ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus

Mécanique

Abderrahmane Ben Yagoub and Ridha Ziani

Bearing fault prognosis based on Cyclical Remaining Useful Life (CRUL)

Volume 353 (2025), p. 1477-1495

Online since: 19 December 2025

https://doi.org/10.5802/crmeca.332

The Comptes Rendus. Mécanique *are a member of the*
Mersenne Center for open scientific publishing
www.centre-mersenne.org — e-ISSN : 1873-7234

ACADÉMIE DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus. Mécanique
2025, Vol. 353, p. 1477-1495
https://doi.org/10.5802/crmeca.332

Research article

Bearing fault prognosis based on Cyclical Remaining Useful Life (CRUL)

Abderrahmane Ben Yagoub [*, a] **and Ridha Ziani** [a]

[a] Laboratory of Applied Precision Mechanics, Institute of Optics and Precision Mechanics, Setif 1 University Ferhat Abbas, Setif 19000, Algeria

E-mail: benaygoub.abde@univ-setif.dz

Abstract. Over the last decade, prognosis has become an interesting line of research in preventive maintenance for a variety of industrial applications. It generally focuses on estimating the Remaining Useful Life (RUL) of mechanical systems. In this context, this paper proposes a new prognosis approach based on cyclical prediction of RUL. The main idea lies in the integration of cyclical prediction of the degradation based on an appropriate adjustment of the data. The proposed degradation model can monitor continuously the evolution of degradation and finally provide an instantaneous estimation of RUL throughout the bearing life. First, the Simple Moving Average (SMA), Cumulative Moving Average (CMA), and Exponential Moving Average (EMA) filters are compared. Then, an exponential degradation model is constructed and fitted to the final n data using both Principal Component Analysis (PCA) and model fitting. Finally, the developed model is evaluated in cyclical estimation of the bearing RUL using vibration data acquired from an endurance test rig. The results demonstrate an accurate prediction throughout each phase of bearing life, which confirms its ability to be applied to the prognosis of bearing failure.

Keywords. Bearings, prognosis, preventive maintenance, cyclical remaining useful life.

Funding. This work was completed in Applied Precision Mechanics Laboratory (Setif 1 University Ferhat Abbas, Algeria). The authors would like to thank the Algerian Ministry of Higher Education and Scientific Research (MESRS) and The Directorate-General for Scientific Research and Technological Development (DGRSDT) for ensure financial support for PRFU research project N. A01L01UN190120220003.

Manuscript received 24 June 2025, revised 22 August 2025, accepted 15 October 2025.

1. Introduction

Rotating machines play a crucial role in industrial systems and they are used ubiquitously in different manufacturing applications. With the rapid development of modern industries, there is an excessive demand for increasing the availability of production equipment. Using these machines in unfavorable conditions, such as humidity and excessive loads, can lead to breakdowns, which can result in huge maintenance costs, degradation in production level, and potential risks to human safety [1].

Rolling Element Bearings (REBs) are one of the essential components in rotating machinery. They are used to allow the rotation of the shaft with respect to a fixed structure [2]. REBs are present in more than 90 % of the rotating machines [3,4]. Statistics show that the failure of these elements represents about 45 % to 55 % of breakdowns causes in rotating machines [5–7]. Such failures can be catastrophic, leading to monetary losses and interrupting industrial processes.

* Corresponding author

Due to their essential nature, the research on bearing condition monitoring has attracted considerable attention in the scientific community. Researchers are focused on understanding mechanisms of bearing deterioration and developing effective maintenance solutions that can be implemented to minimize unplanned failures, downtime, and improve overall performance of the system [8–10]. This research domain proposes a great variety of diagnosis and prognosis techniques that can be applied to assess machines health state. These techniques include: oil debris analysis, vibration-based condition monitoring, acoustic emission, infrared thermography, etc. [11].

Among the proposed techniques, vibration analysis has established itself as one of the most used condition-based monitoring strategies due to the wealth of information it contains and its effectiveness in early detection of typical defects [12,13].

In the last decade, Prognostics and Health Management (PHM) has become a major research focus in the field of preventive maintenance [14]. As a field of engineering, PHM aims to provide an integrated evaluation of the health status of machines or systems as a whole. One of the main elements of PHM is the Remaining Useful Life (RUL) estimate, which anticipates future degradation and helps plan interventions [15]. RUL estimation is mainly based on three approaches: data-driven approach, model-based approach, and experience-based approach [16]. Data-driven approaches use data provided by monitoring sensors which allow to create trends and predict RUL (i.e. learn the degradation model and predict the future health state of the system and its corresponding RUL) [17,18], model-based approach also named "physics of failure prognosis" use analytical model and mathematical equations to predict the future development of the system [19]. Finally, the experience-based approach is based on the use of simple reliability functions [20,21].

Generally, traditional approaches based on statistics are mainly used in the fields of demographics, economics, and medicine, but they are not sufficient for predicting the life expectancy of industrial components such as rolling bearings [22,23]. On the other hand, a series of new interesting researches combining statistics with artificial intelligence (AI) and model-based methods have been proposed to deal with the PHM of bearings [24]. For example, Zhang et al. [25]developed a Mixture Weibull Proportional Hazard Model (MWPHM) in order to predict the multiple failure modes of bearings. The parameters of this mixed model are estimated through the combination of monitoring data and historical lifetime of all failure modes. Caesarendra et al. [26] combined Relevance Vector Machines (RVM) and Logistic Regression (LR) to evaluate failure degradation and prediction form. Based on run-to-failure datasets, LR is used to estimate bearing failure deterioration. After that, RVM is used to forecast the likelihood that specific machine component units would fail. Ali et al. [27] proposed a data-driven prediction method that combines a neural network called Simplified Fuzzy Adaptive Resonance Theory Map (SFAM) and Weibull Distribution (WD). This method is proposed to deal with nonlinear time series where seven categories of output were studied: healthy bearings and six bearing wear conditions.

Guo et al. [28] proposed a hybrid approach for bearing defect prediction. First, a nonlinear health indicator (HI) was formed using full ensemble empirical mode decomposition (EEMD) with kernel principal component analysis and an adaptive noise to accurately and convincingly reflect the bearing health status. Then, multi-domain features were extracted from the vibration signals, and a two-channel transformer network with a convolutional block attention module was applied to create the HI for the remaining bearing. In addition, the 3σ criterion was used to set the health status monitoring interval and determine the first prediction time using a random-effect nonlinear Wiener process.

Wang et al. [29] combined Principal Component Analysis (PCA) and Improved Logistic Regression Model (ILRM) to solve the problems of difficult model establishment and difficult estimation of the remaining life of the rolling bearings. PCA was used for dimensionality reduction and usually to extract sensitive features from the original features. Then ILRM was used to create a model that reflects the degradation trend by eliminating the effect of fluctuations.

Given these literature reviews above, we can conclude that AI and data-driven models are considered the most appropriate for RUL prediction. In fact, the data-driven approach is simpler to implement than the model-based approach, especially for complex systems. Moreover, it is considered as the solution which makes the trade-off between implementation cost, model complexity, and precision. On the other hand, studies conducted in [28,29] are interesting in regard to the following aspects:

- The first one combines several advanced techniques to obtain an accurate estimate of the health status of bearings and predict their RUL. Although it is a hybrid approach for fine modeling, it requires complex management of parameters and resources, which sometimes does not allow to follow exact failure predictions.
- The second one is simple to implement due to its low cost of calculations (less resource-intensive) in addition to the reduction of noise and individual variations. However, the nature of regression may limit the model's ability to capture highly nonlinear dynamics inherent in complex degradation processes.

The shortcomings of both approaches are that they do not allow us to verify whether the most recent predictions are accurate, nor to observe precisely the life cycles of bearings.

