F_j + 2F_{zr} + F_{jr}, \tag{9}$$

where F_m—m.m.f of the PM; F_δ—m.m.f of the air gap; F_z—m.m.f. of the stator magnetic core teeth; F_j—m.m.f. of the stator magnetic core back; F_{zr}—m.m.f. in the radial length of the rotor; F_{jr}—m.m.f. in the axial length of the rotor.

Taking into account that $F_\delta = \frac{1}{\mu_0} B_\delta \delta$, then:

$$\frac{1}{\mu_0} B_\delta \delta = \frac{F_m + 2Iw - 2F_z - F_j - 2F_{zr} - F_{jr}}{2}, \tag{10}$$

M.m.f. of the PM is defined as follows:

$$F_m = H_{cB} l_m \tag{11}$$

Taking into account the temperature dependence of the energy characteristics of PM:

$$F_m = H_{cB}\, l_m \left(1 - \frac{k_{Hc}(\Theta_{PM} - 20)}{100}\right), \tag{12}$$

where $H_{cB}(\Theta)$—RMS values of the coercive force of the PM; Θ_{PM}—the temperature of the PM; k_{Hc}—tension temperature coefficient.

It should be noted that the tension temperature coefficient can be assumed to be constant only when the temperature of the PM is 60–80°C (for intermetallic alloys *NdFeB* and *SmCo*). At temperatures outside this range, this ratio has a nonlinear dependence.

The PM temperature in the steady state operation of the HMB is determined on the basis of the thermal equivalent circuit, **Figure 7**b.

Functions approximating the actual magnetization curve of soft magnetic material from which the HMB magnetic core and shaft are made is used when taking into account the HMB magnetic core saturation:

$$H_z = \alpha_1 \mathrm{sh}\, \beta_1 B_z \tag{13}$$

$$H_j = \alpha_1 \mathrm{sh}\, \beta_1 B_j \tag{14}$$

$$H_{zr} = \alpha_2 \mathrm{sh}\, \beta_2 B_{zr} \tag{15}$$

$$H_{jr} = \alpha_2 \mathrm{sh}\, \beta_2 B_{jr} \tag{16}$$

where α_1, β_1—approximation coefficients for the soft magnetic material of the HMB magnetic core; α_2, β_2—approximation coefficients for the soft magnetic material of the shaft; B_z—flux density in the magnetic core teeth; B_j—flux density in the magnetic core back; B_{zr}—flux density on the shaft in the radial direction and B_{jr}—flux density on the shaft in the axial direction.

Then, using the obtained expression and real magnetization curve of the HMB magnetic core material, it is possible to create HMB characteristic taking into account the saturation (the dependence of the force of gravity from the current).

In **Figure 8**, as an example, dependence of the force from a current is made based on the saturation and for various ambient temperatures. All dependencies are built in static mode, transient thermal and electromagnetic processes when making the dependencies were not considered.

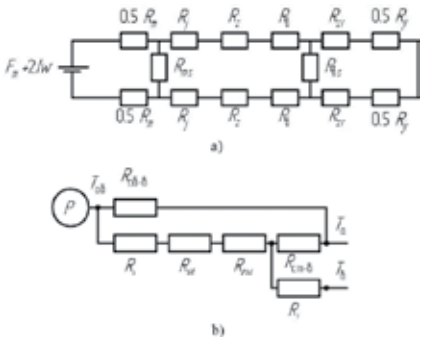


Figure 7. Equivalent circuit of the magnetic circuit HMB: (a) equivalent circuit of the magnetic circuit; (b) equivalent circuit of the thermal circuit. Here, F_m—m.m.f. of the PM; I—current in the AMB winding; w—AMB winding number turns; R_m—the magnetic resistance of the PM; R_j—the magnetic resistance of the AMB magnetic core back; R_z—the magnetic resistance of the AMB magnetic core teeth; R_δ—the magnetic resistance of the HMB air gap; R_{zr}—the magnetic resistance of the rotor radial length; R_{jr}—the magnetic resistance of the rotor axial length; $R_{\delta s}$—the magnetic resistance of the air gap scattering; R_{ms}—the magnetic resistance of the PM scattering; R_i—thermal resistance of the winding insulation; R_{sl}—thermal resistance of the stator; R_{PM}—thermal resistance of the PM insertion.

http://dx.doi.org/10.5772/67144

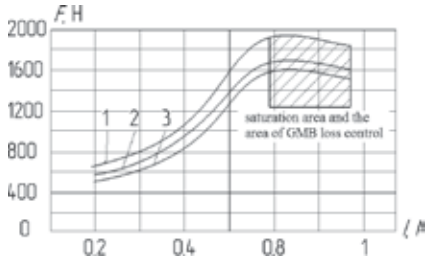


Figure 8. Dependence of the HMB tractive force from a current value (taking into account changes in ambient temperature and magnetic core saturability). Here, 1—at a temperature of 20°C; 2—at a temperature of 60°C; 3—at a temperature of 90°C.

From these curves, it is seen that the HMB is losing its control in the magnetic core saturation area and at high temperatures. This is due to a significant nonlinearity dependence of HMB forces from the current and the magnetic flux of the PM. AMB and HMB control system is usually built on the linearization of these dependencies. Loss controllability area occurs at 0.8 A. In this area, HMB tractive force remains practically unchanged as current increases, since the magnetic core reaches saturation. At a significant saturation, HMB tractive force is slightly reduced, which causes a significant increase in stator back and teeth m.m.f. Steel 2421 was used for making dependence.

4. Computer modeling of dynamic electromagnetic processes in HMB

The developed mathematical apparatus can be used to study the general physical processes in HMB, as well as for engineering calculation of basic geometric dimensions of HMB and AMB considering nonlinear electromagnetic and thermal processes. At the same time, developed mathematical apparatus does not allow making selection of the most rational radial HMB design with magnetic inserts. To solve these problems, the computer simulation methods of the magnetic field of various HMB and AMB designs are more appropriate to use.

Software complex Ansoft Maxwell was used to solve this problem, where two main radial HMB designs with magnetic inserts considered, **Figure 1**, and AMB design present for comparing.

Overall dimensions and constructive parameters of the researched designs are presented in **Table 5**.

Comparison of the considered HMB designs was made under the same weight and overall dimensions, output power and materials properties on the following criteria: the magnitude of the force in the air gap of the HMB (main energy characteristic), stiffness when the rotor is displaced by 60% of the air gap. The forces in the air gap were also compared in the absence of current in the windings. Comparison results are presented in **Figure 9**.

Comparison of HMB and AMB characteristics produced in relative units, the characteristics ascribed to the AMB. The AMB strength and stiffness were taken as 1, and the HMB characteristics are already determined from this base value.

Design	Parameter					
	Number of poles	Air gap (mm)	Active length (mm)	Bore diameter	The outer diameter of the stator	Weight (kg)
HMB with a radially magnetized PM inserts	8	0.5	60	30	60	0.7
HMB with a tangentially magnetized PM inserts	8	0.5	60	30	60	0.7
AMB	8	0.5	60	30	60	0.7

Table 5. HMB and AMB researched designs.

