machines

MDPI

Article

Unsupervised Remaining Useful Life Estimation of Tool-Holder Bearings from Sparse-Autoencoder Reconstruction Error and Exponential Degradation Modeling

Giuseppe Dipace [1,*], Emiliano Mucchi [2] and Gianluca D'Elia [3]

[1] M.T. S.r.l., Via Casino Albini 480, 47842 San Giovanni in Marignano, Italy
[2] Department of Engineering, University of Ferrara, Via Saragat 1, 44122 Ferrara, Italy; mccmln@unife.it
[3] Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, 42122 Modena, Italy; gianluca.delia@unimore.it
* Correspondence: dpcgpp@unife.it

Abstract

This paper presents a hybrid approach for tool-holder bearing prognosis, combining a Sparse Autoencoder with an Exponential Degradation Model. The objective is to enable early detection of wear and degradation, allowing for timely maintenance interventions and minimizing unplanned downtime. The Sparse Autoencoder is employed as an unsupervised anomaly detection tool, trained solely on defect-free data to learn normal bearing behavior and identify deviations through reconstruction errors. These errors serve as a health indicator, capturing early signs of degradation. To estimate the Remaining Useful Life, the Exponential Degradation Model is applied to the health indicator, fitting an exponential curve to predict the degradation rate and forecast the time remaining until failure. The experimental study involved two tool-holders operating under identical conditions. Tool-holder 1 was used for SAE development and for the offline calibration of the health-indicator thresholds, whereas the complete fixed configuration was subsequently applied to tool-holder 2 without SAE retraining or threshold recalibration. The results obtained on tool-holder 2 provide a preliminary within-domain evaluation of the proposed framework on an unseen component. A comparative analysis with three baseline models—Principal Component Analysis (PCA), a Naive Constant Model (NCM), and a supervised Long Short-Term Memory (LSTM) network—further contextualized the results. The Sparse Autoencoder consistently outperformed PCA and NCM in terms of accuracy and robustness, while requiring only a short defect-free baseline for training. The LSTM achieved the lowest Root Mean Square Error values overall, but its dependence on full life-cycle data and higher computational demands limit its practicality in industrial scenarios. Overall, the main contribution of this study lies in demonstrating that a simple, interpretable, and data-efficient hybrid framework can achieve reliable Remaining Useful Life prediction without relying on large supervised datasets, thus bridging the gap between deep learning performance and industrial applicability.

Keywords: anomaly detection; Remaining Useful Life; predictive maintenance; machine learning; vibration analysis

Check for updates

Academic Editor: Jun Zha

Received: 5 June 2026
Revised: 7 July 2026
Accepted: 10 July 2026
Published: 13 July 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) license.

1. Introduction

Early detection and accurate wear prediction in tool-holder bearings are critical to ensuring the reliability and efficiency of the manufacturing process. In this context, predictive

maintenance strategies have gained substantial attention due to their ability to forecast the Remaining Useful Life (RUL), allowing timely interventions that prevent unexpected breakdowns and optimize operational costs.

RUL estimation methods are generally classified into three categories: model-based approaches, data-driven techniques, and hybrid models. Model-based approaches rely on physical laws and failure mechanisms to simulate degradation behavior. These methods often require detailed knowledge of the operational conditions and failure processes, but they offer interpretable insights into the bearing's health status [1]. In contrast, data-driven techniques use historical and real-time sensor data to model the degradation patterns of bearings, providing an empirical framework for predicting RUL without requiring explicit knowledge of the failure mechanisms [2]. Hybrid approaches combine both data-driven and model-based strategies to leverage the strengths of each, enabling more accurate and robust RUL predictions [3].

Data-driven bearing prognosis commonly combines vibration-derived time- and frequency-domain features with machine learning models. Representative approaches include CNNs and LSTMs for automated representation learning and temporal modeling, transfer learning methods intended to improve applicability to unseen operating domains, and hybrid degradation models for RUL forecasting [4–8]. Time–frequency processing also remains important for revealing fault-related spectral changes in non-stationary vibration signals [9,10].

In addition to the prognostic algorithm, industrial deployment requires a reliable data path connecting sensors and machine tools to acquisition, processing, and monitoring services. Tapia et al. [11] implemented a scalable platform for real-time machine-tool monitoring, demonstrating the relevance of connected and scalable data-processing infrastructures. Such platforms provide the enabling infrastructure within which condition-monitoring and prognostic algorithms can operate. The present work focuses specifically on the analytical layer of this chain, transforming sequential vibration features into an updated RUL estimate through a lightweight SAE–Exponential Degradation Model framework.

Despite the impressive progress achieved by data-driven models, several open challenges remain, including the demand for extensive labeled datasets and the limited interpretability of complex learning architectures. To address these challenges, increasing attention has been directed toward approaches that balance predictive capability with computational efficiency and transparency. In this context, the present study introduces a lightweight hybrid strategy for bearing RUL estimation that integrates a Sparse Autoencoder (SAE) with an Exponential Degradation Model. The SAE, trained on vibration features from healthy bearings, serves as an unsupervised anomaly detector whose reconstruction error acts as a health indicator. This indicator is subsequently modeled through an exponential degradation law to estimate the RUL in a simple yet interpretable manner. The proposed approach aims to bridge the gap between high-performance data-driven methods and practical, real-time applications in industrial condition monitoring.

The contribution of this study lies in the integration of a lightweight SAE-based anomaly detector with an interpretable exponential degradation model for RUL prediction. Related approaches have adopted different strategies for constructing health indicators and forecasting degradation. Xu et al. [12] combined a moving-average stacked autoencoder with exponential smoothing, whereas She et al. [13] employed a regularized sparse autoencoder and a particle filter. Nie et al. [14] proposed a supervised similarity-feature-fusion method in which time-, frequency-, and time–frequency-domain descriptors are screened and fused through PCA to construct a health indicator, which is subsequently supplied to a one-dimensional CNN trained with run-to-failure RUL labels. In contrast, the present framework uses six physically informed time- and frequency-domain features, trains a

single SAE exclusively on defect-free observations, and defines the reconstruction error as the health indicator. No similarity-feature construction, PCA-based fusion, or supervised CNN regression is employed. The growing health-indicator history is instead processed through an Exponential Degradation Model, and the RUL is obtained from the predicted crossing of fixed condition thresholds. Accordingly, the proposed approach emphasizes low training-data requirements, computational simplicity, and transparent sequential updating rather than the direct supervised RUL regression adopted by Nie et al. [14].

The experimental study was conducted on two tool-holders operating under identical conditions. The first developed ball pitting after 80 h, whereas the second exhibited an inner-ring raceway defect after 82 h. Tool-holder 1 was used to develop the SAE and to calibrate the health-indicator thresholds offline. The SAE parameters, feature-normalization statistics, and thresholds were then frozen and applied without modification to tool-holder 2. Consequently, tool-holder 1 represents the development and calibration case, whereas tool-holder 2 provides a preliminary within-domain evaluation of the fixed framework on an unseen component. Broader claims regarding generalization remain preliminary and require future validation on additional tool-holders, operating conditions, and degradation modes.

The remainder of this paper is structured as follows: Section 2 introduces the fundamental concepts of SAEs and the Least Squares Method for exponential regression modeling. Section 3 describes the experimental setup, signal acquisition procedures, and feature extraction used for model training and validation. Section 4 presents the application of the proposed approach to two case studies, analyzes the results, and discusses the model's accuracy and consistency within the tested conditions. Finally, Section 6 summarizes the key findings and outlines the potential implications for predictive maintenance in industrial settings.

2. Theoretical Background

This section provides the theoretical foundations underlying the proposed RUL estimation approach. It first introduces SAEs, a type of neural network used for unsupervised feature extraction and anomaly detection. Subsequently, the section outlines the application of the Least Squares Method (LSM) to fit exponential regression models, which are used to model degradation trends over time for RUL prediction.

2.1. Sparse Autoencoders

SAEs are a specialized type of neural network used for unsupervised learning, aimed at reducing data dimensionality while capturing essential features. They belong to the broader class of Autoencoders, which are neural networks designed to learn efficient data representations by attempting to reconstruct input data at their output layer. However, SAEs differ from traditional Autoencoders by introducing sparsity constraints in their latent space, promoting more selective activation of neurons, and encouraging the model to focus on the most important features [1].

An SAE consists of two primary components: an encoder that compresses the input data into a latent representation, and a decoder that reconstructs the input from this compressed form. During training, the network updates its parameters to minimize the reconstruction error, commonly measured by the Mean Squared Error (MSE) between the input and output. In addition to minimizing the reconstruction loss, SAE incorporates regularization terms, such as sparsity regularization and L_2 regularization, in order to enhance the generalization capacity of the learned representations. Sparsity regularization encourages the network to learn efficient and compact features by preventing trivial identity mappings, in which the network merely replicates the input without learning useful

patterns [15]. This is achieved by penalizing neurons that are frequently activated. The average activation of each neuron across the training set is computed and constrained to be low, encouraging neurons to respond strongly only to specific features present in a limited number of inputs.

A common method for enforcing sparsity is the Kullback–Leibler (KL) divergence, which measures the difference between the average activation $\hat{\rho}_j$ of neuron j and desired sparsity level ρ [16,17]. The KL divergence term is expressed as:

$$\Omega_{sparsity} = \sum_{j=1}^{M} \rho \log\left(\frac{\rho}{\hat{\rho}_j}\right) + (1-\rho)\log\left(\frac{1-\rho}{1-\hat{\rho}_j}\right) \tag{1}$$

The KL divergence is minimized when $\hat{\rho}_j$ approaches the desired sparsity level ρ and increases as they diverge. Minimizing the KL divergence term ensures that the learned representation is sparse. However, a network could reduce the sparsity regularizer by increasing the values of weights or decreasing the activations, potentially leading to overfitting or instability. To prevent this, L_2 regularization is applied, penalizing large weight values and promoting a more balanced distribution of learning across neurons [15,18].

The overall cost function for training an SAE, combining reconstruction loss with sparsity and L_2 regularization, is given by:

$$E(\mathbf{w}) = \frac{1}{N}\sum_{n=1}^{N}(\mathbf{x}_n - \hat{\mathbf{x}}_n)^2 + \lambda\Omega_{weights} + \beta\Omega_{sparsity} \tag{2}$$

where λ and β are hyperparameters controlling the influence of L_2 and sparsity regularization, respectively. This formulation ensures a balanced trade-off between accurate reconstruction and regularization, resulting in meaningful and generalizable features.

2.2. Anomaly Detection with Sparse Autoencoders

SAEs are particularly well-suited for anomaly detection due to their ability to capture the regular patterns of normal data. To detect anomalies, the SAE is first trained on a dataset representing normal operating conditions. During training, the model adjusts its weights to minimize the reconstruction error. The sparsity constraint forces the network to focus on the most informative patterns in the data, which are representative of normal behavior [19].

Once trained, the SAE is tested on new data that may contain both normal and anomalous instances. For each input, the model attempts to reconstruct the data and measures the reconstruction error, typically by using MSE [20]:

$$hi_n = \frac{1}{D}\sum_{i=1}^{D}(x_{in} - \hat{x}_{in})^2 \tag{3}$$

where x_{in} is the value of the i-th feature of the n-th observed data, and $\hat{x}_{in}$ is its corresponding reconstruction [21]. If the MSE for a data point exceeds a predefined threshold, the point is flagged as anomalous. This threshold can be determined empirically based on training data distribution or using statistical techniques [22]. Normal data yields low reconstruction errors, while deviations from learned patterns manifest as high errors [23]. In summary, SAEs provide an effective anomaly detection framework by leveraging reconstruction error as a health indicator. The sparsity constraint enhances sensitivity to deviations from normal behavior, making SAEs a powerful tool for identifying faults and detecting outliers in real-world applications [24].

