Unsupervised Remaining Useful Life Prediction for Bearings with Virtual Health Index

Gilbert Cheng*, Sean Lau*, Nicholas Tam*, Ze Kai Wu, Adah Hu, Yan Nei Law
Hong Kong Industrial Artificial Intelligence & Robotics Centre
Hong Kong
{gilbertcheng,seanlau,nicholastam,zkwu,adahhu,ivylaw}@hkflair.org

Edmond Lai and Ming Ge, *Senior Member, IEEE*
Hong Kong Productivity Council
Hong Kong
{edmondlai,mingge}@hkpc.org

***Abstract*—A particular interest in achieving Prognostics Health Management (PHM) for bearings has been developed in the scientific community and the industry, as they are critical components in generators and turbines. Majority of state-of-the-art methods used in prediction of Remaining Useful Life (RUL) require large amounts of run-to-failure data for training. While these methods offer accurate prediction, the usage barrier is particularly high to small-scale, downstream sector companies due to the significant amount of data needed. The goal of this paper is to demonstrate a novel unsupervised method to address this problem. The algorithm takes advantage of Convolution Neural Network (CNN) encoder-decoder to infer Virtual Health Indices (VHI) which are representative of the degradation pattern. Additionally, thresholds for these indices are determined with only End-of-Life (EOL) data, removing the need for run-to-failure experiments. The RUL is then obtained through the inference of the VHI. The suggested method is tested on selected data from the XJTU-SY bearing dataset, offering promising prediction results, reducing the barrier of usage for RUL algorithms.**

Keywords—prognostics health management, bearing failure, remaining useful life, virtual health indicator, deep neural network.

I. Introduction

The surge in popularity of IIoT (Industrial Internet of Things) has increased access to data for power electronics industries. This allows for development of algorithms for Predictive Maintenance (PM), a practice of monitoring machinery by various sensors to schedule appropriate maintenance actions taken at the optimal time. A particular component of interest is the bearing. Failure of bearings often cause catastrophic failure. Predicting the failure time of bearings is therefore a critical task in realizing predictive maintenance [1-4].

Sensors such as accelerometers are often used for machine condition monitoring. These raw signals are often converted to spectrograms for analysis. A particular type of spectrogram, the Mel-spectrogram, has been studied as an effective tool for bearing failure analysis. Hong *et al.* utilized Mel-spectrograms in combination with a Convolutional Neural Network (CNN) and Long-Short Term Memory (LSTM)-based model for bearing and equipment anomaly [5]. Kulevome *et al.* illustrated the superiority of using Mel-spectrograms over traditional spectrograms in bearing fault classification [6]. In this paper, Mel-spectrogram will further be applied in RUL prediction.

To reduce the dimensionality for RUL prediction, VHIs are extracted from the Mel-spectrograms to model the behaviour of the degrading component. These indices are often constructed through statistical manipulation or deep learning networks [7]. While VHIs are exemplary indicators of degradation, using them presents two particular challenges: First, the VHI has to be selected based on its ability to reflect an underlying degradation trend. Monotonicity and trendability are popular metrics for this purpose [8,9]. The second challenge is the selection of a threshold value signalling the end-of-life. As VHI lacks physical meaning, it is difficult to select a corresponding threshold. Various approaches have been suggested to tackle this problem. Beaulieu *et al.* identified common EOL patterns in the C-MAPSS turbofan engine dataset [10]. Liu *et al.* demonstrates using the mean values of extracted health indicators close to failure to be more accurate for RUL prediction [11]. It is reasonable to infer that the last few samples in each EOL datasets can be indicative of failure, and the values of VHI constructed from these samples can be used as the threshold correspondingly.

A. Contributions

This paper addresses some drawbacks of different state-of-the-art methods in RUL prediction and proposes an alternative approach; RUL algorithms can be broadly divided into supervised and unsupervised methods. Supervised methods require run-to-failure data to train [12,13], while the unsupervised method presented in [10] still requires a significant amount of EOL data to model. Long sequence of EOL data is challenging to obtain without run-to-failure experiments. For this purpose, an unsupervised RUL prediction method which requires only a small sample of EOL data from failed bearings is proposed, eliminating the need to spend time for destructive run-to-failure experiments.

