MethodsX
Open access

Article Template

MethodsX
Open access
Article Template

Article information

Article title

Enhanced Bearing Health Indicator Extraction Using Slope Adaptive Signal Decomposition for Predictive Maintenance.

Authors

Dev Bhanushali[a],Pooja Kamat[a], Harsh Dhiman[b]

Affiliations

[a] Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, Maharashtra, India
[b]Ernst & Young LLP India

Corresponding author's email address and Twitter handle

Pooja.kamat@sitpune.edu.in

Keywords

Predictive Maintenance; Bearing Health Monitoring; Signal Processing; Slope Adaptive Signal Decomposition (SASD); Remaining Useful Life (RUL); Machine Health Indicator;

Related research article

For a published article:

Abstract

*This study introduces the **Slope Adaptive Signal Decomposition (SASD)** algorithm, a novel method for extracting enhanced intrinsic machine health indicators from vibration data. Leveraging advanced signal processing techniques such as dynamic Savitzky-Golay filtering, segmentation, and trend-based recalibration, SASD achieves superior noise attenuation while preserving critical trends. Applied to the PRONOSTIA platform's bearing datasets, SASD produces refined health indicators suitable for predictive maintenance tasks, including Remaining Useful Life (RUL) estimation. The extracted features are evaluated using deep learning models like GRU, LSTM, and hybrid architectures, as well as conventional regression approaches, demonstrating SASD's effectiveness in improving prediction accuracy. Among these, GRU achieves the highest performance with near perfect prediction accuracy and the highest MAE scores with **0.0037** in condition 1 bearings, **0.0083** in conditions 2 bearings and **0.0127** in condition 3 bearings. This method bridges the gap between signal processing and data-driven prognostics, enabling robust bearing health monitoring under varying operational conditions.*

Highlights:

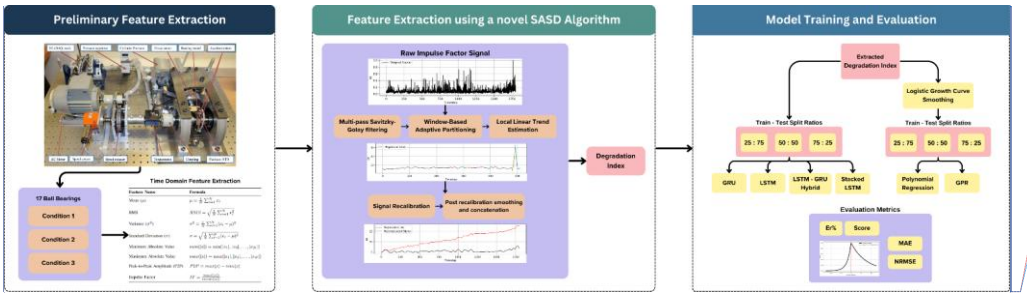
- *Dynamic signal smoothing and segmentation for enhanced trend extraction.*
- *SASD-based health indicators outperform traditional preprocessing techniques.*
- *Validated through RUL prediction using both statistical and deep learning models.*

Graphical abstract

Commented [PK1]: Add a line about best performing model and its accuracy achieved

Commented [DB2R1]: added

MethodsX
Open access
Article Template



Commented [PK3]: Are the train test split different ratios or this concept explained in the paper

Commented [DB4R3]: I have mentioned the 3 split ratios page 10 line 21 and added a new line about averaging results across the split ratios on page 11 line 17

Specifications table

Subject area	Engineering
More specific subject area	*Deep Learning*
Name of your method	*Slope Adaptive Signal Decomposition for bearing life estimation*
Name and reference of original method	*P. Nectoux et al., "PRONOSTIA : An experimental platform for bearings accelerated degradation tests," IEEE International Conference on Prognostics and Health Management, PHM'12., Jun 2012, Denver, Colorado, United States., pp. 1–8, 2012* [1]
Resource availability	*Data: Data available on request* *Software: Google Colab, Scipy library*

Background

Condition-based monitoring (CBM) has undergone profound changes from simple, traditional manual inspections to advanced automated methodologies [2], [3]. Initial CBM techniques were based on simple visual inspections and threshold-based parameter evaluations, like vibration and temperature, for determining the health of equipment [4], [5]. Deep learning has recently emerged as a game-changer in machinery health monitoring, especially when associated with data-driven approaches, such as Remaining Useful Life (RUL) prediction [6]. Deep learning models can independently learn to extract features and map complex temporal dependencies, making them highly effective in applications such as anomaly detection and RUL estimation [7], [8]. Recently, such models demonstrated superiority over classic techniques using predictive maintenance by learning complicated patterns in large datasets of operationally observed signals [9]. The accuracy and reliability of these models depend only on the quality of the input signals[10]. The two essential practices governing successful predictive maintenance are health monitoring and estimation of the Remaining Useful Life (RUL) for machine components [11], [12]. Bearings as a part of rotating machinery determine the reliability and efficiency of industrial systems [13]. Meanwhile, vibration signal analysis emerges as a powerful tool for diagnosing bearing health, but issues such as noise interference and trend distortion deter the extraction of meaningful indicators [14]. To achieve this, this paper introduces the Slope Adaptive Signal Decomposition (SASD) algorithm, a novel signal processing method that is capable of extracting enhanced health indicators from raw vibration data. The SASD algorithm is rich in dynamic smoothing, segmentation, and trend-based recalculation, ensuring noise attenuation while preserving trends. This method, applied to the PRONOSTIA platform[1], generates reliable health indicators. Advanced machine learning models then evaluate these indicators