Motivated by these limitations, the aim of this paper is to propose a prediction model that can be updated regularly, continuously monitor the evolution of degradation, and finally provide an instantaneous estimation of RUL throughout the bearing life. The developed method, that we named Cyclic Remaining Useful Life (CRUL), fits into data-driven prognostic approaches, more precisely into trend-based prognostics methods (extrapolation of statistical trends). The main idea lies in the integration of cyclical monitoring of the prognosis combined with degradation modeling based on an appropriate adjustment to recent data. The latter is done after comparative filtering and dimensional reduction by PCA, which allows an adaptive and continuous prediction of the RUL, even in the presence of limited run-to-failure data.

In order to implement this method a rigorous selection and comparison of smoothing filters: Simple Moving Average (SMA), Cumulative Moving Average (CMA) and Exponential Moving Average (EMA) allow to choose the best filter to capture authentic dynamics, thus improving the quality of the input data for modeling. Extracting a representative indicator by PCA facilitates understanding and monitoring the degradation. By fitting the model on a recent portion of the data, the method ensures that it reflects the current state of the system, which is crucial for accurate prediction of RUL (highlighting recent trends based on the fitting of the last n data). Thus, cyclical monitoring makes it possible to regularly update the model, continuously predict the evolution of degradation and consequently ensure an improvement of RUL estimation.

The rest of this paper is organized as follows: Section 2 explains the proposed approach for bearing fault prognosis; Section 3 describes the bearing dataset used in this study and provides a detailed description of its characteristics; the obtained results, highlighting the effectiveness of the method, are presented in Section 4; Section 5 draws conclusions of this work and outlines future perspectives.

2. The proposed prognostic procedure

In this paper, we propose to construct a new model for bearing fault prognosis. This model is based on an appropriate Moving Average Filter, PCA, and a model fitting. First, vibration data covering all life cycle of bearing is collected and processed to extract different features. After that, PCA is employed to reduce the dimensionality and to identify the principal component with the highest variance. Finally, model fitting is applied to ensure the inclusion of specific trends that could aid in forecasting future outcomes. As mentioned in the introduction, the goal of the proposed approach is to benefit from the advantages of each method composing this model to achieve high prognostic accuracy in such way a precise degradation model is built by enabling a cyclic estimation of RUL. This method that we named Cyclic Remaining Useful Life (CRUL) is not limited to find the RUL of the bearing, but also to verify and to predict each "life cycle" of the bearing using the whole of the dataset, unlike old methods which consider that certain data are unimportant or irrelevant. Figure 1 illustrates the flowchart of the proposed procedure. The principal steps are summarized as follows.

Step 1: Vibration data is acquired using sensors.
Step 2: Calculate various temporal and frequency-based features.
Step 3: Conduct a comparative evaluation of Moving Average Filters. This step aims to identify the optimal filter that accurately captures the true dynamics inherent in the data.
Step 4: Apply PCA to identify the principal component that will be used to reduce dimensionality of the dataset while keeping information as much as possible.
Step 5: Construct and fit the appropriate model using model fitting. This ensures the inclusion of specific trends that could aid in forecasting future outcomes.
Step 6: Apply the degradation model to the entire dataset. This step provides flexibility and the advantage of selecting the number of cycles for prediction, allowing precise results regarding the health state of the bearing and its remaining lifetime. More details about the CRUL will be presented in Section 4.2.

The main functions used to construct the proposed degradation model are presented in the following sections.

2.1. *Moving average filters*

In Step 3, to ensure both temporal precision and smooth tracking of continuous degradation for our model of cyclic prediction, three filters are compared: Exponential Moving Average (EMA) [30], Cumulative Moving Average (CMA) [31] and Simple Moving Average (SMA) [32]. These filters (also called rolling average or running average) are time series filters which calculate running weighted sum of time series. They were widely applied to smooth feature data in bearing faults diagnostic [33] and prognostic [34]. In this study, the final selection of the best filter (i.e., which reflects the actual moves of the real data) is determined following a comparative evaluation of the filters with regard to the following predefined criteria: (1) residual variance (noise reduction), (2) quality of exponential fit, (3) Sum of Squared Errors (SSE) or R^2, and (4) lag. In other words, the established criteria address: the variance value (the smaller it is, the better the filter eliminates noise), the minimal SSE or maximum R^2 (shows that the filter makes the degradation progression more "exponential," and consequently more exploitable by the proposed CRUL model), and the lag (which, if too large, induces a delay in fault detection).

The mathematical equations of the established criteria are presented below:

$$\mathrm{Var}_{\mathrm{residu}} = \mathrm{Var}\big(x(t) - x_{\mathrm{filtered}}(t)\big) \tag{1}$$

$$\widehat{x}_{\mathrm{fit}}(t) = Ae^{\alpha t} + B, \tag{2}$$

$$\mathrm{SSE} = \sum_{i=1}^{N} \big(x_{\mathrm{filtered}}(t_i) - \widehat{x}_{\mathrm{fit}}(t_i)\big)^2, \tag{3}$$

$$\mathrm{R}^2 = 1 - \frac{\mathrm{SSE}}{\sum\big(x_{\mathrm{filtered}}(t_i) - \overline{x}_{\mathrm{filtered}}\big)^2}, \tag{4}$$

