Università degli Studi di Ferrara

DOCTORAL COURSE IN
ENGINEERING SCIENCE

Cycle XXXV

Director: Prof. S. Trillo

A Combined Numerical and Experimental Approach for Rolling Bearing Modelling and Prognostics

Scientific/Disciplinary field ING/IND 13

Candidate
Dott. Alberto Gabrielli

Supervisor
Prof. Giorgio Dalpiaz

Co-Supervisor
Eng. Gianluca D'Elia

Years 2019/2022

.

PREFACE

I graduated *cum laude* in 2019 in Mechanical Engineering at the University of Ferrara discussing the thesis entitled "Implementation of a numerical model for shape memory alloy wires" under the supervision of Prof. R. Rizzoni. Immediately after, I joined with a postgraduate fellowship the research group led by Prof. G. Dalpiaz at the same university. Then, after eight months I started my Ph.D. in Engineering Science, again under the supervision of Prof. G. Dalpiaz.

During the first year of my Ph.D., my research activity was focused on the definition of a numerical finite-element procedure for the estimation of the non-linear radial stiffness of rolling-element bearings. This quantity, in fact, is a crucial input parameter that is inserted into dynamic models of complex machinery and it greatly affects their outcome. As a consequence, the load-displacement relationship of these components should be determined as accurately as possible, e.g. through the finite-element method. Therefore, the inquiry aimed at finding a procedure for the generation of computationally efficient load-dependent meshes, i.e. providing an acceptable compromise between the accuracy of the results and the computational time. This study was first published in 2021 in the journal Mathematical Problems in Engineering [1]:

[1] A. Gabrielli, M. Battarra, and E. Mucchi. "A Critical Analysis of Finite-Element Modeling Procedures for Radial Bearing Stiffness Estimation." In: *Mathematical Problems in Engineering* 2021 (2021), p. 9955398. ISSN: 1024-123X. DOI: 10.1155/2021/9955398.

Further results and additional considerations were later added and presented the following year at the ISMA 2022 conference held in Leuven, Belgium [2]:

[2] A. Gabrielli, M. Battarra, and E. Mucchi. "A numerical finite-element method for radial bearing stiffness estimation based on load dependent meshing." In: *Proceedings of ISMA 2022 - International Conference on Noise and Vibration Engineering and USD 2022 - International Conference on Uncertainty in Structural Dynamics, pp. 1527–1541*. Leuven, Belgium, 2022.

During the second year, I tackled the problem of the lumped parameter modelling of defective rolling-element bearings. In this regard, it was observed that the estimation of several quantities that are inserted in these models is commonly characterized by a high degree of uncertainty. Hence, effort was dedicated to the development of a robust procedure capable of determining the values of these unknown parameters. This goal was achieved by proposing a multi-objective optimization procedure, in which the objective functions to be minimized were global indicators that took into account the discrepancy between signal features estimated numerically and experimentally. The former were obtained by means of a lumped parameter model of a defective bearing. The latter were obtained through an extensive experimental study realised on a dedicated test bench which was set-up at the Engineering Department of the University of Ferrara. The described investigation led to a publication, in 2023, in the journal Mechanical Systems and Signal Processing [3]:

[3] A. Gabrielli, M. Battarra, E. Mucchi, and G. Dalpiaz. "A procedure for the assessment of unknown parameters in modeling defective bearings through multi-objective optimization." In: *Mechanical Systems and Signal Processing* 185 (2023), p. 109783. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2022.109783.

Additionally, the experimental data were stored as a dataset in a online repository hosted by Mendeley Data [4]:

[4] A. Gabrielli, M. Battarra, E. Mucchi, and G. Dalpiaz. *Acceleration signals of rolling element bearings with artificial defects*. Mendeley Data, V1. 2022. DOI: 10.17632/8wdzm5gwng.1.

The last part of my research, which was carried out during the third and final year of my Ph.D., involved the investigation and the development of physics-based models devoted to the prognostics of rolling-element bearings. The previous effort on the dynamic modelling of faulty bearings was extended and integrated into a novel technique for the estimation of a parameter related to the evolution of bearing degradation over time, namely the equivalent damaged volume. A procedure for the estimation of this quantity from real bearing deterioration histories was introduced. Then, it was further employed to propose two psychics-based prognostic models, which were aimed at predicting

the future health state of a bearing on the basis of actual data recorded during its working life. To this end, the aforementioned bearing test bench at the University of Ferrara was utilized to perform a series of run-to-failure experiments, i.e. tests in which the entire operative life of a bearing was monitored from the very beginning until the eventual failure of the component.

Finally, in addition to my Ph.D. research topic, I also had the chance to perform several experimental activities and conduct numerical analyses involving the modelling of complex mechanical systems. These activities allowed me to solve real industrial problems and enhance my problem-solving capabilities in the NVH field. Among those, I worked on a numerical vibro-acoustic methodology for the estimation of the vibratory and acoustic level of a gearbox employed on agricultural equipment. The endeavour aimed at generating the NVH digital twin of the real component, which was experimentally validated through comparison with data acquired on a dedicated test bench. This work was presented at the ISMA 2020 conference [5]:

[5] A. Gabrielli, F. Pizzolante, E. Soave, M. Battarra, C. Mazzeo, M. Tarabra, E. Fava, and E. Mucchi. "A numerical model for NVH analysis of gearboxes employed on agricultural equipment." In: *Proceedings of ISMA 2020 - International Conference on Noise and Vibration Engineering and USD 2020 - International Conference on Uncertainty in Structural Dynamics, pp. 3191 - 3203*. Leuven, Belgium, 2020.

To conclude this brief summary of my Ph.D. experience, I would like to spend a few words to acknowledge the people that worked me in the last three - almost four - years. Firstly, Prof. G. Dalpiaz, that gave me the opportunity to pursue this path under his supervision. Then, I want to express my heartfelt gratitude to Prof. Emiliano Mucchi, Eng. Gianluca D'Elia and Eng. Mattia Battarra for the precious help and the time they spent discussing with me the arguments detailed in this thesis, plus many more related to the field of NVH analysis. Finally, a special thanks goes to all people that shared this experience with me, starting with my office mates, my co-workers in the research group and, most importantly, the friends I have met along the way. Needless to say, this work would have never seen the light of day without the support of all of you. Sincerely, you have my deepest gratitude.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AFE	Acoustical Finite-Element
AI	Artificial Intelligence
ANN	Artificial Neural Network
AR	Autoregressive
ATV	Acoustic Transfer Vector
BPFO	Ball Passing Frequency Outer ring
CAD	Computer-Aided Design
CBM	Condition-Based Maintenance
CNN	Convolutional Neural Network
CWRU	Case Western Reserve University
DDM	Data-Driven Model
DOF	Degree Of Freedom
EOL	End Of Life
EDV	Equivalent Damaged Volume
EDM	Electrical Discharge Machining
EHL	Elastohydrodynamic Lubrication
EMA	Experimental Modal Analysis
FE	Finite-Element
FFNN	Feed-Forward Neural Network
FPT	First Predicting Time
GGHMM	Generalized Gaussian Hidden Markov Model
GPR	Gaussian Process Regression
HMM	Hidden Markov Model
HS	Health State

IG	Inverse Gaussian
IMS	Intelligent Maintenance Systems
ISO	International Organization for Standardization
LP	Lumped Parameter
MAC	Modal Assurance Criterion
MFPT	Machinery Failure Prevention Technology
NF	Neuro-Fuzzy
NVH	Noise, Vibration and Harshness
OA	Operative Analysis
PBM	Physics-Based Model
PDF	Performance Degradation Dictionary
PDF	Probability Density Function
PE	Paris-Erdogan model
PH	Proportional Hazard
PHM	Prognostics and Health Management
PML	Perfectly Matched Layer
REB	Rolling Element Bearing
RMS	Root Mean Square
RUL	Remaining Useful Life
RVM	Relevance Vector Machine
SFE	Structural Finite-Element
SVM	Support Vector Machine
SVR	Support Vector Regression
TT	Time-to-Threshold
VC	Varying Compliance

NOMENCLATURE

Latin symbols - scalars

A^*	Quadratic sum of the first five BPFO harmonics
A_{BPFO}	Amplitude of the BPFO
A_{BPFOe}	Amplitude of the BPFO in the envelope spectrum
a	Contact area semi-width for roller-race contact; ellipse major semi-axis for ball-race contact
a_{ref}	Reference acceleration for dB computation
B	Ring width
$BPFO$	Characteristic frequency for an outer ring defect
b	Ellipse minor semi-axis for ball-race contact
C_0	Basic static load rating
C_r	Basic dynamic load rating
c	Damping coefficient
d	Ball trajectory inside a localised defect
d_b	Ball trajectory inside an extended defect
d_α	Ball trajectory inside a localised defect, aligned with contact angle direction
$d_{\alpha,b}$	Ball trajectory inside an extended defect, aligned with contact angle direction
E	Young's modulus
E_{eq}	Equivalent Young's modulus
$env()$	Signal envelope
F	Force
F_r	Radial load

f_c	Contact force exchanged between rolling elements and races
$\bar{f}_c$	Quasi-static contact force exchanged between rolling elements and races
f_d	Contact damping force
f_l	Low-pass frequency
f_m	Meshing frequency
f_n	Natural frequency
f_{rot}	Frequency of rotation
f_s	Sampling frequency
g	Global mesh size
H_{or}	Sum of the first five BPFO harmonics in the envelope spectrum
$\mathcal{H}()$	Hilbert transform operator
h	Radial clearance
h_d	Defect depth, smooth surface
h_r	Defect depth deviation due to rough surface
I	Identity matrix
i	Vector index
J	Moment of inertia
J_r	Radial load integral
j	Imaginary unit
k	Stiffness
k_b	Bearing stiffness term
k_c	Contact stiffness
$k_{c,b}$	Contact stiffness of a ball in contact with a single race
$k_{c,r}$	Contact stiffness of a cylinder in contact with a single race
k_{cl}	Linearised contact stiffness
$k_{m,B}$	Average meshing stiffness of tooth profile *Baseline*
$k_{m,M}$	Average meshing stiffness of tooth profile *Mod1*

L_{10h}	Bearing rating life, expressed in hours
l	Local mesh size
l_{eff}	Roller effective length
MAC	MAC value
m	Mass
m_b	Ball mass
N	Signal length
n	Load-deflection exponent
n_b	Number of rolling elements on one row
n_r	Number of rows
n_{rep}	Number of test repetitions
n_{subs}	Number of subdivisions along shaft axis for meshes with hexahedral elements
n_{tests}	Number of tested operative conditions
P	Dynamic bearing load
p	Acoustic pressure
p_{ref}	Reference acoustic pressure for dB computation
Q	Load on rolling element
Q_{max}	Maximum load on rolling element
R	Surface roughness scale factor
R^2	Coefficient of determination
RMS	RMS value
r	Radius
r_b	Ball radius
r_c	Cylinder radius
$r_{eq,b}$	Equivalent radius of contact in case of ball-race contact
$r_{eq,c}$	Equivalent radius of contact in case of roller-race contact
r_g	Ring groove radius
r_i	Inner ring race radius
r_o	Outer ring race radius

r_p	Pitch radius
s	Bearing life exponent
TT	Time-to-threshold value
t	Time
t_c	Current time
t_{fs}	Time in which the final stage of bearing life starts
t_{EOL}	Time at bearing end of life
t_{th}	Instant in time in which the equivalent damaged volume threshold is crossed
V	Signal feature
$\overline{V}$	Average value of a signal feature
V_{eq}	Equivalent damaged volume
V_{th}	Threshold value for the equivalent damaged volume
ν	Unknown parameter
ν_s	Speed of sound
X	RMS value computed by the EDV algorithm
x	Displacement in x direction
x_b	Gear backlash
x_d	Dynamic transmission error
x_r	Radial displacement
y	Displacement in y direction
W	Input power
w	External load
w_a	Semi-width of the FE contact area for roller-race contact; length of the major semi-axis in the FE model for ball-race contact
w_b	Length of the minor semi-axis in the FE model for ball-race contact
Z	Signal
Z_a	Analytic signal

Latin symbols - vectors and matrices

$\mathbf{ATM}$	Acoustic transfer matrix
$\mathbf{atv}$	Acoustic transfer vector
$\mathbf{C}$	Damping matrix
$\mathbf{f_c}$	Contact force vector
$\mathbf{f_d}$	Contact damping force vector
$\mathbf{h_d}$	Vector of defect depths
$\mathbf{K}$	Stiffness matrix
$\mathbf{K_b}$	Linearised bearing stiffness matrix
$\mathbf{K_l}$	Linearised stiffness matrix
$\mathbf{k_m}$	Meshing stiffness vector
$\mathbf{M}$	Mass matrix
$\mathbf{p}$	Acoustic pressure vector
$\mathbf{v_n}$	Normal velocity vector
$\mathbf{x}$	Displacement vector
$\dot{\mathbf{x}}$	Velocity vector
$\ddot{\mathbf{x}}$	Acceleration vector
$\mathbf{w}$	External load vector

Greek symbols - scalars

α	Bearing contact angle
α_c	Rayleigh damping coefficient which multiplies the mass term
β	Coefficient of proportionality between damping and stiffness matrices
β_c	Rayleigh damping coefficient which multiplies the stiffness term
Γ	Sign function for gear backlash
γ	Coefficient that nullifies traction forces
$\overline{\gamma}$	Coefficient that nullifies quasi-static traction forces

Δ	Difference between numerical and theoretical stiffness estimate
ΔA^*	Percentage difference between numerical and experimental values of A^*
ΔRMS	Percentage difference between numerical and experimental values of RMS
Δt	Time step
ΔV	Percentage difference between numerical and experimental features
$\Delta\varphi_d$	Angular extent of a bearing fault
δ	Relative contact displacement of the inner ring with respect to the outer ring
$\dot{\delta}$	Relative contact velocity
$\overline{\delta}$	Quasi-static relative contact displacement of the inner ring with respect to the outer ring
δF	Small force disturbance
$\delta\varphi$	Angular sampling
$\overline{\epsilon}$	First simplified elliptical integral
ζ	Damping ratio
η	Coefficient of proportionality between local and global mesh size
$\overline{\kappa}$	Ellipticity parameter
Λ	Constant term in the relationship $k_r(m_r)$
λ	Eigenvalue
ν	Poisson's ratio
ξ	Numerator of $k_r(m_r)$
$\overline{\xi}$	Second simplified elliptical integral
ρ	Density
σ	Slippage term
τ	Gear ratio
φ	Rolling element angular position

φ_0	Initial angular position of the cage
$\varphi_{b,in}$	Angular location in which a rolling element starts contact with the defect bottom surface
$\varphi_{b,out}$	Angular location in which a rolling element ends contact with the defect bottom surface
φ_d	Angular location of the centre of a localised defect
$\varphi_{d,in}$	Angular location of the entry of a localised defect
$\varphi_{d,out}$	Angular location of the exit of a localised defect
φ_l	Load zone angular extent
ϕ_{slip}	Slippage angle
χ	Denominator of $k_r(m_r)$
ω	Eigenfrequency
ω_c	Cage angular speed
ω_{in}	Gearbox input shaft speed
ω_s	Shaft speed
ω_{out}	Gearbox output shaft speed

Greek symbols - vectors and matrices

$\boldsymbol{\gamma}$	Unknown parameters vector
$\boldsymbol{\Delta\varphi}_d$	Vector of bearing fault angular extents
$\boldsymbol{\theta}$	Objective functions vector
$\boldsymbol{\psi}$	Mode shape vector

Subscripts & superscripts

$+$	Value increased by a small disturbance
$-$	Value decreased by a small disturbance
exp	Experimental value
i	Term related to the inner ring
j	j-th rolling element
k	k-th element in a summation
max	Maximum value

min	Minimum value
num	Numerical value
o	Term related to the outer ring
p	p-th test repetition
r	Term related to the high-frequency resonator
x, y, z	Cartesian directions
I	First body in contact
II	Second body in contact

1

INTRODUCTION

1.1 INTRODUCTION TO ROLLING ELEMENT BEARINGS

Rolling-element bearings (REBs) are one of the most frequently employed components in rotating machinery [6]. They are load carrying components which permit to eliminate the sliding friction between two machine elements by introducing rolling elements in between them. As a consequence, they allow for smooth operation while reducing the production of heat, energy consumption and wear [7]. Their ease of mounting and dismounting allows for an efficient exchange of the components. Besides, they are available in a number of shapes and dimensions, and their cost is usually limited. The combination of all these factors explain their wide employment in the field of mechanics.

REBs are commonly composed by four parts, i.e. inner ring, outer ring, rolling elements and cage. An example of bearing assembly is shown in Figure 1. The two rings, which are mounted on the components to be uncoupled, are separated by a number of rolling elements. The rings are designed so that the rolling elements are able to move on paths called raceways or races. A bearing may have one or more raceways on each ring to increase the total number of rolling elements. These three components are made of steel, which is usually hardened on the surfaces in contact. In addition, a cage must be inserted to ensure proper operation. In fact, this component permits to maintain a fixed distance between the rolling elements and keep them evenly spaced. Compared

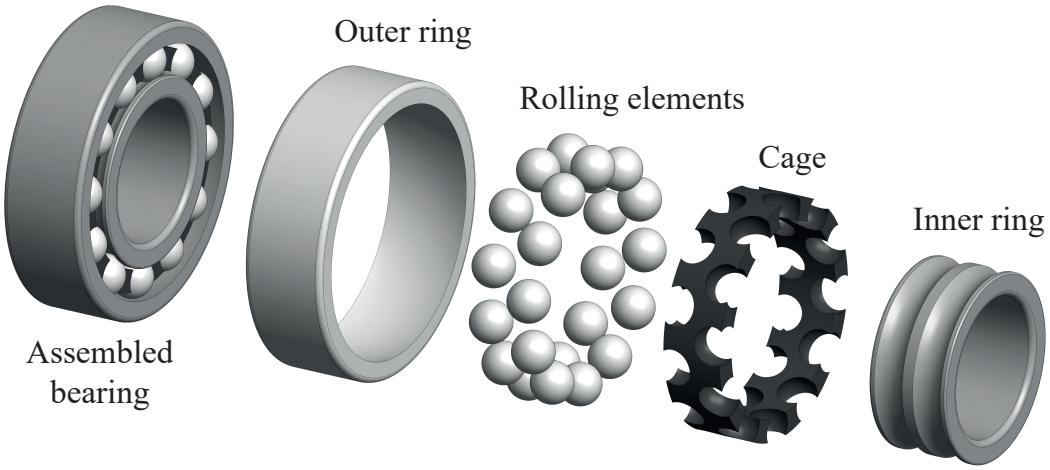

Figure 1: Exploded view of a double-row self-aligning ball bearing.

to the other parts, the cage is made of softer materials, e.g. aluminium or polymers such as polyamide and polytetrafluoroethylene. Nonetheless, steel cages are available for certain applications, for instance in high-temperature environments. In the majority of cases, the inner ring rotates, while the outer ring remains stationary. In this context, the structure that supports the outer ring is usually referred to as housing. However, there are scenarios in which the opposite situation takes place, or where both rings are able to rotate, e.g. in inter-shafts bearings [8]. Finally, REBs are lubricated by either oil or grease to further decrease the friction and dissipate heat.

Rolling bearings may be roughly subdivided in two major categories. i.e. ball and roller bearings, depending on the type of rolling element that is employed. Balls provides low rolling friction, thus allowing to accommodate high shaft speeds. However, because of this characteristic they should not be employed for applications involving heavy loads. On the other hand, rollers are characterized by an higher friction, making them more suitable to carry heavy loads, but they are indicated for slower speeds. Moreover, bearings may be further classified on the basis of the supported load in radial and thrust bearings. Radial bearings are able to carry loads which act perpendicularly to the shaft. On the contrary, thrust bearings may accommodate axial loads only. Some radial bearings are also capable of supporting a combination of axial

loads in one or two axial directions. Similarly, some thrust bearings may carry radial loads. Metric bearing dimensions are standardized according to ISO (International Organization for Standardization) standards. They are denoted by their boundary dimensions, i.e. bore diameter, outside diameter, width or height and chamfer dimensions. The standardization of these widely employed components is essential to guarantee their interchangeability.

Among ball bearings, the most popular are single-row deep-groove ball bearings. They are characterized by raceway grooves which have a curvature radius in the order of 51.5 to 53 % of the ball diameter. They are designed to carry radial loads, but they may work also under of combinations of radial and axial loads. Moreover, they are able to maintain good performance even at high speeds. Double-row ball bearings are similar but have two raceways instead of one. They have a greater load capacity compared to their single-row counterpart. Other common types of REBs are the angular-contact ball bearings, which are specifically designed to accommodate combinations of trust and radial loads. They are also indicated to support heavy thrust loads, and their groove radius is in the order of 52 to 53 % of the ball diameter. They are also available in a single or double-row fashion. The two-rows design allows to carry thrust loads in either directions, even combined with a radial load. Furthermore, self-aligning ball bearings are special types of double-row ball bearings in which the outer raceway is a portion of a sphere. As a consequence, they are internally self-aligning but, due to the low conformity of the outer raceway, they have a reduced low-carrying capacity. Despite this, they are commonly employed in applications that may involve a certain degree of misalignment. Finally, thrust ball bearings are employed to carry axial load and are suitable for high speed operations.

Concerning roller bearings, they provide a larger load carrying capacity compared to ball bearings of the same size. In this regard, they are also characterized by an increased fatigue life and an higher stiffness. Cylindrical roller bearings are the most common type of bearing in this category. They are are able to support high radial loads and may operate at high speeds. As the name suggests, the rollers are characterized by a cylindrical shape. They might be partially or fully crowned to reduce the stress at the edges and offer a slight protection against misalignment. Even for these type of bearings, double-row

variants are available. Needle roller bearings, instead, have a similar design but rollers are designed to have a remarkably higher length compared to their diameter. They are ideal for applications in which the available space is low. In fact, they are sometimes inserted directly on hardened shafts without inserting the two rings. Differently, tapered roller bearings are able to accommodate combinations of radial and axial loads. They are not suitable to run at high speeds due to friction at the guide flanges of the rings, but they guarantee an high load capacity. Furthermore, these components are separable: therefore, they must be always mounted in pairs. It is also worth mentioning the peculiar design of spherical roller bearings, in which the outer raceway is a portion of a sphere. Therefore, these REBs are internally self-aligning. They are able to perform well under very high loads but are not suited for high-speed operations. Finally, roller thrust bearings may be found in a number of variants, i.e. with cylindrical, spherical or tapered rollers. These types of bearing, however, are not indicated for high shaft speeds. They may be replaced by needle roller bearings if the available space is limited, but at the expense of a lighter load capacity.

REBs cover a major role in the dynamic behaviour of the systems in which they are mounted [9]. As a consequence, it is crucial to be able to properly characterize their fundamental properties. Indeed, in the context of bearing modelling, one of the most important properties is the bearing stiffness. In the most general case, a bearing is denoted by 6 degrees of freedom (DOFs). Thus, REB stiffness is completely defined by a 6x6 matrix [10]. However, in the majority of cases it is usually sufficient to be able to characterize either the radial or axial stiffness of these components, as these are the quantities needed for employment in further models. Bearing stiffness values, in fact, may be utilized for several purposes. For instance, they might be provided as input values to model bearings in finite-element models of a mechanical system. Similarly, these values may be inserted in a lumped parameter model of complex machinery [11]. In this context, components might be connected to the ground or among them through springs which idealize the bearings behaviour. Despite the importance of REBs stiffness, its value is not provided by any bearing constructor. As a result, stiffness must be "manually" evaluated by the analyst when this parameter is needed. However, the estimation of the bearing stiffness is a non-trivial process which involves the modelling of a

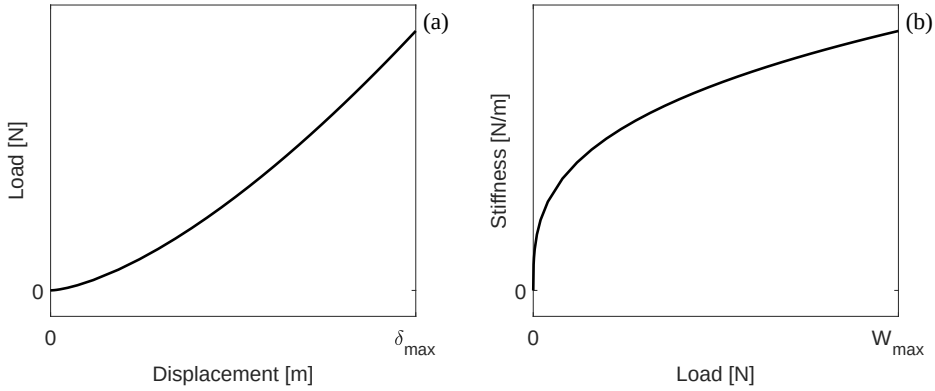

Figure 2: Typical stiffness characteristics of a rolling bearing for a given position of rolling elements: (a) load - displacement relationship; (b) stiffness - load relationship.

number of contacts among complex geometries. Furthermore, this quantity is non-linear, i.e. the displacement does not vary linearly with the load, as it may be appreciated in Figure 2.a. In particular, Figure 2.b shows that stiffness presents a steep trend for lower loads and assumes a more linear trend for higher loads. As a result, stiffness needs to be evaluated at different loads in order to determine the non-linear load-displacement relationship within the load range of interest. In addition, stiffness is also time-dependent, as the load-deflection relationship depends on the position of the rolling elements with respect to the load direction. In fact, if the load is stationary, the stiffness varies over time during the bearing service life since the position of the rolling elements with respect to the direction of the load constantly changes. This variation is periodic if the rotation speed is constant, and it is equal to the the spacing between two consecutive rolling elements. This phenomenon is usually referred to as Varying Compliance (VC) effect in the literature [12]. It is particularly relevant in the dynamic behaviour of machinery, since the periodic stiffness variation induces vibrations on the system even when the bearing does not present any type of fault.

REBs stiffness characteristics is also influenced by the amount of radial clearance, as it modifies the angular extent of the load zone [6]. This translates to a variation of the number of loaded rolling elements and the maximum

load on the most loaded one. In particular, a positive clearance reduces the stiffness due to an increase in the load zone extent, while a negative clearance leads to an increase in stiffness caused by a diminished angular span of the load zone. Furthermore, stiffness also depends on the magnitude of the rotation speed, but only for high velocities, for which inertial effects may not be neglected [13]. This dependency, in fact, is negligible for low and moderate speed values [10, 14]. Finally, REBs stiffness is influenced by temperature [15] and further depends on axial and radial preloads [16]. Concerning this latter aspect, axial preload is commonly applied in angular-contact ball bearings in order to reduce their non-linear characteristic [6]. Radial preload, on the other hand, is employed to obtain a greater number of loaded rolling elements and consequently reduce the maximum load applied on the most loaded one. In conclusion, the value of this peculiar parameter is influenced by several factors. Therefore, authors have developed numerous methods to evaluate REBs stiffness, including experimental techniques, analytical and finite-element models [17]. However, its estimation is indeed challenging and it is characterized by a high degree of uncertainty.

Moreover, despite all the advantages introduced by the employment of REBs, these components are subjected to one important limitation. In fact, two of the most prominent bearing analysts, i.e. T.A. Harris and M.N. Kotzalas, state, in their famous effort "Rolling Bearing Analysis - Essential Concepts of Bearing Analysis" [6], that:

"[...] even if rolling bearings are properly lubricated, properly mounted, properly protected from dirt and moisture, and otherwise properly operated, they will eventually fail because of fatigue of the surfaces in rolling contact."

In other words, they declare that bearing failure is an unavoidable phenomenon that must be always taken into account, as this event takes place independently from the application: a bearing will inevitably fail after a certain number of working cycles due to rolling contact fatigue. The local stress in the contact zones, in fact, is remarkably higher compared to other structural components [18]. By neglecting lubrication effects, it possible to assume that its magnitude and distribution are governed by the Hertz contact theory [19]. Therefore, since the bearing components are cyclically loaded, a subsurface crack eventually

generates inside one of them and progressively propagates to the surface producing a pit or a spall. Then, bearing defects are classified on the basis of their location as outer, inner, rolling element and cage faults [20]. Furthermore, they may subdivided in localised, extended and distributed defects [21]. Localised defects include spalls, cracks, pits, debris and lubricant impurities. Extended defects are generated from localised defects to due to wear of the fault edges caused by the cyclic passage of the rolling elements in the defect area. Finally, distributed defects comprise surface roughness and waviness, misaligned raceways and off-sized rolling elements [22].

A damaged bearing induces vibrations on the system that may possibly lead to the failure of the entire machinery. As a matter of fact, the wide employment of REBs translates to the fact that they are one the most frequent reasons for machine breakdown [23]. Consequently, throughout the years these components have been extensively studied by a number of researchers, in order to thoroughly understand their properties and their dynamic behaviour. Within this framework, particular effort has been devoted to the development of dynamic models aimed at describing the dynamics of both healthy and faulty REBs. These models allow to determine the vibratory signature produced by an healthy or defective bearing. In this latter case, the model should be able to assess the vibration signal for a variety of defect shapes and operative conditions, i.e. shaft speeds and loads. Typical experimental acceleration signals of healthy and damaged bearings are depicted in Figure 3.a and Figure 3.b, respectively. The defect, in this case, is a stationary outer ring fault. These plots highlight the different vibratory response of the system when a localised defect appears in the bearing. In particular, Figure 3.b. shows that the presence of the fault introduces distinct pulses which are caused by the impact of the rolling elements with the defect edges. To model these phenomena, LP models are employed by the vast majority of researchers [20]. Indeed, these type of dynamic models proved to be extremely effective in taking into account the complexity of the contact phenomenon and the time-varying characteristics of bearings. In this regard, several models have been proposed in the last decades. However, this modelling approach still provides challenges to be overcome for the interested scholars. For instance, it is often difficult to reliably estimate all parameters needed as input in these kind of models. This is due to the fact that there are quantities which are intrinsically difficult to measure, e.g. damping

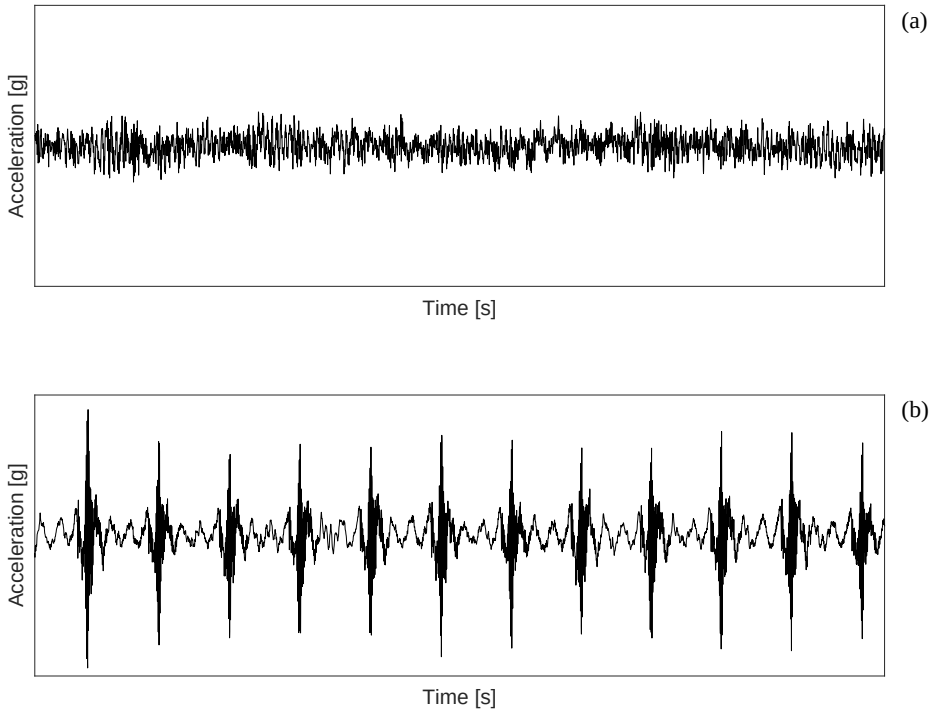

Figure 3: Examples of acceleration signal produced by a REB: (a) healthy bearing; (b) bearing with a localised defect on the outer race.

values, or associated with physical quantities that are not directly measurable, e.g. stiffness of artificial high-frequency resonators introduced for modelling purposes [24].

Dynamic models are extremely important for the development of diagnostic schemes and prognostic procedures [25]. Diagnostic techniques involve the detection and identification of a defect during the operative life of the component, while prognostic models aim at estimating the Remaining Useful Life (RUL) of the system [26]. These two aspects are integrated within the so-called Prognostics and Health Management (PHM) of mechanical systems, which is a topic that has gained increasing attention from the industrial world in recent years [27]. In particular, it has led to the rise in popularity of Condition Based

Maintenance (CBM) approaches for the maintenance of industrial plants and mechanical components. CBM is a maintenance strategy which consists in the real-time monitoring of the health condition of the system in order to plan the maintenance based on its actual health state [28]. This approach differs sensibly from traditional maintenance methods, i.e. preventive and reactive maintenance. The former consists in relying on historical data and life estimation computations to replace the components at scheduled intervals. In this case, however, the risk is to substitute the monitored components way before they show signs of damage, leading to a waste of resources and excessive machine downtime. The latter approach suggests to intervene only when the failure occurs. This approach is particularly dangerous since it may lead to the catastrophic failure of the whole system if machinery is not promptly stopped. On the contrary, a CBM approach greatly increases the efficiency of the maintenance strategy compared to traditional methods. As a matter of fact, CBM allows to sensibly reduce machine downtime and consequently decrease the costs associated with maintenance procedures [27]. Although CBM introduces further costs due to the implementation of the required technology, in the long-term the benefits remarkably overcome the starting additional expense. Additionally, CBM permits to increase the overall security of the mechanical system due to the reduction of unexpected failures and diminish the mean maintenance time due to the possibility to localize the fault from the recorded data [29].

CBM may applied to REBs, provided that suitable prognostic models are available. In this regard, the life of a REB may be subdivided in different Health States (HSs) on the basis of the time-varying trend of a selected health indicator. A common and simple indicator to detect the change in HS is the RMS value of the signal, as the progression of fault severity leads to an increase of this parameter. By observing the RMS evolution over time, bearing operative life may be divided in two or more different HSs depending on the degradation history of the bearing [28]. In order to illustrate these different behaviours, two degradation histories extracted from the well-known IMS dataset [30] are depicted in Figure 4. Specifically, Figure 4.a shows an example of degradation trend with two HS, i.e. healthy stage and unhealthy stage. In the first part, only random fluctuations occur in the RMS since no fault is present in the system. In the second HS, the RMS starts to increase as soon as the bearing

begins to deteriorate. Therefore, RUL should be calculated when the bearing enters in the second stage, as there are no sufficient information to compute its value in the first HS. The initial time from which RUL may be computed is referred to as first predicting time (FPT), which also marks the moment when bearing enters in the unhealthy stage. Differently, Figure 4.b shows a RMS history subdivided in three stages, i.e. an healthy stage, a degradation stage and a critical stage. While the first stage is the same as in the previous case, the two subsequent ones are not. In the degradation stage, the RMS trend is characterized by an "increase-decrease-increase" trend. This phenomenon was studied by T. Williams et al. [31], which defined it as a "healing" process of the REB. In this stage, the defect is initially generated as small spalls and cracks which are subsequently smoothed by the repeated passage of the rolling elements. Eventually, as the damage expands to a broader area, the bearing shifts to the third and last stage, i.e. the critical stage, which is characterized by a rapid increase of the RMS.

In light of the presented discussion, this thesis aims at covering several aspects within the context of bearing modelling. The effort is carried out in order to improve the available methods and to propose novel approaches to tackle REB modelling both in case of static and dynamic simulations. In particular, this work deals with three different topics and their associated challenges:

- It investigates bearing stiffness estimation techniques through the aid of finite-element software. Therefore, effort is placed on providing numerical instruments to determine this crucial quantity which greatly affects the dynamic behaviour of mechanical systems. To this end, a technique for the generation of computationally efficient load-dependent meshes is developed.
- It inquires the dynamic modelling of bearings with faults. In particular, it proposes a robust procedure devoted to the assessment of unknown parameters in REBs lumped parameter models in presence of localised defects. Major effort is placed on the discussion of dynamic modelling of REBs by means of dedicated LP models.
- Lastly, it examines potential numerical methods for bearing prognostics. In this regard, a parameter related to the evolution of the defect during bearing life is introduced, namely the Equivalent Damaged Volume (EDV). An algorithm capable of estimating its value from real bearing degradation

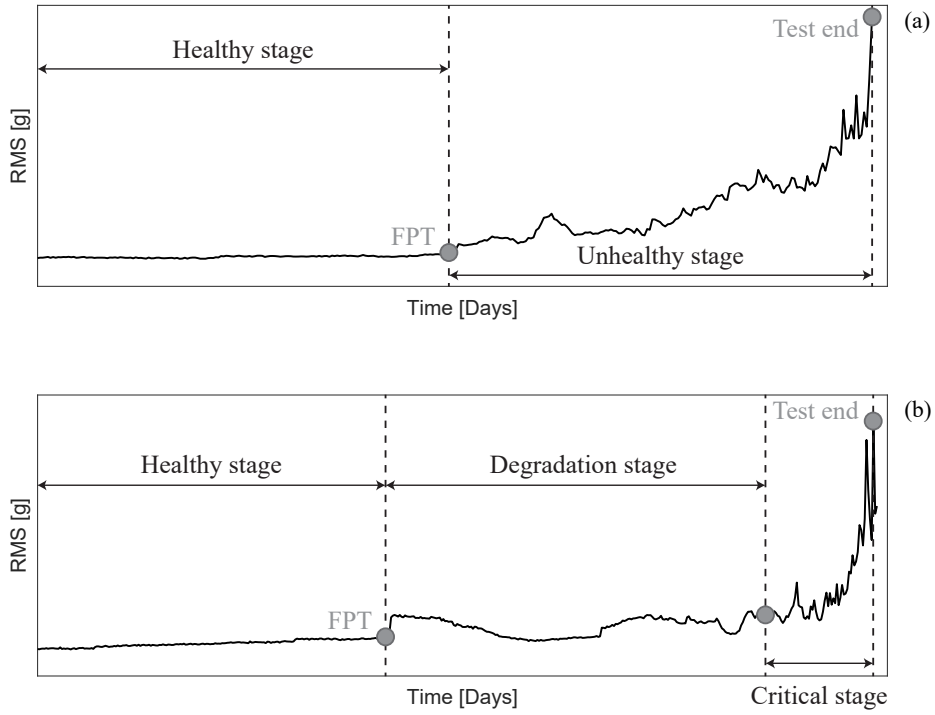

Figure 4: RMS trends representing different degradation processes of a REB: (a) two-stages life; (b) three-stages life.

histories is detailed. Notably, the developed method involves the integration and expansion of the LP model for localised defects in order to also consider extended defects. Finally, two prognostic models based on the EDV algorithm are proposed.

To support the findings of the research, this thesis also provides experimental results for a number of tests which were carried out on a dedicated bearing test rig. Specifically, two types of tests have been performed: stationary tests on faulty bearings with artificial defects and run-to-failure tests on initially healthy bearings. The former are employed to validate the procedure for the estimation of the parameters of the LP model with localised defects, while the latter are exploited to evaluate the efficiency of the EDV algorithm and the associated prognostic models.

Before proceeding with the discussion related to the three aforementioned subjects, this introductory chapter provides a thorough review of the state-of-the-art for these topics. Then, research objectives are outlined along with the structure of the thesis.

1.2 STATE-OF-THE-ART IN BEARING MODELLING

REBs are one of the most frequently employed components in rotating machinery. As such, they have been extensively studied by a number of researchers throughout the years, and a vast literature regarding their characteristics is currently available. The necessity to deeply understand these components arises from the fact that the noise, vibration and harshness (NVH) behaviour of mechanical systems is greatly affected by REBs dynamics. In fact, it is well known that these components cover a major role in transmitting vibrations through the system in which they are inserted [9]. As a consequence, it is crucial to accurately estimate parameters related to REBs, especially if they are needed to be employed in dynamic models of complex machinery [32]. Within this framework, one the most important quantities to be determined is the bearing stiffness, whose estimation has been tackled by means of several techniques. Moreover, scholars devoted a considerable effort in the dynamic modelling of REBs, especially to describe the characteristics of faulty bearings. This allowed to develop accurate REB models that were employed for both diagnostics and prognostics purposes. Concerning this last matter, the field of REBs prognosis gained remarkable traction in the last decades thanks to the potential application of these models in industrial contexts.

Therefore, an extensive literature review on bearing modelling is proposed in this section. In particular, the survey is divided in three parts, which correspond to the topics covered in the thesis. First, the subject of bearing stiffness estimation is covered, mainly focusing on FE methods. Then, attention is placed on bearing dynamic modelling for either healthy and faulty REBs. In this regard, particular effort is dedicated to the description of LP models. To conclude, the final part is devoted to REBs prognostics. This closing section reports examples of PBMs, DDMs and hybrid models.

1.2.1 *Stiffness estimation methods*

Radial stiffness is one of the main parameters that needs to be evaluated in a REB. In fact, it is an important quantity that greatly affects the outcome of dynamic models. For instance, bearing stiffness values are needed for geartrain models, in which connection between gears and frame is realized through REBs [11]. Despite the amount of literature on the topic, the inherent complexity of these components still provides challenges to overcome for researchers and designers when determining stiffness values. As a result, over the years the estimation of the non-linear stiffness-load relationship has been faced by a vast number of experimental, analytical and numerical approaches.

Experimental techniques are commonly divided into direct and indirect methods, depending on the employed procedure. The former requires the direct measurement of the displacement [33], while the latter necessitates other techniques, e.g. modal analysis [34]. Concerning the first approach, R. Madoliat and M. F. Ghanati [35] used a general purpose tensile testing machine to apply the load on the bearing. An ad-hoc housing was constructed to accommodate the bearing inside the loading device and a shaft was fitted in the bearing bore. Tests were first performed without the bearing to determine the fixture stiffness. Then, when the REB was mounted, bearing stiffness was determined by subtraction. Example of application of the second approach may be found in the work of N.J. Ali and J.M. García [36]. The authors developed a test rig in which the load was applied radially and axially through two pneumatic pumps. The system was then excited by an impact hammer, and the frequency response of the structure was employed to extract the stiffness. Moreover, W. Jacobs et al. [37, 38] presented a test rig aimed at determining bearing characteristics through modal analysis. The test bench design was optimized to measure bearing stiffness under highly varying loads both in static and operative conditions.

Concerning the analytical approaches, formulae are either based on a rigorous mathematical procedure, typically based on the Hertz contact theory [6, 39, 40], or derived from experimental measurements [41, 42]. In this context, it is noteworthy the double-effort of T.C. Lim and R. Singh [43, 44]. In the first part, they provided multiple methods to analytically determine bearing stiffness

and complemented the work with parametric studies to assess the influence of several parameters on the results. Besides, in the second part, the previous formulation was integrated with a discrete system model which included LP and FE techniques. They further expanded the theory in subsequent works to study the dynamic behaviour of a geared rotor system [45] and to estimate the coupling loss factor between shaft and casing by means of statistical energy analysis [46]. Afterwards, A. Gunduz and S. Rajendra [47] introduced an approach based on the Hertzian theory aimed at the estimation of the stiffness matrix for double row angular contact ball bearings. In particular, they provided results for the back-to-back, face-to-face, and tandem arrangements. An enhanced version which is also employable for duplex bearings was subsequently presented by T. Xu et al. [48]. Moreover, D. Petersen et al. [49] further included a localised defect in their modelling. Their approach was exploited for defect size estimation in operative conditions by analysing the excitation produced by the change in stiffness caused by the fault. Based on their work, H. Cheng et al. [50] extensively studied the time-varying stiffness introduced by the presence of a defect. Additionally, they further investigated the VC phenomenon to examine its coupling with the local fault. B. Fang et al. [51] inspected the stiffness characteristic under speed-varying conditions. For this purpose, they employed an analytical model derived from the formulation of T.J. Royston and I. Basdogan [52]. Later, the same authors [53] investigated the influence of the off-diagonal terms in the stiffness matrix. In order to perform such a task, they proposed a model for ball bearings which took into account the influence of rotating speed and loads. Then, they analysed the influence of stiffness terms on the coupled rotor system dynamics. Furthermore, Q. Niu et al. [54] proposed a 5 DOFs analytical model for stiffness matrix computation and employed it to assess the effect of different operative conditions and the variations of bearing parameters.

Finally, among the numerical techniques, the most common approach involves the employment of in-house or commercial FE software to model the bearing under exam. Within the context of REB modelling, FE simulations are usually exploited to determine rings displacement and contact stresses. A common issue in this kind of modelling, however, lies in the complexity of the contact phenomenon. In fact, the need for a sufficiently fine mesh in the proximity of the contact area has a significant impact on simulation time and the excessive

number of required mesh elements might lead to an excessive computational burden [55]. Therefore, researchers strived to simplify the problem by using a variety of approaches. The most straightforward solution to this concern consists in reducing the original 3D model to a planar 2D problem. To this end, H. Zhao [56] described a contact algorithm to model roller bearing contacts in 2D simulations and evaluated the influence of various parameters on load distribution. N. Demirhan and B. Kanber [57] investigated stress and displacement distribution on a roller bearing by introducing a 2D mesh with plane strain option to simplify the model. Model validation was carried out by comparison against theoretical and experimental data. X. Hao et al. [15] proposed a 2D FE model that considered temperature effects and clearance change to examine their influence on displacement, stress and bearing stiffness. The relative displacement between inner and outer ring obtained on a test rig was employed for validation. Good agreement was found between FE and experimental results, while analytical formulae departed from them. Although 2D approaches allow to obtain good stiffness estimate, this kind of meshing is not applicable to every bearing type. For instance, the computational domain of deep-groove ball bearings, self-aligning ball bearings and tapered roller bearings may not be reduced to a bi-dimensional plane. Therefore, 3D approaches are more suitable to tackle the problem of REBs modelling as a whole without losing generality in the procedure.

Concerning 3D simulations, several methodologies were established. In particular, a number of scholars focused their effort in proposing methods to reduce the number of contacts in REBs models. Within this framework, L. Kania [58] employed truss elements to replace rollers in slewing bearings. This technique allowed to greatly reduce simulation time at the expense of additional effort in the pre-processing phase. A similar approach was proposed by A. Daidié et al. [59]. They analysed load distribution by employing non-linear traction springs instead of truss elements. The efficiency of this method encouraged other researchers [60, 61, 62, 63] to exploit it for slewing bearings modelling. Techniques devoted to removing contacts in lower size bearings were also studied by L. Molnar et al. [64]. In particular, they proposed two methods to avoid modelling contacts between rollers and rings in needle bearings. The first one was similar to the one developed by A. Daidié et al. [59], as rollers were replaced by springs. The second one consisted in substituting the entire

volume between rings with a fictitious equivalent material replicating the same behaviour of the row of rollers. Obtained radial displacements were compared with the full model involving contacts, demonstrating their capability to considerably reduce the computational time.

Despite these techniques being successful in speeding-up the solution process, other researchers followed different strategies to model the phenomena occurring within the contact area. Y. Guo and R.G. Parker [10] developed a procedure involving a combined surface integral and finite element method to solve the contact problem in rolling and ball bearings. They computed radial, axial and tilting stiffness to obtain a fully populated 6x6 matrix including cross-coupling terms. Results were compared against data available from the literature. F. Massi et al. [65] set-up 2D and 3D simulations of ball bearings to compute the contact stress due to specific boundary conditions and relate them with bearing degradation. They reduced the size of the problem by modelling a portion of the bearing, with only one roller in contact with the two races. The authors noticed that the relative error between 3D and 2D simulations was relevant due to conforming contact between ball and races. R. Lostado et al. [17] studied the contact stress in tapered roller bearings. They developed a procedure to adjust the original mesh by generating subsequent non-linear sub-models with increasingly smaller mesh densities. Relative displacement between raceways was also analysed and compared to experimental data, showing good agreement. The procedure was utilized by R. Martinez et al. [66] in combination with machine learning techniques to determine the optimal working conditions of the device. S. Murer et al. [67] presented a FE model of their experimental set-up to assess the relevancy of using capacitive probes for in-situ measurements of bearing deflection. S. Li [68] developed a software to compute the contact stress in ball and roller bearings by exploiting a novel contact algorithm. Stress distribution on contact areas were found to be different from results reported in previous studies [10] and analytical formulae.

As it may be observed from the proposed literature review, there is a tendency to reduce the size of the computational domain of the problem whenever REB modelling is involved. Typical strategies include taking advantage of symmetry planes [17, 56, 58, 59, 60, 63, 64, 66], removing unloaded rollers [56] and replacing contacts with equivalent elements [58, 59, 60, 61, 62, 63].

Unloaded rollers are kept by most researchers but removed by others, e.g. H. Zhao [56] and S. Murer et al. [67]. Cage is commonly neglected, but e.g. Murer et al. [67] accounted for its effect by employing rigid connectors between rollers. Load is applied on the center of the shaft [15, 56, 57, 64], on rings [59, 63, 65] or on a central node rigidly connected to the nodes on the surface of the inner ring [60, 61, 62]. Problems are solved by employing 2D and 3D approaches, usually exploiting quadrilateral or hexahedral elements, respectively. Meshing with tetrahedral elements is rare, and differences between linear and parabolic elements are not addressed. Convergence check are regularly performed, but some researchers as N. Demirhan and B. Kanber [57] only tested it for one load value, albeit convergence rate depends on applied load, especially at low force values. Concerning the post-process, radial bearing stiffness may be computed considering the approach of bearing rings [15, 17, 57] or the displacement of the shaft axis [64].

In conclusion, the reviewed methods proved to be successful in different aspects of bearing analysis and in particular to estimate REB stiffness. However, the proposed survey highlighted the large number of different available methodologies and the consequent lack of uniformity in the employed approaches.

1.2.2 *Dynamic modelling of faulty bearings*

As underlined in the introductory section of this chapter, REBs are prone to the development of defects after a certain number of working cycles. In fact, faults eventually generate in one of the bearing components because of rolling fatigue generated by cyclic passage of the rolling elements in the loaded zone of the bearing. This phenomenon leads to the onset of subsurface cracks that gradually propagates to the surface, therefore producing spalls or pits. A local fault greatly affects the vibration of the system and might remarkably alter its behaviour or possibly lead to its failure. Hence, throughout the years a number of researchers have tackled the problem of developing dynamic models capable of determining the vibratory response of systems comprising bearings subjected to this peculiar condition.

P. McFadden and J. Smith [69] were amongst the first to work on the topic by proposing an analytical model of the phenomenon. They studied the effect of point defects and modelled the response of the system by convolving the exponential decay response function with a series of impulses at the passing frequency of the rolling elements. As a result, this kind of modelling approach is commonly referred to as impulse train model. They later extended the formulation to multiple point defects [70]. Subsequently, N. Tandon and A. Choudhury [71] considered pulses of finite width whose shape were assumed to be rectangular, triangular of half-sinusoidal. Y.F. Wang and P.J. Kootsookos [72] also exploited an impulse train model to model bearing presence in machinery and estimate the bearing induced vibrations on an industrial component. D. Brie [73] improved the method by introducing a slight period variation between the impulses which allowed to better represent experimental observations. Therefore, this modified model was named quasi-periodic impulse train model, as it included minor random slip between impulses. R.B. Randall et al. [74] took advantage of this approach and analysed the results by means of cyclostationary techniques. The periodic and quasi-periodic impulse train models allowed to correctly predict the spectral components associated with the localised fault. However, a major drawback of both models is that they not take into account the VC effect of the bearing assembly.

Based on these investigations, the majority of subsequent works dealt with the development of LP models capable of describing the response of systems comprising faulty bearings. In these models, the influence of the defect was inserted by either adding a constant displacement to the balls rolling over the fault or by defining the contact forces generated by the balls when striking the edges or the bottom of the defective area. Within this context, S. Sassi et al. [75] employed a 3 DOFs model to assess vibration due to point defects by considering the shock behaviour of the fault. They also introduced a noisy response resulting from sliding friction and other possible disturbances. A. Rafsanjani et al. [76] analysed the effect of local defects on the stability of a rotor-bearing system. They modelled the faults as a series of impulses separated by their characteristic frequency. Their amplitude was related to the angular velocity and the loading condition at the point of contact. The model was employed by the authors to present the conditions for stable and unstable response to subdivide the main routes to chaotic motion. Differently, M. Behzad

et al. [77] dealt with the problem by considering the stochastic excitation generated by the surface roughness in the healthy and defective contact areas between rolling element and races. H. Cao et al. [78] focused on the behaviour of high-speed ball bearings by exploiting Gupta's model [79], thus associating 6 DOFs to each moving element. In addition, they also discussed the influence of single, multiple and compound defects on a cylindrical roller bearing in a subsequent paper [80]. Besides, G. Kogan et al. [81] proposed a 3D dynamic model to simulate faulty duplex bearing whilst underlining the differences in the response of corresponding single bearings. S. Khanam et al. [82] discussed a technique to generate the forcing function induced by the motion of rolling elements inside an inner race defect. The entry and exit events were formulated differently in order to take into account the impulsive nature of the latter. Based on the work of C. Sunnersjö [12], R. Yang. et al. [83] exploited the harmonic balance method to determine the vibration of a rotor-ball bearing system and discussed the presence of super-harmonics of the defect frequency in the resulting spectrum. Y. Qin et al. [84] developed a model for a high-speed faulty angular contact ball bearing where the defect influence was inserted through a B-spline fitting displacement excitation. They indicated that the duration of one impulse increased with larger fault extension and decreasing rotation speed. Moreover, acceleration amplitude was found to be more sensible to defect size compared to shaft rotation frequency. T. Gao and S. Cao [85] studied the paroxysmal impulse waveform generated by faulty inter-shaft bearing and demonstrated their dependence from certain speed ratios. R. Yang et al. [86] studied the influence of the resonance characteristics and rotor eccentric excitations during transients due to raceway faults. Despite the noteworthy contribution of the mentioned papers, a common characteristic of these modelling approaches is the arbitrary choice of damping values. In fact, they are commonly set in order to match some experimental data related to a few test conditions or on the basis of values employed in previous researches.

Within the context of LP modelling of REBs, many researchers focused their efforts on the definition of an analytical expression to account for the path travelled by the rolling elements inside a fault. Contrarily to other models, in which the fault was modelled by means of an additional constant displacement, some authors strived in defining an analytical expression capable of simulating the trajectory of the ball inside the defect. In this regard, N. Sawalhi and R.

Randall [24] improved the 2 DOFs model of S. Fukata et al. [87] and proposed a LP model of a gearbox with a defective self-aligning ball bearing. They described the modelling of inner, outer and rolling element defects. This demonstrated the effectiveness of LP models for the further integration with other components schematized in a similar fashion, e.g. geartrains [11]. Their formulation included the effective path travelled by the balls in the defective area. Other than the masses of inner and outer rings, they also added a fictitious additional mass to model a high resonant mode of the system. Its mass and stiffness were tuned on the basis of a resonance frequency observed on an experimental signal. However, no further indications on the selection of their values and on the influence on the system response were provided. Later, D. Petersen et al. [88] enhanced N. Sawalhi and R.B. Randall's model and discussed the effect of the rapid stiffness changes induced by defects with sharp edges on the system response. Their work was also further studied by L. Cui et al. [89] which investigated the relationship between the defect size and vibration response, therefore assessing the fault severity on the basis of the time between impacts. Moreover, F. Larizza et al. [90] employed the model developed by D. Petersen et al. to analyse the static and dynamic stiffness of the system. They reported that stiffness values rapidly changed for load values above a certain threshold dependent on defect shape and load value. The authors also examined the influence of entry and exit defect slopes in large outer ring defects by computing the contact forces through Love's equation [91]. M. Patil et al. [92] described the defect as half-sinusoidal and employed their model to numerically assess the effects of different defect size. J. Liu et al. [93], instead, exploited a piecewise function whose shape depended on the ratio between ball diameter and the characteristic dimension of the defect, i.e. its length and width. By employing a different approach, A. Moazen Ahmadi et al. [94] avoided making any assumption on the ball trajectory inside the defect. Instead, they considered the finite size of the rolling element, i.e. each ball had its associated DOF and its own dimension. This allowed to better describe the events taking place as the balls roll over the defect. A different formulation was proposed by J. Liu et al. [95], which exploited the quarter-space method to calculate the contact stiffness near the defect edges. The authors also advanced the definition for spalls with a shoulder, which were described as the combination of a sinusoidal and exponential function [96]. Moreover, they added the effect of a deformable interface between the

outer ring and the housing in a consequent work [97]. A further technique was exploited by F. Kong et. al [98], who utilized the estimated Hertzian contact area between rolling elements and races in the proximity of the contact area to determine the actual deflection of the contacting bodies. This approach denoted a change in the contact force before the entry in the defective area. More recently, Y. Jiang et al. [22] thoroughly discussed the modelization of 3D rectangular-like defects. Their inquiry led to establish the larger influence of the circumferential extent of the defect compared to its axial width. Latterly, S. Gao et al. [99] proposed a novel formulation for an asymmetric defect on an angular contact ball bearing by taking into account different defect dimensions and positions with respect to the path followed on a healthy bearing. They suggested that it is possible to mitigate the vibration response by reducing the circumferential extension of the defect or by diminishing the radial-axial load ratio. The cited researchers provided remarkable contributions on the analytical definition of the ball trajectory inside the defective area. However, even in these works the choice of damping values was based upon arbitrary selection.

In other works, the introduction of elastohydrodynamic lubrication (EHL) allowed for a different definition of the forces arising due to the presence of an intermediary fluid film between the balls and the raceways. Within this framework, Y.H. Wijnant et al. [100] were among the first to introduce this aspect in bearing modelling by means of a non-linear spring-damper model. Afterwards, J. Sopanen and A. Mikkola [101, 102] introduced a 6 DOFs model which included the elastohydrodynamic fluid film and a number of non-idealities, i.e. rings waviness, shaft misalignment and localised defects. D. S. Shah an V. Patel [103] discussed the influence of lubrication on their localised defect model. P. Yan et al. [104] included the effect of EHL and reported the impact of load and shaft speed on the lubricating film and consequently on the system stiffness. The effect of the EHL condition on stiffness and damping was also investigated by M. Luo et al. [105]. Furthermore, S. Mufazzal et al. [106] introduced the lubrication theory on a 2 DOFs model with a circular defect to analyse the amplitude of the first four harmonics of the defect characteristic frequency due to different defect size. Differently, V. Parmar et al. [107] included EHL along with race waviness to understand the influence of the number of waves. In addition, V. Parmar et al. [108] also examined,

in a ensuing paper, the response of spherical rolling-element bearings with localised defects under misalignment. It was observed that the rolling element could follow either an offset or an inclined trajectory while passing over the defect. In addition, they reported the effect of load and misalignment angle.

LP techniques were not the only methods utilized for the modelling of faulty REBs. For instance, a few authors exploited dynamic FE techniques to produce models of defective bearings. The efficiency of the FE method in bearing simulations, in fact, was previously underlined in Section 1.2.1. In particular, Z. Kiral and H. Karagülle [109, 110] described a bearing-housing FE model discretized by 3D tetrahedral elements. In this model, the loading condition was assumed to be generated by shaft unbalance. Consequently, the radial load distribution continuously changed around the outer ring circumference. The FE model was employed to estimate vibrations induced by both healthy and faulty bearings. Y. Shao et al. [111] proposed a similar model which included inner ring, outer ring and rolling element localised faults. They determined that, for the same defect dimension, the outer ring defect produced the highest vibration while the rolling element fault induced the lowest. Afterwards, A. Utpat [112] constructed a 3D FE model of a deep-groove ball bearing and inserted a localized defect either in the inner or the outer ring. The author reported the acceleration response on a node located on the outer surface of the outer ring caused by different defect sizes. S. Singh et al. [113] successfully implemented a FE technique to investigate the entry and exit events taking place when a rolling element interacts with the defect edges. Finally, a mixed formulation was employed by M. Tadina et al. [114], which modelled the inner ring as a 2 DOFs lumped mass while the outer ring was discretized by planar FEs. Therefore, differently from other LP models, the outer ring was made deformable. They employed this peculiar modelling approach to determine the effect of various faults under different shaft speeds.

Some researchers also investigated the problem of extended defects modelling. However, this matter received less attention compared to localised defect, and a fewer number of publications on the subject may be found in the literature. Within this context, N. Sawalhi and R.B. Randall [115] extended their previous work [24] in order to also take into account this type of fault. To perform such a task, they proposed a procedure for the generation of a rough surface

within the angular span separating the entry and exit edges of the defect. In fact, in their modified model the rolling elements are allowed to touch the bottom of the defect, which is modelled as an area characterized by a non-zero roughness. Later, their work was further extended by D. Petersen et al. [88]. They employed the model to assess the different bearing stiffness values due to the presence of localised and extended defects. Notably, they showed that an extended defect induces a slow variation of the stiffness terms compared to a narrow localised defect.

To conclude, this review exhibited various approaches commonly applied to model the dynamic behaviour of REBs with localised and extended defects. Although the majority of the detailed approaches produced remarkable results, a common issue found during this research was the lack of sufficient explanation for the choice of several parameters involved in the construction of the models. In particular, the selection of damping values and the inertial and stiffness properties of the resonant masses were found to not be thoroughly illustrated in most of the reviewed works.

1.2.3 *Prognostic models*

Prognostics tackles the problem of estimating the RUL of machinery, which is defined as the time length which separates the current time and the end of the useful life of a component [116]. The useful life is usually considered to have reached its end when the monitored health indicator crosses a fixed failure threshold [117]. This threshold may be either considered a constant line or a probability distribution. Although the second approach is generally more accurate, since it takes into account the system uncertainties, it also introduces an higher degree of complexity. As a result, the first approach is employed in most publications [28].

RUL prediction methods may subdivided in three major categories [27]: physics-based models (PBMs), data-driven models (DDMs) and hybrid models. PBMs describe the degradation process and the failure modes through mathematical modelling of the system physics. The development of these models demands a thorough understanding of the system properties, its degradation

mechanisms and its possible failure modes [118]. On the other hand, DDMs exploit statistical and machine learning techniques to assess the current and future HS of the system. Differently from PBMs, they do not rely on the knowledge of the system physics but only on the history of data acquired on the system. This category may be subdivided in two further groups, i.e. statistical model-based approaches and Artificial Intelligence (AI) approaches, depending on the chosen prediction technique [28]. Finally, hybrid models refers to approaches that combine both PBMs and DDMs in order to take advantage of the benefits of both methodologies. In some cases, however, some hybrid models are the combinations of different models that represent the same phenomenon, so to increase the robustness of the health assessment. These peculiar models are usually referred to as ensemble models [119]. This section reviews a number of works which concern these three types of prognostics models.

1.2.3.1 *Physics-based models*

Prognostic PBMs assess the future HS of the system by building a mathematical model based on the physics of the system. The parameters of the physics model are usually correlated to the stress on the component and the properties of the material [27]. PBMs are able to perform RUL prediction by tracking the degradation process of the system. Therefore, a complete understanding of both the degradation mechanisms and the failure modes of the components under exam must be acquired before constructing a PBM. In this regard, it is worth clarifying that degradation mechanisms are processes which lead to the failure of the mechanical system, while failure modes explain why and how a function of the system may be no longer fulfilled [120]. Besides, degradation models require as input information about the current severity of the system deterioration. These data are acquired by means of sensors mounted on the monitored machinery.

The Paris-Erdogan (PE) model is one of the most widely employed physics models in the RUL prediction of machinery. It was first proposed by P. Paris and F. Erdogan [121] in order to represent the development of crack growth. Cracks, in fact, are a typical consequence of fatigue damage that ultimately brings the system to its failure. According to T. Tinga [120], the process of

crack growth may be subdivided in three regions. In the first one, the crack grows slowly and its length is not considered to be significant. In the second region, the crack dimensions become significant and its size steadily increases. Lastly, in the third phase the crack becomes unstable and rapidly propagates until failure. Within this context, the PE model describes the second region of crack growth. Its validation is based on a number of experimental observations that allowed the two researchers to relate the number of cycles and the crack length through the combination of material parameters and the stress range undergone by the component. Although the PE model does not take into account the instability of the crack and its initial length, its implementation has been proved successful in PBMs. In fact, throughout the years their model was extended and refined, and a number of researchers applied it to the field of machinery and REBs prognostics.

In this regard, Y. Li et al. [122] proposed a model similar to PE. In their formulation, the number of cycles was related to the defect area instead of the crack length. They developed an adaptive algorithm to fine tune the parameters through on-line comparison with measured defect size. They subsequently refined the model in a consequent effort [123] by introducing a stochastic term in the form of a lognormal random variable. A similar approach was adopted by M.R. Hoeprich [124] which introduced an equation similar to PE in order to represent the growth of defect areas in REBs. Their model was extended by M.N. Kotzalas and T.A Harris [125] to include thermal effects and later validated by an experimental campaign carried out by S. Li et al. [126]. Y. Lei et al. [127] and J. Wang et al. [128] modified the PE formulation into an empirical model for RUL assessment. The former presented a method based on particle filter that integrated measurements and physics model into a state-space framework to accurately consider the nonlinearity and the uncertainty of the degradation process. The latter described a procedure to integrate multiple features and correlate them with the degradation history. Similarly, L. Liao et al. [129] proposed a combination of features in order to define a global one which was able to describe the deterioration process. For this purpose, they exploited a genetic algorithm to extract a feature with the optimal monotonicity. J. Sun et al. [130] enhanced the PE model to a state-space model and proposed the optimal degradation state estimation through the employment of a Bayesian framework. D. Xu et al. [131] developed two modified versions of the PE

model for application in rolling bearings. M. Corbetta et al. [132] described an approach to estimate the crack progression by combining a PBM with a particle filter for parameters estimation. In a subsequent work [133], the authors integrated the PE model into a probabilistic method which exploited a sequential Monte Carlo sampling combined with the probability density function of the model parameters. Similar approaches where also presented by F. Cadini et al. [134] and E. Zio and G. Peloni [135]. More recently, Y. Lu et al. [136] proposed a PBM for REBs, which employed the realized volatility of the signal as an indicator for the system HS. L. Saidi et al. [137] integrated the PE law along with a Kalman smoother to perform prognosis of high-speed REBs mounted on wind turbines.

Despite the popularity of the PE model, several other models based on other approaches have been proposed. J. Qiu et al. [138] developed a stiffness-based prognostic model which related the natural frequencies and the vibration amplitude to the failure lifetime of a REB. They described the damage accumulation by either a linear damage rule, damage curve approach and a double linear damage rule. Differently, F. Sadeghi et al. [18] proposed a model for the evaluation of rolling contact fatigue in REBs. In a similar fashion, some authors employed contact stress analysis to describe the wear phenomenon on bearing components. For instance, C.D. Begg et al. [139] and D. Chelidze and J.P. Cudumano [140] based their models on system dynamics, while S. Marble and B.P. Morton [141] took advantage of contact stress analysis. P.K. Gupta and E.V. Zaretsky [142] proposed stress-based life models for REBs and were able to develop a formulation which related bearing life to the maximum subsurface shear stress and the amount of stressed volume. A completely different technique was utilized by T. Slack and F. Sadeghi [143] who developed a 2D explicit FE model to model the crack subsurface initiation and its propagation to the surface. In order to perform such a task, the pressure distribution was continuously moved across the domain so to replicate the complete time history of the subsurface stress. Although applied for other components, FE analysis was also employed in combination with a modified PE law by M.J. Pais and N.H. Kim [144] to predict the fatigue crack growth in aerospace panels. On the other hand, an original methodology was proposed by L. Cui et al. [145]. In fact, they built a so-called performance degradation dictionary by performing a large number of numerical run-to-failure simulations through

the employment of a REB dynamic model based on the work of N. Sawalhi and R.B. Randall [24]. In each simulation, different fault parameters were employed to generate diverse degradation histories. RUL assessment was then performed by taking advantage of the similarity theory [146] which allowed to consider the uncertainty of the RUL estimation through comparison with experimental data. T. Wang et al. [147] presented a probabilistic framework in which fault presence was detected by means of spectral comparison between modelled and experimental signals. Then, RUL was computed through a least-square method and Bayesian inference was used to determine its distribution. Their unsupervised approach allowed to avoid manual parameter setting and their consequent fine-tuning. B. Yan [148] introduced a two-stage physics-based Wiener process. It took into account the fatigue crack mechanism and its growth, the component structure, assembly accuracy and the conditions of the working environment. Interestingly, the first region of crack growth was also considered in their model. Finally, it is worth citing that other crack growth formulations have been employed to model this phenomenon in other types of mechanical components. For instance, C.H. Oppenheimer and K.A. Loparo [149] utilized the Forman crack growth law ([150]) to estimate the RUL for rotor shafts prognostics.