for RUL estimation. This research aims to fill the gap between signal preprocessing and data-driven maintenance strategies by offering a scalable solution to bearing health monitoring with effectiveness in areas**.** Techniques like proposed SASD can bridge this gap, as they bring enhanced health indicators with advanced signal preprocessing, thereby paving the path for improved outcomes of CBM in different operational conditions. The SASD algorithm framework is showcased in Figure 1.

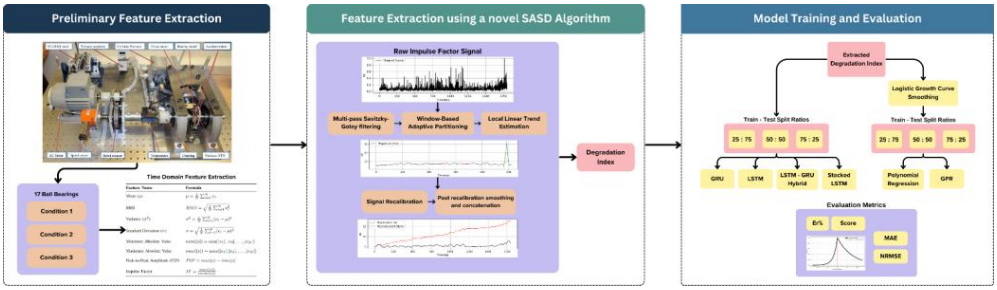


Figure 1: Slope Adaptive Signal Decomposition (SASD) algorithm

Method details

This section introduces the experimental description, dataset preparation as well as the proposed novel Slope Adaptive Signal Decomposition algorithm for extracting an enhanced intrinsic machine health indicator from vibration data by implementing a series of advanced signal processing techniques[15]. The core algorithm involves a succession of preprocessing, segmentation, trend extraction, and signal modification operations. Each stage is mathematically formulated to balance noise attenuation and trend preservation. The pseudocode corresponding to this method is outlined in **Algorithm 1** and is elaborated upon in detail in the subsequent subsections.

Experimental description and feature extraction

The bearing vibration signals used in this manuscript were collected from the PRONOSTIA platform [1]. There are a total of 17 bearing datasets across 3 different operating conditions as described in **Table 1.** The vibration data was collected through two accelerometers placed at 90◦ to each other on the bearing's external race at a sampling frequency of 25.6 kHz.

We start by extracting essential time domain features from the raw vibration signals of bearings. The vibration data is subjected to a non-overlapped sliding window mechanism where each window captures a segment of the horizontal and vertical components of the signals captured by the corresponding accelerometers. From each window, a series of statistical features are computed as described mathematically in Table 2. We will then process these extracted features further using our novel SASD algorithm which is described in detail in the next subsection.

TABLE 1: Datasets of PRONOSTIA platform

	Condition 1	**Condition 2**	**Condition 3**
Radial load (N)	4000	4200	5000
Speed (r/min)	1800	1650	1500

Training datasets	Bearing 1–1	Bearing 2–3	Bearing 3–1
	Bearing 1–2	Bearing 2–5	Bearing 3–2
Testing datasets	Bearing 1–3	Bearing 2–1	Bearing 3–3
	Bearing 1–4	Bearing 2–2	
	Bearing 1–5	Bearing 2–4	
	Bearing 1–6	Bearing 2–6	

TABLE 2: Extracted time domain features

Feature Name	Formula
Mean (σ)	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$
RMS	$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$
Variance (σ^2)	$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$
Standard Deviation (σ)	$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$
Minimum Absolute Value	$min(\lvert x\rvert) = \min(\lvert x_1\rvert, \lvert x_2\rvert, \dots, \lvert x_N\rvert)$
Maximum Absolute Value	$max(\lvert x\rvert) = \max(\lvert x_1\rvert, \lvert x_2\rvert, \dots, \lvert x_N\rvert)$
Peak-to-Peak Amplitude (P2P)	$P2P = max(x) - min(x)$
Impulse Factor	$IF = \frac{max(\lvert x\rvert)}{mean(\lvert x\rvert)}$

