

Article

An Unsupervised Data-Driven Framework for Bearing Failure Prognosis via Health Stage Clustering and Artificial Neural Network-Based Remaining Useful Life Estimation

Charafeddine Khamoudj ¹ , Fatima Benbouzid-Si Tayeb ¹, Karima Benatchba ¹ and Mohamed Benbouzid ^{2,3,*} 

¹ Laboratoire des Méthodes de Conception de Systèmes (LMCS), École Nationale Supérieure d'Informatique (ESI), BP68M, OuedSmar, Algiers 16270, Algeria; c_khamoudj@esi.dz (C.K.); f_sitayeb@esi.dz (F.B.-S.T.); k_benatchba@esi.dz (K.B.)

² Institut de Recherche Dupuy de Lôme (UMR CNRS 6027), University of Brest, 29238 Brest, France

³ Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China

* Correspondence: mohamed.benbouzid@univ-brest.fr; Tel.: +86-332-9801-8007

Abstract

Reliable bearing-failure prognosis in induction machines remains a critical research challenge, as it directly impacts system availability, maintenance efficiency, and overall operational safety. To address this challenge, it is essential to develop an online prognostic system capable of continuously assessing bearing health and predicting future failures in real time. This paper proposes a novel unsupervised data-driven prognostic framework for induction machine bearings that integrates advanced signal processing techniques for the preprocessing step, data clustering to construct bearing health stage (HS), artificial neural network (ANN) forecasting using a designed health indicator (HI) based on the latest historical observations, and a fine-tuning model to improve the estimation of remaining useful life (RUL) for induction machine bearings using vibration and temperature signals provided by the PRONOSTIA and NASA-IMS experimentation platform. The results show that the proposed approach is an effective way for bearing RUL estimation.

Keywords: bearings failure prognosis; induction machines; prognostics and health management; health stages; health indicators; artificial neural networks; data clustering; remaining useful life; unsupervised data-driven approach

1. Introduction

Failure prognosis is an advanced approach in condition monitoring of industrial systems that aims to enhance reliability, performance, and safety by integrating predictive analytics into maintenance strategies. Industrial maintenance has traditionally adhered to either corrective or preventive models. In the corrective model, components are repaired after failure, while in the preventive model, they are replaced or serviced at scheduled intervals. The corrective approach often relies on hardware redundancy for critical components [1], resulting in increased costs and excessive stock requirements. The preventive model, particularly time-based maintenance (TBM), schedules actions. The preventive maintenance model is typically categorized into two main strategies: time-based maintenance (TBM) and condition-based maintenance (CBM). TBM schedules maintenance actions at fixed intervals, regardless of the component's actual condition. This approach can lead to inefficiencies such as over-maintenance, resulting in unnecessary costs, or under-maintenance, which can cause unexpected downtime if actions are not appropriately



Academic Editors: Tibor Krenicky, Ján Pitel' and Kamil Židek

Received: 31 January 2026

Revised: 24 February 2026

Accepted: 27 February 2026

Published: 4 March 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

timed [2]. On the other hand, CBM is a more advanced preventive maintenance model that addresses these challenges by focusing on the actual condition of the components. Rather than relying on fixed schedules, CBM determines when maintenance is necessary based on real-time data collected through sensor signal analysis. By diagnosing the state of components, CBM allows targeted interventions, reducing unnecessary costs associated with over-maintenance and preventing costly downtime from under-maintenance. This model aims to detect failures before significant damage occurs, improving both the efficiency and reliability of the maintenance process [3].

Recently, predictive maintenance has emerged as an intelligent strategy that relies on the real-time condition of equipment. In contrast to preventive maintenance, which schedules regular maintenance interventions regardless of the equipment's condition, predictive maintenance uses the prognostics and health management (PHM) process to schedule maintenance interventions [4,5]. The PHM process involves the integration of various technologies, such as sensors, data analytics, machine learning (ML), and advanced algorithms based on artificial intelligence (AI), to assess the current health state and to predict its future behavior [6]. Concisely, the system health state is continuously (or periodically) observed by analyzing signals collected from embedded sensors (or inspection information), such as vibration, temperature, acoustic, electric, magnetic field, and other signals. The measured signals provide valuable data about the condition of critical components that serve as input for the prognostic algorithms to predict failures at an early stage by estimating the component's time before failure, known as the remaining useful life (RUL) [7]. This proactive approach facilitates early detection of potential failures, reducing downtime, optimizing maintenance schedules, and enhancing the operational efficiency and lifespan of equipment. The development of effective PHM approaches is a significant research field. Indeed, PHM gains are heavily reliant on decision making based on prognosis information, a process known as post-prognostic decision (PPD) [8].

Significant prognostics approaches are designed by an effective health indicator (HI) and a reliable prognostic model because these two factors are key aspects for accurate RUL estimation [9]. Building an effective HI that can accurately describe the degradation process is a prerequisite to designing a prediction model. Health indicators are constructed by selecting the appropriate characteristics and critical information from measured signals to train the prediction model [10,11].

Prognostics approaches can be categorized into two main types: model-based and data-driven approaches. Model-based methods involve building a specific mathematical model for the critical components. This approach is based on a priori knowledge of the system's normal operating mode and the different failure types. It uses fault indicators (residuals), which reflect the presence of a malfunction. Each residual is linked to a failure type; this configuration forms a matrix known as the fault signature [12,13]. Such models require extensive experimentation and validation. While model-based methods can be reliable when accurate models are developed, their performance is highly dependent on the quality of the model. This introduces significant challenges in complex systems, where accurate mathematical modeling may not be feasible. Data-driven approaches are more efficient in this context, as they use AI and ML models to characterize the degradation behavior of monitored components without requiring explicit knowledge of their physical behavior. These approaches can be classified into two main groups: cumulative degradation and direct remaining useful life (RUL) mapping [7].

In cumulative degradation prognostic approaches, RUL estimation is based on the history of data degradation, applying various regression models, such as time series analysis [14–22], and ML techniques [23–27]. On the other hand, direct RUL mapping

approaches focus on constructing a degradation evolution over time, based on historical run-to-failure experiments for training [28,29].

Vibration signals remain the most widely used data for monitoring rotating mechanical equipment, particularly for detecting bearing failures, as they provide information that reflects overall structural performance [30]. However, these signals are often nonstationary, nonlinear, and noisy [31], which makes it challenging to apply time series analysis for forecasting future vibration values and predicting failures. Forecasting models based on time series tend to produce significant errors when degradation trends exhibit nonstationary, nonlinear behavior or include noise [32].

Several previous statistical studies have highlighted that bearings are the most failure-prone components in induction machines [33,34]. This sensitive component should, therefore, be specifically monitored to avoid disturbance and loss caused by unplanned shutdowns. Constructing and validating a mathematical model to describe the operating behavior of induction machines proves inherently difficult, prompting researchers to rely increasingly on data-driven approaches as an alternative solution in the literature to address this issue. These approaches leverage ML techniques to identify patterns of bearing failures over time. Therefore, this paper proposes a novel data-driven approach for predicting bearing failures in induction machines, using vibration and temperature signals for direct RUL mapping.

Although numerous approaches have been proposed for bearing failure prognosis, the proposed approach introduces several fundamental aspects that enhance its novelty and set it apart from prior approaches. Unlike the semi-supervised models of Wang et al. [35] and Zhu et al. [36] that require manual labeling of degradation states, this paper proposes an unsupervised CMO-based clustering approach that enables data-driven classification of normal and faulty operating modes by constructing a representative degradation model directly from training data, thus eliminating the need for prior failure assumptions. This constitutes a key improvement over the approach of Chelmiah et al. [37], which depends on predefined failure models. Furthermore, the proposed dual-ANN architecture with an adjustment-ANN represents a novel departure from the single-model frameworks used in most existing research [38–40]. Unlike the deep learning implementations that process the entire historical data sequence, as in [40], the presented approach specifically isolates and utilizes only data from the faulty operating mode, leading to more focused and accurate predictions. This selective training strategy addresses a critical limitation in previous approaches that fail to differentiate between normal operation and early degradation phases. The proposed methodology distinguishes itself from prior work through the integration of multiple complementary techniques: unsupervised health stage construction that automatically identifies the faulty operating mode, which refines data to focus exclusively on the most informative degradation phases in the training phase, and a complementary fine-tuning mechanism that adapts predictions based on observed error patterns in the test phase. This orchestration of methodologies represents a significant advancement over existing methods that typically implement these techniques in isolation. The proposed approach starts with collecting and preprocessing run-to-failure vibration and temperature signals to train a forecast-ANN model. This model uses a designed health indicator (HI) based on the latest historical observations to forecast future values recursively until detecting failure. To enhance prediction performance, two techniques are used: HS construction and fine-tuning. The first one uses data clustering to extract just faulty mode signals for the forecast-ANN training. The fine-tuning technique consists of forecasting results' improvement by defining an adjustment-ANN, which is applied during the test step.

The key contributions and novel aspects of the proposed approach can be summarized as follows:

- A designed health indicator (HI): It consists of a forecast-ANN trained by a window of historical data to forecast future values recursively until detecting failure.
- An unsupervised health-stage (HS) construction: A CMO-based unsupervised clustering is introduced to automatically refine the faulty operating mode data, focusing exclusively on the most informative degradation phases in the training phase, which enhances the quality of the degradation patterns.
- A fine-tuning mechanism: It consists of designing an adjustment-ANN trained by the application of the forecast-ANN during the test step by calculating the model error.
- An integrated workflow rather than isolated techniques: The proposed approach orchestrates unsupervised health-stage construction during training with a complementary fine-tuning via the adjustment-ANN, enabling an adaptive forecasting process.

The proposed testing approach is conducted using run-to-failure experimental data provided by the PRONOSTIA platform [41] and the NASA-IMS datasets for bearing [20,42].

The rest of the paper is organized as follows: Section 2 introduces the state of the art of component RUL estimation. Section 3 is dedicated to the proposed data-driven ANN-based approach. The performance analysis and experimental results are given in Section 4. Finally, Section 5 concludes this work.

2. Critical Review on Advances in Component RUL Estimation

Predicting failures in electromechanical systems is essential for ensuring the availability of production tools and minimizing losses caused by unexpected downtime. One of the most critical components in these systems is the electrical machine, particularly the induction machine, which is widely used in industrial applications. However, its operation introduces electrical and mechanical constraints, making it susceptible to various internal faults that must be detected and predicted at an early stage.

RUL estimation for components involves calculating their degradation over time until failure, based on historical health data. This process relies on failure prognostics, which support the planning of necessary maintenance actions to prevent unexpected breakdowns. In the literature, failure prognostic approaches are generally classified into two main categories: model-based and data-driven approaches, depending on the type of failure prediction model employed.

The model-based approach involves constructing mathematical models of the monitored system or its critical components to describe failures and their progression. This method requires extensive knowledge of the physics underlying the system or components. The models use measurable indicators, such as corrosion, current insulation, and friction, to represent wear and degradation. Failure prognosis and RUL estimation are performed by comparing the modeled indicators with real-time measurements, often using dynamic ordinary differential equations (ODEs) or partial differential equations (PDEs) [43].

For instance, Ref. [3] developed a mathematical model for wind turbines, collecting data under both normal and faulty operating conditions. This data is used for health state classification, and RUL estimation is achieved by measuring Euclidean distances among clusters and determining the degradation velocity. Ref. [25] proposed a model-based approach using a Kalman smoother to estimate RUL for wind turbine drivetrain. The method uses a Kalman filter to measure HI and calculate the varying degradation trend, specifically the rate of crack growth. RUL is then calculated using Paris's law as a fault propagation model. Ref. [44] used Weibull distribution function model to propose a three-parameter WPHM approach (Weibull-distribution Proportional Hazards Model) for bearing failures prediction, using the Pearson correlation coefficient to find the covariates between parameters that can reflect the bearing state as HI. Ref. [9] enhanced the WPHM approach using four prognostic metrics—monotonicity, robustness, trendability, and consistency—within

weight coefficients for each of them to construct an HI to improve bearing RUL estimation. Ref. [12] proposed a model-based approach for induction machine prognosis that relies on magnetic flux and angular speed indicators, using a state observer. The angular speed of the rotor is estimated by calculating the proportional–integral (PI) control from the observer.

Model-based prognostic approaches are effective when the monitored system is relatively simple, making it feasible to develop an accurate mathematical model. However, this is rarely the case in modern industrial systems [45], which are often complex and operate under varying conditions. Additionally, failures can arise due to external factors such as climatic conditions, including high temperatures and humidity, which can accelerate wear. Even under the same conditions used to construct the model, component degradation may evolve in a nonlinear or exponential manner, as seen with the nonstationary degradation of induction machine bearings. These challenges make it difficult to build accurate degradation models for such systems, which gradually shifts modern RUL studies from model-based to data-driven approaches [46].