Despite the demonstrated efficiency of PBMs in tackling PHM problems, they present some limitations. Most notably, their performance depends on the degree of understanding of the failure mechanisms and the estimation of model parameters. In fact, several degradation mechanisms present high difficulties in their modelling, making this approach unsuitable for complex machinery for which it may be challenging to fully understand the physics of damage [28]. As a result, although PBMs have been proven to be successful in REBs prognosis, they are not suitable in every PHM application. Furthermore, in contrast with DDM approaches, PBMs are non scalable between different systems. However, differently from DDMs, they do not rely on having large amount of previously acquired data on the system being monitored. In addition, PBMs permit to model the physics of the degradation mechanism and consequently allow for an adequate modelling of the system under exam, which may be employed for other purposes other than prognosis, e.g. certification of machinery [27].

1.2.3.2 *Data-driven models*

DDM prognostic model are based on data historically collected on the system [20]. Scholars have developed a large number of techniques, which may be broadly subdivided in two major categories: statistical approaches and AI approaches. Statistical approaches, often referenced as empirical model-based approaches [28], aim at estimating the RUL by constructing statistical models on the basis of experimental observations and empirical experience. Their typical outcome is a conditional probability density function (PDF) of the RUL based on experimental data [116]. Differently, AI approaches employ AI techniques to foresee the future degradation of the system. These methods work as "black-boxes" where data is fed to the AI and prognostic information is obtained as output. [151]. Therefore, the main characteristic that differentiates these methodologies from PBMs is that both approaches do not rely on the knowledge of the system physics. As a result, they are scalable between different systems, but they require large training data to attain reliable results as opposed to PBMs. Popular statistical DDMs include autoregressive (AR) models, random coefficient models, Markov models, Wiener process models, Inverse Gaussian (IG) models, gamma process models and Proportional Hazard (PH) models. Concerning AI approaches, often employed methods encompass Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Neuro-Fuzzy (NF) systems, Support Vector Machine (SVM), Relevance Vector Machine (RVM) and Gaussian Process Regression (GPR).

STATISTICAL APPROACHES Among the statistical approaches, AR models are widely employed for PHM purposes. Their underlying assumption is that the future state of the system may be seen as a combination of a linear function of past observations and random error terms [26]. This approach was utilized in bearing prognosis by W. Caesarendra et al. [152], who introduced the AR moving average model to improve the accuracy of the estimate. For the same purpose, Y. Qian et al. [153] extracted input data for the AR model by employing a Kalman filter. Later, Y. Qian and R. Yan (2015) [154] refined the previous effort by combining a multi-order AR model with a particle filter algorithm. Differently, L. Cui et al. [155] employed Kalman filtering to automatically match different degradation phases. The functions at the base of the procedure were linear in the normal deterioration phase and non linear in the final stage of the component life. The proposed algorithm was able to

autonomously switch between the two functions. In a subsequent effort, the same authors [156] further expanded the available techniques by introducing a time-varying particle filter algorithm constructed by a combination of an adaptive selection rule and a sliding window. Through this procedure they were able to estimate bearings RUL. The main strength of AR models is their easiness of implementation. However, their performance is highly dependent on the quality of available data.

Another common type of models are random coefficient models, which describe the randomness of the deterioration process by introducing normally distributed random coefficients. In this regard, N. Gebraeel and J. Pan [157] developed an exponential model which comprised random errors. Model parameters were determined by means of a Bayesian approach. Their model was later enhanced by Y. Wang et al. [158] to further ameliorate the parameter estimation process. X. Jin et al. [159] tackled bearing prognostics by proposing a model built as the sum of two exponential functions. First, an AR filter was employed to eliminate spectral components unrelated to faults. Subsequently, statistically unhealthy signals were analysed to determine the RUL of the component. Although these models consent to determine the PDF of the RUL, the base assumption of Gaussian randomness restricts their applicability.

Markov models are extensively employed in PHM thanks to their capability to provide an efficient compromise between computational time and accuracy of RUL estimation [160]. They are based on the hypothesis that the degradation process of machinery may be described by a finite number of transitions between different states that follow the principle of the Markov chain [161]. However, the vast majority of these techniques fall in the category of the so-called Hidden Markov Models (HMMs). In fact, in contrast with traditional Markov models, HHMs describe the evolution process of "hidden", non-directly observable states which are typical of machinery in working conditions. To overcome this restriction, the relationship between physical observations and the actual degradation is defined by means of a stochastic approach. Examples of applications of HMMs are described in the works of R.B. Chinnam and P. Baruah [162] and E. Ramasso and T. Denoeux [163]. Recently, E. Soave et al. [164] improved the HHM technique by introducing a Generalized Gaussian Hidden Markov Model (GGHMM). To this end, they considered possible

distribution variations among different states by defining the probabilistic function as a mixture of generalized Gaussian distributions. The GGHMM method was successfully employed for the assessment of bearing degradation histories. Variants in the form of hidden semi-Markov models were also proposed to improve the flexibility of HMMs, as shown by K. Medjaher et al. [165] for bearing prognosis and by Q. Liu et al. [166] for predicting future HSs of hydraulic pumps. Although their demonstrated effectiveness for RUL estimation, HMMs rely on the assumption of Markov property, i.e. the future state does not depend on previous observations but only on the current state. Therefore, their employment is somewhat limited in real applicative scenarios.

Furthermore, Wiener process models are built by combining a drift term plus a diffusion term which follows Browning motion [28]. First proof of their application in PHM may be found in the work of K.A. Doksum and A. Hóyland [167]. Then, examples of employment in REB prognosis may be found in the works of N. Li et al. [168] and Y. Wang et al. [169]. Moreover, J. Wen et al. [170] employed a nonlinear Wiener process model to assess bearing RUL and quantify the associated uncertainties. Similarly to HMMs, the application of Wiener process models is limited by the need to assume Markov property, which is an hypothesis that is not always applicable in industrial applications. Despite this issue, they allow for an efficient description of the temporal variability of the deterioration history.

Similar to the previous models, IG models are based on the assumption that the degradation history is characterized by independent increments. However, in this case, it is hypothesized that they follow an IG distribution instead. In this regard, it is worth mentioning the works of N. Chen et al. [171] and W. Peng et al. [172]. Likewise, gamma process models are based on a assumption similar to IG models but increments rather follow a gamma distribution. Both approaches are based on the Markov hypothesis and are therefore limited in their use. Furthermore, they may only describe monotonic processes. Despite these limitations, they accurately take into account the time variability of the degradation and are able to simultaneously take into account various random effects.

Finally, PH models were introduced by D.R. Cox [173] by assuming that it was possible to subdivide the hazard rate of a system, i.e. the rate of fault for a system at a certain operating time, into two multiplicative factors, namely a base hazard function and a covariate function. In other words, the hazard rate was hypothesized as the product of a deterministic and a stochastic component. His method was applied by D. Banjevic and A.K.S. Jardine [174] to determine the RUL through a Weibull PH model. Later, L. Wang et al. [175] proposed a PH model to forecast bearing RUL. First, they extracted suitable features from on-line measurement and then fed them to a neural network which allowed to predict their future values. Eventually, a PH model was constructed to assess the survival function. In general, PH models allow to obtain accurate predictions but at the expense of a remarkable computational burden.

AI APPROACHES In the field of PHM, ANNs are one of the most popular AI approaches. They are constructed as several nodes linked in a complex structure aimed at reproducing the working process of the human brain. Within this context, the most popular ANN is the feed-forward neural network (FFNN), whose potential has been demonstrated by the works of C. Lu et al. [176] and L. Xiao et al. [177]. The latter authors successfully employed an ANN for the estimation of bearing HSs. Notably, their procedure allowed to predict the future degradation without the need to rely on prior histories of failures nor suspensions. Moreover, C. Sbarufatti et al. [178] added sequential Monte Carlo sampling to determine the RUL of fatigue cracks. M. Behzad et al. [77] also took advantage of a FFNN and utilized a feature based on the high-frequency vibration of a ball bearing for REBs prognosis. In recent years, authors began to employ a ANN variant called Convolutional Neural Network (CNN). Among those, Y. Shang et al. [179] presented a CNN to learn spatial features from the monitored data of a REB and, after further manipulations, were able to extract the degradation trends. Besides, W. Li et al. [180] exploited a CNN to tackle the problem of bearing prognosis under variable operative conditions. ANNs are well-suited to learn complex non-linear relationships and are therefore indicated for complicated systems. However, they require a large number of training data to ensure reliable estimates.

On the other hand, a NF approach may be described as fuzzy-logic system in which membership functions are optimized by ANNs [181]. This methodology

was implemented in the field of prognostics by J.S.R. Jang et al. [182] and was later employed by a number of authors, including C.Chen et al. [183] and D. Zurita et al. [184]. Similarly to ANNs, these techniques require a massive amount of training data.

A different approach is employed by SVM, which is a method based on the statistical theory developed by V.N. Vapnik [185]. They consist in learning algorithms employable for classification and regression analysis. In the context of REBs prognosis, they have been exploited for degradation prediction by S. Dong and T. Luo [186] and F. Sloukia et al. [187]. Nevertheless, the most frequent application in PHM is found in the form of support vector regression (SVR), as demonstrated in the works of E. Fumeo et al. [188], T.H. Loutas et al. [189] and M. Zhao et al. [190]. The aforementioned authors utilized SVR to estimate bearings RUL. However, SVM approaches have one limitation, i.e. they provide point prediction instead of a PDF. To overcome this issue, RVM were introduced [191]. They are based on the same concepts of SVM, but they are able to provide a probabilistic prediction instead. Both SVM and RVM are advisable for problems denoted by a small numbers of training data. Per contra, the major downside of these techniques is associated with the estimation of model parameters and the choice of kernel functions.

Eventually, GPR is a supervised machine learning framework that may be exploited for data regression [192]. Among GPR models for REBs prognosis, it is worth mentioning the efforts of S. Hong et al. [193] and S.A. Aye and P.S. Heyns [194]. These kind of models are characterized by a remarkable computational burden but permit to obtain satisfactory results for both small and large training sets.

1.2.3.3 *Hybrid models*

Hybrid models combine two or more different models in order to take advantage of the capabilities of diverse approaches. These kind of models are considerably rarer in the literature compared to PBMs and DDMs [28], mainly because of the inevitable increase in complexity introduced by the implementation of different models. On this subject, this paragraph reports some

hybrid-models worth of mention. Among those, F. Di Maio et al. [195] integrated a RVM with a random coefficient model to perform an exponential regression that allowed to calculate and continuously update the residual life estimation for bearing prognosis. A. Soualhi et al. [196], instead, detailed a procedure for REB prognosis based on a combination of a HMM and a NF system. While the former detected the imminence of the subsequent degradation stage, the latter predicted the intervening period before the next HS. A different approach was employed by J. Yu [197], who combined a stochastic model to construct the degradation history and employed particle filtering for RUL estimation. Furthermore, Y. Qian et al. [198] proposed a multi-scale approach in which a modified PE model was integrated with a phase space warping technique for bearing prognostics. While the PE model, i.e. the PBM, described the crack propagation on a slow-time scale, the phase space warping, i.e. the DDM, characterized the dynamical behaviour of the component on a fast-time scale. The combined procedure involved the construction of a tracking metric which related measurements on the fast-time scale to the degradation process on the slow-time scale. To conclude, an extensive literature survey concerning multi-model approaches in PHM may be found in the work of J.J. Montero Jimenez et al. [199].

The aforementioned discussion on prognostic models showed a large variety of approaches that are typically employed for RUL estimation. For each kind of method, advantages and disadvantages have been reported.

1.3 RESEARCH OBJECTIVES

The proposed literature review highlighted the broad number of available publications in the field of bearing modelling. Despite only covering a portion of the massive amount of papers that deal with this topic, the survey demonstrated the remarkable advance in bearing modelling and the complex techniques developed by researchers in the last decades. However, it also underlined possible improvements in some modelling aspects, which are consequently taken as motivation to carry out the research detailed in this thesis. Specifically, this work aims at answering the following three questions:

- Is it possible to develop a robust and efficient procedure for the finite-element simulation of rolling-element bearings? Can the proposed procedure allow for an accurate estimation of REB stiffness terms?
- What is an effective way to determine the model parameters that, in LP models of defective bearings, are hardly measurable or denoted by an high degree of uncertainty?
- Are there means to perform bearing life prognosis by utilizing a PBM developed on the basis of parameters extracted from a dedicated LP dynamic model?

These questions are answered one by one in a different chapter of this thesis. In particular, the following original solutions are offered in the dissertation:

- The finite-element simulation of REBs is tackled by proposing a procedure for the generation of load-dependent meshes of the component under exam. The method is developed with the primary goal of determining the stiffness of different types of REBs. The main contribution to the subject is the definition of mesh element dimensions on the basis of analytical formulae and in the subsequently proposed methodology for the estimation of stiffness terms.
- The unknown parameters in LP models of defective bearing are determined through a dedicated multi-objective optimization technique in which the objective functions are features calculated from bearing acceleration signals. The technique involves the comparison of experimentally computed values with numerical features evaluated by means of a LP model. The proposed technique aims at determining the best set of parameters that are able to characterize the system vibration under a variety of operative conditions and defect dimensions. Therefore, the novelty of the procedure lies in the proposed technique and its peculiar implementation in the context of bearing LP models.
- Bearing prognostics is tackled by proposing two PBMs based on a novel degradation-related parameter, namely the Equivalent Damaged Volume (EDV). This quantity is determined through comparison of experimental features with numerically generated maps of the same feature. In this work, the numerical maps are supplied by computing the RMS values associated

to different combinations of angular extent and depth of the defect. This process gives as output the EDV values, which may be further taken as input for further prognostics models. As a result, the original aspect of the proposed endeavour is the development of PBMs on the basis of a peculiar parameter which has been specifically introduced in this work.

1.4 ORGANIZATION OF THE THESIS

The thesis is subdivided as follows. Firstly, Chapter 2 describes the bearing test bench that was set-up at Engineering Department of the University of Ferrara. This chapter also details the two types of performed tests, i.e. stationary tests with localised defects and run-to-failure tests. The results of these tests are employed for the subsequent validation of the proposed numerical methods. Chapter 3 tackles the problem of bearing FE modelling and stiffness estimation. Specifically, it reports the proposed advanced meshing procedure for the generation of load-dependent grids. First, application of the Hertzian theory to REBs is explained. Then, the chapter details the types of elements to be employed for meshing, their optimal size, exploitable modelling hypotheses and computational domain reduction strategies. Demonstration of the capabilities of the proposed modelling procedure are provided by computing the stiffness curves for two different reference bearings. Results are compared against analytical models, whose formulations are also reported. Finally, the chapter closes by evaluating the influence of additional factors on the computations, i.e. inclusion of the cage, load direction and magnitude of the radial clearance. The subsequent Chapter 4 presents a methodology for the assessment of unknown parameters in the context of LP modelling of faulty bearings. A theoretical LP formulation is provided to test the proposed procedure. The LP model is integrated with the proposed technique in order to determine its efficiency. To this end, the results of the stationary tests detailed in Chapter 2 are employed. An extensive description of the proposed procedure and the obtained model parameters is given. Moreover, Chapter 5 tackles the topic of REBs prognostics. In particular, it details two approaches to generate PBM models based on a peculiar degradation-related parameter named equivalent damaged volume. The full algorithm that allows to extract this parameter from real bearing degradation histories is described. For this purpose, data obtained from the

run-to-failure experiments introduced in Chapter 2 are utilized to demonstrate the capabilities of the proposed models. Discussion about the employability of the proposed methodology and its potential use in REB prognostics are thoroughly detailed. Eventually, Chapter 6 summarizes the detailed work and provides some final remarks, also including possible further developments.

As an addition, Appendix A reports a supplementary work that is not strictly related to the field of bearing modelling. However, it is included in this thesis as it is a contribution that was presented by the author at the ISMA 2020 conference in Leuven, Belgium [5]. Moreover, it is somewhat related to the numerical models described in the previous chapters, as it describes the generation process of the digital twin of a gearbox employed on agricultural equipment through the combination of LP and FE models.

2

DESCRIPTION OF THE BEARING TEST BENCH AT THE ENGINEERING DEPARTMENT OF THE UNIVERSITY OF FERRARA

2.1 INTRODUCTION

Experimental validation is the typical approach employed to determine the effectiveness of analytical and numerical models. In an ideal situation, data should be acquired on the effective system where the modelled component is mounted. This methodology would allow to test the part under exam in a condition where the effective load and constraints are applied. However, this operation is not always possible, especially in an industrial scenario, mainly due to the inevitable machine downtime that would follow. A dedicated experimental campaign, in fact, would require to stop the machinery in order to mount the sensors and proceed with the testing. As a consequence, test benches that replicate the effective operative conditions are usually employed both in the academic and industrial field.

Within this context, vibratory signals of REBs have been included by researchers in a number of datasets acquired on different test benches. Among them, the data collected on the test rig of the Intelligent Maintenance Systems (IMS) of the University of Cincinnati (Ohio) is one of the most employed in the literature [30, 200]. The test bench consisted in four double row spherical roller bearings installed on a shaft, which was connected to an AC motor by rub belts. The rotation speed was kept constant at 2000 RPM. A radial load equal to

27.7 kN was applied on the shaft by means of a spring mechanism. Vibration signals of all bearing were captured through an accelerometers mounted on each casing. This experimental setup was employed to perform three run-to-failure tests. For each one, 1 second of signal was acquired every 20 minutes with a sampling rate equal to 20 kHz. Another famous dataset is related to the PRONOSTIA test rig owned by the FEMTO-ST Institute of Besançon, France [201]. This test bench has been employed to perform accelerated run-to-failure experiments on a deep-groove ball bearing. In their test rig, the motor transmitted the rotary motion through a gearbox which delivered the motion to a secondary shaft. The load was applied via a pneumatic jack. The dataset consists of a total of seventeen sets of data for three different combinations of shaft speed and applied load values, i.e. 1800 RPM and 4000 N, 1650 RPM and 4200 N and 1500 RPM and 5000 N. Vibration of the test bearing was monitored through two accelerometers mounted perpendicularly to each other on the outer race. Signals were acquired at a sampling frequency of 25.6 kHz. However, only 2560 samples were recorded every 10 seconds. Moreover, the dataset of the Case Western Reserve University, also known as CWRU dataset [202], is widely employed for the validation of diagnostic models. The test bench consisted in a electric motor that drove a shaft in which a torque transducer and encoder were mounted. Torque was applied to the shaft through an electronically controlled dynamometer. Two deep-groove ball bearings were selected as test bearings. They were mounted on the fan-end and the drive-end of the motor. In these tests, defects were not naturally developed. Rather, they were artificially seeded on their inner and outer rings. The faults were generated via electrical-discharge machining (EDM) and their diameter ranged from 0.18 mm to 0.71 mm. Test were run at an approximate constant speed ranging from 1720 to 1797 RPM. No radial load was applied: therefore, the only radial load acting on the system was the static gravitational load [203]. Vibration signals were measured in several locations in the proximity and far-off from the motor bearings. Sampling frequency was set either to 12 kHz or 48 kHz. Recently, B. Wang et al. made available vibration signals collected on their bearing test rig [128, 204]. The experimental platform allowed to conduct accelerated degradation tests under different operating conditions. The system included an AC motor connected to a shaft, which was mounted on two heavy duty roller bearings. The bearing under test was a deep-groove ball bearing. A total of fifteen bearings were tested under three different operative

conditions. Radial load, which was applied on the test bearing housing, was generated through an hydraulic loading system. Shaft speed, imparted by the motor, was kept constant during the tests. The combinations of loads and speeds were the following: 2100 RPM and 12 kN, 2250 RPM and 11 kN and 2400 RPM and 10 kN. Vibration signals were acquired via two accelerometers mounted orthogonally to each other on the housing of the bearing under test. Finally, 1.28 s were recorded at intervals of 1 min with a sampling frequency of 25.6 kHz. Another dataset is provided by KAT datacenter of the German University of Paderborn [205, 206]. The test rig included a drive motor, a torque measurement shaft and a load motor [207]. This test bench was employed to test both healthy and artificially damaged grooved ball bearings. Tests were run under four different combinations of rotational speed, applied torque and radial force values. For each one, two possible values were chosen, i.e. 900 and 1500 RPM, 0.1 and 0.7 Nm and 400 and 1000 N. A total of thirty-two tests were conducted: six were on healthy bearings, twelve on artificially damaged REBs and fourteen on bearings damaged through accelerated life testing. For each test, 20 measurements of 4 seconds each were performed. Finally, a dataset for diagnosis and fault detection was also provided by the Society for Machinery Failure Prevention Technology (Illinois), also known as MFPT dataset [208]. They employed roller bearings, which were tested under different initial health conditions, i.e. healthy bearing, artificial inner and outer race faults and naturally developed defects. Load conditions ranged from 0 to 1334 N, while shaft speed was kept constant at 25 Hz [209]. This brief review showed just some of the publicly available datasets and the employed test rigs. Many more test benches and design have been proposed by researchers, but only few of them uploaded the acquired dataset on on-line repositories. Therefore, the reported survey attempted to cover some of the most famous ones.

Due to the need to validate bearing models, a dedicated bearing test bench has been set-up in the laboratory of the Engineering Department of the University of Ferrara. Its design allows to replicate real conditions that are applied on the bearing during its operational life. The test rig is tailored with the objective to perform two different types of tests, i.e. stationary tests on faulty bearings and run-to-failure experiments. The former allows to determine the vibratory signal generated by a bearing with some kind of artificial defect inserted on

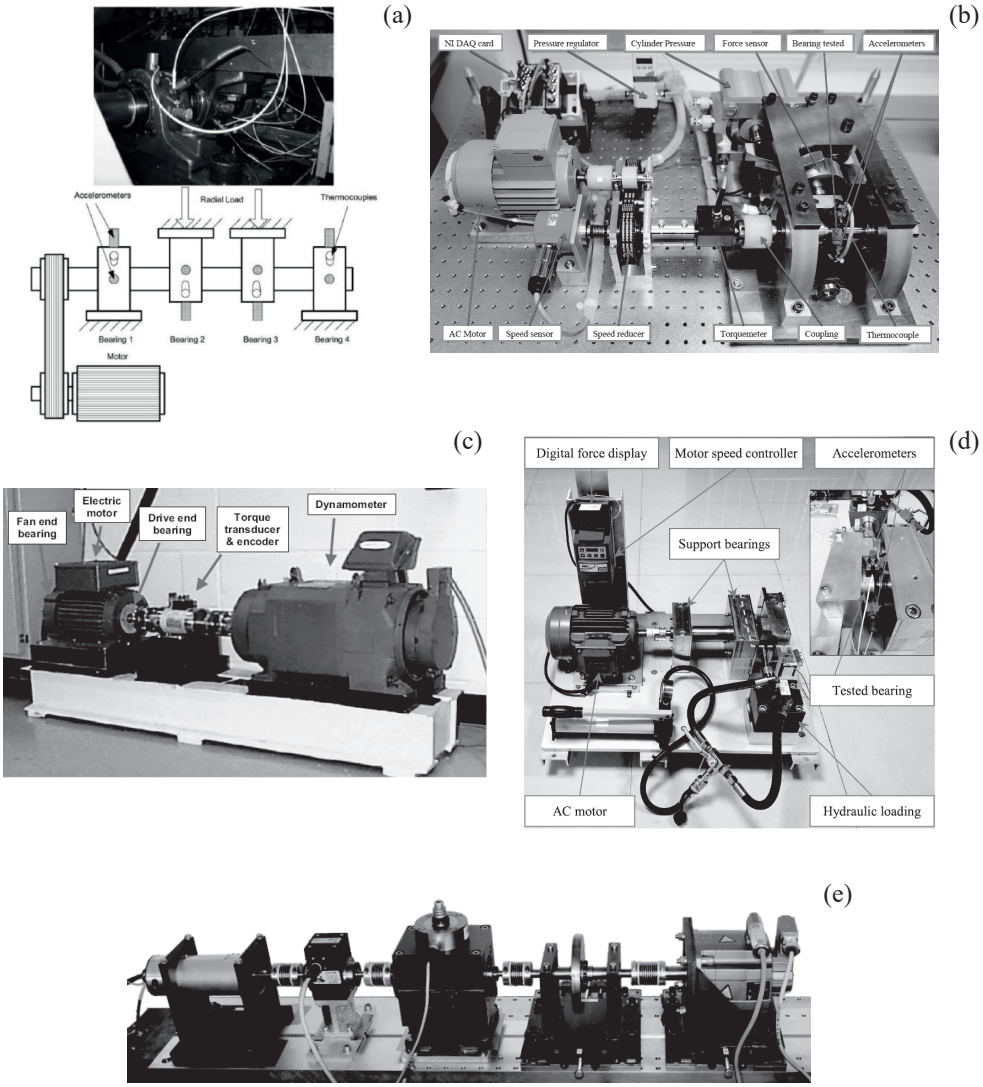

Figure 5: Photos of bearing test benches, taken from the literature: (a) IMS [30]; (b) FEMTO-ST [201]; (c) CWRU [202]; (d) B. Wang et al. [128]; (e) Paderborn University [205].

one of its components, i.e. rings, rolling elements or cage. These tests may be run for a variety of applied loads and shaft rotation frequencies. The acquired data may be employed to validate dynamic models of REBs with faults. The latter permit to acquire the vibratory signal of a REB during its entire operative

life. In this case, no artificial defects are introduced on the bearing under test. In fact, the goal is obtain natural defects that may appear during the real operational life of the component. The acquired signals may be utilized for the validation of diagnostic and prognostic models. In light of these considerations, this section describes the proposed test bench design and the two types of conducted tests.

2.2 TEST BENCH DESCRIPTION

The test rig is displayed in Figure 6 and schematized in Figure 7. The bearing under test is a self-aligning ball bearing, model *1205 ETN9*, as shown in Figure 8. Its dimensions are reported in Table 1. The REB is mounted on a shaft, which is connected through a flexible coupling to an electric motor. The rotation frequency of its output shaft is controlled by an inverter. The shaft is supported by two spherical roller bearings, model *22207 EKS*, connected to the frame. The test bearing is enclosed in a casing, and the load is applied by means of a lever system, oriented so that the resulting force is vertical and acts radially on the bearing. A load is exerted at the end of the longer arm by varying the preload of an extension spring which connects the lever to the ground. The lever system amplifies the magnitude of the applied force to the other side, which is connected to the bearing casing through a steel truss. A load cell is inserted in this component in order to measure the effective value of the resulting force, which is labelled as w in Figure 7. During the tests, the acceleration signal is acquired through a piezoelectric accelerometer model PCB 356B21 mounted on top of the casing. Moreover, during the run-to-failure experiments, two further accelerometers are mounted on the support bearings for monitoring purposes.

2.3 PERFORMED TESTS

2.3.1 *Stationary tests for bearings with artificial defects*

For the stationary tests, artificial defects were generated on a number of test bearings. In particular, nine bearings of this type were considered, and an outer race defect was seeded on each one. These defects are rectangular in shape

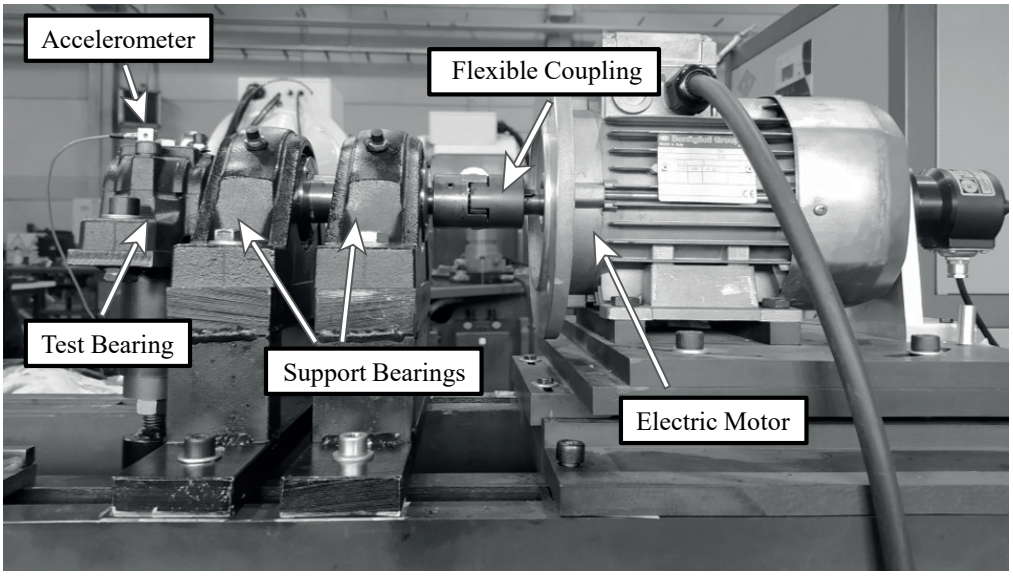

Figure 6: Bearing test bench at the Engineering Department of the University of Ferrara.

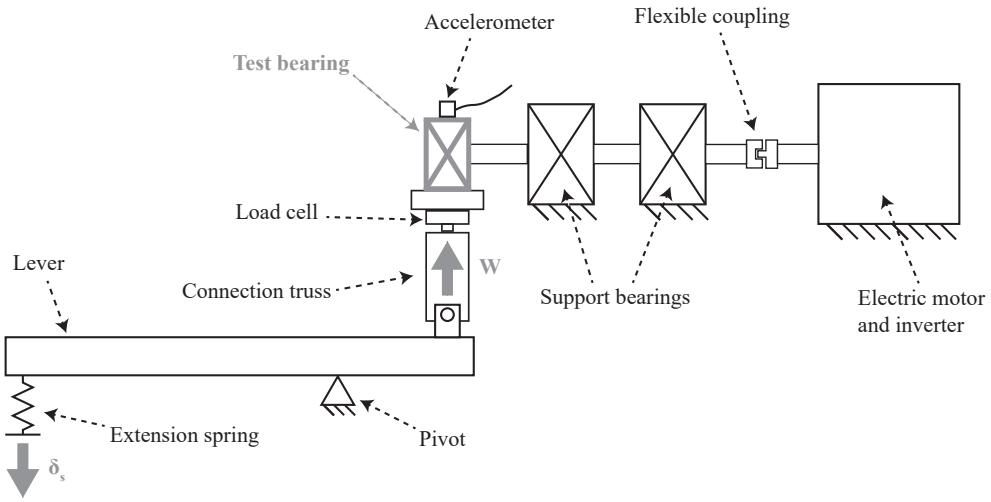

Figure 7: Schematic depiction of the bearing test bench.

and have been generated through EDM. Therefore, they are characterized by three main dimensions, i.e. their depth, circumferential length and axial width. The depth is the same for all defects and, due to the curvature of the outer

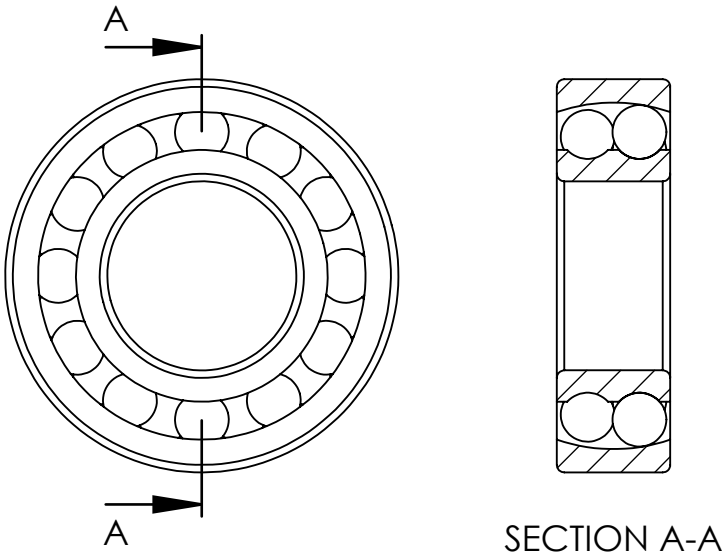

Figure 8: Geometry of the self-aligning ball bearing *1205 ETN9* employed as test bearing on the test bench.

Table 1: Design parameters of the self-aligning ball bearing model *1205 ETN9*.

Description	Symbol	Value
Inner ring groove radius	r_{gi}	3.66 mm
Inner ring race radius	r_i	16.64 mm
Outer ring radius	r_o	22.91 mm
Ball radius	r_b	3.56 mm
Contact angle	α	10.2°
Number of balls on each row	n_b	12

raceway, its minimum value is 0.06 mm towards the center of the ring while its maximum amounts at 1.6 mm at the other extremity. The width is also constant for every bearing and it is equal to 6 mm.

The chosen depth and width prevent the rolling elements to contact either the bottom or the side edges of the defect. In fact, during the modelling stage this

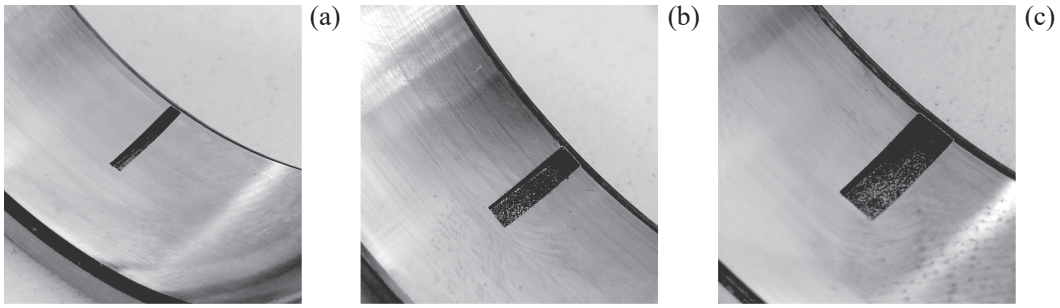

Figure 9: Artificial defects seeded on the outer rings: (a) Defect width D1; (b) Defect width D2; (c) Defect width D3.

defect depth allows to focus on the modelling of the trajectory of the balls and their free flight from the leading to the trailing edge of the defect. Such kind of artificial defects are commonly employed in the literature for model validation [24, 82, 89, 103, 105, 106, 210]. The importance of this type of defect shape is further demonstrated by L. Cui et al. [145], which employed this defect geometry to simulate the initial stages of defect propagation on a prognostic model.

On the other hand, the circumferential length is different for each bearing. Three nominal dimensions, i.e. 0.9 mm, 1.6 mm and 2.5 mm are assumed and labelled as D1, D2 and D3, respectively. The same dimension is replicated three times on three different bearings. Consequently, each bearing is labelled with a letter *a*, *b* or *c* to differentiate among them, as depicted in Figure 9. These defect extensions are similar to dimensions typically found in the literature. In fact, N. Sawalhi and R. Randall [24] used one defect with width 0.8 mm. S. Khanam et al. [82] employed several dimensions ranging from 0.35 mm to 2.02 mm. L. Cui et al. [89] included defects with widths up to 5 mm. D. S. Shah and V. Patel [103] considered defects as wide as 1.5 mm. A. Chen and T. R. Kurfess [210] considered three fault dimensions with widths 0.794 mm, 1.135 mm and 1.530 mm, respectively. M. Luo et al. [105] also proposed three different defects dimensions, i.e., 1 mm, 1.45 mm and 2 mm. S. Mufazzal et al. [106] generated artificial defects ranging from 0.5 mm to 2 mm, although characterized by a circular shape.

Table 2: Dimensions of the defects seeded on each bearing. The width values measured on each bearing have an uncertainty equal to 2 μm.

Values for each bearing			Average values		
ID	Width [mm]	Angular extent [°]	ID	Width [mm]	Angular extent [°]
D1.a	0.932	2.372			
D1.b	0.907	2.308	D1	0.928	2.362
D1.c	0.945	2.405			
D2.a	1.671	4.253			
D2.b	1.664	4.236	D2	1.664	4.236
D2.c	1.658	4.218			
D3.a	2.507	6.382			
D3.b	2.513	6.397	D3	2.501	6.366
D3.c	2.483	6.320			

Finally, the defects are located so that only the balls of one row are able to roll inside the defect. This scenario is a well consolidated strategy for testing defective self-aligning ball bearings [24, 88, 108]. Some authors tackled the problem of multiple and compound faults [80, 211, 212], but that case is not considered in the performed experimental campaign.

The effective length and width of the artificial faults were measured through a coordinate-measuring machine to determine the effective shapes and dimensions of the defects. The results are depicted in Figure 10. They are represented by considering the projection of each coordinate point on a plane parallel to the original rectangular shape, which resides on a circular surface. The effective width of each one has been measured by considering the distance between the opposite edges at the theoretical line of contact between the rolling elements and the raceway. The average values in terms of defect widths and angular extent are reported in Table 2. It is worth noting that the average values are very close to the nominal values. Furthermore, the scatter around each average value is considerably low for all defects, since all measurements deviate from the average by less than 2.3 %.

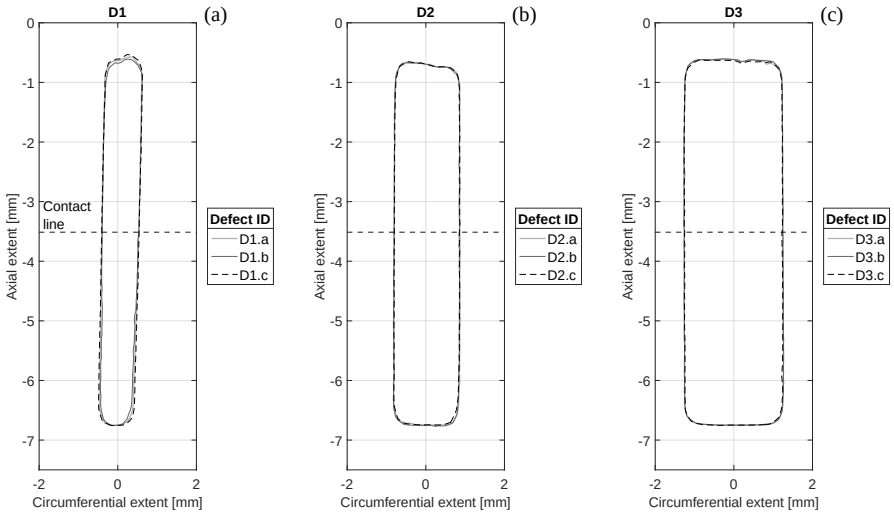

Figure 10: Measured defect dimensions. In each plot, the x-axis is aligned with the external border of the bearing. The employed measuring machine has an accuracy equal to (0.48 + l) µm, where l is the measuring distance expressed in meters. (a) Fault size D1; (b) Fault size D2; (c) Fault size D3.

The faulty bearings were tested for different combinations of applied load values and shaft rotation frequencies. In particular, two static load values equal to 1000 N and 2000 N and three rotation speed values equal to 20 Hz, 30 Hz and 40 Hz were selected, leading to six tested conditions for each bearing under test. Moreover, each defect width was replicated three times on three different bearings, thus leading to eighteen tests on the same defect dimension. Finally, three different defect widths were generated. As a results, a total of fifty-four tests were run. Each signal was acquired for 15 seconds at a sampling frequency of 51.2 kHz. The signals were then low-passed at 9.5 kHz for the subsequent analyses described in Chapter 4. The vibration signals were acquired through a LMS SCADAS Mobile M06 equipped with 56 analog channels, a 24 bit AC/DC converter and an anti-aliasing filter. The raw acceleration signals are publicly available: in fact, they have been uploaded in a online data repository hosted by Mendeley Data [4].

2.3.2 *Run-to-failure tests*

Run-to-failure tests involve the acquisition of the vibratory signal through the entire operative life of the bearing under test. In this way, it is possible to understand the effective vibration of the component in both healthy and unhealthy conditions. Moreover, run-to-failure tests induce the generation of natural defects in the bearing, which are the ones that are effectively produced during real operations. Therefore, the resulting signals also contain information about the damage progression and may consistently differ from stationary tests with artificial defects with controlled dimensions.

The main issue related to run-to-failure tests is the extensive time duration of this kind of tests. In fact, bearing life may be roughly estimated according to the well-known L_{10h} value [6]:

$$L_{10h} = \frac{a_l \pi}{1800 \omega_s} \left(\frac{Cr}{P} \right)^s \cdot 10^6 \tag{1}$$

which allows to determine the fatigue life, expressed in hours, that 90 % of the bearing population will endure. In Eq. 1, ω_s is the shaft rotation speed, C_r is the basic dynamic load rating of the bearing, P is the equivalent dynamic bearing load and s is an exponent which is equal to 3 for ball bearings and 10/9 for roller bearings. Within this context, C_r is the load that, if applied to the bearing, would lead to a rating life equal to 10^6 revolutions. For the bearing reported in Table 1, C_r = 12.2 kN. Moreover, for a radial bearing, P is equal to the applied radial load if the REB is not subjected to any axial load. Additionally, Eq. 1 presents a further term a_l, which is a coefficient that allows to take into account other life influencing factors. Its value is provided by international standards, e.g. ISO 281 [7]. This equation leads to two important considerations. Firstly, that bearing failure is a statistically distributed phenomenon. Therefore, the effective length of run-to-failure test may vary even under identical operative conditions. As a consequence, for the purpose of run-to-failure tests, the estimated L_{10h} allows to perform rough estimation on the expected test length under given conditions. Secondly, L_{10h} value depends on the applied load and the rotation speed of the shaft. In particular, it greatly depends on the magnitude of the applied load. To better underline this concept, Figure 11 shows the values of L_{10h}, expressed in days,

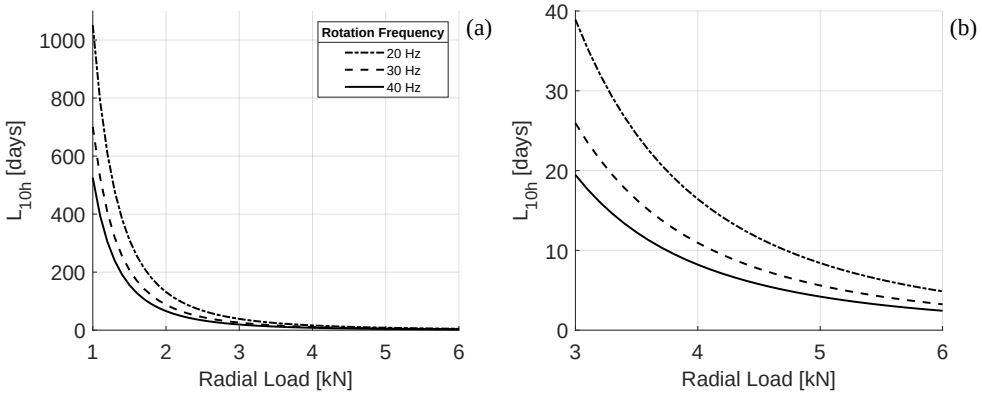

Figure 11: Expected life of the bearing *1205 ETN9* under different combinations of radial load and shaft speed values: (a) Load range 1-6 kN; (b) Load range 3-6 kN.

for the bearing under test as computed for different combinations of loads and speeds. Specifically, Figure 11.a reports the bearing rating life between 1 and 6.6 kN. The plot demonstrates that bearing life decreases exponentially with increasing load, as highlighted by Eq. 1. In particular, the expected life ranges from more than 500 days for 1 kN of load to less than 5 days for 6 kN. In the vast majority of applications, bearings are chosen to work in the left part of the graph, so to have the highest possible expected life and minimize their required maintenance.

However, working at low loads and shaft speeds is not suitable for testing on a bearing test rig. In particular, low loads would lead to excessive test times, in the order of hundreds of days of continuous operation. In order to overcome this issue, accelerated tests are performed. In this kind tests, a large load is applied in order to induce a fast degradation of the component. This is a standard kind of testing procedure employed by researchers to generate faulty bearing datasets, as reported in Section 2.1. This aspect is emphasized by Figure 11.b, which depicts L_{10h} values between 3 and 6 kN. Within this range, the bearing life ranges approximately from 20 to 2.5 days at a shaft speed equal to 40 Hz. Therefore, tests performed inside this range of loads allow to attain an acceptable testing time and permit to carry out multiple tests in a relatively short amount of time. Finally, it is worth underlining that,

Table 3: Operative conditions, expected life and test length for the six tested bearings.

ID	Load [kN]	Shaft speed [Hz]	L_{10h} [days]	Test time [days]
E1	4	40	8.2	17
E2	4	40	8.2	6.9
E3	4	40	8.2	8.3
E4	3	40	19.5	20.9
E5	4.7	40	5.1	6
E6	5	40	4.2	1.8

for the bearing under exam, most of the values in this load range are higher than the basic static load rating C_0 of the bearing, which is equal to 3.3 kN. C_0 is defined as the load that generates a permanent deformation of raceways and rolling elements equal to approximately 1/10000 of the rolling element diameter. This is another common characteristic of these kind of tests [201, 204, 213].

In light of these considerations, six run-to-failure tests have been performed, as reported in Table 3. All tests have been carried out at the same shaft speed of 40 Hz. Three tests, namely E1, E2 and E3, were performed under the same load equal to 4 kN. In the other three cases, i.e. E4, E5 and E6, applied loads were equal to 3 kN, 4.7 kN and 5 kN, respectively. Table 3 also reports the expected life and the total test time. Concerning the latter, tests were stopped when the maximum peak in the acceleration signal reached 20 g. This is a common test end criterion previously employed by several researchers for run-to-failure tests [201, 204, 213]. At the end of the tests, all bearings were characterized by an extended defect on the outer raceway, as shown in Figure 12.

During the tests, vibration signals were continuously acquired through a NI cRio controller, equipped with a 9234 module provided with 4 analog channels. Sampling frequency was set to 25.6 kHz, and 5 seconds of signal were acquired every 5 minutes. As a consequence, 12 signals were stored for each hour of test. Differently from stationary tests, two further accelerometers were also mounted on the support bearings in order to monitor their health state. In

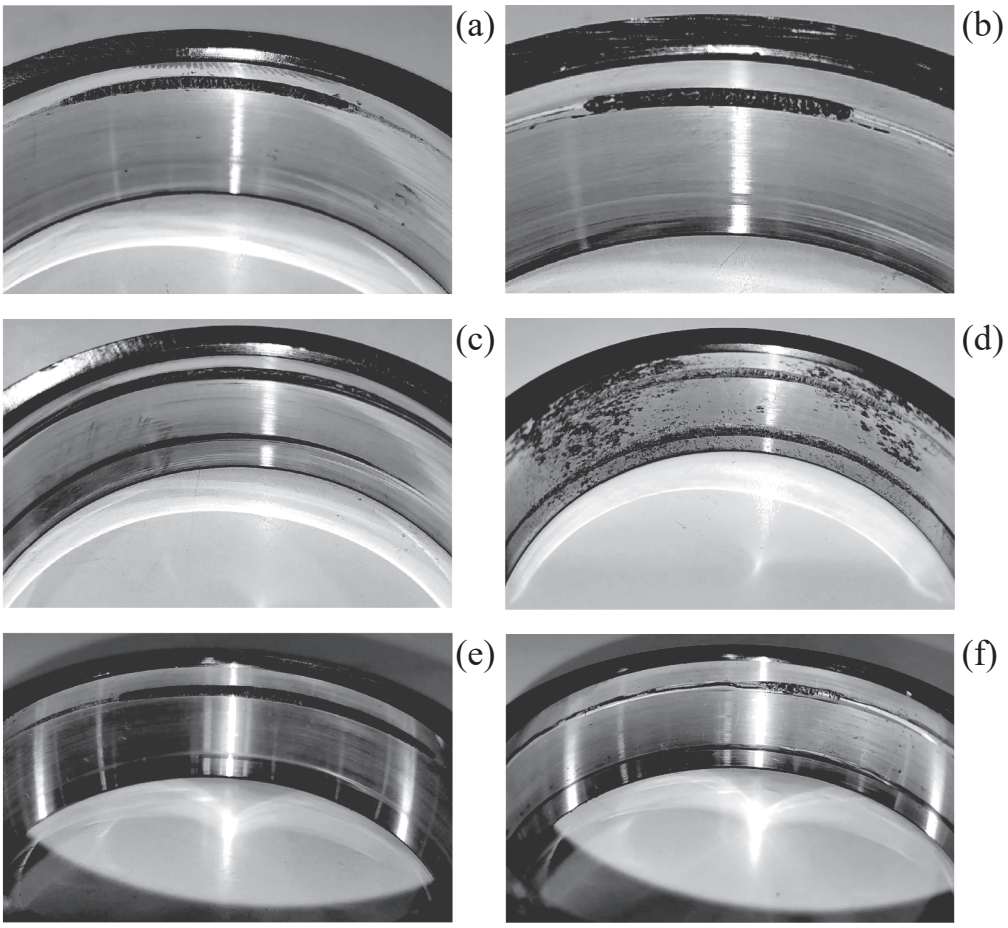

Figure 12: Photos of the bearing outer rings at the end of the run-to-failure tests: (a) E1 (4 kN); (b) E2 (4 kN); (c) E3 (4 kN); (d) E4 (3 kN); (e) E5 (4.7 kN); (f) E6 (5 kN).

fact, although they are characterized by a way higher C_0 compared to the test bearing, i.e. 86.5 kN, it is crucial to control their behaviour during these long tests in order to avoid unexpected failures. Therefore, the signals of these two accelerometers were also acquired for monitoring purposes.

3

NON-LINEAR RADIAL BEARING STIFFNESS ESTIMATION BASED ON LOAD DEPENDENT MESHING

3.1 INTRODUCTION

As detailed by the literature review in Section 1.2.1, several methods have been developed to provide numerical estimates of the bearing stiffness. However, the investigation showed the lack of uniformity in the employed approaches and the possible improvements that could be carried out, especially in the context of FE simulations. Therefore, this chapter discusses a numerical FE procedure to determine the radial stiffness of REBs by means of load-dependent meshes. The proposed methodology involves the generation of a dedicated mesh for each load condition, so that the element size is determined based on the estimated contact area dimensions obtained by means of the Hertz contact theory. Through this approach, a fast and efficient method for generating REBs meshes employable for FE simulations is obtained. In fact, the detailed technique allows to estimate a reasonable dimension for the elements in the contact area, in order to capture the contact phenomenon while balancing the computational time and the accuracy of the results. The influence of element types and mesh size on the solution is thoroughly discussed. Hypotheses that allow to reduce the computational domain are detailed, i.e. the employment of symmetry planes, the removal of unloaded rollers and the load application method. Stiffness influencing factors, such as cage modelling, load direction and clearance are numerically examined. The robustness of the method is tested on two representative bearings, i.e. a cylindrical roller bearing and a

deep-groove ball bearing. In this way, the efficiency of the proposed approach in tackling different types of contacts is demonstrated without losing generality. It is worth noting, in fact, that the proposed methodology is still applicable for other bearings types. However, attention must be paid when analysing different geometries, as there could be only one symmetry plane available and not two as in the reference bearings, e.g. in tapered roller bearings and self-aligning double-row ball bearings. On the other hand, concerning element type, mesh size, unloaded roller removal and load application method, the proposed procedure may be adapted to all bearing types. Finally, it is worth mentioning that the proposed methodology also aids the bearing analyst by providing guidelines for the selection of several aspects involved in the FE simulation, which include choice of the elements, definition of the mesh size, which symmetries to exploit, how to define the load and which location to prefer to evaluate the displacement.

The chapter is subdivided as follows. Firstly, Section 3.2 describes the reference bearing geometries and hypotheses exploited in the numerical analyses. Information on the solver and the contact algorithm are also provided. Moreover, the application of the Hertz contact theory in REBs modelling is discussed. Section 3.3 deals with the procedure employed to determine the mesh size in the contact area and compares the results obtained with different element types for two reference bearings, i.e. in case of ball-races and roller-races contact. These indications are exploited to compute the radial stiffness in Section 3.4. After an initial discussion on the analytical methods for bearing stiffness estimation, the procedure to generate the load-dependent meshes is described. Additional methods to reduce the computational effort are also outlined. Furthermore, Section 3.5 examines the influence of load direction, cage modelling and clearance on the estimated radial stiffness. Finally, 3.6 sums up the chapter by providing some concluding remarks.

3.2 PROBLEM GENERALITIES

This section opens with the description the bearing geometries taken as reference for this work. Then, an overview on the Hertz contact theory and its application within the context of REBs modelling is provided. Subsequently,

Table 4: Design parameters of the cylindrical roller bearing model *NU 202 ECP* and the deep-groove ball bearing model *6210*.

Description	Symbol	Roller bearing *NU 202 ECP*	Ball bearing *6210*
Number of rolling elements	n_b	11	10
Effective roller length	l_{eff}	5.6 mm	-
Rolling element radius	r_c, r_b	2.75 mm	6.35 mm
Inner ring race radius	r_i	9.65 mm	28.7 mm
Outer ring race radius	r_o	15.15 mm	41.35 mm
Inner and outer ring groove radii	r_g	-	6.55 mm
Ring width	B	11 mm	20 mm

focus is placed on aspects involving the numerical simulations. At first, information regarding solver choice and contact algorithm are reported. Next, the section closes by detailing the modelling hypotheses at the basis of the investigation.

3.2.1 *Reference bearings geometry*

In this study, two different rolling element bearings are considered. Figure 13 shows the 3D geometry of both mechanical components. The associated geometrical data, which refer to nominal dimensions, are reported in Table 4. The first bearing, which is shown in Figure 13.a, is a cylindrical roller bearing, model *NU 202 ECP*. Rollers have a straight profile and are 6 mm wide. Since edges are rounded with a 0.2 mm radius, their effective length reduces to 5.6 mm, which is the length of the ideal contact line between roller and races. A 0.1 mm axial clearance between rollers and flanges is also considered. The other bearing is a deep-groove ball bearing, model *6210*, which is depicted in Figure 13.b. It is assumed that all components of both bearings are made of steel, with Young's modulus $E = 210$ GPa and Poisson's ratio $\nu = 0.3$. These properties are considered as constant values in all simulations.

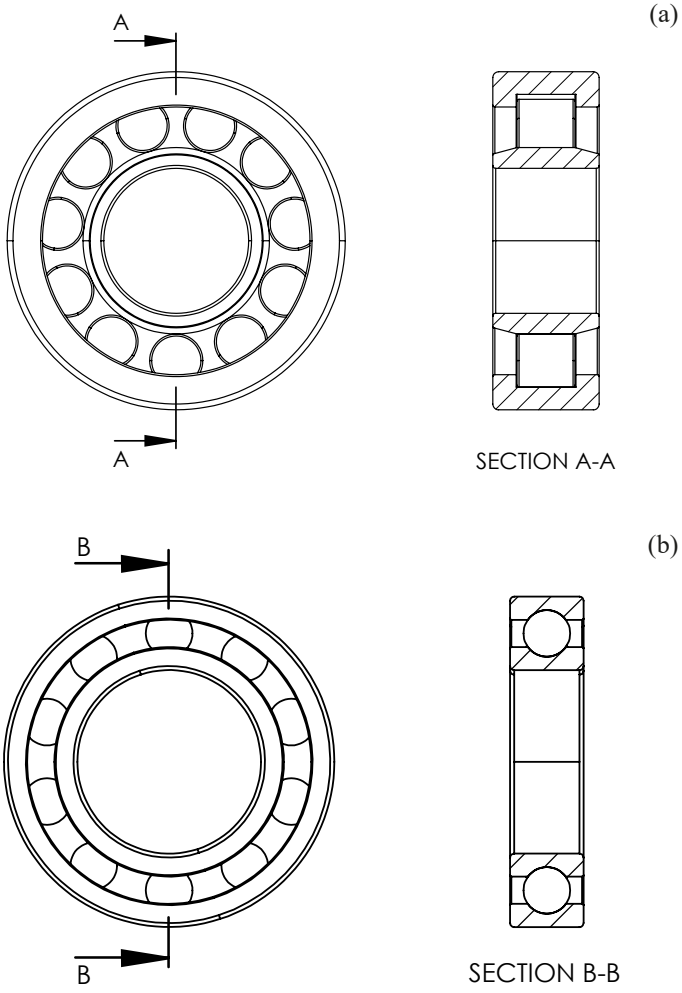

Figure 13: Geometry of the roller bearing taken as reference for the investigation: (a) cylindrical roller bearing *NU 202 ECP*; (b) deep-groove ball bearing model *6210*.

3.2.2 *Application of the Hertz theory for contacts in REBs*

According to the Hertz contact theory, the size of the contact area between two bodies in contact depends on their curvature radii, on their materials and on the magnitude of the applied load [19].

In the case of a cylindrical roller bearing with straight roller profile, the surfaces in contact are cylinders. The shape of the contact area is a rectangle with length equal to l_{eff} and semi-width a, as depicted in Figure 14.a. The latter may be analytically evaluated at each load by introducing an equivalent radius of contact, namely $r_{eq,c}$, and an equivalent Young's modulus, namely E_{eq}. These quantities allow to consider the different geometry and materials of the two bodies. The equivalent radius is evaluated as:

$$r_{eq,c} = \left(\frac{1}{r_I} + \frac{1}{r_{II}} \right)^{-1} \tag{2}$$

where r_I and r_{II} are the radii of curvature of the contacting surfaces. They may assume positive or negative values if the surfaces are convex or concave, respectively. On the other hand, the equivalent modulus is:

$$E_{eq} = \left(\frac{1-\nu_I^2}{E_I} + \frac{1-\nu_{II}^2}{E_{II}} \right)^{-1} \tag{3}$$

where E_I and ν_I are the material properties of the first body and E_{II} and ν_{II} are the material properties of the second body. By knowing r_{eq} and E_{eq}, it is then possible to calculate the semi-width a for an applied load F as:

$$a = \sqrt{\frac{4 r_{eq,c} F}{\pi l_{eff} E_{eq}}} \tag{4}$$

Therefore, a is a function $r_{eq,c}$, E_{eq} and F. On the contrary, the width of the rectangle is equal at all load values to the effective length of contact of the rollers, i.e. l_{eff}. Since the two rings in contact with a roller have different radii, two different contact areas are defined for each rolling element. They are named a_i and a_o for the inner and outer ring contacts, respectively. Besides, for a roller bearing, a_o is always greater than a_i at all load values. Within this framework, however, it is important to observe that the Hertz theory does not account for the edge effects caused by the finite length of the components [62]. Despite this issue, this theory is employed anyway as it provides straightforward formulae that may be used for a rough estimation of the contact area. Finally, it must be considered that a rectangular contact area is produced only if the roller profile is straight. For fully or partially crowned profiles, a different

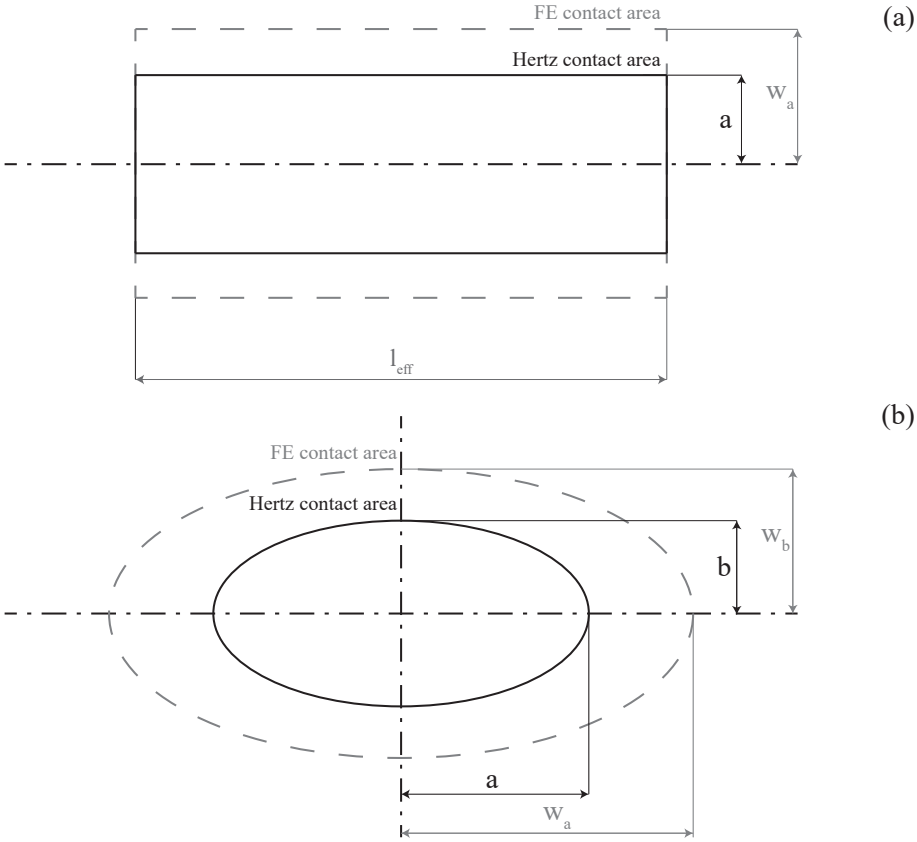

Figure 14: Shape of the contact areas: (a) roller-race contact; (b) ball-race contact.

formulation should be employed to estimate the contact area. However, the meshing procedure and the described methodology remain unaltered.

In the case of a deep-groove ball bearing, the surfaces in contact have different principal relative radii of curvature in orthogonal planes. Consequently, the shape of the contact area becomes an ellipse defined by semi-major axis a and semi-minor axis b, as picture in as depicted in Figure 14.b. Even these quantities are functions of geometry, materials and applied load. However, in this case the equivalent radius, namely $r_{eq,b}$, depends on the radii on the two orthogonal directions x and y, i.e.:

$$r_{eq,b} = \left(\frac{1}{r_x} + \frac{1}{r_y} \right)^{-1} \tag{5}$$

where

$$r_x = \left(\frac{1}{r_{xI}} - \frac{1}{r_{xII}} \right)^{-1} \tag{6}$$

and

$$r_y = \left(\frac{1}{r_{yI}} - \frac{1}{r_{yII}} \right)^{-1} \tag{7}$$

Concerning Eqs. 6 and 7, it is important to note that directions x and y must be assigned so that $r_y > r_x$. Moreover, to estimate a and b for this type of contact, additional quantities must be introduced, namely the simplified elliptical integrals $\overline{\epsilon}$ and $\overline{\xi}$:

$$\overline{\epsilon} = 1.0003 + \frac{0.5968 r_x}{r_y} \tag{8}$$

$$\overline{\xi} = 1.5277 + 0.6023 \ln \left(\frac{r_y}{r_x} \right) \tag{9}$$

and the ellipticity parameter $\overline{\kappa}$:

$$\overline{\kappa} = 1.0339 \left(\frac{r_y}{r_x} \right)^{0.636} \tag{10}$$

Finally, a and b are calculated as:

$$a = \left(\frac{6 \overline{\kappa}^2 \overline{\epsilon} r_{eq} F}{\pi E_{eq}} \right)^{1/3} \tag{11}$$

$$b = \left(\frac{6 \overline{\epsilon} r_{eq} F}{\pi \overline{\kappa} E_{eq}} \right)^{1/3} \tag{12}$$

Similarly to roller contact, two different contact areas are defined for each ball. The semi-axes are named a_i and b_i for the inner ring and a_o and b_o for the outer ring, respectively. Even for a deep-groove ball bearing, the size of the contact area on the outer ring is greater than its inner ring counterpart at all loads. In this case, the effective area of contact may deviate from the Hertz estimate because of the conformity of the races with the balls. However, as in the case of the roller-race contact, the simplicity of the formulae allow for an efficient estimation of the contact parameters.

According to Eqs. 4, 11 and 12, the size of the contact area decreases as the load reduces. As a consequence, in a FE analysis involving contacts, the local mesh in the neighbourhood of the contact areas must be sufficiently small in order to accurately take into account the contact phenomenon. In the meshing procedure described in this chapter, the size of the contact area for each contact pair in the FE model is defined at each load on the basis of analytical estimates. The outcome is an analytical relationship between the dimension of the contact area, i.e. a for the roller-race contact and a and b for the ball-races contact, and the size of the contact area in the FE model, namely w_a and w_b. The size of the contact area in the FE model, in fact, may be larger than the Hertz estimate, as depicted in Figure 14. This observation arises from the fact that employing a larger area may permit to use greater elements, therefore allowing to reduce the number of mesh nodes. Furthermore, the Hertz theory gives only a rough estimate of the effective contact area, as the edge effects and the high conformity of races may lead to deviation in the effective dimensions. Finally, the local mesh size, namely l, will also be defined as proportional to either a or b. To sum up, the main goal of the inquiry is to find coefficients that relate FE model properties and analytical estimates so that $w_a \propto a$ or b, $w_b \propto a$ or b and $l \propto a$ or b.

3.2.3 *Solver choice and contact algorithm*

All simulations described in this chapter are carried out by using Simcenter 3D as pre/post processor and Simcenter Nastran as solver [214]. Nastran is chosen as solver as it is one of the most employed software in the field of FE analysis [215]. Moreover, it is the main solver associated with the pre/post processor exploited in this work, i.e. Simcenter 3D. In fact, Nastran does not have any functionality that allows to graphically observe the model in both the pre and post phases, as the the input and output files are only given as text files. To read these files, additional software is commonly utilised, such as Simcenter 3D, MSC Marc [216] and Autodesk Inventor [217]. Well-known alternatives to Nastran include Ansys [218] and Abaqus [219]. The FE techniques described in this thesis may be also implemented in solution processes involving these solvers, as they allow to consider the contact phenomenon. However, the discussion is centred only on results obtained through the Nastran solver,

without any loss of generality. In particular, Nastran static solution SOL 101 is used to solve the models since it allows to consider the non-linear behaviour given by the contact algorithm and concurrently reduce the computational burden by assuming a linear elastic material and small displacements.

Concerning the contact algorithm, two methods are available, namely segment-to-segment method and node-to-segment method. Previous work by N. El-Abbasi and K.J. Bathe [220] indicated that, while both methods provided stable results, the latter did not pass the patch test [221], leading to discretization errors that did not decrease with mesh refinement. In addition, proof of the successful employment of the segment-to-segment method in bearing simulation may be found in [17] and [222]. These results support the choice of the first approach for REBs simulation. Besides, bilinear Coulomb friction model is chosen to take into account friction-related contributions. Friction coefficient is set equal to 0.05, which indicatively corresponds to a greased contact condition [223]. This value is chosen for the sake of simplicity to represent a common value encountered when grease is employed for bearing lubrication. Approaches based on the bilinear Coulomb friction were utilised in previous studies, as reported by N. Demirhan and B. Kanber [57], R. Lostado-Lorza et al. [17] and R. Fernandez Martinez et al. [66]. In addition, according to Y. Guo and R.G. Parker [10], lubrication has a mild effect on bearing stiffness especially in case of thin fluid film and moderate rolling speed. As a consequence, the effect of friction is not considered to be significant in the analysis, and a simple approach as the one exploited in this work may used to determine the bearing stiffness characteristics. However, in case of particular lubricating conditions, more accurate methods may be utilized to determine the coefficient of friction and the resultant friction force, e.g. by exploiting the methods reported in [224].