Algorithm 1: Signal Adaptive Slope Decomposition

Input:	$X = \{X_1, X_2, \dots, X_N\}$: Vibration signals over time. λ_{win}: Window factor for segmentation. λ_{pre}: Pre-smoothing factor. β_{thresh}: Threshold for slope detection. κ_{elev}: Elevation factor for recalibration. λ_{post}: Post-smoothing factor.
Output:	$S_{proc} = \{s_1, s_2, \dots, s_N\}$: Processed signal.
Apply Savitzky-Golay filter for smoothing:	$S_{smooth}(t) = \text{SGFilter}\left(X, \left\lfloor \frac{N}{\lambda_{pre}} \right\rfloor, 2\right)$
Determine window length T:	$T = \max\left(10, \left\lfloor \frac{N}{\lambda_{win}} \right\rfloor\right)$

Commented [PK5]: Verify all equations fully once again

Commented [DB6R5]: Verified correctness

Segment the smoothed signal into non-overlapping windows:	$W_i(t) = S_{\text{smooth}}(t_i : t_i + T)$
For each window Wi, perform linear regression to estimate local trends:	$W_{\text{reg}}(t) = \beta_i t + \alpha_i$
Recalibrate the signal:	$W_{recal}(t) = \begin{cases} W_i(t), & if\ \beta_i \geq 0 \\ \dfrac{W_i(t)}{\left(\dfrac{W_{reg}(t)}{\kappa_{elev}}\right)}, & if\ \beta_i < 0 \end{cases}$
Post-smooth the recalibrated signal:	$W_{filtered}(t) = \begin{cases} SGFilter\left(W_{recal}(t), \left\lfloor \dfrac{T}{\lambda_{post}} \right\rfloor\right), & if\ \beta_i < \beta_{thresh} \\ W_{recal}(t), & otherwise \end{cases}$
Adjust for continuity:	$S_{\text{recon}}(t) = W_{\text{filtered}}(t) + \left(S_{\text{recon}}(t_{i-1} + T - 1) - W_{\text{filtered}}(t_0)\right)$
Compile the processed signal:	$HI = \{S_{\text{recon}}(1), S_{\text{recon}}(2), \ldots, S_{\text{recon}}(N)\}$
Return:	HI

Slope Adaptive Signal Decomposition algorithm

Formalization and Multi-pass Savitzky-Golay filtering*[16]*

Let the raw vibration signal be represented by $(s(t))$ where $(s(t) \in R)$, and (t) denotes the time index of the signal over (N) discrete points. To remove the high frequency noise, the signal is preprocessed using a Savitzky-Golay filter iteratively to smoothen the data without distorting the original signal structure. This filter fits a low degree polynomial over a sliding window to approximate underlying signal dynamics. It is essentially a local polynomial smoother that can be mathematically formalized in the following manner. For some window length (w_{SG}), polynomial degree (p) and an interval (I_t) described as

$$I_t = \left[t - \frac{w_{SG}}{2}, t + \frac{w_{SG}}{2}\right] \quad (1)$$

the filter approximates signal $(s(t))$ at each timestep (t) by fitting a polynomial $(P(t) \in P_p)$ where (P_p) is the space of polynomials of degree (p) and minimizing the squared error over the window (I_t), formally

$$P(t) = \arg\min_{P \in \mathcal{P}} \sum_{\tau \in I_t} \left(s(\tau) - P(\tau)\right)^2 \quad (2)$$

Given this, the value of $(P(t))$ at each time (t) yields the smoothed signal. The window size, (w_{SG}) in the SASD algorithm is dynamically adapted based on the size of the signal and a pre-smoothing factor (λ_{pre}) in the following manner.

$$w_{SG} = \left\lfloor \frac{\lambda_{pre}}{N} \right\rfloor \quad (3)$$

The need for a dynamically updated window size will be discussed shortly. By applying the SG filter iteratively, we refine the denoising operation progressively through multiple passes. The filtered signal after (n)-th iteration is denoted as $(S_n(t))$. The iterative process can be mathematically described as

$$S_{n+1}(t) = \mathrm{SGFilter}(S_n(t), w_{SG}, p) \qquad (4)$$

where $(n = 0,1, \ldots, N_{iter} - 1)$ and (N_{iter}) is the number of iterations of the filter. Typically, $(p = 2)$ and (N_{iter}) is set to 5.

Signal segmentation and local trend estimation

Following the denoising step, the filtered signal $(S_n(t))$ is partitioned into a series of non-overlapping window segments $(W_i(t))$ for local signal transformation. Just as we previously defined the window size for the Savitzky-Golay filter, the window size for each non-overlapping segment of this step is determined in a similar fashion, which is as follows.

