Mechanical Systems and Signal Processing 215 (2024) 111435

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

# Physics-based prognostics of rolling-element bearings: The equivalent damaged volume algorithm

Alberto Gabrielli *, Mattia Battarra, Emiliano Mucchi, Giorgio Dalpiaz

*University of Ferrara, Engineering Department, Via G. Saragat, 1 44122, Ferrara, Italy*

ARTICLE INFO

Communicated by X. Si

*Keywords:*
Rolling-element bearing
Bearing prognostics
Physics-based model

ABSTRACT

This paper introduces a novel parameter related to bearing degradation, namely the Equivalent Damaged Volume (EDV). An algorithm capable of extracting EDV values from experimental data is detailed. To this end, the proposed technique relies on the comparison between experimental and numerical signals. The former are the result of an extensive campaign of run-to-failure tests performed on a dedicated test rig. On the other hand, the latter are obtained by means of a lumped parameter model that accounts for the angular extent and the depth of either localized or extended defects. The potential of this novel technique is demonstrated by exploiting the values obtained by the algorithm as inputs for a prognostic model. In particular, the developed model aims at predicting the time until a certain threshold on the equivalent damaged volume is crossed, regardless of the applied load and the shaft rotation speed. The advantages and downsides introduced by the proposed approach are thoroughly discussed by providing examples based on real degradation histories.

## 1. Introduction

Rolling-element bearings (REBs) are one of the most frequently employed components in rotating machinery [1]. As such, their failure is a common reason for machine breakdown [2]. In fact, despite all the advantages introduced by their employment, these components are subjected to one important limitation, i.e. they will inevitably fail after a certain number of working cycles due to rolling contact fatigue. This aspect is crucial in the field of Prognostics and Health Management (PHM) of mechanical systems, as a damaged bearing induces vibrations that may possibly lead to the failure of the entire machinery. Consequently, throughout the years these components have been extensively studied by a number of researchers, in order to thoroughly understand their dynamic behaviour and therefore proposing dedicated schemes for their diagnosis and prognosis.

Notably, particular effort has been devoted to the development of dynamic models of faulty REBs. These models are aimed at determining the vibratory signature produced by defective bearings, i.e. at assessing the acceleration signal for a variety of defect shapes and operative conditions. P. McFadden and J. Smith [3,4] were amongst the first to work on the topic by proposing the impulse train model to study the effect of single and multi point defects. Their formulation was later enhanced by N. Tandon and A. Choudhury [5] to consider different pulse shapes and modified by D. Brie [6] to take into account the slight period variation between the impulses. Then, on the basis of these investigations, the majority of subsequent works tackled the problem by employing lumped parameter (LP) formulations. In these kind of models, the defect was usually modelled by describing the path travelled by the rolling elements inside the fault. One of the main contribution to this subject was presented by N. Sawalhi and R. Randall, which introduced a LP model of a self-aligning ball bearing in order to simulate its behaviour in presence of localized [7] and

* Corresponding author.
*E-mail address:* alberto.gabrielli@unife.it (A. Gabrielli).

https://doi.org/10.1016/j.ymssp.2024.111435
Received 20 November 2023; Received in revised form 10 March 2024; Accepted 14 April 2024

**Nomenclature**

| | |
|---|---|
| $\alpha$ | Bearing contact angle |
| $\alpha_c$ | Rayleigh damping coefficient which multiplies the mass term |
| $\beta$ | Coefficient of proportionality between damping and stiffness matrices |
| $\beta_c$ | Rayleigh damping coefficient which multiplies the stiffness term |
| $\delta\phi$ | Angular sampling |
| $\Delta\phi_d$ | Defect angular extent |
| $\delta$ | Relative contact displacement of the inner ring with respect to the outer ring |
| $\dot{\delta}$ | Relative contact velocity |
| $\ddot{\mathbf{x}}$ | Acceleration vector |
| $\dot{\mathbf{x}}$ | Velocity 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 |
| $\mathbf{M}$ | Mass matrix |
| $\mathbf{w}$ | External load vector |
| $\mathbf{x}$ | Displacement vector |
| $\omega_c$ | Cage angular speed |
| $\omega_s$ | Shaft speed |
| $\phi$ | Rolling element angular position |
| $\phi_0$ | Initial angular position of the cage |
| $\phi_d$ | Defect centre angular location |
| $\phi_{b,in}$ | Angular location in which a rolling element starts contact with the defect bottom surface |
| $\phi_{b,out}$ | Angular location in which a rolling element ends contact with the defect bottom surface |
| $\phi_{d,in}$ | Defect entry angular location |
| $\phi_{d,out}$ | Defect exit angular location |
| $c$ | Damping coefficient |
| $C_0$ | Basic static load rating |
| $c_i$ | Inner mass damping |
| $c_o$ | Outer mass damping |
| $c_o$ | Resonator mass damping |
| $D$ | Distance between $(\Delta\phi_d, h_d)$ pairs |
| $d_b$ | Ball trajectory inside a defect |
| $f_c$ | Contact force exchanged between rolling elements and races |
| $f_d$ | Contact damping force |
| $f_s$ | Low-pass frequency |
| $f_s$ | Sampling frequency |
| $h_d$ | Defect depth |
| $h_{dr}$ | Rough surface profile |
| $k_c$ | Contact stiffness |
| $k_i$ | Inner mass stiffness |
| $k_o$ | Outer mass stiffness |
| $k_o$ | Resonator mass stiffness |
| $k_{cl}$ | Linearized contact stiffness |
| $L_{10}$ | Bearing rating life |
| $m$ | Mass |
| $m_i$ | Inner mass |
| $m_o$ | Outer mass |

extended [8] defects. Their work was later expanded by D. Petersen et al. [9], who analysed the influence of defects with sharp edges on the system response. In addition, further studies were conducted by L. Cui et al. [10] to investigate the relationship between the time interval between impacts and the fault severity. A different formulation was proposed by A. Moazen Ahmadi et al. [11],

| | |
|---|---|
| $m_o$ | Resonator mass |
| $n$ | Load–deflection exponent |
| $n_b$ | Number of rolling elements on one row |
| $n_r$ | Number of rows |
| $R$ | Surface roughness scale factor |
| $R^2$ | Coefficient of determination |
| $r_b$ | Ball radius |
| $r_g$ | Ring groove radius |
| $r_i$ | Inner ring race radius |
| $r_o$ | Outer ring race radius |
| $r_p$ | Pitch radius |
| $RMS$ | Root Mean Square value |
| $t$ | Time |
| $t_c$ | Current time |
| $t_{EOL}$ | Time at bearing end of life |
| $t_{fs}$ | Time in which the final stage of bearing life starts |
| $t_{th}$ | Instant in time in which the equivalent damaged volume threshold is crossed |
| $V_{eq}$ | Equivalent damaged volume |
| $V_{th}$ | Threshold value for the equivalent damaged volume |
| $w$ | External load |
| $X$ | RMS value computed by the EDV algorithm |
| $x$ | Displacement in $x$ direction |
| $y$ | Displacement in $y$ direction |

**Acronyms**

| | |
|---|---|
| CBM | Condition Based Maintenance |
| DDM | Data-Driven Model |
| DOF | Degree of Freedom |
| EDV | Equivalent Damaged Volume |
| EOL | End of Life |
| HS | Health State |
| LP | Lumped Parameter |
| PBM | Physics-Based Model |
| PDD | Performance Degradation Dictionary |
| PE | Paris-Erdogan |
| PHM | Prognostics and Health Management |
| REB | Rolling-Element Bearing |
| RUL | Remaining Useful Life |

who proposed an alternative method to establish the path travelled by the rolling elements. In fact, they considered their finite size to determine their interaction with the fault edges and consequently avoid making any prior assumption on the rolling element trajectory. Furthermore, prominent researches in the last years include the works of Y. Jiang et al. [12], who detailed a thorough analytical formulation for 3D rectangular-like defects, and S. Gao et al. [13], which modelled the ball trajectory for asymmetric defects on angular contact ball bearings. Finally, the authors of this paper also contributed to the topic by proposing a technique for the assessment of the unknown model parameters by introducing a dedicated multi-objective optimization procedure [14].

As a matter of fact, dynamic models cover a crucial role in the development of diagnostic schemes and prognostic procedures [15]. 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 [16]. These two aspects are commonly integrated within the PHM of mechanical systems, which is a subject that has gained increasing attention from the industrial world in recent years [17]. 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 (HS) [18], allowing to sensibly reduce machine downtime and consequently decrease the costs associated with maintenance procedures compared to traditional approaches [19].

Concerning prognostic models, they may be broadly subdivided into three major categories [17]: physics-based models (PBMs), data-driven models (DDMs) and hybrid models. Among them, 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 DDMs 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 scenario. The majority of PBMs rely on the Paris-Erdogan (PE) equation [20] and its subsequent enhancements to infer the crack growth over time. In this regard, Y. Li et al. [21,22] proposed a model similar to PE to relate the number of cycles 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. Later, Y. Lei et al. [23] and J. Wang et al. [24] 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. On the other hand, the latter described a procedure to integrate multiple features and correlate them with the degradation history. Similarly, L. Liao et al. [25] proposed a combination of features in order to define a global one which was able to describe the deterioration process. J. Sun et al. [26] enhanced the PE model to a state-space model and proposed the optimal degradation state estimation through the employment of a Bayesian framework. M. Corbetta et al. [27] 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. [28] and E. Zio and G. Peloni [29]. More recently, Y. Lu et al. [30] proposed a PBM for REBs, which employed the realized volatility of the signal as an indicator for the system HS. Finally, L. Saidi et al. [31] combined 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 researchers proposed models based on different approaches. Qiu et al. [32] developed a stiffness-based prognostic model which related the natural frequencies and the vibration amplitude to bearing RUL. F. Sadeghi et al. [33] detailed a model for the evaluation of rolling contact fatigue in REBs. P.K. Gupta and E.V. Zaretsky [34] 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. M.J. Pais and N.H. Kim [35] employed finite-element analysis in combination with a modified PE law to predict the fatigue crack growth. T. Wang et al. [36] 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. B. Yan [37] introduced a two-stage physics-based Wiener process that took into account fatigue crack mechanism and its growth, component structure, assembly accuracy and conditions of the working environment.

Surprisingly, very few works dealt with the prognosis of bearing life by utilizing dynamic LP models. In this regard, an original methodology was recently proposed by L. Cui et al. [38], who employed a LP model of a defective bearing in order to establish a prognostic model based on the similarity theory [39]. To perform such a task, they compared experimental RMS trends with a numerically-generated Performance Degradation Dictionary (PDD). 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 is the authors' opinion that their original idea could be considered as a starting point to develop further PBMs based on a LP formulation. Possible improvements might regard the reduction of the computational burden introduced by the PDD approach, the possibility to take into account different defect propagation velocities rather than a constant one and to avoid assuming the defect dimensions between different stages.

In light of the aforementioned discussion, this paper provides an original contribution 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 into 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 instant in 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. The LP model is formulated in order to take into account the depth and the angular extent of the defect. Then, the potential of this technique is demonstrated by developing a prognostic PBM which takes as input the equivalent damaged volume computed from a reference bearing degradation history. The proposed PBM, in fact, aims at predicting the time until a certain threshold on the equivalent damaged volume is crossed, regardless of the applied load and the shaft rotation speed. As a result, this quantity is named Time-to-Threshold (TT). The results of the research are supported by experimental data acquired on a dedicated test bench for a number of run-to-failure tests.

The paper is outlined as follows. Section 2 describes the LP model and details a procedure for determining the equivalent damaged volume on the basis of experimental data. A PBM based on the values estimated through the EDV algorithm is developed. Then, Section 3 reports the results of the run-to-failure tests which were performed on the dedicated test rig. Thanks to these results, Section 4 details the estimation process of the equivalent damaged volume from the experimental signals and subsequently provides the outcome of the proposed PBM. Finally, Section 5 closes the paper by providing some summarizing remarks.