2.3. Application of the Least Squares Method to Exponential Regression Models

The LSM is a foundational statistical technique used to estimate model parameters by minimizing the sum of squared differences between observed and predicted values. In the context of degradation modeling, LSM can be used to fit exponential trends, which often characterize wear or failure processes. A simplified form of the exponential model can be expressed as follows:

$$hi(t) = \theta e^{vt} \tag{4}$$

where $hi(t)$ is the health indicator at time t, θ is a scaling parameter, and v is the exponential growth or decay rate. The goal is to estimate the parameters θ and v that best fit the observed data. Since Equation (4) is nonlinear in v, directly applying LSM is computationally challenging. To address this, the model can be linearized through a logarithmic transformation. This transformation allows the use of standard linear regression techniques to estimate v and θ [25]. By applying linear regression to the transformed model, the optimal estimates of v and θ can be obtained through the following closed-form expressions [26]:

$$v = \frac{\sum_{n=1}^{N} t_n z_n - \bar{z} \sum_{n=1}^{N} t_n}{\sum_{n=1}^{N} t_n^2 - \bar{t} \sum_{n=1}^{N} t_n} \tag{5}$$

and

$$\theta = e^{\bar{z} - v\bar{t}} \tag{6}$$

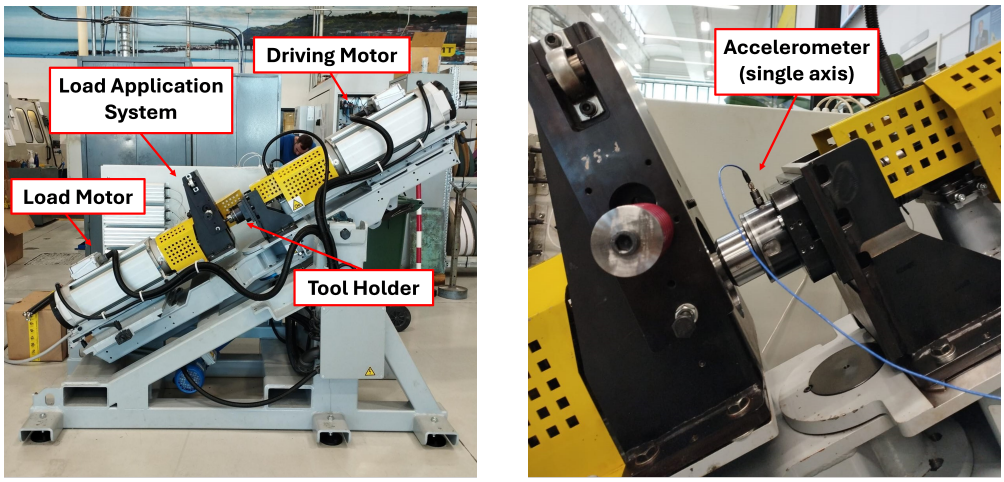
where $\bar{t}$ is the mean of the independent variable t, and $\bar{z}$ is the mean of the transformed dependent variable $z = \ln(hi)$ [27]. In summary, the Least Squares Method can be effectively applied to fit an exponential regression model by leveraging a logarithmic transformation. This allows robust estimation of model parameters and supports real-time tracking of system health over time [28].

3. Methodology

This section outlines the methodology adopted to estimate the RUL of tool-holder bearings. It begins with a description of the experimental setup, followed by an assessment of signal quality. Subsequently, it details the feature extraction process and the architecture, training, and interpretation of the SAE used for anomaly detection and health monitoring. Finally, it presents the procedure developed to translate the health indicator into an online RUL estimate, enabling proactive maintenance decisions.

3.1. Experimental Set-Up

For the purpose of this study, an endurance test bench was developed to reproduce operating conditions representative of tool-holder use during machining. Figure 1 illustrates the test bench, which consists of two electro-spindles, a load application system, and integrated control systems, including inverters and encoders. In this configuration, the first electro-spindle serves as a drive motor to impart a controlled rotational speed to the tool-holder, while the second one acts as a brake, applying a counteracting torque. The load system is positioned between the two spindles, and it can apply forces in axial and radial directions, with a maximum load capacity of 5000 N, enabling the simulation of cutting operations such as milling or drilling. The electro-spindles support speeds up to 10,000 rpm and provide a nominal torque of 110 Nm, allowing a wide range of machining scenarios.



(**a**) Main components of the test bench (**b**) Measurement point

Figure 1. Endurance test bench.

To ensure high-accuracy signal acquisition, a PCB J352C04 (PCB Piezotronics, Model J352C04—specifications: https://www.pcb.com/it/prodotti?m=J352C04 (accessed on 9 July 2026)) single-axis, ground-isolated IEPE accelerometer was used, paired with an NI 9234 (National Instruments, NI 9234—product page: https://www.ni.com/it-it/shop/model/ni-9234.html (accessed on 9 July 2026)) data-acquisition board. Vibration data were acquired continuously at $f_s = 25.6$ kHz; however, for storage and processing, only one 1 s segment (25,600 samples) was kept every 10 min, giving a 600:1 reduction in saved data and defining the rate used during analysis.

The vibration-measurement and data-collection procedure was contextualized with reference to ISO 17359:2018 and ISO 13373-1 [29,30]. ISO 17359 provides general guidance for establishing a machinery condition-monitoring program, whereas ISO 13373-1 addresses the measurement and collection of vibration data. In the present experimental campaign, the sensor position, acquisition hardware, sampling frequency, signal-window duration, and acquisition interval were maintained unchanged throughout both run-to-failure tests to provide a repeatable basis for vibration-trend monitoring.

To apply this vibration-monitoring framework to the present case study, it is first necessary to define the specific mechanical system under analysis. In machining applications, different types of tool-holders can be used depending on the specific operation and on the required mechanical function. For instance, Fernández-Lucio et al. [31] studied a threading holder specifically designed for tapping operations. In contrast, the present work focuses on a radial rotary tool-holder, whose function and internal bearing arrangement are described below. An external view of the investigated radial rotary tool-holder is shown in Figure 2. The cutting tool is mounted on the upper side, whereas the lower side provides the drive interface used to mount the tool-holder in the machine and set it into rotation.

The experimental campaign was carried out on two radial rotary tool-holders, denoted as tool-holder 1 and tool-holder 2 throughout the manuscript, both having the same mechanical configuration. Each tool-holder transmits rotary motion and torque to the cutting tool through an output shaft supported by a pair of identical NSK 7008A5TRSUMP3 precision angular-contact ball bearings mounted in a back-to-back arrangement (DB). These bearings constitute the components monitored in the present prognostic study. The bearing closest to the tool-side output is hereafter referred to as the front bearing. Vibration signals were recorded at a measurement point on the tool-holder casing, as shown in Figure 1b. Each tool-holder was rotated at 3000 rpm and subjected to constant axial and radial loads of 1500 N. The loads were applied to the output shaft at a distance of 110 mm from the

geometric center of the bearing pair. Consequently, the radial load generated a bending moment of 165 N m about the geometric center of the bearing pack. An intermittent duty cycle consisting of nine minutes of rotation followed by one minute of stopping was adopted to reproduce the operating cycle of the tool-holder and to limit overheating. Both the axial and radial loads were maintained constant during the rotation and stopping phases. Each endurance test continued until the vibration Root Mean Square (RMS) exceeded 2.5 g. This threshold was adopted as an application-specific experimental stop criterion for the investigated tool-holder configuration and should not be interpreted as a universal ISO vibration-severity limit. Once this condition was reached, the test was stopped and the tool-holder was disassembled to identify the damaged bearing and the location of the defect. In both cases, the damage was found on the front bearing, i.e., the bearing closest to the tool-side output. The test conditions, total durations, and post-test findings are summarized in Table 1.

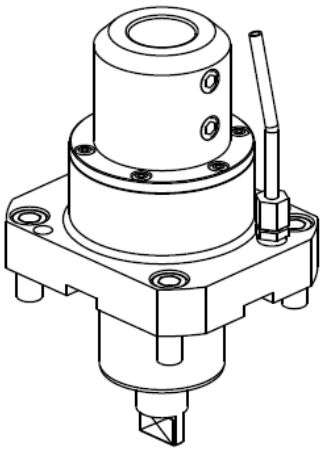
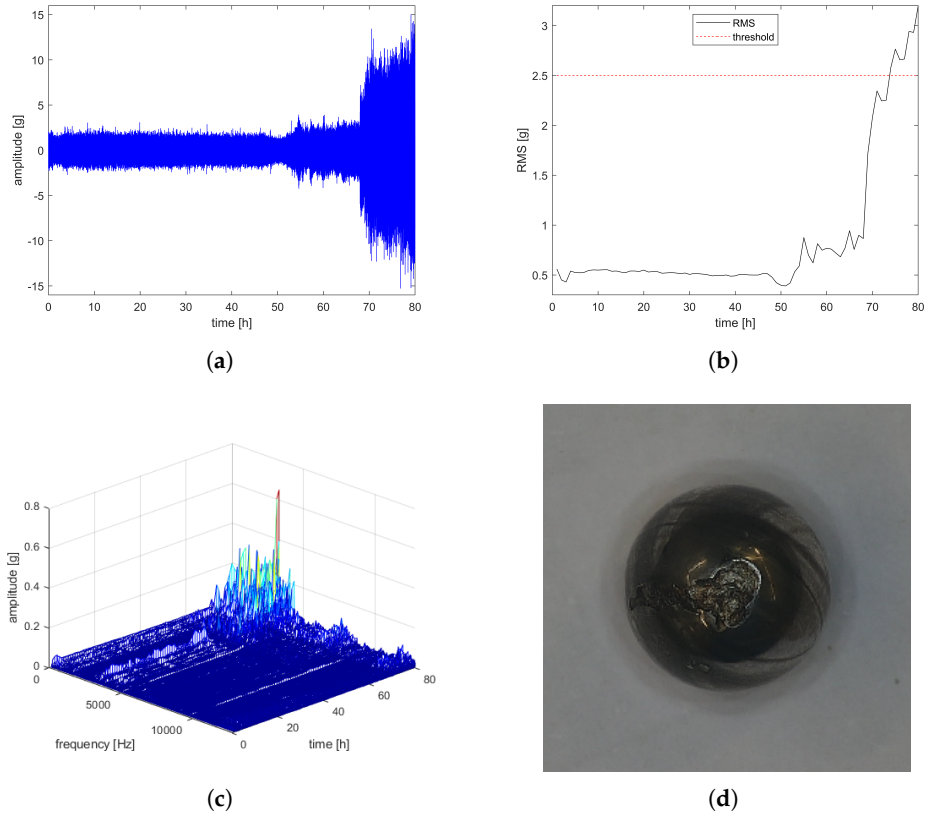


Figure 2. External view of the investigated radial rotary tool-holder. The cutting tool is mounted on the upper side, whereas the lower side provides the drive interface used to mount the tool-holder in the machine and transmit rotation.

Table 1. Test conditions and post-test findings for the two tool-holder endurance tests.

	Tested Cases	
	Case 1	**Case 2**
Bearing model	NSK 7008A5TRSUMP3	NSK 7008A5TRSUMP3
Bearing arrangement	Back-to-back (DB)	Back-to-back (DB)
Speed [rpm]	3000	3000
Axial load [N]	1500	1500
Radial load [N]	1500	1500
Load application distance [mm]	110	110
Bending moment [N m]	165	165
Applied torque [N m]	0	0
Rotation phase [hh:mm]	00:09	00:09
Stop phase [hh:mm]	00:01	00:01
Axial and radial loads during the stop phase	Maintained	Maintained
Testing time [hh:mm]	80:00	82:00
Defect location	Front bearing (tool side)	Front bearing (tool side)
Observed defect	Rolling-element pitting	Inner-ring raceway defect

Before applying the prognostic framework, the acquired vibration signals were examined in the time and time–frequency domains to verify their suitability for degradation monitoring. A time-domain visual inspection of the raw vibration signal from tool-holder 1 (Figure 3a) revealed a stochastic pattern, with amplitudes fluctuating around a near-zero mean. Toward the end of the test, a noticeable increase in amplitude indicated progressive component degradation. This observation was corroborated by the RMS trend (Figure 3b), which showed a stable baseline followed by a sharp rise beyond the critical threshold, indicating failure. To further analyze the signal, Short-Time Fourier Transform (STFT) was applied (Figure 3c). The STFT revealed the evolution of frequency content over time, with increasing spectral amplitude and the emergence of new components toward the end of the test—characteristics typical of defect progression. Visual inspection confirmed this finding, identifying pitting on the rolling elements of the front angular-contact ball bearing, i.e., the bearing closest to the tool-side output, as shown in Figure 3d. A similar vibration evolution was observed for tool-holder 2, as shown in Figure 4. In this case, post-test inspection identified a defect on the inner-ring raceway of the front angular-contact ball bearing, as shown in Figure 4d.