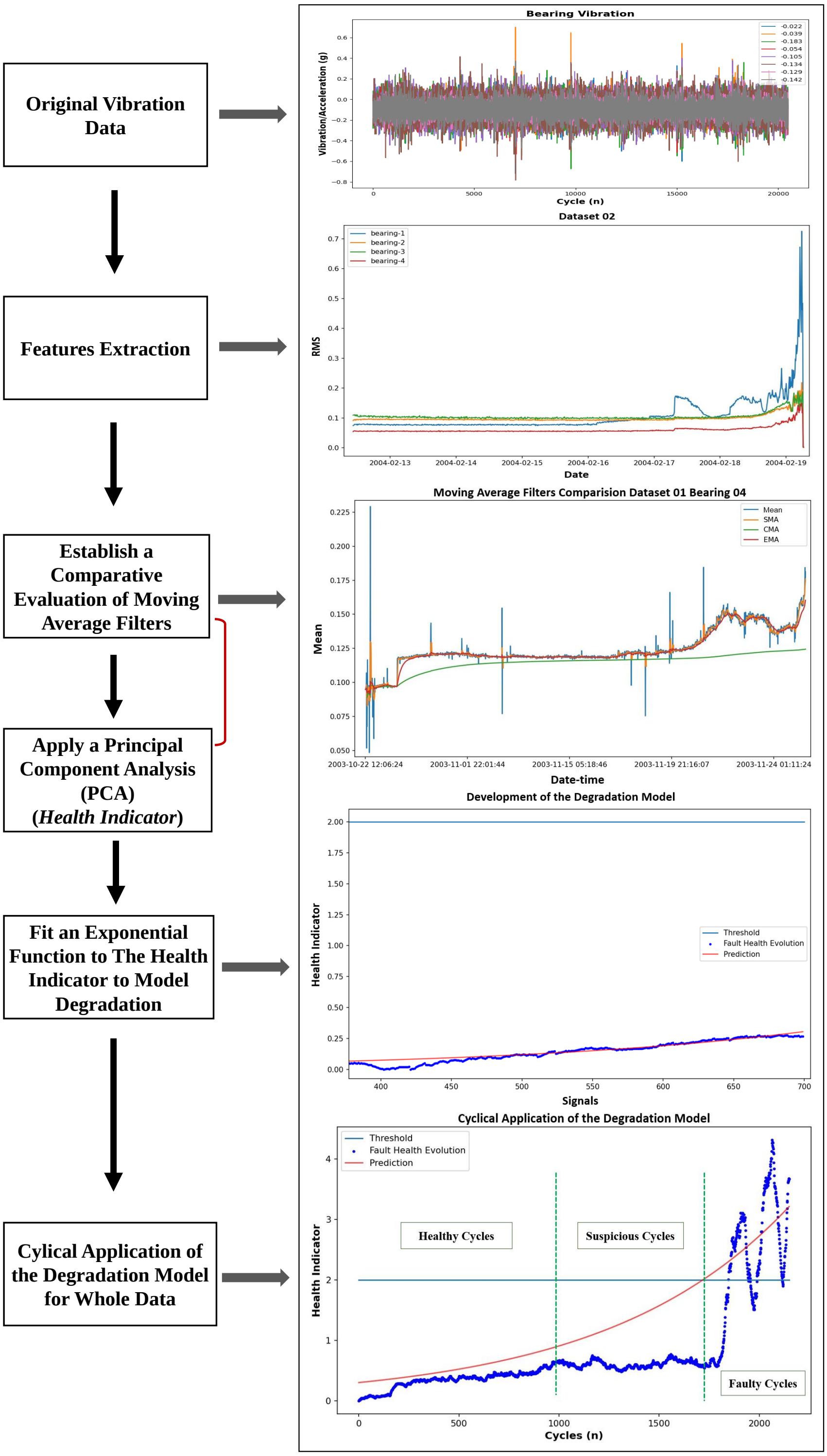


Figure 1. Flowchart of the proposed fault prognosis procedure.

where $x(t)$ is the signal at continuous time t, and $x_{\text{filtered}}(t)$ is the filtered version of the signal; $\widehat{x}_{\text{fit}}(t)$ is the fitted signal from the model with A, α, B as model parameters.

The mathematical expressions of the used moving average filters are presented in the Table 1, and the analysis of the differences between the filters is detailed in Section 4.1.

Table 1. Moving average filters expressions.

Algorithm of moving average filters		
Exponential Moving Average (EMA)	**Cumulative Moving Average** (CMA)	**Simple Moving Average** (SMA)
$\text{EMA}_t = \alpha\, x_t + (1-\alpha)\,\text{EMA}_{t-1}$ $\alpha = \dfrac{2}{n+1}$	$\text{CMA}_t = \dfrac{(t-1)\times \text{CMA}_{t-1} + x_t}{t}$	$\text{SMA}_t = \dfrac{1}{n}\sum_{i=0}^{n-1} x_{t-i}$

2.2. *Health indicator creation*

In Step 4, a health indicator (HI) is created by transforming the selected temporal indicators into a degradation index using principal component analysis. PCA is applied to all the temporal indicators calculated for each study. This late, reduces the dimensionality by projecting data on the first principal component (PC1) that corresponds to the axe that has the highest variance value. For more details about PCA technique, readers can refer to [29,35].

In the proposed approach, the use of a single component (PC1) is motivated by two complementary principles: parsimony (simplicity and interpretability) and the preservation of relevant information for the prognosis. To implement cyclical prediction and ensure the convergence of our degradation model yields steady estimates of the Remaining Useful Life (RUL), we set the requirement of capturing over 95 % of the first principal component (empirical threshold established: PC1 $\geq$ 0.95). Indeed, the first component PC1 is retained as a univariate health indicator when it explains a majority share of the variance. HI will then be used to track the evolution of the data and reduce noise during the next step of model fitting and will contribute to the normalization of the data during the prediction step of the degradation function.

2.3. *Model fitting*

In Step 5, an exponential model is established by fitting an exponential curve in order to predict the evolution of bearing degradation. This function has a direct impact on the health indicator and the degradation model. Thus, to estimate the Remaining Useful Life (RUL), we fit an exponential model to the normalized health indicator H(t) (PC1 derived from principal component analysis). Equation (5) shows the exponential model type:

$$\widehat{H}_{\text{fit}}(t) = Ae^{\beta t} + C, \tag{5}$$

where t is the cycle number and (A, β, C) are free parameters determined by fitting.

The establishment of this model is carried out through the following procedure:

Data selection: The last data of each cycle (healthy, suspect, failure) are used for training.

Exponential model fitting method: The model's parameters adjustment is performed by minimizing the sum of squared errors (SSE) using a 'curve fit' function of python that implements the robust Levenberg–Marquardt algorithm for nonlinear models [36].

Convergence criterion and fit quality: Return the optimized parameters A and B (where A and B represent the limits of the current cycle). First, we proceed to check the convergence through the stability of the cost function (*maxfev*) after a given number of iterations, then, we calculate the final SSE, and the R^2. Finally, the obtained parameters (A, β, C) are used to predict the failure cycle t_{fail}, (H_{thresh} being the health threshold).

Hence, the failure time is calculated as follows:

$$t_{\text{fail}} = \frac{\ln\left(\frac{H_{\text{thresh}} - C}{A}\right)}{\beta}. \tag{6}$$

Algorithm 1 details all operations in pseudocode for a better understanding of the approach.

Algorithm 1 Exponential model fitting for RUL Prediction

Require: $H(t_i)$: normalized health indicator at cycle t_i;
N_{base}: number of last cycles to fit;
H_{th}: critical threshold

Ensure: (A, β, C): fitted parameters
t_{fail}: predicted failure cycle

1: **Model:** $\widehat{H}(t) = A e^{\beta t} + C$
2: **Init:** $(A_0, \beta_0, C_0) = (0.01, 10^{-3}, 0)$
3: **Fit:** CURVE_FIT(exp_model$(t; A, \beta, C)$, $t[N-N_{\text{base}}+1 : N]$, $H[N-N_{\text{base}}+1 : N]$, $p_0 = [A_0, \beta_0, C_0]$, maxfev = 10000)
4: Enforce $\beta \leftarrow |\beta|$ to ensure monotonic growth
5: **Predict:**

$$t_{\text{fail}} = \frac{\ln\big((H_{\text{th}} - C)/A\big)}{\beta}$$

return $(A, \beta, C, t_{\text{fail}})$

2.4. *Predicting degradation function*

For mechanical elements, life degradation is generally divided into three periods: normal, decline, and failure [37,38]. According to Ali et al. [27], in the prognosis of rotating machines, it is possible to segment the degradation into "percentages" or "time slices" to decouple the stationary operating modes. In this study, the notion "life cycle" of the bearing refers to this commonly agreed designation. It is not an individual mechanical rotation, but it refers to a moment or, more precisely, a time interval corresponding to a segment of similar vibratory behavior of the machine. Each cycle t_i can cover two successive measurements (M_i and M_{i+1}), as it can be spread over several measures. The segmentation of these different moments (cycles) is done based on the number of signals and their obtained prediction results. Figure 2 describes the process of cyclic estimations of the RUL. In other words, this segmentation into regular time cycles rather than mechanical turns allows for a precise capture of the degradation evolution under steady-state conditions, without being dependent on speed variations or phase irregularities.

The exploitation of the degradation phase interval that we call a "cycle" allows an instantaneous estimation of the bearing's health status, and therefore a possible prediction of other suspicious or failing cycles. This designation ultimately allows for more accurate precision, given that it ensures the possibility of pinpointing the transition from one cycle as we understand it (as a phase within a time interval) to another cycle (a phase of degradation at another moment).

This process of cyclic estimation of RUL allows to define the nature of the instantaneous cycles (healthy cycle, suspicious cycle, and damaged cycle).