Data-driven prognostic approaches, on the other hand, rely on empirical models to characterize degradation phenomena based on monitoring data. Unlike model-based methods, they do not require explicit knowledge of the physical behavior of the monitored component, making them suitable for systems where constructing an accurate mathematical model is impractical due to physical complexity. Empirical models can be broadly classified into two categories: cumulative degradation and direct remaining useful life (RUL) mapping [7].

RUL estimation in cumulative degradation prognostic approaches is based on data degradation history applying different regression models such as time series analysis, which consists of studying the evolution of a phenomenon over time to provide a forecast model. Ref. [14] used the double exponential smoothing on bearing vibration signals to calculate the performance degradation time. Ref. [19] proposed an ARIMA (auto-regressive integrated moving average) forecast model for mechanical equipment failure prediction using nonstationary vibration signals. After identifying the possible p and q values based on the autocorrelation, the proposed recursive method is composed of three main steps: (1) applying first-order differencing to eliminate stationarity, (2) applying the ARMA(p,q) model, and (3) predicting future values by applying the current configuration. Ref. [17] used the AR (auto-regressive) model for gearbox failure prediction. The proposed method uses empirical mode ensemble decomposition (EEMD) as vibration signal preprocessing to deal with nonlinearity and nonstationarity, then applies the AR(p) forecast model by determining the parameter p . Ref. [18] proposed an auto-regressive moving average (ARMA) forecast model for electric motor failure prediction using electrical voltage and current signals. Ref. [15] proposed an approach based on the Holt–Winter method to deal with multiseasonality time series by estimating and recursively eliminating the seasonality that has the largest period. Ref. [16] used Holt–Winters for bearing performance degradation prediction. The proposed approach uses the preprocessed vibration signals to extract the bearing health state by applying the KJADE algorithm, which represents a combination of kernel function for feature vector extraction and JADE (Joint Approximate Diagonalization of Eigen-matrices) to extract features that reflect the health state. Finally, Holt–Winters exponential smoothing is applied to predict the trend increment, which reflects the progression of bearing degradation.

Due to their deterministic nature, classical time series analysis methods produce identical results when applied to the same time series, offering no inherent mechanism for performance improvement. To overcome this limitation, researchers have developed hybrid approaches that combine these traditional classical methods with AI-based learning techniques. This hybridization leverages the solid mathematical basis of time series

analysis while benefiting from the adaptive learning capabilities of AI algorithms, enabling continuous performance improvement. Among these methods, we can cite the works of [47], which applied expert systems for automatic ARIMA modeling, and Ref. [48], which applied expert systems for the automatic SARIMA modeling. The two proposed methods are based on data transformation into a stationary series in a preprocessing phase using the first-order differencing method; then, the automatic modeling phase, based on rules, makes deductions. Decision trees (DTs) are used in failure prediction as a generalization of AR models, where each leaf node represents an instance of AR(p), which allows the testing of models with different parameters, p. Ref. [20] used DT for bearing failure detection using vibration signals as a time series. Ref. [21] proposed an approach for cooling fan failure prediction using nonstationary rotation speed signals by defining a minimum speed. If the rotation speed is lower than the minimum speed, it loses the cooling effect, which is the faulty mode. The proposed approach hybridizes an ARIMA(p,d,q) model with a neural network using backpropagation. First, the time series is transformed into a stationary process through iterative first-order differencing, applying the operation d times until stationarity is achieved. Next, the stationary data is modeled using an ARMA(p,q) forecast model to generate predictions. Finally, these predictions are refined using an ANN with backpropagation, where both the actual fan speed and the ARMA-forecasted speed serve as inputs. Ref. [22] proposed an approach for predicting electrical insulation degradation based on the identification of the stress factor, which corresponds to the current leaks' increase as HI. The used data are the stator's current leaks signals, where the signals are measured by high-sensitive current transformers (HSCTs), and the electrical voltage using voltage transformers (VT). The obtained signal time series is nonstationary, so a nonparametric approach is used to eliminate trends based on the bilateral moving averages. The authors proposed two prediction models: an ARMA(p,q) model and an ANN model. To select the model, AIC (Akaike information criterion) and BIC (Bayesian information criterion) were used to compute the number of parameters for each model (the number of neurons in the hidden layer for the ANN model and p+q for the ARMA model), where the selected model was the one with the smallest number of parameters. The authors proposed two ANNs for fault identification and failure prediction. For failure prediction, the authors implemented a neural network trained with the RPROP (Resilient Backpropagation) algorithm. This time series forecasting model uses a sliding window approach, where historical observations up to time t-1 ($Z_1, Z_2, \dots, Z_{t-1}$) serve as input, while Z_t represents the output observation. This structure enables the building of a degradation model with a step forecast. The first proposed ANN consists of providing a supervised classification of operating mode (Good, Moisture, Oil/ Dirt and Thermal) to identify the stress type that caused the degradation; however, it is not applied in the prediction process as a data filter.

Despite the solid mathematical basis of the time series models, which gives an efficient calculation, the forecast model leads to large errors when the degradation history is nonstationary, nonlinear, and includes noise [32].

To overcome the limitations of time series models, researchers have focused on alternative solutions. Many have turned to machine learning (ML)-based prognostic approaches. These approaches utilize health indicators for monitored equipment. Such indicators include more health state information and better reflect the degradation trend, which enables more effective training of prediction models.

Ref. [39] used a stacked autoencoder (SAE) structure to fuse the selected features to construct HI, using an LSTM model for RUL estimation. Ref. [23] proposed a data-driven approach to estimate bearing RUL by using two concepts; the first one applies a deterioration exponential function according to the health indicator (HI), and the second one consists of applying a Gaussian mixture model (GMM) as a data clustering method for

class labeling (healthy, deteriorated, critical) to obtain the health stage (HS). Both HI and HS indicators are used in a knowledge-driven process to estimate RUL. Ref. [24] estimated bearing RUL by applying a relevance vector machine (RVM) on the extracted bearing degradation features as a regression method until the failure threshold. Ref. [37] proposed two supervised ML algorithms to estimate bearing RUL: SVM and k-NN, using short-time Fourier transform (STFT) on vibration signals' time-frequency domain to indicate the health condition to construct HI. The results obtained by the two proposed models using the same HI are close, which reveals the importance and impact of constructing an HI. Ref. [26] combined support vector machine (SVM), random forest regression (RFR) and Gaussian process regression (GPR) in a hyper-heuristic algorithm called Sparrow Search Algorithm (SSA) to estimate bearing RUL, to derive maximum advantage from each method. The outcomes indicate the overall performance of SVM in dealing with complex data distributions, regularization, high-dimensional data and hyperparameter sensitivity. In general, the primary limitation of ML approaches lies in their need for extensive learning sessions across various conditions. Recently, Ref. [27] proposed a data-driven approach for diagnosis and prognosis of aeronautical bearing using collaborative selection-based incremental deep transfer learning (CSIDTL) to overcome lack of the pretrained ML patterns. This approach is justified because operating conditions in the target domain often differ significantly from those present in the training dataset. The CSIDTL technique enables ML models to adapt to real-world operational conditions, while simultaneously expanding the knowledge base through the incorporation of newly transferred learning patterns. The proposed approach is enhanced by using long short-term memory (LSTM) adaptive learning rules to overcome data complexity and data change problems. Ref. [49] made a comparative study between LSTM, BiLSTM (Bidirectional-LSTM), GRUs (gated recurrent units), and RF (random forest) predictive models for wind turbine high-speed shaft bearings, using a data aggregation in a time window as a HI. This comparative evaluation of the models' performance revealed that LSTM and BiLSTM outperformed GRU and RF in RUL estimation. An important contribution proposed by [5] combined several cutting-edge techniques. The proposed approach consists of bearing RUL estimation using both HI and HS by solving regression and classification problems, respectively. The solving approach is based on two LSTM models: HI estimator (regression) and HS predictor (classification), involving a transfer learning during the training procedure to fine-tune the RUL estimation. The HI construction based on wavelet transform was used by [38], applying continuous wavelet transform (CWT) on time-domain vibration signals to extract degradation characteristics of bearings and training a CNN prediction model based on global attention mechanism (GAM-CNN) to estimate RUL. Ref. [50] proposed an LSTM model for both failure type detection and RUL estimation using bearing feature fusing as input.

Within the second category of data-driven prognostics, known as direct RUL mapping prognostics, empirical models play a crucial role in predicting RUL. Unlike other methods, this approach does not require the estimation of health status; instead, it involves constructing a degradation evolution model over time using data from historical run-to-failure experiments for training. Limited contributions to the literature exist in this area, and a few notable contributions include the work of [28], which proposed an ANN, which provides the bearing life percentage and then calculates RUL, using run-to-failure vibration data for training after a measured signal fitting step. The inputs of the proposed ANN consist of bearing age and degradation velocity measurements as inputs, where the output represents the failure rate, which indicates the failure probability at a given time and is calculated by a Weibull distribution failure rate function. Ref. [35] proposed a recurrent-CNN using bearing vibration signals' time series as the input to build an RUL estimation model based

on modeling the temporal dependencies of different degradation states, which enables the model to memorize degradation information over time and make connections for each output, not only with the input of the current layer but also with the previous stored state that maintains a memory of all past inputs. Ref. [36] proposed a bearing RUL estimation based on a multiscale convolutional neural network (MSCNN), where signal wavelet transform (WT) technique is applied to construct HI and train the proposed MSCNN as a health degradation regression problem to build a prediction model. Ref. [29] proposed a bearing's direct RUL mapping estimation using an ANN forecast model. The training phase is carried out on test set data, which are filtered by applying a k-means clustering to construct health states, when only features from degradation patterns are extracted to build the ANN model. Ref. [51] proposed a degradation assessment approach for wind turbine bearing using extreme learning machines (ELMs) with Recurrent Expansion (REX) algorithms and Bayesian optimization to adjust the model's parameters. Authors collected 50 days of run-to-failure data to construct bearing degradation patterns and to describe health states. The collected signals were preprocessed by denoising and aggregating on timespan, and then applying linear regression to uncover underlying trends. Ref. [40] applied kernel smoothing density (KS-density) on vibration signals to construct an HI. This extracted degradation data was used as input to train a BiLSTM model for assessing the bearing health and estimating the bearing RUL.

Figure 1 summarizes the different approaches for failure prognosis. The model-based methods allow the proposal of empirical models of system behavior and failure patterns. It replaces mathematical models that require a large amount of knowledge on the system's physics. Data-driven approaches for RUL estimation have gained significant importance in recent years, according to technological advancements in signal processing techniques and the evolution of sensor and instrumentation technologies. These developments have enabled more effective assessments of health conditions and predictions of failures [52].

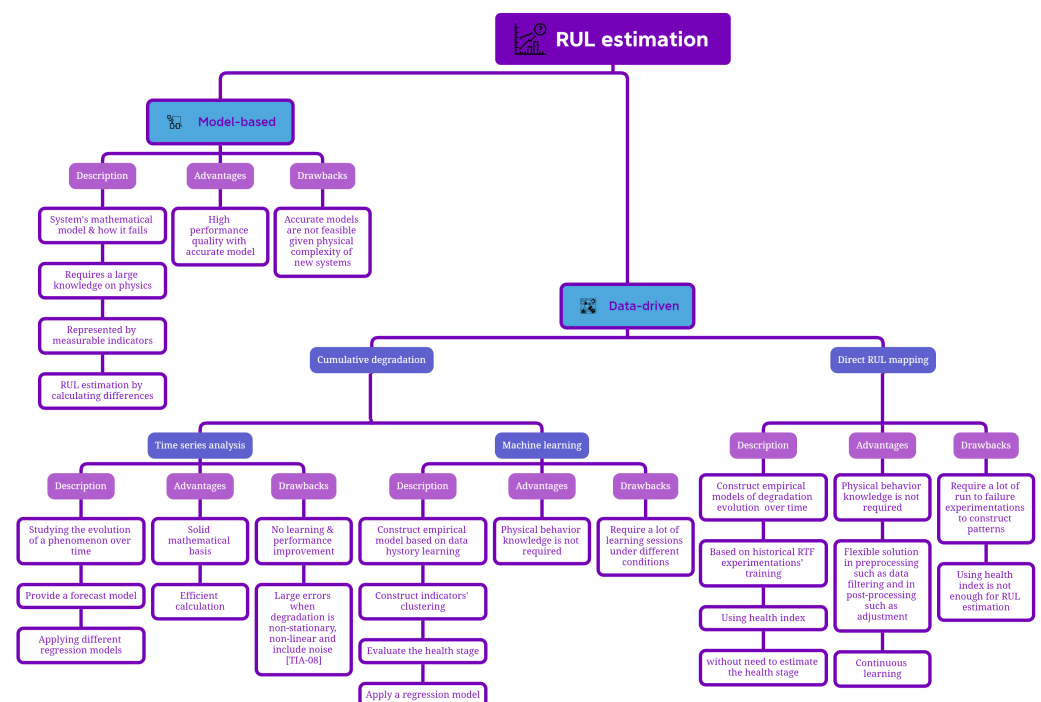


Figure 1. Summary of failure prognosis approaches.