(**a**) (**b**)

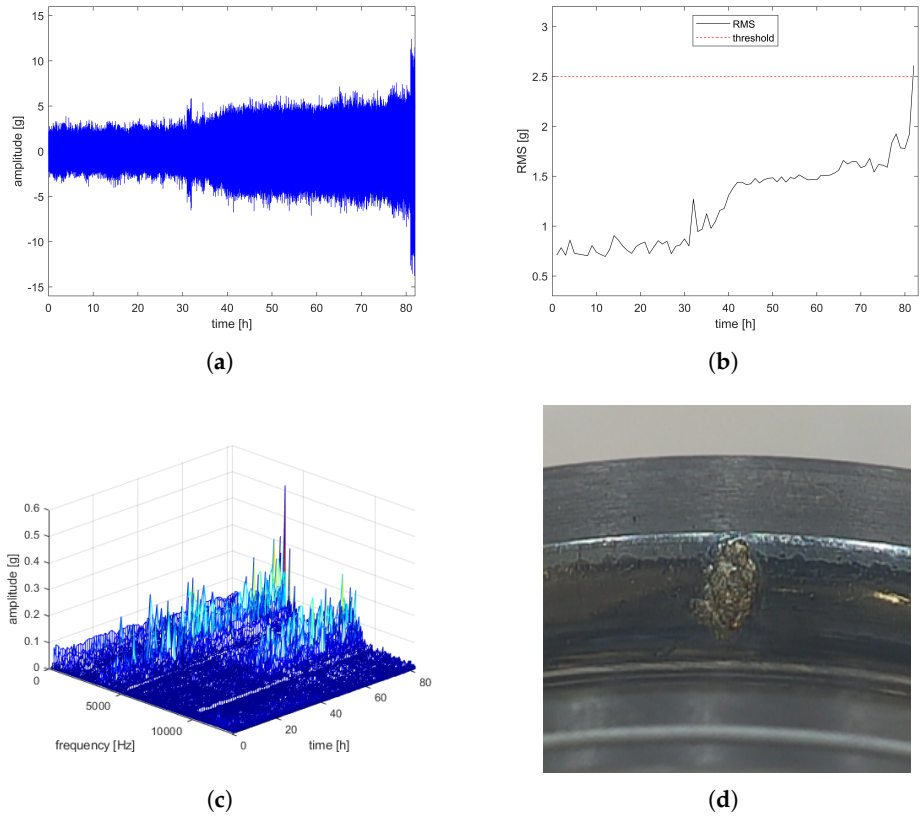
(**c**) (**d**)

Figure 3. Vibration measurements and post-test damage observed for radial tool-holder 1. (**a**) Raw vibration signal; (**b**) Root Mean Square (RMS) values; (**c**) Short-Time Fourier Transform (STFT); (**d**) rolling-element pitting detected in the front bearing (tool side).

3.2. Feature Extraction

To characterize the health condition of the tool-holders, six features were extracted from the vibration data collected during the experimental campaign. These features were computed over 1 s time windows, acquired at 10 min intervals as detailed in Section 3.1, and selected based on their sensitivity to variations in bearing condition. The selected features include both time-domain and frequency-domain indicators, enabling a comprehensive analysis of the vibration behavior. The time-domain features—Root Mean

Square (RMS), Average Amplitude (AA), and Peak-to-Peak ($P2P$)—quantify the amplitude characteristics of the signal. The frequency-domain features—Mean Spectrum Amplitude (MSA), Gravity Frequency (GF), and Mean Square Frequency (MSF)—describe the spectral distribution of signal energy and frequency content. Given a discrete time-series signal $\{s_n,\ n = 0, 1, ..., N-1\}$, the time-domain features are defined as shown in Table 2, where s_n represents the sampling point of the time series.



(**a**) (**b**)

(**c**) (**d**)

Figure 4. Vibration measurements and post-test damage observed for radial tool-holder 2. (**a**) Raw vibration signal; (**b**) Root Mean Square (RMS) values; (**c**) Short-Time Fourier Transform (STFT); (**d**) inner-ring raceway defect detected in the front bearing (tool side).

To compute the frequency-domain features, the signal is first transformed into the frequency domain using the Discrete Fourier Transform (DFT). Since the DFT of real-valued signals yields a conjugate-symmetric spectrum, only the first half of the spectrum is used in the calculation of spectral features [32]. The definitions of the frequency-domain features are also provided in Table 2, where S_k is the spectral amplitude component corresponding to the f_k frequency.

The selected feature set maintains a direct connection with conventional vibration condition-monitoring practice and is consistent with the signal-processing principles described in ISO 13373-2 [33]. RMS is retained as an indicator of the overall vibration level, while AA and P2P provide complementary measures of signal amplitude. MSA, GF, and MSF are derived from the DFT spectrum, which can be computed efficiently using the FFT algorithm, and quantify changes in spectral magnitude and in the distribution of vibration energy over frequency.

The feature selection is also physically motivated. Local rolling-bearing defects generate repeated impacts at frequencies that depend on the damaged component. These impacts excite the structural resonances of the bearing–tool-holder assembly and modify both the overall vibration amplitude and the spectral-energy distribution. The SAE does

not replace the conventional indicators; rather, it combines their multivariate evolution into a scalar reconstruction-error health indicator. The proposed framework is therefore more accurately described as a physically informed data-driven prognostic approach than as a first-principles physical degradation model.

Table 2. Features extracted.

Root Mean Square	Average Amplitude	Peak-to-Peak
$RMS = \sqrt{\frac{1}{N}\sum_{n=0}^{N-1} s_n^2}$	$AA = \frac{1}{N}\sum_{n=0}^{N-1} \lvert s_n \rvert$	$P2P = max(s_n) - min(s_n)$
Mean Spectrum Amplitude	**Gravity Frequency**	**Mean Square Frequency**
$MSA = \frac{2}{N}\sum_{k=1}^{N/2} S_k$	$GF = \frac{\sum_{k=1}^{N/2} f_k S_k}{\sum_{k=1}^{N/2} S_k}$	$MSF = \frac{\sum_{k=1}^{N/2} f_k^2 S_k}{\sum_{k=1}^{N/2} S_k}$

3.3. Sparse Autoencoder for Tool-Holder Bearing Prognosis

This section presents the use of a SAE for detecting wear and degradation in tool-holder bearings—a key factor in ensuring the reliability and efficiency of the manufacturing process. The SAE was developed as an anomaly detection model capable of identifying early signs of bearing degradation using vibration signal analysis. The SAE compresses input features into a lower-dimensional latent space and reconstructs them at the output layer. Its three-layer architecture includes a sparsity constraint on the hidden layer, which encourages the network to focus on the most informative patterns in the data while ignoring irrelevant noise. The network was trained on vibration data from the first 25 h of defect-free operation of tool-holder 1, to learn the normal bearing behavior. Anomalies are then identified by comparing the reconstructed features with the actual values. Any significant deviations between these values signal the onset of bearing degradation.

In this context, the MSE between the original and reconstructed input serves as a health indicator. When the MSE exceeds predefined thresholds, it signals abnormal behavior and potential bearing failure. This approach enables early maintenance intervention, by minimizing downtime and repair costs.

The SAE architecture, shown in Figure 5, comprises three layers. The input layer receives six features extracted from vibration signals: RMS, AA, $P2P$, MSA, GF, and MSF. The hidden layer, which contains three neurons with logarithmic sigmoid (*logsig*) activation functions, serves as a compressed representation of the input data. Finally, the output layer, consisting of six neurons with identity (*purelin*) activation functions, aims to reconstruct the input features. The objective of the SAE is to minimize the reconstruction error between the input and output layers, ensuring that the network captures the critical features of the input data while maintaining a sparse and efficient representation.

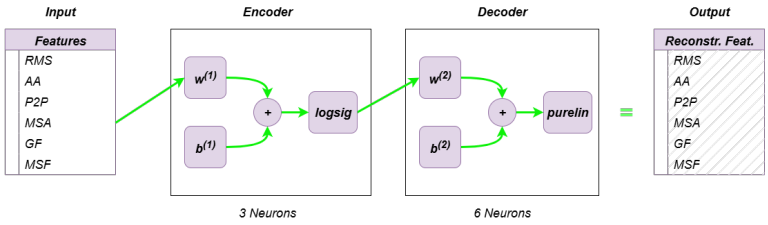


Figure 5. Sparse Autoencoder architecture.

The SAE was trained on the first 25 h of defect-free data from tool-holder 1, capturing the baseline vibration behavior. To improve robustness and mitigate the risk of overfitting to a single subset of healthy data, a blocked K-fold cross-validation approach with weight averaging was employed. Specifically, the first 25 h of defect-free data from tool-holder 1 were divided into $K = 5$ contiguous, non-overlapping folds. For each fold, the autoencoder was trained on the remaining $K - 1$ folds and validated on the excluded fold. The weights and biases from each trained model were then averaged to form a final, consolidated SAE model. This strategy ensures that the model captures consistent healthy features across different data partitions, improving its stability across different data partitions without requiring additional tool-holder experiments.

The regularization coefficients—L_2 regularization coefficient λ, sparsity regularization coefficient β, and target sparsity level ρ—were determined empirically to achieve an effective trade-off between reconstruction accuracy and model reliability. Specifically, several candidate values for λ, β, and ρ were evaluated through a series of preliminary experiments. In each trial, the coefficients were adjusted iteratively, and the corresponding mean squared error (MSE) on the training and validation data was monitored. The final set of coefficients ($\lambda = 0.005$, $\beta = 1$, and $\rho = 0.33$) was selected as it resulted in the lowest MSE while maintaining the ability to accurately reconstruct the vibration features during the first 25 h of defect-free operation.

The K-fold procedure was implemented in MATLAB R2018b, and the five fold-specific SAE models were trained using the Scaled Conjugate Gradient (*trainscg*) algorithm. Training was stopped when one of the following conditions was reached: a maximum of 1000 epochs, an MSE target of 10^{-3} or a minimum gradient threshold of 10^{-5}. Finally, the averaged model was constructed as a feed-forward neural network with a 6-3-6 topology. Table 3 summarizes the training results of the five fold-specific SAE models. For each model, the table reports the final training MSE, the epoch at which training stopped, and the final gradient magnitude. Additionally, it indicates the time intervals used as training and validation sets for each fold. The reported MSE values demonstrate that all models achieved similar and consistently low reconstruction errors, confirming the consistency and reliability of the training procedure. The variability in stopping epochs reflects slight differences in convergence speed among folds, while the final gradient magnitudes indicate that the convergence criteria were properly satisfied.

Table 3. Training summary of the 5 fold-specific Sparse Autoencoder (SAE) models.

SAE Models (Blocked 5-Fold)	Final Training MSE	Epoch at Convergence	Final Gradient Magnitude	Training Interval [h]	Validation Interval [h]
Model 1	0.119	223	8.9×10^{-7}	5–25	0–5
Model 2	0.118	440	8.9×10^{-7}	0–5, 10–25	5–10
Model 3	0.135	618	9.5×10^{-7}	0–10, 15–25	10–15
Model 4	0.134	615	9.7×10^{-7}	0–15, 20–25	15–20
Model 5	0.126	758	9.5×10^{-7}	0–20	20–25

Although additional run-to-failure repetitions on further tool-holders and under different operating conditions would strengthen the statistical validation, such tests could not be performed within the present experimental campaign because of their substantial duration, cost, and destructive nature. The limited availability of complete run-to-failure datasets represents a recognized practical constraint in bearing-prognosis research [34,35]. Within the available dataset, the blocked *K*-fold procedure was used to assess the stability of SAE training across different contiguous portions of the healthy data from tool-holder 1 and to reduce dependence on a single temporal partition. Tool-holder 2 was then kept entirely

separate from model development and used to evaluate the fixed SAE and threshold configuration on an unseen component under the same operating conditions. Nevertheless, this procedure does not replace independent experimental repetitions or establish statistical generalization across different speeds and loads.

To assess the performance of the SAE, Figure 6 presents six plots comparing the actual values of the six input features (*RMS*, *AA*, *P2P*, *MSA*, *GF*, and *MSF*) with the values reconstructed by the SAE during the test on radial tool-holder 1. The actual feature values are represented by solid black lines, while the reconstructed values from the SAE are depicted by solid orange lines.