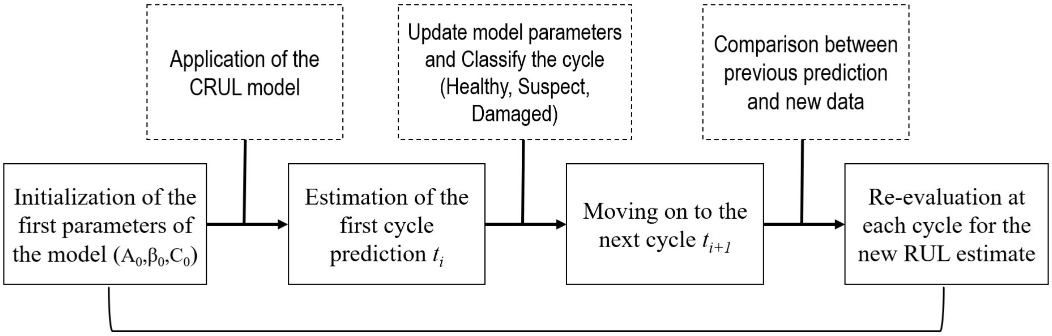


Figure 2. Detailed process of the CRUL model application.

The model established and trained in the previous steps using vibration data is finally ready to be used for prediction (i.e., estimating the cyclic remaining useful life CRUL). The prediction of future degradation (next cycle) is realized by taking into account the obtained parameters (A and B) of the exponential model of the current cycle.

3. Experimental validation

3.1. *Test rig description*

Vibration data used to validate the proposed approach is obtained from the IMS bearing dataset of the University of Cincinnati [39]. This data have been widely used to validate the effectiveness and the robustness of new algorithms for bearing fault prognosis [7,16,26,27,29,35,40–42]. The endurance test rig, presented in Figure 3, has the following characteristics.

Bearings and shaft: Four bearings having the same characteristics were mounted on a shaft. The rotation speed was preserved at 2000 RPM. The shaft was trained by an AC motor using rubber belts.

Load: A spring mechanism was used to generate a radial load of 6000 lbs (26 690 N) on the shaft and bearings. All bearings were force lubricated.

Bearing characteristics: Rexnord ZA-2115 double row bearings were employed. Figure 3 illustrates the configuration of bearings on the test rig.

Sensors: PCB 353B33 High Sensitivity Quartz ICP accelerometers were fixed on the bearing housings.

Failure details: All failures occurred after the designed lifetime of the bearings (exceeding 100 million revolutions).

Table 2. Characteristics of the four bearings Rexnord ZA-2115.

Characteristics	Value
Pitch diameter	2.815 in (71.5 mm)
Rolling element diameter	0.331 in (8.4 mm)
Number of rolling elements per row	16
Contact angle	15.17°
Static load	6000 lbs (26 690 N)

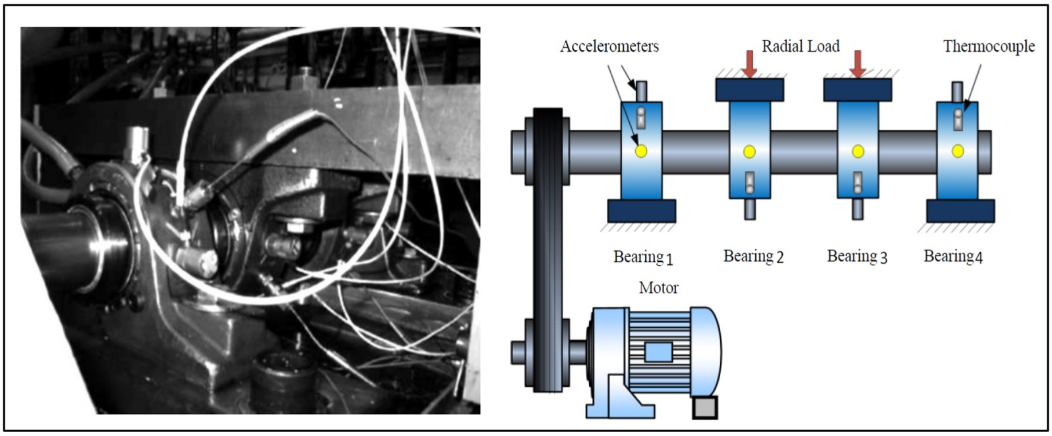


Figure 3. IMS test rig.

Table 3. Faults characteristic frequencies.

Frequency type		Value (Hz)
Shaft Frequency	SF	33.3
Ball Pass Frequency Outer race	BPFO	236
Ball Pass Frequency Inner race	BPFI	297
Ball Spin Frequency	BSF	278
Fundamental Train Frequency	FTF	15

3.2. *Data description*

Three datasets are available on the Nasa web site [43]. This data is composed of multiple files of one second each. Each file consists of 20 480 samples. Even the sampling frequency is declared to be 20 kHz, but when analyzing the data it turns out that the exact value is 20.48 kHz. Data acquisition and collection was done using NI DAQ Card 6062E. Intervals of time stamps are shown in file names. An acquisition of one second is performed every ten minutes except for the first forty-three files of dataset 1 for which an interval of five minutes has been adopted.

By reading the above studies, cited in Section 3.1, as well as in other studies that used this dataset, it appears that the dataset has never been fully processed. Some studies have even relied on an erroneous description of the third dataset, notably due to confusion over folder names and the number of files. Indeed, while the documentation [40] of the dataset mentions a “3rd set” folder containing 4448 files (recorded from March 4, 2004 at 09:27:46 to April 4, 2004 at 19:01:57), the current folder is named “4th test” and actually contains 6324 files, covering a period extending up to April 18, 2004 at 02:42:55. This divergence, which notably induces a defect at the outer race level in Bearing 03, was not noted in the studies that we reviewed, especially since the problem did not manifest itself when we limited ourselves to the 4448 files initially mentioned. This statement inspired us to create our cyclical degradation model. In fact, using a cyclic degradation model can lead to a precise estimation of RUL unlike the previous studies that were unable to predict the bearing’s health condition, even though the fault was present. This will be proved in the next section (Section 4.2).

The detailed description of the datasets is given in Table 4.

Table 4. Datasets description.

Dataset	Announced damages at the end	Number of files	Number of channels	Endurance duration	Duration of recorded signal
1	Bearing 3: inner race Bearing 4: rolling element	2156	8	34 days 12 h (49680 min)	36 min
2	Bearing 1: outer race	984	4	6 days 20 h (9840 min)	16 min
3	Bearing 3: outer race	6324	4	44 days 17 h 15 min (64395 min)	105 min 24 s

4. Numerical analysis and applications

4.1. *Features extraction and filters comparison*

Given the signal acquisition parameters, mainly the acquisition time which is relatively short (1 s every 10 minutes), it is clear that spectral analysis is not a suitable tool. In fact, the frequency resolution of 1 Hz at best is not sufficient to extract signifying information (i.e., information can be lost in 1 Hz). That is the reason behind the choice of the only time domain features given in Table 5.

By analyzing the graphical representations of three different filters CMA, EMA and SMA in Figure 4, a comparative evaluation was established and the following conclusion was returned: it appears that CMA cannot follow the evolution of features because it constantly diverges. The choice will therefore be limited to EMA and SMA. By comparing the evolution of "Mean" Bearing 04 Set 01 and "Std" Bearing 01 Set 02 we can easily conclude that EMA deals better for data filtering than SMA filter. This determination lies in its superior ability to accurately capture and mirror the actual moves observed in the real dataset without being influenced by small fluctuations, which suits us perfectly for our degradation model. The analysis of the results comparing the filters in Table 6 supports this conclusion. Indeed, the numerical data reveal, for the 'mean' indicator of Bearing 04 Set 01, the residual variance, the SSE, and the R^2 of the exponential fit via 'fitexp' function of python, as well as the average lag at the degradation threshold crossing. A reading of Table 6 shows that the CMA minimizes the SSE (0.0216) and maximizes the R^2 (0.7698); this is achieved through a high residual variance (0.000099), hence its decline. Conversely, the SMA offers the lowest variance (0.000027) but presents the least good fit (SSE = 0.0933, R^2 = 0.7271). The EMA serves as a good intermediary, as it combines both a fairly satisfactory noise reduction (variance = 0.000034) and a rather correct exponential fit (SSE = 0.0772, R^2 = 0.7594). Finally, we can note that regarding the average lag: all the filters are null, so this criterion does not differentiate them. At this stage, we are able to select the EMA as the optimal filter for our continuous degradation modeling by CRUL. We can assert that for CRUL modeling, which relies on an exponential fit, the EMA can be preferred given that this filter maintains the overall shape of the degradation (correct fit) and sufficiently filters the fluctuations, providing the best compromise between mitigating undesirable fluctuations and preserving the exponential degradation dynamics exploited by our CRUL method.