Among data-driven contributions in the literature, time series forecast-based models often produce significant errors when applied to nonstationary, nonlinear, and noisy degradation trends [32]. This limitation is particularly relevant for induction machine

bearings, which typically exhibit such complex degradation patterns. Furthermore, developing ML-based empirical models presents additional challenges, requiring extensive training sessions across various operating conditions. Even when an accurate model is successfully constructed, its performance can be decreased significantly when deployed in environments that differ from the training conditions, such as variations in environmental factors that can accelerate bearing wear.

The evolution of data-driven approaches has been particularly significant in recent years, extending beyond bearing failure prediction to various engineering domains with similar challenges. For instance, in civil engineering, data-driven methods have significantly advanced the prediction of wind-induced responses in slender structures [53]. Their comprehensive review highlights several important developments applicable to this domain: the transition from physics-based to hybrid and purely data-driven models, the importance of appropriate feature selection in nonstationary environments, the evolution from traditional ML to deep learning architectures for complex pattern recognition, and the critical challenge of model generalization across different operational conditions. These insights are particularly relevant to bearing failure prognosis, as both domains deal with nonstationary signals, complex system dynamics, and the need to make reliable predictions with limited training data.

Modern data-driven approaches can be further categorized based on their learning approach: supervised, semi-supervised, and unsupervised learning. Supervised learning approaches require labeled training data, where each input is associated with the corresponding output. While effective when historical failure data is abundant, these methods struggle with real-world industrial scenarios where complete run-to-failure datasets are scarce [5]. Semi-supervised approaches attempt to mitigate this limitation by utilizing both labeled and unlabeled data, often employing techniques such as transfer learning [27], self-supervised pretraining, or hybrid models integrating both physics-based knowledge and data-driven learning [54]. Unsupervised approaches, like the one proposed in this paper, represent the most flexible solution, requiring no explicit labels and instead discovering patterns and features directly from the data. This characteristic makes unsupervised methods particularly valuable in practical industrial settings where labeled failure data is often unavailable or insufficient.

Tables 1 and 2 summarize some cited papers according to the used prediction model, HI construction, the used dataset and advantages.

Based on the analysis of existing approaches summarized in Table 1, we can identify specific limitations of each prediction model category when applied to bearing failure prediction: statistical time series models generate substantial errors on nonstationary data (the typical nature of bearing vibration signals in induction machines), model-based approaches require strong a priori knowledge of physical behaviors, which is difficult to obtain for complex systems, and conventional ML methods tend to deteriorate when deployed under domain shift. In light of these findings, this analysis identifies deep-learning-based approaches with direct RUL mapping as the most promising solutions for unsupervised bearing failure prognosis. While architectures such as CNNs, RCNNs, MSCNNs, LSTMs, and BiLSTMs have shown strong fault-feature extraction capability, industrial applications often suffer from limited dataset availability [11]. This constraint, combined with the specific requirements of bearing prognostics, favors the use of ANNs, which require comparatively less training data while remaining effective for time series forecasting: they capture nonstationary degradation patterns, adapt well to sequence prediction, support continuous learning without full retraining, deliver strong performance with relatively small datasets, and seamlessly integrate heterogeneous inputs (e.g., vibration and temperature) without extensive feature engineering.

Table 1. Summary of the cited papers in failure prognosis.

Paper	Category	Prediction Model	HI Construction	Advantages	Dataset
[21]	Time series analysis	Statistical: ARIMA	-	Fine-tuning technique by an ANN model to adjust ARIMA RUL estimation.	DS-Self
[25]	Model-based	Paris's law	Kalman filter.	Hybridizing with data-driven approach for RUL estimation based on Kalman smoother.	DS-1
[35]	Data-driven direct RUL mapping	Deep learning: Recurrent-CNN	temporal dependencies of different degradation states.	Recurrent-CNN can learn short-term and long-term dependencies from time series data.	DS-5
[36]	Data-Driven Direct RUL mapping	Deep learning: Multiscale-CNN	wavelet transform (WT).	The multiscale layer allows the maintenance of the global and local features to enhance the network capacity.	DS-5
[22]	Time series analysis	Statistical: ARMA & ANN	Identification of the stress factor.	Using AIC and BIC for model selection according ARMA and ANN model parameters.	DS-Self
[44]	Model-based	Three parameters WPHM	Pearson correlation coefficient.	Finding the covariates between parameters can reflect the bearing state.	DS-2
[29]	Data-driven direct RUL mapping	Deep learning: ANN	A cumulative degradation function for monotonicity and trendability.	Using k-means clustering to extract degradation patterns as HS.	DS-5
[9]	Model-based	WPHM	Monotonicity, robustness, trendability, and consistency prognostic metrics.	Using weight coefficient for each prognostic metric.	DS-3
[39]	Data-driven ML	Deep learning: CNN-LSTM	Stacked autoencoder (SAE).	The SAE structure allows feature extraction by fusing and reducing dimension.	DS-2, DS-4
[23]	Data-driven ML	Deep learning: CNN-LSTM	Deterioration exponential function to construct HI and Gaussian mixture model to construct HS.	Using both HI and HS to represent health state, and using transfer learning for further meaningful learning.	DS-2
[5]	Data-driven ML	Regression-ML/Classification-ML: LSTM	An exponential function that models the cumulative degradation.	Using both HI & HS and involving a transfer learning during the training procedure to fine-tune the RUL estimation.	DS-5
[37]	Data-driven ML	Supervised-ML: SVM k-NN	Short-time Fourier transform (STFT) and Hilbert marginal spectrum (HMS).	Using STFT and HMS reduces the computational complexity for training.	DS-5
[26]	Data-driven ML	Regression-ML: SVM random forest Gaussian process	Increasing degradation based on monotonicity.	Using a hyper-heuristic algorithm SSA to discover the most suitable hyperparameters for the proposed SVM, RF and GP ML.	DS-6
[40]	Data-driven direct RUL mapping	Deep learning: BiLSTM	Kernel smoothing density.	BiLSTM is a good model to train from extensive data, also data fitting is carried out using Weibull failure rate function.	DS-2

Table 2. Experimental datasets used in the cited papers.

Code	Experimental Platform	Reference
DS-1	Green Power Monitoring Systems, LLC, VT 05753, USA: A real 2 MW wind turbine (Sazelgn)	-
DS-2	IMS, University of Cincinnati. NASA AMES Prognostics Data Repository	[20]
DS-3	XJTU-SY Rolling element bearing accelerated life test datasets	[55]
DS-4	Bearing DataCenter Case Western Reserve University	[56]
DS-5	PRONOSTIA: An Experimental Platform for Bearings Accelerated Life Test	[41]
DS-Self	Self-experimental platform	-

3. Proposed Data-Driven ANN-Based Approach

In this paper, for comprehensive failure prognosis, the proposed approach integrates both health indicator (HI) and health stage (HS). Following [57], HS captures the global degradation trend, while HI reflects local variations in degradation. The selection of the prediction model takes into account data availability, computational constraints, and the need for transfer learning to adapt to changing operational conditions [5]. Within the PHM framework, the methodology targets bearing RUL estimation using vibration and temperature signals, adopting an unsupervised data-driven strategy with direct RUL mapping (Figure 2). Its key contribution lies in combining two complementary elements: (1) HS construction through CMO-based clustering, which discriminates between normal and faulty modes, and (2) HI development via a dual-ANN architecture capable of accurate forecasting without reliance on large datasets. Unlike conventional approaches that train on all available data or rely on complex deep learning models, the proposed method introduces three main innovations: (i) unsupervised clustering to derive meaningful health stages, (ii) training restricted to faulty-condition data to enhance prediction accuracy, and (iii) a dual-ANN structure, consisting of a forecast-ANN to model degradation trajectories and an adjustment-ANN for fine-tuning under varying operating conditions.

The proposed approach follows a four-stage pipeline. First, measured signals are collected from a run-to-failure experimental platform, followed by a preprocessing step to prepare the signals for time series modeling. Second, an HS is constructed using CMO-based data clustering specifically adapted to bearing degradation, keeping only faulty-mode observations to train the proposed ANN. In the third stage, the ANN is used to forecast the evolution of bearing degradation over time, where the constructed HI, represented by the n previous observations, serves as input to predict the next observation $n + 1$. Finally, during the testing phase, a fine-tuning technique is applied through an adjustment-ANN, which significantly improves prediction accuracy. This dual-ANN architecture constitutes a key advantage of the proposed approach, enabling the model to adapt to variations in operating conditions while preserving high predictive performance. The subsequent sections present an in-depth analysis of each stage.

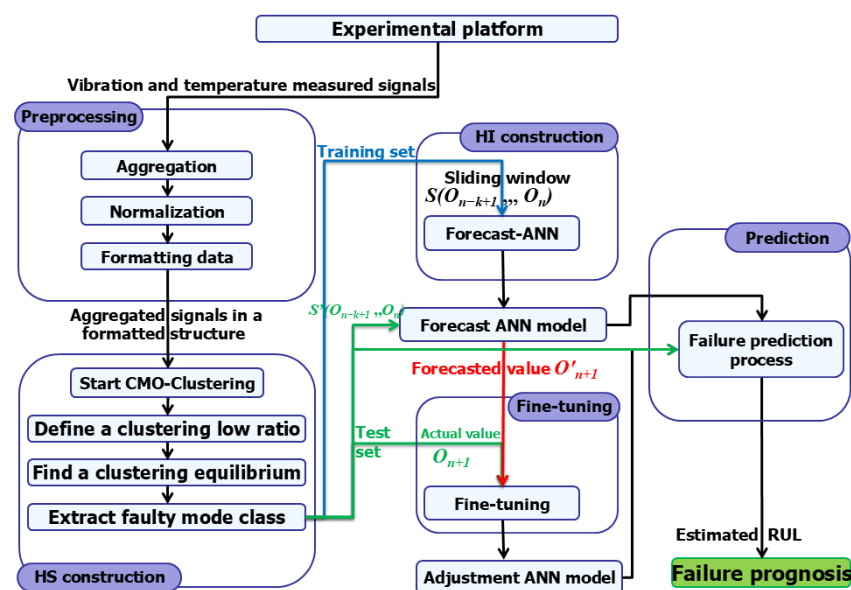


Figure 2. The proposed failure prognosis approach.

Figure 2 illustrates the end-to-end workflow of the proposed unsupervised data-driven approach for bearing failure prognosis. Vibration and temperature signals from run-to-failure experiments are first aggregated and normalized to form structured time series

observations. A CMO-based clustering constructs the health stages (HS) and automatically isolates faulty-mode data to ensure informative training. Using a sliding window of the latest observations as the health indicator (HI), the forecast-ANN recursively predicts the next observation to model the degradation trajectory until failure detection. During testing, an adjustment-ANN refines the forecast, which improves robustness to operating-condition variability. The pipeline outputs an estimation of bearing RUL.

3.1. Data Acquisition and Preprocessing

The data preprocessing step consists of aggregating and normalizing measured signals, and finally making them into structured data. Data normalization is carried out to put the different data (time, vibration and temperature) in the same range and then use vector coefficients to ensure fair treatment of the data. Data aggregation is carried out to eliminate the measured signals noise, thus reducing the problem complexity [58].

Firstly, the vibration and temperature signals are aggregated over a time interval, then represented by unique data computed by one of the aggregation functions: mean, variance, or standard deviation, selected as a parameter in the aggregation procedure. The main constraint is the definition of the time interval, which should be small enough not to lose information and long enough to reduce algorithm complexity and eliminate noise.

Data normalization aims to put the different data (vibration and temperature) in the same range to be able to later correctly give a relative impact to a given vector by using corresponding coefficients. This is important to ensure that all handled data has the same influence or to give more impact by increasing the corresponding coefficient.

The normalization process applies specific transformations to the vibration and temperature vectors according to Equations (1) and (2).

$$\vec{v}_{norm} = \vec{v} \times \bar{c} \times \alpha_v \quad (1)$$

where $\vec{v}_{norm}$ is the normalized vibration vector, $\vec{v}$ is the original vibration vector, $\bar{c}$ is the mean of the temperature vector, and α_v is the vibration normalization coefficient.

$$\vec{c}_{norm} = \vec{c} \times \bar{v} \times \alpha_c \quad (2)$$

where $\vec{c}_{norm}$ is the normalized temperature vector, $\vec{c}$ is the original temperature vector, $\bar{v}$ is the mean of the vibration vector, and α_c is the temperature normalization coefficient.