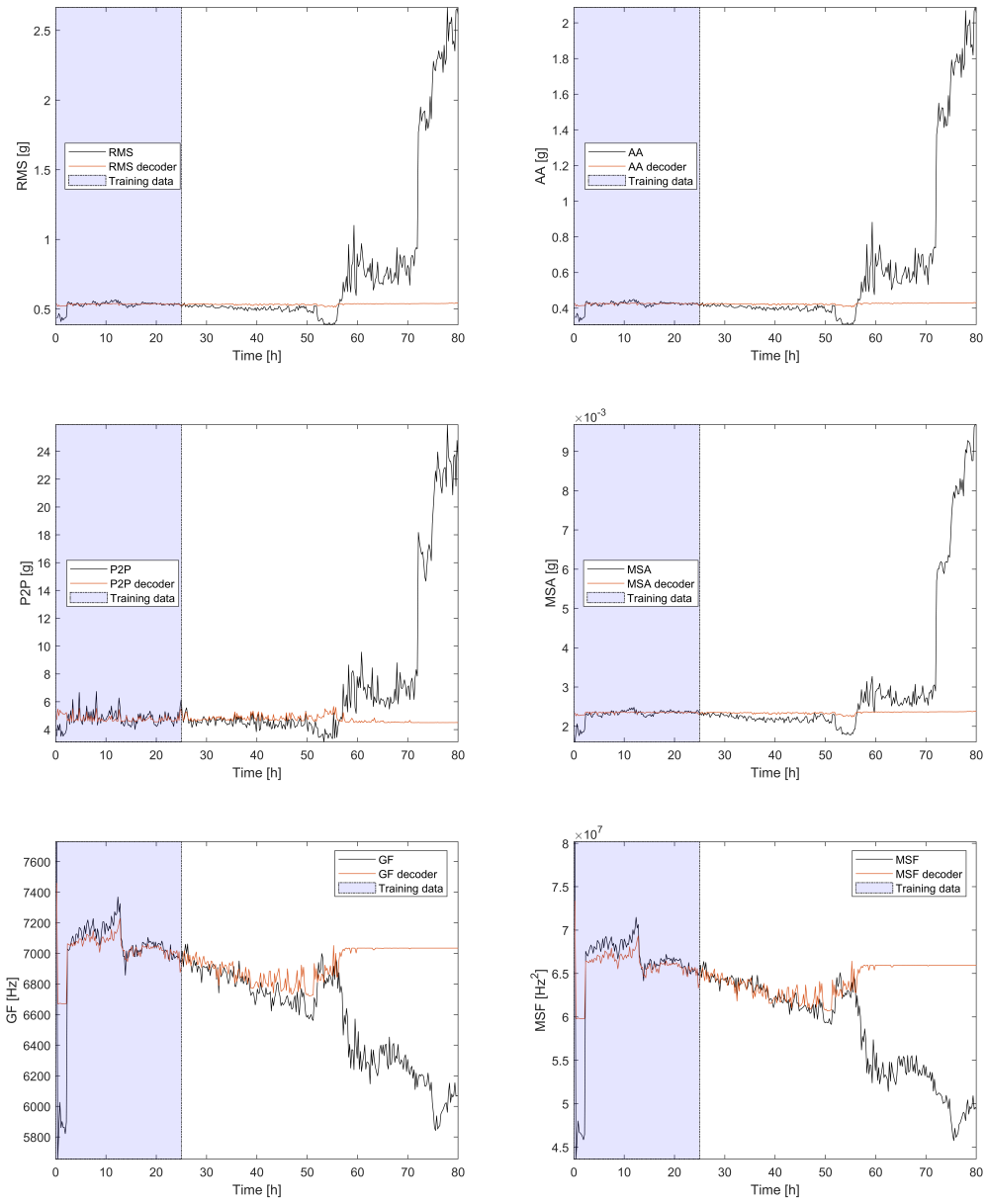


Figure 6. Comparison between the actual values of the six extracted features (*RMS*, *AA*, *P2P*, *MSA*, *GF*, *MSF*) and the values reconstructed by the Sparse Autoencoder (SAE) during the experimental test on radial tool-holder 1.

The region corresponding to the first 25 h of defect-free operation, which was used for training the SAE, is highlighted in purple. During this phase, the SAE accurately reconstructs the input features, reflecting the learned normal behavior. After the training period, deviations between actual and reconstructed values become noticeable. This

indicates the onset of degradation, which becomes more pronounced after 60 h. Specifically, RMS, AA, P2P, and MSA show substantial increases, while GF and MSF decrease—both trends consistent with progressive bearing wear.

Health Indicator

The health indicator is derived from the SAE's reconstruction error, quantified using the MSE as defined in Equation (3). Figure 7 shows the evolution of the health indicator over time for tool-holder 1. Three data-driven decision thresholds were calibrated offline from the empirical cumulative distribution function of the health indicator recorded during the complete 80 h run-to-failure test of tool-holder 1. The lower, nominal, and upper thresholds correspond to the 90th, 93rd, and 95th percentiles of this distribution, respectively. Since the complete degradation trajectory of tool-holder 1 was used to establish these percentile levels, tool-holder 1 must be regarded as the development and threshold-calibration case rather than as an independent validation case. Once calibrated, the three thresholds were frozen and subsequently applied without modification to tool-holder 2. Therefore, tool-holder 2 provides the evaluation of the fixed SAE–threshold configuration on an unseen component operating under the same experimental conditions. The term "online" does not refer to the determination or adaptation of the percentile thresholds during operation. It refers specifically to the sequential update of the Exponential Degradation Model and of the associated RUL estimate using only the health-indicator observations available up to the current monitoring instant. Thus, the thresholds are established offline before deployment, whereas the degradation parameters and RUL estimate are updated online. This separation between development and fixed-parameter evaluation is consistent with recent bearing-prognosis studies in which model generalization is assessed on distinct test bearings while the developed parameters remain unchanged [36]. The percentile levels represent progressively less conservative threshold-crossing criteria. Nevertheless, their numerical values are specific to the present development dataset and operating condition. For deployment on other machines or under different operating regimes, the thresholds would require offline calibration using representative historical run-to-failure data or independent engineering and risk-based acceptance criteria. During the first 50 h, the MSE remains consistently low, reflecting stable operation. Around 50–60 h, small fluctuations appear, marking the early stages of degradation. After 60 h, MSE values rise sharply, indicating significant wear. By hour 75, the health indicator exceeds the upper threshold, confirming failure.

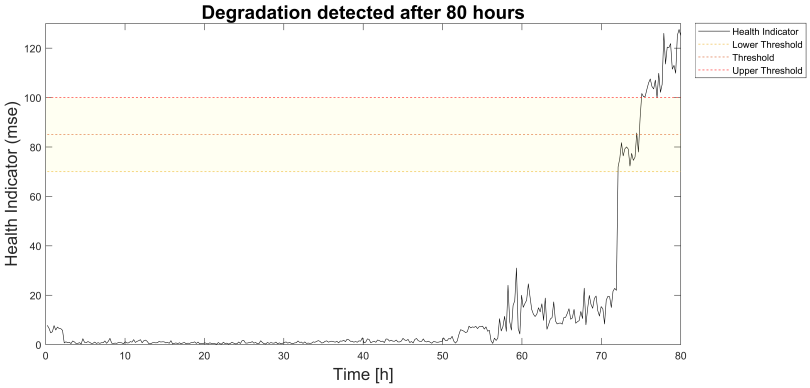


Figure 7. Health indicator (MSE) over time for tool-holder 1, with threshold levels indicating operational boundaries.

3.4. Procedure for Remaining Useful Life Estimation

The following algorithm summarizes the online inference stage used in this work to transform sequential vibration measurements into an updated estimate of a tool-holder bearing's RUL. Before online monitoring, the SAE parameters, the feature-normalization statistics, and the three percentile-based decision thresholds are established offline using the development dataset from tool-holder 1. During online monitoring, these quantities remain fixed. Only the health-indicator history and the parameters of the Exponential Degradation Model are updated whenever a new observation becomes available. Consequently, the RUL estimate at the current monitoring instant is calculated using only the observations acquired up to that instant and does not use future measurements from the monitored tool-holder. Unless otherwise stated, the symbols follow the notation introduced in Section 3.

I. **Signal acquisition and pre-processing**

(I.a) Acquire the acceleration signal $s(t)$ at a sampling frequency of 25.6 kHz.

(I.b) Divide the stream into non-overlapping windows of 1 s (25,600 samples per window).

(I.c) Retain a single 1 s window every 10 min and discard the intermediate ones.

II. **Feature extraction (per window)**

(II.a) Compute the six scalar features listed in Table 2: RMS, AA, $P2P$, MSA, GF, MSF.

(II.b) Standardize each feature with the mean μ_i and standard deviation σ_i obtained from the defect-free training set: $x_{in}^{\text{norm}} = (x_{in} - \mu_i)/\sigma_i$.

III. **Health-indicator generation with the SAE**

(III.a) Feed the normalized feature vector $\mathbf{x}_n^{\text{norm}} \in \mathbb{R}^6$ to the trained SAE (architecture in Figure 5).

(III.b) Reconstruct $\hat{\mathbf{x}}_n = \hat{\mathbf{x}}_n^{\text{norm}}\sigma_i + \mu_i$ and compute the mean-squared reconstruction error:

$$hi_n = \frac{1}{6}\sum_{i=1}^{6}\left(x_{in} - \hat{x}_{in}\right)^2 \tag{7}$$

(III.c) Append hi_n to the cumulative *history* $\mathcal{H} = \{hi_1, hi_2, \ldots, hi_n\}$ (a growing log; no elements are ever discarded).

IV. **Online fitting of the Exponential Degradation Model**

(IV.a) Perform the logarithmic transformation $z_n = \ln(hi_n)$ for every $hi_n \in \mathcal{H}$.

(IV.b) Estimate the parameters θ and v in $hi(t) = \theta e^{vt}$ via the closed-form least-squares solution of Equations (5) and (6).

V. **Prediction of threshold-crossing time**

(V.a) Given an alarm threshold hi_{thr}, solve $hi_{\text{thr}} = \theta e^{vt_{\text{fail}}}$ for t_{fail}:

$$t_{\text{fail}} = \frac{1}{v}\ln\left(\frac{hi_{\text{thr}}}{\theta}\right). \tag{8}$$

(V.b) Let t_{now} be the elapsed time since the start of monitoring. The RUL is then:

$$RUL(t_{\text{now}}) = t_{\text{fail}} - t_{\text{now}}, \qquad RUL \geq 0. \tag{9}$$

VI. **Decision logic**

(VI.a) Iterate over steps I–V continuously. If RUL drops below the maintenance horizon (e.g., $RUL < 5$ h), schedule an intervention.

(VI.b) Compute RUL at multiple thresholds (lower, nominal, upper) to obtain conservative, nominal, and optimistic forecasts.

Algorithm 1 provides pseudocode reflecting the strategy above.

Algorithm 1 Online RUL Estimation with Growing History

Load the trained SAE, the fixed training statistics μ_i, σ_i, and the percentile thresholds hi_{thr} calibrated offline on tool-holder 1.
Initialise the online health-indicator history $\mathcal{H} \leftarrow \varnothing$.
for each 1s window **do**
    Acquire vibration signal.
    Extract and standardise six features $\mathbf{x}_n^{\text{norm}}$.
    Feed $\mathbf{x}_n^{\text{norm}}$ to the trained SAE
    $hi_n \leftarrow$ Compute the MSE using Equation (7)
    Append hi_n to $\mathcal{H}$.
    Estimate θ and v by applying log–linear least squares to the health-indicator observations available in the current history $\mathcal{H}$.
    Compute t_{fail} and RUL with Equations (8) and (9).
    if $RUL <$ maintenance horizon **then**
        Trigger alert and schedule maintenance.
    end if
end for

4. Results and Discussion

This section presents the results of the proposed SAE–Exponential Degradation Model framework. Tool-holder 1 is first examined as the development and threshold-calibration case. The complete fixed configuration is then applied to tool-holder 2, without SAE retraining, modification of the normalization statistics, or recalibration of the percentile thresholds, to provide a within-domain evaluation on an unseen component. The SAE is subsequently compared with three baseline models of increasing complexity: Principal Component Analysis (PCA), a Naive Constant Model (NCM), and a supervised Long Short-Term Memory (LSTM) network. The comparison considers prediction accuracy, sensitivity to the selected threshold, data requirements, and computational efficiency. Finally, the external applicability of the complete framework is investigated using the public PRONOSTIA bearing dataset, which provides multiple run-to-failure experiments acquired under three speed–load combinations.

4.1. Exponential Degradation Model for Remaining Useful Life Estimation

The RUL estimation process goes beyond simply detecting anomalies; it aims to predict the rate of degradation and estimate how long the bearing can operate before failure occurs [37]. This is achieved by integrating anomaly detection methods with degradation models, such as the Exponential Degradation Model. By applying this approach to the health indicator, the system can forecast the bearing's operational lifetime and enable proactive maintenance interventions.

Tool-holder 1 was used both for SAE development and for the offline calibration of the 90th, 93rd, and 95th percentile thresholds. Consequently, the results reported for this tool-holder represent development and calibration performance and must not be interpreted as an independent evaluation of threshold generalization. Nevertheless, at each prediction instant, the Exponential Degradation Model was fitted only to the health-indicator observations acquired up to that instant. Therefore, the sequential RUL-update procedure remains causal even though the threshold levels were established retrospectively during the offline development stage.