$$T_i = \max\left(10, \frac{N}{\lambda_{win}}\right) \qquad (5)$$

This will ensure that the windows are large enough to capture meaningful trends while avoiding over-segmentation that may distort the global structure of the signal. Each segmentation $(W_i(t))$ is defined as

$$W_i(t) = S_n(t_i : t_i + T_i) \quad (6)$$

where (t_i) is the starting index of the (i^{th}) window and (T_i) is the window length. Following this, each window $(W_i(t))$ is subjected to local trend estimation. We assume that each signal segment roughly follows a locally linear model. This trend is then estimated using ordinary least squares regression, fitting the data within each segment to a linear function as described below

$$W_{reg,i}(t) = \beta_i t + \alpha_i \qquad (7)$$

where (β_i) and (α_i) are the slope and intercepts of the (i^{th}) window segment, and are computed by minimizing the sum of squared residuals in the following manner

$$\min_{\beta_i, \alpha_i} \sum_{t=t_i}^{t_i+T_i} \left(W_i(t) - (\beta_i t + \alpha_i)\right)^2 \qquad (8)$$

We then identify the local trend direction by evaluating the estimated slope (β_i). If $(\beta_i \geq 0)$, the local trend is upward, while $(\beta_i < 0)$ indicates a downward trend.

Trend-Based conditional signal modification, Post-recalibration denoising and global signal reconstruction

The estimated regression parameters (β_i) and (α_i) drive the subsequent modification of the signal segment, specifically for cases where the slope parameter (β_i) is negative. The segment undergoes recalibration to emphasize these features in an upward direction. The recalibrated signal, $(\widetilde{W}_i(t))$ can be mathematically expressed as follows

$$\widetilde{W}_i(t) = \begin{cases} W_i(t), & if\ \beta_i \geq 0 \\ \dfrac{W_i(t)}{\left(\frac{W_{reg,i}(t)}{\kappa_{elev}}\right)}, & if\ \beta_i < 0 \end{cases} \qquad (9)$$

where (κ_{elev}) is an elevation factor that scales the recalibration operation. This recalibration process tends to accentuate underlying noise to a very noticeable amount. To reduce noise that may have been introduced, we subject the recalibrated signal ($\widetilde{W}_i(t)$) to smoothing via the Savitzky-Golay filter. This filtering process is performed conditionally based on the slope (β_i) as well as the post-smoothing factor (λ_{post}) in the following manner

$$\widehat{W}_i(t) = \begin{cases} SGFilter\left(\widetilde{W}_i(t), w_{post}, p\right), & if\ \beta_i < \beta_{thresh}\ and\ \lambda_{post} > 0 \\ \widetilde{W}_i(t), & otherwise \end{cases} \quad (10)$$

Here, the threshold (β_{thresh}) governs the extent to which the post-calibration smoothing is applied only when the condition ($\beta_i < \beta_{thresh}$) holds true, and the new segment (w_{post}) is expressed as

$$w_{post} = \left\lfloor \frac{T_i}{\lambda_{post}} \right\rfloor \quad (11)$$

where (λ_{post}) is a tunable parameter controlling the degree of post-recalibration smoothing. After all windows have been processed and smoothed, the global signal can be reconstructed to finally obtain the health indicator (HI). However, since each window with a negative slope (β_i) has been processed locally with no context to neighboring segments, there may be discontinuities at the boundaries between adjacent segments. Continuous transitions can be made by applying vertical adjustments to each segment of the reconstructed signal to obtain ($\hat{s}_i(t)$) which can be mathematically defined as follows

$$\hat{s}_i(t) = \begin{cases} \widehat{W}_i(t) + \left(\hat{s}_{i-1}(t_{i-1} + T_{i-1} - 1) - \widehat{W}_i(t_0)\right), & i > 0 \\ \widehat{W}_i(t), & i = 0 \end{cases} \quad (12)$$

Parameter selection

The SASD algorithm's performance is highly sensitive to the choice of its tunable parameters. After extensive manual analysis, the following parameters were chosen for optimal performance: ($\lambda_{win} = 100$), ($\lambda_{pre} = 45$), ($\beta_{thresh} = 0.4$), ($\kappa_{elev} = 20.1$), ($\lambda_{post} = 3$).

One important observation crucial to extracting features correctly is picking the right value for (λ_{win}). Smaller values of (λ_{win}) lead to larger window sizes, which can obscure underlying trends. There is a possibility of ending up with a flat regression slope due to a segment having both increasing and decreasing trends. Conversely, larger values of (λ_{win}) result in smaller window sizes, which lead to overemphasized local trends, amplified noise in the resulting signal and is a sub-optimal approach overall. Thus, a dynamically scaled window size is needed with a tried and tested minimum bound of 10, as show in **Eqn. 5**. A visual representation of the same can be seen as illustrated in **Fig. 2**. Following signal reconstruction, health indicators derived from this algorithm roughly estimate a linear degradation process across all bearings. Taking this into consideration, we can construct an estimate of the remaining useful life (RUL) of a bearing by subtracting the normalized values of (HI) between 0 and 1 in the following manner.

$$\mathrm{RUL} = 1 - \left(\frac{\mathrm{HI} - \min(\mathrm{HI})}{\max(\mathrm{HI}) - \min(\mathrm{HI})}\right) \quad (13)$$