3.2.4 *Modelling hypotheses*

The proposed investigation exclusively takes into account the major phenomena concurring to the determination of bearing radial stiffness. Based on this approach, the hypotheses at the basis of the present study are hereinafter detailed. The first aspect concerns possible plasticization of the material in the

neighbourhood of the contact areas. Bearing races, in fact, permanently deform at sufficiently high loads, affecting the wear and degradation process of the races [65, 223]. Therefore, only deformations in the elastic range are considered by employing a linear elastic material with undefined yield stress. Concerning boundary conditions, the external outer race of the bearing is assumed to be connected to a rigid frame. As a matter of fact, the adopted hypothesis represents a common scenario where the frame is sufficiently rigid to not interfere with the bearing properties and the applied loads stay within the elastic deformation range. For the sake of completeness, it has to be clarified that actual applications might involve either the installation on compliant frames or the application of an excessive load that may lead to a large deformation of the outer race [6, 17, 225]. In such cases, shaft misalignment may appear [226, 227], causing a significant moment load on the bearing and the loss of one symmetry plane from the system. These peculiar conditions requires dedicated analyses involving different solution schemes and therefore they do not fall within the purpose of the present study. Furthermore, thermal effects and preloading are not considered. As a consequence, simulations are run considering constant properties at room temperature and by applying a simple radial load only. In fact, as described in Section 1.2.1, temperature may affect stress distribution and stiffness [15], while preload is commonly applied in order to modify bearing stiffness characteristics [16]. Finally, it should be noted that static methods such as the one described in this work are employable at low and moderate speed values only [10, 14], since inertia effects cannot be neglected at higher velocities [13].

3.3 MESH SIZE ASSESSMENT AND ELEMENT CHOICE

This section deals with the analysis of the contact stiffness through the FE method. The inquiry is focused on typical types of contact taking place in REBs. First, the procedure employed to address mesh performance is outlined. Then, it is applied to two different cases, i.e. roller-race and ball-race contacts.

3.3.1 *Description of the procedure employed for mesh performance assessment*

The performance of various elements types and mesh dimensions is evaluated in order to determine their most convenient combination to tackle the contact problem in rolling-element bearings. The objective of the investigation is to determine an analytical relationship between the size of the contact area and the local mesh size in a roller-race and in a ball-race contact. In order to perform such a task, contact stiffness is evaluated for different mesh sizes to determine a reasonable dimension that permits to obtain a good estimation of the radial bearing stiffness while maintaining a relatively low computational time. The procedure detailed in the following is applied with a few differences to both bearing types considered in this work.

Firstly, a representative subsystem is extracted from each bearing CAD model. The geometry is first reduced to an angular sector centred on one of the rolling elements, characterized by an angular span equal to the spacing between two consecutive ones. Then, the geometry is further cut by taking advantage of the two available symmetry planes. As a consequence, the considered geometry reduces to a quarter of the rolling element in contact with a portion of both races. Example of the reduced geometries may be observed in Figure 15. In particular, Figure 15.a and Figure 15.b show the representative subsystem of the roller bearing meshed with tetrahedral and hexahedral elements, respectively. Besides, Figure 15.c illustrates the reduced portion of the ball bearing meshed with tetrahedral elements.

Based on the representative subsystems, a number of grids with decreasing mesh size is generated for each element type. The inquiry allows to determine the contact stiffness values attainable with the proposed FE analysis. The investigation regards tetrahedral and hexahedral elements in both their linear and parabolic formulations. In Nastran, tetrahedral elements are named CTETRA4 and CTETRA10. The former are 4-nodes linear elements, while the latter are 10-nodes parabolic elements. Similarly, hexahedral elements are designated as CHEXA8 and CHEXA20 in their linear and parabolic formulations, respectively. Both element types are employed for the roller bearing, while only tetrahedral elements are considered for the ball bearing. For each element type, five meshes are generated. Each one is characterized by different values

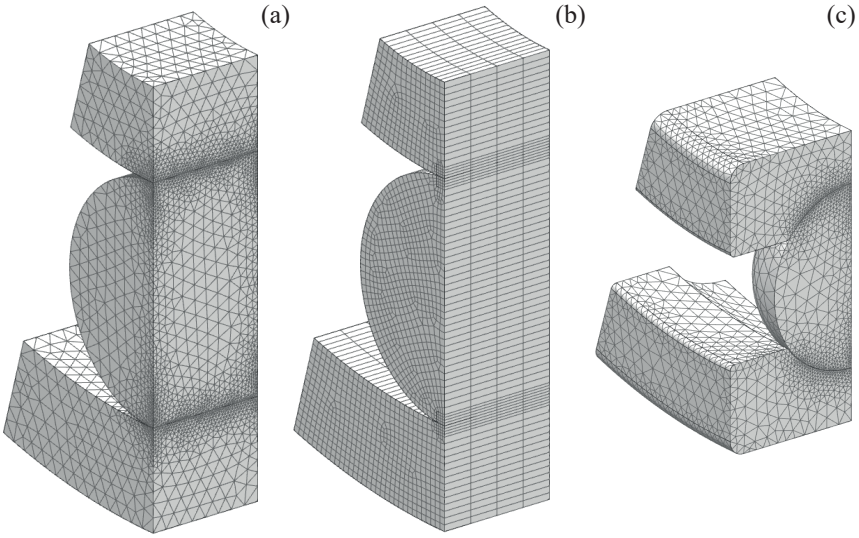

Figure 15: Sub-models and meshes employed for convergence analysis: (a) roller - race contact, tetrahedral elements; (b) roller-race contact, hexahedral elements; (c) ball-race contact, tetrahedral elements.

of local and global mesh size, namely l and g, respectively. The value of l is related to g through a coefficient of proportionality η so that $l = \eta g$. Local mesh size is assigned only on the contact areas generated on rolling elements and races.

Each mesh is tested under the maximum load on a roller produced by two radial loads F_{r1} and F_{r2}. These are chosen as "low" and "high" load values that might act on the bearing. While the magnitude of the former is based on practical assumptions, the latter may be assumed on the basis of the static load rating of the bearing. In absence of clearance, the load acting on the most loaded rolling element may be calculated with the following formula [6]:

$$Q = \frac{F_r}{n_b J_r} \tag{13}$$

where Q is the maximum load on the roller, F_r is the applied load, n_b is the number of rolling elements and J_r is a parameter whose value is equal to 0.2453 for a roller bearing and to 0.2288 for a ball bearing. In this preliminary

analysis, a conservative value is assumed for w, w_a and w_b by considering a contact area larger than the one estimated for Q_2, i.e. the maximum load produced by F_{r2}.

The generated grids allow to compute the contact stiffness due to the applied load Q as

$$k_c = \frac{Q}{\delta^n} \tag{14}$$

where δ is the average radial displacement of the nodes on the inner surface of the inner ring which lie on the plane of maximum load. Besides, exponent n in Eq. 14 is equal to 10/9 for roller bearings and 1.5 for ball bearings. This procedure allows to compare the contact stiffness obtained for meshes characterized by different number of nodes and to assess the values which may be attained with the considered geometry and element types.

Finally, the investigation aims at finding a relationship between the local mesh dimensions and the size of the contact area determined by the Hertz theory. For this purpose, various combinations of contact area and local mesh size are tested. The latter is considered to be proportional to the dimensions a and b evaluated by means of the Hertz theory at loads Q_1 and Q_2. This strategy permits to determine the most convenient combination of element type, mesh size and contact area dimensions in terms of accuracy of the results and computational time. For the roller bearing, the goal is to find a relationship between w, l and a. For the ball bearing, w_a, w_b and l are related to b. Semi-minor axis b is chosen in place of a since it has the lowest value between the two. It is also worth underlining that the contact area dimensions differ for the contacts between the rolling elements and the two rings, as described in Section 3.2.2. The quantities referring to the contact area between rolling elements and inner race are denoted with subscript i, while for the outer race subscript o is employed.

3.3.2 *Roller-races contact*

To study the roller-races contact, the geometry of the reference roller bearing is reduced to an angular sector with extension 16.36°, i.e. $360°/n_b$, where

n_b is the number of rolling elements reported in Table 4. The representative subsystem is depicted in Figure 15.a and Figure 15.b. The flanges of the bearing are also removed to consider a case closer to the theoretical one. Then, a set of five meshes with decreasing mesh size is generated for each element type. This means that a total of 20 meshes are generated by employing CTETRA4, CTETRA10, CHEXA8 and CHEXA20 elements, respectively. Figure 15.a shows one of the tetrahedral meshes, while Figure 15.b depicts one of the hexahedral meshes.

Two radial loads F_{r1} = 100 N and F_{r2} = 10000 N are chosen for the analysis. Each radial load generates a different value of the force acting on the most loaded roller. According to Eq. 13, if the radial clearance is equal to zero then their magnitudes are Q_1 = 37 N and Q_2 = 3706 N, respectively. Therefore, Q_1 and Q_2 are the loads employed for mesh performance assessment of the roller-race contact. The load is applied by providing a sinusoidal load distribution on the inner ring, which resembles the force exchanged with the shaft. The dimensions of the contact area for these two force values are estimated via the Hertz contact theory. These are a_i = 0.0127 mm and a_o = 0.0159 mm at load Q_1, and a_i = 0.127 mm and a_o = 0.159 mm at load Q_2. Because of the geometry, a_o is always greater than a_i. As a result, a conservative value $w_i = w_o$ = 0.25 mm is employed on both rings to generate the rectangular contact area in the preliminary analysis of the meshes. Table 5 reports the local mesh size, the global mesh size and the number of nodes for the meshes generated with tetrahedral elements. Similarly, Table 6 details the characteristics of hexahedral meshes. The global mesh size is also defined by assuming η = 0.14 for tetrahedral elements and η = 0.50 for hexahedral elements in order to provide a smooth transition between different element dimensions. Additionally, only for the hexahedral mesh, the number of elements along shaft axis direction, namely n_{subs}, is set equal to 4, as it may be observed in Figure 15.b. Such a value is chosen as a compromise between the number of nodes along the contact line and the local mesh size. It should be noted, in fact, that if obtaining a certain number of grid points is required, as n_{subs} increases l decreases accordingly. In addition, for two grids with same value of parameter l, computational time will depend on n_{subs} as it affects the total number of nodes and contacting elements. Thus, it is crucial to find a compromise between accuracy and computational time.

Table 5: Statistics of tetrahedral meshes employed to assess element performance in roller-race contact.

	CTETRA4			CTETRA10		
Mesh	l [mm]	g [mm]	Number of nodes	l [mm]	g [mm]	Number of nodes
1	0.05	0.357	9024	0.15	1.071	12826
2	0.025	0.179	32406	0.1	0.714	22306
3	0.02	0.143	49428	0.05	0.357	58767
4	0.01	0.071	213309	0.025	0.179	220487
5	0.0075	0.054	393876	0.02	0.143	341098

Table 6: Statistics of hexahedral meshes employed to assess element performance in roller-race contact.

	CHEXA8			CHEXA20		
Mesh	l [mm]	g [mm]	Number of nodes	l [mm]	g [mm]	Number of nodes
1	0.05	0.100	16405	0.1	0.200	16783
2	0.025	0.050	61975	0.05	0.100	61509
3	0.015	0.030	167465	0.035	0.070	122120
4	0.0125	0.025	242310	0.025	0.050	233825
5	0.01	0.020	373235	0.02	0.040	364085

Figure 16 shows the estimated stiffness values and the computational time obtained with linear and parabolic tetrahedral elements. The plots also show the analytical value of the contact stiffness k_c, which for the roller bearing under exam is a constant value equal to 172549 $N/mm^{10/9}$. According to [6], in fact, the contact stiffness of a single steel roller in contact with a steel ring, namely $k_{c,r}$, is equal to:

$$k_{c,r} = 8.06 \cdot 10^4 l_{eff}^{8/9} \tag{15}$$

As a consequence, the total contact stiffness taking into account both contacting sides is:

$$k_c = \left[\frac{2}{(k_{c,rl})^{(1/n)}}\right]^{-n} \tag{16}$$

since the values for inner and outer ring contacts have the same magnitude. Figure 16.a reveals that, at load Q_1, parabolic elements shows a more stable trend compared to their linear counterpart. The computational time, depicted in Figure 16.c, is similar for both element types. Moreover, Figure 16.b and Figure 16.d demonstrate that they also have comparable performances both in terms of computed stiffness and simulation time. Parabolic elements, however, reach a constant trend in the estimated stiffness for a lower number of nodes.

In light of this, CTETRA10 elements are employed to determine a relationship between the local mesh size and the size of the contact area determined by the Hertz theory. For this purpose, five different combinations of w and l have been tested. For each mesh, these parameters are set as proportional to the semi-width of the contact area a, as reported in Table 7. This implies that the mesh used at load Q_1 has different element dimensions compared to the mesh employed at load Q_2. Furthermore, since $a_o > a_i$, then $l_o > l_i$. Coefficient η is set equal to 0.06 at Q_1 and 0.22 at Q_2. The table reports the percent difference between the numerical and the theoretical stiffness estimates, namely Δ. The results of the inquiry shows that a good compromise between the accuracy of the computed stiffness and the simulation time is obtained for the third mesh, which is denoted by $w = 1.25a$ and $l = 1.25a$. The results obtained with this load-dependent improved mesh are reported in Figure 16. At applied load $Q_1 = 37$ N, maximum percentage difference with respect to the finest generated CTETRA10 grid is 0.1%, while the computational time reduces by 82.6%. At applied load $Q_1 = 3706$ N, instead, stiffness deviates by 1.8% and time decreases by two orders of magnitude. As a result, it is possible to conclude that the chosen mesh dimensions allow to greatly reduce the simulation time, especially at lower loads.

The same comparison is replicated for linear and parabolic hexahedral elements, as reported in Figure 17. Similarly to tetrahedral elements, hexahedral

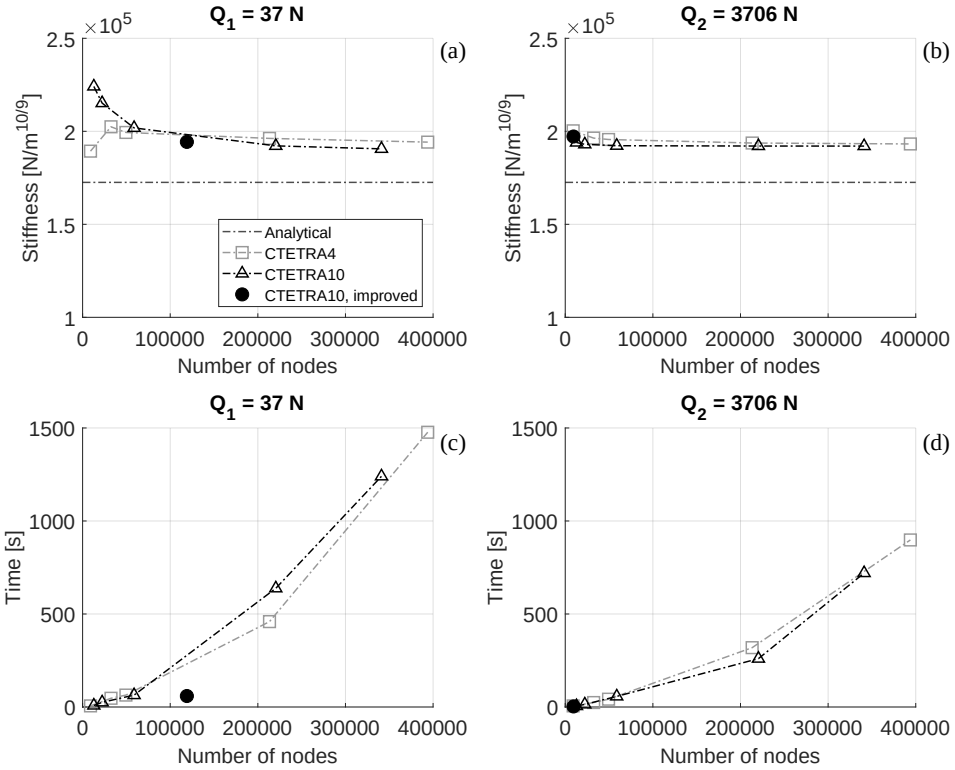

Figure 16: Mesh performance evaluation for tetrahedral elements, roller-races contact: (a) contact stiffness at Q_1; (b) contact stiffness at Q_2; (c) computational time at Q_1; (d) computational time at Q_2.

parabolic elements provide stiffness estimates closer to the value at convergence for a lower number of nodes at load Q_1. At load Q_2, on the other hand, both elements provide a stable trend even for a low number of nodes. The computational time, depicted in Figure 17.c and Figure 17.d, appears to be higher for parabolic elements. However, the difference between element types is negligible for coarser grids with a low number of nodes. In conclusion, parabolic elements are employed to generate the improved mesh even for hexahedral elements.

Table 8 reports the tested combinations of w and l employed to seek a relationship between the dimension of CHEXA20 elements and the size of the

Table 7: Tested combinations of w and l for CTETRA10 elements, roller-races contact.

			Q_1 = 37 N		Q_2 = 3706 N	
Mesh	w [mm]	l [mm]	Δ [%]	t [s]	Δ [%]	t [s]
1	a	a	11.6	65	13.0	4
2	1.5a	0.75a	10.9	129	12.1	10
3	1.25a	1.25a	11.8	64	13.2	3
4	1.5a	1.5a	12.6	59	14.3	2
5	2a	2a	14.5	59	17.6	2

contact area. In this case, coefficient η is set as equal to 0.06 at Q_1 and 0.50 at Q_2. The results shows that the third mesh, characterized by $w = 1.5a$ and $l = 1.5a$, provides a good compromise between computational time and accuracy of the results. The stiffness estimate obtained through the load-dependent improved mesh are reported in Figure 17. With respect to the finest CHEXA20 grid, the stiffness deviates by 0.5% for both loads. Computational time, on the other hand, appears to be reduced by one to two orders of magnitude. Therefore, even in this case, the improved mesh permits to greatly reduce the computational time while maintaining a sufficient accuracy of the results.

It is also noteworthy to compare the performance of hexahedral and tetrahedral improved meshes. At load Q_2, the computational time differs by a small margin. At load Q_1, however, the simulation time with tetrahedral elements is almost three times higher than the one associated with hexahedral elements. This means that for lower mesh dimensions the hexahedral elements provide a sensible improvement in computational time compared to tetrahedral elements. On the contrary, the difference is negligible at higher loads.

3.3.3 *Ball-races contact*

For the reference ball bearing, the geometry is reduced to an angular sector equal to 18°. Moreover, the model is further cut by exploiting the two available symmetry planes. As a consequence, the initial geometry reduces to a quarter

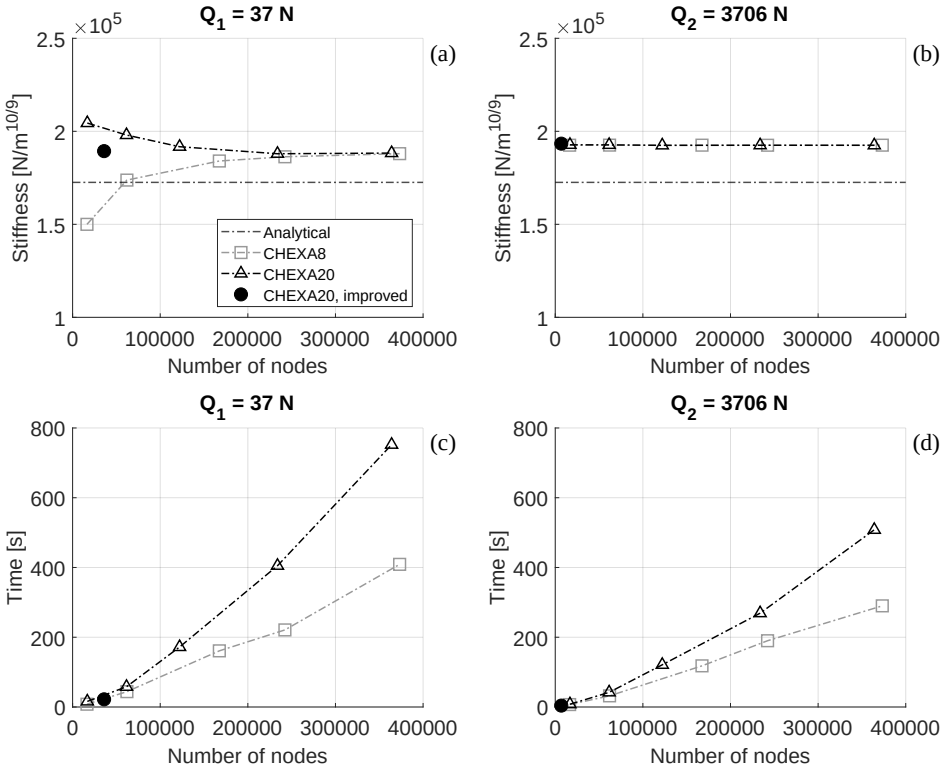

Figure 17: Mesh performance evaluation for hexahedral elements, roller-races contact: (a) contact stiffness at Q_1; (b) contact stiffness at Q_2; (c) computational time at Q_1; (d) computational time at Q_2.

of the original ball in contact with a portion of both races, as depicted in Figure 15.c. By considering two extreme radial load values acting on the bearing equal to 0.5 kN and 23 kN, the two load values employed for mesh evaluation are equal to $Q_1 = 219$ N and $Q_2 = 10052$ N. The dimensions of the contact areas for these force values are $a_i = 0.986$ mm, $b_i = 0.122$ mm, $a_o = 0.994$ mm and $b_o = 0.116$ mm at load Q_1 and $a_i = 3.529$ mm, $b_i = 0.435$ mm, $a_o = 3.558$ mm and $b_o = 0.416$ mm at load Q_2. Conservative values $w_a = 4$ mm and $w_b = 0.7$ mm are employed to generate the elliptical contact area for mesh performance evaluation.

Table 8: Tested combinations of w and l for CHEXA20 elements, roller-races contact.

			Q_1 = 37 N		Q_2 = 3706 N	
Mesh	w [mm]	l [mm]	Δ [%]	t [s]	Δ [%]	t [s]
1	a	0.5a	11.0	51	12.6	21
2	a	a	10.2	31	13.3	4
3	1.5a	1.5a	9.7	22	12.1	4
4	1.75a	1.75a	10.8	16	14.9	3
5	2a	2a	12.2	15	14.8	2

Table 9: Statistics of tetrahedral meshes employed to assess element performance in ball-race contact.

	CTETRA4			CTETRA10		
Mesh	l [mm]	g [mm]	Number of nodes	l [mm]	g [mm]	Number of nodes
1	0.08	0.571	16879	0.2	1.429	27327
2	0.06	0.429	29400	0.15	1.071	36535
3	0.04	0.286	68857	0.1	0.714	73546
4	0.03	0.214	131552	0.07	0.500	146656
5	0.02	0.143	326626	0.05	0.357	302163

Only tetrahedral elements are investigated for ball-races contact. A set of five meshes with decreasing mesh size is generated for each element type, for a total of 10 meshes among linear and parabolic elements. Figure 15.c shows one of the employed tetrahedral meshes. Coefficient η is set equal to 0.14. Local mesh size, global mesh size and number of nodes for all meshes are detailed in Table 9.

Mesh performance is evaluated for all meshes with Eq. 14. The results are reported in Figure 18. Figure 18.a and Figure 18.b show the estimated contact stiffness at load Q_1 and Q_2, respectively, while Figure 18.c and Figure 18.d depict the associated computational time. In addition, the plots also report the

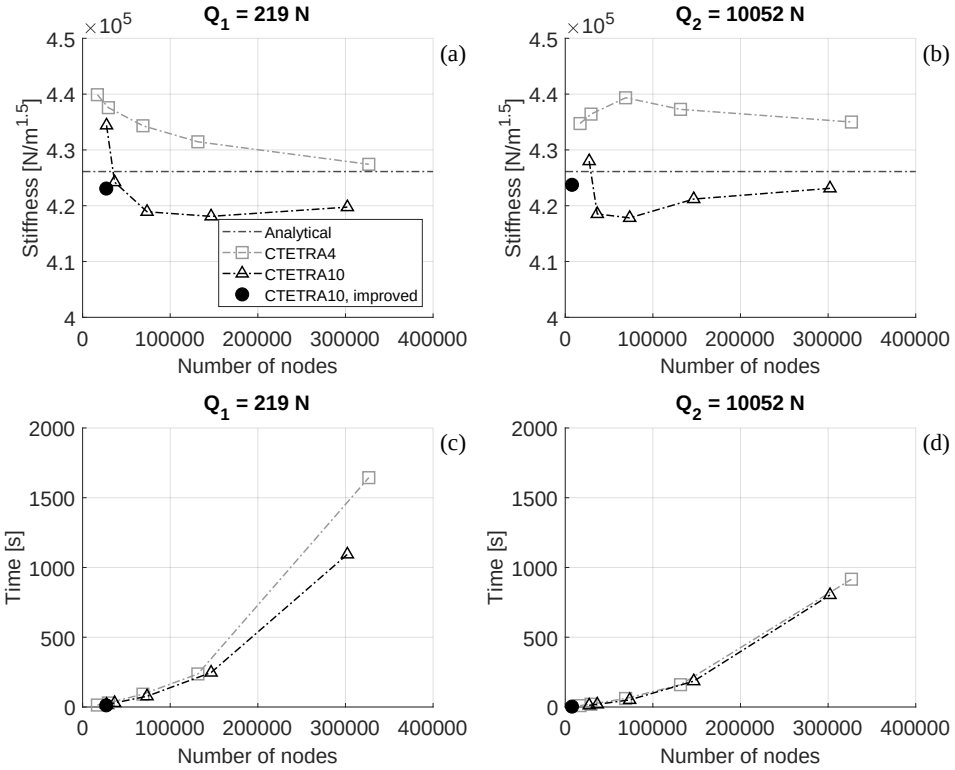

Figure 18: Mesh performance evaluation for tetrahedral elements, ball-races contact: (a) contact stiffness at Q_1; (b) contact stiffness at Q_2; (c) computational time at Q_1; (d) computational time at Q_2.

analytical value of the contact stiffness, namely $k_{c,b}$. According to [19], this quantity may be estimated separately for the inner and outer ring contacts through the following equation:

$$k_{c,bk} = \overline{\xi}\left[\left(\frac{4.5}{\overline{\epsilon}_k r_{eq,k}}\right)\left(\frac{1}{\pi\overline{\kappa}_k E_{eq,k}}\right)^2\right]^{-1/2} \tag{17}$$

where k stands for i or o depending on the considered race in contact. The total stiffness may be subsequently calculated as:

$$k_c = \left[\frac{1}{(k_{c,bi})^{(1/n)}} + \frac{1}{(k_{c,bo})^{(1/n)}}\right]^{-n} \tag{18}$$

Table 10: Tested combinations of w_a, w_b and l for CTETRA10 elements, ball-races contact.

				Q_1 = 219 N		Q_2 = 10052 N	
Mesh	w_a	w_b	l	Δ [%]	t [s]	Δ [%]	t [s]
1	2a	2b	0.5b	-0.3	42	-0.7	9
2	1.5a	1.5b	0.5b	-0.9	31	-0.6	8
3	1.5a	1.5b	b	-0.7	12	-0.6	3
4	1.25a	1.25b	1.25b	3.7	11	-1.1	2
5	1.5a	1.5b	1.5b	5.7	5	5.5	2

Therefore, for the ball bearing under investigation, the resulting stiffness is $k_{c,b}$ = 426124 N/mm$^{1.5}$. It may be observed that both element types provide a good estimate of the contact stiffness compared to analytical results. The computational times are also similar, although they are slightly higher for linear elements at the lowest load. However, parabolic elements provide results closer to converged ones for a lower number of nodes especially at the highest load. Therefore, CTETRA10 elements are chosen to generate the improved mesh. Table 10 reports the tested combinations of w_a, w_b and l for both loads. In this case, since the contact area is defined by two dimensions, l is defined as proportional to the length of the semi-minor axis b. Furthermore, coefficient η is set as equal to 0.06 at Q_1 and 0.14 at Q_2. The proposed investigation shows that by employing the third mesh, which is defined by w_a = 1.5a, w_b = 1.5b and l = b, it is possible to obtain a good compromise between estimated contact stiffness and computational time. Stiffness and computational time values obtained with the two improved meshes may be seen in Figure 18. The deviation from the finest grids are 0.8% and 0.1% at loads Q_1 and Q_2, respectively. In addition, in both cases the computational time decreases by two orders of magnitude. In conclusion, these mesh dimensions allow to considerably reduce the computational effort.

3.4 ESTIMATION OF THE BEARING RADIAL STIFFNESS

This section is dedicated to the description of the procedure adopted to model the reference bearings and to estimate their radial stiffness. The methodology is employed to generate the bearing meshes by exploiting the meshing procedure introduced in Section 3.3. First, the analytical procedure reported by D. Petersen et al. [88] is described. This formulation is utilized as reference to validate the results of the proposed approach. Then, techniques that allow to reduce the computational domain are described. The reduced geometries of both bearings are employed to generate the load-dependent meshes. Finally, the numerically estimated stiffness values are compared against the results of the analytical model.

3.4.1 *Analytical procedure*

The radial stiffness may be analytically computed through a variety of methods. Within the context of this work, the analytical estimate is employed to validate the results of the proposed numerical procedure. The method described in this section was developed by D. Petersen et al. [88] on the basis of the work of N. Sawalhi and R.B. Randall [24].

A single row-ball bearing is schematically depicted in Figure 19. The problem is reduced to a 2D xy plane, in which the only considered components are the two rings and n_b rolling elements, which may be either balls or rollers. Within this framework, load w represents the total load whose maximum value is aligned with the direction of the vector itself. To estimate the radial stiffness of this component, which is subjected to a load in the y direction, it is necessary to evaluate the stiffness component k_{yy}. According to [88], for a fixed position of the rolling elements with respect to the reference system Oxy, its value may be calculated as:

$$k_{yy} = n k_c \sum_{j=1}^{n_b} \overline{\delta}_j^{n-1} \overline{\gamma}_j \cos^2 \phi_j \tag{19}$$

In Eq. 19, $\overline{\delta}_j$ is the quasi-static relative displacement of the j-th rolling element at the angular position ϕ_j and it is equal to:

$$\overline{\delta}_j = \overline{\delta}_x \cos \phi_j + \overline{\delta}_y \sin \phi_j - h. \tag{20}$$

where h is the radial clearance, whilst $\overline{\delta}_x$ and $\overline{\delta}_y$ are the quasi-static relative displacements of the inner ring with respect the the outer ring in the x and y directions, i.e.:

$$\overline{\delta}_x = \overline{x}_i - \overline{x}_o \tag{21}$$

$$\overline{\delta}_y = \overline{y}_i - \overline{y}_o \tag{22}$$

The relative displacements depend on the magnitude of the applied load w. Their values may be evaluated by iteratively solving the equilibrium equation of the system. In fact, by considering that the quasi-static force exchanged by the j-th rolling element with the rings, namely $\overline{f}_{c,j}$, is equal to:

$$\overline{f}_{c,j} = k_c \overline{\delta}_j^n \overline{\gamma}_j \tag{23}$$

then, the static equilibrium of the system may be written as:

$$k_c \sum_{j=1}^{n_b} \overline{\delta}_j^n \overline{\gamma}_j \begin{bmatrix} \cos\phi_j \\ \sin\phi_j \end{bmatrix} = \begin{bmatrix} 0 \\ w \end{bmatrix} \tag{24}$$

where $\overline{\gamma}_j$, which is also present in Eqs. 19 and 23, is a term that takes into account the fact that forces are exchanged only when the rolling elements are under compression, viz.:

$$\overline{\gamma}_j = \begin{cases} 1 & \text{if } \overline{\delta}_j \geqslant 0 \\ 0 & \text{if } \overline{\delta}_j < 0 \end{cases} \tag{25}$$

Therefore, Eq. 24 may be solved iteratively, e.g. by means of the Newton-Raphson method [228], in order to find the values of $\overline{\delta}_x$ and $\overline{\delta}_y$ for a given load value w. Once they are known, it is possible to compute k_{yy} and consequently determine the load-stiffness relationship within the required range of loads.

3.4.2 *Reduction of the computational domain*

For both reference bearings, the geometry is reduced in order to decrease the computational burden of the simulations. To achieve this goal, a series of modelling strategies are introduced. The first one consists in exploiting the

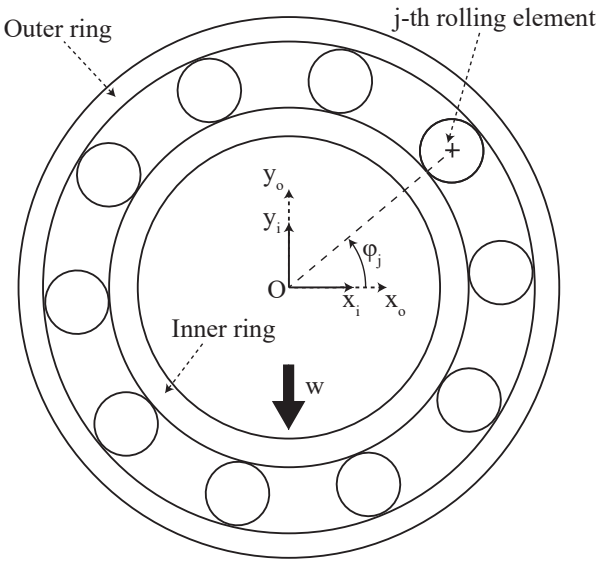

Figure 19: Scheme of the bearing employed for analytical stiffness estimation.

two symmetry planes available for both reference bearings. The first plane is the one passing through the shaft axis on the plane of maximum load. The second plane is the one normal to the shaft axis by halving the bearing width. The latter may be assumed only in absence of misalignment and only for a pure radial load.

In fact, the number of rollers carrying the load from one ring to the other depends on the extent of the load zone, namely ϕ_l, which is equal to [6]:

$$\varphi_l = 2\cos^{-1}\left(\frac{h}{\delta_r}\right) \tag{26}$$

where δ_r is the radial displacement. When $h = 0$, the angular extent of the load zone is 180° as long as $\delta_r \neq 0$. As a result, the number of rolling elements in both reference bearings is reduced to 3, one of which is also cut in half by a symmetry plane. This consideration allows to greatly reduce the computational burden. It is also worth noting that, when $h > 0$, φ_l reduces, while it increases when $h < 0$. In these cases, attention must be paid on how many rollers are effectively loaded to avoid erroneously removing load-carrying rollers.

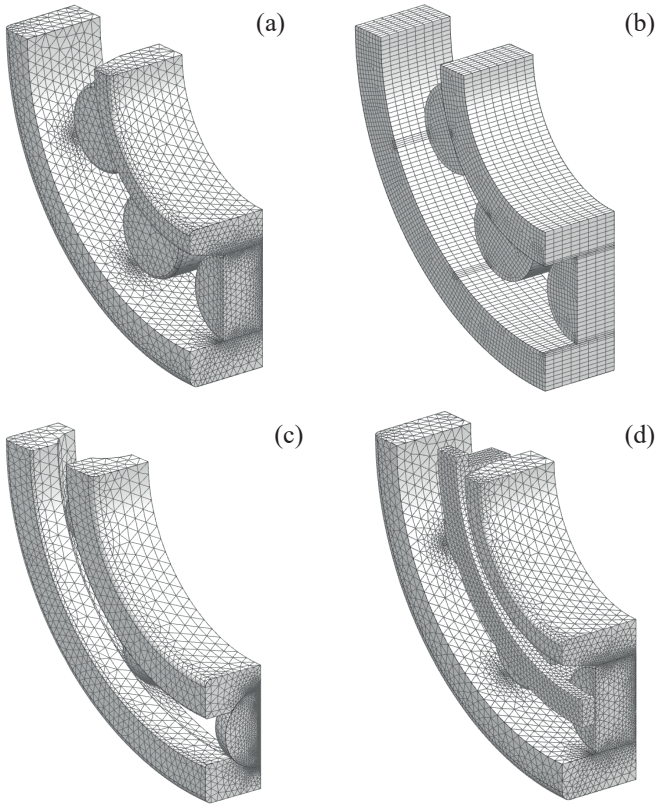

Figure 20: Reduced bearing geometries with different meshes: (a) roller bearing, tetrahedral elements; (b) roller bearing, hexahedral elements; (c) ball bearing, tetrahedral elements; (d) roller bearing with cage.

Due to the above hypotheses, the unloaded portions of the inner and outer ring are also removed. This is done by assuming that they do not give a significant contribution to the radial displacement in the plane of maximum load, i.e. the monitored displacement for subsequent stiffness assessment. All these practical assumptions allow to greatly reduce the number of nodes needed to generate the mesh of the mechanical component. Examples of the resulting FE models are reported in Figure 20.

Concerning the boundary conditions, it is assumed that the outer ring is mounted on an infinitely rigid frame, i.e. all DOFs on the outer surface of the inner ring are restrained. Besides, symmetry conditions are applied on

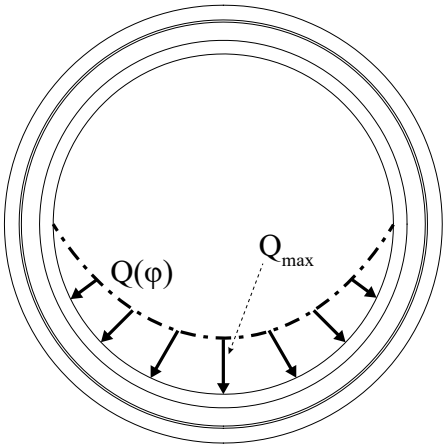

Figure 21: Sinusoidal load distribution acting on a 180° angular sector of the inner surface of the inner ring.

the nodes lying on the symmetry planes. For these nodes, the displacements normal to the planes are constrained.

Finally, attention is devoted to alternative methods to apply the radial load. As seen in Section 1.2.1, researchers have employed various methods to apply the radial load, e.g. on the center of the shaft, on rings or on a central node connected with rigid links to the inner ring. In order to decrease the number of grid nodes, it is suggested to model the shaft contribution by means of an analytical sinusoidal distribution which replicates the load applied by a radially loaded shaft on the inner ring. Such load also depends on the radial clearance h of the bearing. In absence of clearance, however, the load insists on a 180° angular sector as depicted in Figure 21. This permits to remove one contacting body while still taking into account its effect. The efficiency of this assumption is verified in the next section, after the definition of the bearing mesh.

3.4.3 *Generation of the load-dependent meshes*

In a REB, the load-displacement relationship is non-linear. As a consequence, the radial stiffness depends on the magnitude of the applied load. This means

that this parameter must be evaluated at different loads to obtain the stiffness curve over a load range. This procedure is implemented in a FE software by choosing an arbitrary number of loads. For each load, an appropriate mesh is generated. The local and global element dimensions are tuned to be proportional to the dimensions of the contact area as described in Section 3.3.2 and Section 3.3.3 for roller-races and ball-races contact, respectively. The chosen radial loads for both reference bearings are reported in Table 11 along with the corresponding dimensions of the contact areas for the inner and outer ring contacts. For the roller bearing, computational contact area and mesh dimensions depend on the element type. For tetrahedral elements, w_a = 1.25a and l = 1.25a, while for hexahedral elements w_a = 1.5a and l = 1.5a. Furthermore, for the meshes generated with CHEXA20 elements, the number of elements along the shaft axis is set equal to 4 in the rollers and equal to 8 in the rings. In both cases, the length of the rectangular contact area is equal to l_{eff} for every load value. On the other hand, only tetrahedral elements are considered for the ball bearing. In this case, w_a = 1.5a, w_b = 1.5b and l = b. In all cases, a and b are different for inner and outer ring contacts. Finally, the values of η are adapted to the local mesh size at all loads.

Six radial load values have been chosen for both bearings. Figure 20.a and Figure 20.b show an example of tetrahedral and hexahedral meshes applied to the roller bearing, while Figure 20.c depicts a tetrahedral mesh of the ball bearing. Despite the longer setup time needed to generate such a number of grids, the mesh-generation procedure may be automatized in a FE software to consistently reduce the required time. Moreover, the main benefit of the proposed procedure lies in the diminished simulation time due to the tuning of the mesh size according to the applied load value.

Before estimating the radial stiffness, it is worth verifying the load application method. To this end, the efficiency of the technique is demonstrated by evaluating the load distribution on rollers. Hexahedral grids are chosen to compute the numerical estimate, for which a sinusoidal distribution is employed. Obtained results are then compared to distributions computed with other two methods. The first one is the analytical formulation reported in [6]. The load F as a function of the angular coordinate φ, for $h = 0$, is equal to:

$$Q(\varphi) = Q_{max}[\cos(\varphi)]^{10/9} \tag{27}$$

Table 11: Contact areas dimensions for the roller and ball bearings under different loads.

Roller Bearing			Ball Bearing				
F_r [N]	a_i [mm]	a_o [mm]	F_r [N]	a_i [mm]	b_i [mm]	a_o [mm]	b_o [mm]
100	0.013	0.016	500	0.986	0.122	0.994	0.116
500	0.028	0.036	2500	1.684	0.208	1.698	0.199
1000	0.040	0.050	5000	2.122	0.262	2.139	0.250
2500	0.064	0.079	10000	2.674	0.330	2.695	0.315
5000	0.090	0.112	15000	3.061	0.377	3.085	0.361
10000	0.127	0.159	23000	3.529	0.435	3.558	0.416

where Q_{max} is calculated through Eq. 13. The second method consists in modeling the shaft and applying the total load in its central node. Size of the elements in the shaft mesh is the same as in the inner ring. Since the radial clearance is not considered, shaft and inner ring touch on the entirety of the available contact surface. Two radial loads F_{r1} = 0.1 kN and F_{r2} = 10 kN are applied. For the two FE methods, load distribution is calculated for each roller as the sum of the contact forces in the corresponding contact areas. Results are shown in Figure 22, where the roller at angle 0° is the one whose centre is aligned with the direction of maximum load. The sinusoidal load allows to obtain a load distribution which deviates by up to 3% from analytical results. The loaded shaft, on the other hand, shows a similar difference for the rollers at 0° and 32.7° while it deviates by 8.3% at maximum for the roller at 62.4°. Besides, the addition of the shaft causes the computational time to increase by 28.7% for F_{r1} and by 56.8% for F_{r2}. These results show that it is possible to remove the shaft and substitute it with a sinusoidal load in order to eliminate one contacting body and lead to a consistent reduction in computational time.

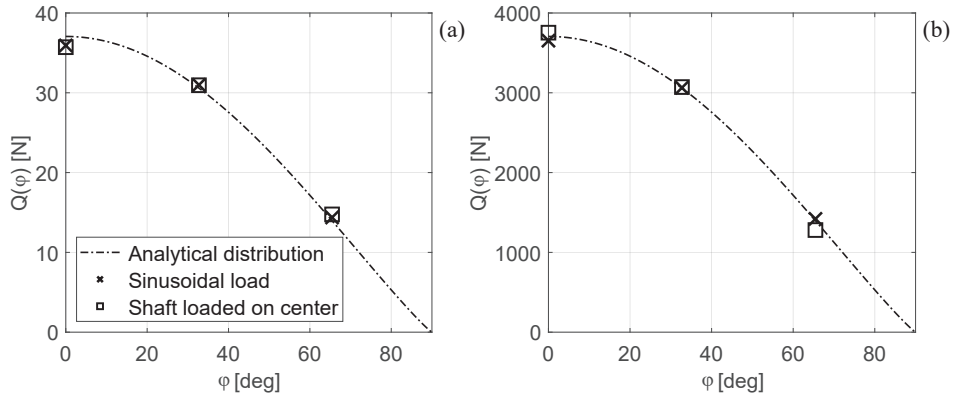

Figure 22: Load distribution on rollers evaluated for two different radial load values: (a) F_{r1} = 0.1 kN; (b) F_{r2} = 10 kN.

3.4.4 *Radial stiffness estimation: comparison between numerical and analytical results*

According to Y. Guo and R.G. Parker [10], stiffness k at a given load F_r may be computed as:

$$k = \frac{\partial F_r}{\partial \delta_r}, \tag{28}$$

where δ_r is the radial displacement generated by the applied load. Consequently, in order to compute the derivatives via FE analysis, it is necessary to evaluate two radial displacement for each value of F_r. By introducing a small disturbance δF, it is possible to rewrite Eq. 28 as:

$$k = \frac{F_r^+ - F_r^-}{\delta_r(F_r^+) - \delta_r(F_r^-)}, \tag{29}$$

where $F_r^+ = F_r + \delta F$, $F_r^- = F_r - \delta F$, and δF is arbitrarily assumed to be 0.01% of F_r at each loading condition. This means that two displacement evaluations are required at each load value. Radial displacement δ_r is evaluated as the average displacement of the nodes on the plane of maximum load on the inner surface of the inner ring.

Concerning the reference roller bearing *NU 202 ECP*, Figure 23.a depicts the load-stiffness curve, determined through the FE method, with either tetrahedral

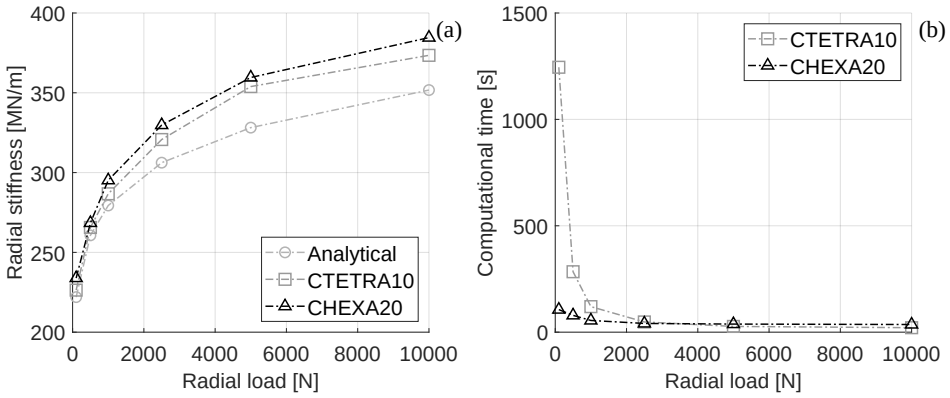

Figure 23: Roller bearing stiffness: (a) load-stiffness curve; (b) computational time.

or hexahedral elements. The results obtained with the two element types are very close, since the maximum percent difference between them is 3.3% at F_r = 100 N. However, Figure 23.b demonstrates that the computational time is remarkably higher for tetrahedral elements when the applied load reduces below 1000 N. For higher loads, on the contrary, the simulation time is similar. Consequently, hexahedral elements are more suitable at lower load values, while at higher loads both element types may be employed without excessively increasing the simulation burden. Furthermore, the results in Figure 23.a are compared against the analytical formulation described in Section 3.4.1 to assess the goodness of the estimated stiffness values. It is possible to observe that the numerical results are higher than the analytical ones at all load values. Nonetheless, the difference with the numerical results is lower than 10% in all cases. Therefore, the proposed procedure is able to efficiently evaluate the bearing radial stiffness of roller bearings.

Similarly, Figure 24.a shows the stiffness curve for the reference ball bearing, which has been modelled with tetrahedral elements only. The results agree well with the analytical formulation, although the results are close at lower loads compared to higher loads. In fact, the differ by 4.6% at 500 N and by 12.3% at 23000 N. Besides, Figure 24.b shows that the computational time increases with lower loads. However, the computational effort is considerably lower compared to the roller bearing meshed with the same element type and is comparable to the simulation times obtained with hexahedral elements.

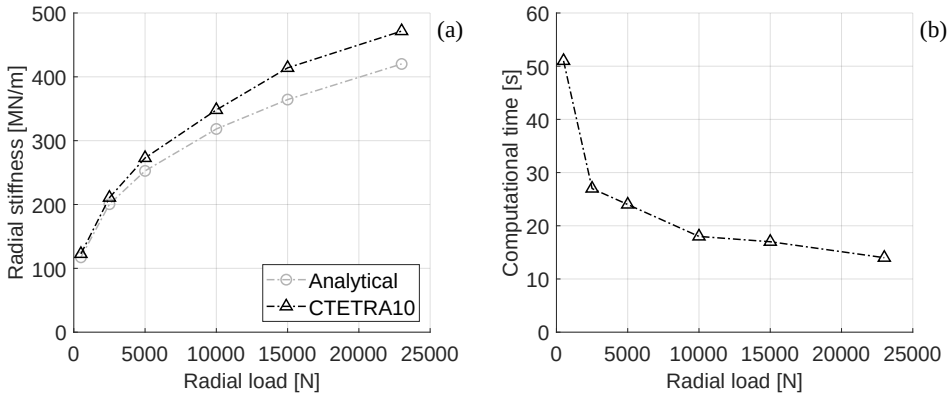

Figure 24: Ball bearing stiffness: (a) load-stiffness curve; (b) computational time.

In conclusion, the proposed procedure also allows to accurately estimate the radial stiffness of ball bearings.

Finally, is worth underlining that the employed approach requires a longer set-up time compared to traditional models as a different grid has to be generated for each load. However, this additional time is compensated by the decreased computational time needed to solve the problem and the lower time required to perform convergence checks, as the element dimension is analytically determined. Moreover, the setup process may be further accelerated by automatizing the procedure in a FE software. Despite the increased number of generated grids, then, the computational time may be lessened compared to traditional approaches.

3.5 ADDITIONAL EFFECTS INFLUENCING THE RADIAL STIFFNESS

This section discusses additional effects that may influence the radial stiffness. Only the roller bearing is considered in this investigation, but the same inquiry may be repeated for the ball bearing. Additional simulations in which the cage is inserted into the model are run in order to determine if it is reasonable to neglect this component in REBs modelling. Influence of roller position is assessed by varying their location with respect to the direction of maximum

load. Lastly, the model is further improved to take into account the radial clearance.

3.5.1 *Inclusion of the cage*

Cage is commonly removed in bearing simulation to lighten the mesh and reduce the number of contacts. Some researchers took it into account by adding rigid connectors between the rolling elements [67]. However, it is worth investigating how much this component might affect the radial stiffness. To this end, a steel cage is modelled and inserted into the model. Parabolic tetrahedral elements are used for all components, as it may be observed in Figure 20.d, since they allow to model the complex cage geometry. The cage is modelled so that each pocket contacts the corresponding roller on the circumferential direction. A 0.1 mm axial clearance is also considered between each rolling element and its cage slot. The mesh is characterized by an average dimension proportional to a at each load through a coefficient of proportion 2η. The results for the models with and without cage are depicted in Figure 25. Figure 25.a shows that the stiffness deviates by 3.7% at the highest load, while for all other loading conditions the difference is lower than 1.4%. On the other hand, Figure 25.b displays that the computational time greatly rises at lower loads if the cage is modelled. Besides, the simulation time is from 2 to 4 times higher at all loads. To sum up, the proposed procedure allows to model the cage. However, its insertion in the model is not suggested as it considerably increases the computational effort without providing a significant change in the computed stiffness.

3.5.2 *Influence of maximum load direction*

The results reported in Section 3.4.4 have been obtained by assuming that the direction of maximum load passes through the middle of one roller. This situation is depicted in Figure 26.a and it is named position 1. However, the position of the rolling elements changes during operation, leading to a subsequent change of the bearing stiffness over time. The influence of load direction is assessed by considering a case in which the plane of maximum load passes through two consecutive rollers. Such a condition, namely position

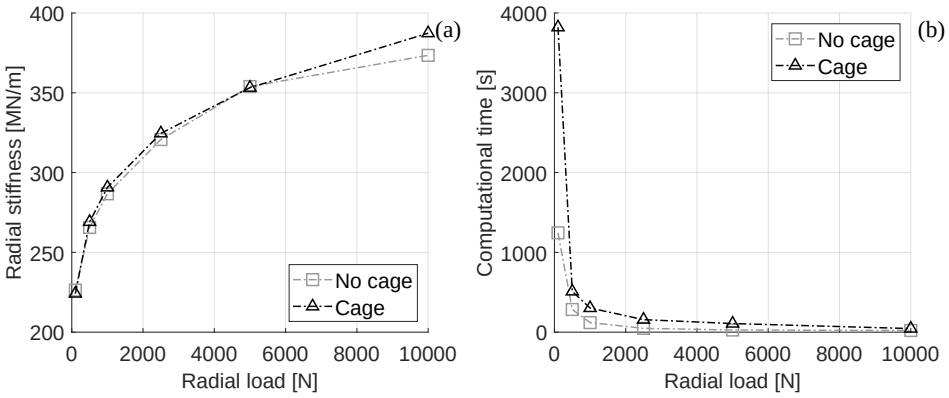

Figure 25: Effect of cage presence on roller bearing stiffness: (a) load-stiffness curve; (b) computational time.

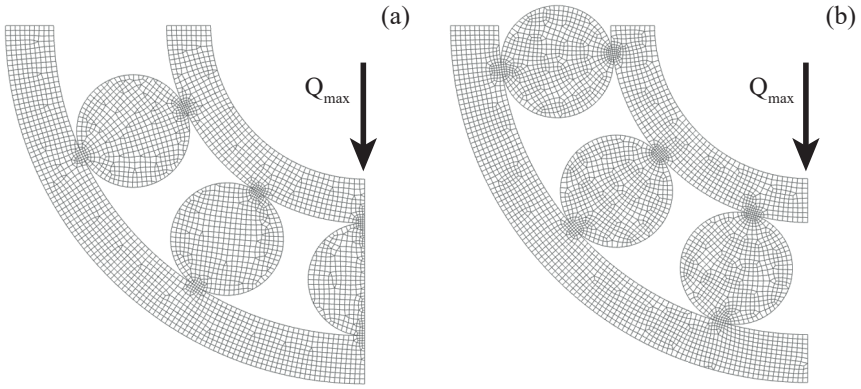

Figure 26: Different directions of the maximum load on rollers: (a) position 1; (b) position 2.

2, is shown in Figure 26.b. The radial stiffness in positions 1 and 2 is evaluated by employing an hexahedral mesh. The simulation results are depicted in Figure 27. The stiffness curve in Figure 27.a shows that there is a slight change in the computed stiffness, up to 7.5% for a radial load equal to 500 N. Besides, the computational time is higher at all loads since there are no sectioned rollers in position 2. Therefore, the presented method allows to take into account different load directions. For the reference bearing, however, the estimated stiffness is not considerably affected.

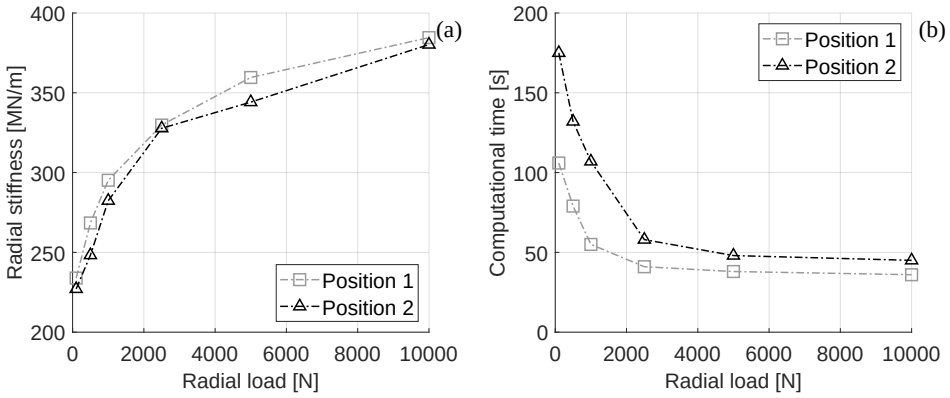

Figure 27: Influence of maximum load direction on roller bearing stiffness: (a) load-stiffness curve; (b) computational time.

3.5.3 *Clearance effect*

Radial clearance, namely h, is an additional factor that affects radial bearing stiffness. Indeed, Eq. 26 shows that for values of h greater than 0 the extent of the load zone reduces, leading to a redistribution of the load among the rollers. In particular, in presence of clearance, the maximum load on roller is different from the one estimated through Eq. 13. In fact, for non-zero values of clearance it is ruled by the following equation [6]:

$$Q_{max} = k_c \left(\delta_r - h\right)^n \tag{30}$$

where k_c is the contact stiffness for either the roller-races or ball-races case. Furthermore, the number of loaded rollers may decrease, depending on the applied load and the angular span between two consecutive rollers. In fact, if the number of rollers decreases, the radial displacement δ_r increases, thus diminishing the radial bearing stiffness. The value of Q_{max} is needed to compute the contact area at each load in order to define the element size for the associated mesh, but Eq. 30 cannot be directly solved since the radial displacement is not initially known as it is equal to

$$\delta_r = h + \left(\frac{Q}{n_b k_c J_r(h, \delta_r)}\right)^{1/n} \tag{31}$$

where J_r is a radial integral depending on h and δ_r. Since J_r depends on the radial displacement, Eq. 31 may be solved by means of an iterative procedure.

Table 12: Clearance effect on contact area, load zone and number of loaded rollers. The number of loaded rollers refers to a bearing in position 1 as in Figure 26.a, for which the maximum number of loaded rollers for $h = 0$ is equal to three.

F_r [N]	a_i [mm]	a_o [mm]	φ_l [°]	Loaded Rollers
100	0.020	0.025	38.0	1
500	0.041	0.051	67.6	2
1000	0.053	0.067	82.4	2
2500	0.076	0.095	104.2	2
5000	0.101	0.127	121.6	2
10000	0.137	0.172	138.2	3

Once δ_r is known for each each radial load, Q_{max} is calculated by employing Eq. 30. Finally, contact area semi-width is computed through Eq. 4 while the loaded angular sector is calculated via Eq. 26.

The proposed meshing method is employed to determine REB stiffness in presence of a positive radial clearance h = 0.02 mm. Load dependent meshes to be employed for FE simulation are generated according to the methods shown in Section 3.4.3. The extent of the load zone, namely φ_l, decreases with lower loads, as reported in Table 12. As a consequence, the number of loaded rollers may reduce if the load is sufficiently low. In fact, as it may observed in Table 12, the number of loaded rollers remains equal to 3 only for F_r = 10000 N. Therefore, the unloaded rollers may be removed from the FE model as they no longer affect the radial stiffness. In addition, the load zone extent reduction generates an higher force on the plane of maximum load. Consequently, the dimensions of the contact area also increase in both the inner and outer race, as detailed in Table 12. Moreover, the analytical sinusoidal load is not applied on a 180° angular sector as in the simulations performed for h = 0 but the angular span is changed at each load according to φ_l values reported in Table 12.

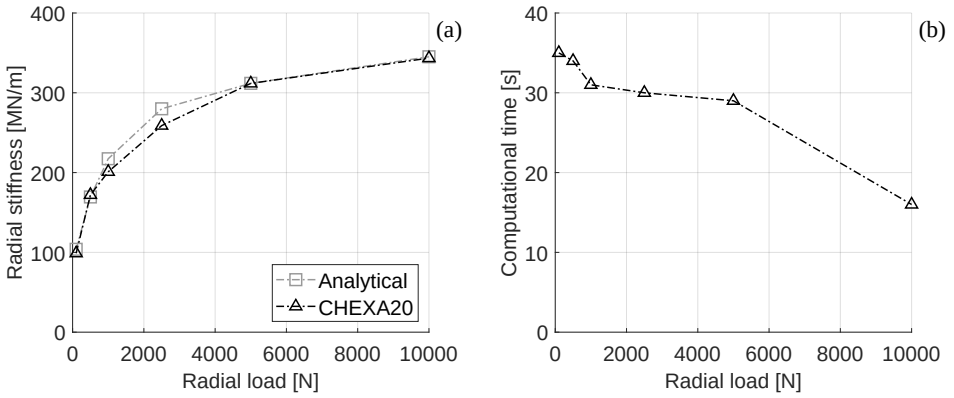

Figure 28: Effect of a radial clearance $h = 0.02$ mm on roller bearing stiffness: (a) load-stiffness curve; (b) computational time.

Figure 28.a shows the estimated radial stiffness along with the analytical results computed through the method described in [88]. The numerical results agree well with the analytical ones at all loads, with a maximum deviation equal to 7.5% at 2500 N. Besides, Figure 28.b details the computational time. Compared to Figure 27.b, the computational effort at lower loads decreases thanks to the removal of unloaded rollers. Moreover, the simulation time also diminishes at higher loads since the contact areas are larger. The investigation suggests that the proposed method may be employed to quantify bearing radial stiffness when $h > 0$. In this case, however, analytical computations must be carried out to determine Q_{max} and φ_l, leading to more calculations and a longer setup compared to the case in which $h = 0$.

3.6 SUMMARIZING REMARKS

This chapter discussed a procedure to evaluate the radial bearing stiffness through the FE method by means of load-dependent meshes. The proposed methodology consists in generating load-dependent grids, so that the mesh size in the contact area is properly tuned according to the dimensions of the analytically estimated contact zone between rolling elements and raceways. This strategy allows to obtain proper element dimensions to capture the contact phenomenon for a wide range of loads, while limiting the number of

grid points and curbing the computational time. Therefore, the main aspect of originality in the described work arises form to this peculiar approach employed in the mesh generation process.

The method was applied to a cylindrical roller bearing and a deep-groove ball bearing. Their meshes were generated by employing either parabolic tetrahedral or parabolic hexahedral elements. For the former, it was found that the most efficient local element dimensions were $l = 1.25a$ and $l = 1.5a$ for tetrahedral and hexahedral elements, respectively. The width of the contact area, w, was equal to l in both cases. For the latter, it was shown that appropriate dimensions were $w_a = 1.5a$, $w_b = 1.5b$ and $l = b$. Concerning the size of the computational domain, it was reduced for both bearings by taking advantage of the available symmetry planes, by removing unloaded rollers and by replacing the shaft with an equivalent sinusoidal load distribution. The estimated stiffness results showed that both considered element types provided similar radial stiffness estimates, although hexahedral elements exhibited considerably higher computational times at low loads. Good agreement was attained between the stiffness estimated with the proposed method and the analytical formulation described in [88], thus denoting the quality of the proposed methodology.

The technique was also successfully employed to determine the effect of cage modelling, load direction and clearance on the numerically evaluated stiffness. These influencing factors were considered by taking as reference the roller bearing. Concerning the effect of the direction of maximum load, the analysis showed a deviation on the computed stiffness by up to 6.7%, which indicates that the inclusion of this effect within the analysis might become mandatory depending on the required degree of accuracy. It must be noted, however, that the total computational time increases if multiple roller positions are simulated. Adding the cage as a meshed body inside the FE model led to a slight change in stiffness, whereas the computational time consistently increased. Therefore, modeling the cage is not suggested if the goal is to attain a low computational time. Finally, clearance effect was assessed by physically inserting it between the components and by changing the sinusoidal load distribution in order to apply the load in a reduced angular sector. Unloaded rollers were also

removed when the load angular span was sufficiently small. Numerical results were in good agreement with analytical formulations.

The meshing procedure described in this chapter permits to obtain the radial stiffness with a relatively low computational effort. Therefore, it may be employed to estimate this parameter for different load conditions, with or without clearance and for any position of the rolling elements. The proposed methodology was tested on a cylindrical roller bearing and a deep-groove ball bearing, but it may also be employed to model other bearing types. In these cases, however, attention should be paid for possible differences in the required modelling. For instance, while the symmetry plane passing through the shaft axis is usually present, the second symmetry plane may not be always available, e.g. for tapered roller bearing and self-aligning double-row ball bearings. Consequently, the symmetry conditions should be always checked before meshing the model as they are not the same for every bearing. Nevertheless, the other assumptions concerning load application, unloaded rollers removal, element type and mesh size are still applicable for all rolling-element bearings subjected to a radial load. The main drawback of the proposed method is the necessity to generate a different mesh for each load condition, thus leading to an increase in the time needed to set-up the problem. However, this is compensated by the efficiency of the method in defining an adequate mesh size by employing element dimensions determined by analytical formulae. In this way, convergence checks at all loads are not required, leading to a consistent reduction in the time needed to obtain a reasonable mesh size. Furthermore, the proposed method is based on a systematic approach that may be straightforwardly applied to different bearing geometries and even automatized in a FE software. In conclusion, although the number of generated grids increases, the total time needed to perform the simulations reduces.

4

MULTI-OBJECTIVE OPTIMIZATION PROCEDURE FOR THE ESTIMATION OF UNKNOWN PARAMETERS IN LUMPED-PARAMETER MODELS OF DEFECTIVE BEARINGS

4.1 INTRODUCTION

Section 1.2.2 exhaustively examined the available methods to simulate the dynamic behaviour of faulty bearings. The proposed survey highlighted that, despite the notable amount of works regarding the dynamic simulations of defective bearings and the remarkable quality of the obtained results, it is commonly observed that the procedures associated with the definition of the value of some model parameters are not always thoroughly described. Most notably, the choice of damping values is often based on prior assumptions and on the subsequent validation with experimental results under one or a few test conditions. Similarly, the approaches involving the additional resonant mass introduced by N. Sawalhi and R. Randall [24] propose to select an associated mass and stiffness in order to reproduce a real natural frequency of the system. However, no further guidelines on the choice of such quantities are provided. These considerations could be extended in general for all the parameters that are hardly measurable or that might have uncertain definitions. Therefore, the objective at the basis of the proposed investigation is to detail a procedure for the robust estimation of the unknown parameters in LP models of defective rolling-element bearings. The proposed method takes advantage of a multi-objective optimization technique, in which the employed objective functions are global indicators that take into account the difference between numerical and

experimental results. The optimization variables, i.e. the unknown parameters, are constrained to only assume values within imposed limits that depend on their physical properties. Moreover, additional analytical relationships based on the physics of the problem are introduced to take into account the dependencies among the parameters.

To demonstrate the efficiency of the technique, a detailed discussion of its application for the validation of a numerical model of a self-aligning ball bearing is provided. The employed numerical model is based on the LP dynamic model introduced by D. Petersen et al. [88]. The defect, however, is modelled by taking advantage of the method proposed by Y. Jiang [22] for deep groove ball bearings and by S. Gao et al. [99] for angular contact ball bearings. Differently from the previous authors, it is extended in order to adapt the defect formulation to self-aligning ball bearings. The model is constituted by six degrees of freedom, two of them referring to the displacement of a fictitious high resonant mass that replicates an high frequency response of the system. The mass and stiffness of this additional component, which are unknown, constitute two of the investigated variable parameters. Additionally, further effort is placed upon the estimation of damping values. In fact, in contrast with the classical approach adopted in bearing dynamics, contact damping is herein defined by means of proportional damping in order to control its influence over the frequency components across the whole spectrum. The optimization technique is employed to determine the most suitable values for the damping for the case under exam, since the definition of this parameter is a common source of uncertainties. Concerning the experimental results, they are obtained through an extensive campaign carried out on a dedicate test bench for various defective self-aligning ball bearings subjected to different values of radial load, rotation frequency and defect size. The results of the inquiry demonstrate the capability of the method to fit a large number of test conditions. Nonetheless, because of the global nature of the chosen indicators, some differences remain for a number of cases. Numerical and experimental results for eighteen different combinations of defect dimension, rotation frequency and applied load values are exhaustively discussed in this chapter.

The chapter is subdivided as follows. Section 4.2 details the numerical model and the proposed optimization procedure. To this end, the analytical con-

straints relating the physical quantities in the model are assessed. Then, Section 4.3 reports the results of the experiments described in Section 2.3.1. Moreover, Section 4.4 details the results of the investigation. First, the results of the optimization technique are illustrated. Subsequently, various analyses on the estimated features and on time and frequency signals are performed. The effect of slippage on the numerical results is also investigated. Finally, the chapter closes with some summarizing remarks in Section 4.5.

4.2 NUMERICAL MODEL

This section discusses the lumped parameter model of a double-row self-aligning ball bearing. Firstly, the formulation for an healthy bearing is detailed. Then, the effect of a localized fault on the outer ring is added to generate the faulty bearing model. This set of equations is later employed in the proposed multi-objective optimization procedure. The section also outlines the optimization technique along with the analytical constraints to be imposed on the unknown parameters.

4.2.1 *Healthy bearing model*

The defective bearing model employed in this work is based on the LP model developed by D. Petersen et al. [88]. On the other hand, the defect shape definition is adapted and extended to the case under exam from the work of Y. Jiang et al. [22]. A similar formulation is also applied by S. Gao et al. [99] for an asymmetric defect in an angular contact ball bearing.