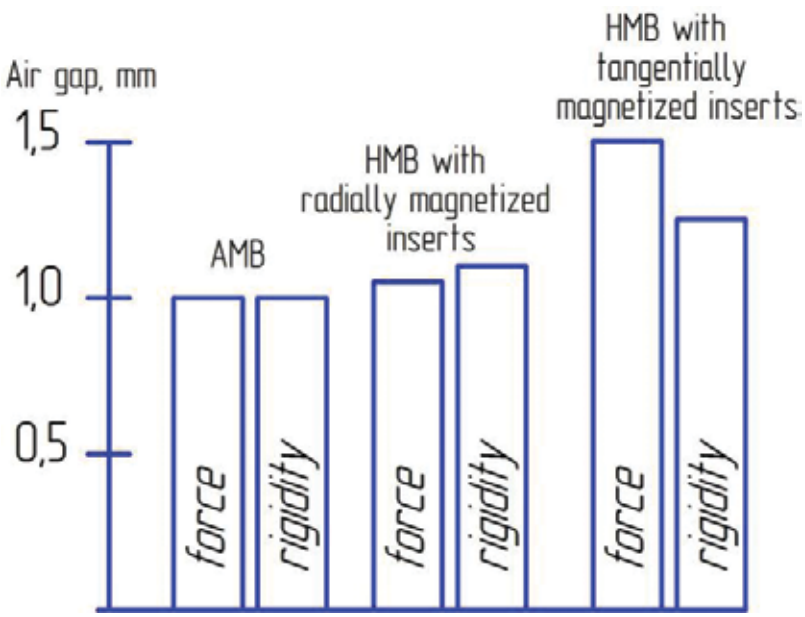


Figure 9. Comparison of the AMB parameters and various HMB designs.

5. Results and conclusions

From the obtained data analysis, it is seen that the maximum stiffness and force value in the air gap has HMB with tangentially magnetized inserts (50 and 40%, respectively, more than AMB indicators for the same weight and overall dimensions). Achieving these characteristics due to the PM inserts will reduce the AMB power consumption at almost two times. The use of a radially magnetized inserts gives little effect: increasing the strength of AMB characteristics by 5–8% and stiffness by 10–12%. In this case, the AMB consumption can be reduced by 8–10%.

HMB with tangential insert provides strength about 130 N in the absence of current in the windings, while the impact strength falls uniformly on the whole rotor. The presence of this force value (25% of the AMB power at maximum current) allows more "gently put" rotor on the bearing harnesses and minimizes the consequences of super heavy transients at AMB failure. HMB with radial inserts provides power of 125–130 N in the absence of power supply, but this force is applied to a small rotor section, and this may lead to a complication of the transition process in case of AMB failure, as it will cause additional "build up" of the rotor.

This chapter also shows that at high temperatures and magnetic core saturation HMB loses control. Thus, generalized approach to the design of AMB and HMB considering nonlinear electromagnetic and thermal dependencies has been developed in this paper.

Author details

Ismagilov Flur Rashitovich, Vavilov Vyacheslav Evgenievich and D.V. Gusakov[1]*

*Address all correspondence to: gusakov.den@mail.ru

Department of Electromechanics, Ufa State Aviation Technical University, Ufa, Russia

References

[1] Super-precision cylindrical roller bearings: High-speed No. 10 series. Publication 07016/1 EN; June 2009

[2] ISO 683-17:1999. Heat-treated steels, alloy steels and free-cutting steels—Part 17: Ball and roller bearing steels. ISO; 1999. 26 p

[3] Schweitzer G. Applications and research topics for active magnetic bearings. In: Schweitzer G, editor. Proceedings of the IUTAM-Symposium on Emerging Trends in Rotor Dynamics, 23-26 March 2009; Indian Institute of Technology, Delhi, India. Springer-Verlag [Online]. Available from: http://www.mcgs.ch/web-content/Delhi_IUTAM_Symp.pdf [Accessed: October 03, 2016]

[4] Zhuravlev Y.N. Active magnetic bearings: theory, calculation, application. In: Zhuravlev Y.N., editor. Saint-Petersburg: University of Technology; 2003. 206 p

[5] Ismagilov FR, Khayrullin IH, Vavilov VE. The hybrid magnetic bearings. Types of structures and development prospects. Safety and Security of Energy. 2016;**1**(32):S.30–34

[6] Johnson D, Pillay P, Malengret M. High speed PM motor with hybrid magnetic bearing for kinetic energy storage. Conference Record. IAS Annual Meeting (IEEE Industry Applications Society), vol. 1; 2001. pp. 57–63

[7] Schetinin VS. Scientific substantiation of creation and development of high-speed spindle units on the gas-magnetic supports of machine tools: The dissertation. Dr. tehn. Sciences: 05 February 2007. Komsomolsk-on-Amure: V.S. Schetinin; 2011. p. 37

[8] Kopitov SM, Kosmynin AV, Ulyanov AV, Schetinin VS, Khvostikov AS. Improved designs of the gas-magnetic supports of high-speed rotary machines. Basic Research. 2013;**10**:25–29

[9] Kosmynin AV. Calculation of bearing capacity of gas-magnetic supports of the high-speed spindle units. Machinery and Tools. 2010;**9**:6–8

[10] Kosmynin VS, Schetinin SM, Kopitov AV, Ulyanov AS, Khvostikov AS. Improving the accuracy of the rotation shafts of high-speed rotor systems. Scientists KnAGTU. 2013;**3-1**(15):47–51

[11] Martynenko YG. Trends in the development of modern gyroscope. Soros Educational Zhurnal. 1997;**11**:120–127

[12] Martynenko YG. Rigid Body Motion in Electric and Magnetic Fields. M. Science; 1988. p. 368

[13] Martynenko YG. Theoretical foundations of electrostatic gyro design with a solid rotor. Beijing: University Tsinhua; 1996. p. 68

[14] Kuvykin VI. The magnetic friction in non-contact bearings. Nizhny Novgorod: Publishing House of the Society "IntelService"; 1997. p. 112

[15] Kozorez VV. Dynamical Systems of Magnetically Interacting Free Bodies. Kiev: Naukova Dumka; 1981. p. 140

[16] Fedorov AE. Dynamics and stability of the electrically charged bodies system: The dissertation candidate. tehn. Sciences: 01 February 2006; Nizhni Novgorod: A.E. Fedorov; 2007. p. 25

Chapter 5

Bearing Fault Detection in Induction Machine Using Squared Envelope Analysis of Stator Current

Valeria Cristina Maria Nascimento Leite,
Jonas Guedes Borges da Silva,
Germano Lambert Torres,
Giscard Francimeire Cintra Veloso,
Luiz Eduardo Borges da Silva,
Erik Leandro Bonaldi and
Levy Ely de Lacerda de Oliveira

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/67145

Abstract

In this chapter, motor current signature analysis based on squared envelope spectrum is applied in order to identify and to estimate the severity of outer race bearing faults in induction machine. This methodology is based on conventional vibration analysis techniques, however, it is, non-invasive, low cost, and easier to implement. Bearing fault detection and identification in induction machines is of utmost importance in order to avoid unexpected breakdowns and even a catastrophic event. Thus, bearing fault characteristic components are extracted combining summation of phase currents, prewhitening, spectral kurtosis and squared envelope spectrum analysis. Experimental results with a 0.37 W, 60 Hz, and three-phase induction machine demonstrated the methodology effectiveness.

Keywords: bearing fault detection, induction machine, motor current signature analysis, squared envelope analysis, spectral kurtosis

1. Introduction

In an industrial scenario, three-phase induction machines have several applications due to their reliability, availability, and cost-effectiveness. Unexpected faults in these machines

could lead to unexpected breakdowns, losses at industrial production, or have catastrophic consequences. In this context, rolling element bearings are responsible for more than 40% of induction machine faults [1]. Rolling bearings are critical mechanical components, which allow relative movement between systems, supporting radial and thrust load. Bearing faults could be associated to contamination, corrosion, inadequate lubrication, installation problems, and misalignment or overloading [2]. In general, a fault affects only one bearing component—inner race, outer race, cage, or ball; as the fault evolves, it spreads to other components; moreover, these faults could be described based on fault mechanism, location, or on a combination of these [3, 4].

Maintenance of electrical machines is an activity of the utmost importance. Moreover, considering a scenario of cost reduction and production efficiency, the development of an effective maintenance program has been gaining more attention and several tools have been implemented to support and encourage best practices. In this sense, advanced methods for data acquisition and processing have been developed in order to allow an effective machine condition monitoring and early fault detection and identification, avoiding unexpected breakdowns and even catastrophic failures, especially for critical systems. Whenever possible, condition monitoring should be done non-invasively and without interrupting machine operation [5–7].