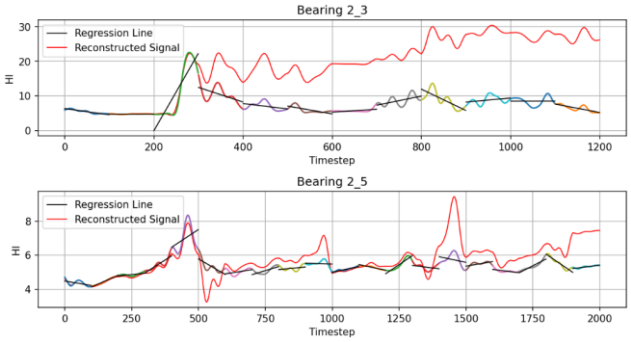


Figure 2: Illustration depicting the impact of λwin on signal trends. A smaller λwin obscures trends (TOP), while the lower plot with a larger λwin amplifies noise (BOTTOM).

Extracted health indicators for bearings subjected to operating condition 1 are illustrated in **Fig. 3.** The extracted health indicators exhibit a strong upward trend indicating degradation of the ball bearing setups and ultimately the degradation of the machine.

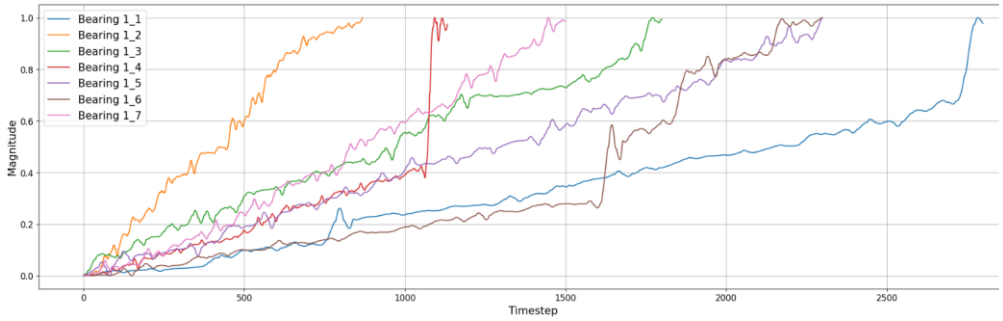


Figure 3: Extracted health indicators of bearings operating under condition 1

Dataset preparation and model training

The Slope Adaptive Signal Decomposition (SASD) algorithm is employed to extract the health indicator ($\widehat{s_i}(t)$) from the pre-computed time domain statistical features. Upon detailed manual analysis, it was observed that the signals derived using the Impulse Factor extracted from the horizontal accelerometer yielded the most informative and high-quality features making it suitable for further predictive analysis tasks relevant to the remaining useful life (RUL) of ball bearings. A sample of the same is illustrated in **Fig. 4.**

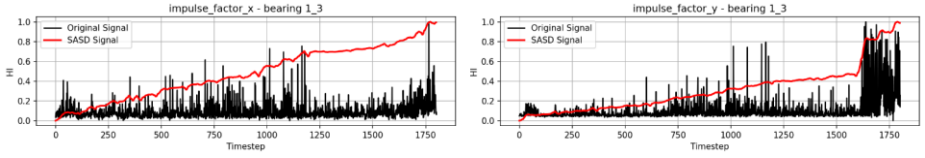


Figure 4: Illustration of the SASD signal sample extracted from the Impulse Factor of Bearing 1 under Condition 3, for both horizontal (left) and vertical (right) accelerometers.

For the purpose of training algorithms, we employ a range of modelling techniques.

I. *Gaussian Process Regression (GPR):* A non-parametric Bayesian approach that is useful for modeling complex, nonlinear relationships and providing uncertainty estimates for RUL predictions.
II. *Long Short-Term Memory (LSTM):* For capturing long-range dependencies in sequential data.
III. *Stacked LSTM networks:* To enhance the capacity of standard LSTMs by stacking multiple layers, allowing for more complex feature extraction and improved learning of intricate patterns in the data.
IV. *Gated Recurrent Units (GRU):* Similar to LSTMs but with a simplified architecture reducing the number of parameters, leading to faster training times.
V. *Stacked LSTM networks:* To combine the strengths of both LSTM and GRU architectures.

To model data using Gaussian Process Regression (GPR), we first enhance the quality of the input data by fitting a Logistic Growth model using ordinary least squares (OLS) to the health indicator for smoothing and representing major trends, eliminating any noise present. The logistic Growth function can be mathematically defined as follows

$$y(t) = \frac{L}{1+e^{-k(t-t_0)}} \qquad (14)$$

where $(y(t))$ is the predicted value at time (t), (L) is the carrying capacity of the curve, (k) is the growth rate and (t_0) is the time at which y is halfway to its inflection point.

The RUL signals used for the deep learning algorithms are directly constructed from the health indicator obtained from the proposed SASD algorithm with no additional preprocessing. All extracted bearing signals are partitioned into 3 groups of train-test sets with split ratios 1:3, 1:1 and 3:1, and are evaluated accordingly.