The model is schematically depicted in Figure 29. It consists of six degrees of freedom, i.e. the planar displacements in the x and y direction of the three masses representing the system components. The first one is denoted with m_i and represents the mass of the inner ring and the shaft. The second mass, namely m_o, takes into account the outer ring and the casing. Lastly, mass m_r corresponds to an additional mass connected with m_o which allows to model a typical response of the system in a high frequency range of the spectrum [24]. This mass considerably affects the system characteristics, as it will be demonstrated in Section 4.4.1. Two planar displacements are associated with

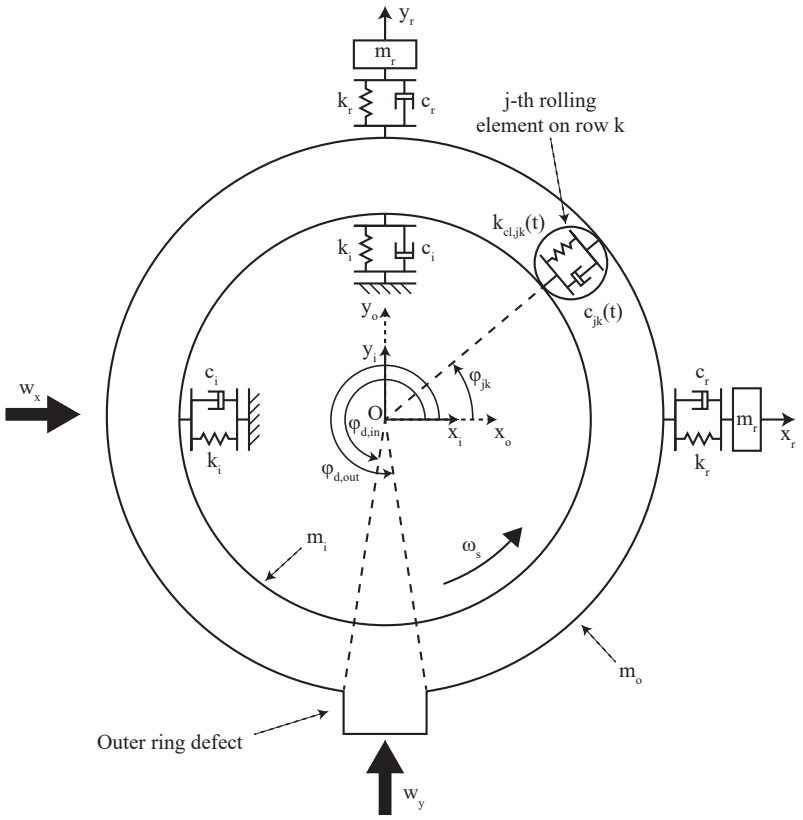

Figure 29: Scheme of the LP model.

each mass, namely x_i and y_i for m_i, x_o and y_o for m_o and x_r and y_r for m_r. The inner ring is connected to the ground through a spring-dashpot system represented by the stiffness k_i and the damping c_i. Similarly, m_r is connected to m_o through a spring with stiffness k_r and damping c_r. In both cases stiffness and damping are assumed to be equal in both x and y direction.

Moreover, each ball is replaced by a nonlinear spring and damper. The force exchanged by a ball with the raceways depends on its position with respect to the loaded area, the relative displacement of the races and the defect position. It is assumed that the cage is infinitely rigid, so that the rolling elements are kinematically constrained to maintain the same distance among them as they

roll. Therefore, the position of the j-th ball on the k-th row may be described as

$$\varphi_{jk}(t) = \varphi_0 + \frac{2\pi(j-1)}{n_b} + \frac{\pi(k-1)}{n_b} + \omega_c t \tag{32}$$

in which φ_0 is the cage initial angular position, while ω_c is the cage speed that is computed as

$$\omega_c = \frac{\omega_s}{2}\left(1 - \frac{r_b}{r_p}\cos\alpha\right) \tag{33}$$

where ω_s is the shaft rotation speed, r_p is the pitch diameter and α is the contact angle. The rolling element in the angular position $\varphi_{jk}(t)$ exerts a force only when under compression. Therefore, the variable $\delta_{jk}(t)$ is introduced to assess the contact deformation of each ball over time. It is defined as:

$$\delta_{jk}(t) = (x_i - x_o)\cos\varphi_{jk}(t) + (y_i - y_o)\sin\varphi_{jk}(t) - h \tag{34}$$

where $(x_i - x_o)$ and $(y_i - y_o)$ are the relative displacements of the inner ring with respect to the outer ring in the horizontal and vertical direction, respectively, while parameter h is the radial clearance. By employing the definition of the contact deformation, the contact force may be written as:

$$f_{c,jk}(t) = \begin{cases} k_c\delta_{jk}^n(t) & \text{if } \delta_{jk}(t) \geqslant 0 \\ 0 & \text{if } \delta_{jk}(t) < 0 \end{cases} \tag{35}$$

where k_c is the contact stiffness and exponent n is equal to to 1.5 for a ball-raceway contact [6]. Similarly, by introducing a contact damping coefficient c, the associated damping force may be defined as:

$$f_{d,jk}(t) = \begin{cases} c\dot{\delta}_{jk}(t) & \text{if } \delta_{jk}(t) \geqslant 0 \\ 0 & \text{if } \delta_{jk}(t) < 0 \end{cases} \tag{36}$$

Coefficient c accounts for the the effect of the lubricant film between the rolling elements and raceways [101]. Its value is commonly set proportional to the linearised stiffness $k_{cl,jk}(t)$ [98], which is defined on the of basis of Eq. 35. In fact, the equation for $\delta_{jk}(t) \geqslant 0$ may be rewritten as

$$f_{c,jk}(t) = k_c\delta_{jk}^n(t) = k_{cl,jk}(t)\delta_{jk}(t) \tag{37}$$

which means that

$$k_{cl,jk}(t) = k_c \delta_{jk}^{n-1}(t) \tag{38}$$

It has to be noted that $k_{cl,jk}(t)$ has the dimension of a stiffness, i.e. N/m, as opposed to the contact stiffness which is measured in $N/m^{1.5}$ for a ball-raceway contact. Due to the above considerations, contact damping $c(t)$ is set proportional to $k_{cl}(t)$ through coefficient β_c. To achieve an enhanced control on this parameter, the contact damping is herein defined by means of Rayleigh damping so that:

$$c_{jk}(t) = \alpha_c m_b + \beta_c k_{cl,jk}(t) \tag{39}$$

where m_b is the ball mass. This formulation permits to obtain a nonlinear relationship between frequency and damping ratio. Moreover, the choice is further justified by noticing that Eq. 39 is usually considered a good representation of the viscous damping [229], as it is the case with the greased contact between rolling elements and raceways.

Furthermore, for the system under exam, mass matrix **M**, damping matrix **C** and stiffness matrix **K** are defined according to the disposition of the components in Figure 29:

$$\mathbf{M} = \begin{bmatrix} m_i & 0 & 0 & 0 & 0 & 0 \\ 0 & m_i & 0 & 0 & 0 & 0 \\ 0 & 0 & m_o & 0 & 0 & 0 \\ 0 & 0 & 0 & m_o & 0 & 0 \\ 0 & 0 & 0 & 0 & m_r & 0 \\ 0 & 0 & 0 & 0 & 0 & m_r \end{bmatrix} \tag{40}$$

$$\mathbf{C} = \begin{bmatrix} c_i & 0 & 0 & 0 & 0 & 0 \\ 0 & c_i & 0 & 0 & 0 & 0 \\ 0 & 0 & c_r & 0 & -c_r & 0 \\ 0 & 0 & 0 & c_r & 0 & -c_r \\ 0 & 0 & -c_r & 0 & c_r & 0 \\ 0 & 0 & 0 & -c_r & 0 & c_r \end{bmatrix} \tag{41}$$

$$\mathbf{K} = \begin{bmatrix} k_i & 0 & 0 & 0 & 0 & 0 \\ 0 & k_i & 0 & 0 & 0 & 0 \\ 0 & 0 & k_r & 0 & -k_r & 0 \\ 0 & 0 & 0 & k_r & 0 & -k_r \\ 0 & 0 & -k_r & 0 & k_r & 0 \\ 0 & 0 & 0 & -k_r & 0 & k_r \end{bmatrix} \tag{42}$$

Subsequently, by grouping all the degrees of freedom into the vector $\mathbf{x}(t)$, so that:

$$\mathbf{x}(t) = \begin{Bmatrix} x_i(t) & y_i(t) & x_o(t) & y_o(t) & x_r(t) & y_r(t) \end{Bmatrix}^T \tag{43}$$

the equations describing the behaviour of the system may be written as

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) + \mathbf{f_c}(t) + \mathbf{f_d}(t) = \mathbf{w} \tag{44}$$

where $\mathbf{f_c}(t)$ is the contact force vector, i.e.

$$\mathbf{f_c}(t) = \sum_{k=1}^{n_r} \sum_{j=1}^{n_b} f_{c,jk}(t) \begin{Bmatrix} \cos\varphi_{jk}(t) \\ \sin\varphi_{jk}(t) \\ -\cos\varphi_{jk}(t) \\ -\sin\varphi_{jk}(t) \\ 0 \\ 0 \end{Bmatrix} \tag{45}$$

$\mathbf{f_d}(t)$ is the contact damping force vector,

$$\mathbf{f_d}(t) = \sum_{k=1}^{n_r} \sum_{j=1}^{n_b} f_{d,jk}(t) \begin{Bmatrix} \cos\varphi_{jk}(t) \\ \sin\varphi_{jk}(t) \\ -\cos\varphi_{jk}(t) \\ -\sin\varphi_{jk}(t) \\ 0 \\ 0 \end{Bmatrix} \tag{46}$$

and **w** contains the external forces applied to the system, that is, in the case of an externally applied load:

$$\mathbf{w} = \begin{Bmatrix} 0 & 0 & w_x & w_y & 0 & 0 \end{Bmatrix}^T \tag{47}$$

where w_x and w_y are the magnitudes of the external loads in the x and y directions, respectively. The damping matrix terms in Eq. 41 are set proportional to the respective stiffness terms by means of coefficients β_i and β_r, i.e. $c_i = \beta_i k_i$ and $c_r = \beta_r k_r$. These cases would be equivalent to Eq. 39 if α_c was set to 0. In this scenario, the damping ratio varies linearly with the frequency. Differently from the contact damping, in fact, these are both cases of structural damping, which is well represented by a linear law.

4.2.2 *Faulty bearing model*

The influence of the defect is introduced by adding the term $d(\varphi_{jk}(t))$ in Eq. 34, which becomes:

$$\delta_{jk}(t) = (x_i - x_o)\cos\varphi_{jk}(t) + (y_i - y_o)\sin\varphi_{jk}(t) - d(\varphi_{jk}(t)) - h \tag{48}$$

This parameter is defined on the basis of the formulation provided by Y. Jiang et al. [22] for a deep-groove ball bearing and exploited by S. Gao et al. [99] for an angular contact ball bearing. In this work, their geometrical definition is applied to a self-aligning ball bearing. In this regard, the outer ring of these components consists in a spherical raceway. Therefore, the race has the same curvature radius r_o in both the circumferential and in the axial direction.

The schematic representation of a rectangular defect centred at an angular location φ_d on one row is depicted in Figure 30. The entry and exit edges, namely $\varphi_{d,in}$ and $\varphi_{d,out}$, respectively, are located at the following angular positions:

$$\varphi_{d,in} = \varphi_d - \frac{\Delta\varphi_d}{2} \tag{49}$$

$$\varphi_{d,out} = \varphi_d + \frac{\Delta\varphi_d}{2} \tag{50}$$

where $\Delta\varphi_d$ is the angular width of the defect. The additional radial displacement due to fault presence of the j-th ball on the k-th row, namely $d(\varphi_{jk}(t))$, may be observed on plane Oyz, as shown in Figure 30.a. However, because the defect is shaped so that the ball only contacts the entry and exit edges, the relationship between the travelled path and the bearing geometry is evaluated on another plane, denoted in Figure 30.b as O'x'y'. Reference frame O'x'y'z' is obtained by rotating Oxyz around its x-axis by an angle equal to the contact angle α. Its value, which actually depends on the applied load [52], is assumed to remain unaltered, i.e. its variation over time is hypothesized to be negligible. Therefore, the axes x and x' are coincident, whilst y and y' are not. The ball displacement on plane O'x'y' is defined by $d_\alpha(\varphi_{jk}(t))$ and it is equal to:

$$d_\alpha(\varphi_{jk}(t)) = \begin{cases} r_b + r_o\cos(\varphi_{jk}(t) - \varphi_{d,in}) - r_o + \\ -\sqrt{r_b^2 - (r_o\sin(\varphi_{jk}(t) - \varphi_{d,in}))^2} & \text{if } \varphi_{d,in} \leqslant \varphi_{jk}(t) \leqslant \varphi_d \\ r_b + r_o\cos(\varphi_{d,out} - \varphi_{jk}(t)) - r_o + \\ -\sqrt{r_b^2 - (r_o\sin(\varphi_{d,out} - \varphi_{jk}(t)))^2} & \text{if } \varphi_d < \varphi_{jk}(t) \leqslant \varphi_{d,out} \\ 0 & \text{otherwise} \end{cases} \tag{51}$$

This formula allows to define the displacement on a radial plane as

$$d(\varphi_{jk}(t)) = d_\alpha(\varphi_{jk}(t))\cos\alpha \tag{52}$$

This additional deflection permits to take into account the effective compression of the rolling elements in the fault area.

4.2.3 *Procedure to determine model parameters*

The numerical model described in 4.2.1 and 4.2.2 requires the estimation of a remarkable number of parameters. The majority of them may be measured with sufficient accuracy, such as the geometrical data and the mass of the components. Nevertheless, a number of parameters are intrinsically uncertain and their accurate estimation is extremely challenging. This observation commonly leads authors to tune these parameters on the basis of their experience or on practical considerations. In light of this observation, the methods detailed in

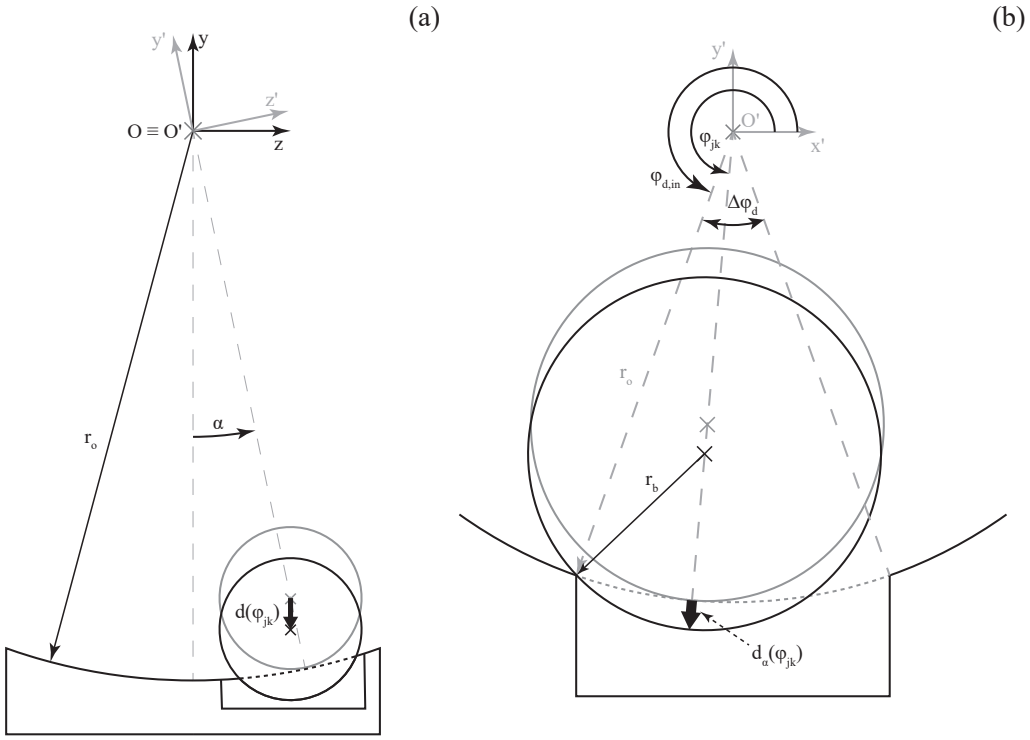

Figure 30: Defect geometry and additional displacement due to outer ring fault. The gray circle indicates the theoretical ball position in absence of the defect, while the black circle accounts for its presence. Two views are shown: (a) Oyz plane, derived from the global reference frame Oxyz; (b) O'x'y' plane, obtained by rotating reference frame Oxyz by an angle α around its x-axis so generate a rotated frame O'x'y'z'.

this chapter aims at describing a procedure devoted to reduce the degree of uncertainty of these parameters. Albeit the precise estimation of such quantities may remain arduous, it is possible to provide guidelines for their choice in absence of specific references. In fact, reasonable limits may be imposed on each parameter on the basis of analytical relationships related to the physics of the problem.

The proposed technique consists in minimizing a multi-objective function in the form

$$\theta = f(\gamma) \tag{53}$$

where θ is a vector containing the objective functions to minimize. These are global indicators that describe the discrepancy between numerical and experimental results for a number of test conditions. Vector γ, instead, comprises the unknown parameters of the model, which are constrained to assume a value comprised within reasonable limits based on the nature of each parameter. The solution of the multi-objective optimization problem consists in a number of points in the objective space so that further decreasing the value of one objective function does not decrease the others, i.e. it is not possible to further improve all objectives at the same time. These points should converge to the Pareto front of the problem [230].

The definition of a global indicator is described by considering a generic signal feature V. The associated numerical and experimental features are indicated as V_{num} and V_{exp}, respectively. Since the value of V depends on the test conditions, in an experimental study these features may be evaluated n_{tests} times, where n_{tests} is the number of test conditions. Additionally, each one may be repeated n_{rep} times, where n_{rep} is the number of times that a single test condition is replicated. Therefore, the average experimental value of the feature V_{exp} for the k-th test condition, i.e. $\overline{V}_{exp,k}$, is:

$$\overline{V}_{exp,k} = \sum_{p=1}^{n_{rep}} V_{exp,kp} \tag{54}$$

where subscript p refers to the p-th repetition of the test. Consequently, the percentage difference between the numerical and experimental feature for the k-th condition is:

$$\Delta V_k(\gamma) = \left| \frac{V_{num,k}(\gamma) - \overline{V}_{exp,k}}{\overline{V}_{exp,k}} \right| \cdot 100 \tag{55}$$

where the k-th numerical feature $V_{num,k}$ depends on the unknowns vector $\boldsymbol{\gamma}$. Finally, the global indicator $\Delta V(\boldsymbol{\gamma})$ is the average value which takes into account all test conditions:

$$\Delta V(\boldsymbol{\gamma}) = \frac{1}{n_{tests}} \sum_{k=1}^{n_{tests}} \Delta V_k(\boldsymbol{\gamma}), \tag{56}$$

This definition implies that n_{tests} simulations must be run for each set of $\boldsymbol{\gamma}$ in order to compute $\Delta V(\boldsymbol{\gamma})$. The value of n_{tests} is influenced by the number of test conditions and the number of test repetitions, i.e. n_{rep}. Attention should be paid to the choice of both quantities. Concerning the number of operative conditions, performing more tests translates to a higher number of conditions to be matched by the optimization algorithm. This may extend the applicability of the results but also leads to a higher computational effort. On the other hand, n_{rep} depends on the number of defective bearings with the same defect widths. The generation of these test samples is usually performed through EDM. Therefore, the time and costs associated with this operation must also be taken into account. Finally, each defect should be measured, since the nominal dimension generated through EDM may be not accurate enough, especially for lower defect dimensions. This further step leads to additional complexities in the process. Due to these observations, a total of 18 operative conditions and three repetitions of each test have been selected in this manuscript to provide an example of application of the proposed procedure. The experimental test setup has been thoroughly discussed in Chapter 2.

In general, any number of function objectives and unknown parameters may be chosen, depending on the requirements of the specific problem. Within the context of rolling element bearings, it is possible to define two objective functions which represent, in different ways, the energy content of the signal. The first feature is the quadratic sum of the first five defect harmonics, namely A^*, and it is defined as:

$$A^* = \sqrt{\sum_{i=1}^{5} A_{BPFO,k}^2} \tag{57}$$

where $A_{BPFO,k}$ is the amplitude of the k-th harmonic of the BPFO for a given value of the rotation frequency. This feature represents the energy content

of the low frequency deterministic phenomena related to the rotation of the components. The other feature is the RMS value of the signal, which is defined by employing the classical definition:

$$RMS = \sqrt{\frac{1}{N}\sum_{k=0}^{N-1} Z(k)^2} \tag{58}$$

where N is the length of the signal and $Z(k)$ is the k-th element of signal Z. This feature accounts for the energy content of the entire signal. Eq. 57 and Eq. 58 are employed to determine the global indicators for all test conditions as detailed in Eq. 56. In the following, they are referred to as ΔA^* and ΔRMS, respectively.

Concerning the unknown parameters, it is possible to define a subset of model parameters whose estimation is characterized by a high degree of uncertainty. Among those, mass m_r and stiffness k_r of the additional resonant mass connected to m_o depend on the properties of the system under exam. Although their values should be tuned in order to model a high natural frequency of the system, no further information on their choice is usually provided in the literature. Nonetheless, it is crucial to properly set their values, as they strongly concur in defining the dynamic response of the system. Uncertainties are also associated with the estimation of the values of α_c and β_c, which are needed to define the contact damping. This parameter is typically inserted as a constant value on the basis of the experience of the authors. Therefore, it may be challenging to estimate the most suitable value for the the case under study in absence of further information. Likewise, the coefficients of the damping matrix c_i and c_r are also set on the basis of practical considerations and are subjected to similar uncertainties. Finally, the nominal value of the radial clearance h is also not known a priori, since it might assume a range of different values even for nominally identical bearings. On the other hand, the estimation of the geometrical data and the mass of the components is generally attainable with sufficient accuracy. Consequently, they are considered as constant values in the proposed procedure. For each unknown parameter, it is possible to define a range of admissible values based on the physical properties of each one. The parameter values that minimize $f(\boldsymbol{\gamma})$, then, have to be sought within these intervals. The estimation of the values which delimit these ranges is exhaustively discussed in the next section.

4.2.4 *Definition of the constraints*

Constraining equations must be supplied for each parameter value inserted into the optimization procedure. Each parameter, in fact, may not take an arbitrary value but it may only assume a value comprised between reasonable limits which depend on the nature of each one. This means that the generic parameter ν must uphold the condition $\nu_{min} \leqslant \nu \leqslant \nu_{max}$, where ν_{min} and ν_{max} are the extreme values of the chosen interval. In the problem described in Section 4.2.3, the parameters m_r, k_r, α_c, β_c, c_i, c_r and h must all respect this condition. Apart from this basic requirement, the interesting aspect is given by the choice to include additional laws linking the unknown parameters. Such relationships may be defined on the basis of the physical model and they allow to drive the optimization avoiding unrealistic computations of the parameters.

Concerning m_r and k_r, their value should be set to model a resonance taking place in a high frequency range of the spectrum. This frequency range is not known a priori, since it depends on the characteristics of the system under examination. Therefore, it should be investigated by analysing the spectrum obtained via experimental testing. In the employed model, the stiffness k_r is assumed to be equal in both the x and y directions. As a result, there are two resonance frequencies of the mass m_r which depend on the bearing stiffness in these two directions. As a first approximation, in fact, the natural frequencies of the model may be estimated by solving an eigenvalue problem [229]:

$$\left(\mathbf{M}^{-1}\mathbf{K_l} - \lambda\mathbf{I}\right)\boldsymbol{\psi} = \mathbf{0} \tag{59}$$

where $\mathbf{K_l}$ is the linearised stiffness matrix, $\mathbf{I}$ is the identity matrix and $\boldsymbol{\psi}$ is the mode shape vector. The linearised stiffness matrix around a given load is defined by [88]:

$$\mathbf{K_l} = \begin{bmatrix} k_{xx}+k_i & 0 & -k_{xx} & 0 & 0 & 0 \\ 0 & k_{yy}+k_i & 0 & -k_{yy} & 0 & 0 \\ -k_{xx} & 0 & k_{xx}+k_r & 0 & -k_r & 0 \\ 0 & -k_{yy} & 0 & k_{yy}+k_r & 0 & -k_r \\ 0 & 0 & -k_r & 0 & k_r & 0 \\ 0 & 0 & 0 & -k_r & 0 & k_r \end{bmatrix} \tag{60}$$

where the terms k_{xx} and k_{yy} are bearing stiffness terms evaluated at a given static load. By introducing the cross coupling terms k_{xy} and k_{yx}, the linearised bearing stiffness matrix $\mathbf{K_b}$ containing these terms is equal to:

$$\mathbf{K_b} = \begin{bmatrix} k_{xx} & k_{xy} \\ k_{yx} & k_{yy} \end{bmatrix} = nk_c \sum_{k=1}^{n_r} \sum_{j=1}^{n_b} \delta_{jk}^{n-1}(t) \begin{bmatrix} \cos^2 \varphi_{jk} & \cos \varphi_{jk} \sin \varphi_{jk} \\ \cos \varphi_{jk} \sin \varphi_{jk} & \sin^2 \varphi_{jk} \end{bmatrix} \quad (61)$$

It is important to note that the bearing stiffness depends on the applied load and on the cage angular position φ_0. However, the numerical analysis is simplified by considering an average value of the stiffness over one period. In fact, it has been shown in Chapter 3 that the stiffness variation introduced by considering different cage positions is smaller than the average value of the stiffness over the ball passing period by at least one order of magnitude. Lastly, the problem is further simplified by neglecting the cross-coupling terms, which are usually two to three order of magnitudes lower than the terms on the matrix diagonal [88].

The solution of the eigenvalue problem allows to determine the natural frequencies and the mode shapes. In this case, however, the goal is to find a relationship relating m_r, k_r and the associated natural frequencies. This allows to express the values of k_r which, for different values of m_r, would produce the requested natural frequency. Such a relationship may be obtained by expanding Eq. 59. The resulting relationship, then, may be expressed as:

$$k_r(m_r) = \frac{\xi(m_r)}{\chi(m_r)} \quad (62)$$

where $\xi(m_r)$ and $\chi(m_r)$ are two terms which linearly depend on m_r:

$$\xi(m_r) = \lambda m_r \left[\lambda^2 m_i m_o - \lambda(k_i m_o + k_b m_o) + k_i k_b\right] \quad (63)$$

$$\chi(m_r) = \lambda m_r(\lambda m_i - k_i - k_b) + \Lambda \quad (64)$$

where k_b may be equal either to k_{xx} or to k_{yy}. As previously mentioned, in fact, two resonance frequencies are associated with mass m_r. The two different possible values of k_b allows to take into account both scenarios. Lastly, coefficient Λ is constant for a given eigenvalue λ:

$$\Lambda = \lambda^2 m_i m_o - \lambda(k_i m_o + k_b m_i + k_b m_o) + k_b k_i. \quad (65)$$

Eq. 62 allows to define the constraining equations for parameters k_r and m_r. In fact, it is possible to assume that they should produce an high resonance frequency comprised between two extreme values $f_{r,min}$ and $f_{r,max}$ so that:

$$f_{r,min} \leqslant f_r \leqslant f_{r,max} \tag{66}$$

The limit values depend on the resonance frequencies that are observed experimentally. If m_r and k_r are also constrained, i.e.:

$$m_{r,min} \leqslant m_r \leqslant m_{r,max} \tag{67}$$
$$k_{r,min} \leqslant k_r \leqslant k_{r,max} \tag{68}$$

then the limiting frequencies $f_{r,min}$ and $f_{r,max}$ are also variable. Consequently, the admissible values of m_r and k_r are further related by the following formula:

$$\frac{\xi_{min}(m_r)}{\chi_{min}(m_r)} \leqslant k_r(m_r) \leqslant \frac{\xi_{max}(m_r)}{\chi_{max}(m_r)} \tag{69}$$

in which ξ_{min}, χ_{min}, ξ_{max}, χ_{max} are computed via Eq. 62 by imposing a natural frequency equal to $f_{r,min}$ for ξ_{min} and χ_{min} and equal to $f_{r,max}$ for ξ_{max} and χ_{max}. In other words, Eq. 69 imposes the same constraint reported in Eq. 66. This relationship must hold for both values of k_b, i.e. k_{xx} and k_{yy}.

Concerning the contact damping, this parameter is controlled by means of a proportional damping approach as stated in Eq. 39. Therefore, its value varies with the linearised contact stiffness $k_{cl,jk}(t)$ and depends on the constant coefficients α_c and β_c, which are unknown. The proportional damping approach allows to define the damping ratio over the whole frequency range under exam by defining two damping ratios at two different frequencies. The damping ratio ζ_n at the natural frequency f_n may be defined as:

$$\zeta_n = \frac{1}{2}\left(\frac{\alpha_c}{\omega_n} + \beta_c \omega_n\right) \tag{70}$$

in which $\omega_n = 2\pi f_n$. According to this definition, the values of α_c and β_c may be determined by imposing two values ζ_1 and ζ_2 at two frequencies f_{n1} and f_{n2}. These natural frequencies may be chosen by selecting two resonance frequencies observable in the experimental spectrum. This means that, for two

assigned frequency values, α_c and β_c are completely defined by the choice of ζ_1 and ζ_2. Therefore, the damping ratios are employed in place of α_c and β_c as unknown parameters in the optimization procedure. This choice is due to the increased physical sensitivity associated with the definition of ζ that allows to provide more reasonable estimate for the admissible intervals compared to α_c and β_c. The values of the damping ratios must be limited in two given intervals:

$$\zeta_{1,min} \leqslant \zeta_1 \leqslant \zeta_{1,max} \tag{71}$$

$$\zeta_{2,min} \leqslant \zeta_2 \leqslant \zeta_{2,max} \tag{72}$$

In addition, to prevent solutions with α_c lower than zero, the damping ratio values must respect the following condition:

$$\frac{\zeta_1}{\omega_1} - \frac{\zeta_2}{\omega_2} > 0 \tag{73}$$

Moreover, according to the literature [101], the coefficient β_c should be set so that is upholds the condition $\beta_{c,min} \leqslant \beta_c \leqslant \beta_{c,max}$, which might be expressed as:

$$\beta_{c,min} \leqslant \frac{2\left(\omega_2\zeta_2 - \omega_1\zeta_1\right)}{\omega_2^2 - \omega_1^2} \leqslant \beta_{c,max} \tag{74}$$

The inequalities in Eq. 73 and Eq. 74 are used in conjunction with Eq. 71 and Eq. 72 to define the complete set of constraints for the contact damping.

The damping coefficients of the damping matrix, namely c_i and c_r, are also defined by exploiting the damping ratios. However, differently from the contact damping, the coefficient which multiplies the mass is set to zero. Consequently, by defining $c_i = \beta_i k_i$ and $c_r = \beta_r k_r$ the associated damping ratios ζ_i and ζ_r become:

$$\zeta_i = \frac{\beta_i \omega_i}{2} \tag{75}$$

$$\zeta_r = \frac{\beta_r \omega_r}{2} \tag{76}$$

where ω_i and ω_r are natural frequencies selected on the basis of experimental observations. In particular, the former should represent a resonance of the

masses attached to the inner ring, while the latter is associated with an high frequency response of the system. Differently from the other parameters, it is sufficient to define the extreme values of these parameters as:

$$\zeta_{i,min} \leqslant \zeta_i \leqslant \zeta_{i,max} \tag{77}$$

$$\zeta_{r,min} \leqslant \zeta_r \leqslant \zeta_{r,max} \tag{78}$$

without the need to provide additional relationships by means of further constraints.

Finally, radial clearance h is another major source of uncertainties. Usually, the manufacturers provide typical intervals within which the nominal values should lie in an unloaded condition. Therefore, it might be assumed that h should be comprised between h_{min} and h_{max}, so that:

$$h_{min} \leqslant h \leqslant h_{max} \tag{79}$$

In conclusion, the system of constraining inequalities that has to be supplied in the optimization procedure is the following:

$$\begin{cases} m_{r,min} \leqslant m_r \leqslant m_{r,max} \\ k_{r,min} \leqslant k_r \leqslant k_{r,max} \\ \frac{\xi_{min}(m_r)}{\chi_{min}(m_r)} \leqslant k_r(m_r) \leqslant \frac{\xi_{max}(m_r)}{\chi_{max}(m_r)} \\ \zeta_{1,min} \leqslant \zeta_1 \leqslant \zeta_{1,max} \\ \zeta_{2,min} \leqslant \zeta_2 \leqslant \zeta_{2,max} \\ \frac{\zeta_1}{\omega_1} - \frac{\zeta_2}{\omega_2} > 0 \\ \beta_{c,min} \leqslant \frac{2(\omega_2\zeta_2 - \omega_1\zeta_1)}{\omega_2^2 - \omega_1^2} \leqslant \beta_{c,max} \\ \zeta_{i,min} \leqslant \zeta_i \leqslant \zeta_{i,max} \\ \zeta_{r,min} \leqslant \zeta_i \leqslant \zeta_{r,max} \\ h_{min} \leqslant h \leqslant h_{max} \end{cases} \tag{80}$$

The constraints in Eq. 80 may be also applied to bearing models involving a different number of DOFs. In fact, their number has no influence on the

constraints on the damping ratio. What may change, however, is the number of inequalities to be added. In fact, in the present work it has been assumed that stiffness k_i and k_r have the same values in both x and y direction. This assumption is not mandatory, and for other systems different values in the two directions may be imposed. On the other hand, the number of DOFs has an effect on the constraint regarding the resonance frequency defined by Eq. 69, especially for models that also include the z direction. In such a case, an additional term k_{zz} is added to the bearing stiffness matrix. Therefore, the relationship imposed by Eq. 69 must be also respected for this term in addition to k_{xx} and k_{yy}. In fact, Eq. 62 is still applicable for problems with higher number of DOFs under the hypothesis that the cross-coupling terms in the bearing stiffness matrix may be neglected. Finally, the optimization procedure may be also applied to models which assign five or six DOFs to the components, as the ones derived from the work of Gupta [79], e.g. [78, 80, 231], or based on other formulations [99, 232, 233]. However, these models also include rotational degrees of freedom. As a result, the solution of Eq. 69 may increase in complexity, as it is the number of inequalities needed for the damping ratios. In particular, the solution of Eq. 69 remains the same only if the cross-coupling terms among linear and rotational DOFs are still negligible. Self-aligning ball bearings are not allowed to carry any moment load. As a consequence, since a bearing of this kind is employed for the experimental tests decribed in Chapter 2, the study of rotational degrees of freedom is neglected in this study. Nonetheless, the proposed procedure does not lose generality and may be also applied to other bearing types as well as for models including a higher number of DOFs, provided that they are adapted accordingly.

4.3 RESULTS OF THE EXPERIMENTAL TESTS

The experimental test rig employed to acquire the data needed for the validation of the optimization procedure has been exhaustively detailed in Section 2.2. Furthermore, Section 2.3.1 described the shape of the defects seeded on the outer races and all test conditions for the stationary tests. Defect dimensions are reported in Table 2. Test parameters, on the other hand, are detailed in Table 13. For each defect width, six combinations of load and shaft speed values are tested. For each one, the characteristic frequency of the defect on

Table 13: Test combinations of operative conditions. Each combination of load and speed is tested for each defect width D1, D2 and D3 and for each replica of every fault. BPFO values refer to the theoretical defect frequency on the outer ring.

Load [N]	Rotation frequency [Hz]	BPFO [Hz]
1000	20	97.8
1000	30	146.7
1000	40	195.6
2000	20	97.8
2000	30	146.7
2000	40	195.6

the outer race, known as ball pass frequency of the outer ring (BPFO), assumes a value which depends on the shaft speed, viz. [234]:

$$\mathrm{BPFO} = \frac{n_b \omega_s}{4\pi}\left(1 - \frac{r_b}{r_p}\cos\alpha\right) \tag{81}$$

where r_p is the bearing pitch diameter. The BPFO values associated with each condition are also reported in Table 13.

Figure 31 shows the time signals for different combinations of values of radial load, shaft rotation frequency and defect width. As a consequence of the considerable amount of runs, only the time signals for a single damage sample are reported. Each signal is represented in a time window corresponding to one cage rotation. The influence of the defect may be clearly observed in each signal, albeit the waveform is greatly influenced by the test parameters. For instance, it is noteworthy to qualitatively compare the time signals for a rotation frequency equal to 40 Hz. In this case, it is possible to observe a drastic change of the waveform when increasing the applied load from 1000 N to 2000 N, especially for the defect dimensions D2 and D3.

In addition to the time signals, a representative spectrum is shown in Figure 32.a. The plot refers to a test conducted for a load value equal to 2000 N

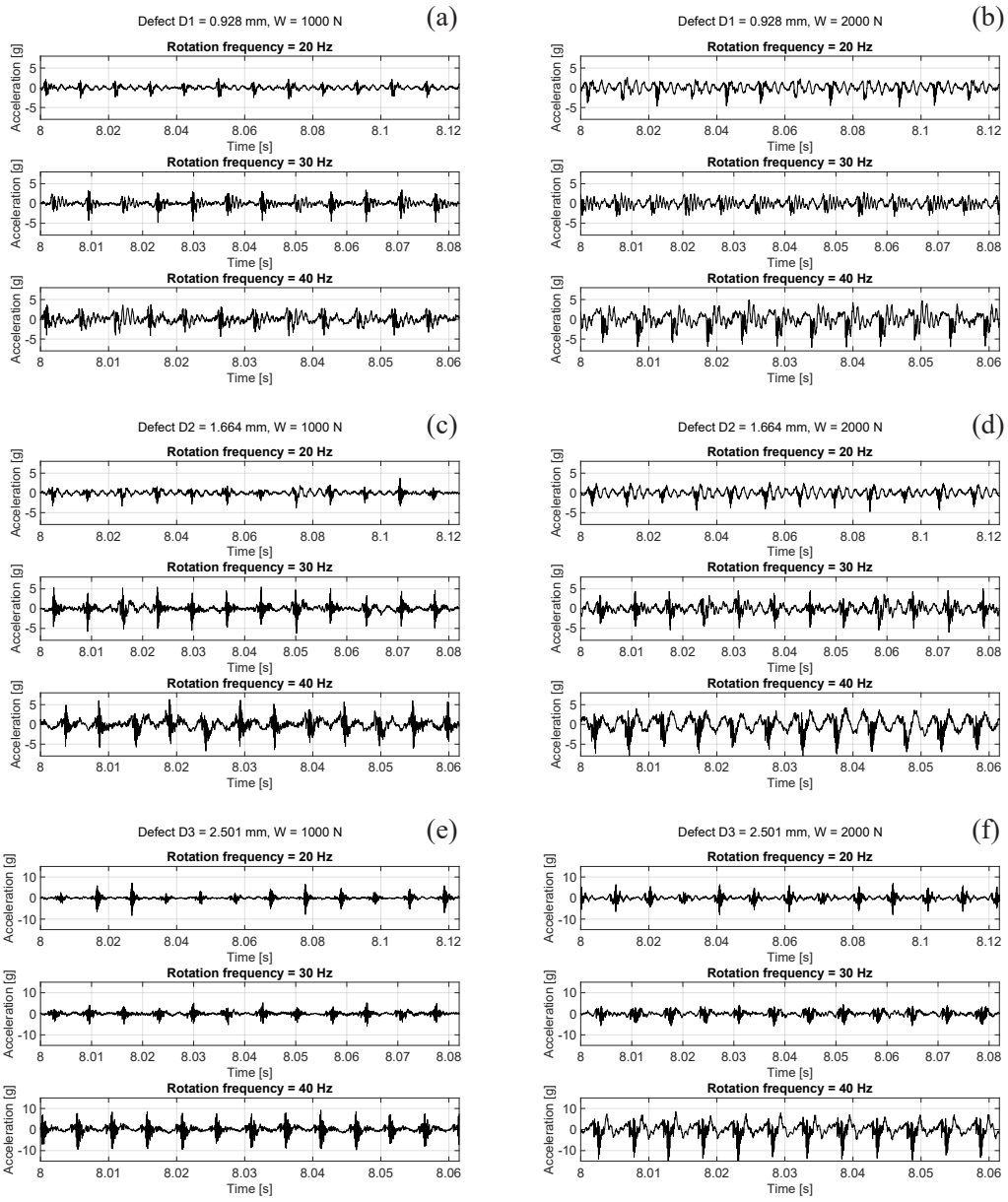

Figure 31: Experimental time signals for different combinations of defect dimension and load w: (a) D1, $w = 1000$ N; (b) D1, $w = 2000$ N; (c) D2, $w = 1000$ N; (d) D2, $w = 2000$ N; (e) D3, $w = 1000$ N; (f) D3, $w = 2000$ N.

and a rotation frequency value equal to 30 Hz for defect sample D2. The analysis of the frequency spectrum shows the presence of the BPFO and its harmonics, that is, 146.7 Hz and its integer multiples. In particular, the zoom in the low-frequency zone of the plot highlights the presence of the first two BPFO harmonics. Alongside them, Figure 32.a also permits to observe the harmonics of the shaft rotation frequency f_{rot}, which may be appreciated only below 300 Hz. On the other hand, BPFO harmonics are particularly noticeable in two frequency zones, i.e. around 400 Hz and 1500 Hz, as depicted in Figure 32.b. These are the two zones showing the highest acceleration magnitudes and might therefore be generated by possible resonances of the system. Notably, Figure 32.b shows that the resonances amplify the response up to the eleventh BPFO harmonic. In addition, Figure 32.c depicts another resonance phenomenon in the frequency interval comprised between 7000 and 9000 Hz. This peculiar behaviour is typically observed in this kind of systems [24, 88, 94].

Moreover, it is worth discussing the trends of the experimental features A^* and RMS introduced in Eq. 57 and Eq. 58, respectively. Figure 33.a, Figure 33.c and Figure 33.e report the RMS values for all the considered test conditions as a function of the rotation frequency. The result of each test has been represented by a cross marker, while their average value is illustrated by a square marker. From the RMS trends shown in Figure 33, it may be seen that increasing the applied static load value produces an increment of the RMS value for all conditions. In addition, increasing the rotation frequency produces the same effect, although with different trends among the defect dimensions. In fact, for the defect characterized by the largest width, i.e. D3, the RMS values are close to each other at 20 Hz and 30 Hz for both loads but they considerably increase at 40 Hz. Figure 33.b, Figures 33.d and Figure 33.f, instead, show the trends of A^*, for which similar conclusions may be drawn. However, the data appear to be more scattered around their average value, especially at a frequency value equal to 40 Hz. Nevertheless, the trends show that A^* consistently rises at 40 Hz compared to the other rotation frequency values. Moreover, its value lowers from 20 Hz to 30 Hz for both loads for defects D1 and D2, thus denoting a different behaviour from the RMS. The peculiar trends of this parameter may be due to the fact that the BPFO changes with the rotation speed. Therefore, the response at the BPFO and its first five harmonics may be influenced by different

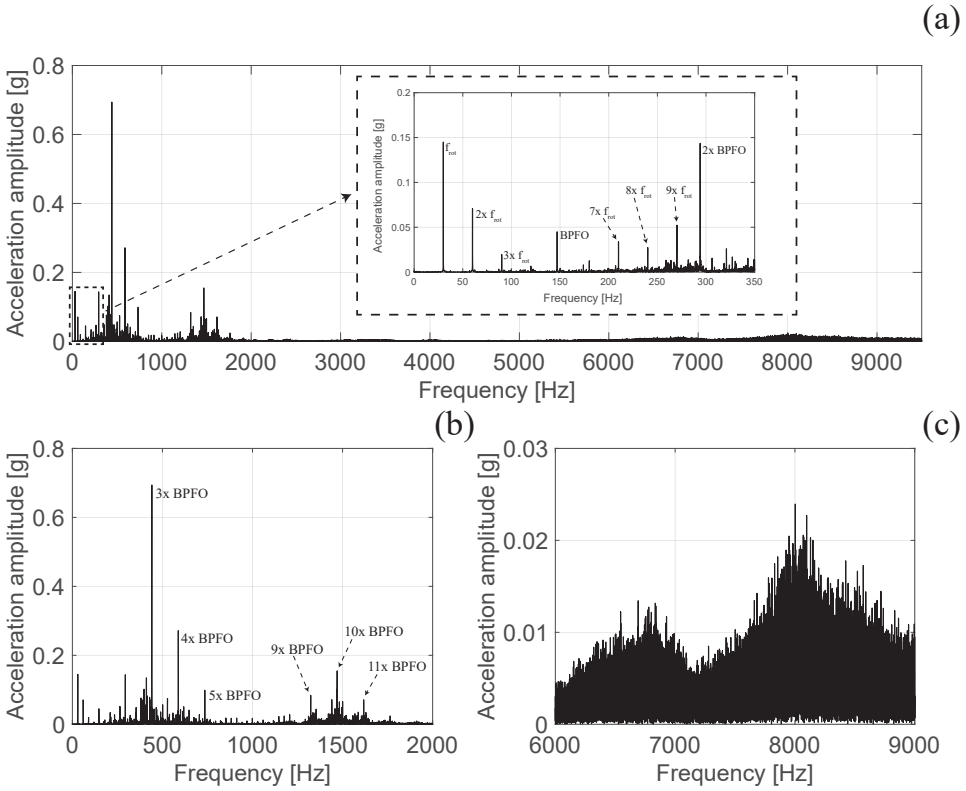

Figure 32: Experimental frequency spectrum for defect D2 under an applied load of 2000 N and a rotation speed of 30 Hz: (a) Spectrum over the frequency range 0 - 9500 Hz; (b) Spectrum in the "low" frequency range 0 - 2000 Hz; (c) Spectrum in the "high" frequency range 6000 - 9000 Hz.

resonances in correspondence of different values of the rotation frequency. As a result, A^* may not necessarily increase with the rotation speed of the shaft.

Additionally, Figure 34 depicts the same results but as a function of the defect width, as some additional considerations may be drawn from these graphs. Concerning the RMS, Figures 34.a, 34.c and 34.e show that increasing defect widths lead to larger RMS values. Besides, the trends of A^* in Figures 34.b, 34.d and 34.f indicate that this parameter decreases at 20 Hz for larger widths while it slightly increases at 30 Hz. The behaviour of this parameter changes

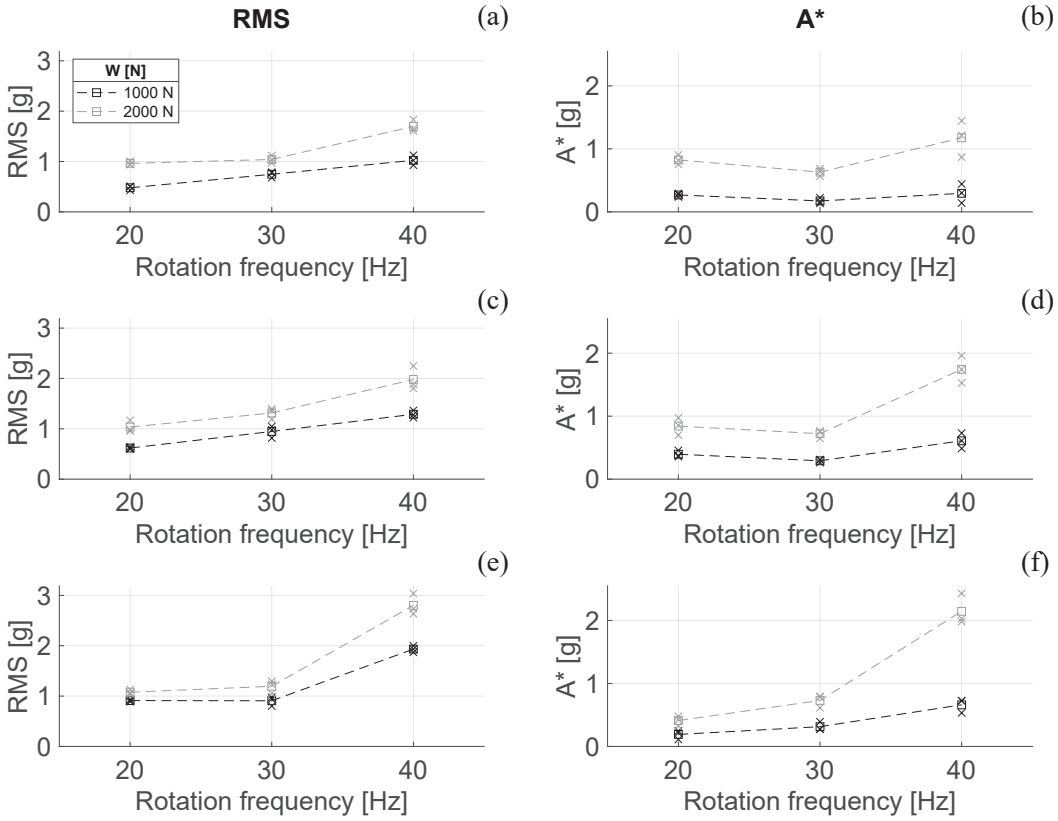

Figure 33: Signal features compared against the rotation frequency. For each rotation frequency, the three cross markers represent a different test run on another bearing with the same defect size. The square marker denotes the average value of the three tests. (a) RMS for defect D1; (b) A^* for defect D1; (c) RMS for defect D2; (d) A^* for defect D2; (e) RMS for defect D3; (f) A^* for defect D3.

drastically at 40 Hz, although the same considerations on data scattering cited for the previous plots may be also applied here.

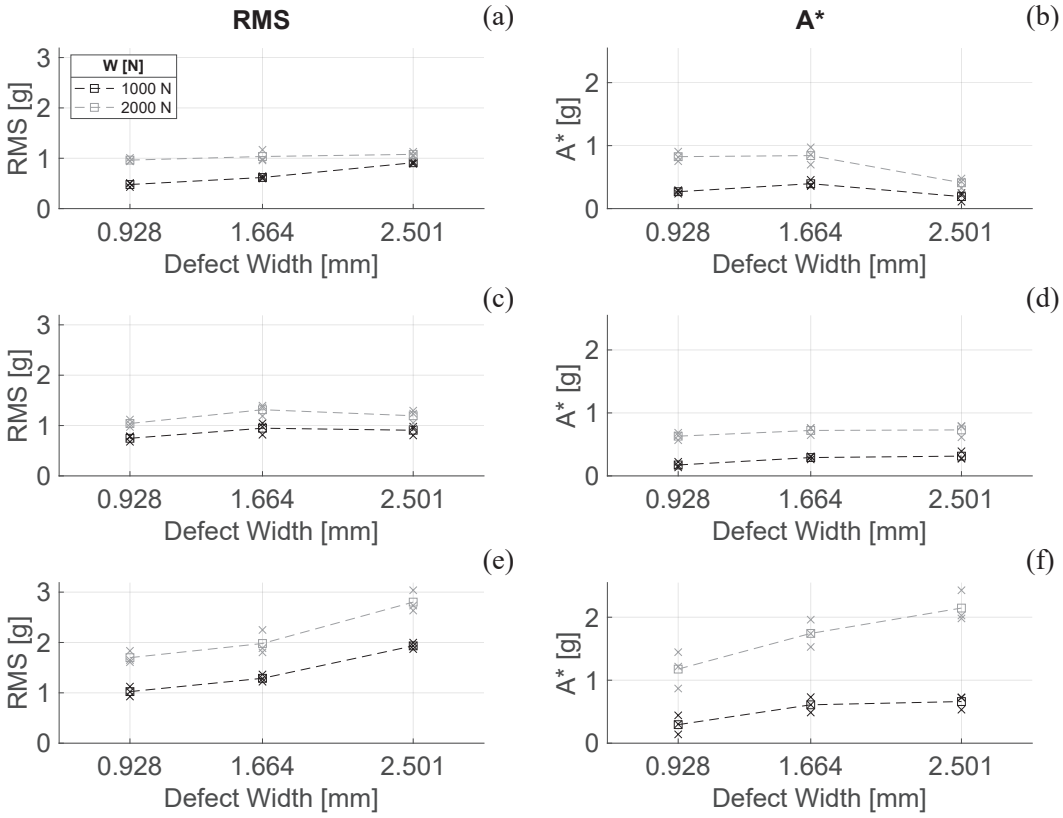

Figure 34: Signal features compared against the defect width. For each defect dimension, the three cross markers represent a different test run on another bearing with same defect size. The square marker denotes the average value of the three tests. (a) RMS at 20 Hz; (b) A^* at 20 Hz; (c) RMS at 30 Hz; (d) A^* at 30 Hz; (e) RMS at 40 Hz; (f) A^* at 40 Hz.

4.4 APPLICATION OF THE MULTI-OBJECTIVE OPTIMIZATION TECHNIQUE TO EXPERIMENTAL SIGNALS

The multi-objective optimization procedure described in Section 4.2.3 is applied in order to determine the unknown parameters in the described numerical model. Firstly, the analytical constraints introduced in Section 4.2.4 are detailed for the case under exam. Then, the results of the multi-objective optimization

are discussed. Time signals, frequency spectra and numerical features are reported to provide a comparison with the experimental results.

4.4.1 *Parameters of the model and definition of the constraints*

The known model parameters are reported in Table 14. The value of stiffness k_i is estimated by means of a FE model of the shaft in which the inner ring is mounted. In this case, in fact, classic beam theory should be avoided since the shaft is short and the support bearings occupy a large part of the shaft itself. The FE model consists of the shaft, meshed with 3D tetrahedral elements, and nonlinear springs replacing the support bearings, as depicted in Figure 35. Both bearing seats on the shaft are connected to a central node by rigid connectors. These nodes are connected to the ground through nonlinear springs whose stiffness is analytically computed via the method described in 3.4. The computational domain is halved by exploiting the system symmetry. Load is applied at the test bearing end by applying a sinuosoidal distribution which replicates the load distribution on the bearing. Nastran static nonlinear solution SOL 401 is employed to solve the model. Then, stiffness is evaluated for different loads by considering the ratio between the magnitude of the applied load and the average displacement of the nodes on the shaft end. The considered nonlinear load-deflection relationship means, however, that the stiffness of the shaft varies with the applied load. As a consequence, by introducing a nonlinear term for the shaft stiffness the complexity of the LP model would substantially increase. Despite this consideration, it is observed that the support bearings are considerably stiffer than the shaft, and the variation of stiffness in the operative range is not excessive. In fact, Table 15 reports that the stiffness has a very low dependence from the load, i.e. lower than 3% for a load ranging from 500 N to 3000 N. Thus, the resulting stiffness may be considered, with a good approximation, a constant value independent from the applied load. The value of the contact stiffness k_c, instead, is analytically computed by exploiting the Hertzian contact theory for the general case of two bodies with different principal relative radii of curvature in orthogonal planes [19], which has been discussed in Section 3.2.2. This value is independent from the applied load and only depends on the materials and the curvature radii of the bodies in contact.

Table 14: Model parameters.

Description	Symbol	Value
Inner ring and shaft mass	m_i	3.48 kg
Outer ring, casing and pedestal mass	m_o	5.79 kg
Ball mass	m_b	1.47 x 10^{-3} kg
Shaft stiffness in the x and y directions	k_i	75 MN/m
Contact stiffness	k_c	5927 MN/$m^{1.5}$

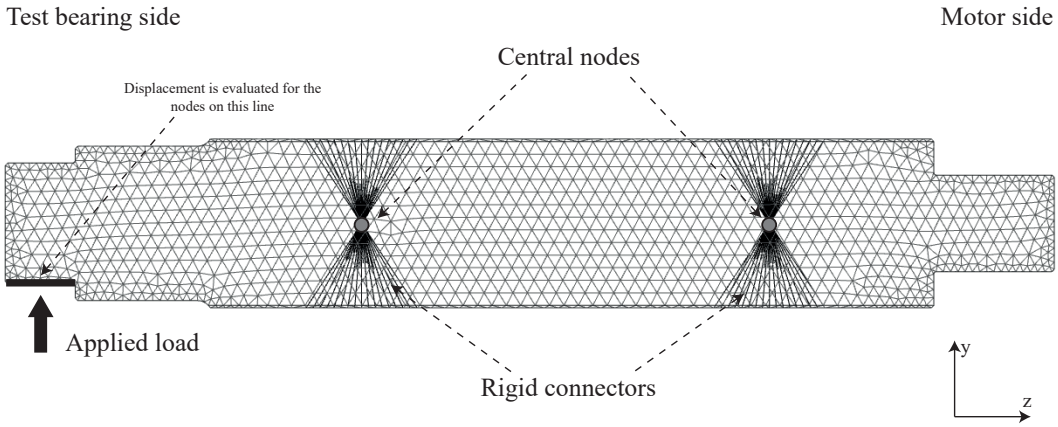

Figure 35: FE model of the shaft.

Table 15: Shaft stiffness estimated through FE analysis.

Load [N]	Shaft stiffness [MN/m]
500	73.72
1000	74.61
2000	75.45
3000	75.93

In order to proceed with the multi-objective optimization, it is first necessary to provide the intervals in which the unknown parameters are allowed to vary. The extreme values of these ranges are reported in Table 16 and they have been

Table 16: Minimum and maximum allowable values for each unknown parameter.

Parameter	Minimum value	Maximum value
m_r [kg]	0.1	5
k_r [GN/m]	0.1	10
ζ_1 [%]	0.1	15
ζ_2 [%]	0.1	15
ζ_i [%]	0.1	15
ζ_r [%]	0.1	15

chosen on the basis of practical considerations. The values of the damping ratios are all assumed to be variable between 0.1 % and 15 %. The values of m_r and k_r have both wide ranges in order to take into a account a variety of possible outcomes. Radial clearance h, on the other hand, has been assumed to be equal to zero in order to reduce the already remarkable computational burden.

To evaluate the further constraint on m_r and k_r, it is crucial to determine the bearing stiffness matrix terms k_{xx} and k_{yy} for the bearing under exam, which may be computed by employing Eq. 61. Bearing stiffness could also be determined by employing the FE procedure detailed in Chapter 3. However, for the purpose of this work, the stiffness must be evaluated for several angular positions of the cage. Therefore, the analytical formulation is utilized in place of the FE method in order to obtain a faster estimate. Figure 36 shows the variation of the stiffness terms for a healthy bearing over the ball pass period for two different applied radial loads. The stiffness change due to cage movement is well-known in the literature as varying compliance (VC) effect [12]. The VC effect leads to a variation of stiffness of the bearing depending on the position of the rolling elements even in a healthy condition. The periodicity is equal to the angular spacing between the balls on opposite rows, i.e. $360°/n_r \cdot n_b = 15°$ in the case under exam. The load values, which correspond to the ones applied on the experiments, increase the stiffness in all cases. Besides, in the considered scenario in which there is only a constant vertical load, the term k_{yy} remains mostly unaltered, as its deviation from its mean value is lower

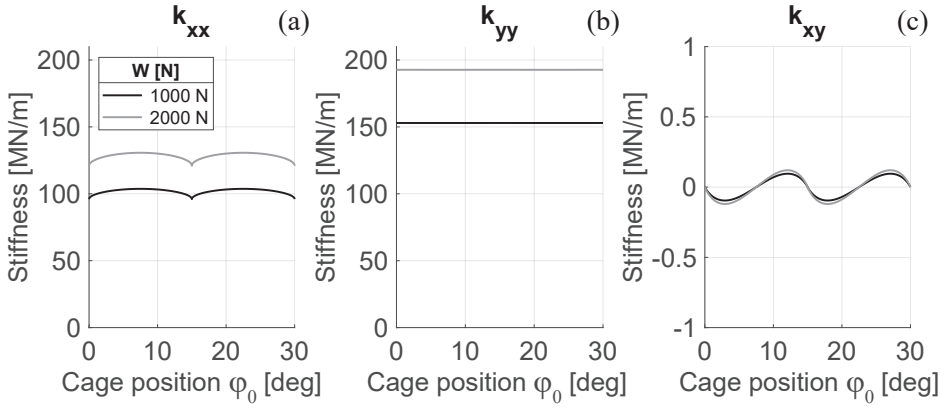

Figure 36: Bearing stiffness terms of an healthy bearing as a function of load and balls position: (a) k_{xx}; (b) k_{yy}; (c) k_{xy}.

than 0.02 %. On the other hand, k_{xx} shows a clear variation over a period, since its minimum and maximum values differ by 7.5 %. The low variation on k_{yy} term is due to the high number of balls which carry the load. These results agree with the stiffness estimates reported in [49, 88]. In addition, Figure 36.c shows one of the cross-coupling terms, k_{xy}. The other term, k_{yx}, is not shown since it assumes the same values. The plot denotes that, in agreement with the hypothesis in Section 4.2.4, the cross-coupling terms are several orders of magnitude lower than the terms on the main diagonal. Therefore, it is possible to neglect them to evaluate the linearised stiffness matrix. Moreover, Figure 37 depicts the stiffness coefficients but for a faulty bearing. Defect dimension D3 has been chosen to better visualize the introduced stiffness variation, but the same behavior may be observed for the D1 and D2 dimensions alike. When one ball enters the defect area, stiffness k_{yy} reduces while k_{xx} increases. For both loads the latter diminishes by 12.5 %.

For the purpose of estimating the natural frequencies of the system, the healthy bearing stiffness is considered. Therefore, the stiffness variations due to a fault are neglected in the study of the resonances. Besides, since k_{xx} is variable, a mean value over the ball pass period is selected in order to simplify the computations. Finally, since the value of k_{xy} is considerably smaller compared to the other terms, it will be neglected in further analyses. The stiffness matrix terms computed in this manner are reported in Table 17.

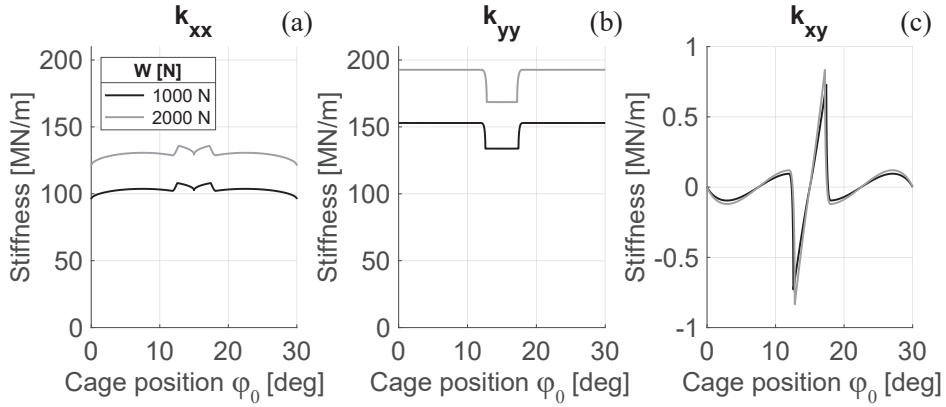

Figure 37: Bearing stiffness terms of a faulty bearing with defect dimensions D3 as a function of load and balls position: (a) k_{xx}; (b) k_{yy}; (c) k_{xy}.

Table 17: Healthy bearing stiffness matrix terms at 1000 N and 2000 N.

Load [N]	1000	2000
k_{xx} [MN/m]	102.3	128.8
k_{yy} [MN/m]	152.3	192.2

By solving the eigenvalue problem in Eq. 59 it is possible to estimate the effect of the variation of m_r and k_r on all the natural frequencies of the system within the limits reported in Table 16. Since the problem has 6 DOFs, there are six natural frequencies. Each one depends on the variable values of m_r and k_r and also on the load. As a result, two of them will pertain to the high frequency resonances of m_r. The numerical results show that the resonance values in the x and y directions are almost identical and that the influence of load on these natural frequencies is negligible. Therefore, the results are depicted in Figure 38 by only plotting one surface corresponding to a radial load value equal to 2000 N. The figure shows that the value of the frequency non-linearly depends on both m_r and k_r. Moreover, the frequency greatly increases when m_r diminishes, in particular for the highest stiffness values.

In addition, Figure 39 depicts the possible values of the other frequencies in the same ranges of mass and stiffness. It may be appreciated how different

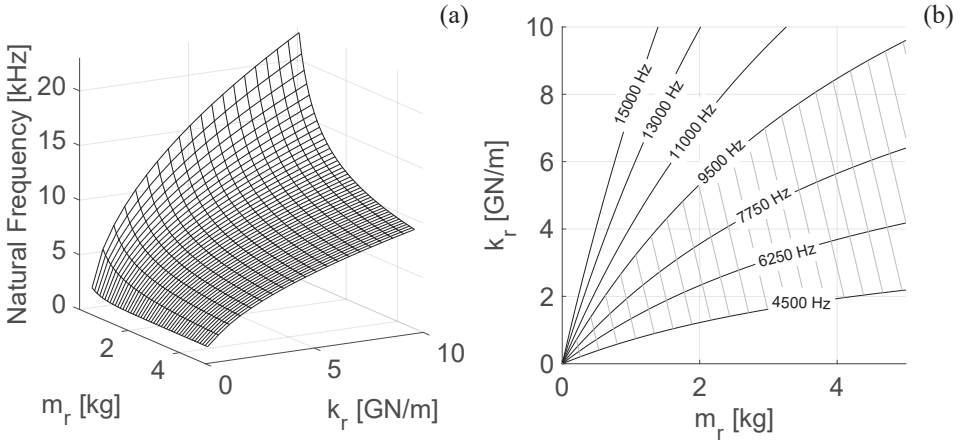

Figure 38: Effect of the values of m_r and k_r on the natural frequency of the high resonant mass for an healthy bearing at the applied static load of 2000 N: (a) Three-dimensional surface plot; (b) Relationship between m_r and k_r for different natural frequencies. The grey hatched area designates the mass and stiffness combinations that produce a resonance frequency comprised between 4500 Hz and 9500 Hz.

values of m_r influence the resonance frequencies, in particular for the first two modes, while modifying k_r does not produce noticeable differences. It is also worth comparing the results for the second and fourth mode to the experimental spectrum that is shown in Figure 32.b. the comparison that there are two resonance zones around 400 Hz and 1500 Hz, in agreement with the experimental observation.

Concerning ζ_1 and ζ_2, the additional constraints in Eq. 73 and Eq. 74 are computed by choosing two frequency values corresponding to the two damping ratios. Based on the experimental observations, the two frequencies f_{n1} and f_{n2} are set equal to 400 Hz and 1500 Hz, respectively. Moreover, the limit values of β_c are defined as $\beta_{c,min} = 0.25 \times 10^{-5}$ s and $\beta_{c,max} = 2.5 \times 10^{-5}$ s [98]. These conditions are plotted in Figure 40. The area denoted with the gray hatching in Figure 40.a represents the set of admissible solutions for Eq. 73 and Eq. 74. In fact, the solid straight lines represent the combinations of ζ_1 and ζ_2 that generate β_c values equal to $\beta_{c,min}$ and $\beta_{c,max}$, while the dash-dot line describes the values which produce $\alpha_c = 0$. Likewise, Figure 40.b presents

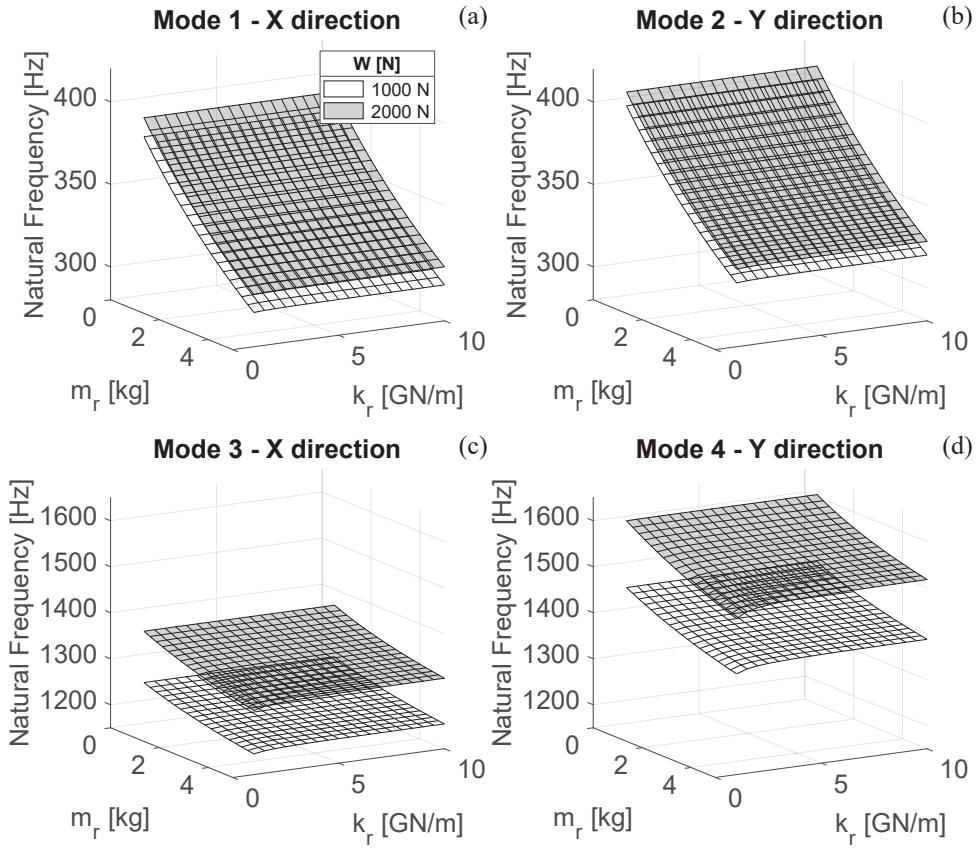

Figure 39: Effect of the values of m_r and k_r on the first four natural frequencies of the healthy bearing. The directions refer to the global reference frame Oxyz. (a) Mode 1; (b) Mode 2; (c) Mode 3; (d) Mode 4.

the same conditions but on a α_c - β_c plane. This plot shows that, within the assumed limits of ζ, the admissible values of α_c increase in case that β_c decreases. Finally, the damping ratio ζ_i is associated with the natural frequency f_i = 400 Hz, whilst ζ_r refers to the natural frequency f_r = 8000 Hz.