Over the years, the concept of maintenance became more comprehensive, reducing fault occurrence and increasing industrial system availability. Besides, requirements of reliability, safety, and criticality were associated with the system or equipment under analysis. Maintenance strategies or schemes can be classified as corrective (run-to-break), preventive (time-based) and predictive (condition-based maintenance) [8]. Corrective maintenance is only performed after an occurrence of a fault and therefore involves unexpected breakdowns, high costs, changes in the production chain, and in addition, it could lead to catastrophic events. Preventive maintenance and interventions occur based onto a scheduled maintenance plan or based on the equipment mean time between failures. Although it is more effective than corrective maintenance, by preventing most failures, unexpected failure may still occur. Additionally, the process cost is still high, especially, the costs associated with labor, inventory, and even with unnecessary replacement of equipment or components [8, 9].

On the other hand, predictive maintenance analyses the equipment condition so that a possible fault can still be identified at an early stage. Predictive maintenance aims to identify a machine anomaly so that it does not result in a fault. Such maintenance involves advanced technique of monitoring, processing, and signal analysis, that are generally performed non-invasively and, in many cases, in real time. In case of induction machines, these techniques can be developed based on vibration, temperature, acoustic emission, or electrical current signal monitoring [9]. It should be noted that the monitoring of such signals or parameters, in order to verify the operating condition of a machine, is called condition monitoring. In fact, condition monitoring aims to not only observe machine current operational condition, but also to predict machine future condition, keeping it under a systematic analysis during the machine's remaining life [8]. In this sense, from a systematic machine condition monitoring, a fault condition can be detected and identified, such that, a diagnosis procedure can be established, whereby properly investigating the fault symptoms and prognosis [10].

http://dx.doi.org/10.5772/67145

In general, the machine's monitored signals are the result of a combination of different sources, which can vary according to machine environmental and operational conditions, monitoring and acquisition systems, among others. Besides, some faults, such as bearing faults, can produce a signal modulation, which give rise to other signals (sidebands and harmonics) [8]. Thus, a great challenge in machine diagnosis consists of separating and identifying these sources. In case of bearing fault detection in induction machines, the motor current signature analysis (MCSA) has emerged as one of the leading condition monitoring techniques. This approach is an advantageous alternative (or complementary) to condition monitoring based on vibration analysis. Many machines already have current monitoring for control or protection purposes, not requiring the installation of other types of sensor, therefore, this approach can be considered non-invasive and cost-effective. Over the years, the spectral estimation using techniques based on Fourier transform has been widely applied for analysis of the stator current [11–13]. This analysis methodology considers the use of stationary signals, that is, considers that the machine is operating at constant speed and load. On the other hand, advanced techniques take into account the nonstationary signals [11].

In this context, this work consists of applying a methodology for analyzing the electrical signature for the diagnosis of point defects in bearings induction motor, based on spectral kurtosis and squared envelope spectrum analysis, in order to increase the fault detection capability even in an incipient stage [14].

2. Bearing fault diagnosis in induction machine

Rolling bearings are one of the most important mechanical components in induction machines. Therefore, it is necessary to assess the health condition of these components, especially by means of signal processing methodologies for bearing fault diagnosis. Bearing fault diagnosis comprises a series of processes performed in order to detect, isolate, and identify the bearing condition based on the machine monitoring [10]. Although there are several techniques for monitoring of bearing condition in induction machine, i.e., vibration, acoustic emission, and ultrasound, this section describes an approach based on electrical stator current analysis or current signature analysis. This approach has been gaining attention since bearing failure causes a modulation in electrical current signal, which can be identified in a similar way, as it is done in vibration analysis [15].

This section aims to describe some of the most used methodologies for induction machine fault detection based on electrical current signature analysis. In this context, it is important to know the machine to be monitored, and often the system in which it is inserted, since practical considerations are essential to allow a proper fault diagnosis. Some of these considerations are mainly related to machine technical specifications; load variations; rotor speed variations; power supply characteristics; failure mode to be analyzed (electrical or mechanical); sensors (physical quantity to be monitored, specification, amount), among others [16, 17].

Bearing fault detection is a technique mainly based on feature extraction from acquired signal, and condition identification based on the analysis of these features [10]. In the case of a

fault localized on the inner or outer race, whenever a rolling element passes through the fault surface, a series of impulses are generated. This almost periodic series of impulses present characteristics that vary with bearing geometry and fault localization; in addition, they excite resonances in the bearing and in the machine structure as a whole [8, 16].

The series of generated impulses are still amplitude modulated as the fault passes by the load zone and they are influenced by the transfer function from the fault to the sensor. The impulses are generated at a rate which varies according to: the fault position (inner race, outer race, and cage), the bearing dimensions, and the machine shaft speed (f_r). Thus, it is possible to estimate the so called bearing characteristic frequencies, i.e., ball pass frequency of the outer race (BPFO), ball pass frequency of the inner race (BPFI), fundamental train frequency (FTF), which is related to cage speed rotation, and ball spin frequency (BSF). The following equations represent these frequencies [3]:

$$\text{BPFO} = \frac{nf_r}{2}\left(1 - \frac{d}{D}\cos\alpha\right) \tag{1}$$

$$\text{BPFI} = \frac{nf_r}{2}\left(1 + \frac{d}{D}\cos\alpha\right) \tag{2}$$

$$\text{FTF} = \frac{f_r}{2}\left(1 - \frac{d}{D}\cos\alpha\right) \tag{3}$$

$$\text{BSF} = \frac{D}{2d}\left[1 - \left(\frac{d}{D}\cos\alpha\right)^2\right] \tag{4}$$

where n corresponds to the number of rolling elements; α is the angle of the load from the radial plane; d is the ball diameter and D is the pitch diameter. When such characteristic frequencies appear (or its amplitude increase) in the analyzed signal spectrum, it is possible to identify a bearing fault and its location [10]. However, it is very difficult to extract these components, since they have low amplitude and are merged with other spectral components and background noise.

Therefore, it is possible to affirm that fault detection based on the current analysis is great a challenge, especially in industrial environments mainly due to low signal-to-noise ratio of the characteristic frequency components associated with these faults, even though several studies have shown promising results in this area [6, 18]. On the other hand, in many situations, motor current signature analysis (MCSA) becomes a useful alternative to traditional fault detection methods, e.g., vibration analysis, particularly considering the sensor installation, risks, costs associated with process, and degree of criticality of the system or machine under analysis [11].

2.1. Motor current signature analysis—MCSA

MCSA is one of the most commonly used techniques to fault detection in induction motors, since it allows identifying electrical and mechanical faults. It performs a spectral analysis of stator electrical current, which is usually monitored at one of three power supply phases. Studies related to mechanical faults effects on motor stator current mainly consider: load torque oscillations, rotating eccentricities, and air gap eccentricity [11, 15, 19].

http://dx.doi.org/10.5772/67145

In case of bearing faults, it is possible to consider that machines inductances can vary due to rotating eccentricities at bearing characteristic frequencies f_C, i.e., BPFO, BPFI, etc., which produces a stator current modulation, described by [11]:

$$f_E = f_s \pm k \cdot f_C \tag{5}$$

where f_E is the frequency related to a bearing fault; f_s is the power supply frequency; and $k = 1, 2, 3, \ldots$ is the harmonic number. Thus, f_C appears in the current spectrum as sidebands.

In this context, it is import to observe that rotor inertia and winding inductances produce an electromechanical filtering effect in stator current, such that, this current is mainly affected by low frequency components [20, 21].

Other studies show that load torque oscillations can occur each time the rolling elements reach a localized fault on the outer or inner race, or when a fault on a rolling element reaches a race. These oscillations cause phase modulations in electrical current as described by Eq. (5) [22].