Method validation

The performance of trained models are assessed using standard error metrics, such as root mean squared error (RMSE) defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (15)$$

where (n) represents the number of samples and $(y_i - \hat{y}_i)$ is the deviation of the predicted RUL datapoint from the actual RUL datapoint at time instant (i), mean absolute error (MAE) defined as

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (16)$$

normalized root mean squared error (NRMSE) defined as

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i)} \qquad (17)$$

Some metrics from are also used to comprehensively assess the performance of the proposed SASD algorithm. The relative percentage error ($E_r\%$) derived from the original work[1]

$$E_{r_i} = \frac{\text{ActRUL}_i - \widehat{\text{RUL}_i}}{\text{ActRUL}_i} \times 100\% \qquad (18)$$

in which (ActRUL_i) and ($\widehat{\text{RUL}_i}$) represent the real RUL and the estimate RUL if the (i^{th}) bearing. The exponential transformed accuracy (see **Fig. 5**) is then adopted to reveal the quality of predictions based on its underestimate or overestimate as

$$ETA_i = \begin{cases} e^{-\ln(0.5)\cdot\left(\frac{E_i}{5}\right)}, & E_{r_i} \leq 0 \\ e^{+\ln(0.5)\cdot\left(\frac{E_i}{20}\right)}, & E_{r_i} > 0 \end{cases} \qquad (19)$$

after which the final score of the RUL estimation is computed as

$$\text{Score} = \frac{1}{\beta}\sum_{i=1}^{\beta} \text{ETA}_i \qquad (20)$$

where (β) is the number of bearings for a given operating condition. The final MAE for a model is computed as the average of all bearings for a given model, similarly NRMSE for a model is computed as the average of NRMSE for all bearings for that model.

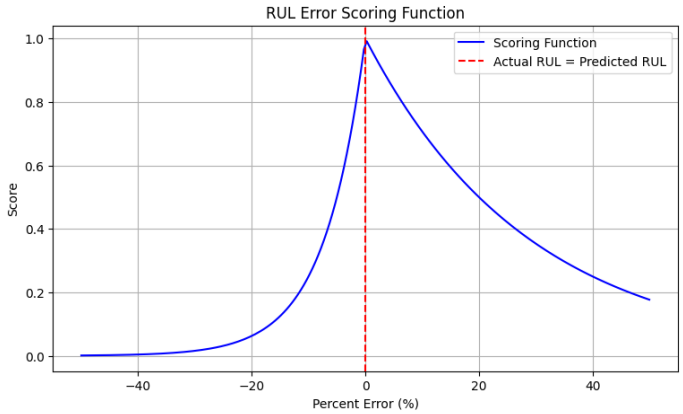


Figure 5: Exponential Scoring function

The aforementioned metrics are computed for all architectures and for each bearing across the three split ratios. The results are then averaged to obtain the final score. **Table 3** lists the evaluation scores for 6 different models across bearings from three operating conditions. It is clear that the GRU model outperforms other architectures in majority of the bearings. This can be seen for all bearings operating under condition 1, condition 2 and in condition 3. Under condition 1, LSTM outperforms other architectures for bearing 1-2 and 1-6 in terms of relative percentage error ($E_r\%$) as well as the average of exponential transformed accuracy across all bearings under that condition. LSTM + GRU outperforms GRU for only bearing 2-4 in terms of relative percentage error ($E_r\%$) equal to (0.57%) and has considerably lower scores for the remaining metrics. It is noteworthy that the GRU performs marginally better than the LSTM in general, relative to other architectures. The RMSE scores can also be visualized for each bearing to compare the relative potency between architectures as illustrated in **Fig. 6**. Bearing 2-7 has turned out to be the only exception in this study due to poor quality of extracted health indicator (HI) resulting in much higher relative percentage errors for all models. Due to this reason the computation of Score, MAE and NRMSE for models estimating RUL of bearings operating in condition 2 will not make use of relative percentage errors computed for this bearing. It can also be observed that the GRU, LSTM, LSTM-GRU hybrid and Stacked LSTM architectures perform worse than

expected compared to simpler modelling methods such as GPR and Polynomial Regression for bearing 2-7. This is due to the inability of deep learning architectures to learn intricate patterns from this specific signal as opposed to the simpler GPR and Polynomial regression capturing general trends and performing better than expected.