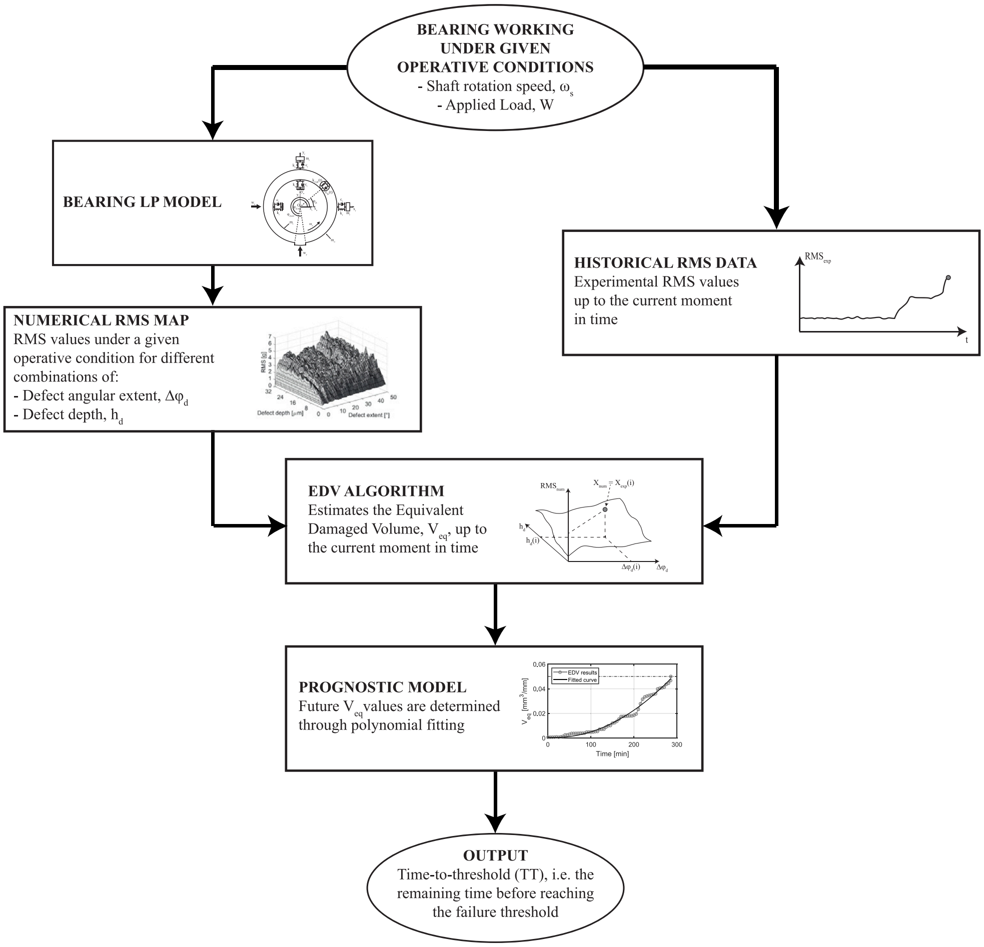


**Fig. 1.** Overview of the proposed prognostic framework.

## 2. Proposed numerical procedure

The LP model at the basis of the EDV algorithm has been described by the authors in a previous work [14]. However, it is further developed in this manuscript to also account for extended defects denoted by a rough surface. In fact, in the previous paper the LP model was described for the case in which the ball could not reach the bottom of the defect. This was done in order to validate the proposed multi-objective optimization procedure with experimental tests performed on rolling bearings in which defects were artificially seeded on their outer rings. Since the present effort tackles the problem of naturally induced defects, this assumption must be modified, as it is necessary to take into account the possible contact of the rolling elements with the bottom surface of the defect. This is done by taking advantage of the theory provided by N. Sawalhi and R. Randall on this subject [8]. In this work, their method is extended and further implemented to take into account a different rolling element trajectory inside the defect. As a consequence, the additional aspects introduced in this paper will be outlined after a brief summary of the model previously described in [14]. After introducing the LP model, the present Section details the EDV algorithm. The method is aimed at determining the equivalent damaged volume over time associated to any experimental run-to-failure test. This peculiar quantity is then employed as the foundation to propose a prognostic PBM.

To better represent the workflow discussed in the manuscript, Fig. 1 shows a schematic representation of the proposed prognostic framework. Consider a REB working under a given set of operative conditions. Acceleration data are continuously monitored on the bearing under exam, in order to collect historical RMS data. The bearing operative conditions are employed as input for a dedicated LP model, which permits to determine the RMS map related to the actual working conditions. Then, the EDV algorithm is employed to extract the equivalent damaged volume up to the current moment in time. This operation is carried out by interpolating the experimental RMS values with the numerical RMS map. Finally, the future degradation trend is estimated from the available equivalent damaged volumes through polynomial fitting. The output of the procedure is the Time-to-Threshold, i.e. the time needed before reaching a critical value of the equivalent damaged volume.

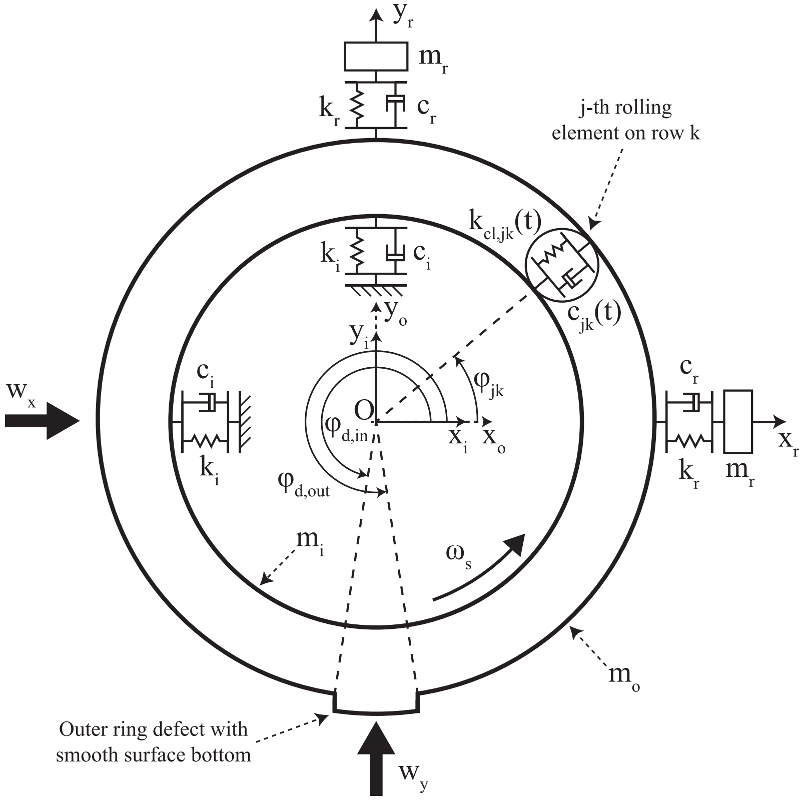


**Fig. 2.** Scheme of the LP model.

### 2.1. Lumped parameter model

The LP model, which is defined for a self-aligning ball bearing, is a planar model characterized by six degrees of freedom (DOF), as depicted in Fig. 2. In fact, the model is defined by three components, each one denoted by two orthogonal planar displacements $x$ and $y$. For each component, parameters $m$, $k$ and $c$ represent the mass, stiffness and damping, respectively. On the other hand, subscripts $i$, $o$ and $r$ refer to quantities associated with inner ring, outer ring and high-frequency resonant mass, respectively.

For a bearing with $n_b$ rolling elements on each row, the position of the $j$th ball on the $k$th row is defined according to [7]:

$$\phi_{jk}(t) = \phi_0 + \frac{2\pi(j-1)}{n_b} + \frac{\pi(k-1)}{n_b} + \omega_c t \tag{1}$$

in which $\phi_0$ is the cage initial angular position and $\omega_c$ is the cage angular speed:

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

where $\omega_s$ is the shaft rotation speed, $r_b$ is the ball radius, $r_p$ is the bearing pitch radius and $\alpha$ is the contact angle.

The contact deformation of a ball is defined on the basis of the relative displacement of the rings [9]:

$$\delta_{jk}(t) = (x_i - x_o)\cos\phi_{jk}(t) + (y_i - y_o)\sin\phi_{jk}(t) - h - d_b(\phi_{jk}(t)) \tag{3}$$

in which $h$ is the radial clearance and $d_b$ is the additional displacement inside the defect area, whose definition will be later discussed. Thanks to Eq. (3) it is possible to evaluate the contact force exchanged between rolling elements and races [12]:

$$f_{c,jk}(t) = \begin{cases} k_c \delta^n_{jk}(t) & \text{if } \delta_{jk}(t) \geq 0 \\ 0 & \text{if } \delta_{jk}(t) < 0 \end{cases} \tag{4}$$

where $k_c$ is the contact stiffness, which may be estimated through analytical or finite-element methods [40], and exponent $n$ is equal to 1.5 for a ball-raceway contact [1]. Furthermore, according to [9], it is possible to determine the contact damping force as:

$$f_{d,jk}(t) = \begin{cases} c_{jk} \dot{\delta}_{jk}(t) & \text{if } \delta_{jk}(t) \geq 0 \\ 0 & \text{if } \delta_{jk}(t) < 0 \end{cases} \tag{5}$$

where $c_{jk}$ is the damping coefficient, defined by means of Rayleigh damping [41] as

$$c_{jk}(t) = \alpha_c m_b + \beta_c k_{cl,jk}(t) \tag{6}$$

where $m_b$ is the ball mass, $k_{cl,jk} = k_c \delta_{jk}^{n-1}(t)$ is the linearized contact stiffness and parameters $\alpha_c$ and $\beta_c$ are the coefficients which define the Rayleigh damping.

As a result, by grouping all DOFs in vector $\mathbf{x}(t)$:

$$\mathbf{x}(t) = \left\{ x_i(t) \quad y_i(t) \quad x_o(t) \quad y_o(t) \quad x_r(t) \quad y_r(t) \right\}^T \tag{7}$$

the system of equations of motion may be written as:

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) + \mathbf{f_c}(t) + \mathbf{f_d}(t) = \mathbf{w} \tag{8}$$

The terms in Eq. (8) are defined as follows. Matrices **M**, **C** and **K** are the mass, damping and stiffness matrix, respectively. Moreover, for a bearing with $n_r$ rows, $\mathbf{f_c}(t)$ is the contact force vector, i.e.:

$$\mathbf{f_c}(t) = \sum_{k=1}^{n_r} \sum_{j=1}^{n_b} f_{c,jk}(t) \left\{ \cos\phi_{jk}(t) \quad \sin\phi_{jk}(t) \quad -\cos\phi_{jk}(t) \quad -\sin\phi_{jk}(t) \quad 0 \quad 0 \right\}^T \tag{9}$$

while $\mathbf{f_d}(t)$ is the damping force vector:

$$\mathbf{f_d}(t) = \sum_{k=1}^{n_r} \sum_{j=1}^{n_b} f_{d,jk}(t) \left\{ \cos\phi_{jk}(t) \quad \sin\phi_{jk}(t) \quad -\cos\phi_{jk}(t) \quad -\sin\phi_{jk}(t) \quad 0 \quad 0 \right\}^T \tag{10}$$

Finally, **w** is the external forces vector:

$$\mathbf{w} = \left\{ 0 \quad 0 \quad w_x \quad w_y \quad 0 \quad 0 \right\}^T \tag{11}$$

where $w_x$ and $w_y$ are the values of external loads in the $x$ and $y$ directions, respectively.

The description of the model is completed by providing the equation which characterizes the ball trajectory inside the outer ring defect, i.e. $d_b(\phi_{jk}(t))$. The formulation is adapted from [12] in order to consider a self-aligning ball bearing. Firstly, the path travelled without the rough surface is described. The extended defect is schematically depicted in Fig. 3. In the plane O'x'y', which is obtained from plane Oxyz by rotating its $x$-axis by an angle equal to $\alpha$, the bottom surface follows a circular path with an offset from the undamaged bearing surface equal to $h_d$, i.e. the depth of the extended defect. The defect is centred at the angular position $\phi_d$ and its angular extent is defined by parameter $\Delta\phi_d$. The ball trajectory $d_b$ on plane Oxyz may be consequently defined as follows:

$$d_b(\phi_{jk}(t)) = \begin{cases} \cos\alpha \left[ r_b + r_o \cos(\phi_{jk}(t) - \phi_{d,in}) - r_o - \sqrt{r_b^2 - (r_o \sin(\phi_{jk}(t) - \phi_{d,in}))^2} \right] & \text{if } \phi_{d,in} \le \phi_{jk}(t) < \phi_{b,in} \\ h_d \cos\alpha & \text{if } \phi_{b,in} \le \phi_{jk}(t) \le \phi_{b,out} \\ \cos\alpha \left[ r_b + r_o \cos(\phi_{d,out} - \phi_{jk}(t)) - r_o - \sqrt{r_b^2 - (r_o \sin(\phi_{d,out} - \phi_{jk}(t)))^2} \right] & \text{if } \phi_{b,out} < \phi_{jk}(t) \le \phi_{d,out} \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

where $r_o$ is outer race radius, $\phi_{d,in}$ and $\phi_{d,out}$ are the entry and exit edges of the defect, viz.:

$$\phi_{d,in} = \phi_d - \Delta\phi_d/2 \tag{13}$$

$$\phi_{d,out} = \phi_d + \Delta\phi_d/2 \tag{14}$$

and $\phi_{b,in}$ and $\phi_{b,out}$ 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 $\phi_{b,in}$ and $\phi_{b,out}$, i.e.:

$$r_b + r_o \cos(\phi_{b,in} - \phi_{d,in}) - r_o - \sqrt{r_b^2 - (r_o \sin(\phi_{b,in} - \phi_{d,in}))^2} - h_d = 0 \tag{15}$$

$$r_b + r_o \cos(\phi_{d,out} - \phi_{b,out}) - r_o - \sqrt{r_b^2 - (r_o \sin(\phi_{d,out} - \phi_{b,out}))^2} - h_d = 0 \tag{16}$$

The proposed rolling element path $d_b(\phi_{jk}(t))$ describes the bottom of the defect as a smooth surface with constant depth. However, experimental evidence [42] shows that the surface is characterized by a significant roughness. Thus, the equation describing $d_b$ between $\phi_{b,in}$ and $\phi_{b,out}$ is further modified to also take into account this aspect. Surface roughness is introduced by applying the procedure described in [8], but only between angles $\phi_{b,in}$ and $\phi_{b,out}$. In 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 $R$. The random profile is defined between the angular locations $\phi_{b,in}$ and $\phi_{b,out}$ and it is characterized by a spatial sampling $\delta\phi$ which is equal to:

$$\delta\phi = \frac{\omega_c}{f_s} \tag{17}$$

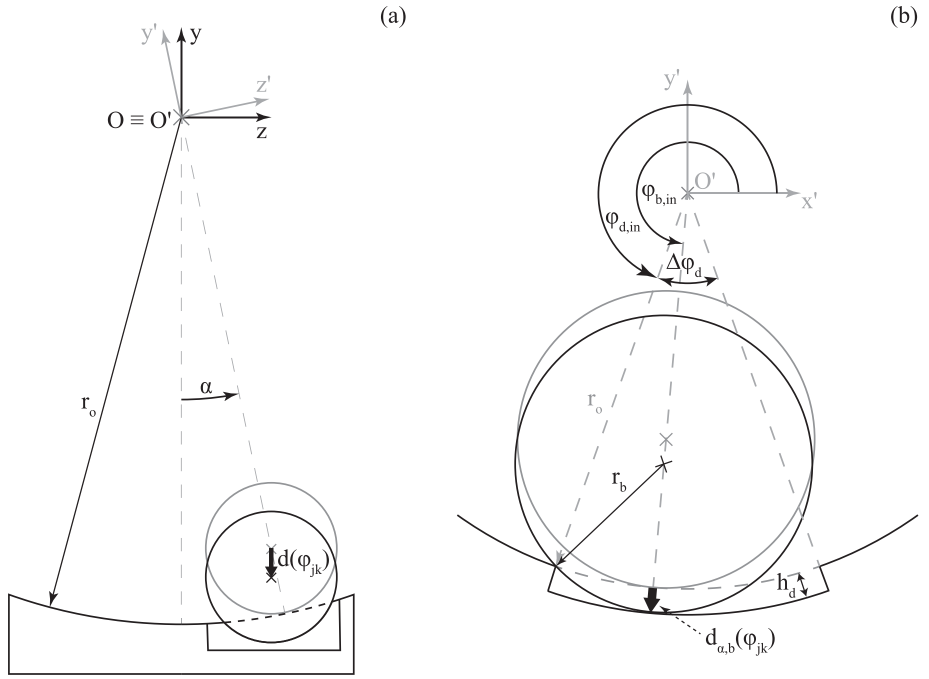


**Fig. 3.** Geometric definition of angular position $\phi_{b,in}$. The same scheme may be replicated for $\phi_{b,out}$ at the exit of the defect.