Specifically, the following contributions are presented:

1. A methodology to obtain VHI without the need of run-to-failure data;
2. An approach to obtain the corresponding failure threshold;
3. An unsupervised method for RUL prediction using the estimated VHI.

The remainder of this paper is organized into the following sections: Section II gives a brief introduction on Mel-spectrograms and CNN encoder-decoders. A description of the architecture of the proposed method is then provided in Section III. Finally, Section IV presents the results on the selected bearings in the XJTU-SY bearing dataset.

* These authors contributed equally to this work.

II. Basic Theory

A. *Mel-Spectrograms*

The spectrogram is a form of representation of vibration signal or sound in short-term power spectrum of a signal transformed by Short Term Fourier Transform (STFT). The Mel-spectrogram is constructed as a projection of the linear frequency scale into the Mel-scale [14]. A representation of the Mel-scale can be found in the form:

$$m = 2595 \log_{10}\left(1 + \frac{f}{700}\right) \tag{1}$$

It has been applied to accelerometer signal processing for bearing fault identification and health degradation.

B. *Convolutional Neural Network Encoder-Decoder*

Convolutional Neural Network (CNN) is a type of artificial neural network vastly used in image processing. A kernel defined with a certain window size moves at a designated stride along the input image to highlight and to summarize the desired features into a feature map.

The encoder-decoder structure accepts an input and a pre-defined output and attempts to learn the mapping between the input and the output. The encoder extracts a latent representation of the input, while the decoder attempts to reconstruct the original input from the latent representation.

III. Proposed Unsupervised RUL Prediction Approach

The benefit of the proposed model lies in its ability to predict RUL without relying on data from run-to-failure experiments. Using VHI also eliminates the need to understand the Physics of Failure (POF). The proposed model requires only a small number of samples of EOL data, without the necessity of including any kind of long-term EOL trends, although it is noted that the more EOL data, the more accurate the model. In order to achieve this, the flowchart of the model is presented in Fig. 1. The model is described in the steps below.

Step 1: Inflection Point Identification

To identify when the bearing displays its first signs of degradation, it is proposed that there is an inflection point, from which point the bearing will transition from a "healthy" state into a "degrading" state. Assuming with no self-recovery or external maintenance, the inflection point is a point-of-no-return in which the bearing will continue to degrade until its EOL. Finding this inflection point is therefore especially critical as it will signal for the proposed model to proceed to next steps.

As shown in Fig. 2, when a bearing is in a healthy state, its raw vibration signals of all axes tend to display a homoscedastic behavior. Once the bearing starts degrading, the variance of the raw vibration signals will increase, and hence will exhibit a heteroscedastic behavior. First proposed by Bartlett [15], Bartlett's Test is a common statistical method to test homoscedasticity.

Bartlett's Test evaluates for the null hypothesis, that for two given groups, the variance of the two groups should be the same, or equivalently, $H_0\colon \sigma_1 = \sigma_2$, with the test statistics, adapted for this work being:

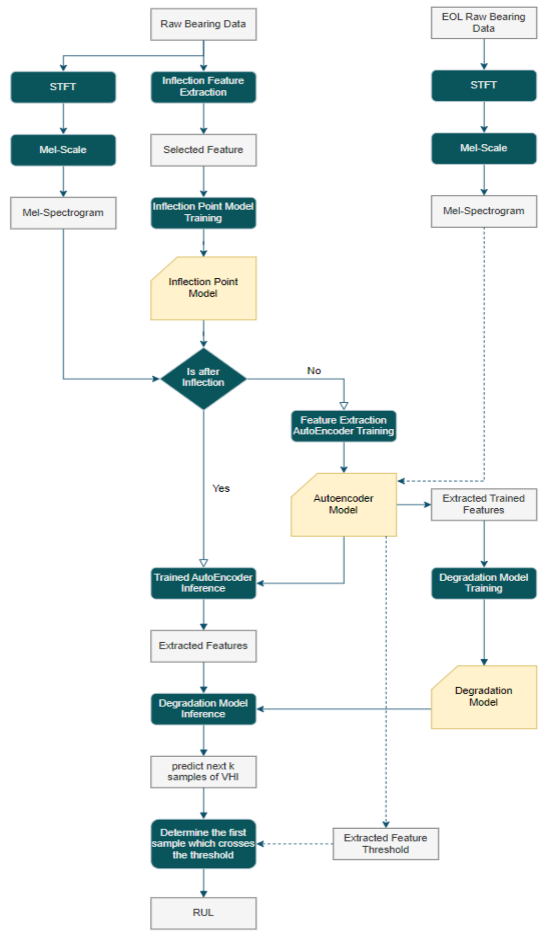

Fig. 1. The flowchart of the proposed solution, its constituent models, and decision points.