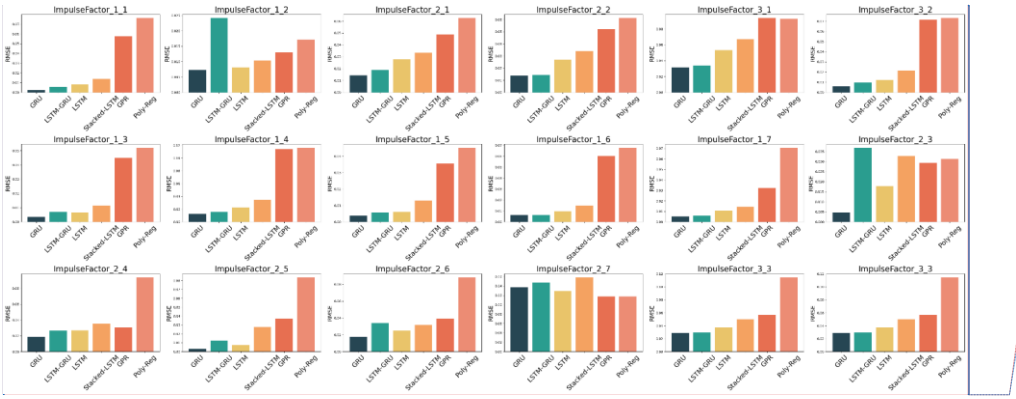


Figure 6: RMSE of HI prediction for all Pronostia bearings

TABLE 3: Performance Metrics of models using the proposed SASD algorithm in different conditions

	Method	E_r%							Score	MAE	NRMSE
		Bearing 1-1	Bearing 1-2	Bearing 1-3	Bearing 1-4	Bearing 1-5	Bearing 1-6	Bearing 1-7			
Condition 1	**GRU**	**0.01%**	0.65%	**0.08%**	-0.37%	**0.08%**	0.52%	**-0.01%**	0.9952	**0.0037**	**0.0086**
	LSTM	0.48%	**0.13%**	-0.20%	0.36%	-0.09%	**-0.07%**	0.04%	**0.9967**	0.0060	0.0154
	LSTM + GRU	-1.00%	2.47%	0.56%	0.80%	0.12%	0.51%	-0.17%	0.9838	0.0078	0.0168
	Stacked LSTM	-1.04%	0.46%	-0.94%	**0.26%**	0.83%	0.10%	-0.17%	0.9905	0.0101	0.0307
	GPR	-1.76%	-0.39%	-1.05%	-1.03%	-1.44%	1.26%	-1.79%	0.8843	0.0476	0.1054
	Poly Reg	-7.95%	17.62%	10.63%	2.61%	10.53%	3.57%	13.41%	0.3435	0.0993	0.2134
		Bearing 2-1	Bearing 2-2	Bearing 2-3	Bearing 2-4	Bearing 2-5	Bearing 2-6	Bearing 2-7	Score	MAE	NRMSE
Condition 2	**GRU**	**0.35%**	**0.90%**	**0.16%**	0.78%	**0.07%**	**0.24%**	-7.16%	**0.9855**	**0.0083**	**0.0206**
	LSTM	1.98%	2.38%	1.65%	0.87%	0.25%	0.50%	-9.02%	0.9567	0.0161	0.0389
	LSTM + GRU	2.22%	-1.37%	3.83%	**0.57%**	1.44%	3.23%	-7.78%	0.9442	0.0197	0.0395
	Stacked LSTM	1.08%	2.81%	3.25%	1.76%	2.50%	-1.86%	-6.54%	0.9463	0.0252	0.0666
	GPR	0.98%	-1.74%	0.46%	12.10%	3.79%	-1.51%	10.06%	0.7914	0.0573	0.1381
	Poly Reg	3.28%	-5.58%	18.31%	39.66%	12.99%	10.46%	-9.97%	0.1445	0.1503	0.3583
		Bearing 3-1	Bearing 3-2	Bearing 3-3					Score	MAE	NRMSE

Commented [PK7]: Is mention about y impulse factor is chosen mentioned in paper

Commented [DB8R7]: Yes, it is implied in Page 9 Line 8

Commented [PK9]: This table is bit confusing. Can in any way you can place best performing model together so it is understandable easily. Especially condition 3 gru performance doesn't seem good

Commented [DB10R9]: The models are placed best to worst (top to down) for each condition based on which one has higher number of better scores.

I have bold formatted the best model under each condition so that its clear. The order is more or less the same for all conditions except condition 3 where LSTM barely outperforms GRU.

Condition 3	**GRU**	1.30%	**0.10%**	**0.90%**	**0.9738**	**0.0107**	0.0335
	LSTM	**0.26%**	-0.15%	2.28%	0.9727	0.0173	0.0625
	LSTM + GRU	-0.52%	0.68%	-1.69%	0.9316	0.0138	**0.0317**
	Stacked LSTM	-1.44%	1.15%	4.25%	0.9553	0.0267	0.0750
	GPR	-0.93%	-0.43%	7.53%	0.9312	0.0650	0.1449
	Poly Reg	1.03%	0.80%	- 19.21%	0.4482	0.1122	0.2571

Limitations

None

Ethics statements

Not Applicable

CRediT author statement

Dev Bhanushali: Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, **Pooja Kamat**: Conceptualization, Methodology, Writing – review & editing, Resources, Supervision, Funding acquisition.

Acknowledgments

Funding: This work was supported by the Research Support Fund (RSF) of Symbiosis International (Deemed University), Pune, India.