Finally, the data are formatted in a step to prepare aggregated and normalized time series data for use in the failure prediction process. In this step, a data structure containing observations $S = (O_1, O_2, \dots, O_i, \dots, O_n)$ is created, where each observation $O_i(t_i, v_i, c_i)$ contains the aggregated and normalized vibration and temperature signals v_i and c_i and the time average value of the corresponding aggregation interval t_i .

3.2. HS Construction

Run-to-failure experimentation platforms provide bearing-measured signals from the start of use through initial failure and into deterioration, when the operating mode is normal for a long time. Using the vibration signals, the bearing degradation trend is linear in normal operating mode; when a fault appears, it becomes nonstationary. To address the challenge of selecting meaningful signals that accurately represent bearing deterioration for ANN model training, Ref. [28] proposed using the Weibull distribution failure rate function to fit the measurements and then using them as inputs to an ANN model. Another interesting approach was used in [22] to filter the faulty mode by an ANN data classifier; however, it was not used to filter the training data for the proposed failure prognosis approach. The proposed HS construction approach distinctly differs from these

methods by automatically identifying and isolating faulty condition data specifically for training purposes.

Training is a central step in developing an effective ANN model, and, therefore, requires high-quality data that accurately represents the correct behavior to ensure optimal results. As the bearing degradation trend has two different behaviors depending on the operating mode, in the normal operating mode, the trend is stable, and it becomes non-stationary when a fault appears. Therefore, unlike conventional approaches that use all available data indiscriminately, a novel selective training strategy is proposed that fits the measured signals, keeping only those that describe the faulty operating mode. For that, after the signals preprocessing phase, the CMO-Clustering [59] is applied, which offers distinct advantages for this task, including its ability to handle overlapping transition regions between normal and faulty states, its robust initialization strategy that minimizes sensitivity to starting conditions, and its straightforward parameterization based on distance ratio principles.

CMO-Clustering is based on the attraction–repulsion mechanism and adaptive distance ratio tuning, which intensifies the attraction force between similar data while maintaining repulsion between dissimilar data. Its application for HS construction uses a low distance ratio, which increases the attraction force in order to obtain a strong class separation, generally giving two classes to represent the normal operating mode and the faulty one, which allows HS construction. Figure 3 represents an example of CMO-Clustering with a low distance ratio; here, the fitted data is the second class, which represents the faulty operating mode.

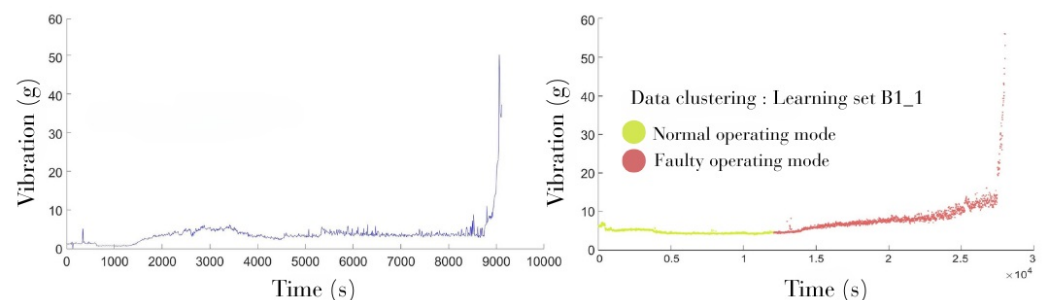


Figure 3. Example of low-ratio CMO-Clustering.

3.3. HI-Based Degradation Forecasting

To estimate the bearing RUL using time series $S(O_1, O_2, \dots, O_i, \dots, O_n)$, where $O_i(t_i, v_i, c_i)$ represents time, vibration, and temperature observation i , a forecast ANN is designed that provides a forecast model; therefore, to make a forecast for O_{n+1} , which represents the ANN output, the last k observations $S'(O_{n-k+1}, \dots, O_n)$ are used as input, where S' represents the health indicator (HI). After extensive experimentation with various architectures, the optimal configuration is selected based on prediction accuracy and computational efficiency. The architecture of the proposed forecast-ANN is as follows:

- Input: An observation window which consists of the last k time series' observations $S'(O_{n-k+1}, \dots, O_n)$ as HI. This creates an input layer with $3 \times k$ neurons (k observations $\times$ 3 features per observation: time, vibration, and temperature).
- Hidden layer: A single hidden layer with 5 neurons using the default hyperbolic tangent (tansig) activation function in Matlab.
- Output: The forecast observation O_{n+1} with 3 neurons (time, vibration, and temperature) and linear activation function.

The forecast-ANN was trained using the following parameters:

- Optimization algorithm: Levenberg–Marquardt backpropagation (trainlm).

- Network architecture: Standard feedforward neural network implemented with Matlab.
- Data split: 80% training, 20% testing, no validation set.
- Performance function: Mean squared error (MSE) by default.

This ANN design allows the model to forecast the next observation using a window of k last observations, which represents a forecast with a horizon $h = 1$ for the HI, as illustrated in Figure 4. An example of the training performance of the proposed forecast-ANN model is shown in Figure 5.

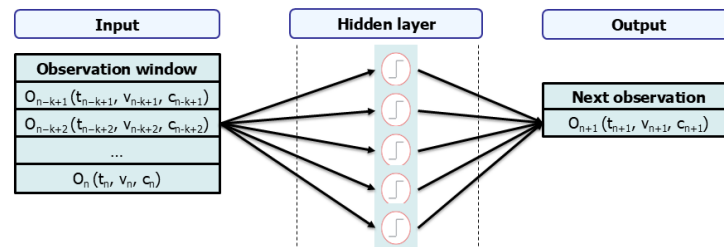


Figure 4. Structure of the proposed forecast-ANN.

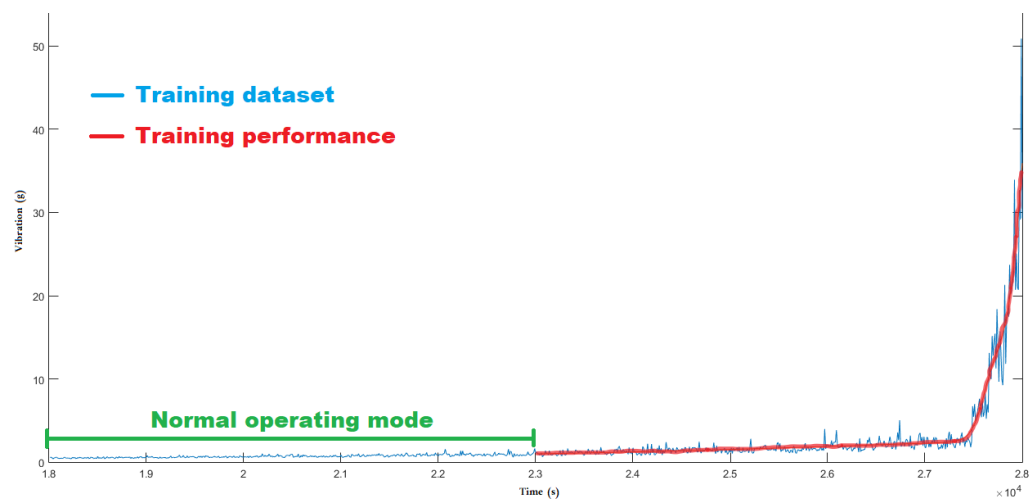


Figure 5. Example of the proposed forecast-ANN training performance.

3.4. Prediction Fine-Tuning

A fine-tuning mechanism is proposed to adapt the failure prediction process into different operating conditions and degradation trend variations to improve the failure prognosis performance. The proposed approach employs a second complementary neural network that specifically learns to correct prediction errors of the primary model dedicated for forecasting. It consists of applying an adjustment-ANN to calculate the forecasting errors in order to design an adjustment model (Figure 6).

This approach differs substantially from fine-tuning methods that typically just update the weights of the original model. Instead, the proposed adjustment-ANN creates a specialized error-correction mechanism that works in tandem with the primary forecast model in the test phase, which increases the adaptability to different operating conditions and where environmental and operational variations are common.

The architecture of the proposed adjustment-ANN is as follows:

- Input: There are two inputs, the actual observation O_i and the forecasted observation O'_i by the forecast-ANN model of the previous observation O_{i-1} . This dual-input design is crucial as it allows the model to directly learn the relationship between predicted and actual values, rather than simply attempting to improve the original pre-

diction. Each input consists of 3 features (time, vibration, and temperature), resulting in a total of 6 input neurons.

- Hidden layer: A single hidden layer with 5 neurons using the default hyperbolic tangent (tansig) activation function in Matlab. This architecture is defined to provide optimal error correction capabilities for the adjustment task.
- Output: Represents the variation between the two values O_i and O'_i . The output layer consists of 3 neurons (for time, vibration, and temperature error corrections) with linear activation functions.

The adjustment-ANN training parameters are as follows:

- Optimization algorithm: Levenberg–Marquardt backpropagation (trainlm).
- Network architecture: Standard feedforward neural network implemented with Matlab.
- Data split: 80% training, 20% testing, no validation set.
- Performance function: Mean squared error (MSE) by default.
- Input structure: Three-dimensional input representing time, vibration and temperature components from the observation period.

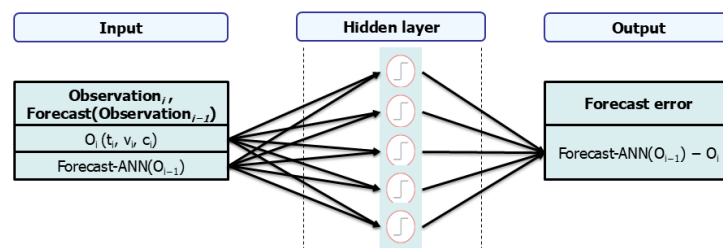


Figure 6. The proposed fine-tuning technique using an adjustment-ANN.

Figure 7 represents a training example for the proposed adjustment-ANN.



Figure 7. Training example for the proposed adjustment-ANN.

The proposed forecast-ANN model consists of forecasting the next observation, whose failure prognosis aim is to estimate the RUL. To do this, an iterative process is proposed which consists of applying the forecast-ANN model on each iteration with fine-tuning and inserting the forecasted values into the time series as a new observation until a stopping criterion is satisfied, which represents the bearing deterioration.

The stopping criterion of the iterative process, denoted as Condition (Current_Observation), is satisfied when the vibration value of the current observation (v_i) ex-

ceeds or equals a predefined threshold. Algorithm 1 outlines a detailed, step-by-step of the failure prediction process.

Algorithm 1 Bearing failure prediction process.

Input: $S(O_{1-l}, O_n)$ // Dataset test
 P // The Forecast-ANN model
 Condition // Stopping criteria
Begin
 Initialize(Q) // Adjustment-ANN
For $i = 1$ **To** $n - 1$
 $p = P(O_i)$
 Training(Q, O_i, p) // Adjustment-ANN training
Next i
 Current_Observation = $S(O_n)$
Repeat
 Forecasted_observation = $P(\text{Current_Observation})$ // Forecasted observation by Forecast-ANN
 Adjusted_Observation = $Q(\text{Forecasted_observation})$ // Adjusted observation by adjustment-ANN
 Insert(S , Adjusted_Observation) // Insert the adjusted value to the time series
 Current_Observation = Adjusted_Observation
Until Condition(Current_Observation)
 RUL = Current_Observation(t) – $O_n(t)$
Return RUL
End

Figure 8 represents an example of RUL estimation.

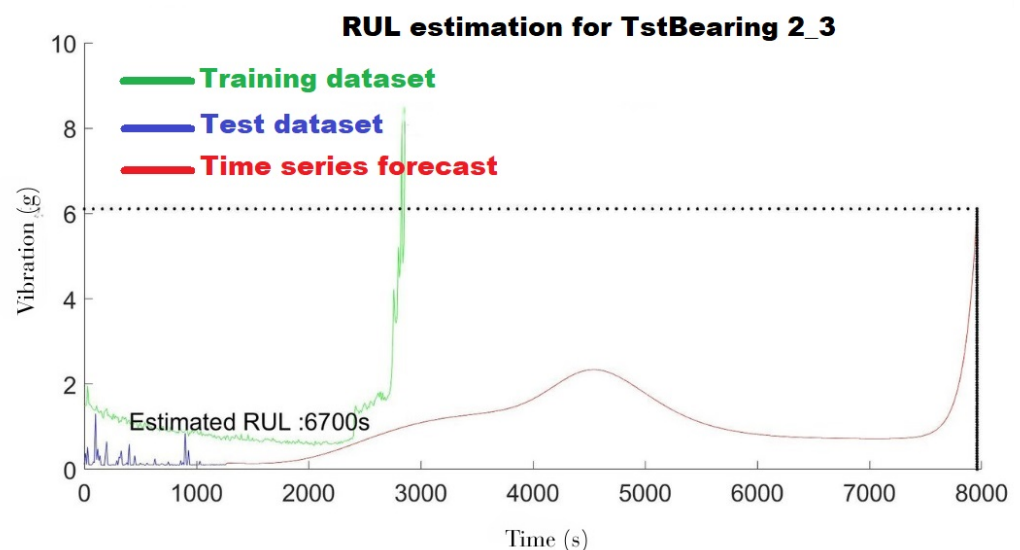


Figure 8. Example of RUL estimation.