Finally, another approach considers that the effect of a localized bearing fault in stator current can be modeled as air gap eccentricity. In this case, a magnetic flux density variation affects stator current as a function of the fault location. Thus, frequencies related to the bearing faults are expressed by [19]:

$$f_{E\,\text{outer race}} = f_s \pm k \cdot \text{BPFO} \tag{6}$$

$$f_{E\,\text{inner race}} = f_s \pm f_r \pm k \cdot \text{BPFI} \tag{7}$$

$$f_{E\,\text{ball}} = f_s \pm \text{FTF} \pm k \cdot \text{BSF} \tag{8}$$

where $f_{E\ outer\ race}$, $f_{E\ inner\ race}$, and $f_{E\ ball}$ are the frequencies related to a fault in outer race, inner race, and ball respectively, which correspond to an amplitude modulation of the fundamental power supply frequency (f_s). It is important to observe that this modulation is caused by a permeance variation on the rotor fundamental magnetomotive force [11].

2.2. Power spectral density

Generally, the MCSA is carried out using classical or nonparametric spectral estimation methods. Nonparametric methods require little information regarding the signal to be analyzed and its computational complexity is low, especially compared to modern spectral estimation methods [16, 23].

Among the most common nonparametric techniques are the periodogram and its refined variations, i.e., Bartlett, Welch, and Daniell methods [22]. Periodogram can be obtained by [23]:

$$\hat{\Phi}_{\text{P}}(\omega) = \frac{1}{N} \left| \sum_{n=0}^{N-1} y(t)\, e^{-j\omega t} \right|^2 \tag{9}$$

where $y(t)$ is signal under analysis and its samples could be represented by $[y(t)]_{t=1}^{N}$.

Mean squared error, represented by the sum of the bias squared and the variance, is a parameter commonly used to evaluate the performance of an estimator. In this sense, bias reduction is obtained by applying a window. In order to reduce periodogram variance, Bartlett method uses an average of several periodograms obtained from different segments of

the signal. In this case, the original signal $[y(t)]_{t=1}^{N}$ with N samples is split into K segments, such that, an average of $L = N/K$ periodograms is computed. Welch method can be seen as evolution of Bartlett method; since the estimation is performed considering that the signal segments are overlapped and windowed. Thus, variance is reduced, but also the resolution [23].

3. Envelope analysis

Bearing faults can be classified as localized (single-point) or extended. Incipient localized faults produce sharp impulses that cover a large bandwidth. These faults, in general, are associated with small pits or spalls. On the other hand, extended faults effect is not so apparent or highlighted in the spectrum and its bandwidth is limited. Brinelling and corrosion are examples of extended bearing faults. It is also possible that a small localized fault becomes an extended fault as the fault evolves over time. Regardless of the type of fault, in general, bearing failure can be detected using envelope analysis [3].

It is also important to observe that signals produced by bearing faults (localized or extended) are typically nonstationary, i.e., signals whose statistical parameters vary in time. More specifically, localized bearing faults signals can be modelled as cyclostationary or pseudocyclostationary [8, 24].

Over the years, the envelope analysis or high frequency resonance demodulation has been widely used for identifying localized faults in rolling bearings. Each time a bearing component strikes the fault surface, a mechanical shock occurs. Consequently, an impulse is generated and structural resonances of the system are excited by it. In addition, these impulses are modulated in amplitude. This way, through the envelope analysis, it is possible to obtain demodulated signals, which are directly related to the bearing condition [8].

The following steps perform envelope analysis. First, digital bandpass filtering of acquired signal in a suitable frequency band, in general, around the machine mechanical resonance is performed. Following, the filtered signal is demodulated. Finally, the resulting signal frequency spectrum is estimated, resulting in the envelope spectrum, whereby it is possible to identify the periodic components associated with a fault in a bearing component [16, 25]. In other words, it is possible to identify the repetition frequency of the impulses caused by a fault simply analyzing the envelope signal spectrum, which, in general, it is not possible by using the raw spectrum [17]. Fourier transform is applied in order to obtain the envelope spectrum.

One of the most used tool for demodulation or envelope extraction is Hilbert transform [26, 27]. First, the acquired signal is bandpass filtered around a machine resonance frequency, and then Hilbert transform is applied. This digital technique reduces the data length and allows flexibility for bandpass filter specification [28].

However, it is important to observe that a suitable frequency band to filter the signal must contain impulses generated by the fault and amplified by machine mechanical or structural resonances [8]. Therefore, one of the main difficulties in using envelope analysis is undoubtedly the choice of an appropriate frequency band for filtering the signal. In order

http://dx.doi.org/10.5772/67145

to circumvent this drawback, algorithms based on spectral kurtosis have been successfully applied, which is discussed later in the chapter.

3.1. Hilbert transform

As mentioned before, bearing fault signals can be seen as amplitude modulated signal, such that, carrier frequency, represented by high frequency resonances are modulated by bearing characteristic frequencies. Hilbert transform can be used for the demodulation process in envelope analysis when modulated signal is proved to be analytic [8].

When envelope analysis is performed based on Hilbert transform, the frequency band to be demodulated can be properly separated from adjacent components that could interfere with the analysis. Impulse response function produced by bearing faults has real and imaginary parts of its corresponding frequency function related by Hilbert transform [8].

In general, signal-to-noise ratio is used as an indication of the frequency band where the modulated signal should be filtered. After filtering, selected frequency band is shifted at low frequencies in the spectrum and padded with zeros to double the length in order to obtain a one-side spectrum. When computing the inverse Fourier transform of this one-side spectrum, an analytic signal is obtained, such that, its imaginary part is the Hilbert transform of the real part. In this way, envelope corresponds to the modulus of real and imaginary parts. However, it is more interesting to analyze the squared envelope, since it can improve signal-to-noise ratio by removing extraneous components in practical situations [28].

3.2. Kurtogram

A rolling bearing fault excites high frequency resonances in the rotating machine, which can produce modulations at bearing characteristic frequencies. Therefore, characteristic frequency components should be demodulated using an optimal selection of frequency and bandwidth (f, Bw) for bearing fault identification based on envelope analysis. In this sense, spectral kurtosis based algorithms, such as kurtogram, aims to find this combination in a computationally efficient way [25].

Initially, spectral kurtosis (SK) was defined based on short-time Fourier transform (STFT) for impulsivity measurement as a function of frequency, and it was mainly applied to sonar signal analysis [17]. Some years ago, SK was also considered and applied for bearing fault analysis [29].

Thus, spectral kurtosis of a signal $x(t)$, i.e., kurtosis value for each frequency (f), can be computed based on the STFT ($X(t, f)$) of this signal, such that [8, 30]:

$$SK(f) = \frac{\langle X^4(t,f)\rangle}{\langle X^2(t,f)\rangle^2} - 2 \tag{10}$$

where $X(t, f)$ corresponds to the envelope as a time-frequency function; $X^2(t, f)$ represents the power spectrum values calculated for each time (t); and the average of all these power spectral values ($\langle X^2(t,f)\rangle$) corresponds to the power spectrum of the analyzed signal as a whole. In addition, the constant factor 2 is subtracted, so that, for Gaussian signal, Eq. (2) turns to zero [8]. In this sense, spectral kurtosis can be understood as a filter so that its value is maximum in the

frequency bands containing impulsive signals and zero for that frequency bands dominated by stationary signals [29].

Since using short-time Fourier transform, parameters, such as window length, can directly affect the spectral kurtosis calculation; therefore, considering an impulsive signal, the window shorter than the spacing between two consecutive pulses and longer than an individual pulse shall provide a maximum kurtosis value. A detailed investigation about the relation between spectral kurtosis value and window length was conducted in Ref. [28]. Additionally, in Ref. [29], it was depicted that the square root of the spectral kurtosis is equivalent to the optimum Wiener filter and it demonstrated a close relation between optimum matched filter and spectral kurtosis value. For envelope analysis, in order to obtain an optimum result, it is of utmost importance to specify properly filter center frequency and bandwidth. For this purpose, the concept of kurtogram emerges as a tool to find the optimum filter for envelope analysis based on spectral kurtosis values. Kurtogram displays the spectral kurtosis values as a function of frequency and windows length, which define the spectral resolution. Experiments showed that the filter set from kurtogram was more efficient for outer race fault detection, when compared with Wiener and matched filters [28].