4.2. *Numerical results*

Dataset 01 — *Bearing 03 (inner race failure).* In the inner race failure case, the results illustrated in Figure 5 and Table 7 show that the predicted values from normal cycle to the suspected cycle

Table 5. Statistical feature definitions [44].

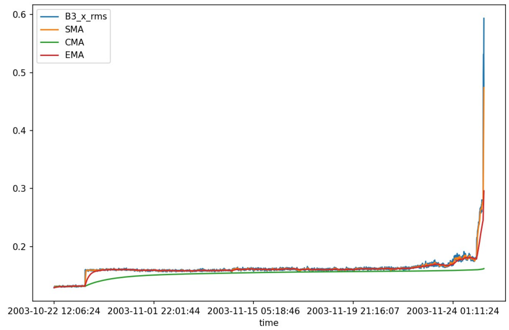
	Feature	Formula
(1)	Mean absolute value	$\bar{x}_{\text{abs}} = \frac{1}{n}\sum_{i=1}^{n}\lvert x_i\rvert$
(2)	Maximum	$X_{\max} = \max(x)$
(3)	RMS (Root Mean Square)	$\text{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$
(4)	Crest factor	$\text{CF} = \frac{X_{\max}}{\text{RMS}}$
(5)	Standard deviation	$\text{Std} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$
(6)	Peak-to-Peak	$\text{PP} = \max(x_i) - \min(x_i)$
(7)	Shape factor	$\text{IF} = \frac{\text{RMS}}{\frac{1}{n}\sum_{i=1}^{n}\lvert x_i\rvert}$
(8)	Impulse factor	$\text{IF} = \frac{\max(\lvert x_i\rvert)}{\frac{1}{n}\sum_{i=1}^{n}\lvert x_i\rvert}$
(9)	Skewness	$S_{\text{ke}} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{(n-1)\,\text{Std}^3}$
(10)	Kurtosis	$K_{\text{ur}} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{(n-1)\,\text{Std}^4}$
(11)	Clearance factor	$\text{Clf} = \frac{\max(\lvert x_i\rvert)}{\left(\frac{1}{n}\sum_{i=1}^{n}\sqrt{\lvert x_i\rvert}\right)^2}$
(12)	Shannon entropy	$\text{Ent} = -\sum_{i=1}^{n} x_i^2 \log(x_i^2)$

Table 6. Filter comparison results for "Mean" Bearing 04 Set 01.

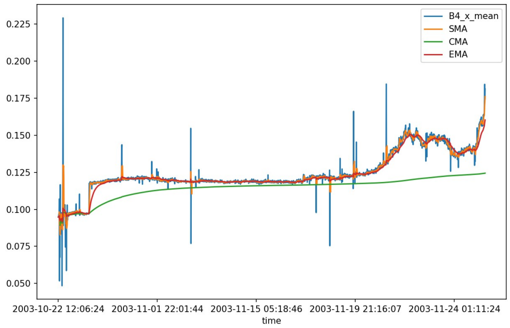
Filter	Var residu	SSE *fitexp*	R^2 *fitexp*	Lag (sample)
SMA	0.000027	0.093318	0.727149	0
CMA	0.000099	0.021625	0.769854	0
EMA	0.000034	0.077277	0.759491	0

(1775–1850) reflect not only the stability of detection but also the early aspect (ability to restore a clear trend). This was further verified at the transition from the suspected cycle to the faulty cycle (inner race failure) with a continuous monitoring (2100–2125) and a precise prediction of 1717.901. The prediction results of this case of study confirm that the proposed approach allows a rapid prognosis, reducing the uncertainty window to a few dozen cycles as well as stable monitoring, faithful to the real evolution of the degradation. The main states of health predicted by the proposed approach are:

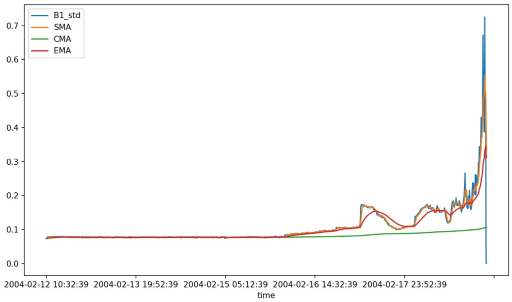
- Cycle 00: Start of prediction with healthy state (normal).



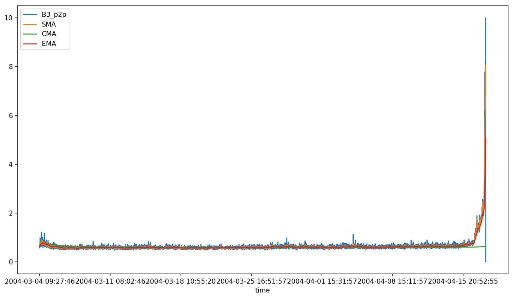
(a) “RMS” Bearing 03 Set 01



(b) “Mean” Bearing 04 Set 01



(c) “Std” Bearing 01 Set 02



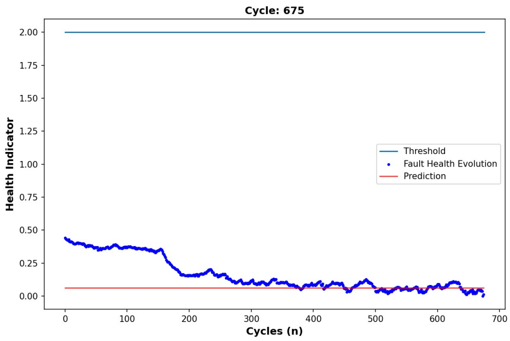
(d) “Peak to Peak” Bearing 03 Set 03

Figure 4. Comparison of filters EMA, SMA, CMA.

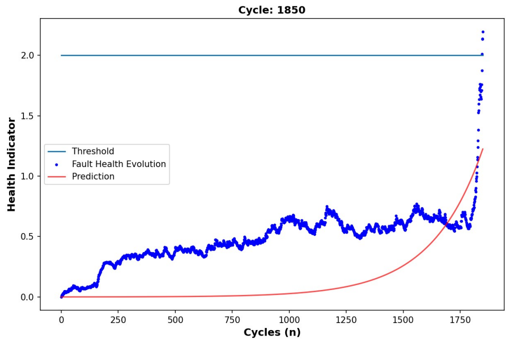
- Cycle 1800: Prediction remains invalid even the transition from normal to suspect state (Real = Suspect). However the prediction becomes valid at cycle 1850: under the Suspect state (1959.19, IsValid = True).
- Cycles 2000–2100: The correct predicted values (Real = Suspect) indicate that the predictions are still valid.
- Cycle 2125: With a successful prediction, the transition to inner race failure (Real = Inner race failure) occurs (cycle 1659.15).
- Cycle 2150: Stable prediction confirming failure state (1717.90, IsValid = True).

To ensure the interpretation of the results, the following definitions are specified for a more fluid reading of the variables in Table 7. The “Prediction” column corresponds to the estimated RUL by the degradation model at each cycle. The “IsValid” column distinguishes valid predictions that are consistent with the expected evolution from unrealistic (or erroneous) values reported as invalid (i.e., an excessively high or negative RUL value). Finally, the “Real” column represents the actual health status of the bearing.