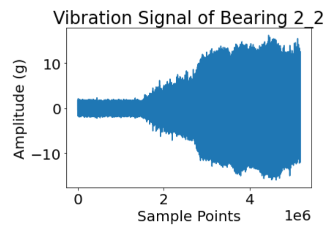

Fig. 2. The typical vibration pattern demonstrated by a bearing over the course of its run-to-failure process.

$$T = \frac{(N-2)\ln\left(S_p^2\right) - \sum_{i=1}^{2}(n_i - 1)\ln\left(S_i^2\right)}{1 + \frac{1}{3}\left(\sum_{i=1}^{n}\left(\frac{1}{n_i - 1}\right) - \frac{1}{N-2}\right)} \sim \chi_1^2 \tag{2}$$

where n_i is the number of samples in the i^{th} group, N is the sum of number of samples from both groups, s_i^2 is the sample variance of the i^{th} group, and s_p^2 is the pooled variance of both groups. The test statistics is to follow a Chi-squared distribution of one degree of freedom.

In this work, Bartlett's Test is repeatedly applied by means of sliding windows, to compare against two groups of consecutive sample periods. Given a group of raw signals s_t sampled at period t, the two groups can be expressed as:

$$g_0 = \{s_0, s_1, s_2, \ldots, s_{t-1}\},\ g_1 = \{s_1, s_2, \ldots, s_{t-1}, s_t\} \tag{3}$$

In other words, one of the groups considers the raw vibration signals of one sampling period ahead. The variances of the two groups are then compared by test statistics T. If $T < \chi^2_{0.05,1}$, then the null hypothesis can be rejected at a chosen significance level of 0.05, inferring that the two groups have significantly different variance values.

This analysis repeats iteratively at every timestep whenever new samples of vibration signals become available. The test statistics is monitored. To ensure that the bearing has truly entered the degradation state, a series of 5 null hypothesis rejections may warrant the declaration of an inflection point.

Step 2: Feature Extraction

The VHI should be an implicit representation of the degradation pattern of the monitored component. To extract the VHI, the raw acoustic data will first be transformed into Mel-spectrograms, which captures both time-domain and frequency-domain information. Spectrograms obtained from the data before the inflection point are assumed to be representative of healthy vibration signals. These will be used to train an undercomplete CNN encoder-decoder, in which the encoder structure will compress the $n \times n$ spectrogram into a vector of length k, providing a latent representation with reduced dimensionality. The decoder structure will attempt to reconstruct the original spectrogram from the latent representation. The overall structure of the CNN encoder-decoder is shown in Fig. 3. Using this method, the vector extracted by the encoder provides a sparse representation of the input Mel-spectrogram, with its entries being implicit yet highly representative features capturing the degradation trend.

Step 3: Feature Selection

By observation, some of the features exhibit a better representation of the degradation trend. This paper introduces a method of dynamic VHI selection. Given a certain extracted and log-transformed feature time series F and the series of time since monitoring per measurement T, the sum of four component scores is used as the final measurement score. This is comprised of an R^2 goodness-of-fit by least squares against a linear model, a traditional indicator of monotonicity (cf. Eqn. (4)), and a trendability measured by using the Pearson correlation coefficient against time (cf. Eqn. (5)). The final measurement score is given by their sum S (cf. Eqn. (6)), with a larger S value representing a more desirable feature. The definition of each score is defined as:

$$S_{monotonicity} = \frac{1}{n}\sum_{i=1}^{n}\left|corr\big(rank(F), rank(T)\big)\right| \tag{4}$$

$$S_{trendability} = \frac{E[FT] - E[F]E[T]}{\sqrt{E[F^2] - (E[F])^2}\sqrt{E[T^2] - (E[T])^2}} \tag{5}$$

$$S = S_{trendability} + S_{monotonicity} + R^2 \tag{6}$$

where n is the number of measured samples in the series, and $E[\cdot]$ denotes the expected value. At each time step, the features extracted from the CNN encoder are log-transformed and ranked according to the score S. The feature with the highest score is used to calculate the RUL at that time step.