Fast kurtogram algorithm was developed as an extension of the kurtogram, especially considering that it was costly and inefficient to analyze all possible combinations of frequency and windows length. Fast kurtogram computes spectral kurtosis using digital filters, instead of short-time Fourier transform, following a dyad-decomposition so-called 1/3-binary tree. This decomposition is similar to discrete wavelet packet transform, where frequency bands are divided into bands with one half of their previous width, but here, divisions by 1/3 are also included [30].

As an alternative for fast kurtogram, the wavelet kurtogram algorithm was developed. In this case, nonorthogonal complex Morlet wavelets are used for signal decomposition and it is considered that the optimum combination center frequency and bandwidth for envelope analysis could be found based on a 1/nth-octave wavelet analysis. In general, the sequence 1/1, 1/2, 1/3, 1/4, 1/8, 1/12, …, 1/*n*th-octave is used, although, any sequence could be applied. Besides, before wavelet decomposition, the original signal power spectral density is prewhitened by an autoregressive model in order to enhance the fault detection into the envelope spectrum. Additionally, applied complex Morlet wavelet was optimized, since several filter banks are tested and the selected for envelope analysis is the one that maximizes the SK. The scheme of signal decomposition by means of filter bank for SK optimization is similar to that one used in kurtogram [17].

Wavelets are used because they present an impulse response with a constant damping ratio, which is more suitable for impulsive signals analysis in comparison with STFT. Besides, complex Morlet wavelet is analytic; therefore, its Fourier transform presents only positive frequencies. Thus, SK for each wavelet filter can be calculated considering that the product of the Morlet wavelet coefficients and their complex conjugate corresponds to the squared envelope of the filtered signal [17]. Here, it is also important to notice that using the quadratic envelope has been more advantageous for bearing signal analysis [28], which will be discussed in the next section.

http://dx.doi.org/10.5772/67145

The SK calculation could be enhanced by prewhitening the spectrum of the signal to be analyzed. Through the prewhitening, signal spectrum becomes almost constant, similar to the white noise spectrum. This process reduces variations that could occur in transient signals spectrum, which can lead to inaccurate SK calculations [17]. An autoregressive model can be used for signal spectrum prewhitening. In this case, the model error corresponds to the noise, but especially to the nonstationary part of the signal, which contains information related to bearing fault. In other words, it is possible to say that a digital filter (linear prediction filter), which is designed based on an autoregressive signal model, predicts the deterministic part of the signal; and the prediction error, which contains an impulsive signal that will be used for machine condition analysis [31].

An autoregressive model (*AR*) of order *p* can be represented by [32]:

$$AR(k) = -\sum_{i=1}^{p} a(i)x(i+k) + \text{error}(k) \tag{11}$$

where $a(i)$, $i = 1, 2, 3, \ldots, p$, corresponds to the linear prediction filter weighting coefficients; *error* (k) is a whitened signal, which is the difference between the original and the predicted signals. Minimum least square error is used to find the coefficients of the linear predictor. Model order (p) will be one that maximizes the kurtosis of the *error*(k), such that, this residual signal will contain fault related impulse signals. Besides, (p) must be smaller (in number of samples) than that the space between two consecutive bearing faults impulses [17].

3.3. Squared envelope analysis

During the envelope analysis, existing random or discrete noise components can make it difficult to identify components related to bearing failure. That is why a major constraint of envelope analysis is related to signal-to-noise ratio. A way to overcome this limitation is by using squared envelope. In this case, envelope spectrum presents a higher harmonic reduction, which cannot be obtained by a common filtering operation [28].

A method for computing squared envelope from an analytic signal was depicted in Ref. [28]. There, squaring envelope process is defined as a convolution of an analytic signal and its complex conjugate. Thus, squared envelope spectrum can be calculated by the convolution of the analytic signal and its complex conjugate corresponding spectra. In this case, spectrum of squared envelope does not present a sum of frequency components, since the analytic signals have only positive frequency components. Besides, the squared envelope spectrum has the same frequency range as if it was calculated using Hilbert transform and zero padding [8].

It is also important to highlight that the integral of spectral correlation of all considered frequencies is equivalent to spectrum of the squared envelope, where the spectral correlation is a two-dimensional Fourier transform calculated on the two-dimensional autocorrelation function [33].

4. Bearing fault detection methodology

Despite major advances in bearing fault detection techniques, such as MCSA, current methodology still has limitations that make it difficult to identify incipient faults, impairing the fault prognosis. Depending on the operational environment and machine specifications, there may be a reduction in the analysis reliability as a whole.

A way to mitigate this problem consists in separation of signals coming from different sources. In general, the components in the machine vibration or current signals have specific characteristics that allow their separation and identification in order to detect changes in machinery health condition. Noise, eccentricity, gear, cavitation, rolling bearing characteristic frequencies, and broken bars are examples of components that may be present in vibration signals or electric current signals [8].

In this scenery, several techniques have been proposed to support signal separation and identification in machine fault detection. Among these techniques, it is possible to mention, for example, time synchronous averaging (TSA), which is used to remove signal components that are not synchronous with rotor speed. In this situation, a minimal disturbance could occur in the resulting signal, but it is necessary an angular sampling for each harmonic family to be separated. This technique removes harmonics, but not lateral modulation bands. Techniques related to noise cancelling, also could be used in order to mitigate noise contamination. In addition, linear prediction filtering could be used to separate the predictable deterministic signal, which must be removed from the original signal in order to highlight the signal component related to bearing fault [1]. Linear prediction was also considered for electrical signature analysis.

Another technique that was evaluated in Ref. [14] to improve the detection of fault related components was the sum of the electric currents. A common operation in three-phase circuit analysis is to obtain the current or voltage phase using information from other phases. In the case of a three-phase induction machine connected to a delta system, considering that the sum of all currents entering a node is equal to the sum of all the currents out of the node (1st Kirchhoff's Law), it is possible to assume that $I_A + I_B + I_C = 0$, where I_A, I_B, and I_C are the measured currents of the phase *A*, *B*, and *C*, respectively. In this sense, the current of any phase (I_A, I_B, or I_C) can easily be defined by the other two. For example, $I_C = -(I_A + I_B)$.

This procedure is similar to the synchronous average calculation. Any mechanical effect related to the machine condition (nominal or under a fault), including periodic or random components, can be observed in any of the three phases' current, or alternatively, in the numerically obtained current, i.e. (I_C). On the other hand, any other uncorrelated random effect will be attenuated using this procedure [14].

This way, the methodology that guided this work follows five steps:

1. Sum of the electric currents.

2. Prewhitening (linear prediction filtering).

3. Spectral kurtosis based algorithm.

4. Squared envelope spectrum.

5. Bearing fault identification based on bearing characteristic frequency detection.

It is also important to highlight that since faults are identified in the envelope spectrum, its amplitude can be used as severity index. Thus, a fault evolution can be analyzed as function of increases in the bearing characteristic frequency amplitude [34].

http://dx.doi.org/10.5772/67145

4.1. Experimental issues

In this section, damaged rolling bearings (model 6203-ZZ) are installed on a three-phase induction motor; for each bearing, stator current signals are acquired and squared envelope spectrum was analyzed in order to detect outer race faults by means of ball pass frequency outer race (BPFO) identification. Rolling bearings were artificially damaged, such that, through holes of 1.0 mm, 2.0 mm, and 3.0 mm diameter were drilled on the outer race to simulate localized faults with different levels of severity. Experiments were performed using 6203-ZZ shielded metric radial bearings, also described as deep groove ball bearing, single row, double shielded, pressed steel cage, normal clearance, prelubricated with grease, with inner (bore) diameter: 17mm; outside diameter: 40mm; and overall width: 12mm.