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 [8], this cut-off frequency $f_l$ is equal to:

$$f_l = \frac{1}{\pi\sqrt{4Rr_b}} \tag{18}$$

Finally, the outcome of this procedure is the rough surface profile $h_{dr}(\phi_{jk}(t))$, which modifies Eq. (12) so that:

$$d_b(\phi_{jk}(t)) = \left[h_d + h_{dr}(\phi_{jk}(t))\right]\cos\alpha \qquad \text{if } \phi_{b,in} \le \phi_{jk}(t) \le \phi_{b,out} \tag{19}$$

Fig. 4 reports examples of the ball path inside the defect in a bearing with outer ring radius $r_o$ = 22.91 mm, ball radius $r_b$ = 3.56 mm and contact angle $\alpha$ = 10.2 deg. Fig. 4.a shows the trajectory followed by the ball in a localized defect characterized by an angular extent $\Delta\phi_d$ = 2 deg and a depth sufficient to avoid contact with its bottom. On the other hand, Fig. 4.b depicts the ball path inside an extended defect with angular span $\Delta\phi_d$ = 20 deg, depth $h_d$ = 20 μm and roughness $R$ = 10 μm. In Fig. 4.b, the dash-dot line represents the path with the smooth bottom, while the solid line describes the path due to a rough surface. Fig. 4.b also permits to assess that the effective path has a deviation around $h_d$ lower than the imposed $R$ value. In fact, the low-passing allows to fit the ball to the path it can practically reach. The efficiency of this method was successfully demonstrated in later works by D. Petersen et al. [9,43] to characterize extended defects and discuss their influence on bearing vibrations.

### *2.2. Equivalent damaged volume algorithm*

Taking the aforementioned LP model as a starting point, this section introduces a novel procedure for the estimation of a parameter related to the evolution of bearing life. This quantity, namely the equivalent damaged volume per unit of width, is aimed at tracking the bearing health deterioration over time. 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 $\Delta\phi_d$ 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 algorithm determines an equivalent volume $V_{eq}(i)$ per unit width equal to:

$$V_{eq}(i) = r_o\Delta\phi_d(i)h_d(i) \tag{20}$$

where $\Delta\phi_d(i)$ and $h_d(i)$ are the defect dimensions at time $t(i)$. If $r_o$ and $h_d$ are expressed in mm and $\Delta\phi_d(i)$ is expressed in radians, then $V_{eq}(i)$ assumes the unit of mm$^3$/mm. In order 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

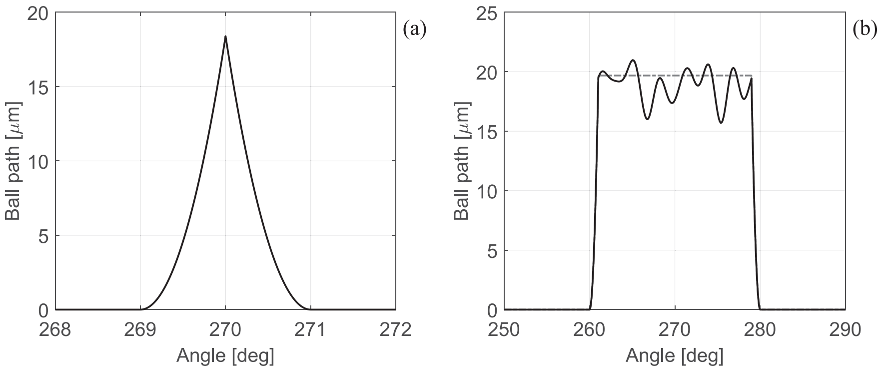


**Fig. 4.** Examples of trajectories followed by the ball inside a defect: (a) localized defect; (b) extended defect. The solid and dash-dot lines represent the paths on the rough and smooth surfaces, respectively.

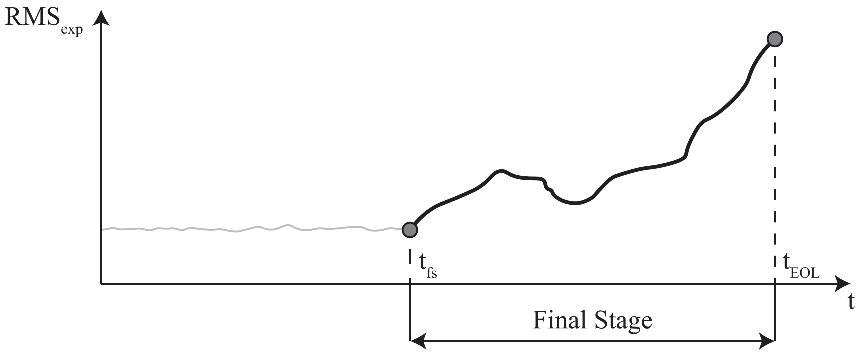


**Fig. 5.** Final stage of bearing life.

compared to its other dimensions [12]. 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 provide increasing values of $\Delta\phi_d$ and $h_d$ during REB degradation, which is a realistic assumption considering that the defect should increase its dimensions as the deterioration progresses. Finally, the choice to use RMS values in spite of other statistical indicators is justified by the behaviour of the indicator itself, that can be consistently correlated to the fault evolution both in terms of experiments and numerical predictions. According to the Parseval Theorem, the RMS of a signal contains information regarding the energy content of the signal. Based on this principle, when a defect appears within a bearing, the energy content of the radial acceleration signal is expected to increase. This is due to the fact that the defect inside the bearing acts as an internal excitation for the system. Therefore, when a defect appears, we may think of it as a new and additional excitation affecting the bearing. Since an additional excitation appears, the energy content of the radial vibration of the bearing is expected to increase. Based on this consideration, the RMS is assumed to be a potential indicator of the presence of defects. Thus, as the defect increases, the internal excitation increases and therefore the energy content of the radial vibration of the bearing is expected to increase. This is the reason why the RMS can be considered as an indicator of the severity of the defect, which is the fundamental property for indicators that need to follow the evolution of the defect.

The proposed algorithm considers only the final part of the bearing life. 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 [2]. 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. Fig. 5 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 2.1. Consequently, RMS values are calculated by defining two vectors, i.e. $\mathbf{h_d}$ and $\boldsymbol{\Delta\phi_d}$, which both range from a minimum to a maximum value, so that $\mathbf{h_d} = [h_{d,min}; h_{d,max}]$ and $\boldsymbol{\Delta\phi_d} = [\Delta\phi_{d,min}; \Delta\phi_{d,max}]$. These limiting values may be assigned on the basis of experimental observations. For instance, the maximum angular extent $\Delta\phi_{d,max}$ may be set according to typical values of defect size observed at the end of bearing life. Then, for each pair of values $(h_d, \Delta\phi_d)$, 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 a previous paper by the authors [14]. Although the method has been validated for localized 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. $\phi_d$, is assumed to be aligned with the direction of the radial load. This is a reasonable assumption that allows to reduce the number of unknowns in the problem. Finally, it is worth underlining that the RMS map may be generated for any kind of bearing defect, i.e. inner ring, outer ring, rolling element or cage.

The main steps of the EDV algorithm are summarized in the flowchart reported in Fig. 6. 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, the algorithm 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 $\Delta\phi_d$ at each time step. In between, the procedure is subdivided as follows:

1. The index $i$ is initialized 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 Fig. 7.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 $min(X_{num}) \leq X_{exp}(i) \leq max(X_{num})$. 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 $\Delta\phi_d$ and $h_d$ for each one, as depicted in Fig. 7.b. The corresponding volume is also calculated through Eq. (20).
5. For each $X_{exp}(i)$ value, several pairs ($\Delta\phi_d$, $h_d$) 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 ($\Delta\phi_d(i)$, $h_d(i)$) which provides the lowest possible volume. This condition is represented in Fig. 7.c and it may be written as:

$$V_{eq}(i) = arg\ min(X_{num} : X_{num} = X_{exp}(i)) \tag{21}$$

   The assumption ensures the minimum relative variation with respect to the equivalent volume $V_{eq}(i-1)$ at instant $t(i-1)$.
6. In case there is more than one pair that minimizes the volume, the algorithm stores the pair ($\Delta\phi_d$, $h_d$) closest to the pair evaluated at time $t(i-1)$, i.e. ($\Delta\phi_d(i-1)$, $h_d(i-1)$). In other words, the chosen pair is the one that minimizes the distance $D(\Delta\phi_d, h_d)$ in the $\Delta\phi_d$ - $h_d$ plane:

$$D = \sqrt{\left(\Delta\phi_d - \Delta\phi_d(i-1)\right)^2 + \left(h_d - h_d(i-1)\right)^2} \tag{22}$$

   This condition may be evaluated only if $i > 1$. When $i = 1$, in fact, there are no prior points available. Therefore, in this case only, the stored pair is the one closest to the origin:

$$D = \sqrt{\Delta\phi_d^2 + h_d^2} \tag{23}$$

   Finally, if there are multiple pairs that also minimize the distance, the pair with the lowest $\Delta\phi_d$ is selected.
7. If $t(i) = t_{EOL}$, the algorithm is stopped. Instead, if $t(i) < t_{EOL}$, the procedure continues.
8. To ensure the increment of damage progression over time, numerical RMS values that are associated with values of $\Delta\phi_d$ and $h_d$ lower than $\Delta\phi_d(i)$ and $h_d(i)$ are removed from the numerical RMS map. As a consequence, in the subsequent time step the available pairs ($\Delta\phi_d$, $h_d$) are only the ones in the right-most part of the plot, as shown in Fig. 7.d. This assumption also guarantees that $V_{eq}(i+1) \geq V_{eq}(i)$, i.e. the damaged volume increases over time. However, there might be cases where at least one suitable pair ($\Delta\phi_d$, $h_d$) 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. To avoid premature stop, in this particular case the search range is extended back to previously removed values until one value that satisfies the condition $X_{num} = X_{exp}(i)$ is found. This is done by extending the search domain by one unit in both coordinates $\Delta\phi_d$ and $h_d$ until the first pair of damage dimensions that ensures the condition 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. $V_{eq}(i) < V_{eq}(i-1)$. However, 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. The goodness of this hypothesis is examined in Section 4.
9. 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, Fig. 7.e and 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 (8).

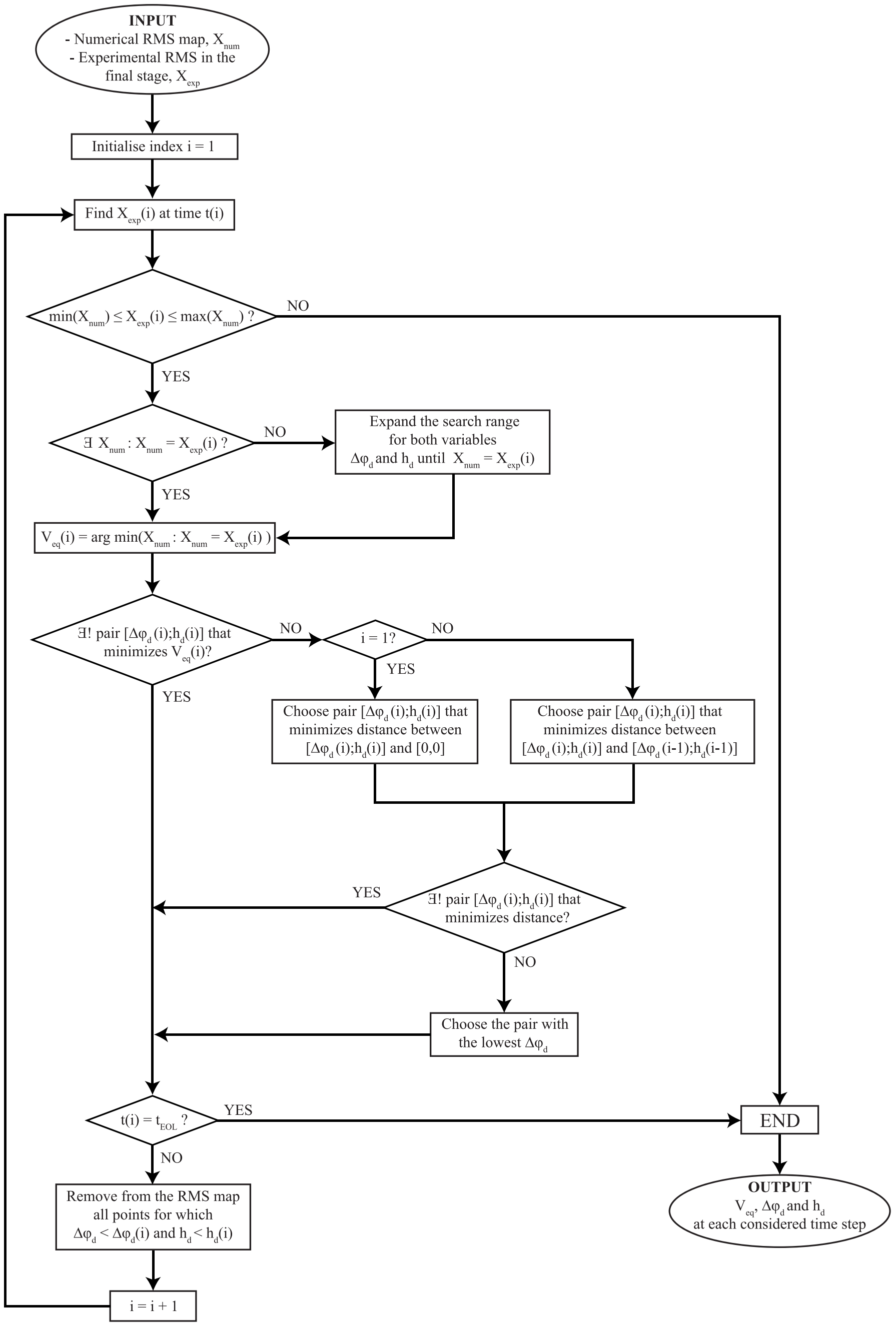


**Fig. 6.** Flowchart of the EDV algorithm.

### 2.3. Development of a prognostic PBM based on the EDV algorithm

The equivalent damaged volume values obtained through the EDV algorithm may be employed for the development of prognostic PBMs. To this end, this section details a novel PBM aimed at predicting 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:

$$TT = t_{th} - t_c \tag{24}$$

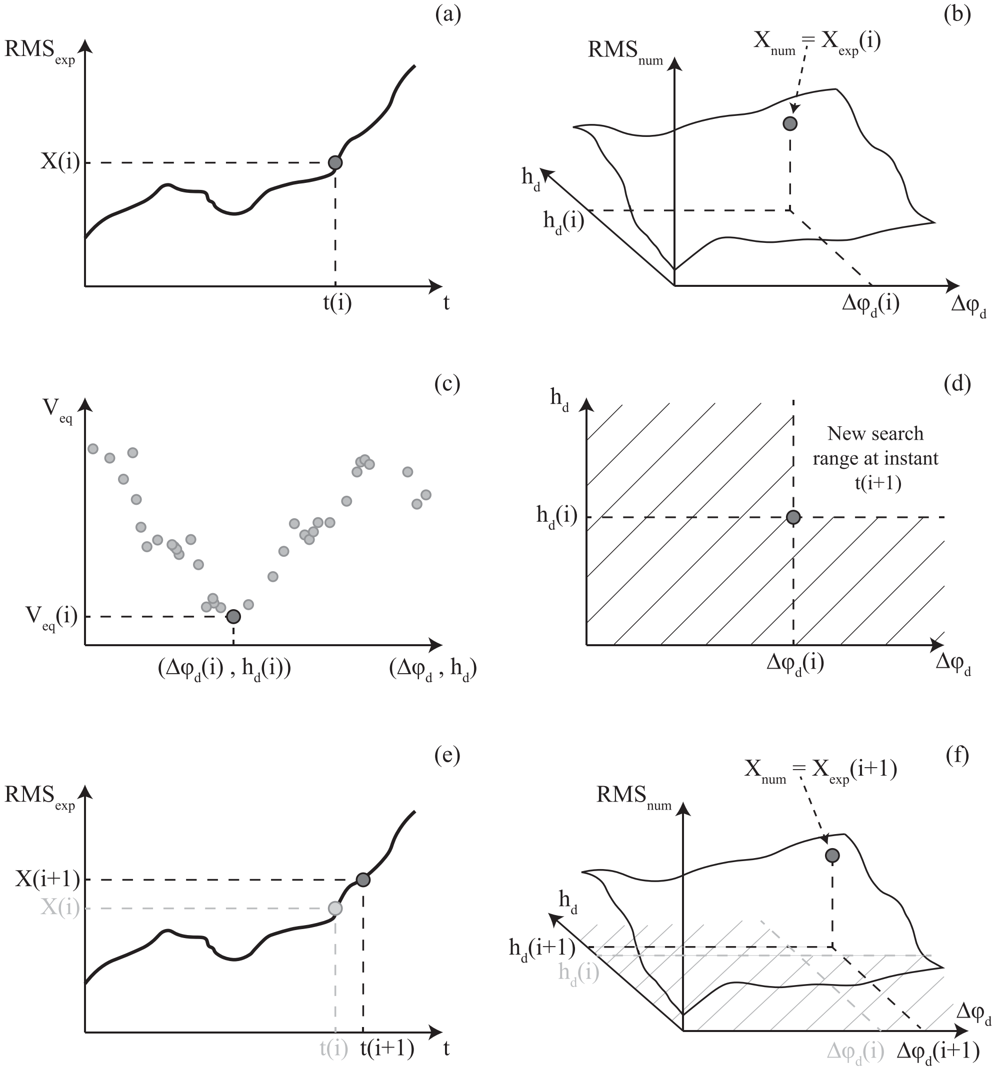


**Fig. 7.** Steps of the EDV algorithm: (a) Consider the experimental RMS value $X(i)$ at time $t(i)$; (b) Find all points in the RMS map that satisfy the condition $X_{num} = X_{exp}(i)$; (c) Determine the pair ($\Delta\phi_d$, $h_d$) that minimizes the equivalent damaged volume $V_{eq}$; (d) Remove 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 Fig. 6 is achieved; (f) RMS map at time $t(i+1)$, in which the area removed in Figure (d) is highlighted by a grey hatching.

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. 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:

$$V_{eq} = V_{eq}(t) \quad \text{if } t_{fs} \leq t \leq t_{th} \tag{25}$$

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 polynomials of order two to determine the function that better fits the data. Furthermore, it is imposed that the polynomial must pass through the first equivalent damaged volume estimated at time $t_{fs}$ to ensure that the best-fitting curve passes through the first point of the degradation history. If this relationship provides satisfactory results for different cases, then it may applied to other working conditions, under the hypothesis that the damage propagation follows the same relationship under every operative condition. Therefore, $t_{th}$ may be estimated during operation by fitting the available data up to the current time with a polynomial of the adequate order as observed from experiments under other conditions. It should be noted, however, that other type of curves could be employed if the polynomials do not provide an acceptable description of the damage evolution in the case under exam.

The proposed PBM is formulated in the following manner. 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.

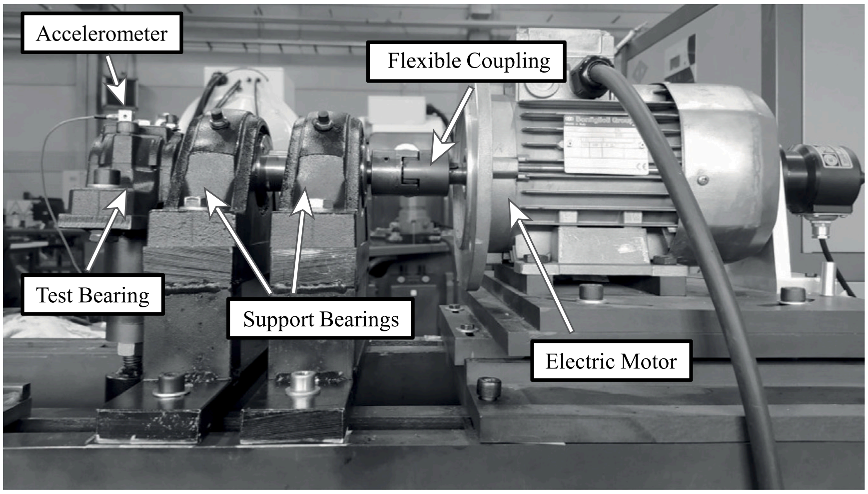


**Fig. 8.** Bearing test bench at the Engineering Department of the University of Ferrara.

**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 | $\alpha$ | 10.2 deg |
| Number of balls on each row | $n_b$ | 12 |

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 utilized 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. (24).

It is worth mentioning that the initial moment in which the bearing health state transitions from the healthy to the final stage must be detected by means of additional tools. For this purpose, diagnostic indicators may be employed [44] to ensure an accurate evaluation of $t_{fs}$.

## 3. Experimental tests

This Section describes the run-to failure tests performed at the University of Ferrara. Firstly, Section 3.1 provides a thorough description of the test rig. Then, Section 3.2 outlines the results of the experiments. The need to perform dedicated experimental tests arises from the consideration that, unfortunately, existing run-to-failure datasets [42,45–47] do not provide sufficient information for the application of the EDV algorithm. Such information includes mass, stiffness and damping values, i.e. the input parameters of the LP model described in Section 2.1. To overcome this issue, the authors decided to perform a series of run-to-failure experiments on a dedicated test bench. In fact, thanks to the possibility to physically access the test rig, it is possible to estimate the unknown model parameters by employing a dedicated procedure developed by the authors in a previous work [14]. Knowledge of these key information allows for the application of the EDV algorithm on data acquired on the test bench.

### 3.1. Test bench description

The test rig is displayed in Fig. 8. Besides, the bearing under test is shown in Fig. 9. It is a self-aligning ball bearing, model *1205 ETN9*, whose dimensions are reported in Table 1. The REB is mounted on a shaft, which is connected through a flexible coupling to an electric motor. An inverter controls the rotation frequency of the shaft, which is supported by two spherical roller bearings connected to the frame. Furthermore, the REB is enclosed in a casing, on which a radial load is applied by means of a lever system.

During the tests, the acceleration signal was acquired through a piezoelectric accelerometer mounted on top of the casing. Data were continuously collected 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 s of signal were acquired every 5 min. As a consequence, 12 signals were stored for each hour

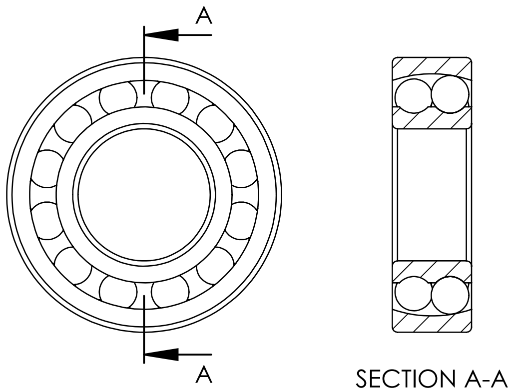


**Fig. 9.** Geometry of the self-aligning ball bearing *1205 ETN9*, employed as test bearing.

**Table 2**
Operative conditions, expected life and test length for the six tested bearings.

| ID | Load [kN] | Shaft speed [Hz] | $L_{10}$ [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 |

of test. Two accelerometers were also mounted on the support bearings in order to monitor their health state. In fact, although they were characterized by a way higher $C_0$ compared to the test bearing, i.e. 86.5 kN, it was crucial to control their behaviour during these long tests to avoid unexpected failures. Therefore, the signals of these two accelerometers were also acquired for monitoring purposes.

### *3.2. Results of the run-to-failure tests*

The test rig is employed to perform a number of run-to-failure tests, i.e. experiments that involve the acquisition of the vibratory signal through the entire operative life of the bearing. The main issue related to run-to-failure tests is the extensive time duration of this kind of tests. In fact, 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. To overcome this issue, it is commonly preferred to perform accelerated tests, which involve the application of a large load to induce a fast degradation of the component. This is a standard kind of testing procedure employed by researchers to generate faulty bearing datasets [45,47,48]. Furthermore, the applied loads are usually greater than the basic static load rating $C_0$ of the bearing, which is equal to 3.3 kN for the bearing used for the tests. As detailed in [1], the load applied on the bearing may exceed $C_0$, provided that it acts continuously during bearing operation. This, in fact, permits to generate permanent deformations which are evenly distributed among races and rolling elements. This condition is known to produce sufficiently smooth operation [1]. Therefore, for stationary tests in which load and speed values do not change, it is possible to employ loads greater than the $C_0$ in order to fasten bearing degradation.

In light of these considerations, six run-to-failure tests were performed, as reported in Table 2. All tests were 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 2 also reports the expected life and the total test time. The expected test duration is evaluated by employing the well-known $L_{10}$ formula [1]. Concerning the test end, runs 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 [45,47,48]. At the end of the tests, all bearings were characterized by an extended defect on the outer raceway. For instance, the defect which developed during test E1 may be observed in Fig. 10. The picture highlights the defect characteristics at the end of the test, i.e. a large angular extent and a non-neglectable bottom surface roughness. As shown in Section 2.1, these two aspects are fully considered in the proposed LP model. Further inspection of the bearings also showed that, in each test, the defect was centred on the plane of maximum load. Moreover, the other components were all found to be healthy at the end of the test, i.e. it may be assumed that they were all properly rotating for the entire duration of the test. In conclusion, experimental evidence justifies the decision to focus only on the outer ring defect for the LP model described in Section 2.1.