The Exponential Degradation Model predicts the progression of a health indicator over time, in this case, the MSE obtained from the SAE. The model assumes that degradation follows an exponential trend, and the LSM is employed to estimate the coefficients of the exponential function that best fits the observed data. These coefficients are sequentially updated as new health-indicator values become available. Therefore, although the percentile thresholds are established during the offline calibration stage, the degradation fit and the corresponding RUL estimate are updated online using the growing measurement history. The model was evaluated using data from radial tool-holder 1, which developed a pitting defect after 80 h. By fitting an exponential curve to the evolving MSE, the model predicted the RUL and adjusted its forecast in response to accelerating degradation. Predictions were made using three thresholds—lower, nominal, and upper—offering varying levels of conservativeness.

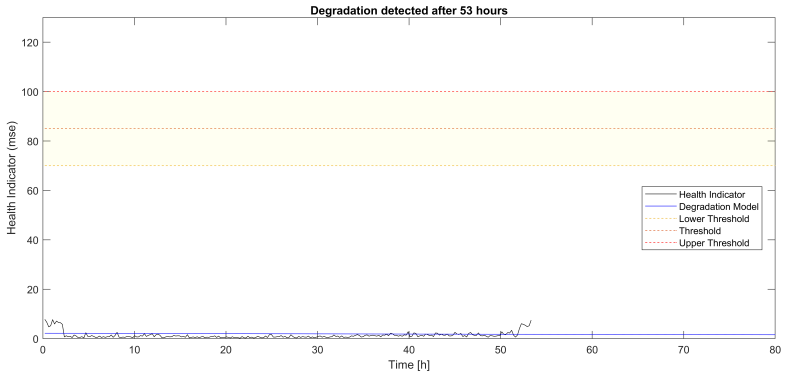
Figure 8 illustrates the evolution of the exponential degradation model as it updates over time in response to new data. The health indicator trend (black line) for radial tool-holder 1 is displayed alongside the exponential degradation model's prediction (blue line), which forecasts the future trajectory of the health indicator. Each subfigure represents the state of the degradation model at a specific times. In Figure 8a, after 53 h, the model predicts a gradual increase in the health indicator, indicating early degradation while remaining within acceptable operational limits. In Figure 8b, after 64 h, the model shows a more pronounced rise in the health indicator, signaling increased wear. In Figure 8c, at 80 h, the model accurately captures the sharp increase in the health indicator (MSE), reflecting critical degradation shortly before the tool-holder reaches failure.

Throughout these figures, the exponential degradation model demonstrates its adaptability to real-time data. In the early stages of testing, the model forecasts minimal degradation, reflecting the stable health of the tool-holder. As the health indicator begins to increase over time, the model dynamically adjusts, predicting a steeper degradation curve. This responsiveness highlights the model's capability to continuously update its predictions based on the observed data, providing an evolving forecast of the system's health. In the final stages of the test, the exponential degradation model fully adapts to the deteriorating condition of the tool-holder, accurately forecasting the increasing health indicator values and predicting that the system will soon reach the upper threshold, signaling imminent failure. The model's ability to offer real-time updates supports predictive maintenance strategies by providing operators with insights into the future state of the tool-holder. In a practical setting, operators can use the model's predictions to anticipate when the bearing is likely to cross critical thresholds, allowing for proactive maintenance or replacement scheduling, thus minimizing downtime and preventing costly failures.

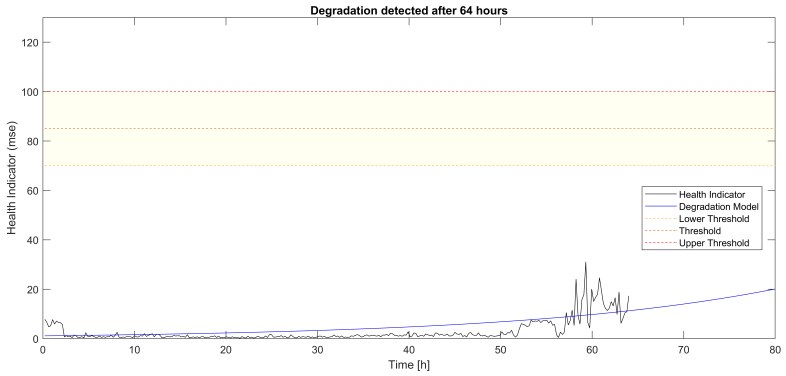
To evaluate the model's performance, Figure 9 compares the RUL of the tool-holder with the RUL predictions generated by the exponential degradation model at various threshold levels. The x-axis represents time in hours, while the y-axis displays the RUL on a logarithmic scale (in hours). The True RUL (black solid line) is calculated as the difference between the current time and the failure time at 80 h. The Predicted RULs for the Lower Threshold (yellow line), Nominal Threshold (orange line), and Upper Threshold (red line) are derived by estimating the time difference between the current time and the predicted failure time, determined by the intersection of the exponential degradation model with each respective threshold. In the early phase (0–53 h), the system shows no signs of degradation, and the model assumes infinite RUL, illustrated by the grey zone. Around 53 h, all three predicted RUL curves begin to diverge from the infinite line, reflecting early signs of wear. From 53 to 70 h, the predictions become more refined, with decreasing variation between thresholds. After 70 h, the degradation accelerates, leading to a sharp decrease in the true RUL, indicating that the tool-holder is approaching failure. The predicted RUL curves

(yellow, orange, and red) follow this sharp decline, with the lower threshold providing the most conservative estimate (reaching near-zero RUL first). By the end, all three predicted RUL curves converge closely with the true RUL, demonstrating that the model accurately captured the rapid degradation and imminent failure of the tool-holder bearing.

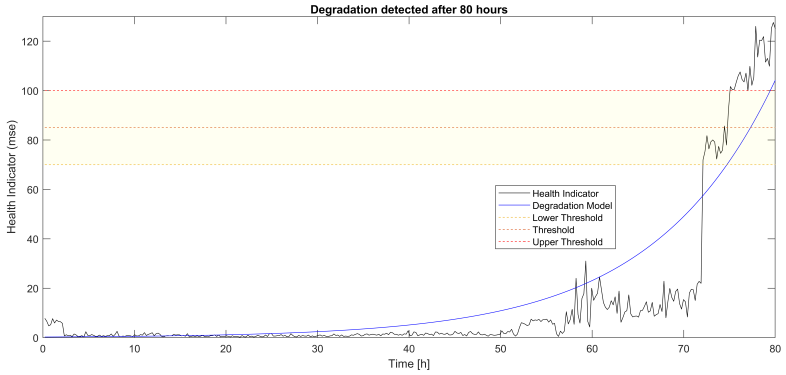
The use of different thresholds (yellow, orange, and red) provides flexibility in decision-making: the lower threshold offers a conservative prediction, while the upper threshold helps maximize tool-holder usage. This multi-threshold approach enables more tailored maintenance strategies depending on the operational priorities.



(**a**) Prediction after 53 h of testing.



(**b**) Prediction after 64 h of testing.



(**c**) Prediction at the end of the test (80 h).

Figure 8. The exponential degradation model fitted the MSE trend for radial tool-holder 1 at various stages.

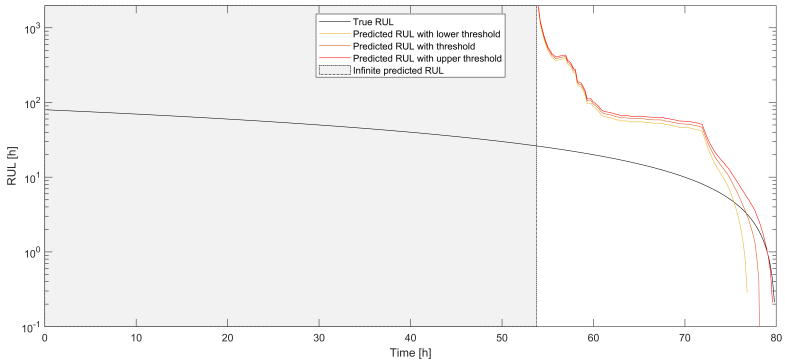


Figure 9. RUL predictions vs. actual RUL for radial tool-holder 1.

4.2. Evaluation on a Second Case Study

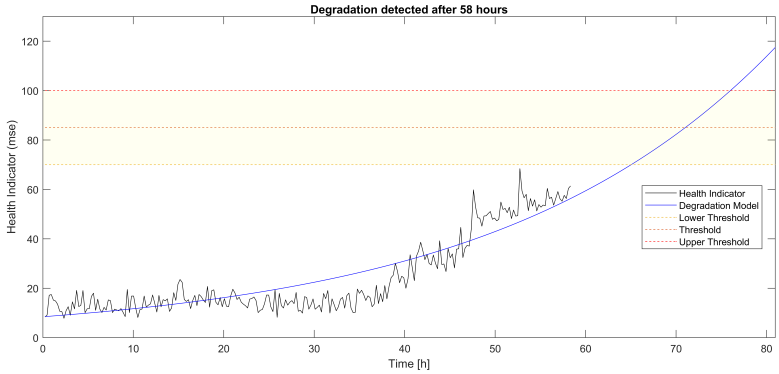
Tool-holder 2 was used to evaluate the complete fixed framework on an unseen component operating under the same experimental conditions. This tool-holder reached the experimental end-of-life condition after 82 h and exhibited an inner-ring defect during the post-test inspection. Neither the SAE parameters, the feature-normalization statistics, nor the 90th, 93rd, and 95th percentile thresholds were recalibrated using data from tool-holder 2. The SAE and percentile thresholds developed from tool-holder 1 were applied unchanged. Only the parameters of the Exponential Degradation Model were updated sequentially as new health-indicator observations from tool-holder 2 became available. Consequently, the results for tool-holder 2 provide the within-domain evaluation of the fixed SAE–threshold configuration, whereas the results for tool-holder 1 represent the development and calibration stage.

Figure 10 represents how the exponential degradation model predicts the MSE trend for the radial tool-holder 2. Each subfigure represents the state of the degradation model at specific times. In Figure 10a, after 58 h, the model shows a modest increase in MSE, reflecting early signs of degradation. In Figure 10b, after 70 h, the model detects a significant rise in MSE, reflecting accelerated wear. In Figure 10c, after 82 h, the model accurately predicts a sharp increase in MSE just before failure, demonstrating its capability to track degradation even in a different tool-holder.

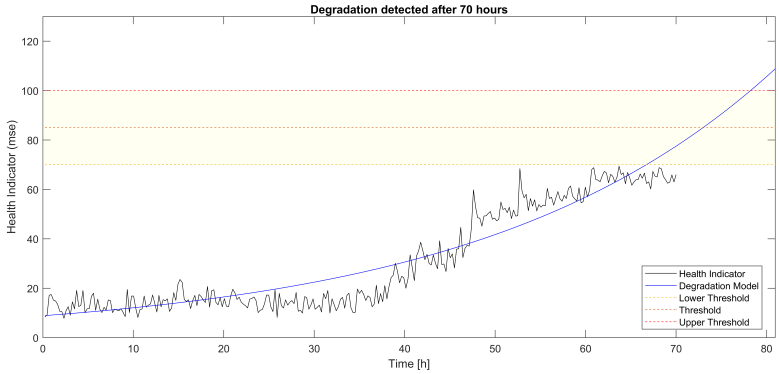
The fitted exponential model captures the transition from the slow initial evolution of the health indicator to its more rapid growth as the bearing approaches the experimental end-of-life condition. This sequential adaptation enables the framework to update the degradation forecast and the associated RUL estimate as new observations become available. The application to radial tool-holder 2 without SAE retraining or threshold recalibration provides preliminary evidence of within-domain transferability under the tested operating conditions. A broader assessment of transferability will require additional tool-holders and operating conditions, since the present evaluation was performed on a single unseen component belonging to the same experimental domain.