Experimental test rig (**Figure 1**) consists of a three-phase squirrel cage induction motor with 0.37 kW power, four poles, and 60 Hz supply frequency, coupled to an electric machine working as a power generator (constant mechanical load), without any speed or torque control. A 24-bit/4-channel data acquisition board (National Instruments NI 9239) and current probes were used to acquire electric current signals at 50 kHz sample rate. Prior to any processing, data was filtered using a low pass filter of 25 kHz.

Two of the three stator currents (I_A and I_B) were measured, and the third one (I_C) was numerically obtained, such that $I_C = -(I_A + I_B)$, and used in the fault detection process. **Figure 2** shows the damaged bearings used in the experiments. Rotational speed was estimated to be 28.80 Hz (1728 rpm), and the characteristic frequency for a fault on the bearing outer race was estimated in (BPFO = 87.93 Hz ± 2%).

The methodology was applied to calculate electric stator current. Following, prewhitening was performed, such that the AR model order was chosen by using the kurtosis maximization criterion of the residual signal. In this work, it is proposed as a simplified methodology, where the healthy

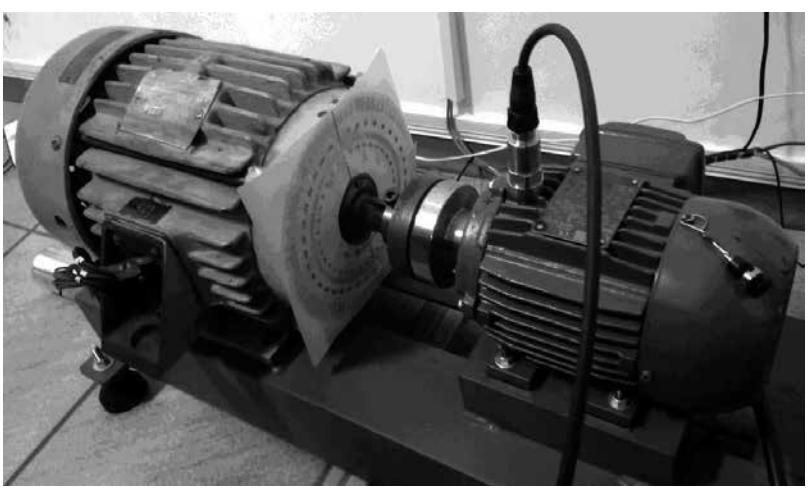


Figure 1. Experimental test rig.

bearing is initially tested and the resulting AR model order is also used for the faulty bearings analysis. Therefore, an AR model order (p = 32) is used for all experiments. Following, fast kurtogram algorithm was applied at five levels of decomposition. It is important to notice that this process, including sampling, signal processing, and feature extraction, lasts about 2 minutes on a modern computer. Although the wavelet kurtogram algorithm has been analyzed, only the results obtained with the fast kurtogram are presented, mainly due to its performance in this application, as explained in Refs. [14, 34]. The signal processing is performed offline using Matlab®.

Figure 2. Damaged bearings used in the experimental tests. From left to right holes of 1.0 mm, 2.0 mm, and 3.0 mm.

Thus, the described methodology was applied for all damaged bearing cases. The fast kurtogram color map was similar to that in **Figure 3**; then, only the resulting squared envelope spectrum was shown for the other experiments. The bandpass filter with center frequency f_C = 6250 Hz and bandwidth Bw = 4167 Hz, at decomposition level (k = 2.6), indicated by black circle in **Figure 3**, was used in all squared envelope calculations, which was very useful for comparisons. In the Figures, an arrow indicates the amplitude of the bearing outer race characteristic frequency (BPFO).

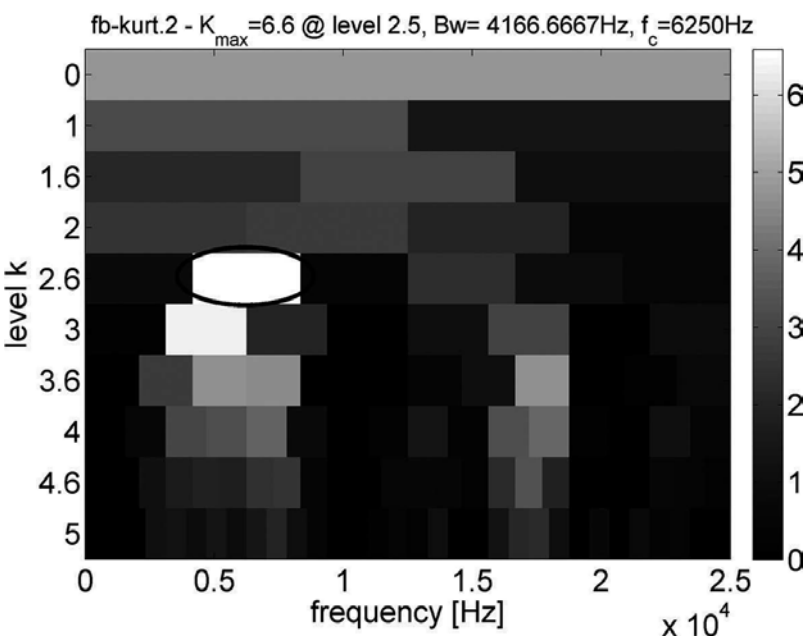


Figure 3. Fast kurtogram color map.

http://dx.doi.org/10.5772/67145

Figure 4 shows the squared envelope spectrum of the electric current for the damaged bearing with the 1.0 mm hole. In this case, the envelope spectrum clearly shows the fault signature around the estimated BPFO, with amplitude $A = 2.9 \times 10^{-9}$.

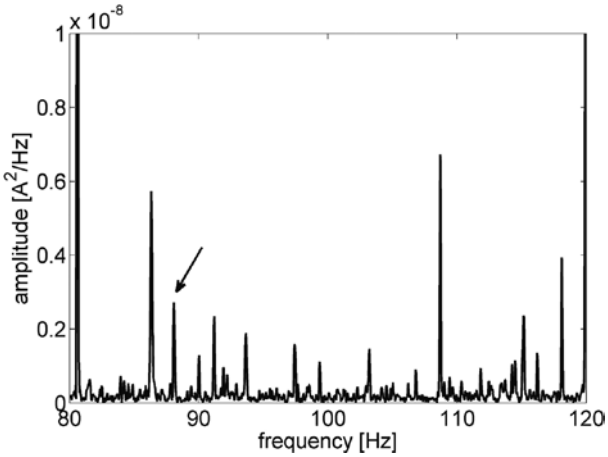


Figure 4. Squared envelope spectrum of the electric current for the damaged bearing with the 1.0 mm hole.

The same procedure was applied to the damaged bearing with 2.0 mm hole, as presented in **Figure 5**. Here, a significant increase in the bearing characteristic fault frequency amplitude ($A = 7.9 \times 10^{-9}$) was observed, confirming the fault effect in the stator current envelope spectrum amplitude.