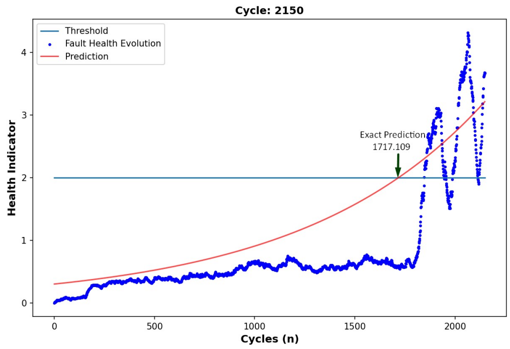
Dataset 01 — *Bearing 04 (roller element failure).* For Bearing 04 (Figure 6 and Table 8), we can see that the prediction of the suspect phase starts from cycle 1050. However, the validation of this state was not confirmed until cycle 1300: a prediction that should be highlighted because this “local evolution” was not considered as failure state and false alarm was avoided. At cycle 1450, the established model crossed a critical phase where the failure state of the rolling element was confirmed at early stage (level 1 of defect). In addition, the generalized defect (defect level 2) was declared at cycle 1850 with a prediction of 1645.049.



(a) Healthy cycle: 675



(b) Suspicious cycle: 1850



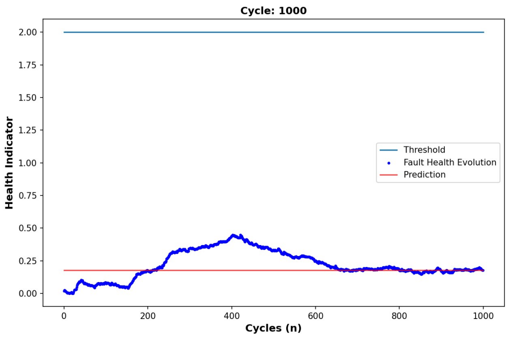
(c) Faulty cycle: 2150

Figure 5. Evolution of the instantaneous prediction for Bearing 03 at selected cycles.

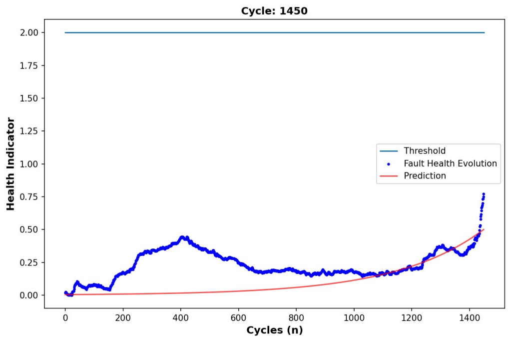
Table 7. Results of cyclical predictions for Set 01 Bearing 03 Inner_race_failure.

Index	Time	Cycle	Prediction	IsValid	Real
675	2003-11-09 00:51:44	675	5.838424e+10	False	Normal
700	2003-11-09 05:01:44	700	2.176056e+12	False	Normal
⋮		⋮	⋮	⋮	⋮
1775	2003-11-22 06:56:56	1775	2.547282e+03	False	Normal
1800	2003-11-22 11:06:56	1800	2.567900e+03	False	Suspect
1850	2003-11-22 19:26:56	1850	1.959186e+03	True	Suspect
⋮		⋮	⋮	⋮	⋮
2100	2003-11-24 20:57:32	2100	1.804905e+03	True	Suspect
2125	2003-11-25 11:47:32	2125	1.659149e+03	True	Inner_race_failure
2150	**2003-11-25 15:57:32**	**2150**	**1.717901e+03**	**True**	**Inner_race_failure**

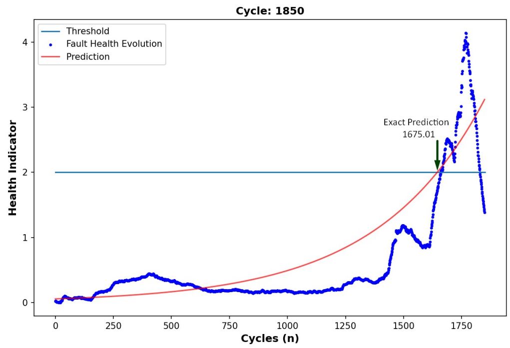
Dataset 02 — *Bearing 01 (outer race failure).* Figure 7 shows the evolution of the predictions given by the proposed method for the second case of study (Dataset 02 with outer race failure). In this figure, a remarkable stability is observed with a dynamic validation mechanism: stable monitoring in normal cycle reflects excellent stability even under significant fluctuations. A progressive trend is then observed with a regular rise from 820.72 (cycle 697) to 983.57 (cycle 961) which illustrates the suspect part. In this case, the model reflects a realistic degradation rather



(a) Healthy cycle: 1000



(b) Suspicious cycle: 1450



(c) Faulty cycle: 1850

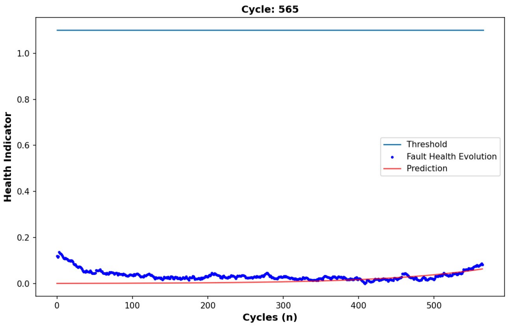
Figure 6. Evolution of the instantaneous prediction for Bearing 04 at selected cycles.

Table 8. Results of cyclical predictions of Set 01 Bearing 04 roller_element_failure.

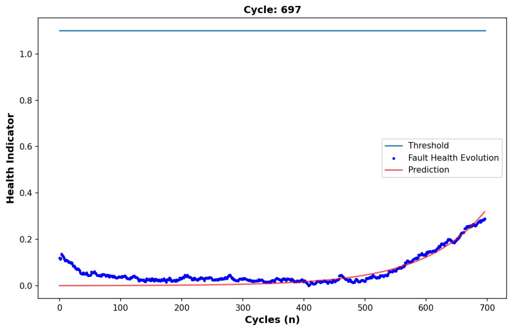
Index	Time	Cycle	Prediction	IsValid	Real
1000	2003-11-15 05:18:46	1000	4.481084e+10	False	Normal
1050	2003-11-15 13:38:46	1050	1.068279e+05	False	Suspect
⋮		⋮	⋮	⋮	⋮
1300	2003-11-17 20:52:30	1300	1.752231e+03	True	Suspect
⋮		⋮	⋮	⋮	⋮
1450	2003-11-19 12:56:07	1450	1.869256e+03	True	roller_element_failure
⋮		⋮	⋮	⋮	⋮
1800	2003-11-22 11:06:56	1800	1.678370e+03	True	roller_element_failure
1850	2003-11-22 19:26:56	1850	1.645049e+03	True	Stage_two_failure
2150	2003-11-25 15:57:32	2150	1.969998e+10	False	Stage_two_failure

than a sudden jump. Around cycle 972, few points (less than ten cycles) were sufficient to switch between suspect state and failure state with an accurate prediction at the end of 913.52 (cycle 983), see Table 9.