4. Validation on Experimental Data

The main challenge in developing a reliable solution for bearing failure prediction in induction machines is the ability to analyze nonstationary time series signals and produce accurate forecasts, especially under faulty operating conditions. This section evaluates the effectiveness of the proposed forecast-ANN model for bearing failure prognosis in comparison with the ARIMA forecasting model. Several configurations were tested before establishing the final architecture. All experiments were implemented and tested on MATLAB 2020 on a 64-bit operating system PC with 8 GB of installed memory and an i7 processor @ 2.00 GHz.

In the following, the selected datasets for evaluation are first described, followed by the metrics used to assess the different methods. Finally, the performance results are analyzed and the conclusions are presented.

4.1. Datasets and Performance Metrics

The datasets used to evaluate the proposed approach consist of the PRONOSTIA bearing degradation dataset [41], which also provides a scoring function based on over- and under-prediction error rates, and the NASA-IMS run-to-failure dataset. These datasets allow us to assess whether the proposed approach maintains its performance and accuracy on nonstationary time series signals collected under different experimental conditions.

The PRONOSTIA experimental platform [41] is dedicated to testing and validating approaches for bearing fault diagnosis and prognosis. It offers datasets generated from run-to-failure experiments, which are carried out until complete bearing failure, defined as reaching a maximum vibration level of 20 g. In this platform, two accelerometers measure the horizontal and vertical vibrations of the bearing, while a thermocouple records its temperature. The vibration and temperature signals are sampled at 25.6 Hz and 10 Hz, respectively. For each experiment, files containing 10-minute segments of recorded signals are produced and stored in dedicated folders. Regarding data management, the datasets are stored in a database using MS SQL Server, with preprocessing and storage procedures developed in VB.Net 2015. A summary of the PRONOSTIA datasets is provided in Table 3.

Table 3. PRONOSTIA datasets operating conditions specifications [41].

Dataset	Operating Conditions		
	Condition 1	Condition 2	Condition 3
Radial Load (N)	4000	4200	5000
Speed (rpm)	1800	1650	1500
Learning sets	B1_1	B2_1	B3_1
	B1_2	B2_2	B3_2
Test set	B1_3	B2_3	B3_3
	B1_4	B2_4	
	B1_5	B2_5	
	B1_6	B2_6	
	B1_7	B2_7	

The measured vibration and temperature signals first undergo preprocessing, including normalization and data aggregation [58]. Normalization is used to map different signal types (time, vibration, and temperature) within a comparable numerical range, allowing each feature to be assigned an appropriate relative weight through dedicated coefficients. Data aggregation is then applied to reduce noise and smooth fluctuations in the raw measurements, thereby decreasing problem complexity by grouping the samples over predefined time windows.

To assess the robustness and generalizability of the proposed approach beyond PRONOSTIA, further validation is performed on the NASA-IMS run-to-failure dataset, which allows us to verify whether the proposed approach retains its performance and accuracy on nonstationary time series acquired under different experimental conditions.

The NASA-IMS run-to-failure dataset for bearing data consists of four bearings mounted on a shaft driven at a constant speed of 2000 rpm with an applied radial load of 6000 lbs. High-sensitivity quartz ICP accelerometers were installed on the bearing housing

(two accelerometers for each bearing that consist of the X and Y axes). Vibration signals were acquired with a sampling frequency of 20 kHz and stored as ASCII files within an interval of 10 minutes between files. Three run-to-failure datasets are provided [20,42]: (i) Set 1 (2156 files) failures occurred in bearing 3 (inner race) and bearing 4 (roller element). (ii) Set 2 (984 files) an outer race failure occurred in bearing 1. (iii) Set 3 (4448 files) an outer race failure occurred in bearing 3. As a preprocessing step, measured vibration signals were aggregated by file into statistical descriptors (average and variance) for each X/Y vibration axis of each bearing. The aggregated signals were stored into an MS-SQL table with fields including set_name, file_name, block_index, at_time, and per-bearing statistical descriptors. For training and test of the proposed approach, the subsets were filtered by set_name, bearing_number, and block_index ranges.

The performance of the proposed approaches was assessed based on the following key metrics:

- The RUL estimation score, denoted as A_i and introduced by PRONOSTIA, is computed based on the error rate as follows [41]:

$$Er_i = 100 \times \frac{RUL_{actual} - RUL_{calculated}}{RUL_{actual}} \quad (3)$$

$$A_i = \begin{cases} \exp^{-\ln(0,5) \cdot (\frac{Er_i}{5})} & \text{if } Er_i \leq 0 \\ \exp^{+\ln(0,5) \cdot (\frac{Er_i}{20})} & \text{if } Er_i > 0 \end{cases} \quad (4)$$

- Mean absolute percentage error (MAPE), a regression-oriented performance metric that provides a meaningful interpretation of result variations by calculating the average of percentage errors. Mathematically, this metric is defined as [49]:

$$MAPE = \frac{1}{n} \sum_i^n \frac{|X_i - X'_i|}{X_i} \quad (5)$$

- Cumulative relative accuracy (CRA), defined as follows:

$$CRA = \frac{1}{n} \sum_i^n \left(1 - \frac{|X_i - X'_i|}{X_i} \right) \quad (6)$$

- Root mean squared error (RMSE), defined as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_i - X'_i)^2}{n}} \quad (7)$$

4.2. Performance Assessment on PRONOSTIA Bearing Degradation Dataset

To validate the proposed failure prognosis framework, two forecasting approaches are investigated: (i) a conventional ARIMA model, representing a strong statistical baseline for time series prediction, and (ii) an adapted ANN-based forecasting architecture specifically designed for nonlinear and nonstationary degradation signals. The objective of this comparative evaluation is to demonstrate the limitations of ARIMA under real bearing degradation conditions and to justify the progressive enhancement of the ANN-based model to achieve robust and accurate remaining useful life (RUL) prediction.

4.2.1. ARIMA Forecast Model's Performance Assessment

The application of the ARIMA forecasting model on the provided signals begins with applying the log function to stabilize variance, followed by the detrend function to eliminate both linear and quadratic trends as a preprocessing step. Next, stationarity is achieved

through iterative first-order differencing until a stationary series is obtained, enabling the application of the ARMA model for forecasting. The resulting outcomes are summarized in Table 4. However, despite this rigorous preprocessing and modeling pipeline, the ARIMA model demonstrates limited predictive capability. Indeed, in 9 out of 11 test cases, the ARIMA model failed to provide RUL, indicating a fundamental limitation in its ability to accurately model the bearing's degradation trends, which are inherently nonstationary and nonlinear.

Table 4. ARIMA application for failures prediction.

Dataset	Test End (s)	Actual RUL (s)	Calculated Failure Time (s)	RUL (s)	Er_i	Score %
Bearing1_3	18,020	5730	18,800	780	86.38	5.01
Bearing1_4	11,390	339	11,720	330	2.65	91.21
Bearing1_5	23,020	1610	Stopping criterion not reached	NA	NA	0.00
Bearing1_6	23,020	1460	Stopping criterion not reached	NA	NA	0.00
Bearing1_7	15,020	7570	Stopping criterion not reached	NA	NA	0.00
Bearing2_3	12,020	7530	Stopping criterion not reached	NA	NA	0.00
Bearing2_4	6120	1390	Stopping criterion not reached	NA	NA	0.00
Bearing2_5	20,020	3090	Stopping criterion not reached	NA	NA	0.00
Bearing2_6	5720	1290	Stopping criterion not reached	NA	NA	0.00
Bearing2_7	1720	580	Stopping criterion not reached	NA	NA	0.00
Bearing3_3	3520	820	Stopping criterion not reached	NA	NA	0.00
Final Score:						8.75

NA: RUL is not available.

Despite the solid mathematical foundation of the ARIMA forecasting model, its performance remains limited for this type of degradation series. This observation is consistent with the findings in [32], which state that “ARIMA models yield large errors when the degradation trend is nonstationary, nonlinear, and noisy”. Such limitations, mainly driven by the nonlinear and nonstationary nature of bearing degradation signals, highlight the need for data-driven models with greater representational power. Accordingly, the following subsection introduces a progressively refined ANN-based forecasting framework designed to overcome these challenges and deliver more accurate RUL estimation.

4.2.2. ANN-Based Approach's Performance Assessment

Due to the inefficiency of the ARIMA forecast model application, a dedicated forecasting ANN model was designed for nonstationary time series. Several approaches were tested and improved step by step before establishing the final solution detailed in Section 3. The proposed ANN-based approach's evolution steps are the following:

- The first proposed ANN model predicts vibration as a function of time. It is trained on time series data of the form $f(t_i) = v_i$, where v_i denotes the vibration value measured at time t_i .
- The second proposed ANN model provides a forecast for the next observation based on the current observation as HI: $O_i(t_i, v_i) \rightarrow O_{i+1}(t_{i+1}, v_{i+1})$.
- The third proposed ANN model provides a forecast for the next observation based on a set of the last k observations using only vibration signals as HI: $[O_{i-k}, \dots, O_i] \rightarrow O_{i+1}(t_{i+1}, v_{i+1})$.
- The last proposed ANN model provides a forecast for the next observation based on a set of last k observations using vibration and temperature signals as HI: $[O_{i-k}, \dots, O_i] \rightarrow O_{i+1}(t_{i+1}, v_{i+1}, c_{i+1})$.

To enhance the performance of the proposed forecasting ANN, two additional steps were used (HS construction and fine-tuning), as follows:

- The HS construction is based on CMO-Clustering to extract data that describes the faulty operating mode to ensure reliable training.
- The fine-tuning consists of correcting the errors of the proposed forecast ANN using an adjustment-ANN.

For clarity in terminology throughout this paper, PRONOSTIA datasets are divided into two types: learning sets, used for the training phase, and test sets, used to evaluate the performance of the final model during the test phase. To evaluate the performance of the proposed approach, $k = 50$ is used to define the number of previous observations used as input for the forecast-ANN. This parameter was determined through experimentation to provide an optimal balance between historical context and prediction accuracy.

The proposed forecasting ANN is a multilayer of interconnected neurons with a default activation function on each one. A feedforward neural network (FNN) based on the backpropagation algorithm learns complex patterns from historical data to make accurate future predictions and adjusts the network's weights iteratively to minimize prediction errors between the actual and predicted outputs [60]. Available training functions for this edition of FNN are the following:

- trainlm: Based on “Levenberg–Marquardt” algorithm;
- trainbr: Based on “Bayesian Regularization” algorithm;
- trainbfg: Based on “BFGS Quasi-Newton” algorithm;
- trainrp: Based on “Resilient Backpropagation” algorithm;
- trainscg: Based on “Scaled Conjugate Gradient” algorithm;
- traincgb: Based on “Conjugate Gradient with Powell/Beale Restarts” algorithm;
- traincgf: Based on “Fletcher-Powell Conjugate Gradient” algorithm;
- traincgp: Based on “Polak-Ribière Conjugate Gradient” algorithm;
- trainoss: Based on “One Step Secant” algorithm;
- traingdx: Based on “Variable Learning Rate Gradient Descent” algorithm;
- traingdm: Based on “Gradient Descent with Momentum” algorithm;
- traingd: Based on “Gradient Descent” algorithm.

In general, Levenberg–Marquardt (trainlm) exhibits fast convergence on small to medium networks, while Bayesian Regularization (trainbr) can improve generalization at the cost of longer training. First-order methods such as trainscg and trainrp scale better to larger problems, with typically slower convergence per epoch but lower memory footprint [61,62]. Indeed, by following an empirical approach in the validation phase of the proposed solution on experimental data, the trainlm training function based on the Levenberg–Marquardt algorithm shows more efficiency in terms of training precision and result accuracy compared to the other functions.

The first proposed ANN model predicts vibration as a function of time, i.e., $f(t_i) = v_i$ (Figure 9). The corresponding failure prognosis results are reported in Table 5.