The resulting RMS histories of the six run-to-failure tests are shown in Fig. 11. In particular, .a, .b and .c represent the tests carried out under a radial load value equal to 4 kN, namely tests E1, E2 and E3, respectively. Furthermore, Fig. 11.d, .e and .f

**Fig. 10.** Extended defect on bearing outer race at the end of test E1.

**Table 3**
$t_{fs}$, $t_{EOL}$ 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 |

refer to the tests E4, E5 and E6 which are characterized by applied loads equal to 3 kN, 4.7 kN and 5 kN, respectively. The total time $t_{EOL}$ of each test corresponds to the final point of each RMS history. 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 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 Fig. 11.b, .d and .e, respectively. Besides, Fig. 11.a highlights that, during test E1, the RMS trend briefly decreases and then starts to increase again. This may be observed even more clearly in Fig. 12.a, which depicts only the final stage of each test. Moreover, tests E3 and E6 show an even different behaviour. Test E3, which is depicted in Figs. 11.c and 12.a, 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, Figs. 11.f and 12.b show 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 to the conclusion of 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 Fig. 11.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, i.e.:

$$H_{or} = \sum_{k=1}^{5} A_k \tag{26}$$

where $A_k$ is the $k$th harmonic of the defect characteristic frequency in the envelope spectrum. The described indicator is introduced in order to demonstrate its capabilities in detecting the defects which typically appear in the experimental test rig. To this end, the experimental values of $H_{or}$ are reported in Fig. 13. 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.

To conclude, Table 3 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:

$$RUL(t_{fs}) = t_{EOL} - t_{fs} \tag{27}$$

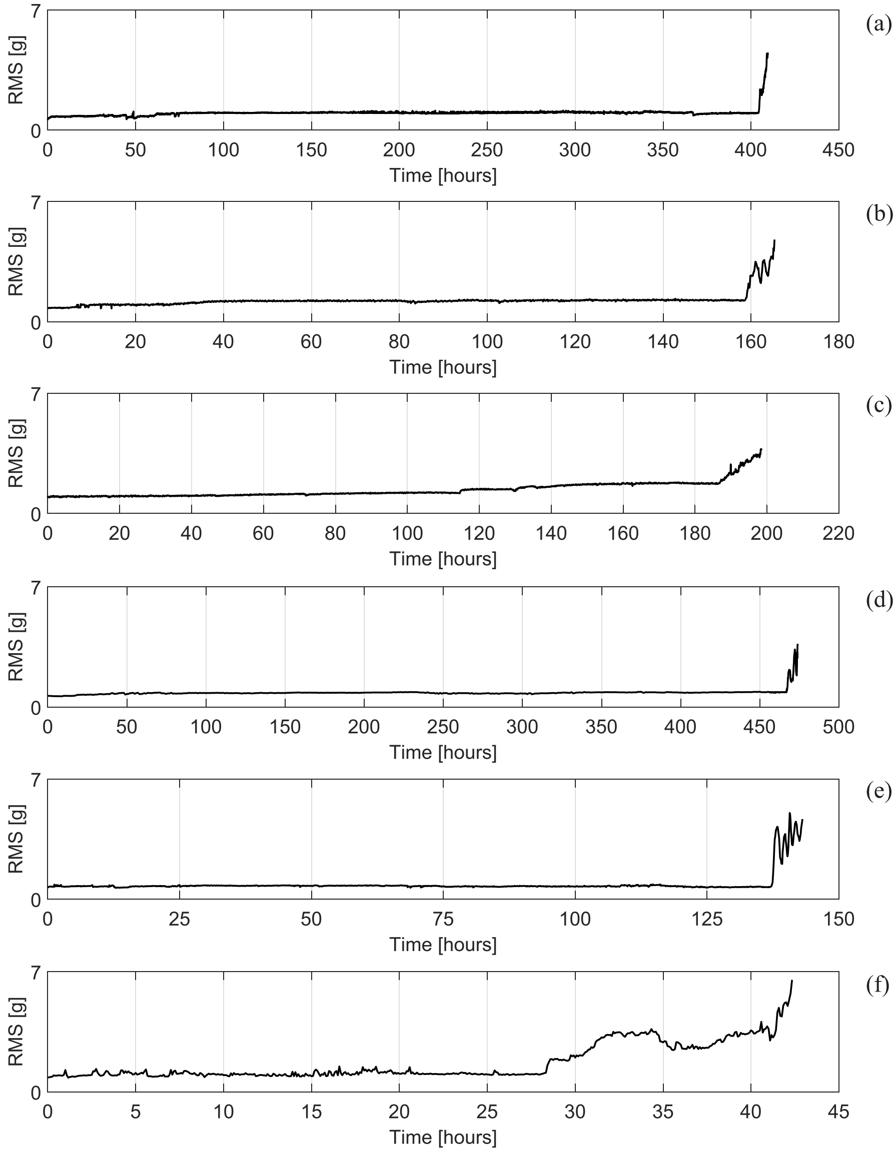


**Fig. 11.** Experimental RMS trends for the entire test length: (a) E1 (4 kN); (b) E2 (4 kN); (c) E3 (4 kN); (d) E4 (3 kN); (e) E5 (4.7 kN); (f) E6 (5 kN).

## 4. REB prognostics through the EDV algorithm

This section describes the application of the methods described in Section 2 through the employment of the experimental data reported in Section 3. Firstly, $V_{eq}$ values are determined by exploiting the EDV algorithm. Then, this parameter is utilized as input for a prognostic PBM.

### 4.1. EDV assessment from test bench data

The EDV algorithm, which has been described in Section 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 Fig. 12.a and .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 2. 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 in Section 2.1. It is worth noticing that the lumped parameter model requires a number of parameters referring to the bearing under exam that make mandatory to know its geometry and material. Table 4 reports the list of required parameter together with the numerical values adopted in the present application. As it may be appreciated, parameters

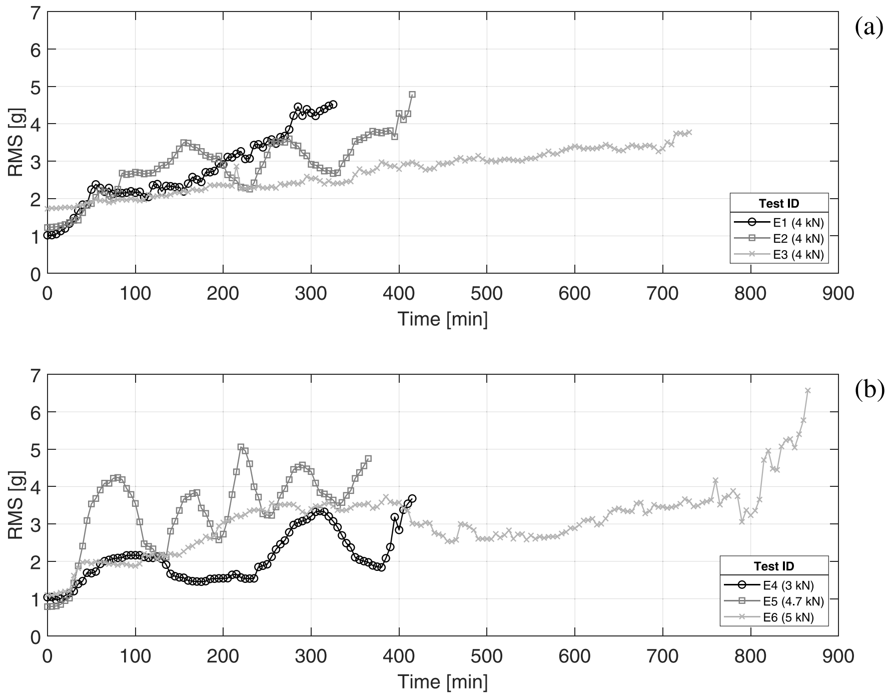


**Fig. 12.** RMS in the final stage: (a) Tests E1, E2 and E3; (b) Tests E4, E5 and E6.

$m_i$, $m_o$ and $m_b$ require to have information regarding the geometry of both the bearing and the machine containing it. Shaft radial stiffness $k_i$ and contact stiffness $k_c$ may be deduced from analytical methods or FE analyses. Within the framework of the present application, $k_c$ has been computed analytically by exploiting the Hertzian contact theory for the general case of two bodies with different principal relative radii of curvature in orthogonal planes. Shaft radial stiffness $k_i$ has been obtained with FE simulations. The remaining terms refer to the properties of the resonator and the damping of the components. Regarding the resonator, according to N. Sawalhi and R. Randall [7,8] who proposed its use for the very first time, it is proposed to select an associated mass and stiffness values in order to reproduce a real natural frequency of the system. However, no further guidelines on the choice of such quantities are available. Regarding the damping of the system, it is well known that lubricated mechanical components may have damping values ranging from the 2% to the 20% and its experimental determination is often very hard. Based on these premises, for the present application the authors deduced the values reported in Table 4 by taking advantage of a multi-objective optimization technique where the unknown parameters, i.e. the damping coefficients and the properties of the resonant mass, are constrained within realistic boundaries. An extended explanation of this approach may be found in Ref. [14]. As a matter of fact, being able to accurately estimate model parameters is a crucial aspect, since the bearing LP model is the foundation of the EDV algorithm. The necessity to fill the lumped parameters dynamic model with the detailed data is the reason why the authors did not have the chance to validate their method against other public datasets, as the ones reported in [41,45,47,48], since no detailed info regarding dimensions and materials of the systems under exam are provided. For this purpose, the employed procedure aims at minimizing the difference between experimental and numerical features across a set of different loads, shaft speeds and defect dimensions. The full explanation of the method, along with its experimental validation, is reported in [14]. Within the context of this work, the resulting model parameters are reported in Table 4. For a given applied load, RMS values are evaluated for several combinations of $\Delta\phi_d$ and $h_d$. The angular extent is limited from 0 to 50 degrees, which 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 [8]. The four resulting RMS maps are shown in Fig. 14. In particular, Fig. 14.a, .b, .c and .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 Fig. 14.a and .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 $\Delta\phi_d$ and $h_d$. Notably, the defect depth influences the RMS value only if the defect is large enough to permit the interaction between balls and surface bottom. This means that, if the angular

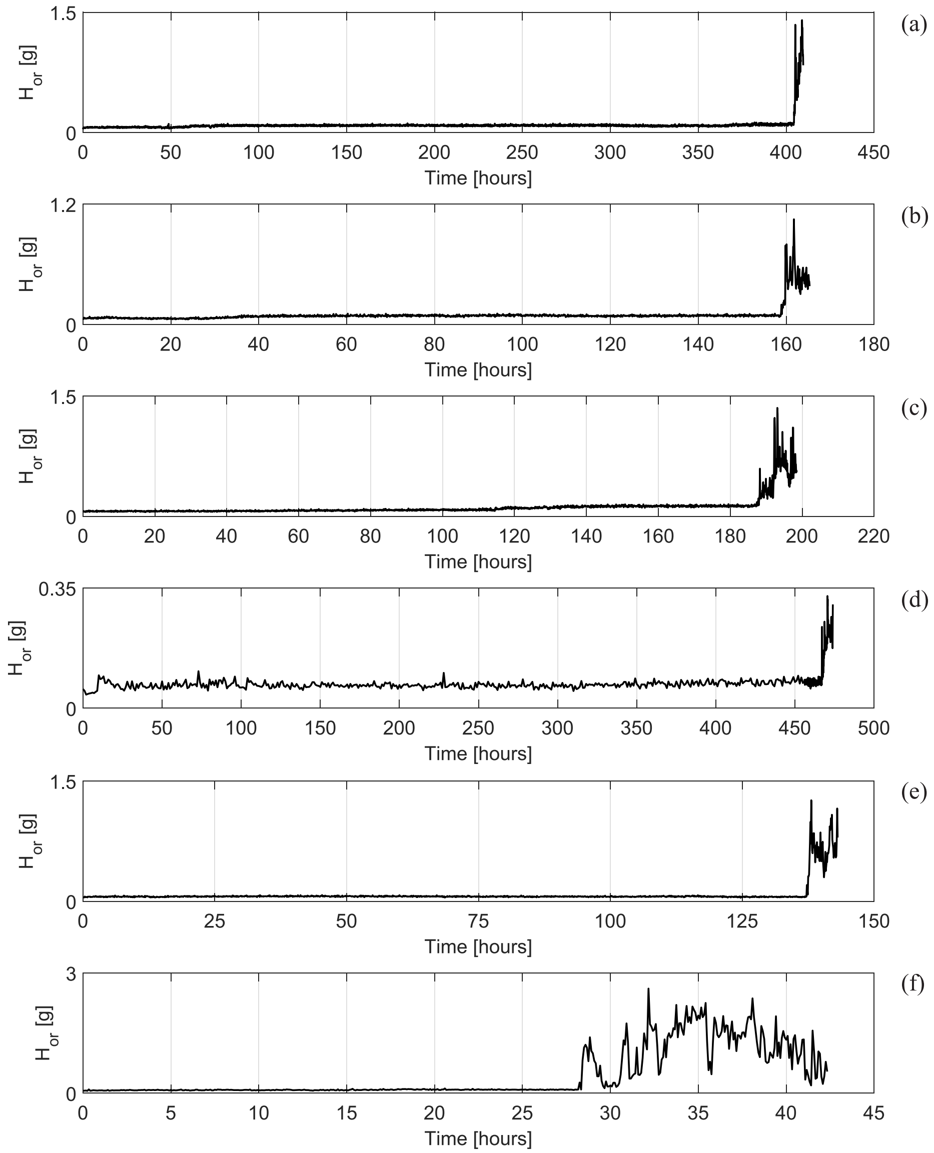


**Fig. 13.** Experimental values of indicator $H_{or}$: (a) E1 (4 kN); (b) E2 (4 kN); (c) E3 (4 kN); (d) E4 (3 kN); (e) E5 (4.7 kN); (f) E6 (5 kN).

extent is sufficiently small, the depth does not give any contribution to the resulting RMS level. 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 $\Delta\phi_d$ and $h_d$ pairs characterized by the highest values. This is in accordance with the experimental evidence that the vibration of the system greatly increases with large extended defects. However, it is worth underlining that RMS values shown in Fig. 14 refer only to the system under exam. In fact, the experiments shown in Section 3 were performed on a dedicated test bench, in which the accelerometer was mounted directly on the housing of the bearing being tested. It this context, it should be noted that actual industrial applications might involve complex machinery. In these cases, the effect of the transmission path from the bearing to the accelerometer should be included within the LP model.