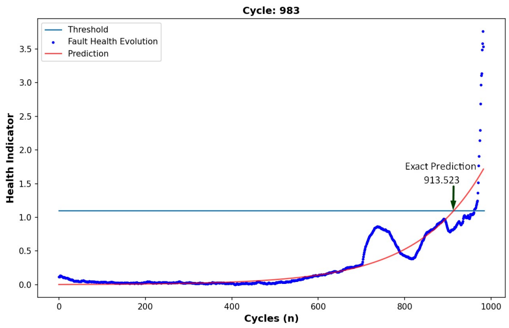
Dataset 03 — *Bearing 03 (outer race failure).* Figure 8 and Table 10 show the prediction results using Dataset 03 with outer race defect in Bearing 01. The results show that the proposed



(a) Healthy cycle: 565



(b) Suspicious cycle: 697



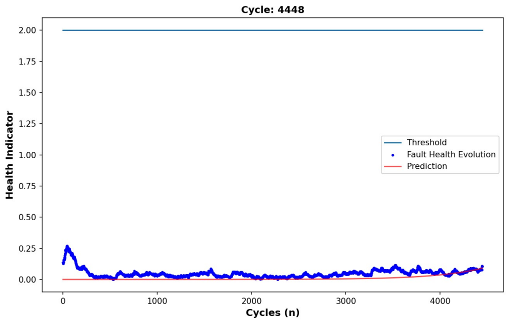
(c) Faulty cycle: 983

Figure 7. Evolution of the instantaneous prediction for Bearing 01 at selected cycles.

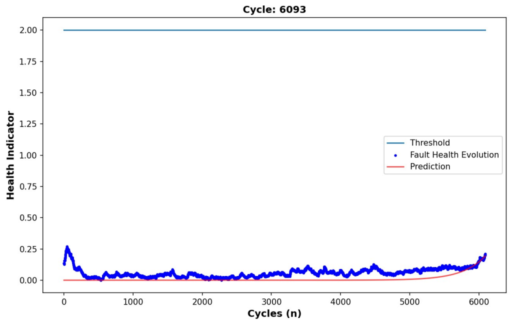
Table 9. Results of cyclical predictions of Set 02 Bearing 01 outer_race_failure.

Index	Time	Cycle	Prediction	IsValid	Real
675	2004-02-17 03:02:39	675	922.476378	True	Normal
697	2004-02-17 06:42:39	697	820.724113	True	Suspect
⋮		⋮	⋮	⋮	⋮
961	2004-02-19 02:42:39	961	983.572733	True	Suspect
972	2004-02-19 04:32:39	972	949.090671	True	Outer_race_failure
983	2004-02-19 06:22:39	983	913.523931	True	Outer_race_failure

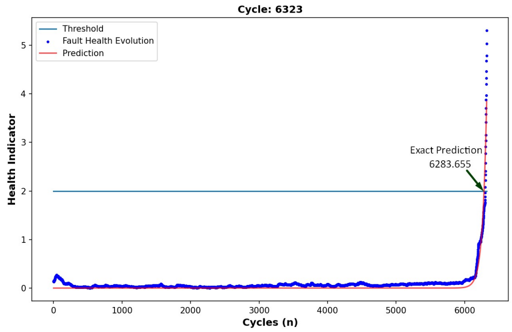
approach was not disturbed by the large amount of data, showing long-term stability with a high tolerance to extreme values. In fact, more than 5000 normal operating cycles were predicted without false alarms or major fluctuations. The transition to the suspect cycle at cycle 6093 shows a very early warning for a bearing. This warning is then confirmed at cycle 6208. The prediction stability allows around cycle 6300 to predict the failure of this bearing with an accurate prediction of 6283.655 at cycle 6323. The fact that this defect has never been highlighted in the previous studies that used this dataset demonstrates the effectiveness of the proposed methodology (ability to discover new failure modes without prior adjustment).



(a) Healthy cycle: 4448



(b) Suspicious cycle: 6093



(c) Faulty cycle: 6383

Figure 8. Evolution of the instantaneous prediction for Bearing 03 at selected cycles

Table 10. Results of cyclical predictions of Set 03 Bearing 03 outer_race_failure

Index	Time	Cycle	Prediction	IsValid	Real
4448	2004-04-04 19:01:57	4448	5990.419	False	Normal
⋮		⋮	⋮	⋮	⋮
5495	2004-04-12 01:41:57	5495	9251.958	False	Normal
⋮		⋮	⋮	⋮	⋮
6070	2004-04-16 08:32:55	6070	6969.723	False	Normal
6093	2004-04-16 12:22:55	6093	6864.935	False	Suspect
⋮		⋮	⋮	⋮	⋮
6208	2004-04-17 07:32:55	6208	6290.845	True	Suspect
⋮		⋮	⋮	⋮	⋮
6277	2004-04-17 19:02:55	6277	6300.985	True	Suspect
6300	2004-04-17 22:52:55	6300	6303.053	True	Outer_race_failure
6323	2004-04-18 02:42:55	6323	6283.655	True	Outer_race_failure

5. Conclusion

In this study a new prognosis approach was proposed. This approach, based on trend analysis of degradation, is named Cyclic Remaining Useful Life (CRUL). The idea is to combine the degradation model based on an appropriate moving average filter with PCA and a model fitting, allowing a cyclic estimation of the RUL. The conduced analysis on bearing datasets containing different types of defects proved that the exponential moving average EMA is undoubtedly the most effective filter to establish a degradation model compared to CMA and SMA models. Combined with an appropriate indicator calculation and a Principal Component Analysis PCA, the prediction algorithm showed precise results in predicting the remaining useful life of the bearings. Firstly, it showed a stable tracking of the trend of degradation with an early and robust detection. Secondly, it has an intelligent validity management (fine-grained management of false positives). Moreover, the reliability of validation was confirmed through different dataset analysis.

The robustness of the proposed approach was tested with three different defects on: inner race, outer race, and rolling element. In addition, for the entire dataset and reported defects, the proposed model was able to predict the life cycles of each bearing, validating its applicability in the prognosis of other defects or machines. Finally, this approach can be improved to be integrated in embedded systems.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Underlying data

The study's supporting data can be found in an open access, public repository [39].

Underlying data for this article is available at https://doi.org/10.57745/0DK1KE (see [45]).

References

[1] D. Neupane and J. Seok, "Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review", *IEEE Access* **8** (2020), pp. 93155–93178.

[2] H. Ocak and K. A. Loparo, "Estimation of the running speed and bearing defect frequencies of an induction motor from vibration data", *Mech. Syst. Signal Process.* **18** (2004), no. 3, pp. 515–533.

[3] S. Buchaiah and P. Shakya, "Bearing fault diagnosis and prognosis using data fusion based feature extraction and feature selection", *Meas.* **188** (2022), article no. 110506.

[4] P. Gupta and M. K. Pradhan, "Fault detection analysis in rolling element bearing: A review", *Mater. Today: Proc.* **4** (2017), no. 2, pp. 2085–2094.

[5] M. Behzad, H. Izanlo, A. Davoodabadi and H. A. Arghand, "Fault detection of rolling element bearing using a temporal signal with artificial intelligence techniques", *J. Theor. Appl. Vibr. Acoust.* **7** (2021), no. 1, pp. 55–71.

[6] M. Behzad, S. Feizhoseini, H. A. Arghand, A. Davoodabadi and D. Mba, "Failure threshold determination of rolling element bearings using vibration fluctuation and failure modes", *Appl. Sci. (Switz.)* **11** (2021), no. 1, article no. 160 (18 pages).

[7] A. K. Mahamad, S. Saon and T. Hiyama, "Predicting remaining useful life of rotating machinery based artificial neural network", *Comput. Math. Appl.* **60** (2010), no. 4, pp. 1078–1087.

[8] D. Ho and R. B. Randall, "Optimisation of bearing diagnostic techniques using simulated and actual bearing fault signals", *Mech. Syst. Signal Process.* **14** (2000), no. 5, pp. 763–788.

[9] P. D. McFadden and J. D. Smith, "Model for the vibration produced by a single point defect in a rolling element bearing", *J. Sound Vib.* **96** (1984), no. 1, pp. 69–82.