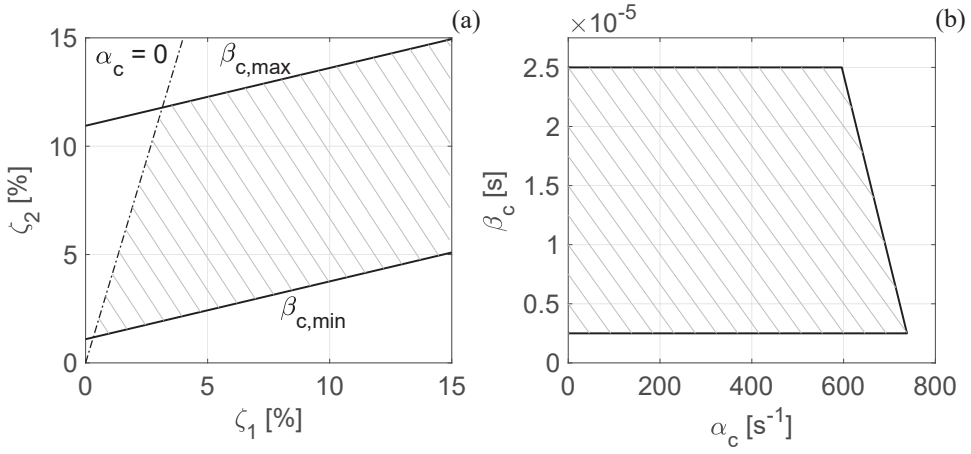

Figure 40: Admissible values of the coefficients related to the contact damping: (a) ζ_1 and ζ_2; (b) α_c and β_c.

4.4.2 *Model parameters estimated through multi-objective optimization*

The problem described in Section 4.2.3 is solved by employing an optimization procedure devoted to the minimization of the vector θ introduced in Eq. 4.2.4. According to the experimental campaign described in Section 2.3.1, 18 different test conditions have been included in the problem. These correspond to two different radial load values, i.e. 1000 N and 2000 N, three rotation speed values, namely 20 Hz, 30 Hz and 40 Hz and three defect widths corresponding to dimensions D1, D2 and D3. Therefore, the numerical model is solved for each combination of these conditions during each iteration of the optimization procedure. The load value is applied on the y direction and is labelled as w in accordance with Figure 7. Therefore, the coefficients in Eq. 47 are $w_y = w$ and $w_x = 0$. The set of equations in Eq. 44 is solved in MATLAB environment [235] through the ordinary differential equation solver *ode45*, which is based on an explicit fourth order Runge-Kutta formula. A time step $\Delta t = 1.9531 \times 10^{-5}$ s is chosen for the numerical integration. For each parameter combination, 1 s of signal is simulated. Solution is obtained in the time domain. The employed time length allows to obtain a regime solution that permits to compute the RMS of the stationary signal after removing the initial transitory. Then, feature A^* is calculated after transforming the signal in the frequency domain. By knowing RMS and A^*, it is possible to compute the objective functions ΔRMS

and ΔA^* for each operative condition by exploiting Eq. 56. Simulations were run on a desktop workstation with 32 GB RAM and an I7 processor. Each condition took about 2.3 s of simulation. As a consequence, it took about 42 s to compute the features for all the operative conditions and defect dimensions. The algorithm needed 152 hours to find the set of optimal parameters.

Figure 41 represents the values of the objective functions of vector θ which lies on the estimated Pareto front. Each point corresponds to a different combination of parameters of the vector γ which complies to the constraints defined in Section 4.2.4. The main consideration that may be done on the results regards the different order of magnitude for the variation of the objectives. In fact, while ΔA^* varies approximately between 35.83 % and 35.9 %, ΔRMS ranges between 17.9 % and 21.3 %. From the graph it is possible to notice that the sets of unknown parameters found by the algorithm greatly affects the RMS value but are unable to further reduce the value of ΔA^* below 35.83 %. The values of ΔA^* constitute a consequence of the high number of operative conditions considered in the experimental campaign, since the global accuracy of the model tends to decrease as the range of working conditions is increased. It is worth noting, in fact, that to the best of author knowledge such a kind of validation across several different operative conditions and defect dimensions is not available in the literature. Closest examples may be found in the works of J. Liu [97] and S. Mufazzal et al. [106]. In the former, the author showed numerical RMS trends depending on fault length for different combinations of load and speed values , but experiments were performed on a single operative condition and for one defect dimension only. In the latter, the authors reported acceleration peaks for different defect dimensions against experimental data, but for only one operative condition. Therefore, within this context, the procedure described in this paper provides the basic framework to allow determining the unknown parameters in situations in which multiple operative conditions must be validated. The points on this peculiar type of front are known as weak Pareto optimals [236]. Differently from the general definition of Pareto optimal, which states that there are no points that simultaneously improve both objective functions, in a cluster of weak Pareto optimals there might exist points which improve one function while not changing the others. Since the variation of ΔRMS is considerably larger than the fluctuation of ΔA^*, it is reasonable to assume a situation similar

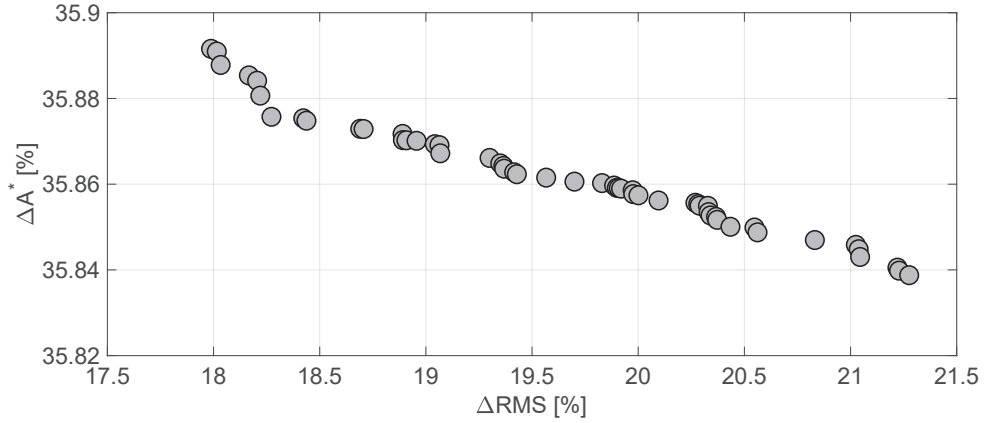

Figure 41: Values of the objective functions of vector θ which form the estimated weak Pareto optimal. Each point corresponds to a different set of parameters γ which solves the optimization problem by providing nondominated points in the feasible objectives space.

to the latter. This type of results, however, should not come as unexpected as this kind of problem is known to possibly produce weak Pareto optimal points [237].

In light of the above considerations, Figure 42 reports the values assumed by each unknown parameter of γ for each point on the Pareto front as a function of the only objective function ΔRMS. It might be reasonably assumed, in fact, that ΔA^* remains approximately constant for all these combinations of parameters. Therefore, in the sebsequent analyses the optimal set of parameters will be defined as the one producing the lowest value of ΔRMS. The associated parameter values are reported in Table 18 along with the associated derived parameters, i.e. the damping terms c_i and c_r, the proportional damping coefficients α_c and β_c and the natural frequency of the high resonant mass f_r.

The plotted trends allow to draw some conclusions concerning the values of the unknown parameters. Figures 42.a and 42.b shows that ζ_i and ζ_r have a low variability, as the former remains slightly below 15 % for all points while the latter is limited between 0.11 % and 0.51 %. Figure 42.c and Figure 42.d reports the possible values of ζ_1 and ζ_2. Based on them, the corresponding

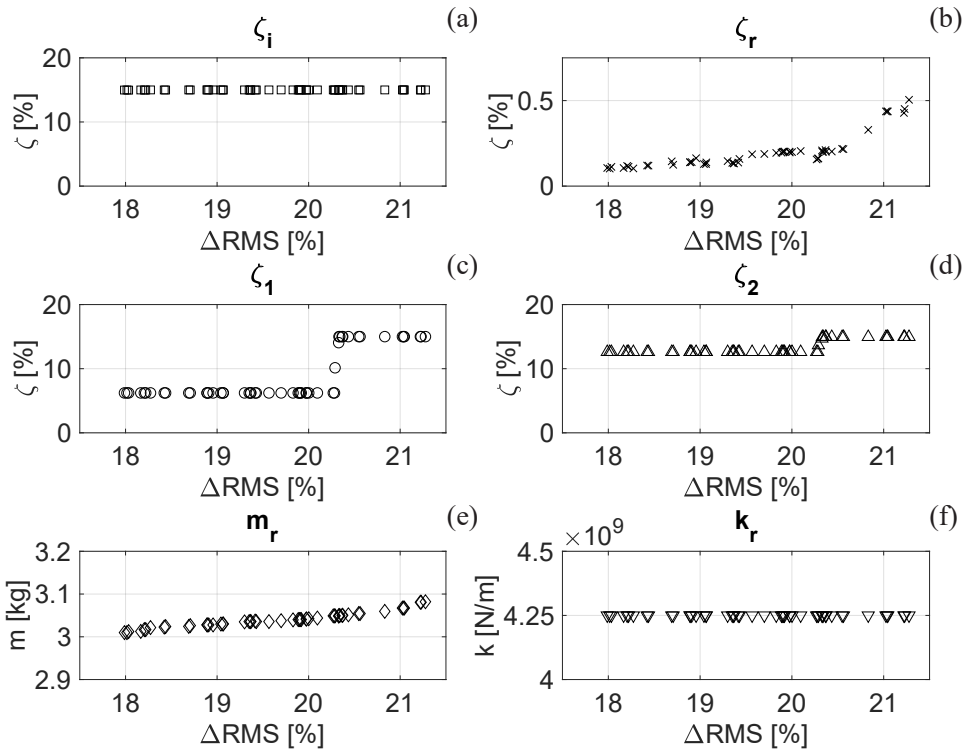

Figure 42: Representation of the unknown parameters estimated through the multi-objective optimization process. Each point represents a different value of the vector θ as a function of ΔRMS. Each plot shows one of the parameters: (a) ζ_i; (b) ζ_r; (c) ζ_1; (d) ζ_2; (e) m_r; (f) k_r.

values of α_c and β_c may be computed through Eq. 70. Interestingly, every combination of ζ_1 and ζ_2 produce a value β_c = 2.5 x 10^{-5}, which is the upper limit defined for this quantity as defined in Section 4.2.4. Coefficient α_c, on the other hand, assumes an appropriate value corresponding to the combination of the two damping ratios. Eventually, Figures 42.e and 42.f depict the values of m_r and k_r, respectively. The former varies between restricted limits, i.e. between 3.0 and 3.09 kg, while the latter assumes a constant value equal to 4.25 x 10^9. These combinations of values generate a number of possible natural frequencies, which are shown in Figure 43. Remarkably, all the estimated natural frequencies assume values approximately between 7.32 kHz and 7.4

Table 18: Combination of parameters that minimizes the value of ΔRMS.

Parameter	Value	Derived parameters	Value
ζ_i [%]	14.95	c_i [Ns/m]	2834.1
ζ_r [%]	0.13	c_r [Ns/m]	186.0
ζ_1 [%]	6.22	α_c [s^{-1}]	154.9
ζ_2 [%]	12.6	β_c [s]	2.5 x 10^{-5}
m_r [kg]	3.01	f_r [Hz]	7388
k_r [N/m]	4.25 x 10^9		

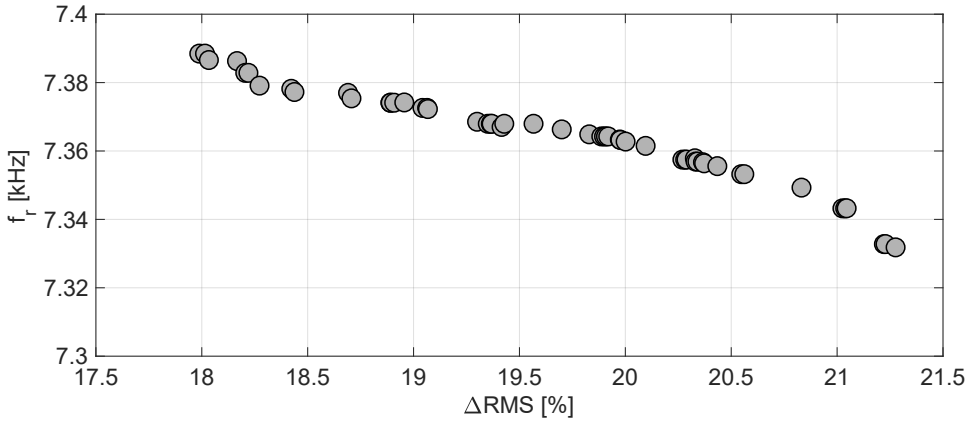

Figure 43: Estimated natural frequencies for each combination of m_r and k_r of each vector γ.

kHz, which are values very close to the resonance zone observed in the experimental tests. In conclusion, the value of ΔRMS is considerably influenced by ζ_r, ζ_1, ζ_2 and m_r. The value ΔA^*, on the contrary, is almost not affected for these combinations of parameter values.

As an additional observation, it is worth underlining that the results described in this section have been obtained by considering all defect dimensions and operative conditions at once. However, it is also possible to perform the optimization procedure by excluding some conditions. In such a case, a more accurate estimation for the case under exam would be obtained, but at the cost

of neglecting the other defect dimensions, leading to a reduction of the accuracy for the other cases. Furthermore, the computational time also decreases, as less conditions need to be matched by the algorithm. For instance, Figure 44 shows the Pareto fronts obtained by considering only a single defect dimension under the six available combinations of operative conditions. In this scenario, the values of ΔA^* and ΔRMS are computed only on those simulated conditions and not on all defect dimensions. From the plot, it is possible to observe the difference in the values of these quantities that may be obtained by considering only one defect at a time. This kind of analysis also leads to different estimates of the unknown parameters, as shown in Table 19. Interestingly, the value of ΔA^* is still considerably high for all defect dimensions, especially for defect D3. This result may be due to the extent of the largest fault considered in this work. In fact, S. Mufazzal. et al [106] noticed that the deviation between the numerical and experimental magnitude of the spectral peaks increases with increasing defect dimensions. Therefore, it is possible that the dimension of defect D3 is too large to produce comparable spectral components between numerical and experimental results, which are the quantities required for the estimation of A^*. Defect dimensions D1 and D2, on the other hand, generate comparable values of ΔA^*. Within this context, it is also interesting to note that Khanam et al. [82] reported that higher variation in the results for larger defect dimensions may be produced by other excitation mechanisms not considered in the model. These observations may justify the high values of ΔA^* that are obtained for defect D3 even when the optimization procedure is applied only on that defect dimension. Finally, it is also worth mentioning that, despite the differences observed on parameter ΔA^*, the value of ΔRMS are comparable in all three cases. To conclude, this example demonstrates that the optimization procedure may be applied on any number of combinations of conditions, but the resulting parameters will differ accordingly. In this work, the optimization procedure has been applied to all available combinations of test conditions and defect widths to show the potential of the procedure to be applied to a variety of operative conditions and defect dimensions at once.

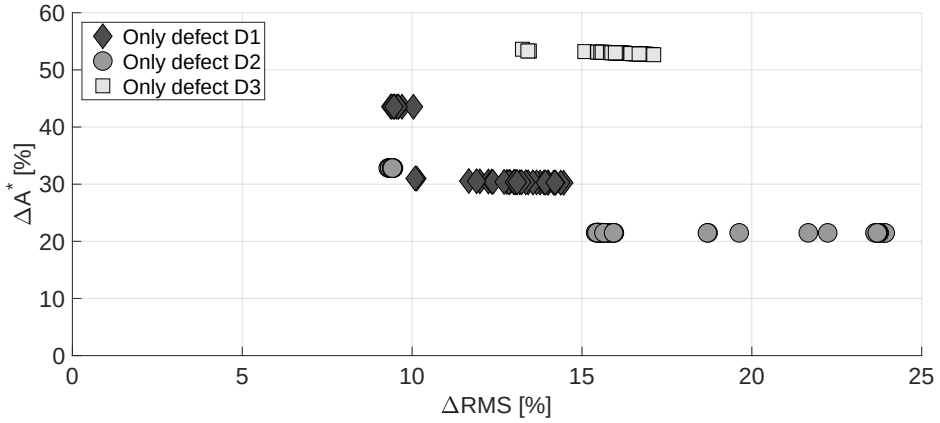

Figure 44: Pareto fronts obtained by performing the optimization procedure on one defect at a time.

Table 19: Combination of parameters obtained by considering one defect dimension at a time. The chosen sets are the ones which produce the points that minimize the distance from the axes origin in Figure 44.

Considered defect	ζ_i [%]	ζ_r [%]	ζ_1 [%]	ζ_2 [%]	m_r [kg]	k_r [N/m]
D1	13.79	15.00	6.22	12.60	2.73	3.97×10^9
D2	13.85	0.27	11.62	11.34	2.96	4.25×10^9
D3	12.08	0.10	15.00	11.19	3.23	3.97×10^9

4.4.3 *Numerical features*

The estimated parameters reported in Table 18 are employed to evaluate the numerical features in order to propose a comparison with the experimental features reported in Section 4.3.

Figure 45 shows the RMS and A^* values computed for all test conditions, both for the numerical and experimental acceleration signals, with respect to the rotation frequency. Concerning the experimental ones, only the average value for the three repetitions of each test is considered. Figure 45.a, Figure 45.c and Figure 45.e report the RMS trends. For defect D1, the RMS values are well captured for all loads and rotation frequencies. This consideration agrees with

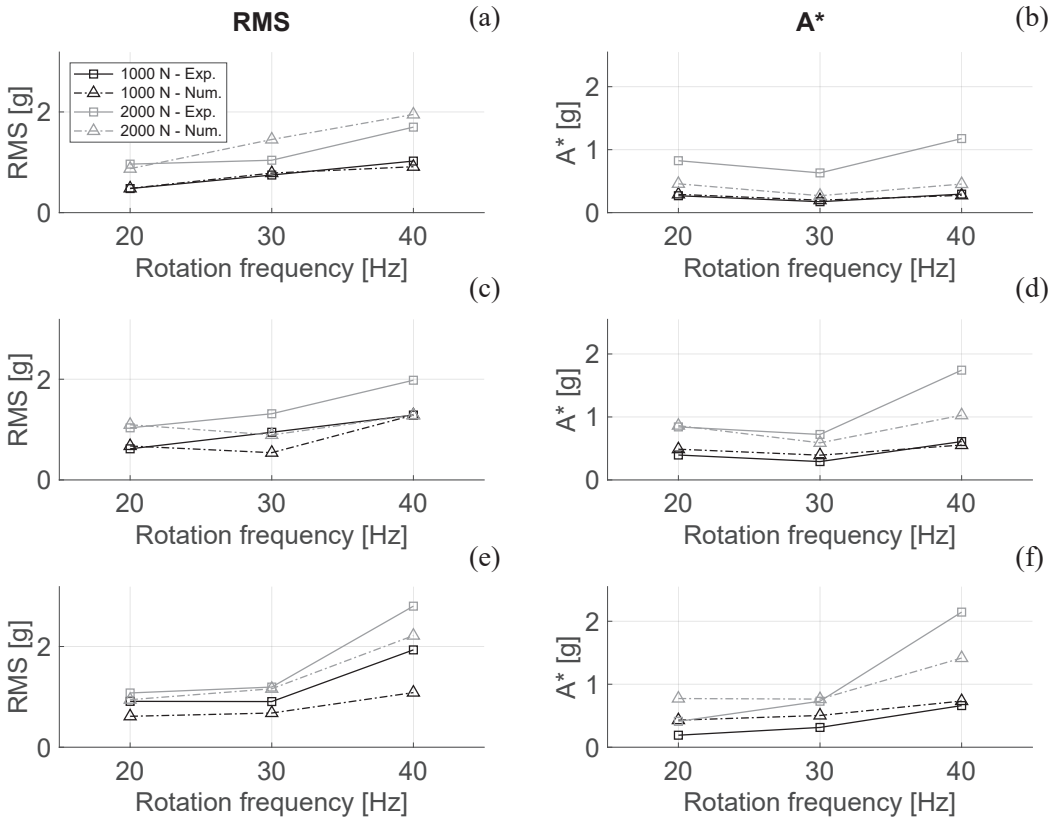

Figure 45: Comparison between numerical and experimental features as a function of rotation frequency: (a) RMS for defect D1; (b) A^* for defect D1; (c) RMS for defect D2; (d) A^* for defect D2; (e) RMS for defect D3; (f) A^* for defect D3.

the observations made on the time signals. For defect width D2, the results agree well at 20 Hz but they show some deviations at the other frequencies. On the contrary, for defect dimension D3 the numerical trends are very similar to the experimental ones, although they assume lower values at all rotation frequency values.

Additional considerations may be provided on A^*, which is depicted in Figure 45.b, Figure 45.d and Figure 45.f. For defect D1, the trends are very similar for a load value equal to 1000 N. Besides, defect width D2 presents trends which are remarkably close for the majority of test conditions. Lastly, the numerical

trends for defect width D3 show a good agreement at a rotation frequency value equal to 30 Hz. It is worth noticing that the optimal parameters found through the multi-objective optimization allows to obtain a lower value of ΔRMS, therefore it is expected that the numerical trends of A^* are more distant from the experimental compared to RMS. The condition that mostly affects these results are the A^* values calculated for defect D1 at a load equal to 2000 N. This deviation is the one that influences the most the value of ΔA^*. The optimization procedure, in fact, allows to obtain a closer match for the other combinations of operative conditions and defect dimensions.

Figure 46 depicts the same features but as a function of the defect width, allowing for additional considerations. For the RMS trends in Figure 46.a, Figure 46.c and Figure 46.e, the experimental and numerical results agree well for all conditions at the rotation frequency of 20 Hz. At 30 Hz, the trends exhibit a deviation for the defect width D2 which alters the trend shape, but the results are similar for the other two defect dimensions. On the other hand, some deviations in the expected trends may be observed for a rotation frequency equal to 40 Hz, for which the best agreement is found for the smallest defect width. Besides, a different behaviour may be noticed for the A^* values shown in Figure 46.b, Figure 46.d and Figure 46.f. For this feature, there is a good correspondence for all cases for a load value equal to 1000 N. For the other load value, the experimental trend of the feature is captured although the magnitude is always lower than expected.

In conclusion, the multi-objective optimization procedure permits to obtain results close to the experimental investigation for different combinations of test conditions. The best results, in terms of agreement between experimental and numerical observations, are obtained when the rotation frequency is equal to 20 Hz or when a defect width D1, i.e. equal to 0.928 mm, is considered. In the first case, results are consistently close to the experimental values for all defect widths. In the second scenario, good agreement is achieved at all rotation frequencies. This kind of behaviour agrees with the observations of previous authors that higher differences with experimental results are usually noticed for increasing defect widths [82, 106]. This discussion underlines the difficulty associated with the definition of model parameters that could take into account multiple test conditions in terms of defect widths, applied loads

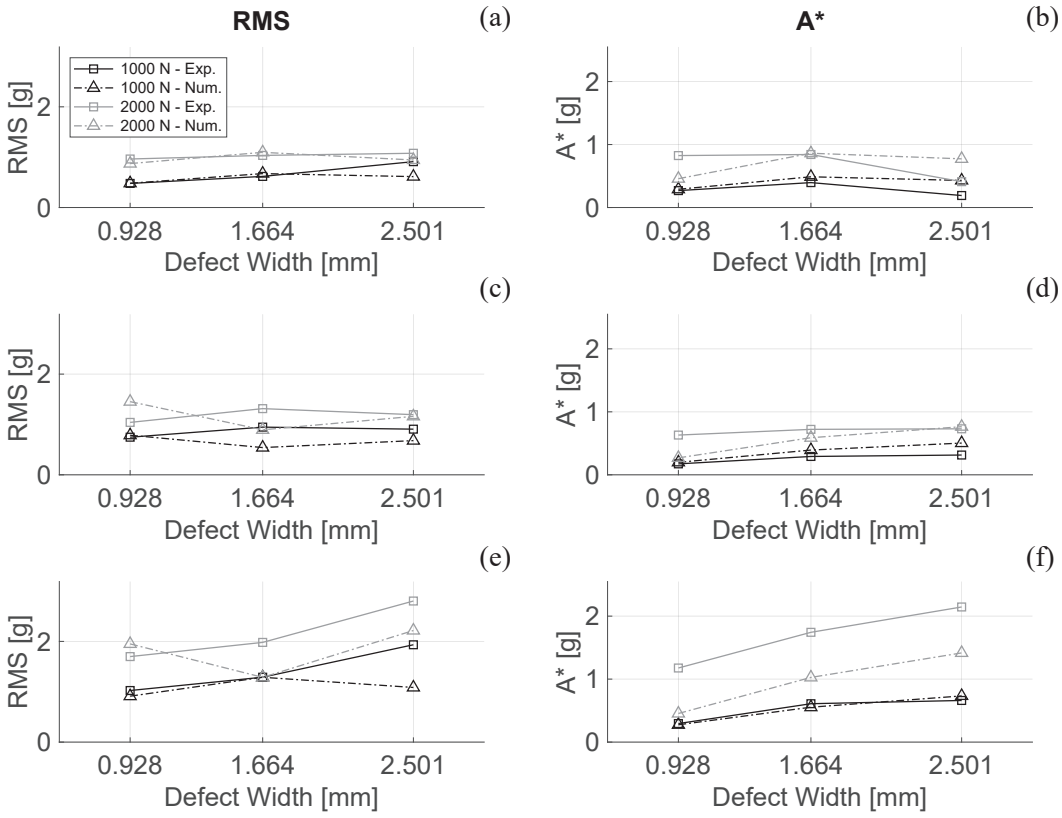

Figure 46: Comparison between numerical and experimental features as a function of defect width: (a) RMS at 20 Hz; (b) A^* at 20 Hz; (c) RMS at 30 Hz; (d) A^* at 30 Hz; (e) RMS at 40 Hz; (f) A^* at 40 Hz.

and shaft rotation frequencies. The proposed technique, which is based on the minimization of global indicator involving all test conditions, it is by definition not capable to simultaneously optimize every test condition as the indicators are average values considering each one of their combinations. However, it enables to propose an estimation for the model parameters that are a priori unknown. Additionally, the study of the time signals and the features trends allows to understand which conditions are better represented by employing the chosen set of parameters.

4.4.4 *Time and frequency analysis*

The numerical model is exploited to evaluate the time signals for all test conditions by employing the set of parameters in Table 18. The time signals are graphed in Figure 47. The limits on the x and y axes are the same as in Figure 31 to permit qualitative comparison between the respective numerical and experimental accelerations. The results show that the numerical model is able to capture a number of test conditions. For defect width D1, the numerical waveform is the most similar to its respective experimental signal for a load value equal to 1000 N at all frequencies. At W = 2000 N, the magnitude of the peaks is mostly captured. For these combinations of defect size, load and frequency values the presence of the defect may be clearly appreciated in the numerical accelerations. For the defects D2 and D3, which are depicted in Figure 47.c, Figure 47.d, Figure 47.e and Figure 47.f, the numerical waveforms at a load value equal to 1000 N denote some deviations from the respective experimental waveforms. In particular, although observable, the defect impulses are more masked in the numerical signals, while they are clearly visible in the experimental accelerations. Nonetheless, for a load value equal to 2000 N the signals are more similar in terms of shape and peak amplitude, especially for the defect width D2. In conclusion, by employing the parameters in Table 14, the best agreement between experimental and numerical results is found for the smallest defect width D1, in accordance with the results in Section 4.4.3.

Additionally, a representative spectrum is reported in Figure 48. The chosen condition is the same as in Figure 32, i.e. the defect dimension D2 = 1.664 mm for a radial load value equal to 2000 N and a rotation frequency of 30 Hz. The frequency content is concentrated at the harmonics of the BPFO, which is equal to 146.7 at f_{rot} = 30 Hz. In the experiments, instead, these components are smeared over the whole spectrum due to the stochastic nature of the real phenomenon [23]. It is worth noting, however, that resonances may be observed approximately around 357.1 Hz, 1543.8 Hz and 7392.6 Hz. In fact, the estimated values of m_r and k_r allow to obtain resonance zones in the frequency spectrum which are remarkably close to the ones observed in the experimental study.

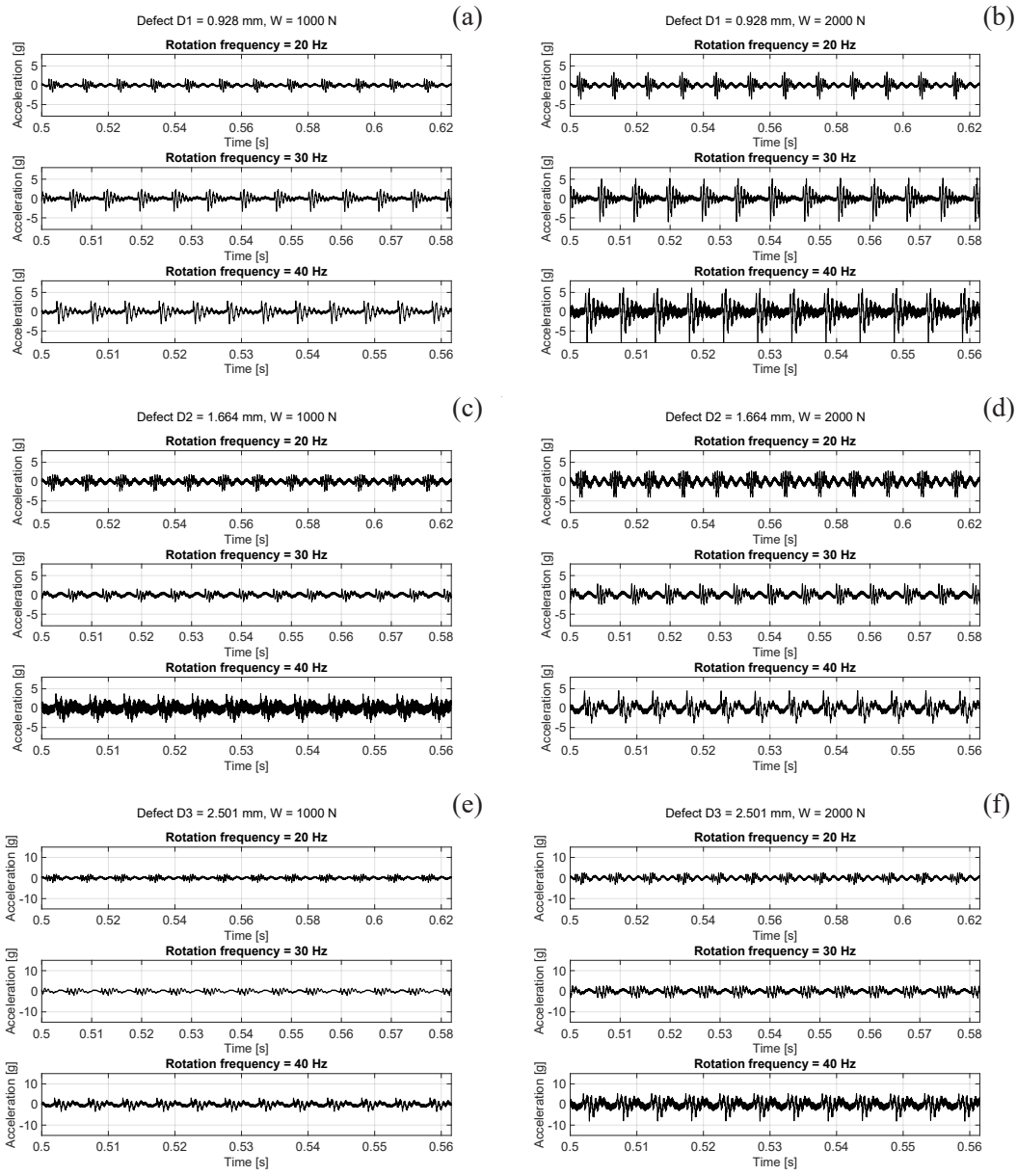

Figure 47: Numerical time signals for different combinations of defect dimension and load w. The limits on the y-axis of each plot are the same as in Figure 31 to allow a comparison between corresponding graphs. (a) D1, $w = 1000$ N; (b) D1, $w = 2000$ N; (c) D2, $w = 1000$ N; (d) D2, $w = 2000$ N; (e) D3, $w = 1000$ N; (f) D3, $w = 2000$ N.

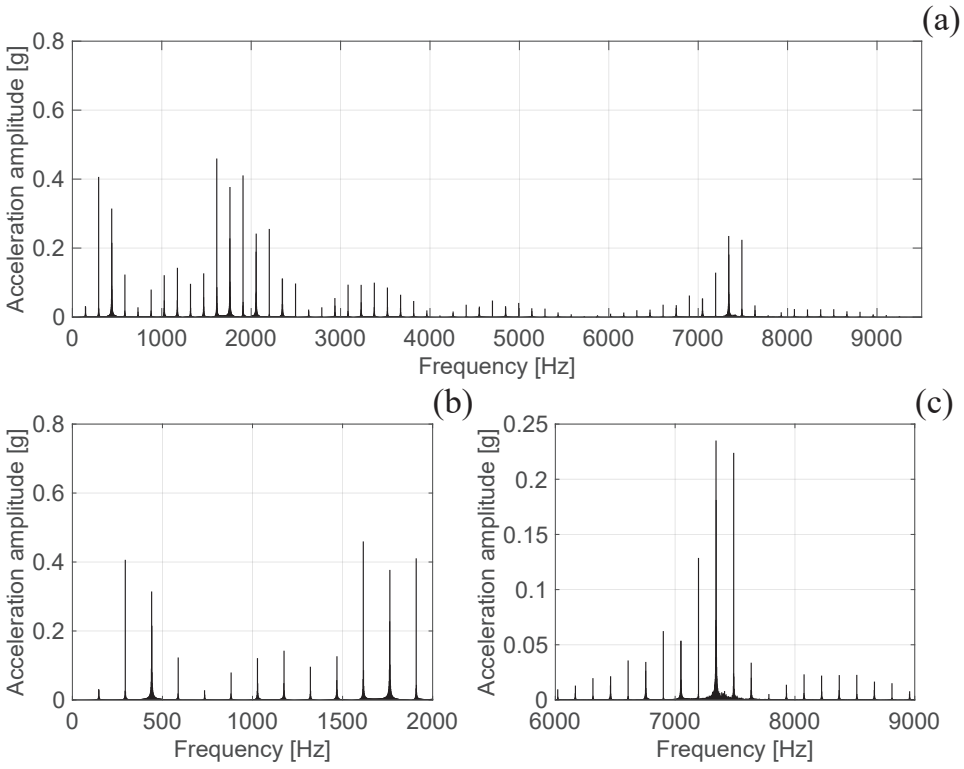

Figure 48: Numerical frequency spectrum for defect D2 under and applied load of 2000 N and a rotation speed of 30 Hz: (a) Spectrum over the frequency range 0 - 9500 Hz; (b) Spectrum in the ”low” frequency range 0 - 2000 Hz; (c) Spectrum in the ”high” frequency range 6000 - 9000 Hz.

4.4.5 *Effect of slippage*

The effect of slippage may be accounted for into the model by adding an additional term $\sigma(t)$ in Eq. 32, so that:

$$\varphi_{jk}(t) = \varphi_{jk}(t) + \sigma(t) \tag{82}$$

$\sigma(t)$ is assumed to be a uniformly distributed variable between two extreme values $-\phi_{slip}$ and ϕ_{slip}, i.e. $\sigma = rand(-\phi_{slip}, \phi_{slip})$. The value of ϕ_{slip} is usually in the order of 0.01 rad and 0.02 rad [24]. To determine the effect on the numerical results, three different value of this parameter are chosen: 0.005 rad, 0.01 rad and 0.02 rad. Simulations are performed by adapting the model to include Eq. 82. Frequency spectra, time signals and the features RMS and A^* are evaluated to compare the effect of this additional influencing factor.

Figure 49.a, Figure 49.b and Figure 49.c depict the frequency spectra obtained by considering the three different slippage angles. In accordance with the previous analyses, the considered test condition corresponds to defect dimension D2, rotation frequency 30 Hz and radial load 2000 N. By comparing these graphs to Figure 48.a, in fact, it is possible to qualitatively assess the effect of ϕ_{slip} on the frequency components. As expected, the introduction of this random component produces a smearing of the frequency components [23], which makes the spectra more resemblant to the experimental ones. More in detail, the effect is particularly evident for the high resonance zone. In that frequency interval the frequency components are not distinctly separated as in the case of no slippage, but are merged around the resonance frequency. Besides, an higher value of ϕ_{slip} leads to an higher peak component. However, the magnitude of the frequency in this zone is still higher than the one observed in the experiments as shown in Figure 32. Concerning the mid resonance zone, the effect of the slippage is not particularly evident for the lowest angle, but its influence noticeably increases for higher values. In particular, for ϕ_{slip} = 0.02 the BPFO harmonics greatly decrease, leading to a greater resemblance with the the experimental spectrum reported in Figure 32. Finally, the effect on the low resonance zone is still observable but substantially less significant.

The differences in the waveforms are observable in the time signals reported in Figure 50. As the value of ϕ_{slip} increases, the impulses are less distinguishable since they appear more blended in the noise introduced by the random

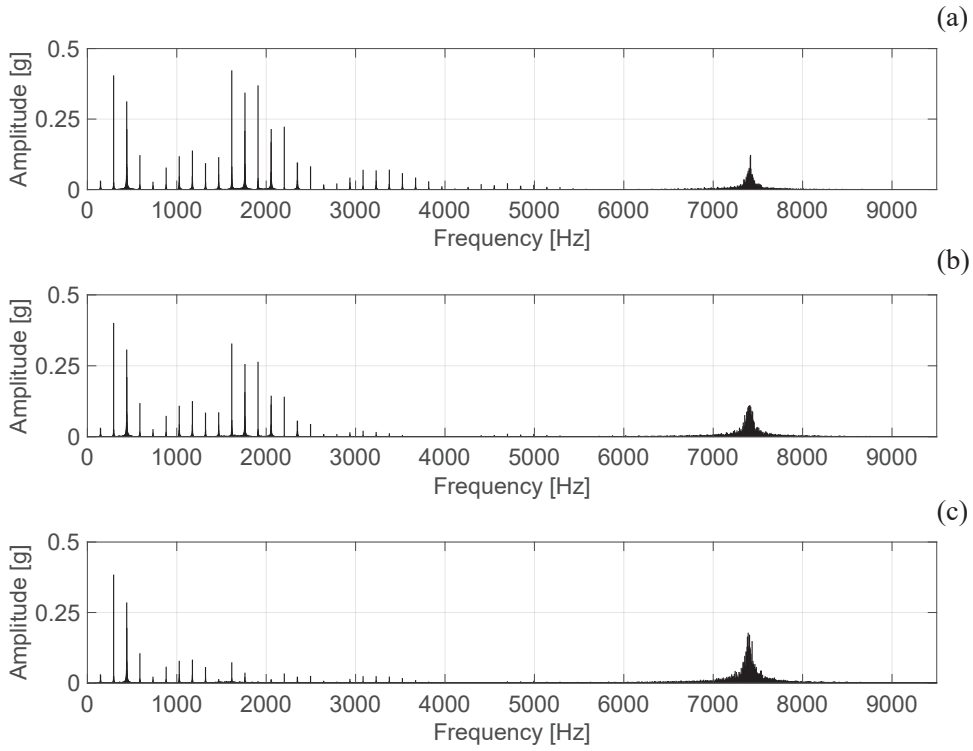

Figure 49: Frequency spectra for the test condition associated with defect width D2, rotation frequency 30 Hz and applied radial load 2000 N for different values of the slippage angle ϕ_{slip}: (a) 0.005 rad; (b) 0.01 rad; (c) 0.02 rad.

variation of the slippage angle. Moreover, because of the random nature of the phenomenon, the signal is no longer periodic as in the case of no slippage. The variability increases by assuming a larger value of ϕ_{slip}.

Eventually, it is questioned whether the slippage angle sensibly affects the global signal features or not. This aspect is inquired in Figure 51.a, Figure 51.c and Figure 51.e for RMS and Figure 51.b, Figure 51.d and Figure 51.f for A^*, respectively. Differently from the frequency and time analysis, only the lowest slip value of 0.005 rad and the highest value of 0.02 rad are reported along with the case of no slippage to enhance the clarity of the figure. By analysing Figure 51.a, it might be noticed that increasing the slip term leads

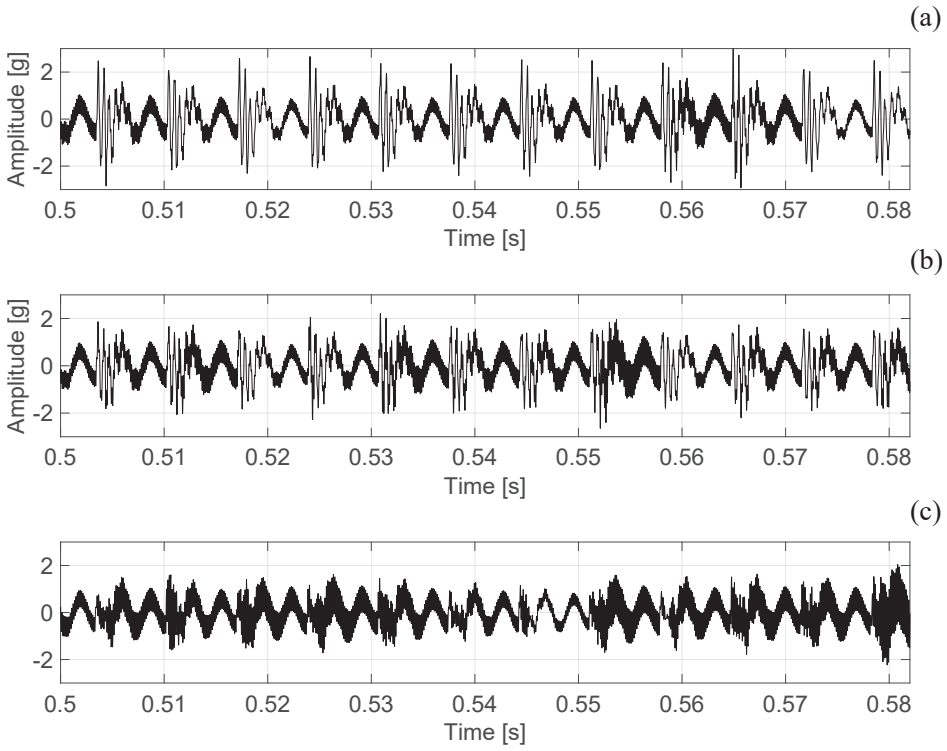

Figure 50: Time signals for the test condition associated with defect width D2, rotation frequency 30 Hz and applied radial load 2000 N for different values of the slippage angle ϕ_{slip}: (a) 0.005 rad; (b) 0.01 rad; (c) 0.02 rad.

to a lower RMS for every test condition. This effect substantially intensifies for a maximum slip angle of 0.02 rad compared to 0.005 rad. Interestingly, the trend shape is preserved in all cases except for the smallest defect, i.e. D1, when ϕ_{slip} is set to 0.02 rad. Therefore, it appears that lowering the defect size magnifies the effect of slippage. This may be due to the fact that for smaller widths the variation of the ball position becomes comparable to the dimension of the defect. In fact, by setting ϕ_{slip} = 0.02 rad the ratio between the slippage angle and the fault extent for defect D1 is equal to 49 %, meaning that at each instant the ball position may vary up to approximately half of the defect width. On the contrary, the indicator A^* is remarkably less affected by ϕ_{slip}. Indeed, increasing values of this coefficient lead to a modest reduction of the

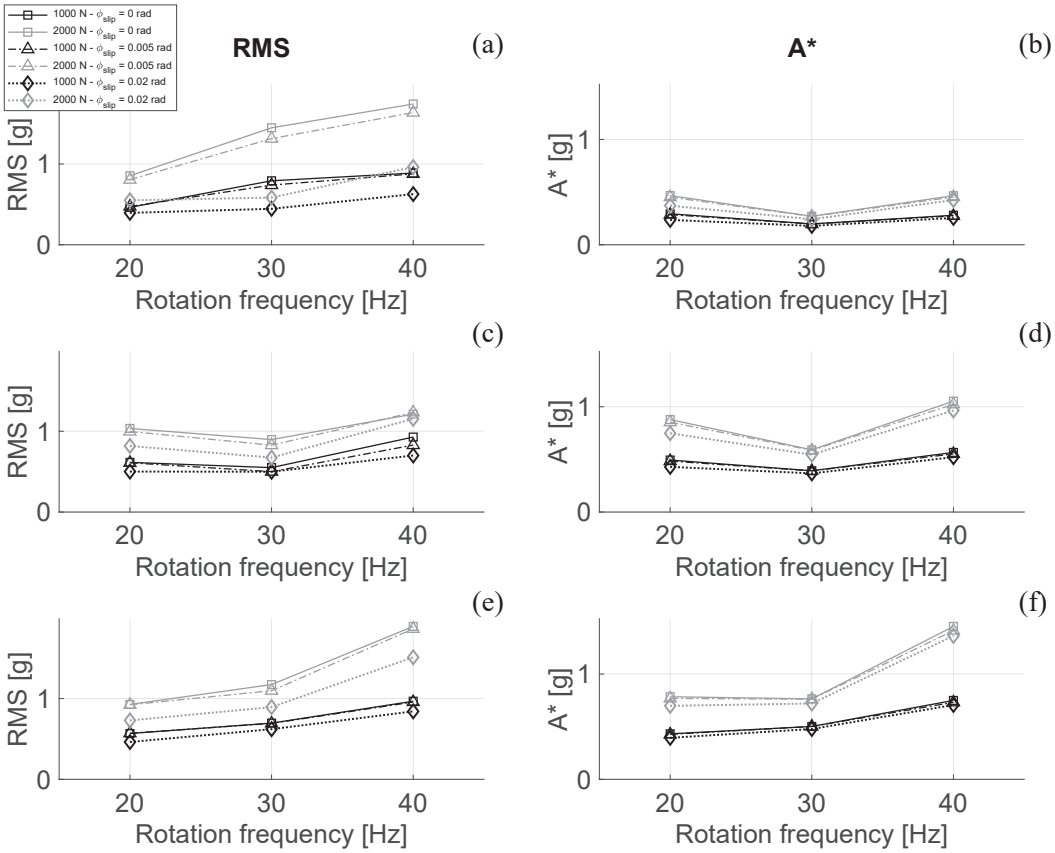

Figure 51: Signal features compared against the rotation frequency for three different slippage conditions and under two applied loads equal to 1000 N and 2000 N: (a) RMS for defect D1; (b) A* for defect D1; (c) RMS for defect D2; (d) A* for defect D2; (e) RMS for defect D3; (f) A* for defect D3.

parameter for all conditions compared to RMS. Since this indicator accounts for the amplitude of the first five BPFO harmonics, this result agrees with the analysis of the frequency spectra illustrated in Figure 49.

4.5 SUMMARIZING REMARKS

The present chapter aimed at the definition of a multi-objective optimization procedure focused on the estimation of unknown model parameters. The technique is applied within the context of simulations of self-aligning ball bearing in presence of defects. The assessment was based on the comparison between global indicators computed on experimental and numerical acceleration signals.

A numerical model of a defective self-aligning ball bearing was set-up based on the model of D. Petersen et al. [88]. The defect definition was adapted to a self-aligning ball bearing by extending the theory proposed by Y. Jiang et al. [22] and S. Gao et al. [99]. Notably, the model is constituted by six degrees of freedom, two of which refer to the displacement of a fictitious high resonant mass that replicates an high frequency response of the system. Differently from previous works, however, the contact damping has been defined by means of Rayleigh damping to achieve a better control over the whole frequency range.

The described multi-objective optimization technique aimed at minimizing two indicators, namely ΔRMS and ΔA^*. The former accounts for the energy of the whole signal, while the latter considers only the first five harmonics of the defect characteristic frequency and represents the energy content of the deterministic phenomena related to the rotation of the components. A set of unknown parameters was chosen among the quantities needed to define the model. Their choice fell on the parameters whose estimation is characterized by an higher degree of uncertainty. Their values were constrained within fixed limits dictated by their physical properties and they were linked by additional analytical relationships based on the physics of the problem. The experimental results, which were needed for the assessment of the indicators, were obtained through an experimental study on a number of defective self-aligning ball bearings. Each one was tested under several test conditions in terms of different values of radial load and rotation frequency.

The outcome of the multi-objective optimization resulted in a number of weak Pareto optimals characterized by a small deviation of ΔA^* and a higher variability of ΔRMS. The chosen parameters set was designated as the one providing

the lowest value of the second indicator. The experimental test results were then compared with the numerical ones by inserting the estimated set of unknown parameters in the numerical model. A good accordance with the experimental results was obtained for all test conditions. In particular, the best agreement was found for all conditions associated with the smallest defect width, i.e. D1 = 0.928 mm, and for all conditions associated with the lowest rotation frequency value, i.e. 20 Hz. These results are in agreement with the limits of this kind of models, which tend to lose accuracy for increasing defect widths [82, 106]. Finally, the frequency analysis of a representative spectrum showed that the frequency content is mainly concentrated at the BPFO harmonics. This effect, which is not present in the experimental observations, is considerably altered by the slippage of rolling elements, namely ϕ_{slip}. Depending on the magnitude of the hypothesized value of ϕ_{slip}, the indicators become consistently lower for increasing values of slippage. The influence of this parameter appears to be more marked for smaller defects as its amplitude becomes comparable to the extent of the defect. Nonetheless, the estimated parameters allowed to capture the resonance phenomena of the real system even neglecting this term in the analysis.

The conducted study highlights the difficulties regarding the choice of the best parameters to be employed in the numerical modelling of REBs. However, within this context it provides guidelines and a robust procedure that allows to estimate the unknown parameters that would otherwise be manually tuned on the basis of the experience of the author or on practical considerations. It is also worth underlining that, although the investigation was performed on a double-row ball bearing, the proposed procedure may be applied for other bearing types, provided that the bearing model is adapted accordingly. In conclusion, the inquiry carried out in the chapter should be considered as a basic framework to deal with problems in which multiple operative conditions must be validated.

5

DEVELOPMENT OF BEARING PROGNOSTIC TECHNIQUES BASED ON EQUIVALENT DAMAGED VOLUME ESTIMATION

5.1 INTRODUCTION

The survey in Section 1.2.3 illustrated the variety of approaches that are commonly employed to deal with the problem of bearing prognostics. Among those, PBMs attracted the attention of researchers thanks to their capability of describing the degradation process of a REB by means of a dedicated mathematical modelling of its physics. Therefore, their main advantage compared to DDM techniques is their potential capability of assessing the future deterioration of the system under exam for different sets of operative conditions, whereas DDMs would need further training data for each considered one. Besides, Section 1.2.3.1 showed that the majority of PBMs rely on the PE equation to infer the crack growth over time. Surprisingly, very few works dealt with the prognosis of bearing life by utilizing dynamic LP models. To this end, a recent work of L. Cui et al. [145] employed a LP model of a defective bearing in order to establish a prognostic model based on the similarity theory [146]. Their effort is noteworthy, but other approaches may be developed on the basis of models analogous to the one described Chapter in 4.

Within this framework, this chapter provides an original take on the problem of REB prognostics by introducing a novel parameter related to bearing degradation, namely the Equivalent Damaged Volume (EDV). This quantity

represents the equivalent defect dimensions per unit of width that, if inserted in the LP model of the defective bearing, would produce the same RMS value as the real feature computed on the experimentally acquired signal. The EDV values are computed through a specifically-designed algorithm. At each time, the experimental RMS value is used to determine the corresponding defect dimensions in the LP model. The algorithm is constructed so it always generates increasing defect volumes over time, except when no suitable values that satisfy this condition are found among the possible fault dimensions. Moreover, the formulation of the LP model described in Chapter 4 is expanded to also take into account the defect depth. In fact, Section 2.3.2 highlighted that faults which were generated in the performed run-to-failure tests are extended rather than localised. Therefore, they need to be modelled by also considering their depth. Then, a first approach to tackle bearing prognostics through the employment of the EDV algorithm is detailed. For this purpose, two PBMs are described. The first one aims at predicting the REB vibration under different operative conditions with respect to a given reference deterioration history. On the other hand, the objective of the second PBM is to predict the time until a certain threshold on the equivalent damaged volume is crossed, regardless of the applied load and the shaft rotation speed. The efficiency of the proposed models is tested by means of the experimental data acquired on the test bench during the run-to-failure tests.

The chapter is subdivided as follows. Section 5.2 introduces the faulty bearing model by extending the LP model previously established for localised defects. On the basis of this model, a procedure for determining the equivalent damaged volume during REB operation is detailed, and two possible physics-based model based on this parameter are developed. Then, Section 5.3 reports the results of the run-to-failure tests described in Section 2.3.2. Thanks to these results, Section 5.4.1 details the estimation process of the equivalent damaged volume from the experimental signals and subsequently provides the outcome of the two PBMs. Finally, Section 5.5 closes the discussion with some final thoughts on the proposed work and its potential further developments.

5.2 DESCRIPTION OF THE PROPOSED PROCEDURE

This section provides a modification to the LP model introduced in Section 4.2.2 in order to also consider extended defects. The improved LP model is then employed to define a novel parameter named Equivalent Damaged Volume (EDV) that is utilized to track the damage progression over time. An algorithm is proposed to compute this peculiar quantity at each instant. Then, the procedure is implemented in a prognostic framework by proposing two different approaches to the problem.

5.2.1 *Faulty bearing model for extended defects*

In the previous chapter, Section 4.2.2 described a LP model of a faulty bearing in which the trajectory of the rolling element inside the defect was determined based on the assumption that the ball was not able to touch the bottom of the defect. This assumption holds true for the analyses proposed in Chapter 4, since the dimensions of the seeded artificial defects were determined *ad hoc* to avoid the contact of the balls with the bottom surface of the defects themselves. However, this hypothesis cannot be introduced in run-to-failure tests. In fact, this type of damage may exist only at the beginning of the unhealthy stage of the bearing life. Subsequent passage of the rolling elements in the defect area further propagates the defective area, increasing its angular extension and modifying its topology from a localised to an extended defect. This is particularly evident from the photos of the damaged bearing shown in Figure 12. In fact, at the end of the tests all outer rings presented an extended defect whose angular span was about 40 to 50 degrees. As a result, it is not possible to model the vibratory response of the system without taking into account the possible contact between the bottom surface of the fault and the rolling elements. Consequently, this section extends the numerical model provided in Sections 4.2.1 and 4.2.2 to further consider the possible contact with the bottom surface of the defect. This is done by changing the ball trajectory described by Eq. 51. The presented method takes inspiration from the work of Y. Jiang et al. [22], who modelled the ball trajectory for an extended defect in the case of a cubic-like defect. However, this shape is different from the ones experimentally observed in the run-to-failure campaign. Therefore, their work is modified

to extended defect with a more realistic shape, i.e. denoted by a curvature comparable to the outer ring radius. Additionally, the proposed model also considers the roughness of the damaged surface. To this end, the procedure described by N. Sawalhi and R.B. Randall [115] is further implemented in the proposed extended defect formulation.

Firstly, the ball trajectory without the rough surface is described. The extended defect is schematically depicted in Figure 52. In the plane O'x'y', which is defined in the same manner as in Section 4.2.2, the bottom surface follows a circular path with an offset h_d from the undamaged bearing surface. Therefore, parameter h_d is the depth of the extended defect. As a result, it possible to rewrite Eq. 51 in order to take into account this additional parameter and consequently find the new trajectory $d_{\alpha,b}$:

$$d_{\alpha,b}(\varphi_{jk}(t)) = \begin{cases} r_b + r_o\cos(\varphi_{jk}(t) - \varphi_{d,in}) - r_o + \\ -\sqrt{r_b^2 - (r_o\sin(\varphi_{jk}(t) - \varphi_{d,in}))^2} & \text{if } \varphi_{d,in} \leqslant \varphi_{jk}(t) < \varphi_{b,in} \\ h_d & \text{if } \varphi_{b,in} \leqslant \varphi_{jk}(t) \leqslant \varphi_{b,out} \\ r_b + r_o\cos(\varphi_{d,out} - \varphi_{jk}(t)) - r_o + \\ -\sqrt{r_b^2 - (r_o\sin(\varphi_{d,out} - \varphi_{jk}(t)))^2} & \text{if } \varphi_{b,out} < \varphi_{jk}(t) \leqslant \varphi_{d,out} \\ 0 & \text{otherwise} \end{cases} \tag{83}$$

where $\varphi_{b,in}$ and $\varphi_{b,out}$ are the angular positions in which the ball begins and finishes being in contact with the bottom surface. Their values may be calculated by equating the defect trajectory equation at the beginning and at the end of the path with the depth h_d of the extended defect. The associated problem, then, consists in solving two non-linear equations in the unknowns $\varphi_{b,in}$ and $\varphi_{b,out}$, viz.:

$$r_b + r_o\cos(\varphi_{b,in} - \varphi_{d,in}) - r_o + \\ - \sqrt{r_b^2 - (r_o\sin(\varphi_{b,in} - \varphi_{d,in}))^2} - h_d = 0 \tag{84}$$

$$r_b + r_o\cos(\varphi_{d,out} - \varphi_{b,out}) - r_o + \\ - \sqrt{r_b^2 - (r_o\sin(\varphi_{d,out} - \varphi_{b,out}))^2} - h_d = 0 \tag{85}$$

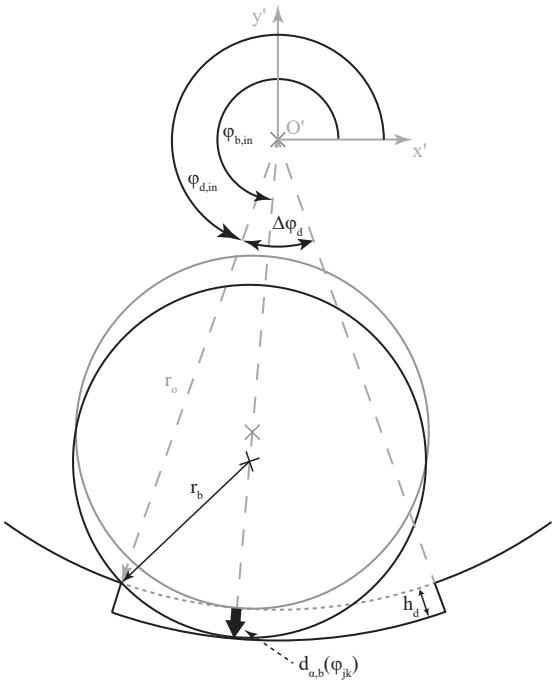

Figure 52: Geometric definition of angular position $\varphi_{b,in}$. The same scheme may be replicated for $\varphi_{b,out}$ at the exit of the defect.

The proposed rolling element path $d_{\alpha,b}(\varphi_{jk}(t))$ describes the bottom of the defect as a smooth surface with constant depth. However, experimental evidence shows that the surface is characterized by a significant roughness. Thus, the equation for $d_{\alpha,b}$ between $\varphi_{b,in}$ and $\varphi_{b,out}$ is further modified to also take into account this aspect. Surface roughness is introduced by applying the procedure described in [115], but only between angles $\varphi_{b,in}$ and $\varphi_{b,out}$. This way, the ball is free to fall inside the defect according to the theoretical path, and then its trajectory is modified by the rough bottom surface until the ball starts to get out of the defect. The procedure to generate the rough surface is subdivided in several steps. First, a random Gaussian noise is generated and scaled according to the value of the surface roughness R. This random profile is defined between the angular locations $\varphi_{b,in}$ and $\varphi_{b,out}$ and it is characterized by a spatial sampling $\delta\varphi$ which is equal to:

$$\delta\varphi = \frac{\omega_c}{f_s} \tag{86}$$

where f_s is the signal sampling frequency. Then, the random noise is low-pass filtered in order to generate a plausible profile that may be touched by the rolling elements. According to [115], this cut-off frequency f_l is equal to:

$$f_l = \frac{1}{\pi\sqrt{4Rr_b}} \tag{87}$$

Finally, the outcome of this procedure is the random profile $h_{dr}(\varphi_{jk}(t))$, which modifies Eq. 83 so that:

$$d_{\alpha,b}(\varphi_{jk}(t)) = h_d + h_{dr}(\varphi_{jk}(t)) \quad \text{if } \varphi_{b,in} \leqslant \varphi_{jk}(t) \leqslant \varphi_{b,out} \tag{88}$$

The final step consists in multiplying $d_{\alpha,b}(\varphi_{jk}(t))$ by the cosine of the contact angle to determine the radial displacement, as in Eq. 52. Therefore, the resulting ball displacement in the fault area is:

$$d_b(\varphi_{jk}(t)) = d_{\alpha,b}(\varphi_{jk}(t)) \cos\alpha \tag{89}$$

Figure 53 reports examples of the ball path inside the defect. Figure 53.a shows the trajectory followed by the ball in a localised defect characterized by an angular extent $\Delta\varphi_d = 2°$ and sufficiently deep to avoid contact with its bottom. On the other hand, Figure 53.b depicts the ball path in a extended defect with $\Delta\varphi_d = 20°$, depth $h_d = 20$ µm and roughness $R = 10$ µm. In Figure 53.b, the dash-dot line represents the path with the smooth bottom, while the solid line describes the path due to a rough surface. Figure 53.b also permits to assess that the effective path has a deviation around h_d lower than the imposed R value. In fact, the profile generation procedure involves the low-passing of the signal to fit the ball to the path it can practically reach. The efficiency of this method was also successfully demonstrated in later works by D. Petersen et al. [49, 88] to characterize extended defects and discuss their influence on bearing vibrations.

5.2.2 *Equivalent damaged volume (EDV) method*

This section details a novel procedure for the estimation of a parameter, related to the evolution of bearing life, which is able to assess the bearing health deterioration over time, namely the equivalent damaged volume per unit of

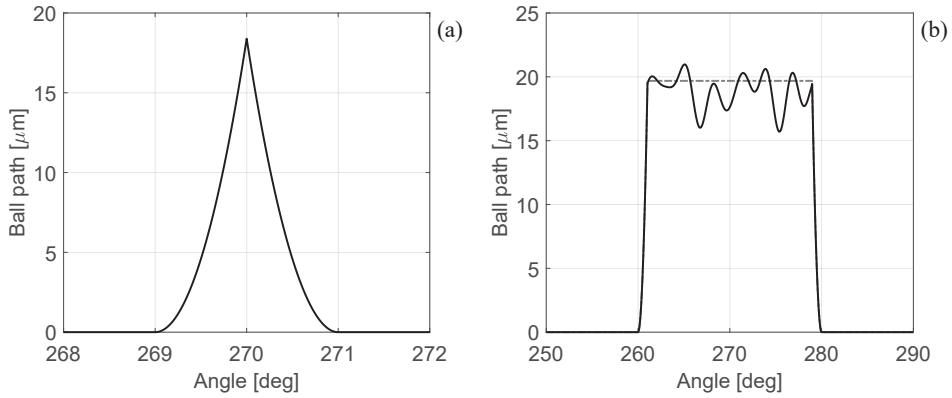

Figure 53: Examples of trajectories followed by the ball inside a defect: (a) localised defect with $\Delta\varphi_d = 2°$ and sufficient depth to avoid contact with the bottom surface; (b) extended defect with $\Delta\varphi_d = 20°$, $h_d = 20$ μm and $R = 10$ μm. The solid and dash-dot lines represent the paths on the rough and smooth surfaces, respectively.

width. Because of this, the proposed technique is named Equivalent Damaged Volume (EDV) method. The algorithm tracks the bearing damage evolution by determining the angular extent $\Delta\varphi_d$ and depth h_d of the equivalent defect that, if given as input to the LP model of the faulty bearing, would give the same RMS as the experimentally measured one at each time step. As a consequence, at the i-th time step $t(i)$ the EDV method determines an equivalent volume $V_{eq}(i)$ per unit width equal to:

$$V_{eq}(i) = r_o \Delta\varphi_d(i) h_d(i) \tag{90}$$

where $\Delta\varphi_d(i)$ and $h_d(i)$ are the defect dimensions at time $t(i)$. If r_o and h_d are expressed in mm and $\Delta\varphi_d(i)$ is expressed in radians, then $V_{eq}(i)$ assumes the units mm^3/mm. To reduce the complexity of the procedure, in fact, the axial width of the defect is not considered in the model. Therefore, it is postulated that the defect maintains a constant area along the bearing axial direction. This assumption is based on the consideration that the axial extent of the defect has a smaller influence on bearing vibrations compared to its other dimensions [22]. Moreover, this procedure is intended to track an *equivalent* damaged volume rather than the effective volume. Consequently, V_{eq} is actually defined as an equivalent volume per mm of width, which explains the peculiar unit of measurement assigned to it. Furthermore, the algorithm is constructed to

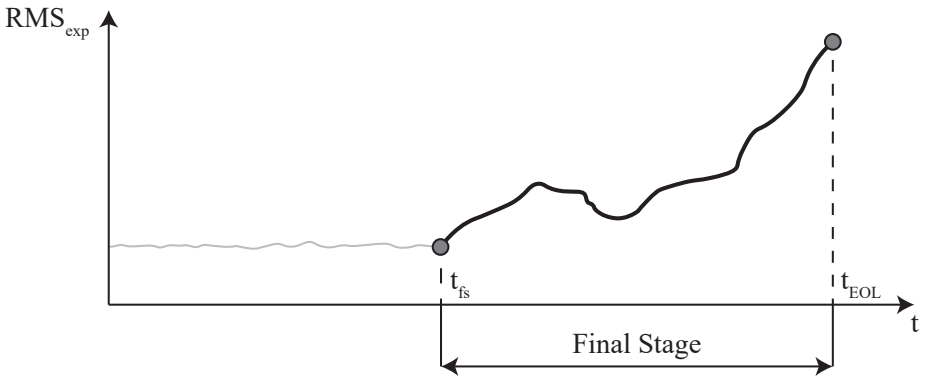

Figure 54: Final stage of bearing life.

provide increasing values of $\Delta\varphi_d$ and h_d during REB degradation, which is a realistic assumption considering that the defect should increase its dimensions as the deterioration progresses.