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 Fig. 12 is enhanced by utilizing a moving average with a sliding window of length equal to 6 data points. Then, the procedure detailed in Section 2.2 is applied to extract the associated $V_{eq}$ values over time. These are reported in Fig. 15.a for tests E1, E2 and E3 and in Fig. 15.b for tests E4, E5, and E6, respectively. The results show that in each test $V_{eq}$ maintains a value which is 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

**Table 4**
LP 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 |
| Resonator mass | $m_r$ | 3.01 kg |
| Ball mass | $m_b$ | $1.47 \times 10^{-3}$ kg |
| Shaft stiffness in the × and y directions | $k_i$ | 75 MN/m |
| Resonator stiffness | $k_r$ | $4.25 \times 10^3$ MN/m |
| Contact stiffness | $k_c$ | 5927 MN/m$^{1.5}$ |
| Inner ring and shaft damping | $c_i$ | 2834.1 N s/m |
| Outer ring, casing and pedestal damping | $c_r$ | 186.0 N s/m |
| Rayleigh damping coefficient proportional to mass | $\alpha_c$ | 154.9 s$^{-1}$ |
| Rayleigh damping coefficient proportional to stiffness | $\beta_c$ | $2.5 \times 10^{-5}$ s |

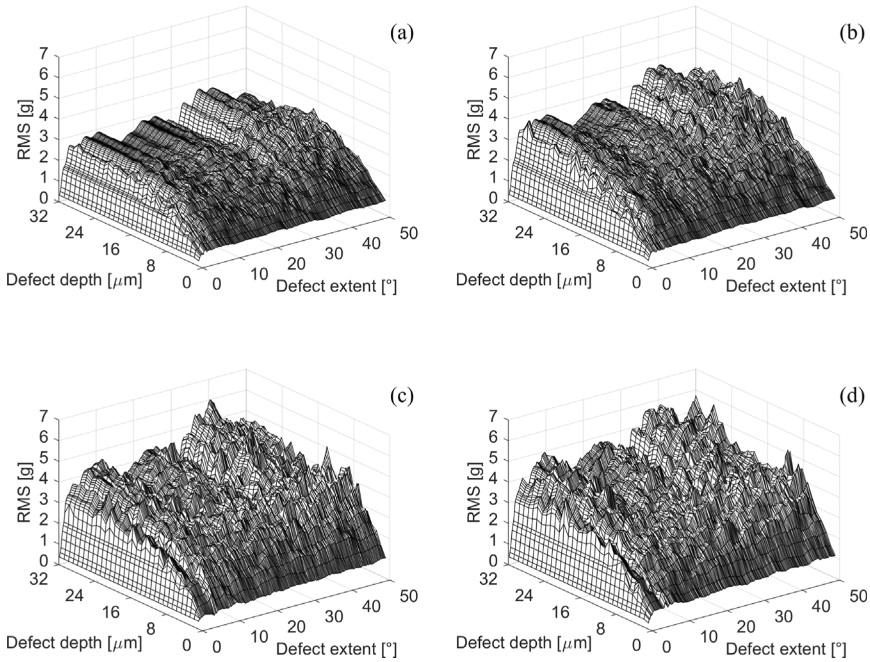


**Fig. 14.** RMS maps for a constant rotation speed equal to 40 Hz and different values of applied load: (a) 3000 N; (b) 4000 N; (c) 4700 N; (d) 5000 N.

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 exception of the final part of test E4. In fact, with reference to Fig. 6, the EDV algorithm was defined in order to have at each instant the possibility 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 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 Fig. 12. This is another important matter that makes the equivalent damaged volume $V_{eq}$ a very promising parameter for further employment in prognostic models.

Finally, it is worth noting that, although run-to-failure experiments were performed for different load values, only a single shaft speed was employed. This choice was arbitrarily made in order to obtain degradation data as quickly as possible with the test bench described in Section 3.1. As a result, the prognostic framework is tested for a single speed only. In spite of this, the proposed EDV algorithm may be employed for different shaft speeds without any kind of modifications, provided that the bearing is working inside its allowable operative range. In fact, if run-to-failure experiments were run at different speeds, the only change would concern the time needed to induce a fault in the bearing and reach the final stage of its life. However, future inquiries on the EDV algorithm might involve testing under different shaft speeds, in order to further justify the validity of this assumption.

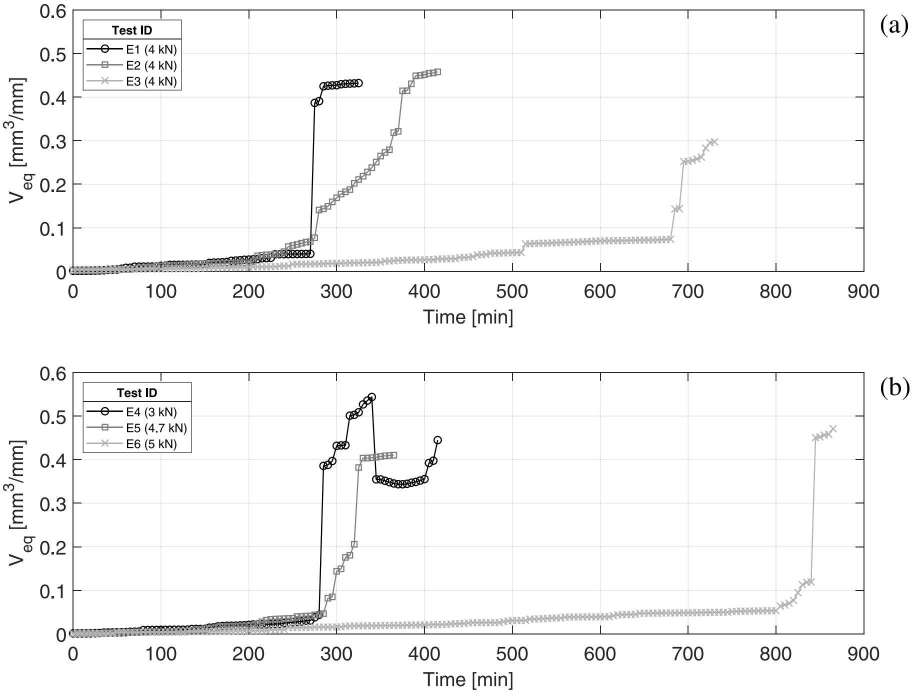


**Fig. 15.** Equivalent damaged volume estimated through the EDV algorithm: (a) Tests E1, E2 and E3; (b) Tests E4, E5 and E6.

### *4.2. Application of the prognostic PBM*

The proposed prognostic model, which is based on the EDV algorithm, involves the prediction of the TT between the actual time and a pre-determined threshold $V_{th}$. While this is not a direct measure of the RUL, Fig. 15 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 2.3 is applied to $V_{eq}$ values retrieved from the experimental measurements. The capabilities of the proposed method are then evaluated by computing the TT over time for all tests depicted in Fig. 15.

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 Fig. 15. However, this choice is arbitrary and it is referred to the bearing under examination. When using the algorithm on other applications with other bearings, the authors suggest to re-tune it on the basis of experimental observations specifically referred to the new bearing geometry. In this context, the authors would like to state also that bearings with similar dimensions and similar materials are expected to have a similar threshold. Moreover, this parameter is quite robust, since a variability of ±30% does not seriously affect the performance of the prediction. Nevertheless, this assumption allows to show the capabilities of the EDV algorithm as the foundation of prognostic models, which is the main focus of the manuscript. In future inquiries, it would be indeed useful to employ different threshold values to further improve the efficiency of the prognostic framework. The resulting $V_{eq}$ histories are reported in Fig. 16.a for tests E1, E2 and E3 and in Fig. 16.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 Fig. 16 are fitted with a polynomial of order two. The fitted curves for each test are shown in Fig. 17. From these 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 goodness-of-fit of the data [49]. The values of $R^2$ are reported in Table 5 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 Fig. 17.d. In fact, the gradient of the degradation trend consistently increases even below 0.04 mm$^3$/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. Further inquiry on the best-fitting 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 2.3 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

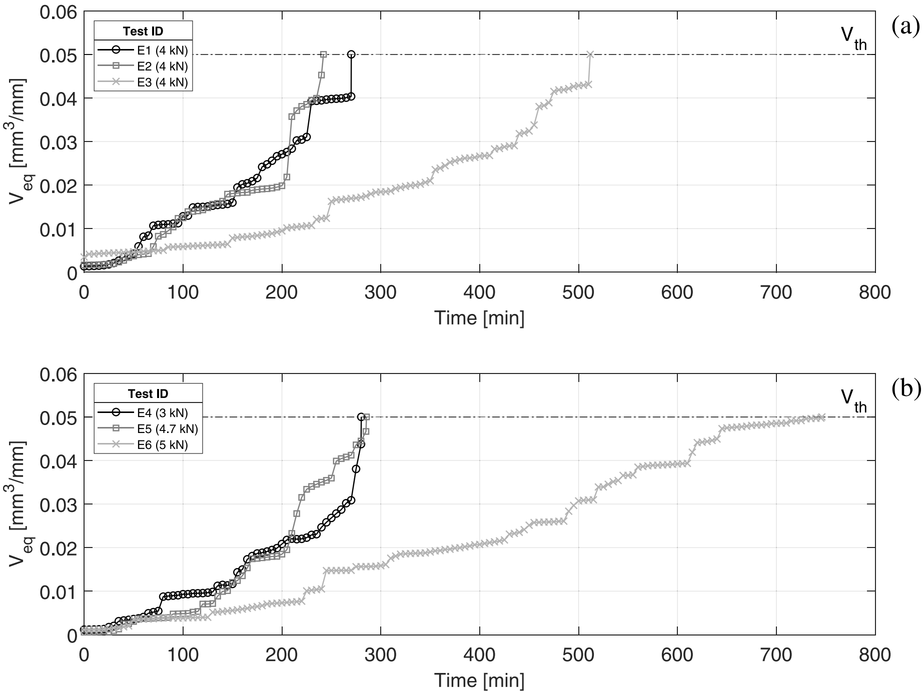


**Fig. 16.** Values of $V_{eq}$ below the threshold $V_{th}$ = 0.05 mm$^3$/mm: (a) Tests E1, E2 and E3; (b) Tests E4, E5 and E6.

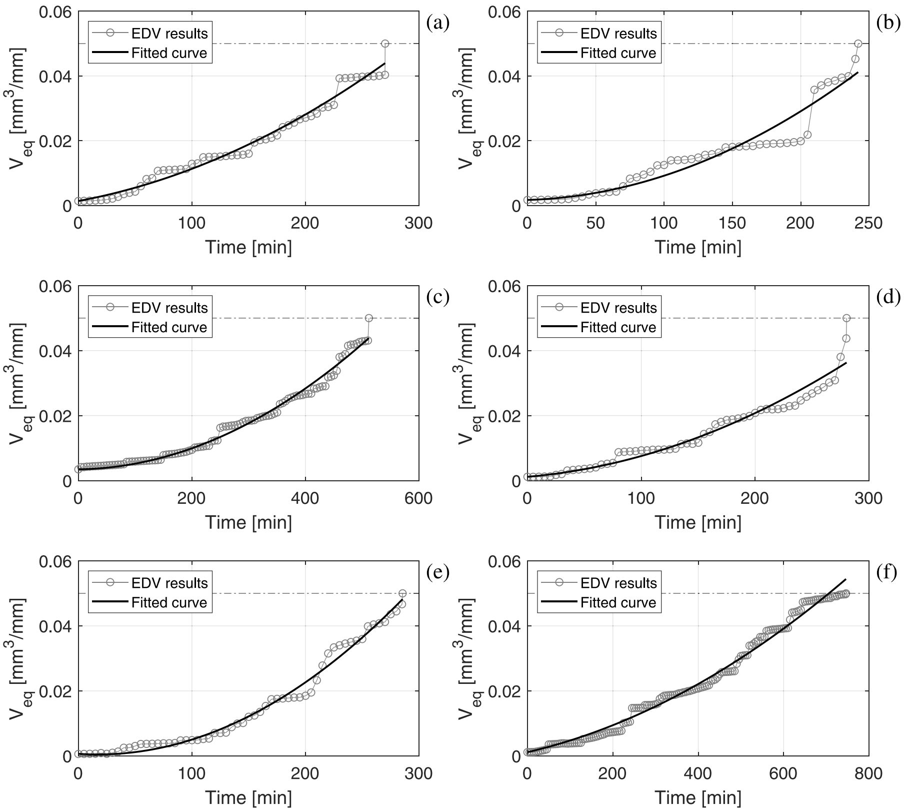


**Fig. 17.** $V_{eq}$ values below threshold fitted with a quadratic polynomial: (a) E1 (4 kN); (b) E2 (4 kN); (c) E3 (4 kN); (d) E4 (3 kN); (e) E5 (4.7 kN); (f) E6 (5 kN).

**Table 5**
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 ($\times 10^{-7}$) | b ($\times 10^{-5}$) | c ($\times 10^{-3}$) | $R^2$ |
|---|---|---|---|---|
| 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 |

**Table 6**
Percentage of time that the predicted TT falls inside ±15% of the total time below threshold.

| Test ID | Time limits [min] | Time when limit is crossed [min] | Percentage of time within limits |
|---|---|---|---|
| E1 | 40 | 145 | 45.3% |
| E2 | 36 | 35 | 85.1% |
| E3 | 77 | 110 | 78.2% |
| E4 | 42 | 105 | 61.8% |
| E5 | 43 | 65 | 76.8% |
| E6 | 112 | 385 | 48.0% |

polynomial. Subsequently, TT is computed by 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 test time progresses. TT values estimated through this procedure are shown in Fig. 18. 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 actual TT values. Fig. 18 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 Fig. 18.b, .c, .d and .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 the very beginning of the final stage may lead to extremely low values of TT, e.g. as in test E3 plotted in Fig. 18.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 min. 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, as it may be seen from the considered plots.