[10] O. G. Gustafsson and T. Tallian, "Detection of damage in assembled rolling element bearings", *ASLE Trans.* **5** (1962), no. 1, pp. 197–209.

[11] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao and D. Siegel, "Prognostics and health management design for rotary machinery systems — Reviews, methodology and applications", *Mech. Syst. Signal Process.* **42** (2014), no. 1–2, pp. 314–334.

[12] J. Hayouni, *Développement d'un modèle de pronostic pour les roulements des éoliennes*, Université du Québec à Rimouski (Canada), 2017.

[13] N. S. Jammu and P. K. Kankar, "A review on prognosis of rolling element bearings", *Int. J. Eng. Sci. Technol.* **3** (2011), no. 10, pp. 7497–7503.

[14] S. Kim, J.-H. Choi and N. H. Kim, "Challenges and opportunities of system-level prognostics", *Sensors* **21** (2021), no. 22, article no. 7655 (25 pages).

[15] K. Medjaher, D. A. Tobon-Mejia and N. Zerhouni, "Remaining useful life estimation of critical components with application to bearings", *IEEE Trans. Reliab.* **61** (2012), no. 2, pp. 292–302.

[16] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni and G. Tripot, "A mixture of gaussians hidden markov model for failure diagnostic and prognostic", in *2010 IEEE International Conference on Automation Science and Engineering*, IEEE, 2010, pp. 338–343.

[17] D. An, N. H. Kim and J.-H. Choi, "Practical options for selecting data-driven or physics-based prognostics algorithms with reviews", *Reliab. Eng. Syst. Saf.* **133** (2015), pp. 223–236.

[18] X.-S. Si, W. Wang, C.-H. Hu and D.-H. Zhou, "Remaining useful life estimation — A review on the statistical data driven approaches", *Eur. J. Oper. Res.* **213** (2011), no. 1, pp. 1–14.

[19] W. K. Yu and T. A. Harris, "A new stress-based fatigue life model for ball bearings", *Tribol. Trans.* **44** (2001), no. 1, pp. 11–18.

[20] A. Palmgren, "The service life of ball bearings", *Z. Vereines Deutscher Inge.* **68** (1924), no. 14, pp. 339–341.

[21] Y. Hu, S. Liu, H. Lu and H. Zhang, "Remaining useful life model and assessment of mechanical products: a brief review and a note on the state space model method", *Chin. J. Mech. Eng.* **32** (2019), pp. 1–20.

[22] K. Medjaher, N. Zerhouni and J. Baklouti, "Data-driven prognostics based on health indicator construction: Application to PRONOSTIA's data", in *2013 European Control Conference (ECC)*, IEEE, 2013, pp. 1451–1456.

[23] T. H. Loutas, D. Roulias and G. Georgoulas, "Remaining useful life estimation in rolling bearings utilizing data-driven probabilistic E-support vectors regression", *IEEE Trans. Reliab.* **62** (2013), no. 4, pp. 821–832.

[24] A. K. Jardine, D. Lin and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance", *Mech. Syst. Signal Process.* **20** (2006), no. 7, pp. 1483–1510.

[25] Q. Zhang, C. Hua and G. Xu, "A mixture Weibull proportional hazard model for mechanical system failure prediction utilising lifetime and monitoring data", *Mech. Syst. Signal Process.* **43** (2014), no. 1–2, pp. 103–112.

[26] W. Caesarendra, A. Widodo and B.-S. Yang, "Application of relevance vector machine and logistic regression for machine degradation assessment", *Mech. Syst. Signal Process.* **24** (2010), no. 4, pp. 1161–1171.

[27] J. B. Ali, B. Chebel-Morello, L. Saidi, S. Malinowski and F. Fnaiech, "Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network", *Mech. Syst. Signal Process.* **56** (2015), pp. 150–172.

[28] J. Guo, Z. Wang, H. Li, Y. Yang, C.-G. Huang, M. Yazdi and H. S. Kang, "A hybrid prognosis scheme for rolling bearings based on a novel health indicator and nonlinear Wiener process", *Reliab. Eng. Syst. Saf.* **245** (2024), article no. 110014 (16 pages).

[29] F. Wang, B. Wang, B. Dun, X. Chen, D. Yan and H. Zhu, "Remaining life prediction of rolling bearing based on PCA and improved logistic regression model", *J. Vibroeng.* **18** (2016), no. 8, pp. 5192–5203.

[30] J. Stuart Hunter, "The exponentially weighted moving average", *J. Qual. Technol.* **18** (1986), no. 4, pp. 203–210.

[31] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, Prentice Hall, 1975.

[32] C. Ya-Lun, *Statistical Analysis: With Business and Economic Applications*, Holt, Rinehart and Winston, 1963.

[33] S. Hansun, "A new approach of moving average method in time series analysis", in *2013 Conference on New Media Studies (CoNMedia)*, IEEE, 2013, pp. 1–4.

[34] Y. Su, C. Cui and H. Qu, "Self-attentive moving average for time series prediction", *Appl. Sci. (Switz.)* **12** (2022), no. 7, article no. 3602.

[35] A. Rehab, I. Ali, W. Gomaa and M. N. Fors, "Bearings fault detection using hidden Markov models and principal component analysis enhanced features", 2021. Online at https://arxiv.org/abs/2104.10519.

[36] J. J. Moré, "The Levenberg-Marquardt algorithm: implementation and theory", in *Numerical analysis: Proceedings of the biennial Conference held at Dundee, June 28–July 1, 1977*, Springer, 2006, pp. 105–116.

[37] Y. Zhang, L. Fang, Z. Qi and H. Deng, "A review of remaining useful life prediction approaches for mechanical equipment", *IEEE Sens. J.* **23** (2023), no. 24, pp. 29991–30006.

[38] N. Li, Y. Lei, X. Liu, T. Yan and P. Xu, "Machinery health prognostics with multimodel fusion degradation modeling", *IEEE Trans. Ind. Electron.* **70** (2022), no. 11, pp. 11764–11773.

[39] J. Lee, H. Qiu, G. Yu and J. Lin, *IMS, University of Cincinnati. "Bearing Data Set", NASA Ames Prognostics Data Repository*, May 29, 2025. Online at https://data.nasa.gov/dataset/ims-bearings (accessed on November 4, 2025).

[40] H. Qiu, J. Lee, J. Lin and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics", *J. Sound Vib.* **289** (2006), no. 4–5, pp. 1066–1090.

[41] A. Soualhi, H. Razik, G. Clerc and D. D. Doan, "Prognosis of bearing failures using hidden Markov models and the adaptive neuro-fuzzy inference system", *IEEE Trans. Ind. Electron.* **61** (2013), no. 6, pp. 2864–2874.

[42] W. Gousseau, J. Antoni, F. Girardin and J. Griffaton, "Analysis of the Rolling Element Bearing data set of the Center for Intelligent Maintenance Systems of the University of Cincinnati", Conference paper: CM2016, 2016. Online at https://hal.science/hal-01715193/document.

[43] NASA, *Prognostics Center of Excellence Data Set Repository*, April 28, 2025. Online at https://www.nasa.gov/intelligent-systems-division/discovery-and-systems-health/pcoe/pcoe-data-set-repository/ (accessed on November 4, 2025).

[44] J. K. Kimotho and W. Sextro, "An approach for feature extraction and selection from non-trending data for machinery prognosis", in *Proceedings of the European Conference of the PHM Society 2014, Vol. 2, No. 1*, PHM Society, 2014.

[45] A. Ben Yagoub and R. Ziani, *Supplementary information to "Bearing fault prognosis based on Cyclical Remaining Useful Life (CRUL)",* Comptes Rendus. Mécanique*, 2025*, V1, 2025.