Figure 11 presents the predicted RUL versus actual RUL for tool-holder 2. As with the previous case, the grey area represents the initial interval during which the estimated exponential growth rate does not yield a finite threshold-crossing time. At approximately 12 h, the evolution of the health indicator allows the model to provide the first finite RUL estimates. Between approximately 12 and 35 h, however, the degradation trend remains weak and only partially developed. Under these conditions, the estimated exponential growth rate is sensitive to small changes in the cumulative health-indicator history, producing comparatively large variations in the predicted RUL. From an operational perspective, the predicted RUL values in this interval remain substantially higher than the maintenance horizon considered in Algorithm 1 and would therefore not trigger an intervention or a

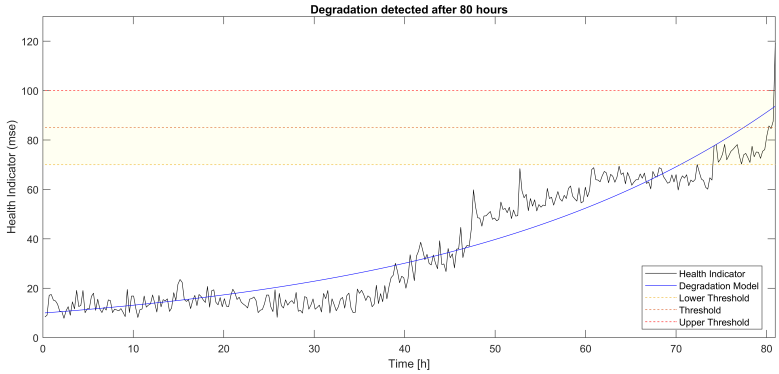
sudden machine shutdown. As further degradation observations become available, the health-indicator trend becomes more clearly defined and the RUL estimates decrease more consistently. After approximately 40 h, the predicted trajectories progressively approach the true RUL, with the lower threshold retaining the most conservative behavior.



(**a**) Prediction after 58 h of testing.



(**b**) Prediction after 70 h of testing.



(**c**) Prediction after 80 h of testing.

Figure 10. The exponential degradation model fitted the MSE trend for radial tool-holder 2 at various stages.

The rolling-element pitting observed in tool-holder 1 and the inner-ring raceway defect observed in tool-holder 2 are associated with different characteristic frequencies and therefore produce different detailed vibration signatures. Nevertheless, both damage mechanisms altered the global amplitude and spectral-energy descriptors introduced in Section 3.2. The proposed framework does not diagnose the fault location by matching individual BPFI, BPFO, BSF, or FTF components; instead, the SAE quantifies departures

of the multivariate feature vector from the learned healthy condition. Consequently, both degradation mechanisms increased the reconstruction-error health indicator and could be tracked using the same prognostic structure. The different RUL trajectories and prediction accuracies obtained for the two tool-holders nevertheless confirm that the detailed response remained dependent on the specific degradation process.

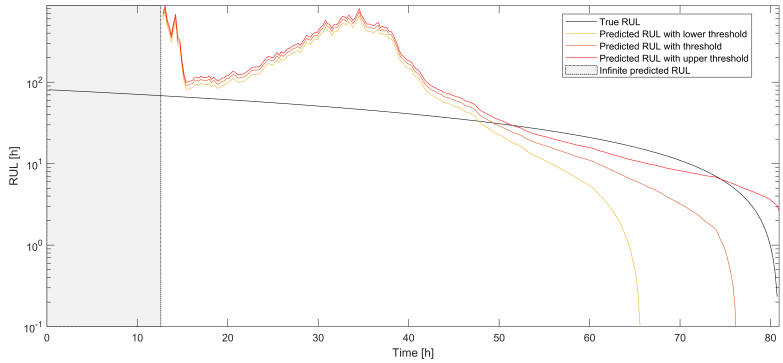


Figure 11. Predicted and true RUL for radial tool-holder 2.

Overall, the two case studies provide preliminary within-domain evidence that the proposed framework can support RUL estimation for the two observed damage modes under the investigated operating conditions. Equivalent performance for outer-race or cage defects, as well as robustness under different operating regimes, remains to be established through broader experimental validation.

4.3. Baseline Models

To rigorously evaluate the added value of the proposed SAE, three baseline models of increasing complexity were adopted for comparison: Principal Component Analysis (PCA), a Naive Constant Model (NCM), and a supervised Long Short-Term Memory (LSTM) network. These models were selected to provide complementary perspectives on the reconstruction and RUL prediction tasks, ranging from simple static assumptions to a high-capacity deep learning approach.

4.3.1. Model Descriptions

PCA represents a well-established, unsupervised method for linear dimensionality reduction. In this study, PCA was trained on the same defect-free dataset used for the SAE, corresponding to the first 25 h of operation of tool-holder 1. After normalization, feature vectors were projected into a three-dimensional latent space, matching the size of the SAE's bottleneck layer. Feature reconstruction was performed by applying the transposed PCA coefficient matrix to the latent representation, and the health indicator was computed as the reconstruction error. Despite its simplicity and low computational footprint, PCA provides a meaningful linear baseline against which more sophisticated models can be evaluated.

The NCM was introduced as a minimal benchmark that does not rely on training or encoding. This model assumes that the system remains indefinitely in its healthy state and reconstructs each input vector using the mean feature values computed from the first 25 h of defect-free operation. For each incoming sample, the health indicator is calculated as the mean squared deviation from this static reference. Although it cannot capture dynamic or nonlinear degradation, the NCM establishes a lower bound for reconstruction-based monitoring and helps assess whether more advanced methods, such as the SAE, learn non-trivial structure beyond simple averaging.

Finally, the LSTM model represents a fully supervised, high-capacity learning architecture designed to directly predict the RUL from sequences of vibration-derived features. The

network was trained using data collected exclusively from tool-holder 1. A sliding window strategy with a sequence length of 10 was adopted, such that each input sample consisted of ten consecutive time steps of six features, associated with the corresponding RUL value at the final step. To ensure coverage of the entire degradation trajectory, the dataset was randomly shuffled and split into 80% training and 20% testing subsets. The architecture comprised one recurrent layer with 32 hidden units, followed by a fully connected layer with 16 neurons and ReLU activation, and a regression output layer. The model was trained using the Adam optimizer for a maximum of 200 epochs, with a mini-batch size of 32 and an initial learning rate of 1×10^{-3}. The MSE was employed as the loss function.

The LSTM baseline operates in a supervised, full run-to-failure regime and directly outputs the RUL, whereas the SAE, PCA and NCM are trained on defect-free data and infer the End-of-Life (EoL) through a threshold-crossing mechanism. Despite this methodological difference, the LSTM is included to provide a supervised reference that reflects the best attainable performance in terms of average RUL accuracy when complete life-cycle data are available. This inclusion serves to contextualize the performance of the unsupervised SAE relative to a high-capacity, data-intensive alternative, without implying direct equivalence in the underlying learning objectives or prediction targets.

4.3.2. Comparative Results and Computational Analysis

Table 4 summarizes the performance of the SAE and the three baselines on tool-holder 1 and tool-holder 2 in terms of End of Life (EoL) and Root Mean Square Error (RMSE). Here, EoL denotes the predicted operational lifetime in hours up to a specified health-indicator threshold, while RMSE quantifies the deviation between predicted and true RUL trajectories over time. It should be noted that some entries in Table 4 are intentionally left blank ("–"). These correspond to cases where the model did not generate a valid EoL estimate because the predicted RUL did not converge toward zero within the observed time horizon.

Table 4. Comparison between SAE and baseline models in End of Life (EoL) prediction of tool-holder 1 and tool-holder 2. Note: "–" indicates that the model did not predict the EoL because the estimated RUL curve did not converge to zero within the observed degradation interval.

		Tool-Holder 1			**Tool-Holder 2**		
Model	**Threshold**	**True EoL [h]**	**Pred. EoL [h]**	**RMSE [h]**	**True EoL [h]**	**Pred. EoL [h]**	**RMSE [h]**
SAE	lower	80	76.59	0.6718	82	65.33	0.9559
	nominal	80	78.08	0.6722	82	75.83	0.5397
	upper	80	79.36	0.6806	82	–	0.2076
PCA	lower	80	78.93	0.6912	82	59.50	0.9417
	nominal	80	–	0.7551	82	64.87	0.4436
	upper	80	–	0.8157	82	75.13	0.2713
NCM	lower	80	–	1.0428	82	–	1.5613
	nominal	80	–	1.1082	82	–	1.6522
	upper	80	–	1.1722	82	–	1.7408
LSTM	–	80	–	0.0892	82	–	0.1361

Results in Table 4 indicate that the SAE consistently outperformed PCA and NCM, producing EoL estimates close to the true failure points and RMSE values below 0.7 h in most cases. For tool-holder 1, SAE predictions were accurate across all thresholds, whereas PCA was reliable only at the lower threshold and failed at the nominal and upper levels.

As expected, the NCM performed worst overall, as its static formulation could not capture progressive degradation, leading to substantially higher RMSE across both tool-holders.

The LSTM achieved the lowest RMSE values, reaching 0.0892 h for tool-holder 1 and 0.1361 h for tool-holder 2, demonstrating the ability of supervised sequence learning to capture complex degradation patterns. These results must be interpreted in light of the fundamental differences in training requirements between LSTM and SAE. The SAE was trained exclusively on the first 25 h of healthy operation, requiring no information about the degradation trajectory, whereas the LSTM relied on 80% of the complete life-cycle data of tool-holder 1, including the final degradation stage that provides the most informative dynamics for RUL prediction. As a supervised model with a more complex architecture, the LSTM depends heavily on large and comprehensive datasets to avoid overfitting. The acquisition of such full life-cycle datasets is highly costly, as each endurance test requires approximately 80 h of operation, incurring significant expenses in terms of time, component wear, and energy consumption. This makes approaches such as the SAE particularly advantageous, as they require only short defect-free baselines yet still provide accurate and robust prognostics.

However, although the LSTM yielded the smallest RMSE values, this metric alone does not guarantee a reliable prediction of the EoL. The RMSE evaluates the average deviation between predicted and true RUL values but remains insensitive to whether the model can capture the final convergence of the degradation trajectory. From a prognostic perspective, anticipating the point at which the RUL collapses to zero is more critical than achieving a minimal average error. As illustrated in Figure 12, the LSTM predictions remained close to the true RUL trajectory during most of the degradation process but did not exhibit the rapid convergence toward zero near failure. In contrast, the SAE successfully predicted the complete degradation trend and provided physically consistent EoL estimates, despite a slightly higher RMSE.

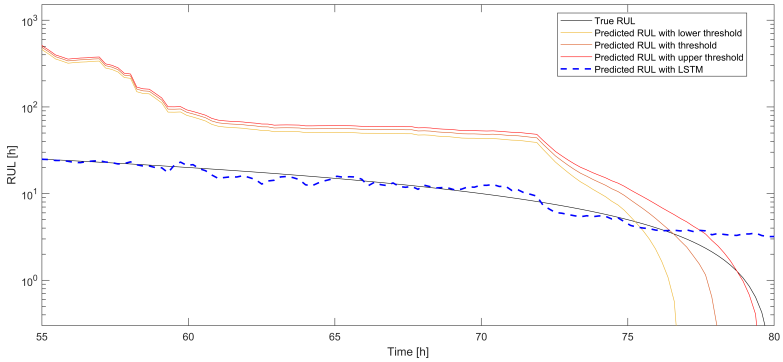


Figure 12. Comparison between the true RUL trajectory and the predicted RUL obtained with the proposed SAE model and the LSTM baseline during the final operating phase of tool-holder 1.

In addition to predictive accuracy, model compactness was assessed in terms of the number of trainable parameters (Params) and the operation count at inference (MACs), as summarized in Table 5. To provide a fair, hardware-agnostic comparison, counts refer to the complete RUL estimation pipeline. For SAE, PCA, and NCM, this includes (i) model traversal to obtain the health indicator and (ii) the exponential degradation model update used to produce the RUL estimate. For the LSTM, only the network forward pass is considered, as it directly outputs the RUL. Computational cost is expressed as multiply–accumulate operations (MACs), with isolated additions regarded as negligible unless otherwise stated.

For the SAE, the encoder and decoder comprise 6×3 and 3×6 fully connected layers, yielding 45 trainable parameters and approximately 36 MACs per forward pass. The

PCA baseline uses the same 6↔3 projection/back-projection dimensionality and requires approximately 36 MACs with no trainable parameters at inference. The NCM computes a mean-squared deviation from the healthy reference and entails about 12 MACs to form the per-sample indicator.

The exponential degradation model is fitted on the growing history $\mathcal{H}$ through a closed-form log–linear least-squares solution. With a sampling scheme of one 1 s window every 10 min over an 80 h lifetime, the history contains at most $N \approx 480$ samples. Computing the sufficient statistics $\sum t_n z_n$ and $\sum t_n^2$ requires approximately $2N$ MACs per update, while $\sum t_n$ and $\sum z_n$ involve additions only. Solving for the slope/intercept and computing the threshold-crossing time adds a negligible constant overhead. Hence, the exponential degradation model update requires roughly 960 MACs per prediction at $N \approx 480$. In practice, maintaining running sums reduces this cost to $O(1)$ per step; the reported values therefore represent an upper-bound scenario.