The proposed algorithm considers only the final part of the bearing life. Therefore, with reference to Figure 4, the bearing must be either in the unhealthy stage or in the critical stage. For the purpose of this work, the working range of the algorithm is named *final stage*. The time at the start of the final stage is named t_{fs}, whilst the time at the end of life (EOL) is referred to as t_{EOL}. The latter is found either by directly monitoring the RMS over time or by utilizing some diagnostic indicator capable of detecting the incipient fault. Several indicators and features may be exploited for this purpose [23]. On the other hand, the former is evaluated by assigning a threshold on one signal feature, and the test terminates once this threshold is crossed. To this end, a commonly chosen parameter is the peak acceleration in the time signal. Figure 54 shows a simplified depiction of the final stage observable in the RMS trend.

The equivalent damaged volume is extracted from the experimental RMS trend by comparison with a numerically generated RMS map, which gives the values of numerical RMS for several combinations of defect depths and circumferential extensions. The map is computed by employing the LP model described in Section 4.2.1 and modified by introducing the defect path detailed

in Section 5.2.1. Consequently, RMS values are calculated by defining two vectors, i.e. $\mathbf{h_d}$ and $\boldsymbol{\Delta\varphi_d}$, which both range from a minimum to a maximum value, so that $\mathbf{h_d} = [h_{d,min}; h_{d,max}]$ and $\boldsymbol{\Delta\varphi_d} = [\Delta\varphi_{d,min}; \Delta\varphi_{d,max}]$. These limiting values may be assigned on the basis of experimental observations. For instance, the maximum angular extent $\Delta\varphi_{d,max}$ may be set according to typical values of defect size observed at the end of bearing life. Then, for each pair of values $(h_d, \Delta\varphi_d)$, the associated numerical RMS is evaluated by employing the LP model of the faulty bearing. The parameters of the dynamic model are tuned according to the procedure described in Chapter 4. Although the method has been validated for localised defects, the estimated parameters are suitable for employment in further models such as the one proposed in this section. In fact, the assessed parameters are characteristic properties of the system under exam. Moreover, the external load and shaft rotation speed are set to reflect the ones in the experiment. Therefore, the RMS map is different for each working condition of the system. In addition, to simplify the problem, the angular position of the centre of the defect, i.e. φ_d, is assumed to be aligned with the direction of the radial load. This a reasonable assumption that allows to reduce the number of unknowns in the problem. Finally, the RMS map is generated by only considering outer race defects. This is in accordance with the faults observed at the end of bearing life in the performed run-to-failure tests, as seen in Figure 12. Nevertheless, the proposed procedure could be implemented in the same manner for other types of defects, i.e located on inner race, rolling elements or cage.

The main steps of the EDV algorithm are schematically depicted in Figure 55 and summarized in the flowchart reported in Figure 56. Initially, the algorithm takes as input the numerical RMS map computed for the experimental working condition and the experimental RMS history in the final stage of the bearing life. The numerical values extracted from the map are named X_{num}, while experimental values are referred to as X_{exp}. At the end, it provides as output the equivalent damaged volume V_{eq} for mm of defect width. Moreover, it supplies the associated values of defect depth h_d and angular extent $\Delta\varphi_d$ at each time step. In between, the procedure is subdivided as follows:

1. The index i is initialised with a value equal to 1.

2. The i-th value of the experimental RMS, i.e. $X_{exp}(i)$, is evaluated at time $t(i)$, as shown in Figure 55.a. Notably, when $i = 1$, $t(1) = t_{fs}$.

3. The numerical RMS map must contain at least one value equal to $X_{exp}(i)$ in order to compute the associated V_{eq}. Therefore, the algorithm may continue only if $min(X_{num}) \leqslant X_{exp}(i) \leqslant max(X_{num})$. Otherwise, the procedure is stopped and the damaged volume is only computed up to time $t(i-1)$. If $X_{exp}(1)$ does not uphold this condition, then no value of V_{eq} is computed and the algorithm is immediately stopped.

4. The algorithm searches on the RMS map all values of X_{num} which are equal to $X_{exp}(i)$ and stores the associated values of $\Delta\varphi_d$ and h_d for each one, as depicted in Figure 55.b. The corresponding volume is also calculated through Eq. 90.

5. For each $X_{exp}(i)$ value, several pairs ($\Delta\varphi_d$, h_d) may exist that satisfy the condition $X_{num} = X_{exp}(i)$. Among all these, it is assumed that the value $V_{eq}(i)$ at time $t(i)$ is the one associated with the pair ($\Delta\varphi_d(i)$, $h_d(i)$) which provides the lowest possible volume. This condition is represented in Figure 55.c and may be written as:

$$V_{eq}(i) = arg\ min(X_{num} : X_{num} = X_{exp}(i)) \tag{91}$$

   The assumption ensures the minimum relative variation with respect to the equivalent volume $V_{eq}(i-1)$ at instant $t(i-1)$.

6. If $t(i) = t_{EOL}$, the algorithm is stopped. Instead, if $t(i) < t_{EOL}$, the procedure continues.

7. To ensure the increment of damage progression over time, numerical RMS values that are associated with values of $\Delta\varphi_d$ and h_d lower than $\Delta\varphi_d(i)$ and $h_d(i)$ are removed from the numerical RMS map. As a consequence, in the subsequent time step the available pairs ($\Delta\varphi_d$, h_d) are only the ones in the right-most part of the plot, as shown in Figure 55.d. This assumption also guarantees that $V_{eq}(i+1) \geqslant V_{eq}(i)$, i.e. the damaged volume increases over time. However, there might be cases where at least one suitable pair ($\Delta\varphi_d$, h_d) existed in the original RMS map, but it has been removed in previous iterations. As a result, in this situation the algorithm is not able to continue and it must be stopped. A similar condition is also reached if one of the limiting values $h_{d,max}$ and $\Delta\varphi_{d,max}$ is reached. In this case, the

next available pairs would be only the ones which correspond to $h_{d,max}$ or $\Delta\varphi_{d,max}$. While these two criteria could be employed as end conditions for the algorithm, another approach is chosen in this work. In particular, when there are no pairs available, the search range is extended back to previously removed values until one value that satisfies the condition is found. In fact, it is reasonable to assume that the RMS value $X_{exp}(i)$ that must be interpolated in the numerical map is sufficiently close to the previous one. Therefore, in order to find this pair, the search area is extended by one unit in both coordinates $\Delta\varphi_d$ and h_d until the first pair of damage dimensions that ensures $X_{num} = X_{exp}(i)$ is detected. The downside of this method is that it always generates an equivalent volume lower than the one in the previous step, i.e. $V_{eq}(i) < V_{eq}(i-1)$. As a consequence, it must be employed with caution to avoid unsatisfactory results. The goodness of this hypothesis is examined in the section dedicated to the experimental validation of the proposed procedure.

8. Index i is increased by one unit, that is, $i = i + 1$, and the procedure is repeated starting from step (2) until one of the stopping criteria is met.

Finally, Figure 55.e and Figure 55.f also show the selection process of $X_{exp}(i+1)$ and $X_{num}(i+1)$, respectively. In particular, the latter highlights with a grey hatching the area removed in step (7).

5.2.3 *Possible applications of the EDV algorithm to REB prognostics*

The EDV algorithm provides the equivalent damage evolution over time in the final stage of bearing life. This information may be employed to develop bearing prognostics techniques, i.e. aimed at determining the RUL of the REB during its operative life. The interesting aspect of models developed in such a manner is that, by being based on the EDV algorithm, they fall in the category of PBM models. As detailed in the literature review proposed in Section 1.2.3, these types of models are usually based on the PE equation. In fact, despite the large number of papers discussing dynamic LP models, very few authors have employed them for bearing prognostics.

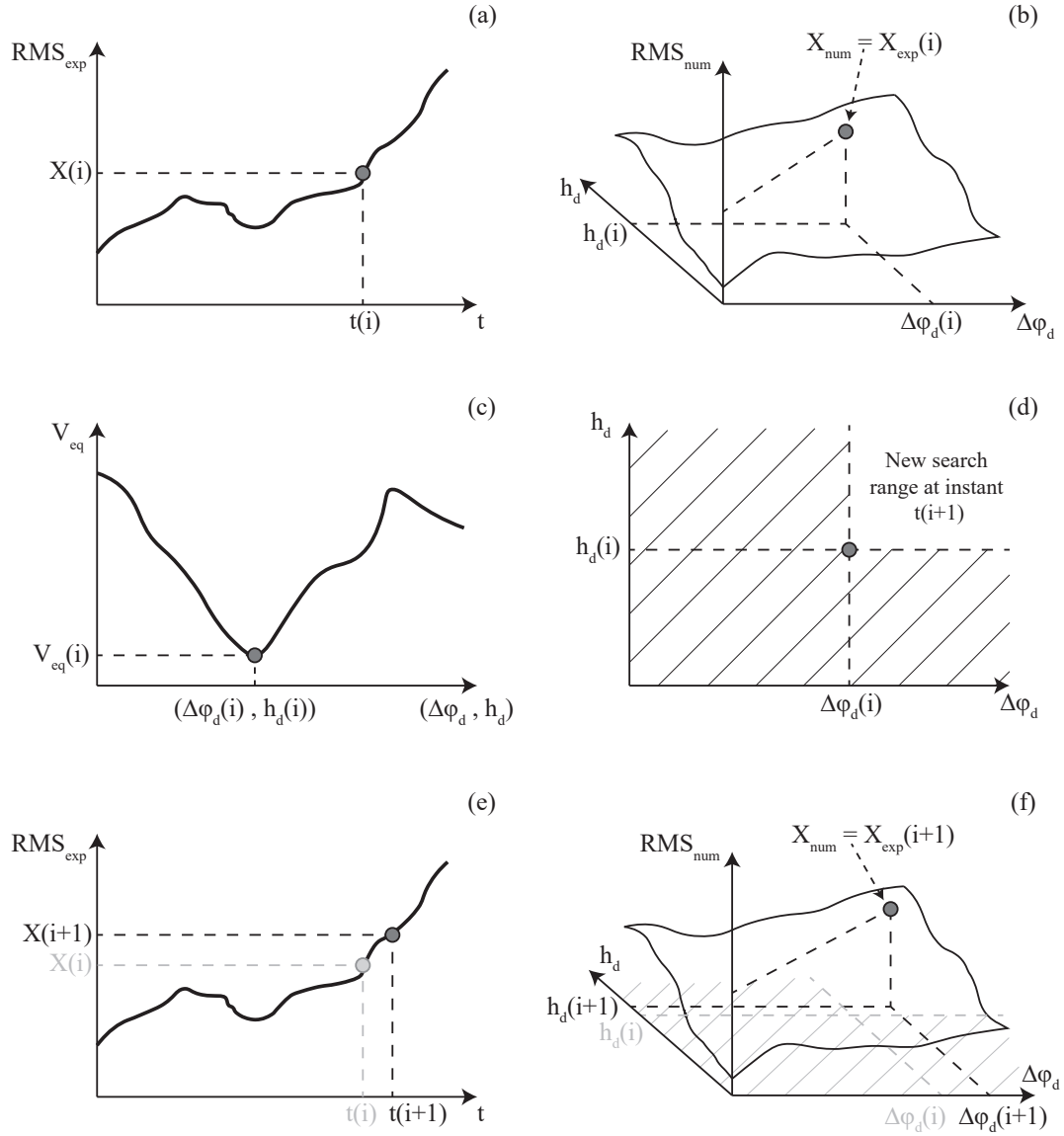

Figure 55: Steps of the EDV algorithm: (a) Consider the experimental RMS value $X(i)$ at time $t(i)$; (b) Find all points in the RMS map that satisfy the condition $X_{num} = X_{exp}(i)$; (c) Determine the pair $(\Delta\varphi_d, h_d)$ that minimizes the equivalent damaged volume V_{eq}; (d) Remove all points with lower angular extent and depth from the RMS map for the next iteration; (e) Repeat the procedure at time $t(i+1)$, until an end criterion described by Figure 56 is achieved; (f) RMS map at time $t(i+1)$, in which the area removed in Figure (d) is highlighted by a grey hatching.

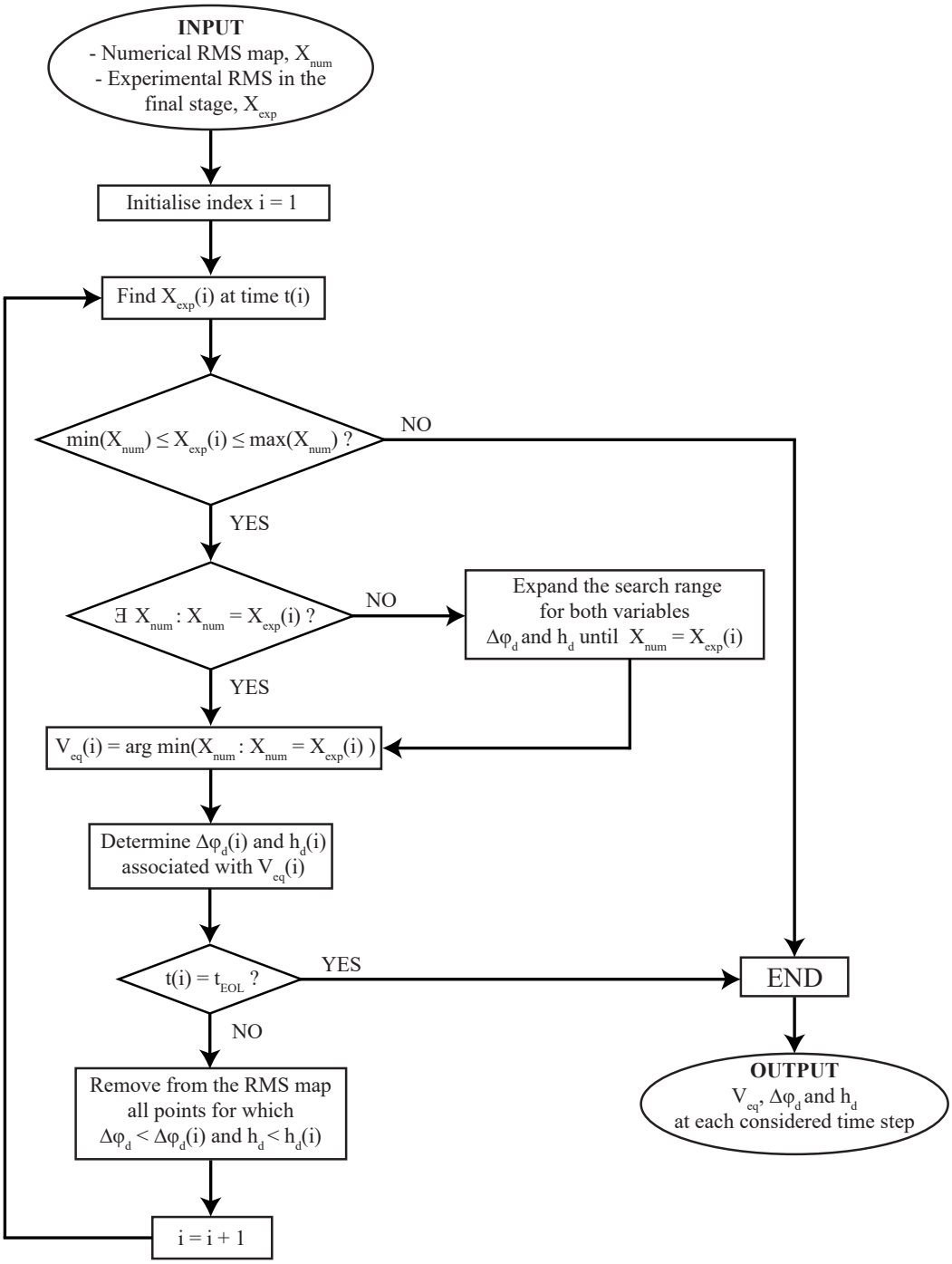

Figure 56: Flowchart of the EDV algorithm.

Within this context, L. Cui et al. [145] recently employed the similarity theory [146] to predict bearing RUL by comparing a numerically generated Performance Degradation Dictionary (PDD) with experimental RMS trends. In their model, the authors subdivided the bearing life in several stages, i.e. steady stage, defect initiation, defect propagation and damage stage. Then, they constructed bearing degradation histories by assigning different defect dimensions and several values of propagation speed in each stage. Therefore, the PDD was generated thanks to a faulty bearing LP model, which was exploited to perform a massive amount of simulations each one characterized by different defects dimensions and propagation velocities. Subsequently, RUL was predicted by comparing the PDD with the experimental trends and evaluating the most probable time that would lead to failure. This method is very interesting, since it provides a different approach to bearing prognostics compared to previous works. However, it presents some downsides. Firstly, the generation process of the PDD is extremely long and is characterized by a remarkable computational burden. Secondly, propagation speeds and defect dimensions between different stages are arbitrarily assumed. More importantly, the defect propagation velocity is hypothesized to always have a constant value. This may not be true in the real case, as the damage might propagate according to different laws other than a linear one. As a result, the proposed approaches take inspiration from the work in [145], in the sense that a dynamic LP model of a faulty bearing is employed to deal with the problem of bearing life prognosis. However, the proposed methodology is remarkably different, since it exploits a LP model with a different definition of the path of the rolling element inside the defect area. Moreover, it introduces a novel technique, namely the EDV algorithm, to be further implemented in prognostic models.

To summarize, this work attempts to provide a first approach to the problem of REB prognostics through the EDV method. Such a task is carried out in the form of two possible approaches, that will be verified with the aid of the results of the run-to-failure experiments described in Section 2.3.2. The first one regards the prevision of the degradation history by applying a previously known damage progression history to different working conditions. The second one aims at predicting the time between the current time and the time in which the value V_{eq} crosses a given threshold. These two approaches are briefly described hereinafter.

5.2.3.1 *The first PBM: prevision of the degradation history under different operative conditions*

The first PBM aims at predicting the vibration history of a REB under a certain set of operative conditions on the basis of the EDV estimated under another working condition. This idea is based on the hypothesis that the applied load and the rotation speed value do not differ significantly from the values referred to the RMS evolution from which the V_{eq} trend over time has been extracted. Therefore, the underlying assumption of this method is that the damage progression under a different load or shaft speed follows the same progression as the original one. This hypothesis, however, certainly does not hold true for a single reference V_{eq} history, because it would lead to failures that all have the same t_{EOL}. Therefore, several degradation histories under one single working conditions are needed to perform a comparison with available data under other conditions. As a result, within the context of this work this simple approach is only exploited to compare the obtained RMS trends under different conditions and understand its possible employment in REB prognostics, without attempting to provide any formulae to compute the RUL.

The proposed procedure is articulated as follows. The EDV method is applied on the entire degradation history of a previously known RMS trend under a certain working conditions in order to estimate associated values of $\Delta\varphi_d$ and h_d at each time. Then, a new degradation history is computed under different load and/or rotation frequency. This is done by computing the new RMS value by exploiting the LP model of the faulty bearing. To this end, at each instant the corresponding values of defect dimensions are employed to determine the system vibration. Eventually, the generated history is compared with the observed RMS trend under the actual working condition, which is different from the reference one. As previously mentioned, this procedure is still under development and should be considered as a first attempt to tackle the problem.

5.2.3.2 *The second PBM: time-to-threshold assessment by estimation of the future values of the equivalent damaged volume*

The goal of the second PBM is to predict the propagation of the equivalent defect that develops in the bearing during the final stage of its life. However, rather than computing the RUL, this approach is applied in order to estimate the time between the current observation time and the instant in which V_{eq} crosses a certain threshold V_{th}. Thus, this parameter is named Time-to-Threshold (TT) and it is defined as:

$$TT = t_{th} - t_c \tag{92}$$

where t_{th} is the instant in which V_{eq} crosses the threshold value V_{th} and t_c represents the current instant in which the TT is computed. Thus, the challenge it to be able to compute t_{th} at a generic time t_c during the final stage of bearing life. This model is based on the assumption that the damage propagation history between t_{fs} and t_{th} is governed by a propagation law in the form:

$$V_{eq} = V_{eq}(t) \quad \text{if } t_{fs} \leqslant t \leqslant t_{th} \tag{93}$$

The function relating t to V_{eq} is determined on the basis of V_{eq} values estimated from the experimental measurements. The experimental t-V_{eq} curves are fitted through polynomial of order two to determine the function that better fits the data. If this relationship provides satisfactory results for different cases, then this means that it may applied to other working conditions, under the hypothesis that the damage propagation follows the same relationship under every operative condition. Therefore, during operation t_{th} may be estimated by fitting the available data up to the current time with a polynomial of the adequate order as observed from experiments under other conditions.

The proposed PBM is formulated according to the following steps. First, a relationship relating time t and the equivalent damaged volume V_{eq} below a certain threshold V_{th} is sought by analysing the V_{eq} values estimated from the available experimental observations. In this work, data are fitted with polynomials of order two, as it will be seen from the analysis of the experimental results in Section 5.4.3. Furthermore, it is imposed that the polynomial must pass through the first equivalent volume estimated at time t_{fs} to ensure that

the best-fitting curve passes through the first point of the degradation history. However, other type of curves, e.g higher-order polynomials or exponentials, could be employed if the polynomials do not provide an acceptable description of the damage evolution in the case under exam. Then, the prevision model is applied for the on-line prediction of the TT. During operation, the monitored bearing will remain in the healthy state for the majority of its working life. Once the bearing enters in the final stage, the EDV algorithm may be utilised to determine the progression of the equivalent damaged volume up to the current time. Then, t_{th} may be predicted by fitting the available points up to t_c with the previously determined relationship between V_{eq} and t. Finally, the TT may be calculated through Eq. 92.

It is worth mentioning that the initial moment in which the bearing transitions from the healthy to the final stage must be detected by means of additional tools. For this purpose, diagnostic indicators may be employed to ensure an accurate evaluation of t_{fs}.

5.3 RESULTS OF THE RUN-TO-FAILURE TESTS

This section discusses the results of the experimental run-to-failure tests described in Section 2.3.2, whose values of applied load, shaft speed and the total time length have been reported in Table 3. The resulting RMS histories of the six run-to-failure tests are shown in Figure 57. In particular, Figure 57.a, Figure 57.b and Figure 57.c represent the tests carried out under a radial load value equal to 4 kN, namely tests E1, E2 and E3, respectively. Furthermore, Figure 57.d, Figure 57.e and Figure 57.f refer to the tests E4, E5 and E6 which are characterised by applied loads equal to 3 kN, 4.7 kN and 5 kN, respectively. As detailed in Section 2.3.2, the tests have been stopped when the peak amplitude of the acceleration in the time signal reached 20 g. Therefore, the total time t_{EOL} of each test corresponds to the final point of each RMS history, as shown in Figure 54.

On the other hand, t_{fs} represents the moment in which the bearing transitions to the final stage of its life. This HS shift may be seen by the steep rise of the RMS values in the final part of each degradation history. This phenomenon is

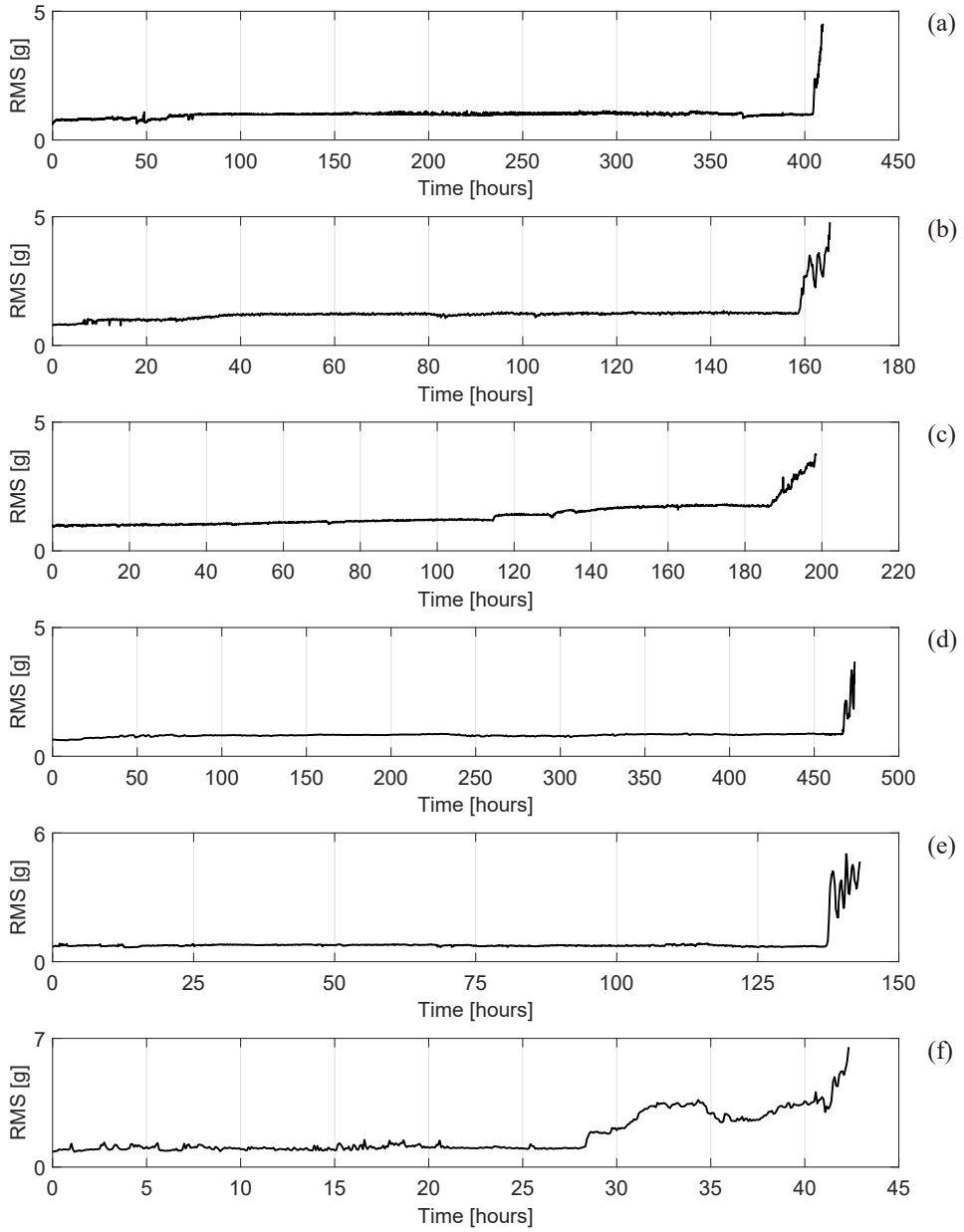

Figure 57: Experimental RMS trends for the entire test length: (a) E1 (4 kN); (b) E2 (4 kN); (c) E3 (4 kN); (d) E4 (3 kN); (e) E5 (4.7 kN); (f) E6 (5 kN).

particularly evident for tests E1, E2, E4 and E5. In fact, the RMS value remains

approximately constant for the majority of the test, then it greatly increases in a short period of time. After this rapid increment, tests E2, E4 and E5 show an oscillating trend with several increases and decreases. This phenomenon may be observed in Figure 54.b, Figure 54.d and Figure 54.e, respectively. Besides, Figure 54.a highlights that, during test E1, the RMS trend briefly decreases and then starts to increase again. Moreover, tests E3 and E6 show an even different behaviour. Test E3, which is depicted in Figure 57.c, experiences a slow rise of its RMS value even before the beginning of the final stage. Then, the degradation consistently increases during the final stage, although with a less steep increase compared to the other tests. Differently, Figure 57.f shows that during test E6 a consistent increase in RMS values is detected at the beginning of the final stage. However, the values descended after a few hours, only to rapidly rise again and eventually bring an halt to the test.

Indeed, the value t_{fs} for each test may be evaluated directly from the RMS history. However, this task may be arduous, because a defect may induce a slow increase of the RMS over time rather than a fast one. Moreover, the RMS increase is not necessarily related to fault initiation. This is particularly evident in test E3, which is shown in Figure 57.c. In fact, in this case the RMS increment starts way before the beginning of the final stage, as it will be demonstrated later in this section. Therefore, in a on-line monitoring scenario, it is suggested to employ diagnostic indicators to efficiently detect the time in which the bearing life transitions to the final stage. Since the discussion of diagnostic indicators for REB monitoring goes beyond the scope of the present dissertation, a simple but effective indicator has been employed in this work. This indicator is named H_{or} and it is defined as the sum of the first five harmonics in the envelope spectrum of the signal. To compute this quantity for each acquired sample, the signal is first band-passed in the high frequency resonance zone. This frequency interval has been characterized in the previous experimental campaign regarding the stationary tests with artificial defects. According to the findings reported in Section 4.3, this peculiar resonance zone is found between 7 and 9 kHz. As a consequence, the signal is filtered in this frequency interval. Then, the envelope of the resulting signal is computed and transformed into the frequency domain. The envelope of the signal is evaluated by first determining the analytic signal $Z_a(t)$, i.e.:

$$Z_a(t) = Z(t) + j\mathcal{H}(Z(t)) \tag{94}$$

where $Z(t)$ is the filtered acceleration signal, j is the imaginary unit and $\mathcal{H}(Z(t))$ is the Hilbert transform of $Z(t)$ [238]. Then, the envelope is the modulus of $Z_a(t)$, that is:

$$env(Z(t)) = \sqrt{Z^2(t) + \mathcal{H}(Z(t))} \tag{95}$$

Finally, H_{or} is computed as:

$$H_{or} = \sum_{k=1}^{5} A_{BPFOe,k} \tag{96}$$

where $A_{BPFOe,k}$ is the k-th harmonic of the BPFO in the envelope spectrum. This quantity differs from parameter A_{BPFO} introduced in Section 4.2.3 since the latter is computed on the raw signal spectrum, while A_{BPFOe} refers to the envelope spectrum of the filtered signal. It is worth mentioning that similar indicators could be also introduced for the other bearing characteristic frequencies related to other fault types. However, only the described indicator is introduced in order to demonstrate its capabilities in detecting the defects which typically appear in the experimental test rig described in Chapter 2. The experimental values of H_{or} are reported in Figure 58. Similarly to RMS, H_{or} rapidly increases at the beginning of the final stage. Notably, this indicator is able to assess the transition for test E3 in a more clear manner compared to the RMS value. In this regard, Figure 59 shows the envelope spectra of two signals. While Figure 59.a shows the envelope spectrum in the first part of bearing life, Figure 59.b depicts the typical spectrum in the final stage. The difference between the two is remarkable, since in Figure 59.b the BPFO harmonics may be distinctly observed in the spectrum. As a result, the employed indicator provides an efficient way to estimate t_{fs}. To summarize the content of this section, Table 20 lists the values of t_{fs}, t_{EOL} and the length of the final stage for each test, i.e. the RUL from the start of the final stage:

$$RUL(t_{fs}) = t_{EOL} - t_{fs} \tag{97}$$

The final stage of each test is also depicted in Figure 60.

5.4 REB PROGNOSTICS BASED ON EDV ESTIMATION

This section describes the application of the methods described in Section 5.2 through the employment of the experimental data reported in Section 5.3.

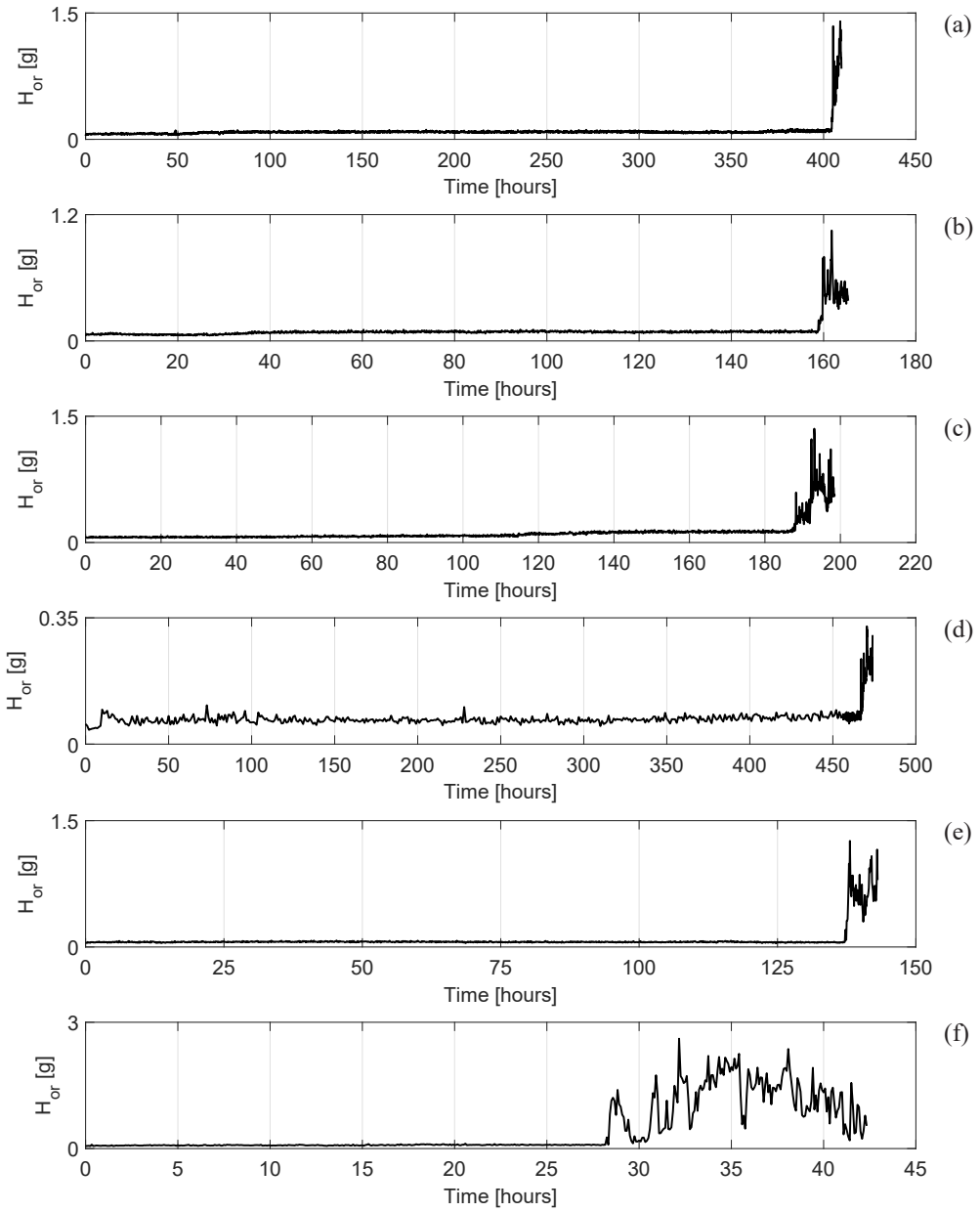

Figure 58: Experimental values of indicator H_{or}: (a) E1 (4 kN); (b) E2 (4 kN); (c) E3 (4 kN); (d) E4 (3 kN); (e) E5 (4.7 kN); (f) E6 (5 kN).

Firstly, V_{eq} values are determined by exploiting the EDV algorithm. Then, this parameter is utilized as input for two prognostic models.

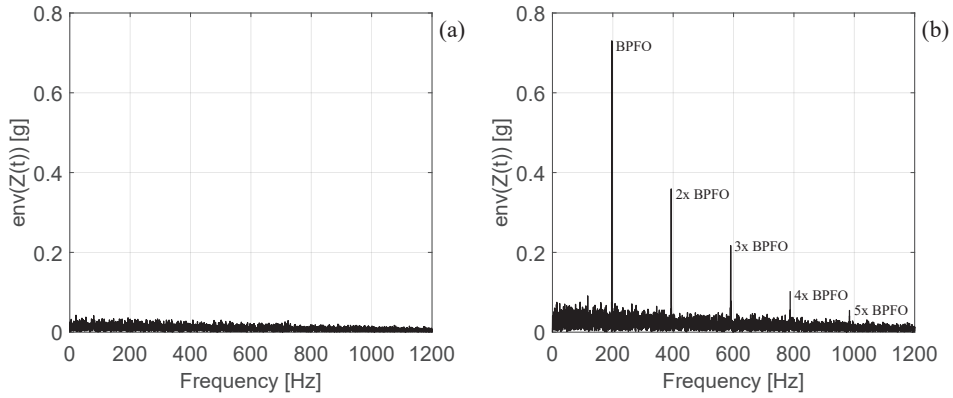

Figure 59: Example of envelope spectra taken from test E3: (a) Normal operation; (b) Final stage.

Table 20: t_{fs}, t_{EOL} and length of the final stage for each test.

ID	t_{fs} [h:min]	t_{EOL} [h:min]	RUL(t_{fs}) [min]
E1	404:15	409:40	325
E2	158:25	165:20	415
E3	186:20	198:30	730
E4	467:00	473:55	415
E5	137:00	143:05	365
E6	28:20	42:25	845

5.4.1 *EDV assessment from test bench data*

The EDV algorithm, which has been described in Section 5.2.2, is applied to the acceleration data acquired during the run-to-failure tests. Thus, the goal is to evaluate the equivalent damaged volume V_{eq} from the final stage of the six degradation histories shown in Figure 60.a and Figure 60.b. To evaluate V_{eq} values associated to each test, it is necessary to compute the RMS maps associated to each operative condition listed in Table 3. Therefore, four maps are generated. Each one is evaluated at the same rotation frequency of 40 Hz, but at four different load values, i.e. 3000 N, 4000 N, 4700 N and 5000 N. This task is performed by running several simulations with the numerical model detailed

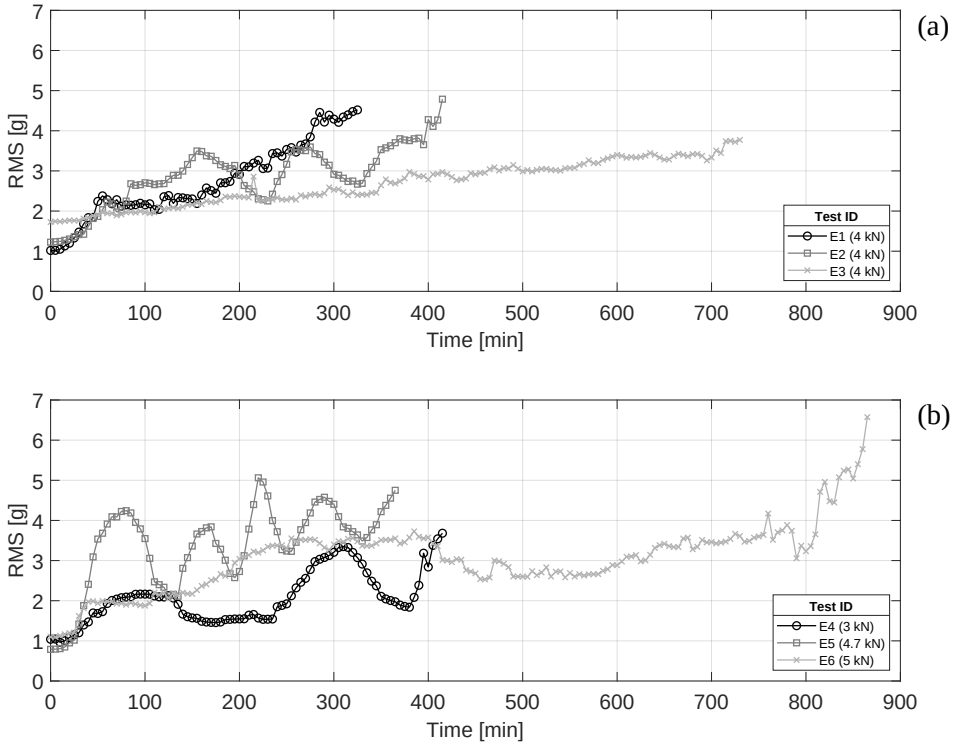

Figure 60: RMS in the final stage: (a) Tests E1, E2 and E3; (b) Tests E4, E5 and E6.

in Section 5.2.1. The parameters of the dynamic model are the ones obtained through the multi-objective optimization technique described in Chapter 4. As a consequence, the model parameters employed in the simulations are the ones reported in Table 18. For a given applied load, the RMS value is evaluated for several combinations of $\Delta\varphi_d$ and h_d. The angular extent is limited from 0 to 50 degrees, since this is approximately the maximum value observed in the experiments at the end of bearing life. On the other hand, the defect depth has been capped at 32 μm since prior simulations showed that RMS values were mostly unaffected for larger depths. Finally, surface roughness R is set to 10 μm, which is a common value employed to characterize a rough surface [115]. The four resulting RMS maps are shown in Figure 61. In particular, Figure 61.a, Figure 61.b, Figure 61.c and Figure 61.d represent the combinations of RMS values at applied loads equal to 3000 N, 4000 N, 4700

N and 5000 N. respectively. These plots highlight that, as expected, the RMS values increase with higher loads. Furthermore, larger loads lead to a less smooth distribution of RMS values in the associated map, as it may be notably observed by comparing Figure 61.a and Figure 61.d. Besides, it is interesting to note that the RMS does not necessarily increase with higher defect dimensions. Rather, each combination provides a different RMS value, which may be lower for increasing values of angular extent and/or defect depths compared to lower fault dimensions. Nevertheless, the generally observed trend is the following. The RMS remains quite stable, regardless of the defect extent, for low profile depths. For deeper defects, the RMS values start to increase and its trend becomes less predictable as each combination of defect parameters leads to different RMS values on the basis of the interaction between $\Delta\varphi_d$ and h_d. In all cases, however, it is observed that the highest RMS values are always obtained in the upper right-most part of the map, i.e. for $\Delta\varphi_d$ and h_d pairs characterized by the highest values. This is in accordance with the experimental evidence that the system vibration greatly increases with large extended defects.

After computing the RMS maps, the EDV algorithm is applied to the acceleration signals acquired on the bearing test rig during the run-to-failure tests. First, the smoothness of the experimental RMS trends shown in Figure 60 is enhanced by utilizing a moving average with a sliding window of length equal to 6 data points. Then, the procedure detailed in Section 5.2.2 is applied to extract the associated V_{eq} values over time. These are reported in Figure 62.a for tests E1, E2 a and E3 and in Figure 62.b for tests E4, E5, and E6, respectively. The results show that in each test V_{eq} maintains a "low" value, approximately lower than about 0.05 to 0.01 mm^3/mm, for the majority of the duration of the final stage. Then, V_{eq} rapidly increases at the very end of the bearing life, when failure is imminent. This result is interesting for multiple reasons. First, based on the experimental observations, this peculiar trend allows to detect the end of bearing life. In fact, a rapid increase of this quantity warns about the proximity of bearing failure. Therefore, this indicator may be also employed for the diagnosis of fast-approaching failure during the final stage of the bearing life. Moreover, all V_{eq} histories seem to follow a common trend. As a result, the possibility to fit the data to some kind of fitting functions is investigated with the aim of developing prognostics procedures. It is also worth noting that the volume does never decrease over time, with the only

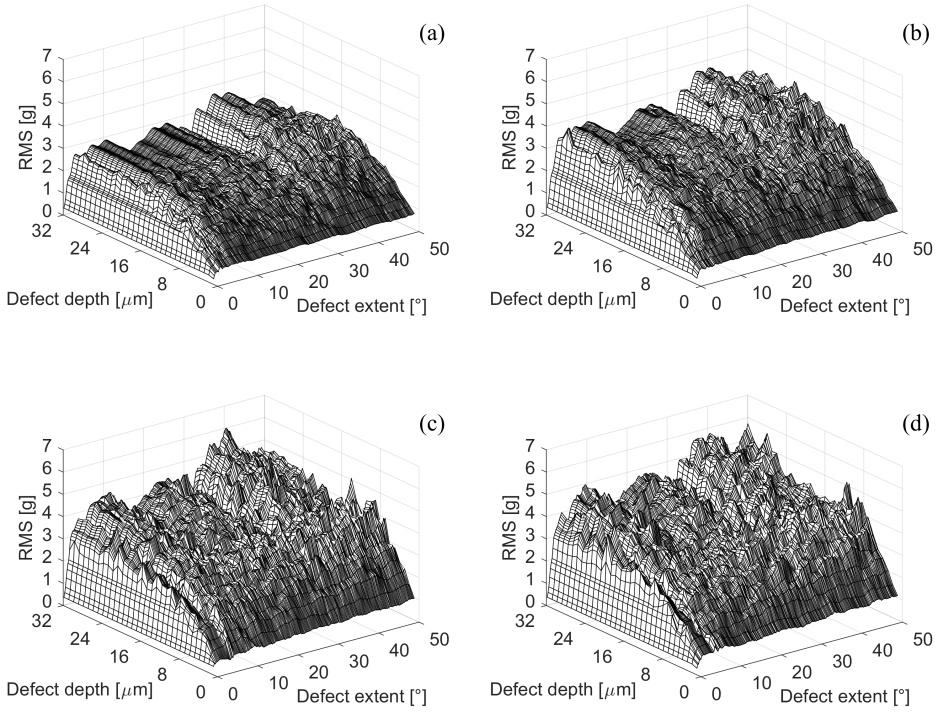

Figure 61: RMS maps for a constant rotation speed equal to 40 Hz and different values of applied load: (a) 3000 N; (b) 4000 N; (c) 4700 N; (d) 5000 N.

exception of the final part of test E4. In fact, with reference to Figure 56, in the EDV algorithm it was inserted the possibility, at each instant, to expand the search range to points which were previously removed from the RMS map. This hypothesis was added into the algorithm in order to ensure its capability to always describe the entire experimental RMS trend, as detailed in Section 5.2.2. However, the results of the EDV algorithm show that this phenomenon does not usually take place, except in one test, and only at the very end of the test itself. Finally, the fact that V_{eq} mostly shows an increasing trend is an aspect of major interest. Indeed, V_{eq} is denoted by an increasing trend, but it is extracted by a non-increasing trend, i.e. the RMS, which in contrast may increase or decrease over time, as seen in Figure 60. This is another impor-

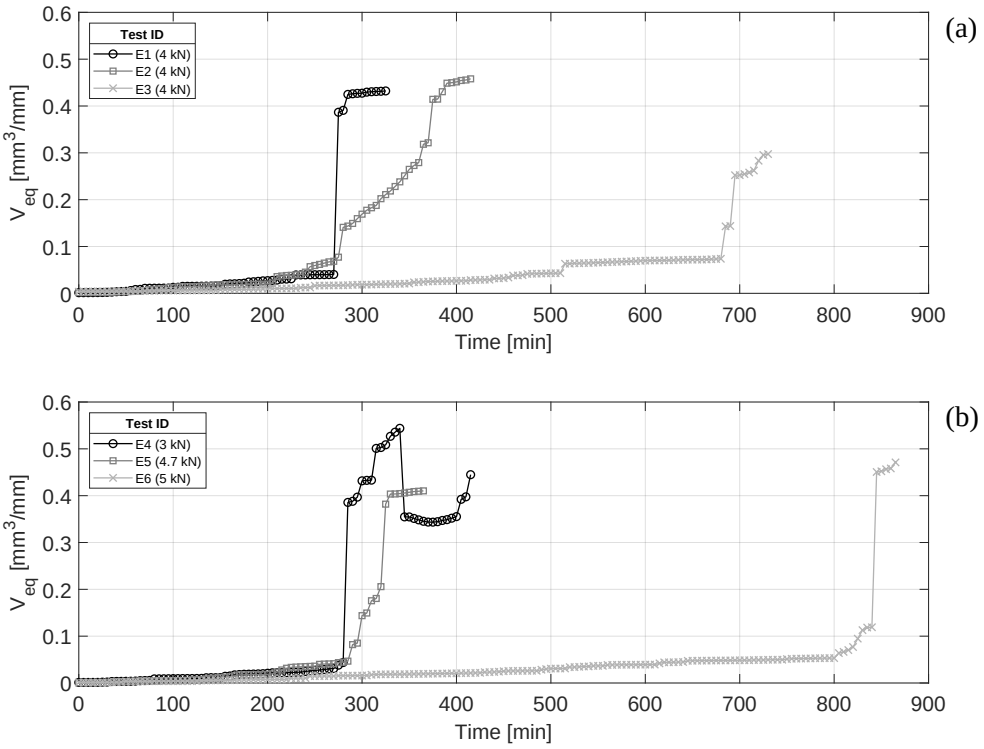

Figure 62: Equivalent damaged volume estimated through the EDV algorithm: (a) Tests E1, E2 and E3; (b) Tests E4, E5 and E6.

tant matter that makes the equivalent damaged volume V_{eq} a very promising parameter for further employment in prognostic models.

5.4.2 *Application of the first PBM: prediction of degradation histories*

The first of the two proposed PBM is a preliminary attempt at providing a prognostic model based on the EDV algorithm. The full procedure at the foundation of this PBM has been described in Section 5.2.3.1. The outputs of the model are the degradation histories at different operative conditions with respect to the reference one. Within this context, the reference histories are associated with the three tests carried out under a radial load equal to

4000 N. Therefore, the pairs ($\Delta\varphi_d$, h_d) related to each point in Figure 62 are used to generate degradation histories at the rotation speed of 40 Hz but under radial load values equal to 3000 N, 4700 and 5000 N, respectively. Consequently, three numerical histories are generated for each load. At each time, the RMS is evaluated by performing a simulation in which defect depth and angular extent are based on the pair ($\Delta\varphi_d$, h_d) computed through the EDV algorithm. The resulting deterioration histories are reported in Figure 63. Specifically, Figure 63.a depicts the numerical and experimental trends at 3000 N, while Figure 63.b and Figure 63.c represent the trends at 4700 N and 5000 N, respectively. As it may be observed from the plots, the numerically generated degradation histories have the same trends as the original experimental ones from which they are derived, but they are scaled on the basis of the new load value. However, as expected, this approach may not be directly utilized for prognostics, as the trends generated in this way will always have the same t_{EOL}. Regardless, it is possible to compare the numerical histories with the real experimental trends. As of now, only three tests at different loads are available. A complete discussion for this model should be performed on the basis of an high number of tests, in order to provide an insightful comparison between numerical and experimental data. For instance, consider Figure 63.c, which compares the RMS trends at 5000 N. The experimental trend is very similar to the numerical trend produced by the degradation history E2 until about 250 minutes from the beginning of the final stage. Therefore, the model developed in this way is not able to capture the different trends over time due to the limited amount of tests. Nevertheless, some useful information may still be obtained by employing this first proposed model. Interestingly, in fact, the RMS value at t_{EOL} is very similar between the experiments and the numerical results. While this information may not be employed for prognostics, since information about the time of failure is missing, it allows to determine which acceleration value is expected to be reached at the end of the test. The RMS values at life end are reported in Table 21. At the lowest load of 3000 N, the final RMS is lower for all numerical trends. At 4700 N, the values are closer, especially for the history generated from the experimental test E2. Finally, results are more dissimilar at 5000 N, as the model underestimates the ending RMS value in all instances.

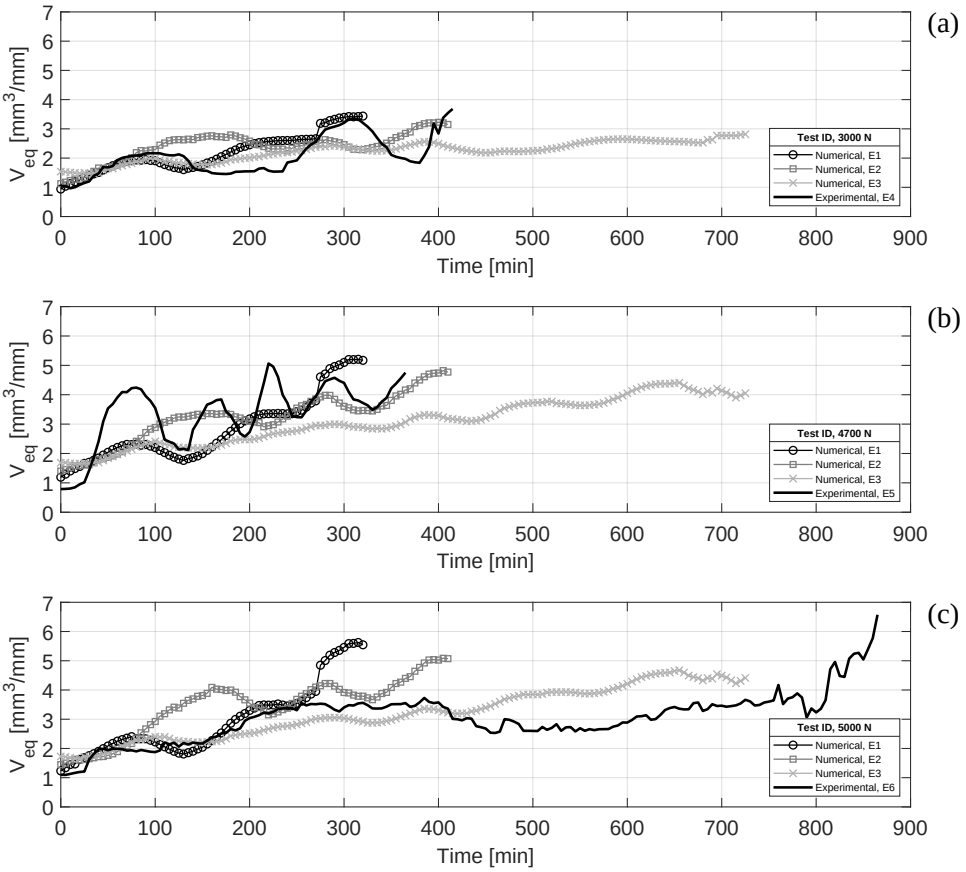

Figure 63: Comparison between numerically generated and experimental degradation histories, for different load values: (a) 3000 N; (b) 4700 N; (c) 5000 N.

5.4.3 *Application of the second PBM: TT assessment on actual vibration data*

The second model based on the EDV method involves the prediction of the TT between the actual time and a pre-determined threshold V_{th}. While this is not

Table 21: RMS values at the end of each test as estimated experimentally and through the numerical model.

	Load value		
Test ID	3000 N	4700 N	5000 N
Numerical, E1 [g]	3.4	5.2	5.5
Numerical, E2 [g]	3.2	4.8	5.1
Numerical, E3 [g]	2.8	4.1	4.4
Experimental [g]	3.7	4.8	6.6

a direct measure of the RUL, Figure 62 shows that the last part of bearing life is always characterized by a steep increase in the value of V_{eq}. Hence, by setting an adequate value of V_{th}, it should be possible to determine the remaining time before the equivalent damaged volume crosses a "critical" threshold that, in terms of time, is very close to the end of life of the component. As a result, the procedure described in Section 5.2.3.2 is applied to V_{eq} values retrieved from the experimental measurements. The efficiency of the proposed method is then evaluated by computing the TT over time for all tests depicted in Figure 62.

Firstly, V_{eq} trends are truncated at a threshold V_{th}. For the purpose of this study, a value $V_{th} = 0.05\ mm^3/mm$ is chosen on the basis of the trends observed in Figure 62. However, this choice is arbitrary. As a consequence, different threshold values might be employed in further investigations. The resulting V_{eq} histories are reported in Figure 66.a for tests E1, E2 and E3 and in Figure 66.b for tests E4, E5 and E6, respectively. In each degradation trend, the final point on the threshold is obtained by linear interpolation between the points immediately before and after V_{th}.

Subsequently, the curves in Figure 66 are fitted with a polynomial of order two. The fitted curves for each test are shown in Figure 65. From this figures, it is possible to observe that the fitting curves are able to provide a good fitting of the data. This is evaluated by computing the coefficient of determination, namely R^2, for each curve. This parameter, in fact, permits to evaluate the

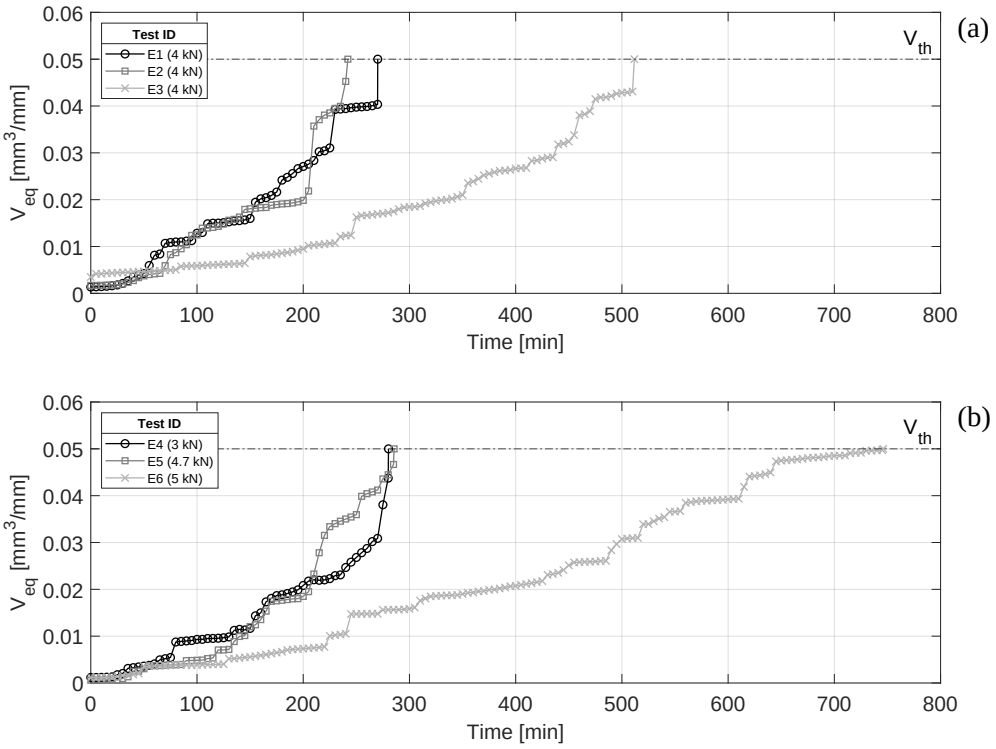

Figure 64: Values of V_{eq} below the threshold V_{th} = 0.05 mm³/mm: (a) Tests E1, E2 and E3; (b) Tests E4, E5 and E6.

goodness-of-fit of the data [239]. The values of R^2 are reported in Table 22 alongside the coefficients a, b and c of the corresponding polynomial $ax^2 + bx + c$. The R^2 values are always higher than 0.92 for each curve, thus denoting a good fitting of the data in all cases. The main downside of this approach is that the fitting curve may overestimate the TT if the values of V_{eq} begin to rapidly increase before the damaged volume reaches the chosen value of V_{th}. This is particularly evident in test E4, which is shown in Figure 65.d. In fact, the gradient of the degradation trend consistently increases even below 0.04 mm³/mm. However, the polynomial fitting behaves sufficiently well for the other cases considered in this study. Therefore, a polynomial of order two is employed for the TT prediction in the next step of the proposed procedure. Even in this case, it is worth underlining that further inquiry on the best-fitting

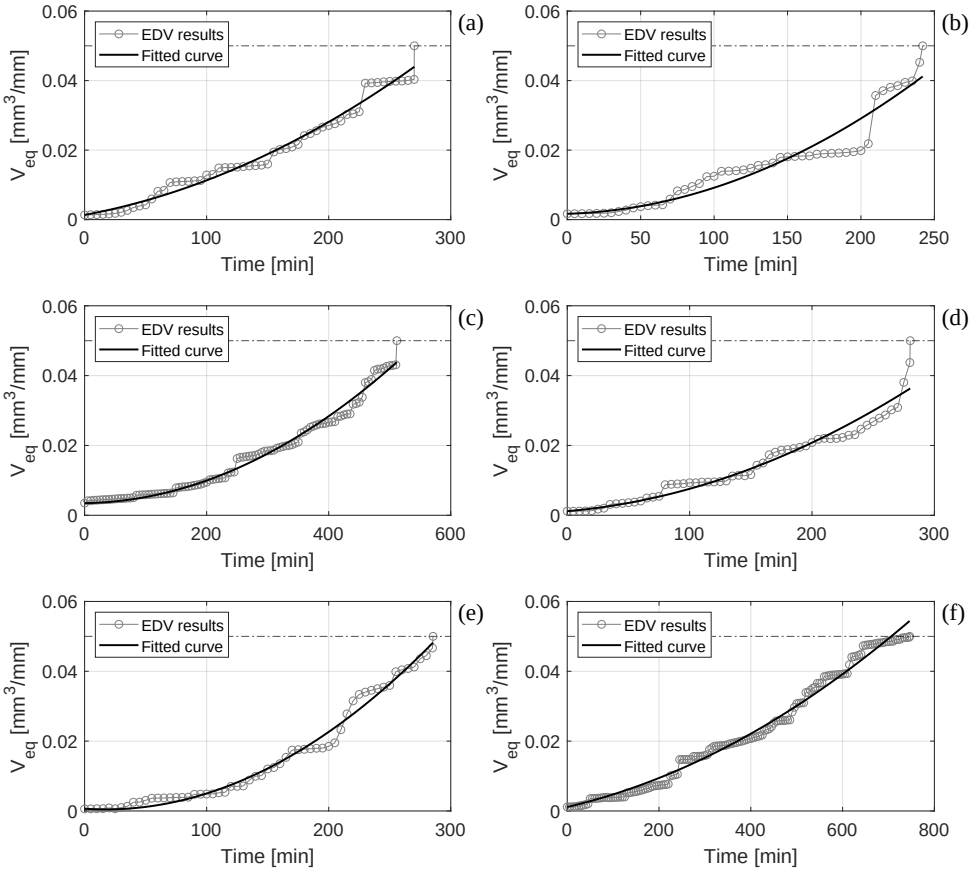

Figure 65: V_{eq} values below threshold fitted with a quadratic polynomial: (a) E1 (4 kN); (b) E2 (4 kN); (c) E3 (4 kN); (d) E4 (3 kN); (e) E5 (4.7 kN); (f) E6 (5 kN).

options and the choice of threshold values for TT assessment may be proposed in future investigations to improve the efficiency of the developed methodology.

Finally, the TT estimation method described in Section 5.2.3.2 is applied to the examined experimental tests. The technique is applied as follows. For each test, TT is predicted at each instant by fitting the estimated V_{eq} up to the considered time with a quadratic polynomial. Subsequently, TT is computed by

Table 22: Values of coefficients a, b and c of the polynomial $ax^2 + bx + c$ and goodness-of-fit parameter R^2 for each test.

Test ID	a (x 10^{-7})	b (x 10^{-5})	c (x 10^{-3})	R^2
E1	3.42	6.53	1.35	0.979
E2	6.22	1.28	1.68	0.921
E3	1.49	0.25	3.48	0.986
E4	3.41	2.99	1.18	0.941
E5	6.62	2.24	0.63	0.987
E6	0.55	3.05	1.14	0.989

evaluating the time t_{th} in which the fitted curve crosses the volume threshold V_{th}. This leads to a different TT estimation at each instant, since more data points are available for fitting as the test time progresses. TT values estimated through this procedure are shown in Figure 66. In each plot, the predicted TT is compared with the actual TT, which is a straight decreasing line with a slope equal to -1. Therefore, the prediction accuracy of the method is evaluated by comparing the predicted values with this reference line: the quality of the estimate is higher if the predicted points are closer to the actual TT. Figure 66 shows that the same behaviour is observed for all curves. In fact, the predicted TT consistently differ from the actual TT at the beginning of the final stage. Later, when more points become available, the quality of the fitting increases. Predicted values are more accurate for certain tests, in particular for tests E2, E3, E4 and E5, which are shown in Figure 66.b, Figure 66.c, Figure 66.d and 66.e, respectively. In these cases, the predicted points all fall around the actual TT line after a certain number of test points. This behaviour is due to the different slopes of the real curve during the degradation process. In fact, the fitting curve detects an accurate t_{th} if the damage progression speed is comparable with the actual velocity up to the considered data sample. In other words, a steep increase of the equivalent damaged volume at very beginning of the final stage may lead to extremely low values of TT, e.g. as in test E3 plotted in Figure 66.c. In this test the initial TT estimates are remarkably far from the actual ones, but they are sensibly closer to the real values after approximately 150 minutes. In fact, as the degradation progresses, the curve stabilizes and

converges to the actual data distribution. This becomes more evident the further the degradation evolution has proceeded, has it may be seen from the considered plots. On the other hand, Figure 66.a and Figure 66.e show that for tests E1 and E6 the estimated TT is always lower than the real one. While the estimation for E1 is denoted by low values that follow the real distribution, results are considerably worse for test E6. Indeed, Figure 66.e shows that test E6 produces the worst results with this method. This may imputable to the incorrect choice of threshold values or fitting function for this particular test, which instead provided satisfactory results in all other cases. This aspect is investigated by testing different threshold values and subsequently estimating the associate TT for test E6. The results of this additional study are depicted in Figure 67.a, Figure 67.b and Figure 67.c for threshold values equal to 0.03 mm^3/mm, 0.08 mm^3/mm and 0.1 mm^3/mm, respectively. By comparing these trends with the results obtained for V_{th} = 0.05 mm^3/mm, it is possible to examine the differences introduced by the choice of another threshold value. By employing the lowest V_{eq}, the predicted trend is closer to the expected one. Differently, the predicted trend change for the other two values. In particular, by changing from 0.08 to 0.1 mm^3/mm the curve shifts towards higher TT values.

In conclusion, the chosen threshold may greatly affect the outcome of the model. Although good estimates are obtained with a value of V_{eq} set to 0.05 mm^3/mm, different and possibly better results were obtained by utilizing diverse values in test E6. Moreover, other kinds of fitting functions should be investigated in further inquiries. In spite of this, the proposed methodology already shows a good potential for application in REB prognostics. In fact, it is worth underlining that PBM models may be applied to several operative conditions, since they are based on the physics of the problem rather than on previously acquired data on the system. As a result, the two proposed PBMs demonstrate the capabilities of the EDV method within the context of bearing prognostics.

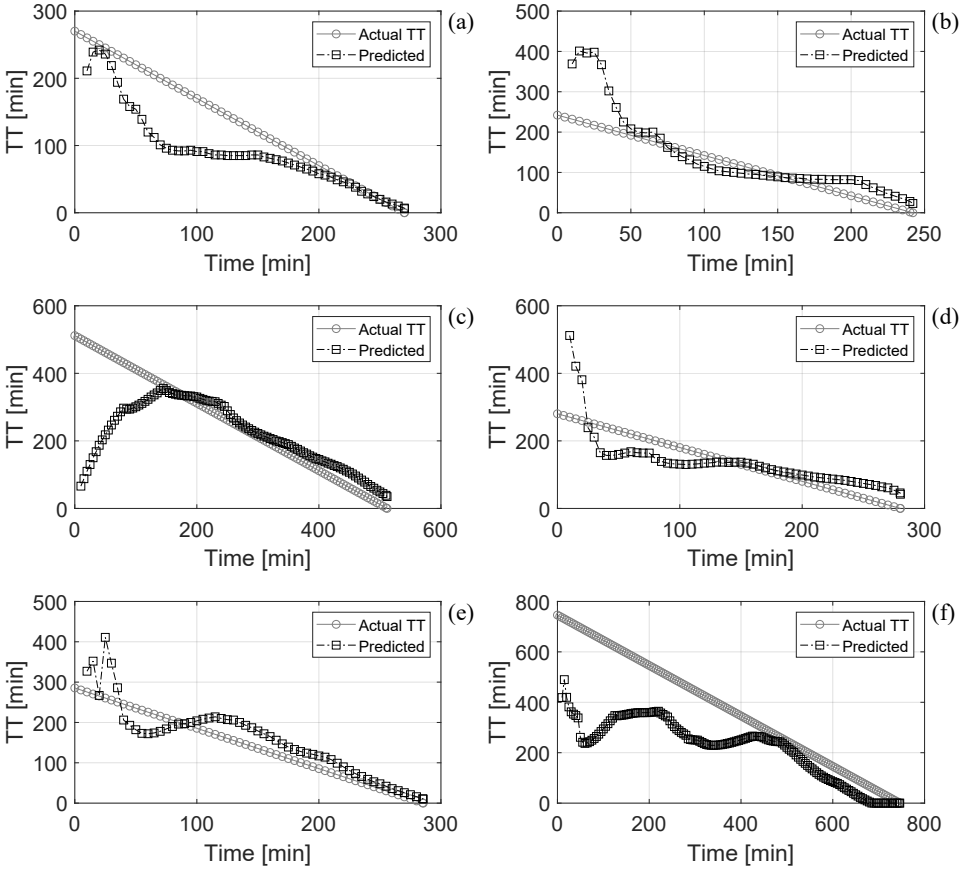

Figure 66: TT computed through the proposed method: (a) E1 (4 kN); (b) E2 (4 kN); (c) E3 (4 kN); (d) E4 (3 kN); (e) E5 (4.7 kN); (f) E6 (5 kN).