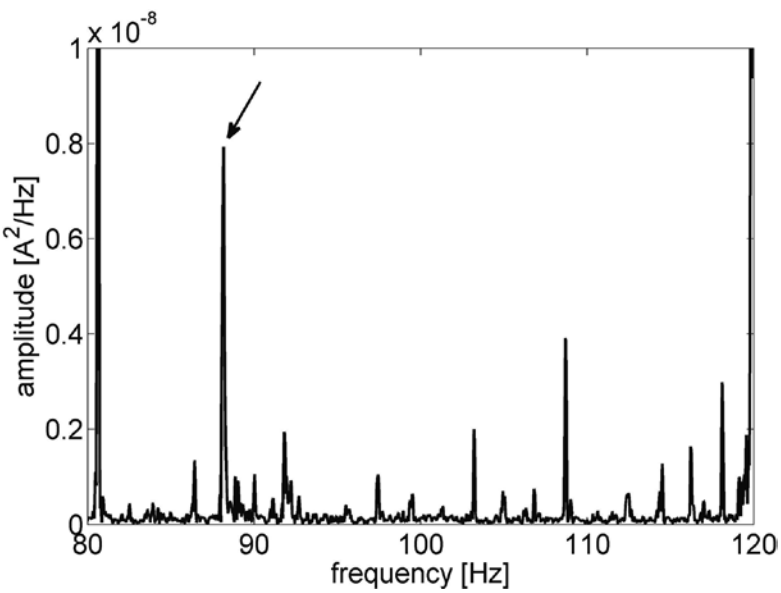


Figure 5. Squared envelope spectrum of the electric current for the damaged bearing with the 2.0 mm hole.

The last experiment assessed the damaged bearing with 3.0 mm hole. In this case, it is important to observe a change in the envelope spectrum graphic scale (**Figure 6**), due to the increase in amplitude ($A = 11.3 \times 10^{-9}$) in the observed fault frequency.

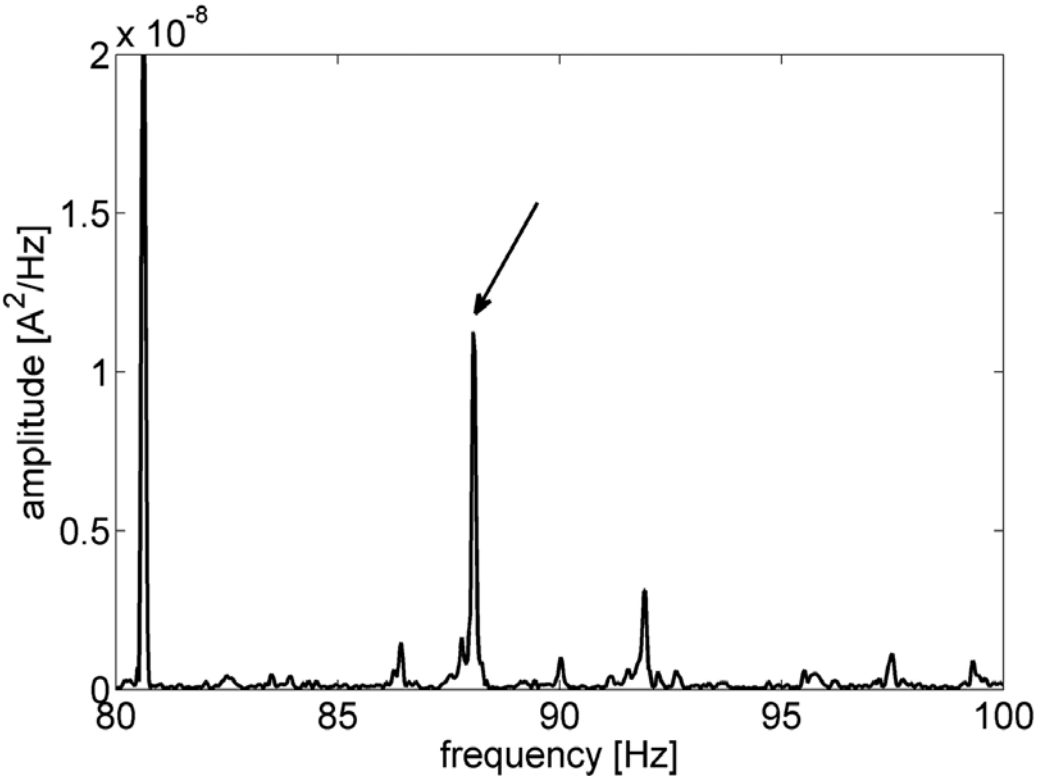


Figure 6. Squared envelope spectrum of the electric current for the damaged bearing with the 3.0 mm hole.

The obtained results validate the methodology, and therefore, the involved theoretical concepts. A BFPO frequency (at 88.1 Hz) was detected for each damaged bearing experiment, strongly indicating a bearing outer race fault. Besides, the characteristic frequency amplitude increases with the fault severity, which could be used as a prognosis indication. In the envelope spectra, it was also observed that as the amplitude of BPFO increased, the amplitude of another frequency component decreased. Thus, as in Ref. [11], it is possible to conclude that, although the stator current analysis is more complex than the vibration analysis, it is an important alternative to bearing fault detection in induction motors, mainly due to its advantages related to cost, availability and applications.

5. Conclusions and comments

This work describes a methodology to enhance MCSA for bearing fault detection and identification in induction machines by combining electrical currents sum, prewhitening based on linear

prediction filtering, spectral kurtosis, and squared envelope analysis. This methodology is based on successful methodologies and algorithms, initially proposed to be applied to vibration signals. An experimental test of such methodology was depicted using a test rig where artificially damaged bearings were created in order to simulate faults at different severity levels. Results show that the methodology improves MCSA in comparison with traditional spectrum analysis. Besides, the methodology provides an indication of fault severity based on bearing characteristic frequency (e.g. BPFO) amplitude in squared envelope, which can be used for prognosis purposes. For real industrial applications, the authors believe that this methodology could be easily carried by a professional predictive maintenance team, given adequate equipment and analysis software.

Acknowledgments

The authors would like to thank the National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES), Minas Gerais Research Foundation (FAPEMIG), and Brazilian Electricity Regulatory Agency Research and Development (ANEEL R&D) for supporting this work.

The authors would like to thank the Professor J. Antoni for providing the Fast Kurtogram Matlab® code.

Author details

Valeria Cristina Maria Nascimento Leite[1]*, Jonas Guedes Borges da Silva[2], Germano Lambert Torres[2], Giscard Francimeire Cintra Veloso[3], Luiz Eduardo Borges da Silva[3], Erik Leandro Bonaldi[2] and Levy Ely de Lacerda de Oliveira[2]

*Address all correspondence to: valeria.nleite@gmail.com

1 Institute of Aeronautics and Space, São José dos Campos, Brazil

2 PS Solutions, Itajubá, Brazil

3 Itajuba Federal University, Itajubá, Brazil

References

[1] IEEE Motor Reliability Working Group. Report on large motor reliability survey of industrial and commercial instalations. IEEE Transactions on Industry Applications. 1985;**21**(4):853–864. DOI: 10.1109/TIA.1985.349532

[2] Harris T.A., Kotzalas M.N. Advanced Concepts of Bearing Technology. 5th ed. Boca Raton: CRC Press; 2006. 368 p.

[3] Randal R.B., Antoni J. Rolling element bearing diagnostics—a tutorial. Mechanical Systems and Signal Processing. 2011;**25**(2):485–520. DOI: 10.1016/j.ymssp.2010.07.017

[4] Moyer C.A., Ai X. Rolling Element Bearings. In: Bhushan B., editor. Modern Tribology Handbook. 1st ed. Boca Raton: CRC Press; 2001. p. 28.

[5] Lu B., Durocher D.B., Stemper P. Predictive maintenance techniques. IEEE Industry Applications Magazine. 2009;**15**(6):52–60. DOI: 10.1109/MIAS.2009.934444

[6] Zhou W., Habetler T.G., Harley R.G. Bearing fault detection via stator current noise cancellation and statistical control. IEEE Transactions on Industrial Electronics. 2008;**55**(12):4260–4269. DOI: 10.1109/TIE.2008.2005018

[7] Jung J.H., Lee J.J., Kwon B.H. Online diagnosis of induction motors using MCSA. IEEE Transactions on Industrial Electronics. 2006;**53**(6):1842–1852. DOI: 10.1109/TIE.2006.885131

[8] Randall R.B. Vibration-Based Condition Monitoring: Industrial, Aerospace and Automotive Applications. 1st ed. Hoboken: Wiley; 2011. 308 p.