The supervised LSTM uses an input size of 6, a hidden size of 32, and a sequence length of 10, resulting in 5537 trainable parameters and approximately 50,448 MACs per prediction (including the recurrent core and the two fully connected layers). These figures substantiate the characterization of the proposed framework as computationally efficient in an *end-to-end* sense, since the exponential degradation model overhead remains two orders of magnitude smaller than the LSTM forward pass.

Table 5. Hardware-agnostic complexity metrics for end-to-end RUL inference. Params: trainable parameters. MACs: multiply–accumulate operations per prediction. EDM: exponential degradation model.

Model	Params	MACs (Model)	MACs (Total = Model + EDM)
SAE	45	∼36	∼996
PCA	0	∼36	∼996
NCM	0	∼12	∼972
LSTM	5537	∼50,448	∼50,448

Notes. (i) Feature extraction is shared by all methods and therefore omitted. (ii) Constant-cost transcendental functions (e.g., ln) are not counted in MACs.

A complementary analysis of computational efficiency is reported in Table 6. The reported values correspond to the overall time required to estimate the RUL for a single data sample. For SAE and PCA, this time includes both model traversal (feature compression and reconstruction) and the subsequent application of the exponential degradation model. For the NCM, the reported time reflects the computation of the health indicator as the deviation from the mean reference vector, followed by exponential degradation modeling. In the case of the LSTM, the reported traversal time directly corresponds to the complete RUL estimation, since the network output provides the prediction without the need for an external degradation model. All samples from tool-holder 1 and tool-holder 2 were processed individually for this analysis. Measurements were performed on a desktop workstation equipped with an Intel CPU (x86 Family 6 Model 158 Stepping 10, GenuineIntel) and 5.48 GB of RAM, running Windows 10 Version 10.0.26100.4652, and using MATLAB R2018b. All computations were executed on the CPU only, without GPU acceleration or parallel computing enhancements, in order to reflect a standard industrial deployment scenario. The results indicate that SAE, PCA, and NCM achieve traversal times within a fraction of a millisecond per sample, confirming their suitability for real-time deployment. By contrast, the LSTM required several milliseconds per sample, reflecting its sequential processing and higher architectural complexity. These findings further emphasize that the SAE strikes a favorable balance between predictive accuracy and computational efficiency,

making it a practical and effective solution for industrial prognostics where complete run-to-failure datasets are rarely available.

Table 6. Comparison of computational times (compression and reconstruction) for SAE and baseline models.

Model	Average Time [ms]	Standard Deviation [ms]	Median Time [ms]	Minimum Time [ms]	Maximum Time [ms]
SAE	0.0533	0.1707	0.0334	0.0152	2.5038
PCA	0.0522	0.1537	0.0331	0.0150	2.4460
NCM	0.0485	0.1361	0.0314	0.0135	1.9679
LSTM	9.2705	2.2151	8.7966	8.4237	24.4075

4.4. External Validation on the PRONOSTIA Dataset

To assess the applicability of the proposed framework beyond the proprietary toolholder experiments, an additional evaluation was performed using the PRONOSTIA bearing dataset [38]. PRONOSTIA is a publicly available accelerated bearing run-to-failure benchmark released in the context of the IEEE PHM 2012 Prognostic Challenge and associated with the FEMTO-ST Institute. The dataset files and the original challenge documentation are available through public repositories, since the original FEMTO-ST challenge page is no longer active (https://github.com/wkzs111/phm-ieee-2012-data-challenge-dataset, accessed on 7 July 2026). The dataset contains 17 accelerated run-to-failure bearing experiments distributed across three constant operating conditions: 1800 rpm and 4000 N; 1650 rpm and 4200 N; and 1500 rpm and 5000 N. The official dataset division includes six complete learning bearings, with two bearings for each operating condition, and eleven test bearings whose vibration histories are truncated before the actual EoL. Each vibration record contains 0.1 s of acceleration data sampled at 25.6 kHz and acquired at 10 s intervals. In the present analysis, the vertical acceleration channel was used.

A separate condition-specific model was developed for each of the three operating conditions using only the corresponding two complete learning bearings. The first 10% of each learning-bearing trajectory was assumed to represent healthy operation and was used to calculate the feature-normalization statistics and train the SAE. The complete trajectories of the same learning bearings were subsequently used to calibrate the 90th, 93rd, and 95th percentile thresholds. The resulting SAE parameters, normalization statistics, and thresholds were then frozen before evaluation on the corresponding test bearings.

The PRONOSTIA evaluation was conducted without benchmark-specific tuning. No test-bearing data were used for SAE training, hyperparameter selection, threshold calibration, or component-specific adjustment. For each of the eleven test bearings, the health indicator and Exponential Degradation Model were updated using only the measurements available up to the official truncation instant. The predicted EoL was calculated by adding the estimated RUL to the operating time already elapsed at that instant.

Table 7 reports the actual and predicted EoL values obtained using the three percentile thresholds. The error ΔEoL_{P93} is defined as the nominal predicted EoL minus the actual EoL. Negative values therefore indicate conservative predictions, whereas positive values indicate optimistic predictions.

Table 7. Actual and predicted End of Life at the official PRONOSTIA test-set truncation instant. The symbol ∞ indicates that the estimated degradation rate did not produce a finite threshold-crossing time.

Operating Condition	Bearing	Actual EoL [h]	Pred. EoL P90 [h]	Pred. EoL P93 [h]	Pred. EoL P95 [h]	ΔEoL_{P93} [h]
1800 rpm 4000 N	Bearing1_3	6.5944	3.3393	3.6433	4.2331	−2.9511
	Bearing1_4	3.2553	1.8162	1.9606	2.2407	−1.2947
	Bearing1_5	6.8389	9.4149	10.2943	11.9998	+3.4554
	Bearing1_6	6.7972	11.7904	12.9649	15.2431	+6.1677
	Bearing1_7	6.2722	4.4696	4.8642	5.6297	−1.4080
1650 rpm 4200 N	Bearing2_3	5.4278	∞	∞	∞	–
	Bearing2_4	2.0833	5.3017	5.4042	5.5630	+3.3209
	Bearing2_5	6.4167	∞	∞	∞	–
	Bearing2_6	1.9444	4.5122	4.6175	4.7805	+2.6731
	Bearing2_7	0.6361	0.4846	0.4945	0.5098	−0.1416
1500 rpm 5000 N	Bearing3_3	1.2028	0.3561	0.4401	0.5347	−0.7627

The corresponding graphical comparison is shown in Figure 13.

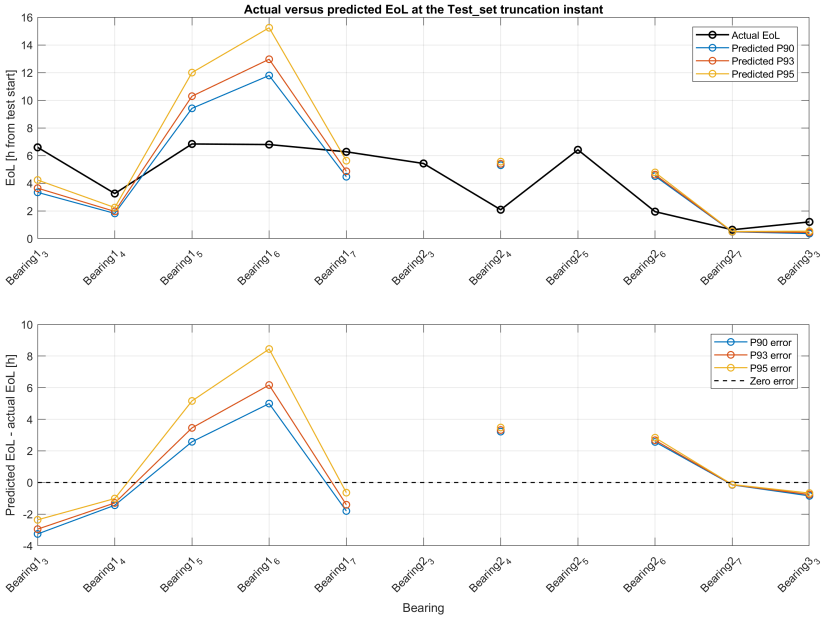


Figure 13. Comparison between the actual and predicted EoL values at the official truncation instant of the eleven PRONOSTIA test bearings. The lower panel reports the prediction error, defined as predicted EoL minus actual EoL. Missing predictions correspond to infinite threshold-crossing estimates.

The framework produced a finite EoL estimate for nine of the eleven test bearings, corresponding to a prediction coverage of 81.8%. Considering the nominal P93 threshold, five of the nine finite estimates were conservative and four were optimistic. The median signed EoL error was −0.14 h, indicating a slightly conservative central tendency, whereas the mean signed error was +1.01 h because of a limited number of larger optimistic errors. Particularly close conservative predictions were obtained for Bearing2_7, for which the nominal estimate anticipated the actual EoL by approximately 0.14 h, and for Bearing1_7, for which the P95 estimate anticipated the actual EoL by approximately 0.64 h.

The two non-finite estimates were obtained for Bearing2_3 and Bearing2_5, both belonging to the second operating condition. At the official truncation instant, the available

health-indicator histories did not provide a positive degradation rate sufficient to predict a finite intersection with the calibrated thresholds.

Recent PRONOSTIA studies have reported lower RUL errors by employing supervised deep architectures, extensive time–frequency feature engineering, degradation-stage segmentation, and model-specific optimization [39–41]. Compared with these tuned approaches, the error margins obtained by the present lightweight framework are larger. Nevertheless, the comparison must account for the substantially different modeling objectives and data requirements. The present approach was trained using only an initial healthy portion of the learning bearings, retained the same SAE architecture and percentile rules used for the tool-holder analysis, and was applied to the eleven test bearings without test-set tuning or component-specific recalibration. The external evaluation therefore provides additional evidence that the proposed framework can be applied beyond the two proprietary tool-holders without benchmark-specific tuning, while retaining its lightweight, interpretable, and data-efficient formulation.

5. Limitations and Future Work

The experimental campaign was limited to two tool-holders tested under identical operating conditions. Moreover, the 90th, 93rd, and 95th percentile thresholds were calibrated retrospectively using the complete run-to-failure trajectory of tool-holder 1. These thresholds must therefore be interpreted as application-specific offline calibration parameters rather than as universal limits or quantities determined autonomously during online monitoring.

Tool-holder 1 represents the development and calibration case, whereas tool-holder 2 is the only proprietary component evaluated with the SAE parameters, normalization statistics, and thresholds fixed in advance. For industrial deployment, the thresholds should be established offline using a larger and representative development population, independent run-to-failure histories, or engineering and risk-based acceptance criteria.

The additional PRONOSTIA analysis broadened the external assessment to eleven test bearings and three speed–load combinations, thereby extending the empirical evidence beyond the original proprietary dataset. Nevertheless, further experimental validation on additional tool-holders remains necessary to establish the statistical robustness of the framework across a wider range of components and degradation modes.

Future work will therefore extend the proprietary experimental campaign to additional machines, operating conditions, and degradation modes, enabling a broader statistical assessment and a more robust calibration of the percentile levels. Although the PRONOSTIA dataset includes three different speed–load combinations, each bearing was tested under a constant operating regime and each condition was modeled separately. Consequently, the present results do not demonstrate that a single set of feature-normalization statistics and percentile thresholds remains directly applicable when speed, load, or feed rate vary dynamically during operation. Variations in the operating regime, together with cutting shocks and structural disturbances encountered during actual machining, may alter the distributions of the extracted features and of the resulting health indicator even in the absence of additional degradation. Future developments will therefore investigate operating-regime identification, condition-dependent feature normalization and percentile thresholds, and validation during transitions between operating conditions. Portability across domains and field deployments will also be investigated to align the prognostic outputs with production and risk constraints.

6. Conclusions

This study presented a lightweight framework for tool-holder bearing prognosis based on a Sparse Autoencoder combined with an Exponential Degradation Model. The SAE was trained only on the first 25 h of defect-free operation of tool-holder 1, using the reconstruction error of six vibration-derived features as a scalar health indicator. The condition thresholds were calibrated offline, whereas the degradation parameters and the corresponding RUL estimates were updated online through a sequential log–linear Least Squares procedure as new observations became available.