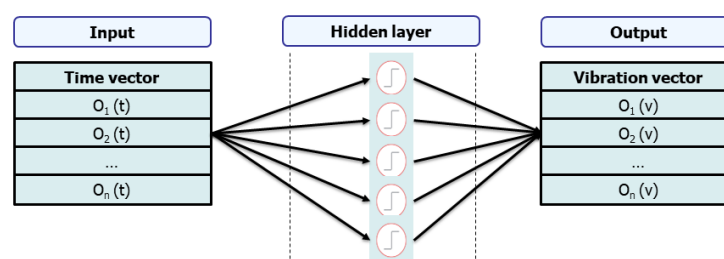
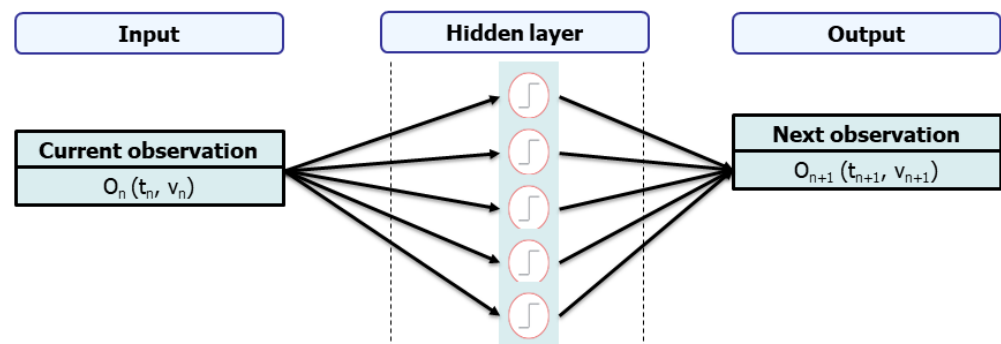


Figure 9. ANN model providing vibration depending on time.

Table 5. RUL prediction results of the first proposed ANN.

Dataset	Test End (s)	Actual RUL (s)	The First ANN Application		
			RUL (s)	Er_i	Score %
Bearing1_3	18,020	5730	2650	53.75	15.52
Bearing1_4	11,390	339	10	97.05	3.46
Bearing1_5	23,020	1610	3168	−96.77	0
Bearing1_6	23,020	1460	2457	−68.29	0.01
Bearing1_7	15,020	7570	4011	47.01	19.6
Bearing2_3	12,020	7530	10	99.87	3.14
Bearing2_4	6120	1390	2100	−51.08	0.08
Bearing2_5	20,020	3090	10	99.68	3.16
Bearing2_6	5720	1290	2380	−84.5	0
Bearing2_7	1720	580	5020	−765.52	0
Bearing3_3	3520	820	82879	−10,007.2	0
Final score:					7.72

The results obtained with this forecasting ANN are limited, as the prediction of vibration over time is insufficient. It depends on HI at a given time. From this, a second forecast ANN is proposed which consists of providing a forecast model for the next observation based on the current observation as HI (Figure 10). The associated failure prognosis results are presented in Table 6.

**Figure 10.** Second ANN providing forecast based on current observation.

Upon evaluating this forecasting model, we observed noticeable improvement over the first one. However, its performance remains limited, as the prediction of the next observation cannot rely solely on the current point; it must also capture the underlying degradation history. This motivated the design of an enhanced ANN that incorporates the last k observations as health indicators (HIs) to forecast the next point. Its architecture is shown in Figure 3, and the associated results appear in Table 7.

It is worth noting that the findings from the final ANN demonstrate its capability to estimate the bearing RUL. Although computing the score using the method proposed in [41] is challenging, the predicted RUL values are, nevertheless, close to the actual RUL.

Figure 8 shows a failure prognosis example using the final approach. The final results of the three proposed ANNs are subject to improvement, using HS construction by data clustering in order to pass just the faulty operating mode data, and using the adjustment-ANN as fine-tuning during the test step.

Table 6. RUL prediction results of the second proposed ANN.

Dataset	Test End (s)	Actual RUL (s)	The Second ANN Application		
			RUL (s)	Er_i	Score %
Bearing1_3	18,020	5730	4260	25.65	41.1
Bearing1_4	11,390	339	7760	−2189	0
Bearing1_5	23,020	1610	1520	5.59	82.39
Bearing1_6	23,020	1460	6560	−349.32	0
Bearing1_7	15,020	7570	7210	4.76	84.8
Bearing2_3	12,020	7530	4830	35.86	28.86
Bearing2_4	6120	1390	4450	−220.14	0
Bearing2_5	20,020	3090	1790	42.07	23.27
Bearing2_6	5720	1290	1040	19.38	51.09
Bearing2_7	1720	580	5070	−774.14	0
Bearing3_3	3520	820	4280	−421.95	0
Final score:					28.32

Table 7. RUL prediction results of the final proposed ANN.

Dataset	Test End (s)	Actual RUL (s)	The Final ANN Application		
			RUL (s)	Er_i	Score %
Bearing1_3	18,020	5730	6170	−7.68	34.49
Bearing1_4	11,390	339	190	43.95	21.8
Bearing1_5	23,020	1610	1520	5.59	82.39
Bearing1_6	23,020	1460	1680	−15.07	12.38
Bearing1_7	15,020	7570	7200	4.89	84.42
Bearing2_3	12,020	7530	6700	11.02	68.25
Bearing2_4	6120	1390	1440	−3.6	60.73
Bearing2_5	20,020	3090	4390	−42.07	0.29
Bearing2_6	5720	1290	1030	20.16	49.73
Bearing2_7	1720	580	4060	−600	0
Bearing3_3	3520	820	940	−14.63	13.15
Final score:					38.88

4.3. Performance Assessment on NASA-IMS Run-to-Failure Dataset

Unlike PRONOSTIA, where all tests are run-to-failure, NASA-IMS datasets are not all run-to-failure; this motivated us to make some adaptations to the proposed approach and to use other metrics that do not rely on full RUL availability. For example, in Set 1, only Bearing 3 (inner race defect) and Bearing 4 (roller element defect) go down to failure. For the NASA-IMS dataset, the adaptation is performed while keeping the core method unchanged. Specifically, we used the average and variance computed on each aggregated file as features for the ANN input. The average of the aggregated signals describes the bearing health state, whereas the variance characterizes the evolution of degradation over time. We retain the HI construction using n previous observations as the window to forecast

the next one, where inputs are sliding windows built from these two features. An example of the resulting RUL estimation is illustrated in Figure 11.

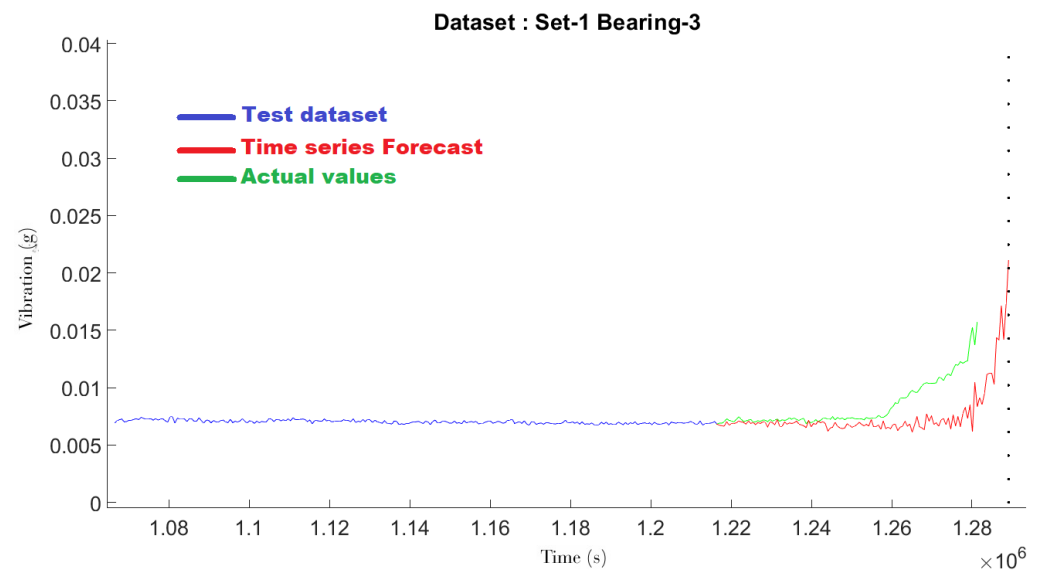


Figure 11. Example of RUL estimation on NASA-IMS using the proposed approach.

On NASA-IMS, the proposed ANN maintains a favorable accuracy/robustness trade-off. Representative results show low RMSE and MAPE on the test set, and a stable cumulative CRA over time, indicating sustained relative accuracy as degradation progresses.

To evaluate the performance of the proposed approach during the test phase, the same failure type is used for both the training and test datasets. The model is trained on Set-*i*/Bearing-*k* and tested on Set-*j*/Bearing-*l* with the same failure type, using CRA, RMSE, MAPE, and PRONOSTIA score metrics for all tests where run-to-failure data is available. Table 8 reports the performance outcomes for each dataset and bearing.

Table 8. NASA-IMS results per dataset and bearing using the proposed approach.

Dataset	Channel	Pronostia Score (%)	CRA	RMSE	MAPE (%)
Set-1	Bearing-1	79.27	0.9880	0.0008	1.20
Set-1	Bearing-2	80.82	0.9838	0.0007	1.61
Set-1	Bearing-3	60.45	0.9781	0.0018	2.19
Set-1	Bearing-4	87.33	0.9839	0.0012	1.61
Set-2	Bearing-1	76.75	0.9337	0.0052	6.62
Set-2	Bearing-2	77.68	0.9408	0.0033	5.92
Set-2	Bearing-3	52.38	0.9713	0.0032	2.87
Set-2	Bearing-4	61.26	0.8924	0.0017	10.70
Set-3	Bearing-1	72.56	0.9155	0.0022	8.45
Set-3	Bearing-2	78.58	0.9463	0.0011	9.76
Set-3	Bearing-3	51.37	0.9577	0.0025	4.23
Set-3	Bearing-4	68.67	0.9370	0.0059	6.30

Overall, for tests involving the same failure type, the global performance metrics are as follows: PRONOSTIA score 44.02%, CRA 0.9308, RMSE 0.01735, and MAPE 6.92%. These results indicate that the proposed approach not only achieves accurate predictions of the bearing RUL but also maintains consistent performance across different datasets within the NASA-IMS benchmark. The high CRA value reflects the model's ability to closely follow the actual degradation trend, while the low RMSE and MAPE values demonstrate precise quantitative predictions with minimal average error. Furthermore, the PRONOSTIA score

confirms the reliability of the predictions in a prognostics context. Taken together, these metrics highlight the approach's portability, robustness, and potential for generalization across bearings and operating conditions, confirming its suitability for practical predictive maintenance applications.

4.4. Performance Comparison

To assess the effectiveness of the proposed ANN-based approach, a comprehensive comparative analysis was conducted against recent state-of-the-art data-driven methodologies for bearing failure prognosis. The comparison includes the following:

- A data-driven approach based on Recurrent-CNN with direct RUL mapping [35].
- A supervised machine learning approach based on SVM [37].
- A supervised machine learning approach based on KNN [37].
- A data-driven approach based on Multiscale-CNN with direct RUL mapping [36].
- A data-driven approach based on Double-CNN with direct RUL mapping [63].
- A data-driven approach based on LSTM with direct RUL mapping [39].
- A data-driven approach based on GAM-CNN with direct RUL mapping [38].
- A data-driven approach based on BiLSTM with direct RUL mapping [40].

This evaluation highlights the relative performance of the proposed method compared to advanced deep learning and traditional machine learning techniques.

In this study, there is only the last observation, which represents the failure time of the test dataset, so the MAPE of the obtained results was calculated with $n = 1$. Table 9 shows the performance comparison for bearing RUL estimation using PRONOSTIA datasets [41].

Table 9. MAPE performance metric comparison of some data-driven approaches.

Dataset	[35]	[37]-SVM	[37]-KNN	[36]	[61]	[39]	[38]	[40]	Proposed Approach
Bearing1_3	-	-	-	0.1743	-	-	-	-	0.0713
Bearing1_4	-	-	-	0.8691	-	-	-	-	0.4395
Bearing1_5	-	-	-	0.5971	-	-	-	-	0.0559
Bearing1_6	0.2219	-	-	0.1628	-	-	-	-	0.1310
Bearing1_7	0.2581	-	-	0.1955	-	-	-	-	0.0489
Bearing2_3	-	-	-	-	-	-	-	-	0.1102
Bearing2_4	-	-	-	-	-	-	-	-	0.0347
Bearing2_5	-	-	-	-	-	-	-	-	0.2961
Bearing2_6	0.1884	-	-	-	-	-	-	-	0.2016
Bearing2_7	0.1547	-	-	-	-	-	-	-	0.8571
Bearing3_3	-	-	-	-	-	-	-	-	0.1277
Final score (MAPE):	0.2058	0.2590	0.2700	0.3998	0.2375	-	-	-	0.2158
Final score (RMSE):	-	-	-	-	-	0.0326	0.0297	0.0267	-

The results clearly demonstrate that the proposed approach achieves competitive or superior performance despite using a simpler network architecture and requiring less training data. One can notice that only approach A1 [35], based on Recurrent-CNN with direct RUL mapping strategy, achieves a final score of 0.2058, which is slightly better than the proposed approach's score of 0.2158. However, Ref. [35] does not report scores for all tests (only for Bearings 1_6, 1_7, 2_6, and 2_7). The proposed approach demonstrates notable performance in several specific test cases (*Bearing1_3*, *Bearing1_5*, *Bearing1_7*, *Bearing2_4*) with MAPE values below 0.08, substantially outperforming all compared methods.