[9] Girdhar P., Scheffer C. Practical Machinery Vibration Analysis and Predictive Maintenance. 1st ed. Oxford: Elsevier; 2004. 272 p.

[10] Rai A., Upadhyay S. A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. Tribology International. 2016;**96**:289–306. DOI: 10.1016/j.triboint.2015.12.037

[11] Blodt M., Grajon P., Raison B., Regnier J. Mechanical Fault Detection in Induction Motor Drives through Stator Current Monitoring—Theory and Application Examples. In: Zhang W., editor. Fault Detection. 1st ed. Rijeka: InTech; 2010. Ch. 21, pp. 451–488. DOI: 10.5772/9072.

[12] Bonaldi E.L., de Oliveira L.E.L., Lambert-Torres G., Borges da Silva L.E. Proposing a Procedure for the Application of Motor Current Signature Analysis on Predictive Maintenance of Induction Motors. In: 20th International Congress & Exhibition on Condition Monitoring and Diagnosis Engineering Management; June 13-15, 2007; Faro, Portugal. Faro, Portugal: COMADEM Press; 2007. DOI: 10.13140/RG.2.1.2836.7605

[13] Frosini L., Bassi E. Stator current and motor efficiency as indicators for different types of bearings faults in induction motors. IEEE Transactions on Industrial Electronics. 2010;**57**(1):244–251. DOI: 10.1109/TIE.2009.2026770

[14] Leite V.C.M.N., Borges da Silva J.G., Veloso G.F.C., Borges da Silva L.E., Lambert-Torres G., et al. Detection of localized bearing faults in induction machines by spectral kurtosis and envelope analysis of stator current. IEEE Transactions on Industrial Electronics. 2015;**62**(3):1855–1865. DOI: 10.1109/TIE.2014.2345330

[15] Schoen R.R., Habetler T.G., Kamran F., Bartheld R.G. Motor bearing damage detection using stator current monitoring. IEEE Transactions on Industry Applications. 1995;**31**(6):1274–1279. DOI: 10.1109/28.475697

[16] Borges da Silva J.G. Modeling and Treatment of Motor Electrical Signature Machines Signals for Fault Diagnosis Enhancement (in Portuguese) [dissertation]. Itajuba, Brazil: Itajuba Federal University; 2014. 170p.

[17] Sawalhi N. Diagnostics, Prognostics and Fault Simulation for Rolling Element Bearings [thesis]. New South Wales, Australia: The University of New South Wales; 2007. 354 p. Available from: http://handle.unsw.edu.au/1959.4/40544

[18] Bellini A., Filippetti F., Tassoni C., Capolino G.A. Advances in diagnostic techniques for induction machines. IEEE Transactions on Industrial Electronics. 2008;**55**(12):4109–4126. DOI: 10.1109/TIE.2008.2007527

[19] Blodt M., Granjon P., Raison B., Rostaing G. Models for bearing damage detection in induction motors using stator current monitoring. IEEE Transactions on Industrial Electronics. 2008;**55**(4):1813–1822. DOI: 10.1109/TIE.2008.917108

[20] Bonaldi E.L., Borges da Silva L.E., Lambert-Torres G., de Oliveira L.E.L., Assuncao F.O. Using Rough Sets Technique as a Fault Diagnosis Classifier for Induction Motors. In: 28th Annual Conference of the Industrial Electronics Society, IECON 2012; November 5–8, 2002; Sevilla, Spain. New Jersey: IEEE Press; 2002. pp. 3383–3388. DOI: 10.1109/IECON.2002.1182941

[21] Immovilli F., Bellini A., Rubini R., Tassoni C. Diagnosis of bearing faults in induction machines by vibration or current signals: A critical comparison. IEEE Transactions on Industry Applications. 2010;**46**(4):1350–1359. DOI: 10.1109/TIA.2010.2049623

[22] Blodt M., Chabert M., Regnier J., Faucher J. Mechanical load fault detection in induction motors by stator current time frequency analysis. IEEE Transactions on Industry Applications. 2006;**42**(6):1454–1463. DOI: 10.1109/TIA.2006.882631

[23] Stoica P., Moses R. Spectral Analysis of Signals. 1st ed. Upper Saddle River: Prentice-Hall; 2005. 452 p.

[24] Antoni J., Randall R.B. Differential diagnosis of gear and bearing faults. Journal of Vibration and Acoustics. 2002;**124**(2):165–171. DOI: 10.1115/1.1456906

[25] Guo Y., Na J., Li B., Fung R.F. Envelope extraction based dimension reduction for independent component analysis in fault diagnosis of rolling element bearing. Journal of Sound and Vibration. 2014;**333**(13):2983–2994. DOI: 10.1016/j.jsv.2014.02.038

[26] Randall R.B. Developments in Digital Analysis Techniques for Diagnosis of Bearings and Gears. In: Australian Acoustical Society Adelaide, Fifth International Congress on Sound and Vibration; December 15–18, 1997. Adelaide, Australia, vol. 1, pp. 133–149.

[27] El Bouchikhi E.H., Choqueuse V., Benbouzid M., Antonino-Daviu J.A. Stator Current Demodulation for Induction Machine Rotor Faults Diagnosis. In: 2014 First International Conference on Green Energy ICGE 2014; March 25–27, 2014; Sfax, Tunisia. New Jersey: IEEE Press; 2014. pp. 176–181. DOI: 10.1109/ICGE.2014.6835418

[28] Ho D., Randall R.B. Optimisation of bearing diagnostic techiniqueusing simulated and actual fault signals. Mechanical Systems and Signal Processing. 2000;**14**(5):763–788. DOI: 10.1006/mssp.2000.1304

[29] Antoni J., Randall R.B. The spectral kurtosis: application to vibratory surveillance and diagnostics of rotating machines. Mechanical Systems and Signal Processing. 2006;**20**(2):308–331. DOI: 10.1016/j.ymssp.2004.09.002

[30] Antoni J. Fast computation of the kurtogram for the detection of transient faults. Mechanical Systems and Signal Processing. 2007;**21**(1):108–124. DOI: 10.1016/j.ymssp.2005.12.002

[31] Sawalhi N., Randall R.B., Endo H. The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis. Mechanical Systems and Signal Processing, 2007;**21**(6):2616–2633. DOI: 10.1016/j.ymssp.2006.12.002

[32] Sawalhi N., Randall R.B. Spectral Kurtosis Optimization for Rolling Element Bearings. In: Eighth International Symposium on Signal Processing and Its Applications; August 28–31, 2005; Sydney, Australia. New Jersey: IEEE Press; 2005. pp. 839–842. DOI: 10.1109/ISSPA.2005.1581069

[33] Randall R.B., Antoni J., Chobsaard S. The relationship between spectral correlation and envelope analysis in the diagnostic of bearing faults and other cyclostationary machine signals. Mechanical Systems and Signal Processing. 2001;**15**(5):945–962. DOI: 10.1006/mssp.2001.1415

[34] Leite V.C.M.N., Borges da Silva J.G., Borges da Silva L.E., Veloso G.F.C., Lambert-Torres G., et al. Experimental bearing fault detection, identification and prognosis through spectral kurtosis and envelope spectral analysis. Electric Power Components and Systems. Taylor & Francis, 2016;**44**(18):2121–2132. DOI: 10.1080/15325008.2016.1209705.

Edited by Pranav H. Darji

In the twenty-first century, bearings are expected to perform better in the form of various operating conditions, that is from low speed to extremely high speed and from low load to huge load applications. The expectations from the field of bearing technology are great. During the recent years, we have been witnessing the development of a new generation of mechanical systems that are highly miniaturized and very sophisticated, yet extremely robust. Technological progress creates increasingly arduous conditions for rolling mechanisms.

ISBN 978-953-51-4821-0

Photo by SafakOguz / iStock

IntechOpen