Declaration of interests

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

Supplementary material *and/or* additional information [OPTIONAL]

None

References

[1] P. Nectoux *et al.*, "PRONOSTIA : An experimental platform for bearings accelerated degradation tests," *IEEE International Conference on Prognostics and Health Management, PHM'12., Jun 2012, Denver, Col- orado, United States.*, pp. 1–8, 2012.

[2] M. Kostrzewski and R. Melnik, "Condition Monitoring of Rail Transport Systems: A Bibliometric Performance Analysis and Systematic Literature Review," *Sensors 2021, Vol. 21, Page 4710*, vol. 21, no. 14, p. 4710, Jul. 2021, doi: 10.3390/S21144710.

[3] M. M. U. Z. Siddiqui and A. Tabassum, "Condition-based monitoring techniques and algorithms in 3d printing and additive manufacturing: a state-of-the-art review," *Progress in Additive Manufacturing 2024*, pp. 1–48, Nov. 2024, doi: 10.1007/S40964-024-00816-5.

[4] A. Kumar, C. P. Gandhi, H. Tang, W. Sun, and J. Xiang, "Latest innovations in the field of condition-based maintenance of rotatory machinery: a review," *Meas Sci Technol*, vol. 35, no. 2, p. 022003, Nov. 2023, doi: 10.1088/1361-6501/AD0F67.

[5] Y. Liu, L. Guo, H. Gao, Z. You, Y. Ye, and B. Zhang, "Machine vision based condition monitoring and fault diagnosis of machine tools using information from machined surface texture: A review," *Mech Syst Signal Process*, vol. 164, p. 108068, Feb. 2022, doi: 10.1016/J.YMSSP.2021.108068.

[6] S. Khaleghi, M. S. Hosen, J. Van Mierlo, and M. Berecibar, "Towards machine-learning driven prognostics and health management of Li-ion batteries. A comprehensive review," *Renewable and Sustainable Energy Reviews*, vol. 192, p. 114224, Mar. 2024, doi: 10.1016/J.RSER.2023.114224.

[7] W. He *et al.*, "Progress in prediction of remaining useful life of hydrogen fuel cells based on deep learning," *Renewable and Sustainable Energy Reviews*, vol. 192, p. 114193, Mar. 2024, doi: 10.1016/J.RSER.2023.114193.

[8] S. Wang, Y. Fan, S. Jin, P. Takyi-Aninakwa, and C. Fernandez, "Improved anti-noise adaptive long short-term memory neural network modeling for the robust remaining useful life prediction of lithium-ion batteries," *Reliab Eng Syst Saf*, vol. 230, p. 108920, Feb. 2023, doi: 10.1016/J.RESS.2022.108920.

[9] S. Plakias and Y. S. Boutalis, "Fault detection and identification of rolling element bearings with Attentive Dense CNN," *Neurocomputing*, 2020, doi: 10.1016/j.neucom.2020.04.143.

[10] P. Kamat, S. Kumar, S. Patil, and K. Kotecha, "Anomaly-informed remaining useful life estimation (AIRULE) of bearing machinery using deep learning framework," *MethodsX*, vol. 12, p. 102555, Jun. 2024, doi: 10.1016/J.MEX.2024.102555.

[11] M. Paramesha, N. L. Rane, and J. Rane, "Big Data Analytics, Artificial Intelligence, Machine Learning, Internet of Things, and Blockchain for Enhanced Business Intelligence," *Partners Universal Multidisciplinary Research Journal*, vol. 1, no. 2, pp. 110–133, Jul. 2024, doi: 10.5281/ZENODO.12827323.

[12] M. Soori, B. Arezoo, and R. Dastres, "Internet of things for smart factories in industry 4.0, a review," *Internet of Things and Cyber-Physical Systems*, vol. 3, pp. 192–204, Jan. 2023, doi: 10.1016/J.IOTCPS.2023.04.006.

[13] P. Kundu, S. Chopra, and B. K. Lad, "Multiple failure behaviors identification and remaining useful life prediction of ball bearings," *J Intell Manuf*, vol. 30, no. 4, pp. 1795–1807, 2019, doi: 10.1007/s10845-017-1357-8.

[14] P. Kamat, S. Kumar, and R. Sugandhi, "Vibration-based anomaly pattern mining for remaining useful life (RUL) prediction in bearings," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 46, no. 5, pp. 1–20, May 2024, doi: 10.1007/S40430-024-04872-4/METRICS.

[15] B. Hou, D. Wang, T. Xia, Z. Peng, and K. L. Tsui, "Difference mode decomposition for adaptive signal decomposition," *Mech Syst Signal Process*, vol. 191, p. 110203, May 2023, doi: 10.1016/J.YMSSP.2023.110203.

[16] G. Zhang *et al.*, "Optimized adaptive Savitzky-Golay filtering algorithm based on deep learning network for absorption spectroscopy," *Spectrochim Acta A Mol Biomol Spectrosc*, vol. 263, p. 120187, Dec. 2021, doi: 10.1016/J.SAA.2021.120187.