In addition, Fig. 18 also shows the comparison with a benchmark method, namely *RMS method*. The method involves the direct employment of the RMS in place of $V_{eq}$ as input for the prognostic PBM. For this purpose, RMS values up to the failure threshold are fitted with a quadratic polynomial. Then, TT values are computed in the same manner as if $V_{eq}$ was employed through the previously outlined procedure. This comparison is done in order to justify the need to extract $V_{eq}$ by employing the EDV algorithm, rather than directly employing a statistical indicator of the signal such as the RMS value. Indeed, Fig. 18 shows that the RMS method is not capable of accurately assessing TT values at any point during the final stage of bearing life. This is mainly due to the fact that the RMS trend is not monotonically increasing, as illustrated in Fig. 12. On the other hand, because of the characteristics of the EDV algorithm, $V_{eq}$ values always increase with time.

To further analyse the results, Fig. 19 depicts the difference between the actual TT and the TT evaluated through the proposed PBM for each test. In every performed experiment, the difference sensibly reduces after a certain amount of time. This may be evaluated by assigning a limit value and subsequently estimating the time needed to reduce the TT difference below this value. To this end, in each test the limit value is set as the 15% of the time between $t_{fs}$ and the instant in which the equivalent damaged volume reaches the threshold $V_{th}$. The time limits computed in this manner are reported in Table 6 along with the time in which $V_{eq}$ transitions below the assigned limit. These limits are also shown in Fig. 19 with dash-dot lines. In addition, Table 6 also reports the percentage of points that fall within the 15% limits. For tests E1 and E6, the difference is slightly below 50%. On the other hand, for the other tests the number of points within limits is higher, up to 85.1% for test E2. These results allow to assess the goodness of the proposed methodology in providing a reliable estimate of TT values.

The critical spall dimension, i.e. 0.05 $mm^3/mm$, is chosen to fit the data with a quadratic polynomial and assess time-to-threshold values. However, this decision is based on the results of an experimental campaign carried out for one specific bearing, namely a self-aligning ball-bearing characterized by the dimensions reported in Table 1. For other kinds of bearing, i.e. characterized by different dimensions and geometries, the critical spall dimension may vary. This aspect, however, may only be calibrated through experiments. For example, the size of the equivalent critical spall should be related to a condition in which bearing operation is no longer acceptable. For the run-to-failure experiments described in Section 3, an equivalent damage equal to 0.05 $mm^3/mm$ corresponds to RMS values which are deemed not tolerable in terms of noise and vibration. Furthermore, the progression of $V_{eq}$

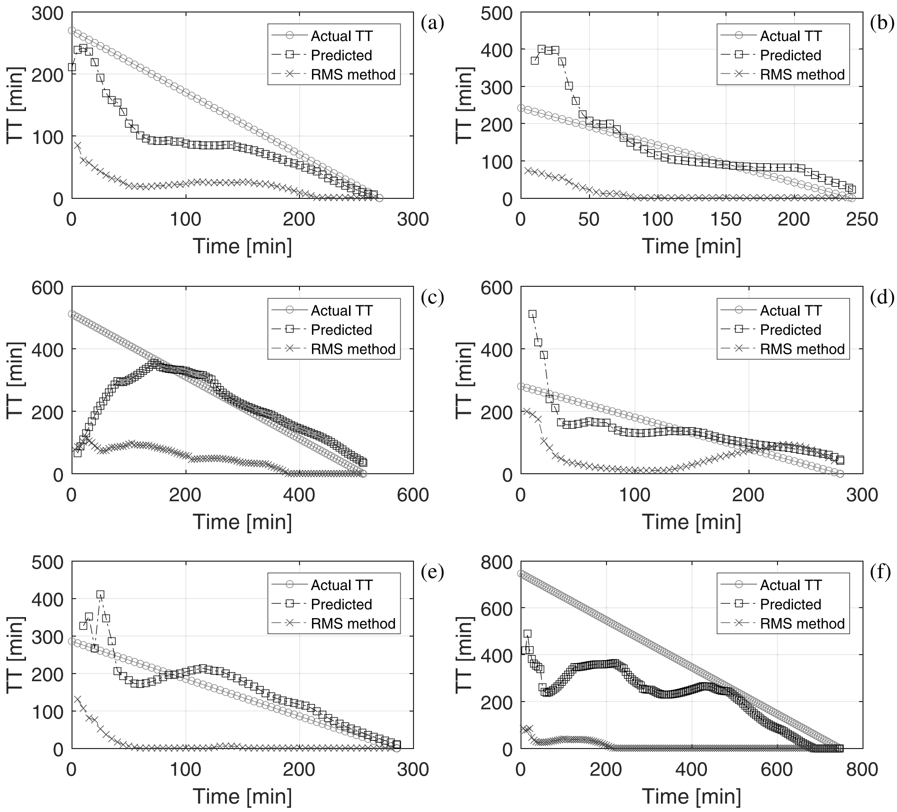


**Fig. 18.** TT computed through the proposed method. The results are compared against the actual TT and the TT estimated by the RMS method: (a) E1 (4 kN); (b) E2 (4 kN); (c) E3 (4 kN); (d) E4 (3 kN); (e) E5 (4.7 kN); (f) E6 (5 kN).

values over time was found to be well-fitted by a quadratic-polynomial. For applications involving other kinds of bearing, the validity of both assumptions should be checked to further ameliorate the quality of TT prediction.

With the purpose to further analyse the results, it is worth noticing that the choice of the fitting type is not mandatory and the user is free to use the preferred fitting approach. Indeed, the EDV methodology is suitable to work with whatever fitting approach, as long as the chosen fitting method provides a statistically good representation of the data. In this context, the influence of the fitting method is hereinafter evaluated by repeating the estimation with other three fitting approaches: namely a $3^{rd}$-order polynomial, a 4th-order polynomial and the exponential function. Fig. 20 reports the results of the analysis, where the estimated TT for the four different fitting schemes is compared against the experimental results. As it may be appreciated, both 3rd and 4th order polynomials struggle to provide a stable and accurate estimation of the TT. On the other hand, 2nd order polynomial and exponential have a similar capabilities, with the former performing generally better than the latter, in particular in Fig. 20(a), (b), (e) and (f). The reason at the basis of the obtained results stands on the fact that the EDV algorithm has the fundamental property of producing estimations of the Equivalent Damage Volume characterized by a monotonically increasing behaviour with a hardening trend. Then, the algorithm requires to apply a fitting method on this outcome to predict the TT. The type of fitting is not mandatory, but since the EDV is a monotonically increasing function, it is highly recommended to use a proper fitting method, such as the 2nd order polynomial and the exponential.

The reported investigation highlights the capabilities of the developed approach. However, to conclude the discussion on this matter, it is worth underlining some relevant aspects of the EDV algorithm and its potential further developments within the field of bearing prognostics. In this regard, this endeavour only dealt with outer races defects. This was based on the observation that the defect which generated during 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 associated to a different combination of load and speed. It is evident that a similar approach requires a large initial computational effort in order 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 computations 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 a single one.

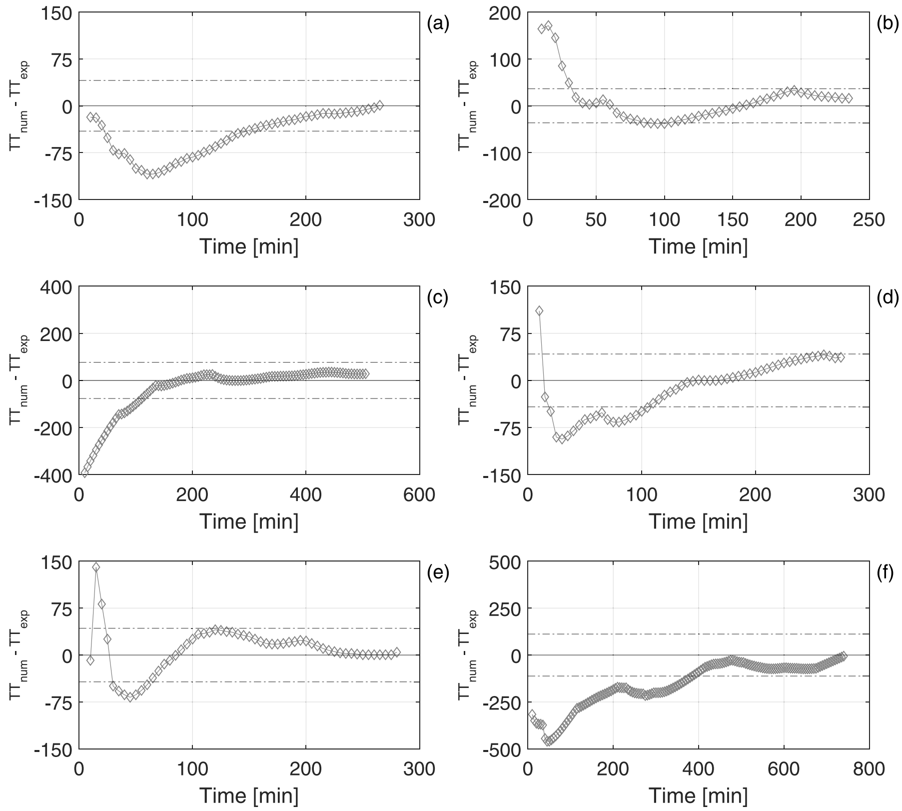


**Fig. 19.** Difference between actual and predicted TT: (a) E1 (4 kN); (b) E2 (4 kN); (c) E3 (4 kN); (d) E4 (3 kN); (e) E5 (4.7 kN); (f) E6 (5 kN).

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. In this regard, possible improvements to the EDV algorithm might include studies regarding complex machinery instead of a dedicated test bench. In actual industrial applications, in fact, the effect of the transmission path should not be neglected. Therefore, to extend the capabilities of the proposed technique, further investigation might concern the inclusion of the transmission path from the bearing to the accelerometer in the LP model. Concerning actual application, it is also worth noting that, at the current stage of development, the proposed prognostic framework may be applied to systems for which prior histories of faulty bearings were acquired through a sensor mounted in the same location. These data are necessary to tune the model parameters and to set the failure threshold. Consequently, the EDV algorithm is more well-suited for applications in which faulty bearing data may be acquired, e.g. bearings installed in factory machinery.

## 5. Conclusions

This paper introduced a novel parameter related to the evolution of bearing degradation during the final stage of REB life. This quantity, named *Equivalent Damaged Volume* (EDV), was exploited to propose a prognostic PBM for the assessment of the rolling bearing health state. 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. Then, it provides as outputs the equivalent damaged volume per unit of width $V_{eq}$ and the associated defect dimensions $\Delta\phi_d$ and $h_d$ for each point extracted from the experimental RMS history.

Concerning the prognostic PBM, its objective 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 the equivalent damaged volume crosses a pre-defined threshold. The prediction was based on the assumption that the damage evolution followed a quadratic polynomial function over time. 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. Despite the fact that the work carried out in this paper should be only considered a first approach to the problem of bearing prognostics through the EDV algorithm, it provides an original contribution on the subject. In fact, it introduces a modelling approach based on an indicator, i.e. the equivalent damaged volume, that is calculated by means of a dynamic model of a faulty bearing with localized or extended defects. Indeed, this is the main aspect of novelty of the detailed study.

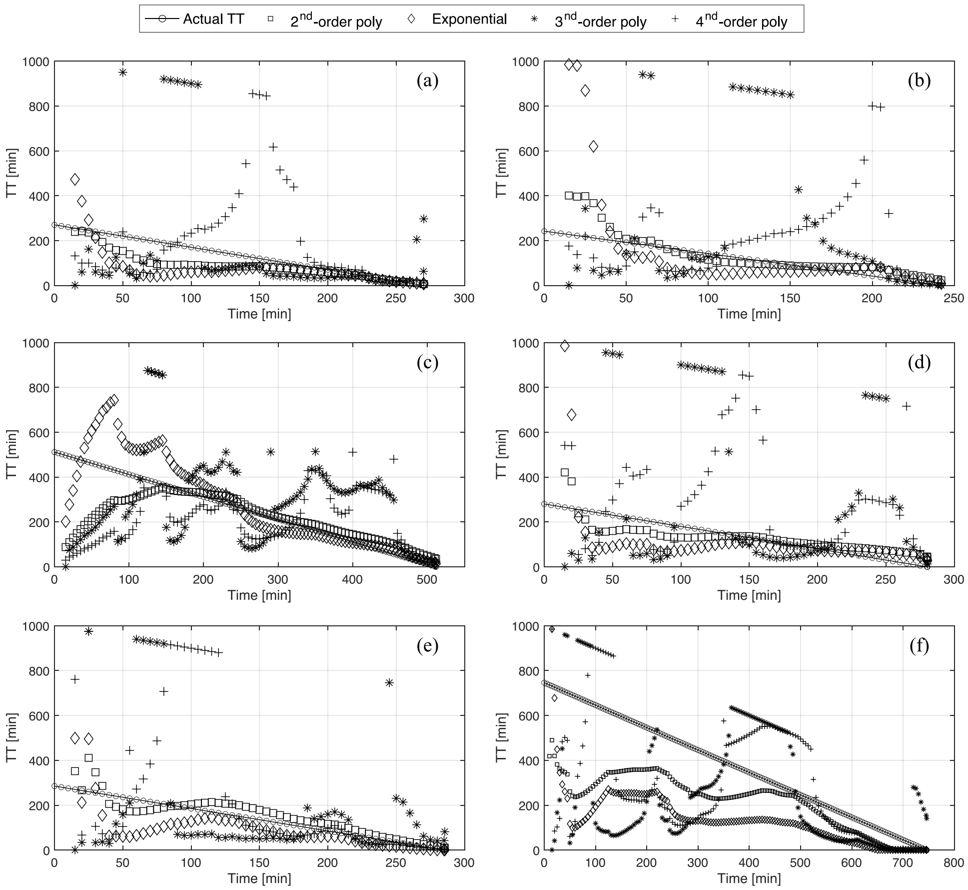


**Fig. 20.** TT estimated through the proposed method with four different fitting approaches and compared against the actual TT: (a) E1 (4 kN); (b) E2 (4 kN); (c) E3 (4 kN); (d) E4 (3 kN); (e) E5 (4.7 kN); (f) E6 (5 kN).