5.5 SUMMARIZING REMARKS

This chapter introduced a novel parameter related to the evolution of defect dimensions during the final stage of bearing life. This quantity, named equivalent damaged volume, was exploited to propose two physics-based prognostics models for the assessment of rolling bearing health state.

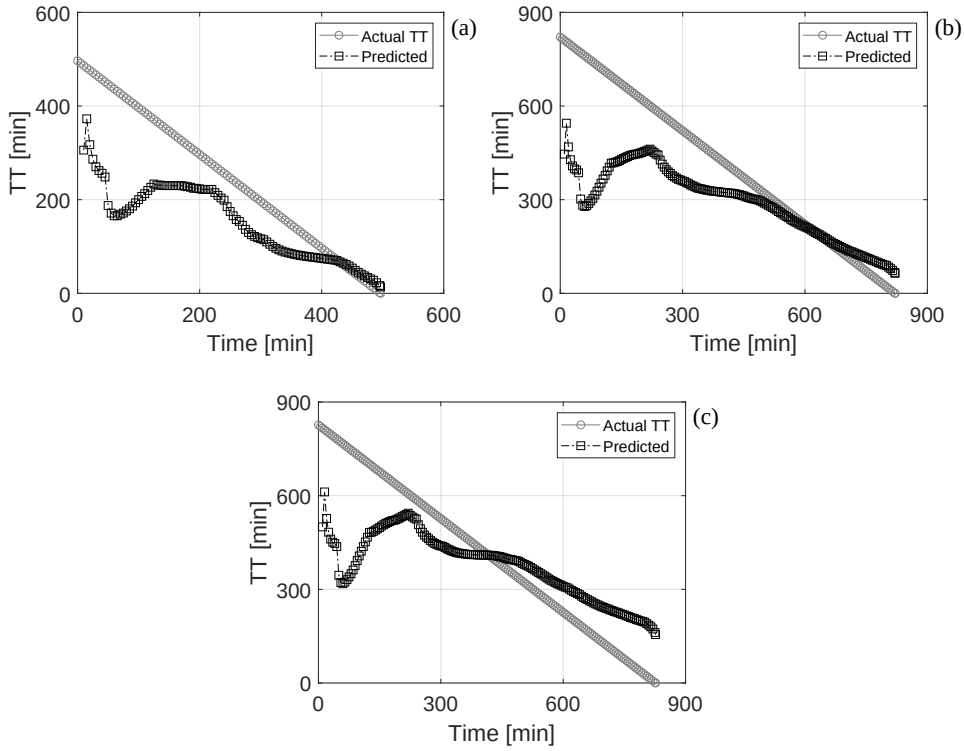

Figure 67: TT computed for test E6 for various threshold values: (a) V_{th} = 0.03 mm^3/mm; (b) V_{th} = 0.08 mm^3/mm; (c) V_{th} = 0.1 mm^3/mm.

The faulty bearing dynamic model, which was previously described in Section 4.2.2, was extended to also consider extended defects. This task was performed in order to replicate the experimental evidence that the defects produced during the experimental campaign where all characterized by a large angular extent at the end of bearing life. Thus, to account for this aspect in the model, the ball path trajectory inside the defect was modified by introducing the possibility of contact with the bottom of the defect, which was modelled as a rough surface. Besides, the model parameters were tuned according to the procedure to determine their unknown values described in Chapter 4.

Then, the algorithm that allows to compute the EDV values was detailed. The proposed algorithm takes as input the numerical RMS map computed at the

experimental working condition and the experimental RMS history in the final stage of the bearing life. The RMS map is a surface in which, for a given load and shaft rotation speed, at each point corresponds a different combination of the dimensions that characterize the modelled defect, i.e. angular extent and depth. The numerical RMS values were computed through the enhanced LP model, which is able to consider both dimensions. The output of the algorithm is the equivalent damaged volume per unit of width V_{eq} and the associated defect dimensions $\Delta\varphi_d$ and h_d for each point extracted from the experimental RMS history.

The EDV algorithm was then exploited to propose two potential PBM prognostic models. The first PBM aimed at the prediction of RMS trends under different conditions with respect to a reference degradation history. Due to the limited number of available test data, this methodology was only employed to compare the obtained histories and to perform a rough quantitative estimate on the RMS values at end of the bearing life. On the contrary, the objective of the second PBM was to predict the propagation of the equivalent defect that develops in the bearing during the final stage of its life. This approach was applied in order to evaluate the time-to-threshold, i.e. the time between the actual time and the instant in which V_{eq} crosses a pre-defined threshold. The prediction was based on the assumption that the damage evolution follows a quadratic polynomial function over time. This was verified by fitting V_{eq} values with the assumed fitting functions. Then, the method was applied to determine TT values during operation by fitting the available data points up to the actual time with the chosen polynomial. The efficiency of the method was demonstrated by estimating the time-to-threshold on signals acquired during the run-to-failure experiments. However, the main downside of this technique is that it provides unrealistic estimate at the beginning of the final stage due to the low number of sample points. However, as more data become available during bearing operation, the quality of the fitting increases and so does the accuracy of the estimate. Moreover, different fitting functions and other threshold values shall be further investigated in future works to enhance the capabilities of the proposed PBM. Despite the fact that the work carried out in this chapter should be only considered a first approach to the problem of bearing prognostics, it provides an original contribution on the subject. In fact, it introduces two modelling approaches based on an indicator, i.e. the

equivalent damaged volume, that is calculated by means of a dynamic model of a faulty bearing with localised or extended defects. Indeed, this is the main aspect of novelty of the detailed study.

Finally, to conclude the chapter, it is worth underlining some relevant aspects of the EDV algorithm and its potential further developments. In this regard, this endeavour only dealt with outer races defects. This was based on the observation that the defect observed in the experimental tests were all located in the outer raceways. However, in general, defects may also develop in the other components of the bearing, i.e. inner ring, rolling elements and cage. The EDV algorithm may be also applied to these type of damages, provided that the RMS maps are computed by modifying the LP model accordingly. Besides, the EDV algorithm might be also employed to compute V_{eq} values under varying operative conditions. However, in this case, the complexity of the procedure greatly increases. In fact, several RMS maps must be generated, each one for a different combination of load and speed. It is evident that a similar approach requires a large initial computational effort to generate all the required RMS maps. Nevertheless, for a given system, this process must be performed only the first time, since for subsequent EDV analyses the RMS maps will be already available. Notably, it is not possible to generate a RMS map for every operative condition, since these maps may only be defined at discrete load or shaft speed values. Consequently, values in-between must be somehow extracted through interpolation between the RMS values estimated from two maps instead of one. Finally, it is also worth emphasizing that the main characteristic of prognostic models based on the EDV algorithm is that, by being physics-based, they may be employed for any given combination of load and shaft speed values. Therefore, these models show a remarkable potential for employment in real-case scenarios.

6

FINAL REMARKS

This thesis discussed a number of numerical techniques that may be employed to address several aspects related to the modelling of rolling-element bearings. In particular, the dissertation covered the following topics:

- The estimation of bearing stiffness values through the finite-element method, by proposing a computationally efficient methodology for the generation of load-dependent meshes.
- The dynamic LP modelling of rolling bearings with localised defects, by establishing a procedure for the estimation of unknown model parameters.
- The development of physics-based prognostic models, through the introduction of a degradation-related indicator based on a LP model of a defective bearing.

All these subjects where thoroughly analysed in a dedicated chapter each. Results were validated either against analytical models retrieved from the literature or from actual experimental data acquired on a specifically-designed test bench.

Firstly, Chapter 1 introduced the arguments of the thesis and provided an exhaustive literature review on the state-of-the-art in bearing modelling. The survey highlighted the major aspects of interest from the massive amount of works which discuss about the modelling of these widely-employed mechanical components. Moreover, the chapter outlined the research objectives

and illustrated the organization of the thesis. Then, Chapter 2 described the bearing test bench that was set up at the Engineering Department of the University of Ferrara. The test rig was employed to perform two type of tests: stationary tests for bearing with artificial localised defects and run-to-failure tests. Concerning the former, the tests involved the acquisition of acceleration signals of defective bearings under several combinations of values of applied load, shaft rotation speed and defect dimension. To this end, artificial defects were seeded on the outer raceway of a number of self-aligning ball bearings through electro-discharge machining. Differently, run-to-failure tests regarded the acquisition of bearing vibration signals throughout the entire life of the components. In this regard, accelerated tests were performed by exerting a high load on the bearing under test. A total of six run-to-failure experiments were accomplished, all at the same rotation speed but with different applied loads. Every test led to the generation of an extended defect on the outer raceway of each tested bearing. The results of both test types were employed for the validation of the modelling techniques delineated in the subsequent chapters.

Chapter 3 focused on the finite-element simulation of REBs. Within this context, it discussed the development of a computationally-efficient method for the generation of load-dependent meshes. The technique was primarily aimed at proposing an efficient tool for the estimation of bearing radial stiffness, but it may be also employed to model these components in more complex FE assemblies. The developed methodology involved the generation of a dedicated mesh for each load condition under exam. Interestingly, element dimensions were determined through analytical formulae rather than performing several convergence checks on subsequently finer grids. In fact, the main difficulty associated with this kind of simulations is the need of a sufficiently fine mesh in the proximity of the contact zones. This may lead to an excessive computational burden, especially at lower loads, for which contact surfaces reduce accordingly. Therefore, element dimensions in the neighbourhood of the contact areas between rolling elements and races were determined by means of the Hertzian theory. This well-known theory provided a simple and straightforward approach to the problem for a rough estimation of the expected contact areas. The proposed technique was then applied to two reference bearings, i.e. a roller bearing and a deep-groove ball bearing, in order to demonstrate

its efficiency for two different types of contacts between raceways and rolling elements. The inquiry ultimately led to the definition of the best element types to be employed for these simulations and the associated element dimensions for different load values. The accuracy of the stiffness estimate was compared against the results of an analytical model retrieved from the literature, thus validating the procedure. Moreover, additional effects that might influence bearing stiffness values were also investigated. Specifically, the effect of cage modelling, load direction and clearance were examined. By inserting the cage into the model as a meshed body, the computational time greatly increased in spite of a slight variation in the estimated stiffness. This result suggests to avoid modelling this component for the purpose of stiffness evaluation. On the other hand, a different direction of maximum load led to a stiffness deviation of up to 6.7%. Consequently, this effect might be considered by the interested analyst depending on the required degree of accuracy. Finally, clearance effect was studied by introducing a small gap between the components and by correspondingly modifying the load distribution. Clearance remarkably affected the stiffness values, in accordance with the results of the analytical model. In conclusion, the main aspect of originality of this work lies in the definition of an analytical procedure for the assessment of the optimal mesh elements dimensions for the estimation of bearing stiffness through FE analysis.

Subsequently, Chapter 4 inquired a numerical procedure for the estimation of the unknown parameter in LP models of defective REBs. In fact, it was observed that several parameters which are commonly inserted in these models are hardly measurable or rather denoted by a high degree of uncertainty. Notably, these involve the selection of damping values and the definition of the properties of the high-frequency resonant masses. Hence, the investigation was aimed at detailing a procedure for the estimation of these parameters on the basis of available experimental data. This problem was solved by introducing a multi-objective optimization technique in which the objective functions were global indicators that took into account the difference between numerical and experimental results for a variety of operative conditions and defect dimensions. The numerical features were determined through a LP model of a defective bearing, while experimental ones were calculated on the basis of the experimental signals acquired on a dedicated bearing test rig. The unknown parameters were the optimization variables of the problem.

These were constrained in order to only be able to assume plausible values within fixed boundaries depending on their physical properties. Furthermore, additional analytical relationships relating these quantities were added to take into account all possible dependencies between the considered parameters. To apply the proposed technique, a LP model of a self-aligning ball bearing was generated. Its formulation was based on previous works [22, 88, 99], but it was adapted to take into account a different ball path inside the defect area. This kind of model was employed to replicate the same bearing that was mounted on the bearing test bench. However, the LP model and the subsequent optimization procedure may be also extended to other bearing types. Moreover, contact damping was also differently defined by means of Rayleigh damping to achieve an enhanced control of this parameter over the whole frequency range of interest. Six unknown parameters were chosen, i.e. the two values needed to define the contact damping, the damping value of the inner ring, and the damping, stiffness and mass of the high-frequency resonant mass. The multi-objective optimization technique, then, aimed at minimizing two indicators, namely ΔRMS and ΔA^*, which took into account the energy of the whole signal and the energy associated to defect-related frequencies, respectively. The outcome of the procedure resulted in a number of weak Pareto optimals in which the variability in ΔRMS was considerably higher than the deviation of ΔA^* values. Therefore, the optimal set of parameters was estimated to be the one providing the lowest value of the first indicator. Comparison with the experimental acceleration signals denoted a good agreement, especially for all conditions associated with the lowest rotation frequency value, i.e. 20 Hz, and for all conditions associated with the smallest defect width, i.e. 0.928 mm. This result is in agreement with previously reported studies, which showed that these kind of models tend to loose accuracy for higher size of the localised defect. Furthermore, the effect of rolling element slippage was also inquired. The influence of the slip term appeared to be more marked for lower defect dimensions. Nevertheless, the parameters estimated by the optimization procedure allowed to capture the system vibration even by neglecting this additional term in the analysis. Ultimately, the main aspect of novelty introduced by the developed methodology consists in the peculiar approach adopted for the determination of the unknown model parameters. For this purpose, the technique provided satisfactory results, as highlighted by the comparison with experimental data.

Finally, Chapter 5 discussed about bearing prognostics. Specifically, effort was placed on the development of physics-based models to assess the bearing health state during its operative life. In this regard, a novel indicator named Equivalent Damaged Volume (EDV) was proposed. This parameter was defined as the equivalent volume per unit of width whose associated defect dimensions, if inserted into the LP model of the bearing, would give the same numerical RMS as the experimental one. The employed LP model was based on the one introduced in Chapter 4, but it was further improved to also take into account extended defects. This type of defect was modelled by adding the possible contact of the rolling elements with the bottom surface of the defect, which was described as a rough surface. The need to introduce this additional formulation arises from the experimental evidence that the bearing defects at the end of the run-to-failure tests all fell in the category of extended defects rather than their localised counterpart. These latter type of faults, in fact, usually appear only at the beginning of the damage stage. Then, subsequent passage of the rolling elements over the defect edges leads to the enlargement of the fault until it transitions to an extended defect. The proposed EDV algorithm took as input the RMS of the experimental signal during the final stage of the bearing life and the numerical RMS map calculated under a specified operative condition. The beginning of the final stage was evaluated by introducing a diagnostic indicator, i.e. the sum of the first five bearing characteristics frequencies in the envelope spectrum. The rapid increase of this indicator, in fact, paralleled the relatively slow increment of the RMS value, thus allowing for a better prediction of the time in which the bearing shifted from its normal operative life to the final stage. Besides, the RMS map were constructed by running several simulations with the LP model, each one characterized by different combinations of defect depth and angular extent. As a result, the output of the algorithm were the equivalent damaged volume V_{eq} and its associated dimensions $\Delta\varphi_d$ and h_d over time for each acquisition of the experimental signal. Interestingly, the EDV method was constructed in order to always generate an increasing volume over time, except when no suitable values to satisfy this condition were available in the RMS map for a given signal sample. The proposed algorithm was exploited to compute V_{eq} values of the experimental signals acquired during the run-to-failure campaign. Based on the estimated values, a first approach to bearing prognostics was attempted by proposing two prognostic PBMs that exploited this peculiar parameter. The first

PBM involved the prediction of bearing degradation histories under different conditions with respect to a reference one. This method was based on the assumption that the damage process followed the same trend under a different operative condition. However, this idea was limited by the low number of performed experimental tests. In fact, the same bearing may undergo several different degradation histories even under the same conditions of applied load and shaft rotation speeds. Therefore, this first approach shall be investigated in future efforts, presumably by considering a larger experimental data pool. On the contrary, the second PBM allowed to obtain more promising results. In fact, it was noted that experimental V_{eq} values had a slow increment prior to a certain threshold value. After that value of V_{eq} was crossed, the damaged volume rapidly increased until failure occurred. Therefore, the second PBM was based on the assumption that the trend of the equivalent volume below a certain threshold V_{th} could be predicted by fitting the data with a quadratic polynomial. As a a consequence, this approach aimed at estimating the time-to-threshold value rather than the RUL of the bearing. The efficiency in predicting the TT was demonstrated by applying the proposed method to the run-to-failure experiments. However, the major downside of this technique was that the quality of the fitting depended on the number of available data points in the final stage. In fact, the accuracy of the estimated TT was found to be low at the beginning of this phase, but it consistently increased as more data became available during each test. Therefore, further improvements to this model might include the analysis of other threshold values and the employment of different functions to fit the experimental data. To conclude, the main contribution of the closing chapter of this dissertation is the definition of a degradation-related parameter, i.e. the equivalent damaged volume per unit of width, which is employed to propose novel approaches for bearing prognostics.

A

APPENDIX: A NUMERICAL MODEL FOR NVH ANALYSIS OF GEARBOXES EMPLOYED ON AGRICULTURAL EQUIPMENT

A.1 INTRODUCTION

The appendix of this thesis deals with the problem of gearbox modelling applied to a real industrial problem. In particular, this chapter describes the generation process of a digital twin aimed at the noise, vibration and harshness (NVH) analysis of a gearbox employed on agricultural equipment. These components, in fact, may produce high levels of noise during their operational life. The noise is generated by gear meshing, and it is further amplified by the resonances of the case [240]. Consequently, noise levels during operation may be too high and exceed acoustic tolerance limits. As a matter of fact, the acoustic emission is hardly estimable during the design process, which is finalized at obtaining a certain gear ratio for the gearbox. Nonetheless, the overall acoustic pressure level will depend both on the choice of the gears and on the properties of the case. Therefore, by assuming a fixed design for the case, the generated noise will depend on the appropriate selection of gears that are able to achieve the required gear ratio. In particular, their teeth profile highly affects the meshing phenomenon and the associated vibration. As a result, noise reduction may be accomplished by performing an extensive campaign of experimental testing. However, this means measuring acoustic pressure levels for every possible combination of gears and teeth profile. Indeed, the process is very time-consuming and requires a large amount of resources.

A faster way to optimize the gearbox design consists in generating a digital twin of the component itself, i.e. a digital replication of the physical entity [241] which allows to test different gear configurations without creating a physical prototype of the gearbox for each test. In this case, the experimental tests must be carried out only to validate the initial baseline model. In fact, subsequent design modifications may be numerically evaluated by changing its input parameters [242]. Within this framework, this appendix describes the generation process of a digital twin of a real gearbox employed on agricultural equipment. The main objective is to provide a tool capable of assessing the overall acoustics in operational conditions. The proposed digital twin is a combination of a lumped-parameter (LP) model, a structural finite-element (SFE) model and an acoustical finite-element (AFE) model. A similar procedure, involving the use of these three types of models, was developed in previous works [243, 244, 245]. The LP model is used to obtain the reaction forces on the bearings during working conditions. Reaction forces are employed as an input for the further SFE dynamic model to evaluate the dynamic response of the gearbox case, which is then exploited to set-up the AFE model. This last model allows to estimate the noise generation in terms of overall acoustic pressure. Numerical results are validated through comparison against experimental data acquired on a real gearbox. Testing activities were carried out at Comer Industries facility in Reggiolo, where specimens and test benches were set.

In the following, the developed models are presented and experimentally assessed. Section A.2 describes the system under examination and the three numerical models. Section A.3 reports the experimental setup and the validation process of the models based on the acquired data. Eventually, Section A.4 closes the work by providing some summarizing remarks.

A.2 DIGITAL TWIN OF THE GEARBOX

The developed digital twin is a combination of three sub-models: a LP model, a SFE model and an AFE model of a gearbox employed on agricultural equipment. The output of the combined LP/SFE/AFE model is the overall acoustic pressure due to specified working conditions and gear design. The following

sections describe the system under exam and all the implemented numerical models.

A.2.1 *Description of the system*

The system under study is schematically depicted in Figure 68. The gearbox case contains three shafts, five spur gears and six bearings. The shafts are named driver, intermediate and driven shafts, respectively. On the other hand, the gears are denoted by numbers from 1 to 5. Gear ratio of the system, i.e. $\tau = \Omega_{out}/\Omega_{in}$, may take the values 1 or 1.8, depending on the selected meshing gears. In fact, τ is equal to 1 when the meshing takes place between gears 1-3 and 3-5, while $\tau = 1.8$ when the meshing occurs between gears 2-4 and 3-5. The gear ratio may be manually changed by employing a coupling mounted on the driver shaft. Within the context of this work, results will be reported only for $\tau = 1$, i.e. $\Omega_{in} = \Omega_{out}$. This gear ratio is achieved due to the fact that gears 1, 3 and 5 have the same number of teeth, i.e. 37. For the purpose of this work, only the condition $\tau = 1$ is investigated.

A.2.2 *Lumped parameter model*

The dynamic behaviour of the moving components inside the case is described by means of a non-linear lumped-parameter model. The LP model allows to estimate the reaction forces on the bearings due to gear meshing by simulating the dynamic effect of the motion of the components inside the case. Therefore, it consists in the subdivision of the system into masses and inertias connected to each other by elastic and viscous damping elements. The model considers the effects of time-varying mesh stiffness, the non-linearity of the meshing phenomenon and a constant bearing stiffness. It also takes into account the backlash, which is necessary to allow better lubrication, reduce wear and limit interference due to geometrical construction errors. However, the backlash induces torsional vibrations that can cause the detachment between teeth [246]. Consequently, in this model, for a given direction of pinion rotation the forces are exchanged along the direct line of action. Despite this, in case contact is lost, the driven wheel may impact the pinion on the opposite side of the tooth. By considering that only spur gears are present in the system, translations

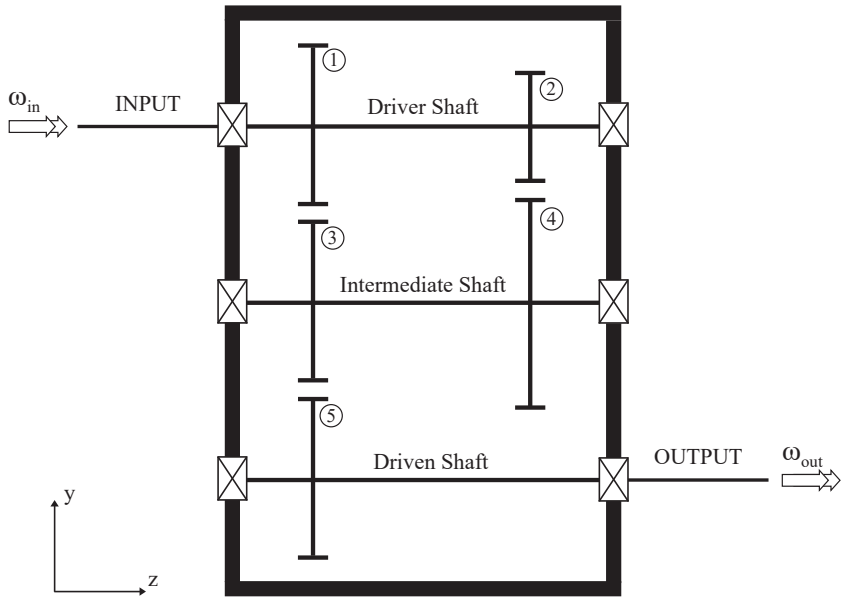

Figure 68: Schematic depiction of the gearbox.

and rotations were monitored just in the plane transverse to the gears, as no displacement will occur in the z direction. As a result, as it may be appreciated in Figure 69, the LP model is composed by nine degrees of freedom, i.e. the x and y translations and the angular displacement θ around the shaft rotation axis for each gear. The equation governing the motion of the system is:

$$\mathbf{M}\ddot{\mathbf{x}} + \Gamma(x_d, x_b)\left[\mathbf{C}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} - \frac{1}{2}\mathbf{k_m}x_b\right] = \mathbf{w} \tag{98}$$

where

$$\Gamma(x_d, x_b) = \begin{cases} 1 & \text{if } x_d > x_b/2 \\ 0 & \text{if } -x_b/2 \leqslant x_d \leqslant x_b/2 \\ -1 & \text{if } x_d < -x_b/2 \end{cases} \tag{99}$$

while **M**, **C** and **K** are the mass, damping and stiffness matrices, respectively; x_d is the dynamic transmission error; x_b is the gear backlash; $\mathbf{k_m}$ is the time-varying meshing stiffness; $\ddot{\mathbf{x}}$, $\dot{\mathbf{x}}$, **x** are the acceleration, velocity and displacement

vectors, respectively; **w** is the external forces vector. Damping is defined as Rayleigh damping, i.e. proportional to mass and stiffness matrices, as detailed in Eq. 39. Different strategies may be followed for the choice of parameters α_c and β_c. As a preliminary approach, they may be fixed according to experience. Then, they may be calibrated on the most important linearised resonances of the geartrain, which may be obtained experimentally. In this study, the first option is adopted. As a result, different damping coefficients are considered for each damper in the LP model.

According to Figure 68, since the investigation is carried out only for gear ratio $\tau = 1$, the considered gears are the number 1, 3 and 5. Since each gear has three DOFs, the total number of DOFs in the model is nine. For the i-th gear, $k_{b,i}$ and $c_{b,i}$ are the bearing stiffness and damping, respectively. For the latter, coefficient $\alpha_{c,i}$ and $\beta_{c,i}$ are also introduced. Besides, $k_{m,ij}(t)$ is the time-varying meshing stiffness between the i-th and j-th gears, while $c_{m,ij}(t)$ is the associated damping. The associated Rayleigh damping parameters are named $\alpha_{c,ij}$ and $\beta_{c,ij}$, respectively. Parameters m_1 and j_1 are the mass and inertia of the components mounted on the driver shaft. Similarly, m_3 and J_3 have the same meaning for the intermediate shaft, while m_5 and J_5 refer to the driven shaft. When meshing takes place between gears 1 and 3, gear 2 is not rotating with the driver shaft. Its rotary motion, in fact, is accomplished by gear 4 that is mounted on the intermediate shaft. As a consequence, the rotary inertia of gear 2 is considered as part of J_3. On the other hand, the oscillating mass of gear 2 is accounted for by m_1. The proposed numerical model is implemented in MATLAB [235] and it is employed to determine the reaction forces on bearings for a given operative condition.

A.2.3 *Structural finite-element model*

A FE analysis is carried out in order to estimate the case vibration due to the reaction forces exerted on bearings by gear meshing. The model is developed in Simcenter 3D [214], where the pre and post-process phases are carried out, while the chosen solver is Simcenter Nastran. In the proposed SFE model, only the external case of the gearbox is meshed. The operation is performed by exploiting 4-nodes 3D tetrahedral elements, as shown in Figure 70.a. On the

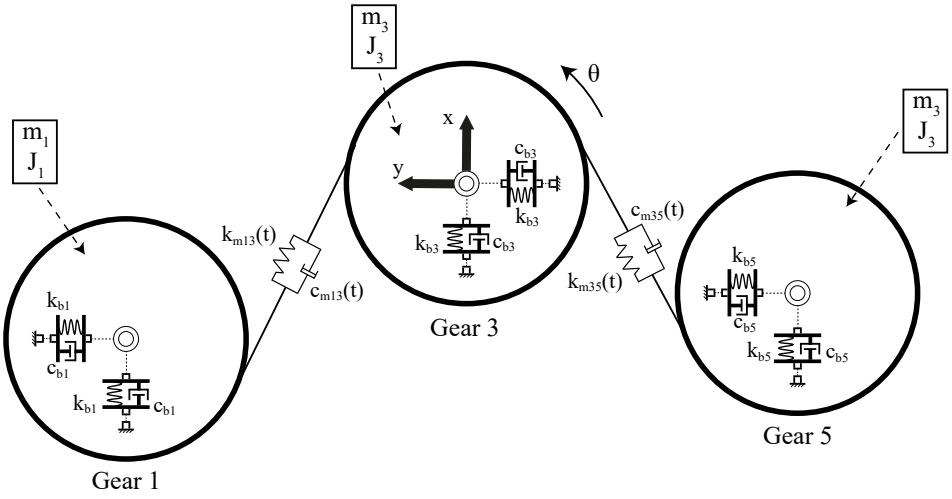

Figure 69: Scheme of the geartrain LP model.

contrary, all the internal components are replaced by concentrated masses, which are elements defined only by their centre of mass and inertial properties, i.e. mass and moments of inertia. As it may be appreciated in Figure 70.b, these concentrated masses are connected to the respective bearing housings by rigid elements, namely RBE2 in Nastran. Thus, the central node, i.e. the concentrated mass, is connected to all the nodes on the surface of the corresponding bearing housing. Concerning the boundary conditions, Figure 70.c shows that fixed constraints are applied on the bottom of the case, in place of the screws that would tighten it to the plate below. A fixed constraint implies that all the degrees of freedom for the involved nodes are restrained. The SFE model is employed to perform a numerical modal analysis and a dynamic simulation. The former allows to validate the mesh of the model through comparison with experimental results, and it is solved by exploiting Nastran solution SOL 103. The dynamic model, on the other hand, is solved by utilizing the modal frequency response solution, i.e. SOL 111 in Nastran. The output forces of the LP model are used as input for this model to assess the structural vibration of the case. Results are validated by comparison against acceleration data acquired during testing.

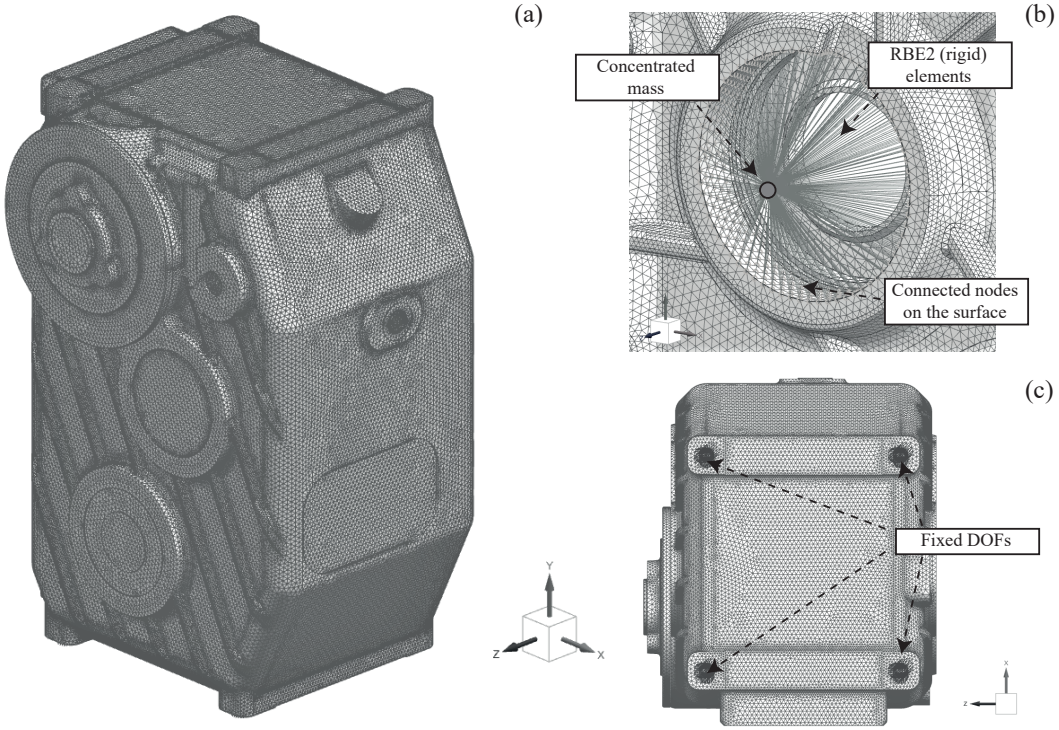

Figure 70: Gearbox mesh: (a) whole case; (b) rigid elements connecting a concentrated mass to the case; (c) constrained nodes.

A.2.4 *Acoustical finite-element model*

An acoustical finite-element model is developed to estimate the overall acoustic level due to the vibration of the case induced by gear meshing. The proposed AFE model exploits the acoustic transfer vector (ATV) approach. This method consists in computing a set of functions which establish a relationship between the normal velocity of the nodes on the surface of the vibrating structure and the acoustic pressure at defined locations [247], named microphone points. For each one of them, a frequency-dependent vector $\mathbf{atv}(f)$ is computed so that:

$$p(f) = \mathbf{atv}^{T}(f)\mathbf{v_n}(f) \tag{100}$$

where $p(f)$ is the pressure at a specified location and $\mathbf{v_n}(f)$ is a vector containing the normal velocities of each node on the radiating surface. Both these quantities are also frequency dependent. Furthermore, by assuming that there

are m microphone points and n nodes on the vibrating surface, the vector $\mathbf{atv}(f)$ is computed for all m point in order to generate an $m \times n$ acoustic transfer matrix $\mathbf{ATM}(f)$ that satisfies the equation:

$$\mathbf{p}(f) = \mathbf{ATM}(f)^T \mathbf{v_n}(f) \tag{101}$$

where $\mathbf{p}(f)$ is a vector of length m containing the frequency-dependent acoustic pressures at each microphone point. The main characteristic of the ATV matrix $\mathbf{ATM}(f)$ is that it does not depend on the structural response. In fact, it only depends on the properties of the acoustic domain and on the position of the microphone points. This peculiar characteristic translates to the fact that, for a given fluid - radiating surface interface, matrix $\mathbf{ATM}(f)$ is constant and needs to be computed only once. As a result, various loading conditions may be computed by using the same ATV matrix, thus leading to a consistent reduction in analysis time compared to other approaches.

To compute $\mathbf{ATM}(f)$, it is necessary to create a 3D acoustic mesh for the fluid surrounding the gearbox case. As the problem concerns the solution of exterior acoustics, the size of the domain is reduced by exploiting a perfectly matched layer (PML), which is an artificial layer that absorbs all the incoming incident waves instead of reflecting them back into the fluid, independently from their frequency and their direction [248]. The PML allows to truncate the size of the acoustic domain, thus reducing the number of 3D elements needed to create the mesh. The resulting grid is shown in Figure 71. The layer is applied on the entire external surface of the mesh excluding its bottom side, where an infinite plane is defined instead. In fact, according to Figure 72, such a plane allows to take into account the presence of the steel plate located below the case, which behaves like a reflecting surface. To this end, the infinite plane allows to impose a symmetry boundary condition, i.e. the normal velocity on the plane is enforced to be equal to zero. Finally, microphone points may be placed anywhere outside the layer, as shown in Figure 71. In the acoustical model, they are located in the same location as the corresponding microphones in the experimental tests. This model is solved through Nastran solution SOL 108.

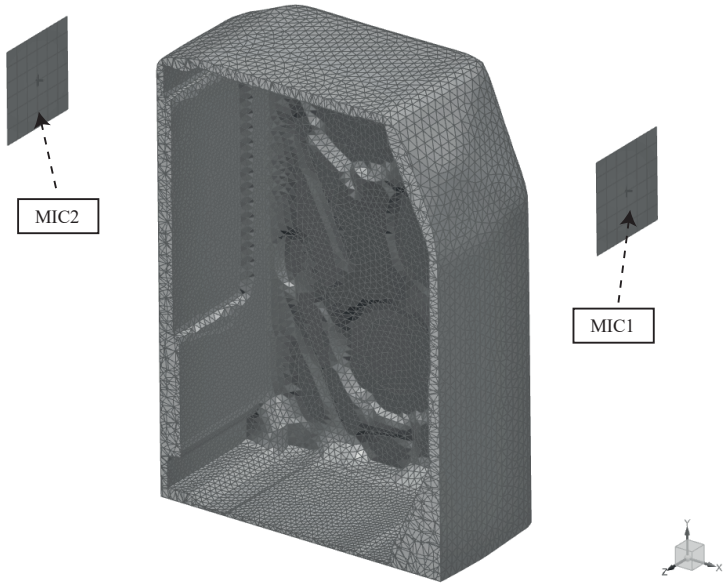

Figure 71: Acoustic mesh and microphone points.

A.3 VALIDATION OF THE DIGITAL TWIN

This section outlines the validation process of the models which constitute the digital twin of the gearbox. The experimental setup employed to acquire the data during operation is described. Then, the results of the LP, SFE and AFE models are reported along with the associated experimental validations.

A.3.1 *Experimental setup*

The experimental setup is shown in Figure 72. The driver shaft is connected to the output shaft of the engine of an agricultural tractor by an universal joint, while the driven shaft is connected to a braking system. The case is fixed on a steel plate. The reference frame depicted in the bottom left of the figure, employed for the experimental tests, has the same orientation as the ones utilised in all the numerical models. The performed experimental tests may be divided in two categories, i.e. experimental modal analysis (EMA) and operative analysis (OA). The EMA was carried out by using the roving hammer based method. For this purpose, 31 excitation points and 3 response

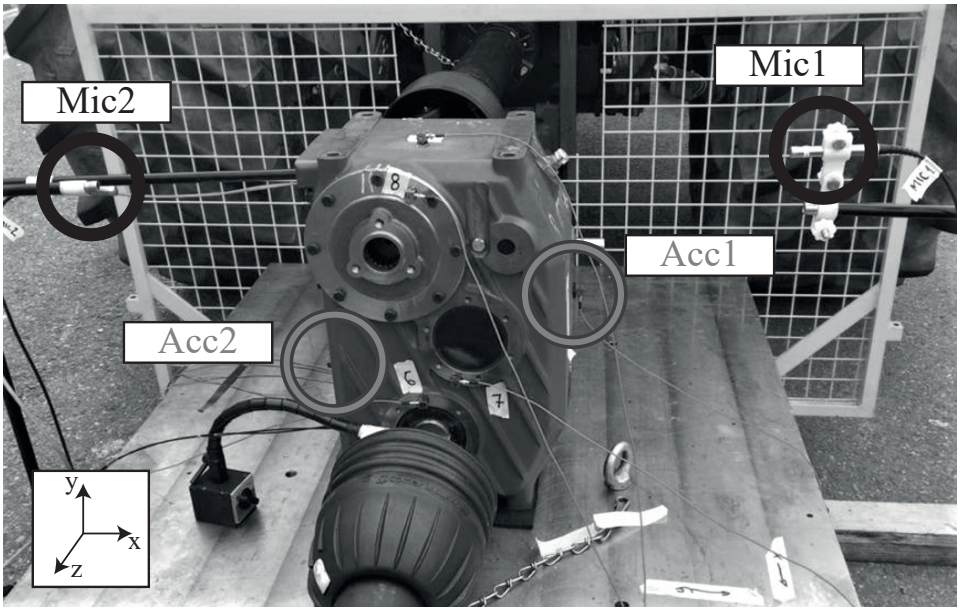

Figure 72: Experimental setup.

points distributed on the surfaces of the case were employed. Excitations were applied with an impact hammer model PCB 086D05, while the responses were measured by piezoelectric triaxial accelerometers model PCB 356B21. The results of this analysis are used to validate the SFE model. On the other hand, the OA has been realized in firing condition by applying a load on the driver shaft through the output shaft of the engine. Acceleration on the case was measured by using two piezoelectric triaxial accelerometers model PCB 356B21 (Acc1 and Acc2 in Figure 72) fixed on two opposite faces of the case. At the same time, the acoustic pressure was measured with two microphones model PCB 378B02 in front of the accelerometers, at a distance of 20 cm from the corresponding surfaces (Mic1 and Mic2 in Figure 72). The rotational speed of the input and output shafts was acquired by using two optical tachometers placed on the corresponding shafts.

Tests were run for one working condition and for two different teeth profiles, here called *Baseline* and *Mod1*. The latter consists in a modification of the geometry of the former, i.e. the original profile. This peculiar profile is not described for confidentiality reasons. By changing the profiles, two different

Description	Symbol	Value
Input power	W	50 kW
Gear ratio	τ	1
Rotational speed of the driver shaft	Ω_{in}	506.4 RPM
Rotational speed of the driven shaft	Ω_{out}	506.4 RPM
Rotational frequency of the shafts	f_{rot}	8.4 Hz
Meshing frequency of gear pairs 1-3 and 3-5	f_m	310.8 Hz

Table 23: Tested working condition.

time-varying meshing stiffness are generated. As a consequence, different acoustic emissions are produced. The working condition utilized for testing is reported in Table 23. It is defined by the rotational speed of the driver shaft Ω_{in}, input power W and gear ratio τ. The table also reports the rotational frequency f_{rot} of the shafts and the gear meshing frequency f_m of gears 1-3 and 3-5. It is worth noting that for both gear pairs only one value is listed since for $\tau = 1$ all shafts have the same speed, thus leading to an equal gear meshing frequency.

Pressure levels measured during OA are reported in Figure 73 for both tested profiles. From these graphs it may be noticed how the highest contribution to the overall level comes from the frequency interval consisting in the 1/3 octave bands whose centre frequencies range from 630 Hz to 2500 Hz, i.e. from 562 Hz (lower frequency limit of the 630 Hz band) to 2818 Hz (upper frequency limit of the 2818 Hz band). This observation allows lowering the computational burden for the SFE and AFE dynamic model by decreasing the frequency range of the analysis from 0 - 20 kHz to the reduced interval 562 - 2818 Hz. To conclude, it should be noted that profile *Mod1* provides a reduction in the overall levels. In fact, reductions of 3.5 dB(A) on Mic1 and 4.9 dB(A) on Mic2 are obtained by changing from the original to the modified teeth profile. These levels are A-weighted and computed over the reduced frequency range 562 - 2818 Hz for a reference pressure $P_{ref} = 2 \cdot 10^{-5}$ Pa.

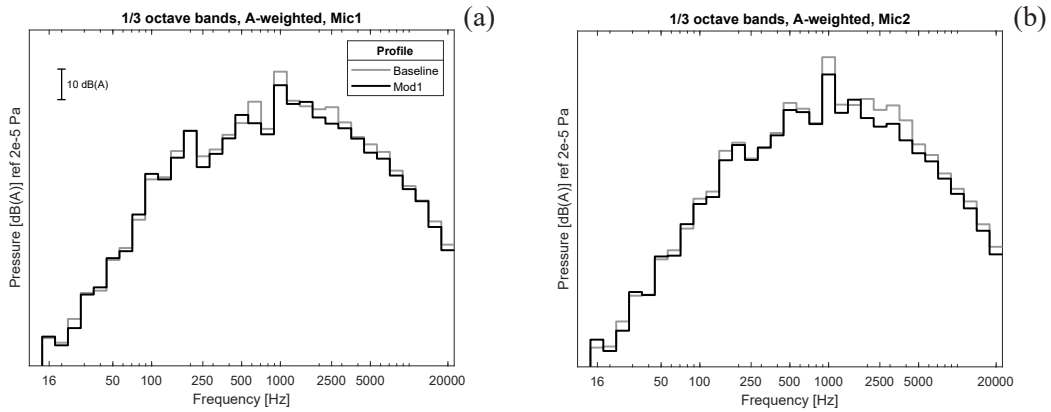

Figure 73: Acoustic levels sensed by the microphones: (a) Mic1; (b) Mic2. Reference pressure is $p_{ref} = 2 \cdot 10^{-5}$ Pa. (Values on y-axis not shown for confidentiality reasons).

A.3.2 *LP model results*

The value of the parameters used for the LP models are reported in Table 24. The values of masses and inertias are determined based on the considerations reported in Section A.2.2. Since all six bearings are the same model only one value, namely k_b, is reported in Table 24 To simplify the problem, the nonlinear characteristic of the bearing is neglected, thus leading to a constant value of radial stiffness. The meshing stiffness, on the other hand, is the source of excitation in the model and it is therefore time-varying. The average value over time for the two profiles *Baseline* and *Mod1* are shown in the table. Then, in the LP model, the corresponding meshing stiffness values $k_{m13}(t)$ and $k_{m35}(t)$ are equal to either one of them depending on the testing condition under exam. Damping values, on the contrary, assume different values on each bearing and meshing gear pair. As detailed in Section A.2.2, damping is defined through Rayleigh damping. As a consequence, Table 24 reports the values coefficients α_c and β_c. While α_c retains the same value for each component, β_c does not. The values of β_c, in fact, are chosen in order to consider the damping effect of the oil inside the case.

Resulting reaction forces on bearings are exploited as input solicitation for the subsequent SFE analysis. The forces generated by gear meshing are depicted

Description		Symbol	Value
Mass	Gear 1	m_1	15.88 kg
	Gear 3	m_3	16.85 kg
	Gear 5	m_5	10.69 kg
Moment of inertia	Gear 1	J_1	$2.36 \cdot 10^{10}$ kg·m²
	Gear 3	J_3	$5.38 \cdot 10^{10}$ kg·m²
	Gear 5	J_5	$1.72 \cdot 10^{10}$ kg·m²
Bearing stiffness	All	k_b	$1.82 \cdot 10^9$ N/m
Average meshing stiffness	*Baseline*	$k_{m,B}$	$7.73 \cdot 10^8$ N/m
	Mod1	$k_{m,M}$	$7.26 \cdot 10^8$ N/m
β_c coefficients	Gear 1	$\beta_{c,}$	$1 \cdot 10^{-4}$ s
	Gear 3	β_{c3}	$5 \cdot 10^{-4}$ s
	Gear 5	β_{c5}	$6 \cdot 10^{-4}$ s
	Meshing 1-3	β_{c13}	$1 \cdot 10^{-4}$ s
	Meshing 3-5	β_{c35}	$5 \cdot 10^{-4}$ s
α_c coefficient	All	α_{ci}, α_{cij}	1 s^{-1}

Table 24: Values assigned to the components of the LP model.

in Figure 74 for both planar directions, i.e. x and y. However, the shafts are mounted on two bearings each. Therefore, in the SFE model, the total force on each gear is subdivided on the two respective bearings on the basis of the distance between the gear itself and the bearings. Finally, from Figure 74 it is also possible to observe that profile modification from *Baseline* and *Mod1* leads to a strong reduction of bearing reaction forces.

A.3.3 *SFE model results*

The SFE model is generated according to the methods described in Section A.2.3. Material of the case is cast iron with Young's modulus $E = 90$ GPa, Poisson's ratio $\nu = 0.3$ and density $\rho = 7250$ kg/m³. Before proceeding with the dynamic analysis, the structural mesh is validated against the EMA results.

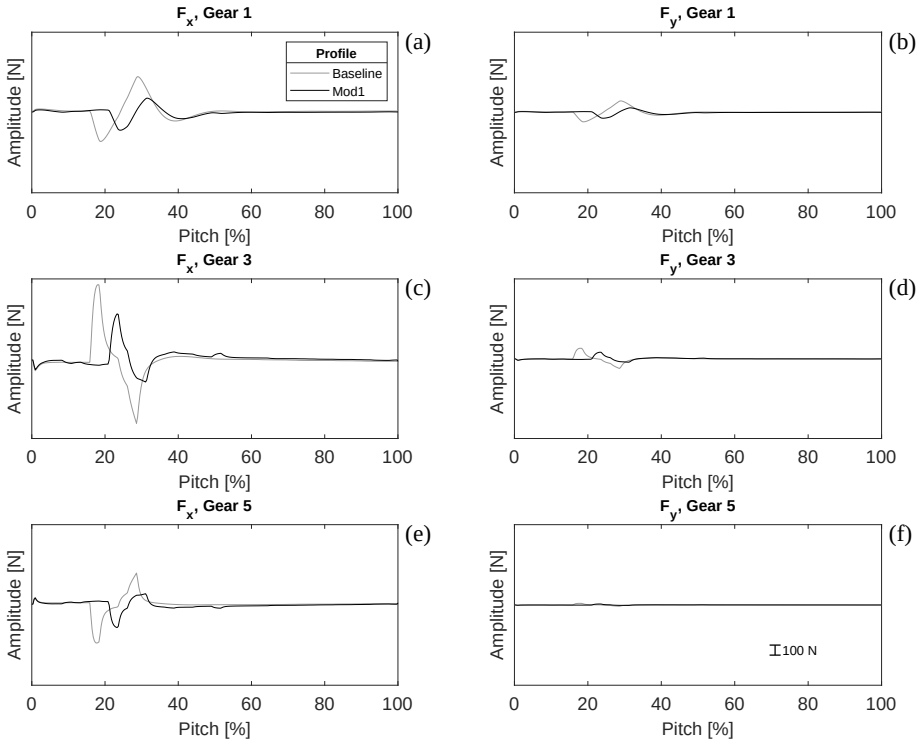

Figure 74: Reaction forces on bearings as a function of the gear pitch: (a) Gear 1, x direction; (b) Gear 1, y direction; (c) Gear 3, x direction; (d) Gear 3, y direction; (e) Gear 5, x direction; (f) Gear 5, y direction. (Values on y-axis not shown for confidentiality reasons).

This purpose is fulfilled by comparing the natural frequencies and the mode shapes obtained numerically against EMA results. In agreement with the acoustic measurements reported in Section A.3.1, focus is placed only on the frequency range with the highest acoustical emission, i.e. the one located between 562 and 2818 Hz. In particular, validation is carried out for the first two experimental modes in this interval which are found at 828 Hz and 1221 Hz. The results obtained with SOL 103 are reported in Table 25, while the numerical and experimental mode shapes are shown in Figure 75.a for the first mode and Figure 75.b for the second mode. It is worth nothing that, for both modes, the obtained numerical natural frequencies are close to the

	Natural Frequency [Hz]		
	EMA	Numerical	MAC
Mode 1	828	835	0.66
Mode 2	1221	1143	0.73

Table 25: Comparison between the first two natural frequencies obtained experimentally (EMA) and numerically (Nastran SOL 103).

experimental ones, as the frequency difference between them is 0.9% for mode 1 and 6.4% for mode 2. Furthermore, the modal assurance criterion (MAC) takes fairly high values for both of them, since it is equal 0.66 and 0.73 for the first and second mode, respectively. These results denote a good accordance between the experimental and numerical results, hence validating the SFE model and allowing to further employ it for the dynamic analysis.

The modal frequency response solution, i.e. SOL 111 in Nastran, is employed to solve the dynamic model. Forces are applied on the bearing housings by providing their magnitude and phase and by distributing them on the entire surface of each housing. Since the axial forces are neglected, a total of twelve forces acting on the system, six of them in the x direction and the other six in the y direction, are considered so that two forces acting on each bearing are accounted as described in Section A.3.2. As for the modal analysis, the dynamic analysis is carried out only in the frequency range of interest located between 562 and 2818 Hz. Furthermore, frequency-dependent modal damping is introduced in this analysis. Its values, estimated by the EMA, are reported in Table 26. The applied forces are the ones estimated by the LP model for the working conditions reported in Table 23. The output of the model of major interest is the vibration of the case, namely its acceleration. To validate the results, numerical acceleration results are compared with experimental results obtained during the operative analysis by the accelerometers Acc1 and Acc2 (Figure 72). Acceleration is numerically evaluated by taking the x component of the acceleration from two nodes on the case shown in Figure 76. In fact, they are located in the same position as the accelerometers in Figure 72. Results are then compared in terms of reduction of overall acceleration levels in the

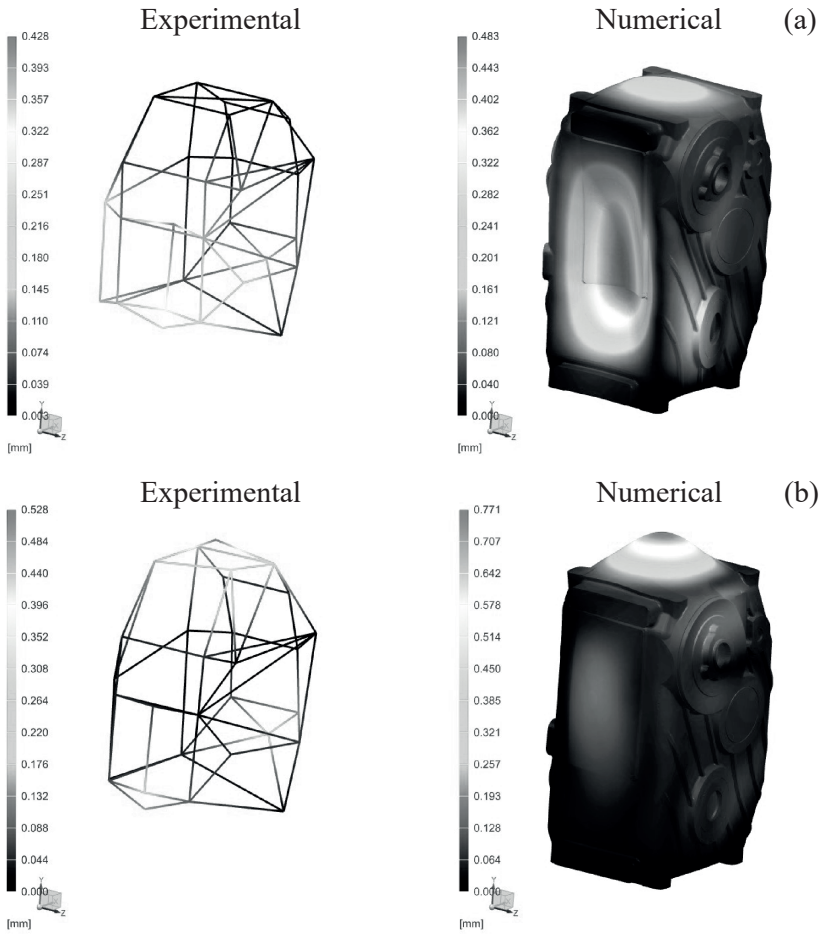

Figure 75: Comparison between the first two mode shapes obtained experimentally (EMA) and numerically (Natran SOL 103): (a) Mode 1; (b) Mode 2.

frequency range 562 - 2818 Hz as reported in Table 27 and in 1/3 octave band spectrum as shown in Figure 77. The results show that the dynamic model is able to capture the reduction due to the modified profile, as the order of magnitude of the dB reductions is similar for both accelerometers. In fact, for accelerometer Acc1, the model predicts a reduction of 7.3 dB, while the experimentally measured reduction is 5.3 dB; for Acc2 the results are closer, as the model gives a reduction of 5.5 dB instead of the experimental reduction of 6.1 dB. Since interest is placed in the assessment of the overall levels, these

Frequency [Hz]	Modal Damping [%]
562	7.6
828	7.6
1221	3.7
1423	1.8
2154	0.5
2818	2.3

Table 26: Modal damping values estimated by the EMA .

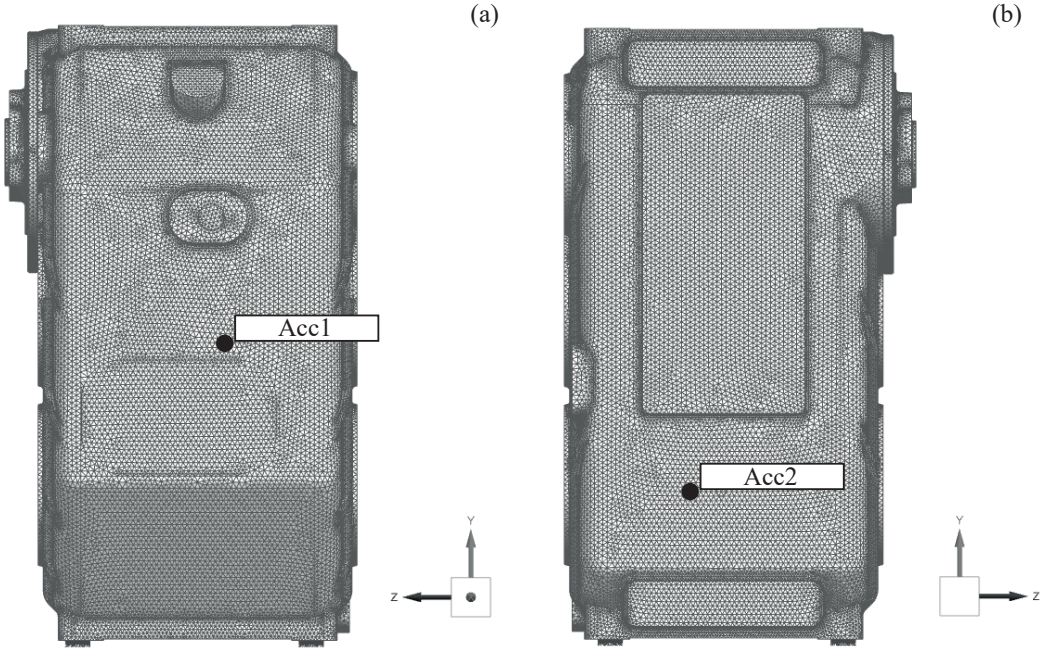

Figure 76: Location of the nodes corresponding to the two accelerometers: (a) Acc1; (b) Acc2.

results validate the dynamic model. Consequently, it may be subsequently exploited for the AFE model.

	Acc1		Acc2	
	Experimental	Numerical	Experimental	Numerical
dB reduction	-5.3	-7.3	-6.1	-5.5

Table 27: Reduction of the overall acceleration levels, x component, computed in the highest acoustic emission frequency range (562 - 2818 Hz). Reference acceleration is a_{ref} = 1 g.

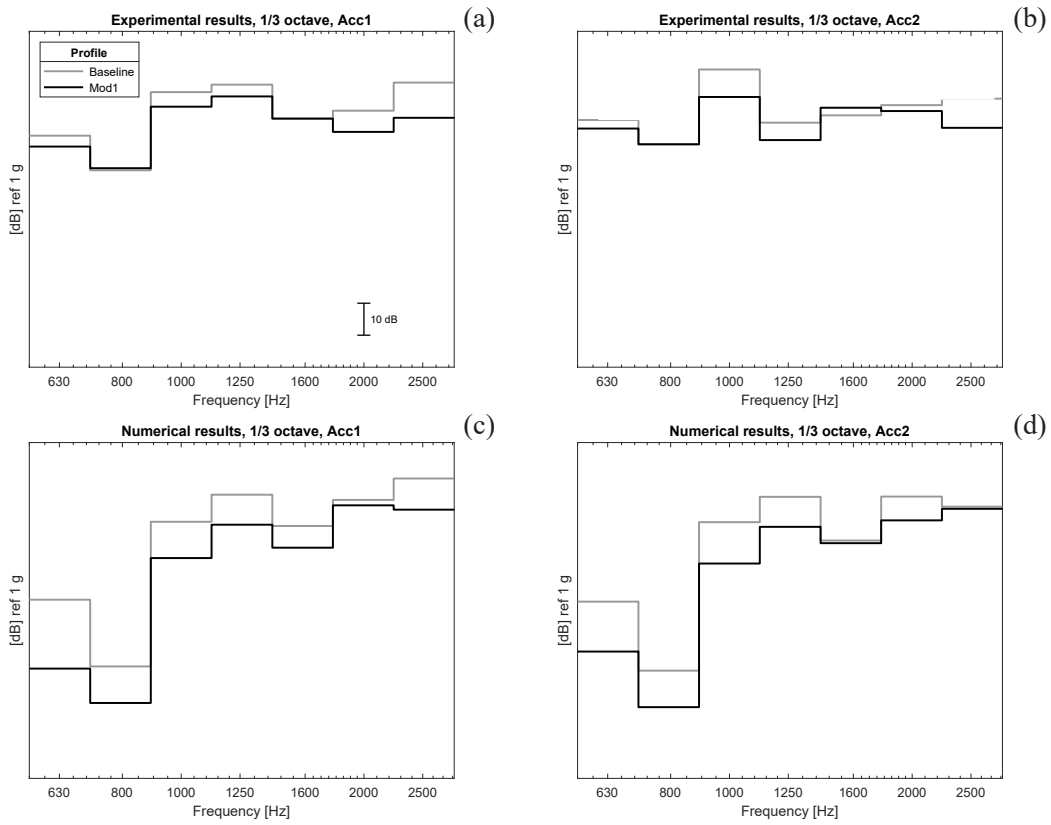

Figure 77: Experimental and numerical acceleration levels in 1/3 octave band spectrum, x component, for both teeth profiles: (a) Acc1, experimental; (b) Acc2, experimental; (c) Acc1, numerical; (d) Acc2, numerical. (Values on y-axis not shown for confidentiality reasons).

	Mic1		Mic2	
	Experimental	Numerical	Experimental	Numerical
dB(A) reduction	-3.5	-3.3	-4.9	-4.9

Table 28: Reduction of the overall acoustic pressure levels, A-weighted, computed in the highest acoustic emission frequency range (562 - 2818 Hz). Reference pressure is $p_{ref} = 2 \cdot 10^{-5}$ Pa.

A.3.4 *AFE model results*

The AFE model detailed in Section A.2.4 allows to compute the ATV matrix that relates the normal velocity on the surface to the acoustic pressure at defined microphone points. Therefore, it can be combined with the results of the dynamic model to obtain the acoustic pressure at locations Mic1 and Mic2 shown in Figure 71. The acoustic fluid is assumed to be air having constant properties, in particular density $\rho = 1.20$ kg/m^3 and speed of sound $v_s = 343$ m/s.

The reductions of overall acoustic pressure levels due to to teeth profile modification are reported in Table 28. Moreover, Figure 78 compares the experimental and numerical results in 1/3 octave band spectrum. According to these results, the model is able to capture the relative reductions of the acoustic pressure when the tooth profile is changed. In fact, for Mic1 the reduction is 3.5 dB(A) in the experiments and 3.3 dB(A) for the numerical model. Moreover, the same reduction is achieved for both the experimental and the numerical results for Mic2, i.e. 4.9 dB(A). This ensures the quality of the results obtained by the combined LP/SFE/AFE model. In conclusion, the proposed digital twin for the gearbox is validated, as it is able to capture the overall reduction of sound pressure levels in the selected frequency range due to modifications in the geartrain design.

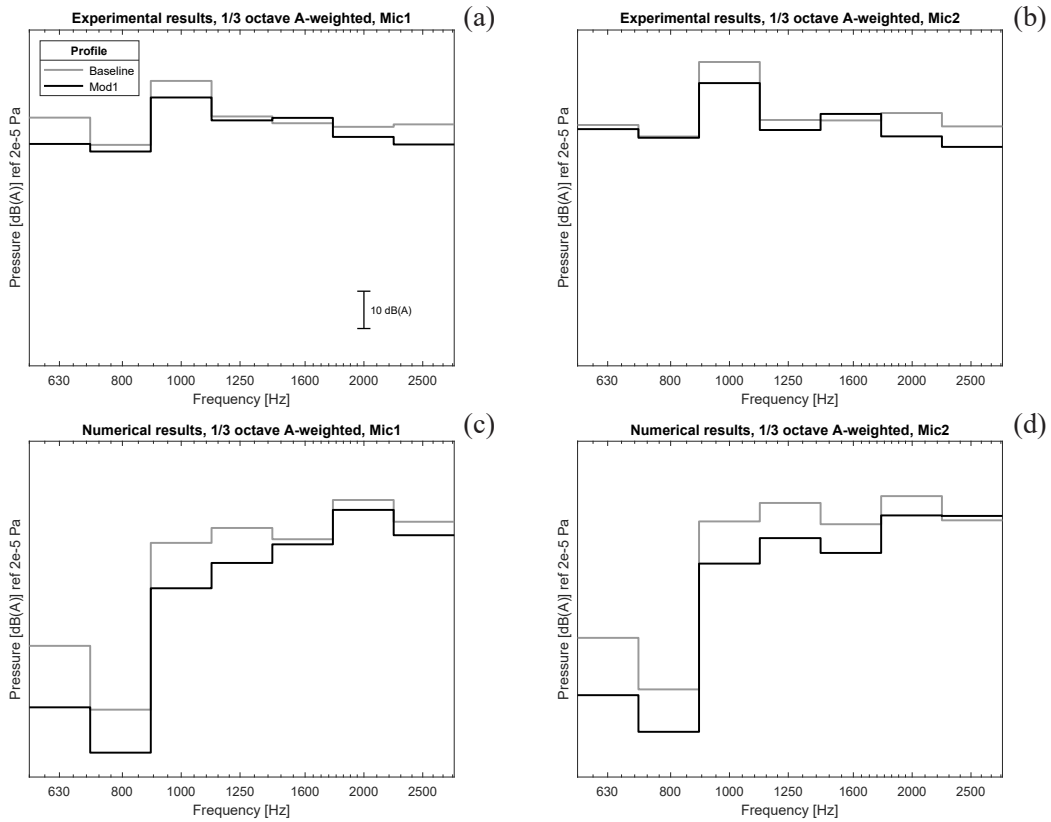

Figure 78: Experimental and numerical A-weighted acoustic pressure levels in 1/3 octave band spectrum, for both teeth profiles: (a) Mic1, experimental; (b) Mic2, experimental; (c) Mic1, numerical; (d) Mic2, numerical. (Values on y-axis not shown for confidentiality reasons).

A.4 SUMMARIZING REMARKS

This appendix detailed a combined LP/SFE/AFE model forthe vibro-acoustic analysis of gearboxes. The model was exploited to generate the NVH digital twin of a real gearbox employed on agricultural equipment. The LP model simulated the dynamic behaviour of the components inside the case, while the SFE model allowed to predict its vibration. Finally, the AFE model permitted to determine the overall acoustic pressure level at microphone points. Results were experimentally assessed, and a good accordance was found between the numerical and the experimental data. In fact, the SFE model of the case

allowed to accurately identify the first two natural frequencies in the highest acoustic emission interval. Furthermore, the dynamic SFE model was capable of capturing the reduction of overall acceleration levels due to different teeth profiles. A similar result was achieved by the AFE model in terms of overall acoustic pressure level.

The validation process allows to assert the combined LP/SFE/AFE model as a digital twin of the gearbox, that is, a tool that may accompany the designer in the choice of the optimal solution for the noise generation problem. The main advantage of this model is the capability of being able to easily estimate the acoustic pressure level generated by various working conditions, i.e. input speeds, power, gear ratio and teeth profiles, in order to choose the optimal design which reduces the noise produced by gear meshing. In fact, by using the ATV method, the ATV matrix must be computed only once provided that the fluid-radiating surface interface does not change. In other words, further computations are not needed if the external geometry of the case is not modified. As a consequence, this methodology speeds-up the optimization process since only the LP and the SFE models have to be solved again to obtain the normal velocities on the surface of the case that are needed for the AFE model. Also, even the numerical modal analysis (SOL 103) has to be performed only once. In fact, once the structural mesh is validated, it shall not be modified in subsequent analyses with the same case.

In conclusion, the proposed digital twin is suitable to estimate overall levels and it can be employed as an instrument to evaluate the best design among a set of possible combinations of gear types to determine which one generates the lowest noise. Its main strength, in fact, resides in the capability to assess the variations in terms of dB when the magnitude of the excitation source, i.e. gear meshing, changes.

BIBLIOGRAPHY

[1] A. Gabrielli, M. Battarra, and E. Mucchi. "A Critical Analysis of Finite-Element Modeling Procedures for Radial Bearing Stiffness Estimation." In: *Mathematical Problems in Engineering* 2021 (2021), p. 9955398. ISSN: 1024-123X. DOI: 10.1155/2021/9955398.

[2] A. Gabrielli, M. Battarra, and E. Mucchi. "A numerical finite-element method for radial bearing stiffness estimation based on load dependent meshing." In: *Proceedings of ISMA 2022 - International Conference on Noise and Vibration Engineering and USD 2022 - International Conference on Uncertainty in Structural Dynamics, pp. 1527–1541*. Leuven, Belgium, 2022.