It is worth noting that the evaluation revealed a high MAPE (0.8571) for the *Bearing2_7* test case, indicating a potential challenge with this particular dataset, which may

be attributed to unique degradation patterns not represented in the training data or to measurement anomalies specific to this test case. However, rather than arbitrarily excluding this challenging case, we report both the complete results (MAPE of 0.2158) and an analysis of the impact of this potential outlier. If we exclude the score from test Bearing 27, the final score of the proposed approach improves to 0.1389, significantly outperforming all the compared approaches. This dual reporting approach is consistent with rigorous statistical practice in prognostic research, where outlier detection and handling must be transparent and justified [64].

To further validate the robustness and generalizability of the proposed approach, we conducted an additional comparative analysis on the NASA-IMS bearing dataset. This cross-platform validation is crucial as it demonstrates the approach's effectiveness under different experimental conditions and data acquisition. Table 10 presents a quantitative comparison with recent state-of-the-art methods, using RMSE as the primary metric for fair comparison.

Table 10. RMSE comparison on NASA-IMS dataset.

Reference	HI Construction	Prediction Model	RMSE
[39]	Stacked autoencoder	RNN	0.0245
[40]	KS-density	BiLSTM	0.0198
[38]	Feature extraction	GAM-CNN	0.0223
[50]	Parallel feature fusion	CNN	0.0267
Proposed approach	Sliding observation window	Dual-ANN	0.01735

As shown in Table 10, the proposed approach achieves the lowest RMSE (0.01735) among all compared methods on the NASA-IMS dataset. This performance gain is particularly significant given that the proposed approach uses a simpler architecture with only two input attributes (average and variance) within a sliding observation window compared to the complex feature engineering and deep architectures employed by competing methods. The BiLSTM approach [40], which represents the current state of the art, achieves an RMSE of 0.0198, while traditional LSTM and CNN-based methods yield higher errors. This demonstrates that the innovative combination of unsupervised health stage clustering with targeted ANN forecasting can outperform sophisticated deep learning architectures while maintaining lower computational complexity and reduced data requirements.

These comparative results provide strong empirical evidence of the novelty and effectiveness of the proposed approach. While complex deep learning models like CNN and LSTM variants have demonstrated promising results, they typically require extensive training data and computational resources. In contrast, this contribution achieves competitive or superior performance through the innovative use of unsupervised clustering for health stage construction and a dual-ANN architecture for forecasting and fine-tuning, making it particularly suitable for practical industrial applications with limited failure data availability.

4.5. Ablation Analysis

The proposed approach can be viewed as a thoughtfully engineered integration of multiple complementary techniques: health state (HS) construction through CMO-Clustering, health index (HI) formulation via feature selection and normalization, prediction fine-tuning using a dual-ANN architecture, and specific variants of ANN configurations optimized for time series forecasting. To validate the contribution of each of these key components, we conducted comprehensive ablation analysis. This analysis provides empirical evidence for the design choices and quantifies the performance impact

of each component. Table 11 presents the results of these experiments, where components were selectively removed or replaced to measure their individual contribution to the final approach performance.

Table 11. Summary of ablation studies.

Configuration	Final Score (%)	MAPE	Relative Performance
Final proposed approach	38.88	0.2158	100%
The first proposed ANN (time input)	7.72	0.3906	19.85%
The second proposed ANN (single observation input)	28.32	0.2740	72.84%
Without health stage (using all data)	21.63	0.3108	56%
Without temperature features (vibration only)	35.72	0.2331	91.87%
Without fine-tuning (only forecast-ANN)	27.44	0.2784	71%

The results clearly demonstrate the essential contribution of each component to the overall system performance:

- **Historical observations vs. time-based modeling:** The evolution from the first ANN (using time as input) to the second approach (using single observations) showed a strong improvement in performance (from 19.85% to 72.84% relative performance). The results show that the time-based training function is not adapted for bearing degradation trend modeling, which is fundamentally nonstationary. Further enhancement to the final approach (using sequences of observations) improved performance of over 80% compared to the first time-based approach. Using a sequence of previous observations rather than only the current observation improved performance by 27.16%, demonstrating the importance of historical data.
- **Health stage construction:** Removing this component and using all available data for training (both normal and faulty modes) resulted in a 44% performance decrease. This shows that fitting the training data specifically on faulty mode significantly improves the failure predictive performance.
- **Multimodal feature:** The results show that using vibration signals alone provides 91.87% relative performance, and incorporating temperature data alongside vibration signals improves performance by 8.23%. This small enhancement confirms the correlation between mechanical issues (detected through vibration) and temperature increases. However, even with this slight improvement, temperature data offers complementary information, particularly in early degradation phases where thermal changes can precede mechanical manifestations reflected by vibration changes.
- **Fine-tuning mechanism:** The adjustment-ANN contributed a 29% performance improvement, confirming the effectiveness of the error correction approach compared to using only the forecast-ANN. This demonstrates the value of the dual-ANN architecture in adapting to specific degradation patterns.

These ablation studies not only justify the architectural choices but also provide insights into the relative importance of each component, which could guide future research.

5. Conclusions and Future Works

Accurate RUL estimation is essential for effective bearing monitoring, helping to prevent unexpected machine shutdowns and reduce maintenance costs. This paper proposes a novel unsupervised ANN-based approach that forecasts bearing degradation using time series of vibration and temperature signals. The prediction process iteratively applies the forecast-ANN until a stopping criterion corresponding to failure is reached, estimating the remaining useful life. The approach integrates multiple innovations: health stage construc-

tion via CMO-clustering to separate normal and faulty modes, a dual-ANN architecture combining forecast- and adjustment-ANN models for enhanced accuracy, and a multimodal design leveraging both vibration and temperature data. Ablation studies confirmed the critical role of intelligent health stage construction and demonstrated that vibration data alone provides strong predictive power. The proposed method requires less training data and computation than complex deep learning models while delivering accurate and consistent failure prognosis across multiple conditions, outperforming state-of-the-art approaches and offering practical industrial applicability.

It should be noted that the proposed framework is compatible with real-time implementation, as the computationally intensive stages (health stage construction and ANN training) are performed offline, while the online phase only involves signal preprocessing, sliding-window HI construction, and feedforward ANN inference, which can be executed with low latency. This motivates future work to extend the scope of the proposed approach to integrate a complete PHM process, enabling seamless information flow across diagnostic, detection and prognostic phases. The objective is to develop a unified system capable of detecting incipient faults, classifying their type and severity, predicting their progression, and providing accurate RUL estimation. Given the unsupervised nature of the proposed approach, the second dimension of extension consists of generalizing it to a broader range of mechanical systems, particularly rotating machinery such as gears, shafts, and rotors.

Author Contributions: Conceptualization, C.K. and M.B.; methodology, C.K.; software, C.K.; validation, C.K., F.B.-S.T. and K.B.; formal analysis, C.K.; investigation, C.K.; resources, M.B.; data curation, C.K.; writing—original draft preparation, C.K.; writing—review and editing, C.K., F.B.-S.T., K.B. and M.B.; visualization, C.K.; supervision, F.B.-S.T., K.B. and M.B.; project administration, M.B.; funding acquisition, M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available. The PRONOSTIA dataset is available at <https://www.nasa.gov/content/prognostics-center-of-excellence-data-set-repository> (accessed on 15 January 2026). The NASA-IMS dataset is available at <https://www.kaggle.com/datasets/vinayak123tyagi/bearing-dataset> (accessed on 15 January 2026).

Acknowledgments: The authors would like to thank the AS2M Department of FEMTO-ST Institute for providing datasets of the PRONOSTIA experimental platform for this study. The authors also acknowledge the Center for Intelligent Maintenance Systems (IMS), University of Cincinnati, for providing the bearing dataset available through the NASA Open Data Portal, which was used for cross-platform validation.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PHM	Prognostics and Health Management
RUL	Remaining Useful Life
HI	Health Indicator
HS	Health Stage
ANN	Artificial Neural Network
TBM	Time-Based Maintenance
CBM	Condition-Based Maintenance
ML	Machine Learning

AI	Artificial Intelligence
CMO	Cooperative Multiobjective
ARIMA	Auto-Regressive Integrated Moving Average
LSTM	Long Short-Term Memory
CNN	Convolutional Neural Network
SVM	Support Vector Machine
MAPE	Mean Absolute Percentage Error
RMSE	Root Mean Squared Error
CRA	Cumulative Relative Accuracy

References

1. Silva, M.; Josias, G.; Rejane, S.; Moura, C. Discrete Event System Decision: An Approach to the Corrective Maintenance of a Machine Based on Colored Petri Net. *J. Mechatron. Eng.* **2023**, *6*, 10–22. [\[CrossRef\]](#)
2. Syamsundar, A.; Naikan, V.; Wu, S. Estimating maintenance effectiveness of a repairable system under time-based preventive maintenance. *Comput. Ind. Eng.* **2021**, *156*, 107278. [\[CrossRef\]](#)
3. deJonge, B.; Teunter, R.; Tinga, T. The influence of practical factors on the benefits of condition-based maintenance over time-based maintenance. *Reliab. Eng. Syst. Saf.* **2017**, *158*, 21–30. [\[CrossRef\]](#)
4. Atamuradov, V.; Medjaher, K.; Dersin, P.; Lamoureux, B.; Zerhouni, N. Prognostics and health management for maintenance practitioners-Review, implementation and tools evaluation. *Int. J. Progn. Health Manag.* **2017**, *8*, 1–31. [\[CrossRef\]](#)
5. Berghout, T.; Benbouzid, M. A Systematic Guide for Predicting Remaining Useful Life with Machine Learning. *Electronics* **2022**, *11*, 1125. [\[CrossRef\]](#)
6. Magena, C. Machine Learning Models for Predictive Maintenance in Industrial Engineering. *Int. J. Comput. Eng.* **2024**, *6*, 1–14. [\[CrossRef\]](#)
7. Mosallam, A. Remaining Useful Life Estimation of Critical Components Based on Bayesian Approaches. Ph.D. Thesis, Franche-Comté University, Belfort, France, 2014.
8. Iyer, N.; Goebel, K.; Bonissone, P. Framework for post-prognostic decision support. In Proceedings of the 2006 IEEE Aerospace Conference, Big Sky, Montana, 4–11 March 2006; Volume 2006, p. 10. [\[CrossRef\]](#)
9. Gu, Y.; Bi, Q.; Qiu, G. Practical health indicator construction methodology for bearing ensemble remaining useful life prediction with ISOMAP-DE and ELM-WPHM. *Meas. Sci. Technol.* **2021**, *33*, 025007. [\[CrossRef\]](#)
10. Xu, J.; Duan, S.; Chen, W.; Wang, D.; Fan, Y. SACGNet: A Remaining Useful Life Prediction of Bearing with Self-Attention Augmented Convolution GRU Network. *Lubricants* **2022**, *10*, 21. [\[CrossRef\]](#)
11. Xu, Z.; Liu, Q.; Bashir, M.; Liu, J.; Ekere, N. A Novel Health Indicator for Intelligent Prediction of Rolling Bearing Remaining Useful Life based on Unsupervised Learning Model. *Comput. Ind. Eng.* **2023**, *176*, 108999. [\[CrossRef\]](#)
12. Białoń, T.; Niestrój, R.; Michalak, J.; Pasko, M. Induction Motor PI Observer with Reduced-Order Integrating Unit. *Energies* **2021**, *14*, 4906. [\[CrossRef\]](#)
13. Djeziri, M.; Benmoussa, S.; Sanchez, R. Hybrid method for remaining useful life prediction in wind turbine systems. *Renew. Energy* **2018**, *116*, 173–187. [\[CrossRef\]](#)
14. Hu, Y.; Zhang, H.; Li, C.; Liu, S.; Zhang, Y. Exponential smoothing model for condition monitoring: A case study. In Proceedings of the 2013 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE), Chengdu, China, 15–18 July 2013; pp. 1742–1746. [\[CrossRef\]](#)
15. Trull, O.; García-Díaz, J.; Troncoso, A. Initialization Methods for Multiple Seasonal Holt–Winters Forecasting Models. *Mathematics* **2020**, *8*, 268. [\[CrossRef\]](#)
16. Wei, R.; Hu, Y.; He, C.; Cao, Z.; Lu, S.; Liu, F.; Liu, Y. Performance Degradation Prediction of Rolling Bearing based on KJADE and Holt–Winters. In Proceedings of the 2020 International Conference on Sensing, Measurement & Data Analytics in the Era of Artificial Intelligence (ICSMD), Xi'an, China, 15–17 October 2020; pp. 475–478. [\[CrossRef\]](#)
17. Liu, Y.; Qiao, N.; Zhao, C.; Zhuang, J. Vibration Signal Prediction of Gearbox in High-Speed Train Based on Monitoring Data. *IEEE Access* **2018**, *6*, 50709–50719. [\[CrossRef\]](#)
18. Carrión, D.; González, J.; López, G.; Isaac, I. Alternative fault detection method in electrical power systems based on ARMA model. In Proceedings of the 2019 FISE-IEEE/CIGRE Conference—Living the Energy Transition (FISE/CIGRE), Medellín, Colombia, 3–6 December 2019; pp. 1–6. [\[CrossRef\]](#)
19. Yang, Y.; Wu, W.; Sun, L. Prediction of mechanical equipment vibration trend using autoregressive integrated moving average model. In Proceedings of the 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Shanghai, China, 14–16 October 2017; pp. 1–5. [\[CrossRef\]](#)