## CRediT authorship contribution statement

**Alberto Gabrielli:** Writing – original draft, Visualization, Methodology, Investigation, Data curation, Conceptualization. **Mattia Battarra:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Emiliano Mucchi:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Giorgio Dalpiaz:** Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

The employed data are stored in a data repository hosted by Mendeley Data and they will be described by means of a Data In Brief article. Due to limitations on the dataset dimension, the dataset has been split in two parts. The link to each part of the dataset may be found in [50,51].

## References

[1] T. Harris, M. Kotzalas, Essential Concepts of Bearing Technology: Rolling Bearing Analysis, fifth ed., CRC Press, 2006.
[2] R. Randall, J. Antoni, Rolling element bearing diagnostics—A tutorial, Mech. Syst. Signal Process. 25 (2) (2011) 485–520, http://dx.doi.org/10.1016/j.ymssp.2010.07.017.
[3] P. McFadden, J. Smith, Model for the vibration produced by a single point defect in a rolling element bearing, J. Sound Vib. 96 (1) (1984) 69–82, http://dx.doi.org/10.1016/0022-460X(84)90595-9.
[4] P. McFadden, J. Smith, The vibration produced by multiple point defects in a rolling element bearing, J. Sound Vib. 98 (2) (1985) 263–273, http://dx.doi.org/10.1016/0022-460X(85)90390-6.

[5] N. Tandon, A. Choudhury, An analytical model for the prediction of the vibration response of rolling element bearings due to a localized defect, J. Sound Vib. 205 (3) (1997) 275–292, http://dx.doi.org/10.1006/jsvi.1997.1031.

[6] D. Brie, Modelling of the spalled rolling element bearing vibration signal: an overview and some new results, Mech. Syst. Signal Process. 14 (3) (2000) 353–369, http://dx.doi.org/10.1006/mssp.1999.1237.

[7] N. Sawalhi, R. 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, Mech. Syst. Signal Process. 22 (8) (2008) 1924–1951, http://dx.doi.org/10.1016/j.ymssp.2007.12.001.

[8] N. Sawalhi, R. Randall, Simulating gear and bearing interactions in the presence of faults: Part II: Simulation of the vibrations produced by extended bearing faults, Mech. Syst. Signal Process. 22 (8) (2008) 1952–1966, http://dx.doi.org/10.1016/j.ymssp.2007.12.002.

[9] D. Petersen, C. Howard, N. Sawalhi, A. Moazen Ahmadi, S. Singh, Analysis of bearing stiffness variations, contact forces and vibrations in radially loaded double row rolling element bearings with raceway defects, Mech. Syst. Signal Process. 50–51 (2015) 139–160, http://dx.doi.org/10.1016/j.ymssp.2014.04.014.

[10] L. Cui, Y. Zhang, F. Zhang, J. Zhang, S. Lee, Vibration response mechanism of faulty outer race rolling element bearings for quantitative analysis, J. Sound Vib. 364 (2016) 67–76, http://dx.doi.org/10.1016/j.jsv.2015.10.015.

[11] A. Moazen Ahmadi, D. Petersen, C. Howard, A nonlinear dynamic vibration model of defective bearings – the importance of modelling the finite size of rolling elements, Mech. Syst. Signal Process. 52–53 (2015) 309–326, http://dx.doi.org/10.1016/j.ymssp.2014.06.006.

[12] Y. Jiang, W. Huang, J. Luo, W. Wang, An improved dynamic model of defective bearings considering the three-dimensional geometric relationship between the rolling element and defect area, Mech. Syst. Signal Process. 129 (2019) 694–716, http://dx.doi.org/10.1016/j.ymssp.2019.04.056.

[13] S. Gao, S. Chatterton, P. Pennacchi, 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, Mech. Mach. Theory 163 (2021) 104376, http://dx.doi.org/10.1016/j.mechmachtheory.2021.104376.

[14] A. Gabrielli, M. Battarra, E. Mucchi, G. Dalpiaz, A procedure for the assessment of unknown parameters in modeling defective bearings through multi-objective optimization, Mech. Syst. Signal Process. 185 (2023) 109783, http://dx.doi.org/10.1016/j.ymssp.2022.109783.

[15] E. Soave, G. D'Elia, M. Cocconcelli, M. Battarra, Blind deconvolution criterion based on Fourier–Bessel series expansion for rolling element bearing diagnostics, Mech. Syst. Signal Process. 169 (2022) 108588, http://dx.doi.org/10.1016/j.ymssp.2021.108588.

[16] J. Sikorska, M. Hodkiewicz, L. Ma, Prognostic modelling options for remaining useful life estimation by industry, Mech. Syst. Signal Process. 25 (5) (2011) 1803–1836, http://dx.doi.org/10.1016/j.ymssp.2010.11.018.

[17] A. Cubillo, S. Perinpanayagam, M. Esperon Miguez, A review of physics-based models in prognostics: Application to gears and bearings of rotating machinery, Adv. Mech. Eng. 8 (2016) http://dx.doi.org/10.1177/1687814016664660.

[18] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, J. Lin, Machinery health prognostics: A systematic review from data acquisition to RUL prediction, Mech. Syst. Signal Process. 104 (2018) 799–834, http://dx.doi.org/10.1016/j.ymssp.2017.11.016.

[19] E. Soave, G. D'Elia, G. Dalpiaz, Prognostics of rotating machines through generalized Gaussian hidden Markov models, Mech. Syst. Signal Process. 185 (2023) 109767, http://dx.doi.org/10.1016/j.ymssp.2022.109767.

[20] P. Paris, F. Erdogan, A critical analysis of crack propagation laws, J. Fluids Eng. 85 (1963) 528–533.

[21] Y. Li, S. Billington, C. Zhang, T. Kurfess, S. Danyluk, S. Liang, Adaptive prognostics for rolling element bearing condition, Mech. Syst. Signal Process. 13 (1) (1999) 103–113, http://dx.doi.org/10.1006/mssp.1998.0183.

[22] Y. Li, T. Kurfess, S. Liang, Stochastic prognostic for rolling element bearings, Mech. Syst. Signal Process. 14 (5) (2000) 747–762, http://dx.doi.org/10.1006/mssp.2000.1301.

[23] Y. Lei, N. Li, S. Gontarz, J. Lin, S. Radkowski, J. Dybala, A model-based method for remaining useful life prediction of machinery, IEEE Trans. Reliab. 65 (3) (2016) 1314–1326, http://dx.doi.org/10.1109/TR.2016.2570568.

[24] B. Wang, Y. Lei, N. Li, T. Yan, Deep separable convolutional network for remaining useful life prediction of machinery, Mech. Syst. Signal Process. 134 (2019) 106330, http://dx.doi.org/10.1016/j.ymssp.2019.106330.

[25] L. Liao, Discovering prognostic features using genetic programming in remaining useful life prediction, IEEE Trans. Ind. Electron. 61 (5) (2014) 2464–2472, http://dx.doi.org/10.1109/TIE.2013.2270212.

[26] J. Sun, H. Zuo, W. Wang, M. Pecht, Prognostics uncertainty reduction by fusing on-line monitoring data based on a state-space-based degradation model, Mech. Syst. Signal Process. 45 (2) (2014) 396–407, http://dx.doi.org/10.1016/j.ymssp.2013.08.022.

[27] M. Corbetta, C. Sbarufatti, A. Manes, M. Giglio, Real-time prognosis of crack growth evolution using sequential Monte Carlo methods and statistical model parameters, IEEE Trans. Reliab. 64 (2) (2015) 736–753, http://dx.doi.org/10.1109/TR.2014.2366759.

[28] F. Cadini, E. Zio, D. Avram, Monte Carlo-based filtering for fatigue crack growth estimation, Probab. Eng. Mech. 24 (3) (2009) 367–373, http://dx.doi.org/10.1016/j.probengmech.2008.10.002.

[29] E. Zio, G. Peloni, Particle filtering prognostic estimation of the remaining useful life of nonlinear components, Reliab. Eng. Syst. Saf. 96 (3) (2011) 403–409, http://dx.doi.org/10.1016/j.ress.2010.08.009.

[30] Y. Lu, Q. Li, S. Liang, Physics-based intelligent prognosis for rolling bearing with fault feature extraction, Int. J. Adv. Manuf. Technol. 97 (1) (2018) 611–620, http://dx.doi.org/10.1007/s00170-018-1959-0.

[31] L. Saidi, J. Ben Ali, M. Benbouzid, E. Bechhofer, An integrated wind turbine failures prognostic approach implementing Kalman smoother with confidence bounds, Appl. Acoust. 138 (2018) 199–208, http://dx.doi.org/10.1016/j.apacoust.2018.04.005.

[32] J. Qiu, B. Seth, S. Liang, C. Zhang, Damage mechanics approach for bearing lifetime prognostics, Mech. Syst. Signal Process. 16 (5) (2002) 817–829, http://dx.doi.org/10.1006/mssp.2002.1483.

[33] F. Sadeghi, B. Jalalahmadi, T. Slack, N. Raje, N. Arakere, A review of rolling contact fatigue, J. Tribol. 131 (4) (2009) http://dx.doi.org/10.1115/1.3209132.

[34] P. Gupta, E. Zaretsky, New stress-based fatigue life models for ball and roller bearings, Tribol. Trans. 61 (2) (2018) 304–324, http://dx.doi.org/10.1080/10402004.2017.1319524.

[35] M. Pais, N. Kim, Predicting fatigue crack growth under variable amplitude loadings with usage monitoring data, Adv. Mech. Eng. 7 (12) (2015) http://dx.doi.org/10.1177/1687814015619135.

[36] T. Wang, Z. Liu, N. Mrad, A probabilistic framework for remaining useful life prediction of bearings, IEEE Trans. Instrum. Meas. 70 (2021) 1–12, http://dx.doi.org/10.1109/TIM.2020.3029382.

[37] B. Yan, X. Ma, G. Huang, Y. Zhao, Two-stage physics-based Wiener process models for online RUL prediction in field vibration data, Mech. Syst. Signal Process. 152 (2021) 107378, http://dx.doi.org/10.1016/j.ymssp.2020.107378.

[38] L. Cui, X. Wang, H. Wang, H. Jiang, Remaining useful life prediction of rolling element bearings based on simulated performance degradation dictionary, Mech. Mach. Theory 153 (2020) 103967, http://dx.doi.org/10.1016/j.mechmachtheory.2020.103967.

[39] T. Wang, J. Y., D. Siegel, 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, http://dx.doi.org/10.1109/PHM.2008.4711421.

[40] A. Gabrielli, M. Battarra, E. Mucchi, A critical analysis of finite-element modeling procedures for radial bearing stiffness estimation, Math. Probl. Eng. 2021 (2021) 9955398, http://dx.doi.org/10.1155/2021/9955398.

[41] S. Rao, Mechanical vibrations, Pearson Education, Incorporated, 2017.

[42] H. Qiu, J. Lee, J. Lin, G. Yu, Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics, J. Sound Vib. 289 (4) (2006) 1066–1090, http://dx.doi.org/10.1016/j.jsv.2005.03.007.

[43] D. Petersen, C. Howard, Z. Prime, Varying stiffness and load distributions in defective ball bearings: Analytical formulation and application to defect size estimation, J. Sound Vib. 337 (2015) 284–300, http://dx.doi.org/10.1016/j.jsv.2014.10.004.

[44] J. Antoni, P. Borghesani, A statistical methodology for the design of condition indicators, Mech. Syst. Signal Process. 114 (2019) 290–327, http://dx.doi.org/10.1016/j.ymssp.2018.05.012.

[45] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Chebel-Morello, N. Zerhouni, C. Varnier, PRONOSTIA: An experimental platform for bearings accelerated degradation tests, in: Conference on Prognostics and Health Management, 2012, pp. 1–8.

[46] J. Wang, An intuitive tutorial to Gaussian processes regression, 2020, arXiv abs/2009.10862.

[47] A. Rohani Bastami, A. Aasi, H.A. Arghand, Estimation of remaining useful life of rolling element bearings using wavelet packet decomposition and artificial neural network, Iran. J. Sci. Technol. Trans. Electr. Eng. 43 (1) (2019) 233–245, http://dx.doi.org/10.1007/s40998-018-0108-y.

[48] B. Wang, Y. Lei, N. Li, N. Li, A hybrid prognostics approach for estimating remaining useful life of rolling element bearings, IEEE Trans. Reliab. 69 (1) (2020) 401–412, http://dx.doi.org/10.1109/TR.2018.2882682.

[49] C. Champ, D. Shepherd, Encyclopedia of Statistics in Quality and Reliability, John Wiley & Sons, Ltd, 2007.

[50] A. Gabrielli, L. Arpa, M. Battarra, E. Mucchi, "Run-to-failure vibration dataset of self-aligning double-row ball bearings - PART 1", Mendeley Data, V1, 2024, http://dx.doi.org/10.17632/htk59pp5wx.1.

[51] A. Gabrielli, L. Arpa, M. Battarra, E. Mucchi, "Run-to-failure vibration dataset of self-aligning double-row ball bearings - PART 2", Mendeley Data, V1, 2024, http://dx.doi.org/10.17632/zz8hpyx939.1.