The main findings of the study can be summarized in the following four points:

1. The proposed framework was able to track two different degradation processes in the proprietary tool-holder experiments: rolling-element pitting in tool-holder 1 and an inner-ring raceway defect in tool-holder 2. The SAE, normalization statistics, and threshold configuration developed from tool-holder 1 were transferred to tool-holder 2 without retraining or recalibration. With the nominal P93 threshold, the predicted EoL values were 78.08 h and 75.83 h for actual test durations of 80 h and 82 h, respectively, with RUL RMSE values of 0.6722 h and 0.5397 h.
2. The comparison with PCA, NCM, and LSTM clarified the practical trade-off of the proposed method. The SAE-based framework provided more consistent degradation tracking than the simpler reconstruction-based baselines, while requiring only a short healthy baseline rather than labeled run-to-failure data. The supervised LSTM achieved the lowest RUL RMSE values, but required complete life-cycle information, 5537 trainable parameters, and an average inference time of 9.2705 ms per sample, compared with 45 trainable parameters and 0.0533 ms for the SAE-based framework.
3. The external evaluation on the PRONOSTIA dataset further assessed the applicability of the method beyond the proprietary tool-holder tests. Without test-set tuning or bearing-specific recalibration, finite EoL estimates were obtained for nine of the eleven test bearings, corresponding to a prediction coverage of 81.8% across three operating conditions. For the nominal P93 threshold, five predictions were conservative and four were optimistic, with a median signed EoL error of -0.14 h and a mean signed error of $+1.01$ h.
4. The results indicate that the SAE–Exponential Degradation Model can provide an interpretable and computationally efficient solution for condition-based RUL estimation. However, the method remains dependent on application-specific threshold calibration, and non-finite predictions may occur when the available health-indicator history does not exhibit a sufficiently clear positive degradation trend.

Overall, the proposed framework offers a practical compromise between predictive performance, data requirements, computational cost, and interpretability. Nevertheless, broader validation on additional tool-holders, operating regimes, and degradation modes is required before general industrial deployment.

Author Contributions: Conceptualization, G.D. (Giuseppe Dipace), E.M. and G.D. (Gianluca D'Elia); methodology, G.D. (Giuseppe Dipace), E.M. and G.D. (Gianluca D'Elia); software, G.D. (Giuseppe Dipace); validation, G.D. (Giuseppe Dipace), E.M. and G.D. (Gianluca D'Elia); formal analysis, G.D. (Giuseppe Dipace) and G.D. (Gianluca D'Elia); investigation, G.D. (Giuseppe Dipace); resources, G.D. (Giuseppe Dipace); data curation, G.D. (Giuseppe Dipace); writing—original draft preparation, G.D. (Giuseppe Dipace); writing—review and editing, E.M. and G.D. (Gianluca D'Elia); visualization, G.D. (Giuseppe Dipace); supervision, E.M. and G.D. (Gianluca D'Elia); project administration, E.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. The APC was waived by the publisher.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The PRONOSTIA dataset used for the external validation is publicly available at https://github.com/wkzs111/phm-ieee-2012-data-challenge-dataset, accessed on 9 July 2026. The experimental tool-holder data generated and analyzed during the present study are not publicly available owing to confidentiality and commercial restrictions of M.T. S.r.l.

Conflicts of Interest: Giuseppe Dipace is employed by M.T. S.r.l. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Bottou, L. Large-scale machine learning with stochastic gradient descent. In *Proceedings of the COMPSTAT'2010*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 177–186.
2. Si, X.; Wang, X.; Hu, C.; Zhou, D. Remaining useful life estimation—A review on the statistical data-driven approaches. *Eur. J. Oper. Res.* **2011**, *213*, 1–14. [CrossRef]
3. Li, S.; Liu, G.; Tang, X.; Lu, J.; Hu, J. An ensemble deep convolutional neural network model with improved D-S evidence fusion for bearing degradation diagnosis. *Sensors* **2017**, *17*, 1729. [CrossRef] [PubMed]
4. Hu, T.; Guo, Y.; Gu, L.; Zhou, Y.; Zhang, Z.; Zhou, Z. Remaining useful life prediction of bearings under different working conditions using a deep feature disentanglement based transfer learning method. *Reliab. Eng. Syst. Saf.* **2022**, *219*, 108265. [CrossRef]
5. Ding, Y.; Ma, L.; Ma, J.; Suo, M.; Tao, L.; Cheng, Y.; Lu, C. Intelligent fault diagnosis for rotating machinery using deep Q-network based health state classification: A deep reinforcement learning approach. *Adv. Eng. Inform.* **2019**, *42*, 100977. [CrossRef]
6. Liu, S.; Zhou, X.; Wang, Z. Adversarial domain adaptation for RUL prediction of industrial machines. *IEEE Trans. Ind. Inform.* **2021**, *17*, 2733–2742. [CrossRef]
7. Ding, Y.; Ding, P.; Zhao, X.; Cao, Y.; Jia, M. Transfer Learning for Remaining Useful Life Prediction Across Operating Conditions Based on Multisource Domain Adaptation. *IEEE/ASME Trans. Mechatron.* **2022**, *27*, 4143–4152. [CrossRef]
8. Lv, S.; Liu, S.; Li, H.; Wang, Y.; Liu, G.; Dai, W. A hybrid method combining Lévy process and neural network for predicting remaining useful life of rotating machinery. *Adv. Eng. Inform.* **2024**, *61*, 102490. [CrossRef]
9. Jia, Y.; Yao, D.; Yang, J.; Jia, L.; Zhou, T.; Long, Q. A multi-scale wavelet and spectral kurtosis-guided framework for noise-resistant rolling bearing fault diagnosis. *Mech. Syst. Signal Process.* **2026**, *242*, 113662. [CrossRef]
10. Zhong, D.; Guo, W.; He, D. An Intelligent Fault Diagnosis Method based on STFT and Convolutional Neural Network for Bearings Under Variable Working Conditions. In *2019 Prognostics and System Health Management Conference;* ACM: Qingdao, China, 2019; pp. 1–6.
11. Tapia, E.; Lopez-Novoa, U.; Sastoque-Pinilla, L.; López-de Lacalle, L.N. Implementation of a scalable platform for real-time monitoring of machine tools. *Comput. Ind.* **2024**, *155*, 104065. [CrossRef]
12. Xu, F.; Yang, F.; Fan, X.; Huang, Z.; Tsui, K.L. Extracting degradation trends for roller bearings by using a moving-average stacked auto-encoder and a novel exponential function. *Measurement* **2020**, *152*, 107371. [CrossRef]
13. She, D.; Jia, M.; Pecht, M.G. Sparse auto-encoder with regularization method for health indicator construction and remaining useful life prediction of rolling bearing. *Meas. Sci. Technol.* **2020**, *31*, 105005. [CrossRef]
14. Nie, L.; Zhang, L.; Xu, S.; Cai, W.; Yang, H. Remaining useful life prediction for rolling bearings based on similarity feature fusion and convolutional neural network. *J. Braz. Soc. Mech. Sci. Eng.* **2022**, *44*, 328. [CrossRef]
15. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507. [CrossRef] [PubMed]
16. Olshausen, B.A.; Field, D.J. Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vis. Res.* **1997**, *37*, 3311–3325. [CrossRef] [PubMed]
17. Ng, A. *Sparse Autoencoders*; CS294A Lecture Notes; Stanford University: Stanford, CA, USA, 2011.
18. Krogh, A.; Hertz, J.A. A simple weight decay can improve generalization. In Proceedings of the Advances in Neural Information Processing Systems, Denver, CO, USA, 2–5 December 1991; pp. 950–957.
19. Zhao, Y.; Nasrullah, Z.; Li, Z. PyOD: A Python Toolbox for Scalable Outlier Detection. *J. Mach. Learn. Res.* **2019**, *20*, 1–7. [CrossRef]
20. Erfani, S.M.; Rajasegarar, S.; Karunasekera, S.; Leckie, C. High-dimensional and Large-scale Anomaly Detection Using a Linear One-class SVM with Deep Learning. *Pattern Recognit.* **2016**, *58*, 121–134. [CrossRef]
21. Vincent, P.; Larochelle, H.; Bengio, Y.; Manzagol, P.A. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th International Conference on Machine Learning*; ACM: New York, NY, USA, 2008; pp. 1096–1103.

22. Aggarwal, C. *Outlier Analysis*; Springer International Publishing: Berlin/Heidelberg, Germany, 2016.
23. Zhou, C.; Paffenroth, R. *Anomaly Detection with Robust Deep Autoencoders*; Association for Computing Machinery: New York, NY, USA, 2017; pp. 665–674. [CrossRef]
24. Ahmed, M.; Mahmood, A.N.; Hu, J. A Survey of Network Anomaly Detection Techniques. *J. Netw. Comput. Appl.* **2016**, *60*, 19–31. [CrossRef]
25. Draper, N.R.; Smith, H. *Applied Regression Analysis*, 3rd ed.; Wiley: New York, NY, USA, 1998.
26. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed.; Springer: New York, NY, USA, 2009.
27. Montgomery, D.C.; Peck, E.A.; Vining, G.G. *Introduction to Linear Regression Analysis*, 5th ed.; Wiley: Hoboken, NJ, USA, 2012.
28. Weisberg, S. *Applied Linear Regression*, 3rd ed.; Wiley-Interscience: Hoboken, NJ, USA, 2005.
29. *ISO 17359:2018*; Condition Monitoring and Diagnostics of Machines—General Guidelines. International Organization for Standardization: Geneva, Switzerland, 2018.
30. *ISO 13373-1:2002*; Condition Monitoring and Diagnostics of Machines—Vibration Condition Monitoring—Part 1: General Procedures. International Organization for Standardization: Geneva, Switzerland, 2002.
31. Fernández-Lucio, P.; Gil Del Val, A.; Plaza, S.; Pereira, O.; Fernández-Valdivielso, A.; López de Lacalle, L.N. Threading holder based on axial metal cylinder pins to reduce tap risk during reversion instant. *Alex. Eng. J.* **2023**, *66*, 845–859. [CrossRef]
32. Behrendt, M.; de Angelis, M.; Comerford, L.; Zhang, Y.; Beer, M. Projecting interval uncertainty through the discrete Fourier transform: An application to time signals with poor precision. *Mech. Syst. Signal Process.* **2022**, *172*, 108920. [CrossRef]
33. *ISO 13373-2:2016*; Condition Monitoring and Diagnostics of Machines—Vibration Condition Monitoring—Part 2: Processing, Analysis and Presentation of Vibration Data. International Organization for Standardization: Geneva, Switzerland, 2016.
34. Li, H.; Wang, C. Combining First Prediction Time Identification and Time-Series Feature Window for Remaining Useful Life Prediction of Rolling Bearings with Limited Data. *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.* **2024**, *238*, 274–290. [CrossRef]
35. Wen, H.; Zhang, L.; Sinha, J.K. From Envelope Spectra to Bearing Remaining Useful Life: An Intelligent Vibration-Based Prediction Model with Quantified Uncertainty. *Sensors* **2024**, *24*, 7257. [CrossRef] [PubMed]
36. Xu, Z.; Chow, C.W.K.; Rahman, M.M.; Rameezdeen, R.; Law, Y.W. Remaining Useful Life Prediction Across Conditions Based on a Health Indicator-Weighted Subdomain Alignment Network. *Sensors* **2025**, *25*, 4536. [CrossRef] [PubMed]
37. Chandola, V.; Banerjee, A.; Kumar, V. Anomaly Detection: A Survey. *ACM Comput. Surv. CSUR* **2009**, *41*, 15. [CrossRef]
38. Nectoux, P.; Gouriveau, R.; Medjaher, K.; Ramasso, E.; Chebel-Morello, B.; Zerhouni, N.; Varnier, C. PRONOSTIA: An Experimental Platform for Bearings Accelerated Degradation Tests. In Proceedings of the IEEE International Conference on Prognostics and Health Management, Denver, CO, USA, 18–21 June 2012; pp. 1–8.
39. Dhungana, H.; Rykkje, T.; Lundervold, A.S. Bearing Prognostics Using the PRONOSTIA Data: A Comparative Study. *IEEE Access* **2025**, *13*, 49433–49456. [CrossRef]
40. Jin, X.; Ji, Y.; Li, S.; Lv, K.; Xu, J.; Jiang, H.; Fu, S. Remaining Useful Life Prediction for Rolling Bearings Based on TCN–Transformer Networks Using Vibration Signals. *Sensors* **2025**, *25*, 3571. [CrossRef] [PubMed]
41. Wei, X.; Fan, J.; Wang, H.; Cai, L. Remaining Useful Life Prediction Method for Bearings Based on Pruned Exact Linear Time State Segmentation and Time–Frequency Diagram. *Sensors* **2025**, *25*, 1950. [CrossRef] [PubMed]

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.