20. Lee, J.; Qiu, H.; Yu, G.; Lin, J. Rexnord Technical Services, Bearing Data Set. In *NASA Ames Prognostics Data Repository*; NASA Ames Research Center: Moffett Field, CA, USA, 2007.
21. Wang, L.; Wu, Z.; Fu, Y.; Yang, G. Remaining life predictions of fan based on time series analysis and BP neural networks. In *Proceedings of the 2016 IEEE Information Technology, Networking, Electronic and Automation Control Conference*, Chongqing, China, 20–22 May 2016; pp. 607–611. [\[CrossRef\]](#)
22. Guedes, A.; Silva, S. Insulation Failures Prognosis in Electric Machines: Preventive Detection and Time to Failure Forecast. *IET Electr. Power Appl.* **2020**, *14*, 1108–1117. [\[CrossRef\]](#)
23. Berghout, T.; Benbouzid, M.; Mouss, L.H. Leveraging Label Information in a Knowledge-Driven Approach for Rolling-Element Bearings Remaining Useful Life Prediction. *Energies* **2021**, *14*, 2163. [\[CrossRef\]](#)
24. Xu, P.; Tu, Z.; Li, M.; Wang, J.; Wang, X.-B. Remaining Useful Life Prediction for Rolling Bearings based on RVM-Hausdorff Distance. *Meas. Sci. Technol.* **2023**, *34*, 125121. [\[CrossRef\]](#)
25. Saidi, L.; Ben Ali, J.; Benbouzid, M.; Bechhoefer, E. Wind Turbine Drivetrain Prognosis Approach Based on Kalman Smoother with Confidence Bounds. In *Proceedings of the IEEE International Conference on Industrial Technology ICIT 2018At: Lyon, France*, 20–22 February 2018; pp. 1865–1870. [\[CrossRef\]](#)
26. Pandit, R.; Xie, W. Data-driven models for predicting remaining useful life of high-speed shaft bearings in wind turbines using vibration signal analysis and sparrow search algorithm. *Energy Sci. Eng.* **2023**, *11*, 4557–4569. [\[CrossRef\]](#)
27. Berghout, T.; Benbouzid, M. Diagnosis and Prognosis of Faults in High-Speed Aeronautical Bearings with a Collaborative Selection Incremental Deep Transfer Learning Approach. *Appl. Sci.* **2023**, *13*, 10916. [\[CrossRef\]](#)
28. Tian, Z. An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring. *J. Intell. Manuf.* **2012**, *23*, 227–237. [\[CrossRef\]](#)
29. Singh, J.; Darpe, A.; Singh, S. Bearing remaining useful life estimation using an adaptive data driven model based on health state change point identification and K-means clustering. *Meas. Sci. Technol.* **2020**, *31*, 085601. [\[CrossRef\]](#)
30. Delpha, C.; Diallo, D.; Harmouche, J.; Benbouzid, M.; Amirat, Y.; Elbouchikhi, E. Bearing Fault Diagnosis in Rotating Machines. In *Electrical System II, from Diagnosis to Prognosis*; ISTE Wiley: London, UK, 2020; Chapter 4, pp. 123–152.
31. Gao, Z.; Cecati, C.; Ding, S. A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-base approaches. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3757–3767. [\[CrossRef\]](#)
32. Wang, T.; Yu, J.; Siegel, D.; Lee, J. A similarity-based prognostics approach for Remaining Useful Life estimation of engineered systems. In *Proceedings of the 2008 International Conference on Prognostics and Health Management*, Denver, CO, USA, 6–9 October 2008; pp. 1–6. [\[CrossRef\]](#)
33. Sharma, A. A review of fault diagnostic and monitoring schemes of induction motors. *Int. J. Res. Appl. Sci. Eng. Technol.* **2015**, *3*, 1145–1152.
34. Amirat, Y.; Elbouchikhi, E.; Delpha, C.; Benbouzid, M.; Diallo, D. Modal Decomposition for Bearing Fault Detection, Electrical System I, From Diagnosis to Prognosis. In *Electrical System I, from Diagnosis to Prognosis*; ISTE Wiley: London, UK, 2020; Chapter 4, pp. 121–168.
35. Wang, B.; Lei, Y.; Yan, T.; Li, N.; Guo, L. Recurrent convolutional neural network: A new framework for remaining useful life prediction of machinery. *Neurocomputing* **2019**, *379*, 117–129. [\[CrossRef\]](#)
36. Zhu, J.; Chen, N.; Peng, W. Estimation of Bearing Remaining Useful Life Based on Multiscale Convolutional Neural Network. *IEEE Trans. Ind. Electron.* **2019**, *66*, 3208–3216.
37. Chelmiah, E.; McLoone, V.; Kavanagh, D. Low Complexity Non-Linear Spectral Features and Wear State Models for Remaining Useful Life Estimation of Bearings. *Energies* **2023**, *16*, 5312. [\[CrossRef\]](#)
38. Du, X.; Jia, W.; Yu, P.; Shi, Y.; Gong, B. RUL prediction based on GAM-CNN for rotating machinery. *J. Braz. Soc. Mech. Sci. Eng.* **2023**, *45*, 142. [\[CrossRef\]](#)
39. Han, T.; Pang, J.; Tan, A. Remaining useful life prediction of bearing based on stacked autoencoder and recurrent neural network. *J. Manuf. Syst.* **2021**, *61*, 576–591. [\[CrossRef\]](#)
40. Habbouche, H.; Benkedjouh, T.; Amirat, Y.; Benbouzid, M. Rotating machine bearing health prognosis using a data driven approach based on KS-density and BiLSTM. *IET Sci. Meas. Technol.* **2024**, *19*, e12215. [\[CrossRef\]](#)
41. Nectoux, P.; Gouriveau, R.; Medjaher, K.; Ramasso, E.; Morello, B.; Zerhouni, N.; Varnier, C. PRONOSTIA: An Experimental Platform for Bearings Accelerated Life Test. In *Proceedings of the IEEE International Conference on Prognostics and Health Management*, Denver, CO, USA, 18–21 June 2012; pp. 1–8.
42. Qiu, H.; Lee, J.; Lin, J. Wavelet Filter-based Weak Signature Detection Method and its Application on Roller Bearing Prognostics. *J. Sound Vib.* **2006**, *289*, 1066–1090. [\[CrossRef\]](#)
43. Vachtsevanos, G.; Lewis, F.; Roemer, M.; Hess, A.; Wu, B. *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*; John Wiley and Sons Inc.: Hoboken, NJ, USA, 2006.
44. Wang, Y.; Chen, Z.; Zhang, Y.; Li, X.; Li, Z. Remaining useful life prediction of rolling bearings based on the three-parameter Weibull distribution proportional hazards model. *Insight Non-Destr. Test. Cond. Monit.* **2020**, *62*, 710–718. [\[CrossRef\]](#)

45. Zong, T.; Li, J.; Lu, G. Auxiliary model-based multi-innovation PSO identification for Wiener–Hammerstein systems with scarce measurements. *Eng. Appl. Artif. Intell.* **2021**, *106*, 104470. [CrossRef]
46. Singleton, R.; Strangas, E.; Aviyente, S. Extended kalman filtering for remaining-useful-life estimation of bearings. *IEEE Trans. Ind. Electron.* **2015**, *62*, 1781–1790. [CrossRef]
47. Mélard, G.; Pasteels, J. Automatic ARIMA modeling including interventions, using time series expert software. *Int. J. Forecast.* **2000**, *16*, 497–508. [CrossRef]
48. Njimi, H.; Mélard, G.; Pasteels, J. Modélisation SARIMA assistée. In Proceedings of the XXXVèmes Journées de Statistique, Lyon, France, 13–17 May 2003; Volume 2, pp. 731–734.
49. Pandit, R.; Santos, M.; Sierra-García, J. Comparative analysis of novel data-driven techniques for remaining useful life estimation of wind turbine high-speed shaft bearings. *Energy Sci. Eng.* **2024**, *12*, 4613–4623. [CrossRef]
50. Zhang, H.; Shuai, L.; Hu, S. A parallel feature fusion network for simultaneous bearing fault detection and remaining useful life prediction. In Proceedings of the 2023 IEEE International Conference on Mechatronics and Automation (ICMA), Piscataway, NJ, USA, 6–9 August 2023; pp. 2064–2069.
51. Berghout, T.; Benbouzid, M. UBO-EREX: Uncertainty Bayesian-Optimized Extreme Recurrent EXpansion for Degradation Assessment of Wind Turbine Bearings. *Electronics* **2024**, *13*, 2419. [CrossRef]
52. Dongdong, C.; Huixin, L.; Zhang, D. Analysis of transient temperature field of cylindrical roller bearings. *Mach. Des. Manuf.* **2018**, *565*, 62–64.
53. Zhang, Y.; Li, H.; Wang, H. Data-driven wind-induced response prediction for slender civil infrastructure: Progress, challenges and opportunities. *Structures* **2025**, *74*, 108650. [CrossRef]
54. Gholamian, M.; Seryasat, O.R.; Mohammadzadeh, K.; Ghanaatshoar, M.; Alickovic, E. A physics-informed deep learning approach for bearing remaining useful life prediction under variable operating conditions. *Sci. Rep.* **2023**, *13*, 104295. [CrossRef]
55. Lei, Y. XJTU-SY Rolling element bearing accelerated life test datasets: A tutorial. *J. Mech. Eng.* **2019**, *55*, 1–6.
56. Loparo, K. Bearing Data Center, Case Western Reserve University. 2013. Available online: <https://engineering.case.edu/bearingdatacenter/download-data-file> (accessed on 15 January 2026).
57. Lei, Y.; Li, N.; Guo, L.; Li, N.; Yan, T.; Lin, J. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mech. Syst. Signal Process.* **2018**, *104*, 799–834. [CrossRef]
58. Khamoudj, C.; Benbouzid-SiTayeb, F.; Benatchba, K.; Benbouzid, M.; Djaafri, A. A Learning Variable Neighborhood Search Approach for Induction Machines Bearing Failures Detection and Diagnosis. *Energies* **2020**, *13*, 2953. [CrossRef]
59. Khamoudj, C.; Benbouzid-SiTayeb, F.; Benatchba, K.; Benbouzid, M. Classical mechanics-inspired optimization metaheuristic for induction machines bearing failures detection and diagnosis. In Proceedings of the IECON, Beijing, China, 29 October–1 November 2017; pp. 3803–3808.
60. Liu, H. On the Levenberg–Marquardt training method for feed-forward neural networks. In Proceedings of the 2010 Sixth International Conference on Natural Computation, Yantai, China, 10–12 August 2010; Volume 1, pp. 456–460. [CrossRef]
61. Ampazis, N.; Perantonis, S.J. Levenberg–Marquardt Algorithm with Adaptive Momentum for the Efficient Training of Feed-forward Networks. In Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN 2000): Neural Computing: New Challenges and Perspectives for the New Millennium, Como, Italy, 24–27 July 2000; Volume 1, pp. 126–131. [CrossRef]
62. Troiano, M.; Nobile, E.; Mangini, F.; Mastrogiuseppe, M.; Barbaro, C.C.; Frezza, F. A Comparative Analysis of the Bayesian Regularization and Levenberg–Marquardt Training Algorithms in Neural Networks for Small Datasets: A Metrics Prediction of Neolithic Laminar Artefacts. *Information* **2024**, *15*, 270. [CrossRef]
63. Yang, B.; Liu, R.; Zio, E. Remaining Useful Life Prediction Based on a Double-Convolutional Neural Network Architecture. *IEEE Trans. Ind. Electron.* **2019**, *66*, 9521–9530. [CrossRef]
64. Hodkiewicz, M.; Batsioudis, Z.; Radomiljac, T.; Ho, M. Statistical best practices for prognostic research in maintenance applications. *Int. J. Progn. Health Manag.* **2020**, *11*, 1–18.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.