[3] A. Gabrielli, M. Battarra, E. Mucchi, and G. Dalpiaz. "A procedure for the assessment of unknown parameters in modeling defective bearings through multi-objective optimization." In: *Mechanical Systems and Signal Processing* 185 (2023), p. 109783. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2022.109783.

[4] A. Gabrielli, M. Battarra, E. Mucchi, and G. Dalpiaz. *Acceleration signals of rolling element bearings with artificial defects*. Mendeley Data, V1. 2022. DOI: 10.17632/8wdzm5gwng.1.

[5] A. Gabrielli, F. Pizzolante, E. Soave, M. Battarra, C. Mazzeo, M. Tarabra, E. Fava, and E. Mucchi. "A numerical model for NVH analysis of gearboxes employed on agricultural equipment." In: *Proceedings of ISMA 2020 - International Conference on Noise and Vibration Engineering and USD 2020 - International Conference on Uncertainty in Structural Dynamics, pp. 3191 - 3203*. Leuven, Belgium, 2020.

[6] T.A. Harris and M.N. Kotzalas. *Essential Concepts of Bearing Technology: Rolling Bearing Analysis, Fifth Edition*. CRC Press, 2006. ISBN: 9780429621024.

[7] SKF. *Rolling bearings*. SKF Group 2018.

[8] P. Gao, L. Hou, R. Yang, and Y. Chen. "Local defect modelling and nonlinear dynamic analysis for the inter-shaft bearing in a dual-rotor system." In: *Applied Mathematical Modelling* 68 (2019), pp. 29–47. ISSN: 0307-904X. DOI: https://doi.org/10.1016/j.apm.2018.11.014.

[9] S.W. Hong and T. Canh. "Rolling-element bearing modeling: A review." In: *International Journal of Precision Engineering and Manufacturing* 17 (Dec. 2016), pp. 1729–1749. DOI: 10.1007/s12541-016-0200-z.

[10] Y. Guo and R.G. Parker. "Stiffness matrix calculation of rolling element bearings using a finite element/contact mechanics model." In: *Mechanism and Machine Theory* 51 (2012), pp. 32 –45. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2011.12.006.

[11] F. Pizzolante, M. Battarra, G. D'Elia, and E. Mucchi. "A rattle index formulation for single and multiple branch geartrains." In: *Mechanism and Machine Theory* 158 (2021), p. 104246. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2021.104246.

[12] C.S. Sunnersjö. "Varying compliance vibrations of rolling bearings." In: *Journal of Sound and Vibration* 58.3 (1978), pp. 363–373. ISSN: 0022-460X. DOI: https://doi.org/10.1016/S0022-460X(78)80044-3.

[13] C. Lei, F. Li, B. Gong, and X. Jia. "An Integrated Model to Characterize Comprehensive Stiffness of Angular Contact Ball Bearings." In: *Mathematical Problems in Engineering* 2020 (Apr. 2020), pp. 1–12. DOI: 10.1155/2020/4951828.

[14] T.A. Harris and M.N. Kotzalas. *Advanced Concepts of Bearing Technology: Rolling Bearing Analysis, Fifth Edition*. CRC Press, 2006. ISBN: 9781420006582.

[15] X. Hao, X. Gu, X. Zhou, X. Liao, and Q. Han. "Distribution characteristics of stress and displacement of rings of cylindrical roller bearing." In: *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 233.12 (2019), pp. 4348–4358. DOI: 10.1177/0954406218820551.

[16] J. Liu, C. Tang, H. Wu, Z. Xu, and L. Wang. "An analytical calculation method of the load distribution and stiffness of an angular contact ball bearing." In: *Mechanism and Machine Theory* 142 (2019), p. 103597. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2019.103597.

[17] R. Lostado-Lorza, Roberto Fernandez, and Bryan Mac Donald. "Determination of the contact stresses in double-row tapered roller bearings using the finite element method, experimental analysis and analytical models." In: *Journal of Mechanical Science and Technology* 29 (Nov. 2015), pp. 4645–4656. DOI: 10.1007/s12206-015-1010-4.

[18] F. Sadeghi, B. Jalalahmadi, T.S. Slack, N. Raje, and N.K. Arakere. "A Review of Rolling Contact Fatigue." In: *Journal of Tribology* 131.4 (Sept. 2009). ISSN: 0742-4787. DOI: 10.1115/1.3209132.

[19] G. Stachowiak and A.W. Batchelor. *Engineering Tribology*. Engineering Tribology. Elsevier Science, 2011. ISBN: 9780080531038.

[20] I. El-Thalji and E. Jantunen. "A summary of fault modelling and predictive health monitoring of rolling element bearings." In: *Mechanical Systems and Signal Processing* 60-61 (2015), pp. 252–272. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2015.02.008.

[21] S. Singh, C.Q. Howard, and C.H. Hansen. "An extensive review of vibration modelling of rolling element bearings with localised and extended defects." In: *Journal of Sound and Vibration* 357 (2015), pp. 300–330. ISSN: 0022-460X. DOI: https://doi.org/10.1016/j.jsv.2015.04.037.

[22] Y. Jiang, W. Huang, J. Luo, and W. Wang. "An improved dynamic model of defective bearings considering the three-dimensional geometric relationship between the rolling element and defect area." In: *Mechanical Systems and Signal Processing* 129 (2019), pp. 694–716. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2019.04.056.

[23] R.B. Randall and J. Antoni. "Rolling element bearing diagnostics—A tutorial." In: *Mechanical Systems and Signal Processing* 25.2 (2011), pp. 485–520. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2010.07.017.

[24] N. Sawalhi and R.B. Randall. "Simulating gear and bearing interactions in the presence of faults: Part I. The combined gear bearing dynamic model and the simulation of localised bearing faults." In: *Mechanical Systems and Signal Processing* 22.8 (2008), pp. 1924–1951. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2007.12.001.

[25] E. Soave, G. D'Elia, M. Cocconcelli, and M. Battarra. "Blind deconvolution criterion based on Fourier–Bessel series expansion for rolling element bearing diagnostics." In: *Mechanical Systems and Signal Processing* 169 (2022), p. 108588. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2021.108588.

[26] J.Z. Sikorska, M. Hodkiewicz, and L. Ma. "Prognostic modelling options for remaining useful life estimation by industry." In: *Mechanical Systems and Signal Processing* 25.5 (2011), pp. 1803–1836. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2010.11.018.

[27] A. Cubillo, S. Perinpanayagam, and M. Esperon Miguez. "A review of physics-based models in prognostics: Application to gears and bearings of rotating machinery." In: *Advances in Mechanical Engineering* 8 (Aug. 2016). DOI: 10.1177/1687814016664660.

[28] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin. "Machinery health prognostics: A systematic review from data acquisition to RUL prediction." In: *Mechanical Systems and Signal Processing* 104 (2018), pp. 799–834. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2017.11.016.

[29] J.H. MacConnell. "ISHM and Design: A review of the benefits of the ideal ISHM system." In: *2007 IEEE Aerospace Conference*. 2007, pp. 1–18. DOI: 10.1109/AERO.2007.352834.

[30] J. Lee, H. Qiu, G. Yu, and J. Lin. "Rexnord Technical Services, 'Bearing Data Set', IMS, University of Cincinnati, NASA Ames Prognostics Data Repository." In: (2007). URL: http://ti.arc.nasa.gov/project/prognostic-data-repository..

[31] T. Williams, X. Ribadeneira, S. Billington, and T. Kurfess. "Rolling bearing diagnostics in run-to-failure lifetime testing." In: *Mechanical Systems and Signal Processing* 15.5 (2001), pp. 979–993. ISSN: 0888-3270. DOI: https://doi.org/10.1006/mssp.2001.1418.

[32] P.K. Gupta. *Advanced Dynamics of Rolling Elements*. Springer-Verlag, 1984.

[33] T. L. H. Walford and B. J. Stone. "The Measurement of the Radial Stiffness of Rolling Element Bearings under Oscillating Conditions." In: *Journal of Mechanical Engineering Science* 22.4 (1980), pp. 175–181. DOI: 10.1243/JMES_JOUR_1980_022_035_02.

[34] J. Kraus, J. J. Blech, and S. G. Braun. "In Situ Determination of Rolling Bearing Stiffness and Damping by Modal Analysis." In: *Journal of Vibration, Acoustics, Stress, and Reliability in Design* 109.3 (July 1987), pp. 235–240. ISSN: 0739-3717. DOI: 10.1115/1.3269426.

[35] R. Madoliat and M. F. Ghanati. "Theoretical and Experimental Study of Spindle Ball Bearing Nonlinear Stiffness." In: *Journal of Mechanics* 29.4 (Aug. 2013), pp. 633–642. ISSN: 1811-8216. DOI: 10.1017/jmech.2013.48.

[36] N.J. Ali and J.M. García. "Experimental studies on the dynamic characteristics of rolling element bearings." In: *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology* 224.7 (2010), pp. 659–666. DOI: 10.1243/13506501JET698.

[37] W. Jacobs, R. Boonen, P. Sas, and D. Moens. "Measuring the rigid body behaviour of a deep groove ball bearing setup." In: vol. 1. 2012, 715 – 725.

[38] W. Jacobs, R. Boonen, P. Sas, and D. Moens. "Estimating Rolling Element Bearing Stiffness Under Different Operational Conditions Through Modal Analysis." In: *Advances in Condition Monitoring of Machinery in Non-Stationary Operations*. Ed. by G. Dalpiaz, R. Rubini, G. D'Elia, M. Cocconcelli, F. Chaari, R. Zimroz, W. Bartelmus, and M. Haddar. Springer Berlin Heidelberg, 2014, pp. 103–113. ISBN: 978-3-642-39348-8.

[39] A. B. Jones. "A General Theory for Elastically Constrained Ball and Radial Roller Bearings Under Arbitrary Load and Speed Conditions." In: *Journal of Basic Engineering* 82.2 (June 1960), pp. 309–320. ISSN: 0021-9223. DOI: 10.1115/1.3662587.

[40] M.F. While. "Rolling Element Bearing Vibration Transfer Characteristics: Effect of Stiffness." In: *Journal of Applied Mechanics* 46.3 (Sept. 1979), pp. 677–684. ISSN: 0021-8936. DOI: 10.1115/1.3424626.

[41] A. Palmgren, G. Palmgren, and inc SKF Industries. *Ball and Roller Bearing Engineering*. SKF Industries, 1959.

[42] E.P. Gargiulo Jr. "Simple Way to Estimate Bearing Stiffness." In: *Machine Design* 52.17 (1980), pp. 107–110.

[43] T.C. Lim and R. Singh. "Vibration transmission through rolling element bearings, part I: Bearing stiffness formulation." In: *Journal of Sound and Vibration* 139.2 (1990), pp. 179–199. ISSN: 0022-460X. DOI: https://doi.org/10.1016/0022-460X(90)90882-Z.

[44] T.C. Lim and R. Singh. "Vibration transmission through rolling element bearings, part II: System studies." In: *Journal of Sound and Vibration* 139.2 (1990), pp. 201–225. ISSN: 0022-460X. DOI: https://doi.org/10.1016/0022-460X(90)90883-2.

[45] T.C. Lim and R. Singh. "Vibration transmission through rolling element bearings. Part III: Geared rotor system studies." In: *Journal of Sound and Vibration* 151.1 (1991), pp. 31–54. ISSN: 0022-460X. DOI: https://doi.org/10.1016/0022-460X(91)90650-9.

[46] T.C. Lim and R. Singh. "Vibration transmission through rolling element bearings, part IV: statistical energy analysis." In: *Journal of Sound and Vibration* 153.1 (1992), pp. 37–50. ISSN: 0022-460X. DOI: https://doi.org/10.1016/0022-460X(92)90625-8.

[47] A. Gunduz and R. Singh. "Stiffness matrix formulation for double row angular contact ball bearings: Analytical development and validation." In: *Journal of Sound and Vibration* 332.22 (2013), pp. 5898–5916. ISSN: 0022-460X. DOI: https://doi.org/10.1016/j.jsv.2013.04.049.

[48] T. Xu, L. Yang, W. Wu, Y. Han, E. Xu, and K. Wang. "The stiffness characteristics analysis of the duplex angular contact ball bearings based on a comprehensive multi-degree-of-freedom mathematical model." In: *Applied Mathematical Modelling* 106 (2022), pp. 601–626. ISSN: 0307-904X. DOI: https://doi.org/10.1016/j.apm.2022.02.017.

[49] D. Petersen, C. Howard, and Z. Prime. "Varying stiffness and load distributions in defective ball bearings: Analytical formulation and application to defect size estimation." In: *Journal of Sound and Vibration* 337 (2015), pp. 284–300. ISSN: 0022-460X. DOI: https://doi.org/10.1016/j.jsv.2014.10.004.

[50] H. Cheng, Y. Zhang, W. Lu, and Z. Yang. "Research on time-varying stiffness of bearing based on local defect and varying compliance coupling." In: *Measurement* 143 (2019), pp. 155–179. ISSN: 0263-2241. DOI: https://doi.org/10.1016/j.measurement.2019.04.079.

[51] B. Fang, J. Zhang, K. Yan, J. Hong, and M. Yu Wang. "A comprehensive study on the speed-varying stiffness of ball bearing under different load conditions." In: *Mechanism and Machine Theory* 136 (2019), pp. 1–13. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2019.02.012.

[52] T.J. Royston and I. Basdogan. "Vibration transmission through self-aligning (spherical) rolling element bearings: theory and experiment." In: *Journal of Sound and Vibration* 215.5 (1998), pp. 997–1014. ISSN: 0022-460X. DOI: https://doi.org/10.1006/jsvi.1998.9999.

[53] B. Fang, K. Yan, J. Hong, and J. Zhang. "A comprehensive study on the off-diagonal coupling elements in the stiffness matrix of the angular contact ball bearing and their influence on the dynamic characteristics of the rotor system." In: *Mechanism and Machine Theory* 158 (2021), p. 104251. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2021.104251.

[54] Q. Niu, Y. Li, Y. Zhu, S. Pei, Y. Yin, and D. Wang. "Analytical Determination and Influence Analysis of Stiffness Matrix of Ball Bearing under Different Load Conditions." In: *Machines* 10.4 (). DOI: 10.3390/machines10040238.

[55] C. Natali, M. Battarra, G. Dalpiaz, and E. Mucchi. "A critical review on FE-based methods for mesh stiffness estimation in spur gears." In: *Mechanism and Machine Theory* 161 (2021), p. 104319. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2021.104319.

[56] H. Zhao. "Analysis of Load Distributions Within Solid and Hollow Roller Bearings." In: *Journal of Tribology* 120.1 (Jan. 1998), pp. 134–139. ISSN: 0742-4787. DOI: 10.1115/1.2834176.

[57] N. Demirhan and B. Kanber. "Stress and Displacement Distributions on Cylindrical Roller Bearing Rings Using FEM." In: *Mechanics Based Design of Structures and Machines* 36.1 (2008), pp. 86–102. DOI: 10.1080/15397730701842537.

[58] L. Kania. "Modelling of rollers in calculation of slewing bearing with the use of finite elements." In: *Mechanism and Machine Theory* 41.11 (2006), pp. 1359 –1376. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2005.12.007.

[59] A. Daidié, Z. Chaib, and A. Ghosn. "3D Simplified Finite Elements Analysis of Load and Contact Angle in a Slewing Ball Bearing." In: *Journal of Mechanical Design* 130.8 (July 2008). 082601. ISSN: 1050-0472. DOI: 10.1115/1.2918915.

[60] M. Krynke, L. Kania, and Eugeniusz M. "Modelling the Contact between the Rolling Elements and the Raceways of Bulky Slewing Bearings." In: *Key Engineering Materials* 490 (Jan. 2012), pp. 166–178. DOI: 10.4028/www.scientific.net/KEM.490.166.

[61] J. Aguirrebeitia, M. Abasolo, R. Avilés, and I. Fernandez de Bustos. "General static load-carrying capacity for the design and selection of four contact point slewing bearings: Finite element calculations and theoretical model validation." In: *Finite Elements in Analysis and Design* 55 (Aug. 2012), pp. 23–30. DOI: 10.1016/j.finel.2012.02.002.

[62] P. Göncz, M. Drobne, and S. Glodež. "Computational model for determination of dynamic load capacity of large three-row roller slewing bearings." In: *Engineering Failure Analysis* 32 (2013), pp. 44 –53. ISSN: 1350-6307. DOI: https://doi.org/10.1016/j.engfailanal.2013.02.030.

[63] Hua W., Peiyu H., Bitao P., and Xuehai G. "A new computational model of large three-row roller slewing bearings using nonlinear springs." In: *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 231.20 (2017), pp. 3831–3839. DOI: 10.1177/0954406217704223.

[64] L. Molnár, K. Váradi, G. Bódai, P. Zwierczyk, and L. Oroszváry. "Simplified modeling for needle roller bearings to analyze engineering structures by FEM." In: *Periodica Polytechnica Mechanical Engineering* 54.1 (2010), pp. 27–33. DOI: 10.3311/pp.me.2010-1.05.

[65] F. Massi, N. Bouscharain, S. Milana, G. Le Jeune, Y. Maheo, and Y. Berthier. "Degradation of high loaded oscillating bearings: Numerical analysis and comparison with experimental observations." In: *Wear* 317.1 (2014), pp. 141 –152. ISSN: 0043-1648. DOI: https://doi.org/10.1016/j.wear.2014.06.004.

[66] R.Fernandez Martinez, R. Lostado Lorza, A.A. Santos Delgado, and N.O. Piedra Pullaguari. "Optimizing presetting attributes by softcomputing techniques to improve tapered roller bearings working conditions." In: *Advances in Engineering Software* 123 (2018), pp. 13 –24. ISSN: 0965-9978. DOI: https://doi.org/10.1016/j.advengsoft.2018.05.005.

[67] S. Murer, F. Bogard, L. Rasolofondraibe, B. Pottier, and P. Marconnet. "Determination of loads transmitted by rolling elements in a roller bearing using capacitive probes: Finite element validation." In: *Mechanical Systems and Signal Processing* 54-55 (2015), pp. 306 –313. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2014.07.006.

[68] S. Li. "A mathematical model and numeric method for contact analysis of rolling bearings." In: *Mechanism and Machine Theory* 119 (2018), pp. 61 –73. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2017.08.020.

[69] P.D. McFadden and J.D. Smith. "Model for the vibration produced by a single point defect in a rolling element bearing." In: *Journal of Sound and Vibration* 96.1 (1984), pp. 69–82. ISSN: 0022-460X. DOI: https://doi.org/10.1016/0022-460X(84)90595-9.

[70] P.D. McFadden and J.D. Smith. "The vibration produced by multiple point defects in a rolling element bearing." In: *Journal of Sound and Vibration* 98.2 (1985), pp. 263–273. ISSN: 0022-460X. DOI: https://doi.org/10.1016/0022-460X(85)90390-6.

[71] N. Tandon and A. Choudhury. "An analytical model for the prediction of the vibration response of rolling element bearings due to a localized defect." In: *Journal of Sound and Vibration* 205.3 (1997), pp. 275–292. ISSN: 0022-460X. DOI: https://doi.org/10.1006/jsvi.1997.1031.

[72] Y.F. Wang and P.J. Kootsookos. "Modeling of low shaft speed bearing faults for condition monitoring." In: *Mechanical Systems and Signal Processing* 12.3 (1998), pp. 415–426. ISSN: 0888-3270. DOI: https://doi.org/10.1006/mssp.1997.0149.

[73] D. Brie. "Modelling of the spalled rolling element bearing vibration signal: an overview and some new results." In: *Mechanical Systems and Signal Processing* 14.3 (2000), pp. 353–369. ISSN: 0888-3270. DOI: https://doi.org/10.1006/mssp.1999.1237.

[74] R.B. Randall, J. Antoni, and S. Chobsaard. "The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals." In: *Mechanical Systems and Signal Processing* 15.5 (2001), pp. 945–962. ISSN: 0888-3270. DOI: https://doi.org/10.1006/mssp.2001.1415.

[75] S. Sassi, B. Badri, and M. Thomas. "A Numerical Model to Predict Damaged Bearing Vibrations." In: *Journal of Vibration and Control - J Vib Control* 13 (Nov. 2007), pp. 1603–1628. DOI: 10.1177/1077546307080040.

[76] A. Rafsanjani, S. Abbasion, A. Farshidianfar, and H. Moeenfard. "Nonlinear dynamic modeling of surface defects in rolling element bearing systems." In: *Journal of Sound and Vibration* 319.3 (2009), pp. 1150–1174. ISSN: 0022-460X. DOI: https://doi.org/10.1016/j.jsv.2008.06.043.

[77] M. Behzad, A.R. Bastami, and D. Mba. "A New Model for Estimating Vibrations Generated in the Defective Rolling Element Bearings." In: *Journal of Vibration and Acoustics* 133.4 (Apr. 2011). ISSN: 1048-9002. DOI: 10.1115/1.4003595.

[78] H. Cao, L. Niu, Z. He, and L. Yamin. "Dynamic Modeling and Vibration Response Simulation for High Speed Rolling Ball Bearings With Localized Surface Defects in Raceways." In: *Journal of Manufacturing Science and Engineering* 136 (May 2014). DOI: 10.1115/1.4027334.

[79] P. K. Gupta. "Dynamics of Rolling-Element Bearings—Part IV: Ball Bearing Results." In: *Journal of Lubrication Technology* 101.3 (July 1979), pp. 319–326. ISSN: 0022-2305. DOI: 10.1115/1.3453364.

[80] H. Cao, S. Su, X. Jing, and D. Li. "Vibration mechanism analysis for cylindrical roller bearings with single/multi defects and compound faults." In: *Mechanical Systems and Signal Processing* 144 (2020), p. 106903. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2020.106903.

[81] G. Kogan, R. Klein, A. Kushnirsky, and J. Bortman. "Toward a 3D dynamic model of a faulty duplex ball bearing." In: *Mechanical Systems and Signal Processing* 54-55 (2015), pp. 243–258. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2014.07.020.

[82] S. Khanam, N. Tandon, and J.K. Dutt. "Multi-Event Excitation Force Model for Inner Race Defect in a Rolling Element Bearing." In: *Journal of Tribology* 138 (Aug. 2015). DOI: 10.1115/1.4031394.

[83] R. Yang, Y. Jin, L. Hou, and Y. Chen. "Study for ball bearing outer race characteristic defect frequency based on nonlinear dynamics analysis." In: *Nonlinear Dynamics* 90 (Oct. 2017), pp. 1–16. DOI: 10.1007/s11071-017-3692-x.

[84] Y. Qin, C. Li, F. Cao, and H. Chen. "A fault dynamic model of high-speed angular contact ball bearings." In: *Mechanism and Machine Theory* 143 (2020), p. 103627. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2019.103627.

[85] T. Gao and S. Cao. "Paroxysmal impulse vibration phenomena and mechanism of a dual–rotor system with an outer raceway defect of the inter-shaft bearing." In: *Mechanical Systems and Signal Processing* 157 (2021), p. 107730. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2021.107730.

[86] R. Yang, Z. Zhang, and Y. Chen. "Analysis of vibration signals for a ball bearing-rotor system with raceway local defects and rotor eccentricity." In: *Mechanism and Machine Theory* 169 (2022), p. 104594. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2021.104594.

[87] S. Fukata, E.H. Gad, T. Kondou, T. Ayabe, and H. Tamura. "On the Radial Vibration of Ball Bearings : Computer Simulation." In: *Bulletin of JSME* 28.239 (1985), pp. 899–904. DOI: 10.1299/jsme1958.28.899.

[88] D. Petersen, C. Howard, N. Sawalhi, A. Moazen Ahmadi, and S. Singh. "Analysis of bearing stiffness variations, contact forces and vibrations in radially loaded double row rolling element bearings with raceway defects." In: *Mechanical Systems and Signal Processing* 50-51 (2015), pp. 139–160. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2014.04.014.

[89] L. Cui, Y. Zhang, F. Zhang, J. Zhang, and S. Lee. "Vibration response mechanism of faulty outer race rolling element bearings for quantitative analysis." In: *Journal of Sound and Vibration* 364 (2016), pp. 67–76. ISSN: 0022-460X. DOI: https://doi.org/10.1016/j.jsv.2015.10.015.

[90] F. Larizza, A. Moazen-Ahmadi, C. Howard, and S. Grainger. "The importance of bearing stiffness and load when estimating the size of a defect in a rolling element bearing." In: *Structural Health Monitoring* 18 (Oct. 2018), p. 147592171880880. DOI: 10.1177/1475921718808805.

[91] F. Larizza, C. Howard, S. Grainger, and W. Wang. "A nonlinear dynamic vibration model of a defective bearing: the importance of modelling the angle of the leading and trailing edges of a defect." In: *Structural Health Monitoring* (Nov. 2020), p. 147592172096395. DOI: 10.1177/1475921720963950.

[92] M.S. Patil, Jose Mathew, P.K. Rajendrakumar, and Sandeep Desai. "A theoretical model to predict the effect of localized defect on vibrations associated with ball bearing." In: *International Journal of Mechanical Sciences* 52.9 (2010). Special Issue on Advances in Materials and Processing Technologies, pp. 1193–1201. ISSN: 0020-7403. DOI: https://doi.org/10.1016/j.ijmecsci.2010.05.005.

[93] J. Liu, Y. Shao, and T.C. Lim. "Vibration analysis of ball bearings with a localized defect applying piecewise response function." In: *Mechanism and Machine Theory* 56 (2012), pp. 156–169. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2012.05.008.

[94] A. Moazen Ahmadi, D. Petersen, and C. Howard. "A nonlinear dynamic vibration model of defective bearings – The importance of modelling the finite size of rolling elements." In: *Mechanical Systems and Signal Processing* 52-53 (2015), pp. 309–326. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2014.06.006.

[95] J. Liu and Y. Shao. "Dynamic modeling for rigid rotor bearing systems with a localized defect considering additional deformations at the sharp edges." In: *Journal of Sound and Vibration* 398 (2017), pp. 84–102. ISSN: 0022-460X. DOI: https://doi.org/10.1016/j.jsv.2017.03.007.

[96] J. Liu, H. Wu, and Y. Shao. "A theoretical study on vibrations of a ball bearing caused by a dent on the races." In: *Engineering Failure Analysis* 83 (2018), pp. 220–229. ISSN: 1350-6307. DOI: https://doi.org/10.1016/j.engfailanal.2017.10.006.

[97] J. Liu. "A dynamic modelling method of a rotor-roller bearing-housing system with a localized fault including the additional excitation zone." In: *Journal of Sound and Vibration* 469 (2020), p. 115144. ISSN: 0022-460X. DOI: https://doi.org/10.1016/j.jsv.2019.115144.

[98] F. Kong, W. Huang, Y. Jiang, W. Wang, and X. Zhao. "A Vibration Model of Ball Bearings with a Localized Defect Based on the Hertzian Contact Stress Distribution." In: *Shock and Vibration* 2018 (Apr. 2018), pp. 1–14. DOI: 10.1155/2018/5424875.

[99] S. Gao, S. Chatterton, P. Pennacchi, and F. Chu. "Behaviour of an angular contact ball bearing with three-dimensional cubic-like defect: A comprehensive non-linear dynamic model for predicting vibration response." In: *Mechanism and Machine Theory* 163 (2021), p. 104376. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2021.104376.

[100] Y.H. Wijnant, J.A. Wensing, and G.C. Nijen. "The influence of lubrication on the dynamic behaviour of ball bearings." In: *Journal of Sound and Vibration* 222.4 (1999), pp. 579–596. ISSN: 0022-460X. DOI: https://doi.org/10.1006/jsvi.1998.2068. URL: https://www.sciencedirect.com/science/article/pii/S0022460X98920685.

[101] J. Sopanen and A. Mikkola. "Dynamic Model of a Deep Groove Ball Bearing Including Localized and Distributed Defects, Part 1: Theory." In: *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics* 217(K) (Sept. 2003), pp. 201–211. DOI: 10.1243/14644190360713551.

[102] J. Sopanen and A. Mikkola. "Dynamic model of a deep-groove ball bearing including localized and distributed defects. Part 2: Implementation and results." In: *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics* 217 (Sept. 2003), pp. 213–223. DOI: 10.1243/14644190360713560.

[103] D.S. Shah and V.N. Patel. "A dynamic model for vibration studies of dry and lubricated deep groove ball bearings considering local defects on races." In: *Measurement* 137 (2019), pp. 535–555. ISSN: 0263-2241. DOI: https://doi.org/10.1016/j.measurement.2019.01.097.

[104] P. Yan, C. Yan, K. Wang, F. Wang, and L. Wu. "5-DOF Dynamic Modeling of Rolling Bearing with Local Defect considering Comprehensive Stiffness under Isothermal Elastohydrodynamic Lubrication." In: *Shock and Vibration* 2020 (2020), p. 9310278. ISSN: 1070-9622. DOI: 10.1155/2020/9310278.

[105] M. Luo, Y. Guo, H. Andre, X. Wu, and J. Na. "Dynamic modeling and quantitative diagnosis for dual-impulse behavior of rolling element bearing with a spall on inner race." In: *Mechanical Systems and Signal Processing* 158 (2021), p. 107711. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2021.107711.

[106] S. Mufazzal, S.M. Muzakkir, and S. Khanam. "Theoretical and experimental analyses of vibration impulses and their influence on accurate diagnosis of ball bearing with localized outer race defect." In: *Journal of Sound and Vibration* 513 (2021), p. 116407. ISSN: 0022-460X. DOI: https://doi.org/10.1016/j.jsv.2021.116407.

[107] V. Parmar, H. Vy, and S. Harsha. "Nonlinear vibration response analysis of a double-row self-aligning ball bearing due to surface imperfections." In: *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics* 234 (May 2020), p. 146441932092491. DOI: 10.1177/1464419320924918.

[108] V. Parmar, V. Huzur Saran, and S.P. Harsha. "Effect of dynamic misalignment on the vibration response, trajectory followed and defect-depth achieved by the rolling-elements in a double-row spherical rolling-element bearing." In: *Mechanism and Machine Theory* 162 (2021), p. 104366. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2021.104366.

[109] Z. Kiral and H. Karagülle. "Simulation and analysis of vibration signals generated by rolling element bearing with defects." In: *Tribology International* 36.9 (2003), pp. 667–678. ISSN: 0301-679X. DOI: https://doi.org/10.1016/S0301-679X(03)00010-0.

[110] Z. Kıral and H. Karagülle. "Vibration analysis of rolling element bearings with various defects under the action of an unbalanced force." In: *Mechanical Systems and Signal Processing* 20.8 (2006), pp. 1967–1991. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2005.05.001.

[111] Y. Shao, W. Tu, and F. Gu. "A simulation study of defects in a rolling element bearing using FEA." In: *ICCAS 2010*. 2010, pp. 596–599. DOI: 10.1109/ICCAS.2010.5669813.

[112] A. Utpat. "Vibration Signature analysis of defective deep groove ball bearing by Numerical and Experimental approach." In: *International Journal of Science and Engineering Research* 4.6 (2013), 592 – 598.

[113] S. Singh, U. G. Köpke, C.Q. Howard, and D. Petersen. "Analyses of contact forces and vibration response for a defective rolling element bearing using an explicit dynamics finite element model." In: *Journal of Sound and Vibration* 333.21 (2014), pp. 5356–5377. ISSN: 0022-460X. DOI: https://doi.org/10.1016/j.jsv.2014.05.011.

[114] M. Tadina and M. Boltežar. "Improved model of a ball bearing for the simulation of vibration signals due to faults during run-up." In: *Journal of Sound and Vibration* 330.17 (2011). Computational Aero-Acoustics (CAA) for Aircraft Noise Prediction – Part B, pp. 4287–4301. ISSN: 0022-460X. DOI: https://doi.org/10.1016/j.jsv.2011.03.031.

[115] N. Sawalhi and R.B. Randall. "Simulating gear and bearing interactions in the presence of faults: Part II: Simulation of the vibrations produced by extended bearing faults." In: *Mechanical Systems and Signal Processing* 22.8 (2008), pp. 1952–1966. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2007.12.002.

[116] X.S. Si, W. Wang, C.H. Hu, and D.H. Zhou. "Remaining useful life estimation – A review on the statistical data driven approaches." In: *European Journal of Operational Research* 213.1 (2011), pp. 1–14. ISSN: 0377-2217. DOI: https://doi.org/10.1016/j.ejor.2010.11.018.

[117] X.S. Si, W. Wang, C.H. Hu, M.Y. Chen, and D.H. Zhou. "A Wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation." In: *Mechanical Systems and Signal Processing* 35.1 (2013), pp. 219–237. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2012.08.016.

[118] A. Heng, S. Zhang, A.C.C. Tan, and J. Mathew. "Rotating machinery prognostics: State of the art, challenges and opportunities." In: *Mechanical Systems and Signal Processing* 23.3 (2009), pp. 724–739. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2008.06.009.

[119] R. Polikar. "Ensemble based systems in decision making." In: *IEEE Circuits and Systems Magazine* 6.3 (2006), pp. 21–45. DOI: 10.1109/MCAS.2006.1688199.

[120] T. Tinga. *Principles of Loads and Failure Mechanisms. Applications in Maintenance, Reliability and Design*. Jan. 2013. ISBN: 978-1-4471-4917-0.

[121] P. Paris and F. Erdogan. "A critical analysis of crack propagation laws." In: *J. Fluids Eng.* 85 (1963), 528–533.

[122] Y. Li, S. Billington, C. Zhang, T.R. Kurfess, S. Danyluk, and S. Liang. "Adaptive prognostics for rolling element bearing condition." In: *Mechanical Systems and Signal Processing* 13.1 (1999), pp. 103–113. ISSN: 0888-3270. DOI: https://doi.org/10.1006/mssp.1998.0183.

[123] Y. Li, T.R. Kurfess, and S.Y. Liang. "Stochastic prognostic for rolling element bearings." In: *Mechanical Systems and Signal Processing* 14.5 (2000), pp. 747–762. ISSN: 0888-3270. DOI: https://doi.org/10.1006/mssp.2000.1301.

[124] M.R. Hoeprich. "Rolling Element Bearing Fatigue Damage Propagation." In: *Journal of Tribology* 114.2 (Apr. 1992), pp. 328–333. ISSN: 0742-4787. DOI: 10.1115/1.2920891.

[125] M.N. Kotzalas and T.A. Harris. "Fatigue Failure Progression in Ball Bearings." In: *Journal of Tribology* 123.2 (Mar. 2000), pp. 238–242. ISSN: 0742-4787. DOI: 10.1115/1.1308013.

[126] S. Li, A. Kahraman, N. Anderson, and L.D. Wedeven. "A model to predict scuffing failures of a ball-on-disk contact." In: *Tribology International* 60 (2013), pp. 233–245. ISSN: 0301-679X. DOI: https://doi.org/10.1016/j.triboint.2012.11.007.

[127] Y. Lei, N. Li, S. Gontarz, J. Lin, S. Radkowski, and J. Dybala. "A model-based method for remaining useful life prediction of machinery." In: *IEEE Transactions on reliability* 65.3 (2016), pp. 1314–1326. DOI: 10.1109/TR.2016.2570568.

[128] B. Wang, Y. Lei, N. Li, and T. Yan. "Deep separable convolutional network for remaining useful life prediction of machinery." In: *Mechanical Systems and Signal Processing* 134 (2019), p. 106330. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2019.106330.

[129] L. Liao. "Discovering Prognostic Features Using Genetic Programming in Remaining Useful Life Prediction." In: *IEEE Transactions on Industrial Electronics* 61.5 (2014), pp. 2464–2472. DOI: 10.1109/TIE.2013.2270212.

[130] J. Sun, H. Zuo, W. Wang, and M.G. Pecht. "Prognostics uncertainty reduction by fusing on-line monitoring data based on a state-space-based degradation model." In: *Mechanical Systems and Signal Processing* 45.2 (2014), pp. 396–407. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2013.08.022.

[131] D. Xu, J. Huang, Q. Zhu, X. Chen, Y. Xu, and S. Wang. "Residual fatigue life prediction of ball bearings based on Paris law and RMS." In: 25.2 (2012), 320 – 327. DOI: 10.3901/CJME.2012.02.320.

[132] M. Corbetta, C. Sbarufatti, A. Manes, and M. Giglio. "Sequential Monte Carlo sampling for crack growth prediction providing for several uncertainties." In: July 2014.

[133] M. Corbetta, C. Sbarufatti, A. Manes, and M. Giglio. "Real-Time Prognosis of Crack Growth Evolution Using Sequential Monte Carlo Methods and Statistical Model Parameters." In: *IEEE Transactions on Reliability* 64.2 (2015), pp. 736–753. DOI: 10.1109/TR.2014.2366759.

[134] F. Cadini, E. Zio, and D. Avram. "Monte Carlo-based filtering for fatigue crack growth estimation." In: *Probabilistic Engineering Mechanics* 24.3 (2009), pp. 367–373. ISSN: 0266-8920. DOI: https://doi.org/10.1016/j.probengmech.2008.10.002.

[135] E. Zio and G. Peloni. "Particle filtering prognostic estimation of the remaining useful life of nonlinear components." In: *Reliability Engineering & System Safety* 96.3 (2011), pp. 403–409. ISSN: 0951-8320. DOI: https://doi.org/10.1016/j.ress.2010.08.009.

[136] Y. Lu, Q. Li, and S.Y. Liang. "Physics-based intelligent prognosis for rolling bearing with fault feature extraction." In: *The International Journal of Advanced Manufacturing Technology* 97.1 (2018), pp. 611–620. ISSN: 1433-3015. DOI: 10.1007/s00170-018-1959-0.

[137] L. Saidi, J. Ben Ali, M. Benbouzid, and E. Bechhofer. "An integrated wind turbine failures prognostic approach implementing Kalman smoother with confidence bounds." In: *Applied Acoustics* 138 (2018), pp. 199–208. ISSN: 0003-682X. DOI: https://doi.org/10.1016/j.apacoust.2018.04.005.

[138] J. Qiu, B.B. Seth, S.Y. Liang, and C. Zhang. "Damage mechanics approach for bearing lifetime prognostics." In: *Mechanical Systems and Signal Processing* 16.5 (2002), pp. 817–829. ISSN: 0888-3270. DOI: https://doi.org/10.1006/mssp.2002.1483.

[139] C.D. Begg, C. Byington, and K.P. Maynard. "Dynamic simulation of mechanical fault transition," in: *Proceedings of the 54th Meeting of the Society for Machinery Failure Prevention Technology* (2000), pp. 203–212.

[140] D. Chelidze and J. P Cusumano. "A dynamical systems approach to failure prognosis." In: *J. Vib. Acoust.* 126.1 (2004), pp. 2–8.

[141] S. Marble and B.P. Morton. "Predicting the remaining life of propulsion system bearings." In: *Proceedings of the 2006 IEEE Aerospace Conference* (2006).

[142] P.K. Gupta and E.V. Zaretsky. "New Stress-Based Fatigue Life Models for Ball and Roller Bearings." In: *Tribology Transactions* 61.2 (2018), pp. 304–324. DOI: 10.1080/10402004.2017.1319524.

[143] T. Slack and F. Sadeghi. "Explicit finite element modeling of subsurface initiated spalling in rolling contacts." In: *Tribology International* 43.9 (2010), pp. 1693–1702. ISSN: 0301-679X. DOI: https://doi.org/10.1016/j.triboint.2010.03.019.

[144] M.J. Pais and N.H. Kim. "Predicting fatigue crack growth under variable amplitude loadings with usage monitoring data." In: *Advances in Mechanical Engineering* 7.12 (2015). DOI: 10.1177/1687814015619135.

[145] L. Cui, X. Wang, H. Wang, and H. Jiang. "Remaining useful life prediction of rolling element bearings based on simulated performance degradation dictionary." In: *Mechanism and Machine Theory* 153 (2020), p. 103967. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2020.103967.

[146] T. Wang, Jianbo Y., D. Siegel, and J. Lee. "A similarity-based prognostics approach for Remaining Useful Life estimation of engineered systems." In: *2008 International Conference on Prognostics and Health Management*. 2008, pp. 1–6. DOI: 10.1109/PHM.2008.4711421.

[147] T. Wang, Z. Liu, and N. Mrad. "A Probabilistic Framework for Remaining Useful Life Prediction of Bearings." In: *IEEE Transactions on Instrumentation and Measurement* 70 (2021), pp. 1–12. DOI: 10.1109/TIM.2020.3029382.

[148] B. Yan, X. Ma, G. Huang, and Y. Zhao. "Two-stage physics-based Wiener process models for online RUL prediction in field vibration data." In: *Mechanical Systems and Signal Processing* 152 (2021), p. 107378. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2020.107378.

[149] K.A. Loparo C.H. Oppenheimer. "Physically based diagnosis and prognosis of cracked rotor shafts." In: *Component and Systems Diagnostics, Prognostics, and Health Management II* (2002), pp. 1–11.

[150] W. M. Ostachowicz and M. Krawczuk. "Coupled torsional and bending vibrations of a rotor with an open crack." In: *Archive of Applied Mechanics* 62.3 (1992), pp. 191–201. ISSN: 1432-0681. DOI: 10.1007/BF00787959.

[151] J. Singh, M. Azamfar, F. Li, and J. Lee. "A systematic review of machine learning algorithms for prognostics and health management of rolling element bearings: fundamentals, concepts and applications." In: *Measurement Science and Technology* 32.1 (2020), p. 012001. DOI: 10.1088/1361-6501/ab8df9.

[152] W. Caesarendra, A. Widodo, P.H. Thom, B.S. Yang, and J.D. Setiawan. "Combined Probability Approach and Indirect Data-Driven Method for Bearing Degradation Prognostics." In: *IEEE Transactions on Reliability* 60.1 (2011), pp. 14–20. DOI: 10.1109/TR.2011.2104716.

[153] Y. Qian, R. Yan, and S. Hu. "Bearing Degradation Evaluation Using Recurrence Quantification Analysis and Kalman Filter." In: *IEEE Transactions on Instrumentation and Measurement* 63.11 (2014), pp. 2599–2610. DOI: 10.1109/TIM.2014.2313034.

[154] Y. Qian and R. Yan. "Remaining Useful Life Prediction of Rolling Bearings Using an Enhanced Particle Filter." In: *IEEE Transactions on Instrumentation and Measurement* 64.10 (2015), pp. 2696–2707. DOI: 10.1109/TIM.2015.2427891.

[155] L. Cui, X. Wang, H. Wang, and J. Ma. "Research on Remaining Useful Life Prediction of Rolling Element Bearings Based on Time-Varying Kalman Filter." In: *IEEE Transactions on Instrumentation and Measurement* 69.6 (2020), pp. 2858–2867. DOI: 10.1109/TIM.2019.2924509.

[156] L. Cui, W. Li, X. Wang, D. Zhao, and H. Wang. "Comprehensive Remaining Useful Life Prediction for Rolling Element Bearings Based on Time-Varying Particle Filtering." In: *IEEE Transactions on Instrumentation and Measurement* 71 (2022), pp. 1–10. DOI: 10.1109/TIM.2022.3163167.

[157] N. Gebraeel and J. Pan. "Prognostic Degradation Models for Computing and Updating Residual Life Distributions in a Time-Varying Environment." In: *IEEE Transactions on Reliability* 57.4 (2008), pp. 539–550. DOI: 10.1109/TR.2008.928245.

[158] Y. Wang, C. Deng, J. Wu, and Y. Xiong. "Failure time prediction for mechanical device based on the degradation sequence." In: *Journal of Intelligent Manufacturing* 26.6 (2015), pp. 1181–1199. ISSN: 1572-8145. DOI: 10.1007/s10845-013-0849-4.

[159] Y. Jin X.and Sun, Z. Que, Y. Wang, and T.W.S. Chow. "Anomaly Detection and Fault Prognosis for Bearings." In: *IEEE Transactions on Instrumentation and Measurement* 65.9 (2016), pp. 2046–2054. DOI: 10.1109/TIM.2016.2570398.

[160] E. Soave. "Diagnostics and Prognostics of Rotating Machines Through Cyclostationary Methods and Machine Learning." PhD thesis. University of Ferrara, Department of Engineering, 2022.

[161] J.P. Kharoufeh and S.M. Cox. "Stochastic models for degradation-based reliability." In: *IIE Transactions* 37.6 (2005), pp. 533–542. DOI: 10.1080/07408170590929009.

[162] R.B. Chinnam and P. Baruah. "Autonomous diagnostics and prognostics in machining processes through competitive learning-driven HMM-based clustering." In: *International Journal of Production Research* 47.23 (2009), pp. 6739–6758. DOI: 10.1080/00207540802232930.

[163] E. Ramasso and T. Denoeux. "Making Use of Partial Knowledge About Hidden States in HMMs: An Approach Based on Belief Functions." In: *IEEE Transactions on Fuzzy Systems* 22.2 (2014), pp. 395–405. DOI: 10.1109/TFUZZ.2013.2259496.

[164] E. Soave, G. D'Elia, and G. Dalpiaz. "Prognostics of rotating machines through generalized Gaussian hidden Markov models." In: *Mechanical Systems and Signal Processing* 185 (2023), p. 109767. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2022.109767.

[165] K. Medjaher, D.A. Tobon-Mejia, and N. Zerhouni. "Remaining Useful Life Estimation of Critical Components With Application to Bearings." In: *IEEE Transactions on Reliability* 61.2 (2012), pp. 292–302. DOI: 10.1109/TR.2012.2194175.

[166] Q. Liu, M. Dong, W. Lv, X. Geng, and Y. Li. "A novel method using adaptive hidden semi-Markov model for multi-sensor monitoring equipment health prognosis." In: *Mechanical Systems and Signal Processing* 64-65 (2015), pp. 217–232. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2015.03.029.

[167] K.A. Doksum and A. Hóyland. "Models for Variable-Stress Accelerated Life Testing Experiments Based on Wiener Processes and the Inverse Gaussian Distribution." In: *Technometrics* 34.1 (1992), pp. 74–82. ISSN: 00401706.

[168] N. Li, Y. Lei, J. Lin, and S.X. Ding. "An Improved Exponential Model for Predicting Remaining Useful Life of Rolling Element Bearings." In: *IEEE Transactions on Industrial Electronics* 62.12 (2015), pp. 7762–7773. DOI: 10.1109/TIE.2015.2455055.

[169] Y. Wang, Y. Peng, Y. Zi, X. Jin, and K.L. Tsui. "A Two-Stage Data-Driven-Based Prognostic Approach for Bearing Degradation Problem." In: *IEEE Transactions on Industrial Informatics* 12.3 (2016), pp. 924–932. DOI: 10.1109/TII.2016.2535368.

[170] J. Wen, H. Gao, and J. Zhang. "Bearing Remaining Useful Life Prediction Based on a Nonlinear Wiener Process Model." In: *Shock and Vibration* 2018 (2018), p. 4068431. ISSN: 1070-9622. DOI: 10.1155/2018/4068431.

[171] N. Chen, Z.S. Ye, Y. Xiang, and L. Zhang. "Condition-based maintenance using the inverse Gaussian degradation model." In: *European Journal of Operational Research* 243.1 (2015), pp. 190–199. ISSN: 0377-2217. DOI: https://doi.org/10.1016/j.ejor.2014.11.029.

[172] W. Peng, Y.F. Li, Y.J. Yang, S.P. Zhu, and H.Z. Huang. "Bivariate Analysis of Incomplete Degradation Observations Based on Inverse Gaussian Processes and Copulas." In: *IEEE Transactions on Reliability* 65.2 (2016), pp. 624–639. DOI: 10.1109/TR.2015.2513038.

[173] D.R. Cox. "Regression Models and Life-Tables." In: *Journal of the Royal Statistical Society. Series B (Methodological)* 34.2 (1972), pp. 187–220. ISSN: 00359246.

[174] D. Banjevic and A. K. S. Jardine. "Calculation of reliability function and remaining useful life for a Markov failure time process." In: *IMA Journal of Management Mathematics* 17.2 (Apr. 2006), pp. 115–130. ISSN: 1471-678X. DOI: 10.1093/imaman/dpi029.

[175] L. Wang, L. Zhang, and X.Z. Wang. "Reliability estimation and remaining useful lifetime prediction for bearing based on proportional hazard model." In: *Journal of Central South University* 22.12 (2015), pp. 4625–4633. ISSN: 2227-5223. DOI: 10.1007/s11771-015-3013-9.

[176] C. Lu, L. Tao, and H. Fan. "An intelligent approach to machine component health prognostics by utilizing only truncated histories." In: *Mechanical Systems and Signal Processing* 42.1 (2014), pp. 300–313. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2013.08.025.

[177] L. Xiao, X. Chen, X. Zhang, and M. Liu. "A novel approach for bearing remaining useful life estimation under neither failure nor suspension histories condition." In: *Journal of Intelligent Manufacturing* 28 (Apr. 2015). DOI: 10.1007/s10845-015-1077-x.

[178] C. Sbarufatti, M. Corbetta, A. Manes, and M. Giglio. "Sequential Monte-Carlo sampling based on a committee of artificial neural networks for posterior state estimation and residual lifetime prediction." In: *International Journal of Fatigue* 83 (2016), pp. 10–23. ISSN: 0142-1123. DOI: https://doi.org/10.1016/j.ijfatigue.2015.05.017.

[179] Y. Shang, X. Tang, G. Zhao, P. Jiang, and T. Ran Lin. "A remaining life prediction of rolling element bearings based on a bidirectional gate recurrent unit and convolution neural network." In: *Measurement* 202 (2022), p. 111893. ISSN: 0263-2241. DOI: https://doi.org/10.1016/j.measurement.2022.111893.

[180] W. Li, Z. Shang, M. Gao, S. Qian, and Z. Feng. "Remaining useful life prediction based on transfer multi-stage shrinkage attention temporal convolutional network under variable working conditions." In: *Reliability Engineering & System Safety* 226 (2022), p. 108722. ISSN: 0951-8320. DOI: https://doi.org/10.1016/j.ress.2022.108722.

[181] E. Lughofer. *Evolving Fuzzy Systems: Methodologies, Advanced Concepts and Applications*. Vol. 266. Jan. 2011. ISBN: 978-3-642-18086-6. DOI: 10.1007/978-3-642-18087-3.

[182] J.S.R. Jang, C.T. Sun, and E. Mizutani. "Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence." In: *IEEE Transactions on automatic control* 42.10 (1997), pp. 1482–1484.

[183] C. Chen, G. Vachtsevanos, and M.E. Orchard. "Machine remaining useful life prediction: An integrated adaptive neuro-fuzzy and high-order particle filtering approach." In: *Mechanical Systems and Signal Processing* 28 (2012). Interdisciplinary and Integration Aspects in Structural Health Monitoring, pp. 597–607. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2011.10.009.

[184] D. Zurita, J.A. Carino, M. Delgado, and J.A. Ortega. "Distributed neuro-fuzzy feature forecasting approach for condition monitoring." In: *Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA)*. 2014, pp. 1–8. DOI: 10.1109/ETFA.2014.7005180.

[185] V.N. Vapnik. "An overview of statistical learning theory." In: *IEEE Transactions on Neural Networks* 10.5 (1999), pp. 988–999. DOI: 10.1109/72.788640.

[186] S. Dong and T. Luo. "Bearing degradation process prediction based on the PCA and optimized LS-SVM model." In: *Measurement* 46.9 (2013), pp. 3143–3152. ISSN: 0263-2241. DOI: https://doi.org/10.1016/j.measurement.2013.06.038.

[187] F. Sloukia, M. El Aroussi, H. Medromi, and M. Wahbi. "Bearings prognostic using Mixture of Gaussians Hidden Markov Model and Support Vector Machine." In: *2013 ACS International Conference on Computer Systems and Applications (AICCSA)*. 2013, pp. 1–4. DOI: 10.1109/AICCSA.2013.6616438.

[188] E. Fumeo, L. Oneto, and D. Anguita. "Condition Based Maintenance in Railway Transportation Systems Based on Big Data Streaming Analysis." In: *Procedia Computer Science* 53 (2015). INNS Conference on Big Data 2015 Program San Francisco, CA, USA 8-10 August 2015, pp. 437–446. ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2015.07.321.

[189] T.H. Loutas, D. Roulias, and G. Georgoulas. "Remaining Useful Life Estimation in Rolling Bearings Utilizing Data-Driven Probabilistic E-Support Vectors Regression." In: *IEEE Transactions on Reliability* 62.4 (2013), pp. 821–832. DOI: 10.1109/TR.2013.2285318.

[190] M. Zhao, B. Tang, and Q. Tan. "Bearing remaining useful life estimation based on time–frequency representation and supervised dimensionality reduction." In: *Measurement* 86 (2016), pp. 41–55. ISSN: 0263-2241. DOI: https://doi.org/10.1016/j.measurement.2015.11.047.

[191] P. Wang, B.D. Youn, and C. Hu. "A generic probabilistic framework for structural health prognostics and uncertainty management." In: *Mechanical Systems and Signal Processing* 28 (2012). Interdisciplinary and Integration Aspects in Structural Health Monitoring, pp. 622–637. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2011.10.019.

[192] J. Wang. "An Intuitive Tutorial to Gaussian Processes Regression." In: *ArXiv* abs/10862 (2020).

[193] S. Hong, Z.Zhou, E. Zio, and K. Hong. "Condition assessment for the performance degradation of bearing based on a combinatorial feature extraction method." In: *Digital Signal Processing* 27 (2014), pp. 159–166. ISSN: 1051-2004. DOI: https://doi.org/10.1016/j.dsp.2013.12.010.

[194] S.A. Aye and P.S. Heyns. "An integrated Gaussian process regression for prediction of remaining useful life of slow speed bearings based on acoustic emission." In: *Mechanical Systems and Signal Processing* 84 (2017), pp. 485–498. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2016.07.039.

[195] F. Di Maio, K.L. Tsui, and E. Zio. "Combining Relevance Vector Machines and exponential regression for bearing residual life estimation." In: *Mechanical Systems and Signal Processing* 31 (2012), pp. 405–427. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2012.03.011.

[196] A. Soualhi, H. Razik, G. Clerc, and D.D. Doan. "Prognosis of Bearing Failures Using Hidden Markov Models and the Adaptive Neuro-Fuzzy Inference System." In: *IEEE Transactions on Industrial Electronics* 61.6 (2014), pp. 2864–2874. DOI: 10.1109/TIE.2013.2274415.

[197] J. Yu. "Machine health prognostics using the Bayesian-inference-based probabilistic indication and high-order particle filtering framework." In: *Journal of Sound and Vibration* 358 (2015), pp. 97–110. ISSN: 0022-460X. DOI: https://doi.org/10.1016/j.jsv.2015.08.013.

[198] Y. Qian, R. Yan, and R.X. Gao. "A multi-time scale approach to remaining useful life prediction in rolling bearing." In: *Mechanical Systems and Signal Processing* 83 (2017), pp. 549–567. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2016.06.031.

[199] J. J. Montero Jimenez, Sé. Schwartz, R. Vingerhoeds, B. Grabot, and M. Salaün. "Towards multi-model approaches to predictive maintenance: A systematic literature survey on diagnostics and prognostics." In: *Journal of Manufacturing Systems* 56 (2020), pp. 539–557. ISSN: 0278-6125. DOI: https://doi.org/10.1016/j.jmsy.2020.07.008.

[200] H. Qiu, J. Lee, J. Lin, and G. Yu. "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics." In: *Journal of Sound and Vibration* 289.4 (2006), pp. 1066–1090. ISSN: 0022-460X. DOI: https://doi.org/10.1016/j.jsv.2005.03.007.

[201] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Chebel-Morello, N. Zerhouni, and C. Varnier. "PRONOSTIA: An experimental platform for bearings accelerated degradation tests." In: *Conference on Prognostics and Health Management.* (June 2012), pp. 1–8.

[202] W.A. Smith and R.B. Randall. "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study." In: *Mechanical Systems and Signal Processing* 64-65 (2015), pp. 100–131. DOI: 10.1016/j.ymssp.2013.10.007.

[203] D. Neupane and J. Seok. "Bearing fault detection and diagnosis using Case Western Reserve University dataset with deep learning approaches: A Review." In: *IEEE Access* 8 (2020), pp. 93155–93178. DOI: 10.1109/ACCESS.2020.2990528.

[204] B. Wang, Y. Lei, N. Li, and N. Li. "A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings." In: *IEEE Transactions on Reliability* 69.1 (2020), pp. 401–412. DOI: 10.1109/TR.2018.2882682.

[205] C. Lessmeier, O. Enge-Rosenblatt, C. Bayer, and D. Zimmer. "Data Acquisition and Signal Analysis from Measured Motor Currents for Defect Detection in Electromechanical Drive Systems." In: *Prognostics and Health Management Society (Hg.): European Conference of the PHM Society 2014 Proceedings* (2014).

[206] C. Lessmeier, J.K. Kimotho, D. Zimmer, and W. Sextro. "Condition Monitoring of Bearing Damage in Electromechanical Drive Systems by Using Motor Current Signals of Electric Motors: A Benchmark Data Set for Data-Driven Classification." In: *PHM Society European Conference* 3(1) (2016). DOI: https://doi.org/10.36001/phme.2016.v3i1.1577.

[207] V. Pandhare, J. Singh, and J. Lee. "Convolutional Neural Network Based Rolling-Element Bearing Fault Diagnosis for Naturally Occurring and Progressing Defects Using Time-Frequency Domain Features." In: *2019 Prognostics and System Health Management Conference (PHM-Paris)*. 2019, pp. 320–326. DOI: 10.1109/PHM-Paris.2019.00061.

[208] E. Bechhoefer. "MFPT Dataset, Society For Machinery Failure Prevention Technology." In: (Accessed: Dec. 22, 2022. [Online]). URL: https://mfpt.org/fault-datasets/..

[209] C. Sobie, C. Freitas, and M. Nicolai. "Simulation-driven machine learning: Bearing fault classification." In: *Mechanical Systems and Signal Processing* 99 (2018), pp. 403–419. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2017.06.025.

[210] A. Chen and T. R. Kurfess. "A new model for rolling element bearing defect size estimation." In: *Measurement* 114 (2018), pp. 144–149. ISSN: 0263-2241. DOI: https://doi.org/10.1016/j.measurement.2017.09.018.

[211] V. N. Patel, N. Tandon, and R. K. Pandey. "A Dynamic Model for Vibration Studies of Deep Groove Ball Bearings Considering Single and Multiple Defects in Races." In: *Journal of Tribology* 132.4 (Oct. 2010). ISSN: 0742-4787. DOI: 10.1115/1.4002333.

[212] X. Yuan, Y. Zhu, and Y. Zhang. "Multi-body vibration modelling of ball bearing–rotor system considering single and compound multi-defects." In: *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics* 228.2 (2014), pp. 199–212. DOI: 10.1177/1464419314522372.

[213] A. Rohani Bastami, A. Aasi, and H. A. Arghand. "Estimation of Remaining Useful Life of Rolling Element Bearings Using Wavelet Packet Decomposition and Artificial Neural Network." In: *Iranian Journal of Science and Technology, Transactions of Electrical Engineering* 43.1 (2019), pp. 233–245. ISSN: 2364-1827. DOI: 10.1007/s40998-018-0108-y.

[214] Siemens PLM Software Inc. *Simcenter 3D*. Version 2019.2. 2019.

[215] N.H. Kim. *Introduction to nonlinear finite element analysis*. 2015. DOI: 10.1007/978-1-4419-1746-1.

[216] Z. Javanbakht. *Advanced Finite Element Simulation with MSC Marc*. Jan. 2017. DOI: 10.1007/978-3-319-47668-1.

[217] Autodesk. *Autodesk Inventor*. Version 22.0. San Rafael, CA, 2018.

[218] G.J. De Salvo and J. A. Swanson. *ANSYS engineering analysis system user's manual*. Swanson Analysis Systems, 1985.

[219] M. Smith. *ABAQUS/Standard User's Manual, Version 6.9*. English. United States: Dassault Systèmes Simulia Corp, 2009.

[220] N. El-Abbasi and K.J. Bathe. "Stability and patch test performance of contact discretizations and a new solution algorithm." In: *Computers & Structures* 79.16 (2001), pp. 1473 –1486. ISSN: 0045-7949. DOI: https://doi.org/10.1016/S0045-7949(01)00048-7.

[221] R. H. Macneal and R. L. Harder. "A proposed standard set of problems to test finite element accuracy." In: *Finite Elements in Analysis and Design* 1.1 (1985), pp. 3 –20. ISSN: 0168-874X. DOI: https://doi.org/10.1016/0168-874X(85)90003-4.

[222] R. Lostado-Lorza, Rubén Escribano-Garcia, and Roberto Fernandez. "Optimization of operating conditions for a double-row tapered roller bearing." In: *International Journal of Mechanics and Materials in Design* 12 (May 2015). DOI: 10.1007/s10999-015-9311-4.

[223] D. Tonazzi, E. Houara Komba, F. Massi, G. Le Jeune, J.B. Coudert, Y. Maheo, and Y. Berthier. "Numerical analysis of contact stress and strain distributions for greased and ungreased high loaded oscillating bearings." In: *Wear* 376-377 (2017). 21st International Conference on Wear of Materials, pp. 1164 –1175. ISSN: 0043-1648. DOI: https://doi.org/10.1016/j.wear.2016.11.037.

[224] J. Liu, Y. Xu, and G. Pan. "A combined acoustic and dynamic model of a defective ball bearing." In: *Journal of Sound and Vibration* 501 (2021), p. 116029. ISSN: 0022-460X. DOI: https://doi.org/10.1016/j.jsv.2021.116029.

[225] A. B. Jones and T. A. Harris. "Analysis of a Rolling-Element Idler Gear Bearing Having a Deformable Outer-Race Structure." In: *Journal of Basic Engineering* 85.2 (June 1963), pp. 273–278. ISSN: 0021-9223. DOI: 10.1115/1.3656576.

[226] R. Lostado-Lorza, F.J. Ascacibar, A. Pernía-Espinoza, F. Alba-Elías, and J. Blanco. "Combining regression trees and the finite element method to define stress models of highly non-linear mechanical systems." In: *Journal of Strain Analysis for Engineering Design* 44 (Aug. 2009), pp. 491–502. DOI: 10.1243/03093247JSA497.

[227] R. Lostado-Lorza, R. Escribano-Garcia, R. Fernandez-Martinez, M. Illera-cueva, and B.J. Mac Donald. "Using the finite element method and data mining techniques as an alternative method to determine the maximum load capacity in tapered roller bearings." In: *Journal of Applied Logic* 24 (2017). SI:SOCO14, pp. 4 –14. ISSN: 1570-8683. DOI: https://doi.org/10.1016/j.jal.2016.11.009.

[228] J.P. Dedieu. "Newton-Raphson Method." In: *Encyclopedia of Applied and Computational Mathematics*. Ed. by Björn Engquist. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 1023–1028. ISBN: 978-3-540-70529-1. DOI: 10.1007/978-3-540-70529-1_374.

[229] S.S. Rao. *Mechanical Vibrations*. Pearson Education, Inc., 2017. ISBN: 9780134361307.

[230] A. L. Custódio, J. F. A. Madeira, A. I. F. Vaz, and L. N. Vicente. "Direct Multisearch for Multiobjective Optimization." In: *SIAM Journal on Optimization* 21.3 (2011), pp. 1109–1140. DOI: 10.1137/10079731X.

[231] L. Niu, H. Cao, H.Hou, B. Wu, Y. Lan, and X. Xiong. "Experimental observations and dynamic modeling of vibration characteristics of a cylindrical roller bearing with roller defects." In: *Mechanical Systems and Signal Processing* 138 (2020), p. 106553. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2019.106553.

[232] A. Gunduz, J.T. Dreyer, and R. Singh. "Effect of bearing preloads on the modal characteristics of a shaft-bearing assembly: Experiments on double row angular contact ball bearings." In: *Mechanical Systems and Signal Processing* 31 (2012), pp. 176–195. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2012.03.013.

[233] L. Bizarre, F. Nonato, and K.L. Cavalca. "Formulation of five degrees of freedom ball bearing model accounting for the nonlinear stiffness and damping of elastohydrodynamic point contacts." In: *Mechanism and Machine Theory* 124 (2018), pp. 179–196. ISSN: 0094-114X. DOI: https://doi.org/10.1016/j.mechmachtheory.2018.03.001.

[234] A. Boudiaf, A. Moussaoui, A. Dahane, and I. Atoui. "A Comparative Study of Various Methods of Bearing Faults Diagnosis Using the Case Western Reserve University Data." In: *Journal of Failure Analysis and Prevention* 16.2 (2016), pp. 271–284. ISSN: 1864-1245. DOI: 10.1007/s11668-016-0080-7.

[235] MATLAB. *version 9.5.0 (R2018b)*. The MathWorks Inc., 2018.

[236] B. Pan, Z. Shi, and X. Xu. "Analysis for the Weakly Pareto Optimum in Multiobjective-Based Hyperspectral Band Selection." In: *IEEE Transactions on Geoscience and Remote Sensing* 57.6 (2019), pp. 3729–3740. DOI: 10.1109/TGRS.2018.2886853.

[237] J.S. Arora. "Chapter 16 - Global Optimization Concepts and Methods." In: *Introduction to Optimum Design (Fourth Edition)*. Ed. by J.S. Arora. Boston: Academic Press, 2017, pp. 707–738. ISBN: 978-0-12-800806-5.

[238] Z.K. Peng, Peter W. Tse, and F.L. Chu. "A comparison study of improved Hilbert–Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing." In: *Mechanical Systems and Signal Processing* 19.5 (2005), pp. 974–988. ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2004.01.006.

[239] C.W. Champ and D.K. Shepherd. "Encyclopedia of statistics in quality and reliability." In: John Wiley & Sons, Ltd, 2007.

[240] J. Tuma. "Gearbox noise and vibration prediction and control." In: *International Journal of Acoustics and Vibrations* 14.2 (2009), pp. 99–108. DOI: 10.20855/ijav.2009.14.2242.

[241] A. El Saddik. "Digital Twins: The Convergence of Multimedia Technologies." In: *IEEE MultiMedia* 25 (Apr. 2018), pp. 87–92. DOI: 10.1109/MMUL.2018.023121167.

[242] F. Tao, H. Zhang, A. Liu, and A.Y.C. Nee. "Digital Twin in Industry: State-of-the-Art." In: *IEEE Transactions on Industrial Informatics* 15.4 (2019), pp. 2405–2415. DOI: 10.1109/TII.2018.2873186.

[243] E. Mucchi and G. Dalpiaz. "Numerical vibro-acoustic analysis of gear pumps for automotive applications." In: *International Conference on Noise and Vibration Engineering 2012, ISMA 2012, including USD 2012: International Conference on Uncertainty in Structure Dynamics*. Vol. 5. 2012, pp. 3951–3961. ISBN: 978-162276825-7.

[244] E. Mucchi, A. Rivola, and G. Dalpiaz. "Modelling dynamic behaviour and noise generation in gear pumps: Procedure and validation." In: *Applied Acoustics* 77 (Mar. 2014), 99–111. DOI: 10.1016/j.apacoust.2013.10.007.

[245] Y. Guo, T. Eritenel, T.M. Ericson, and R.G. Parker. "Vibro-acoustic propagation of gear dynamics in a gear-bearing-housing system." In: *Journal of Sound and Vibration* 333.22 (2014), pp. 5762 –5785. DOI: 10.1016/j.jsv.2014.05.055.

[246] A. Kahraman and R. Singh. "Non-linear dynamics of a spur gear pair." In: *Journal of Sound and Vibration* 142.1 (1990), pp. 49 –75. DOI: 10.1016/0022-460X(90)90582-K.

[247] R. Citarella, L. Federico, and A. Cicatiello. "Modal acoustic transfer vector approach in a FEM–BEM vibro-acoustic analysis." In: *Engineering Analysis with Boundary Elements* 31.3 (2007), pp. 248 –258. DOI: 10.1016/j.enganabound.2006.09.004.

[248] F. Hastings, J. Schneider, and S. Broschat. "Application of the perfectly matched layer (PML) absorbing boundary condition to elastic wave propagation." In: *The Journal of the Acoustical Society of America* 100.5 (1996), pp. 3061–3069. DOI: 10.1121/1.417118.