Step 4: Feature Threshold Determination

Beaulieu *et al.* has identified a common EOL pattern [10]. Extending from their findings, this work hypothesizes that for the same failure component under the same working condition, the bearings should demonstrate similar EOL patterns. To verify this result, cosine similarity has been performed on each frequency band of the EOL Mel-spectrogram between pairs of bearings operating under the same working and failure condition.

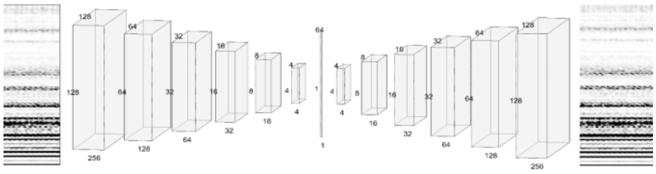

Fig. 3. The specifications of the CNN encoder-decoder employed in this work to extract relevant features from a given spectrogram.

Here, each frequency band of each bearing is represented as a numeric vector after processing into a Mel-spectrogram. Furthermore, each frequency band is compared against its counterpart in a different bearing under the same condition using cosine similarity, which is a similarity measure derived from the cosine of the angle between the said vectors $\vec{A}$ and $\vec{B}$, given as:

$$S_C(\vec{A}, \vec{B}) = \frac{\vec{A} \cdot \vec{B}}{||\vec{A}||\ ||\vec{B}||} \tag{7}$$

whereby $S_c = 1$ indicates the angle between the two vectors is 0, meaning both vectors are the same. Likewise, whereby $S_c = 0$, this indicates orthogonality of the two vectors, and that the two vectors are completely different.

Liu *et al.* has demonstrated that the mean values of extracted health indicators can be accurate in RUL prediction [9]. Extending from Liu's work and based on high cosine similarities among bands observed in empirical results of the case study, it is theorized that a generic EOL spectrogram can be generated by averaging. The last 5% of the spectrograms are selected to generate the EOL spectrogram based on their ability to reflect EOL patterns.

Step 5: RUL Prediction

Previous studies have shown that a raw extracted feature $f(t)$ can be fitted to an exponential model given by [16]:

$$f(t) = a\exp(bt + c) + d \tag{8}$$

where a, b, c, d are parameters to be determined by curve fitting. However, the exponential model is sensitive to noise, and often fails for cases where the trend is slightly non-monotonic [17]. A more robust approach involves taking the log-transformed features and applying least squares fitting to a linear model [10]. Future values of the selected VHI are then projected according to the fitted linear model, and the RUL is estimated as the point at which the VHI is projected to exceed the determined threshold. An exponential moving average filter is applied to smooth out any large jumps in predicted RUL values that may result from the transition between different selected features.

IV. Results

This section presents the findings using the proposed models on the XJTU-SY bearing dataset.

A. Dataset Description

The XJTU-SY bearing dataset is employed as a case study in this work, provided by XiAn JiaoTong University and the ChangXing Sumyoung Technology Company [16]. The run-to-failure bearing dataset is provided in CSV files. Each file contains 32,768 datapoints, representing a sample of 1.28

seconds with a sampling frequency of 25.6 kHz. A sample is collected once every minute. In each CSV file, there are two columns, one for horizontal vibration, and the other for vertical vibration. In the dataset, there are three experiments on five different bearings, each experiment ran under different operating conditions.

For easier identification, the bearings dataset follows a nomenclature of "*x_y*", where *x* specifies the operating condition and failure type (set number), and *y* specifies the bearing in scope. For this work, two bearing sets of different operating conditions were used, each comprises 3 bearings of the same failure type. Due to space limit, only the results of Condition II are presented, its specifications shown in Table I. The complete results are posted online[1].

TABLE I. BEARING SET II

Set	Working Conditions		Failure Type	Candidate Bearings		
	Rotation Speed (rpm)	*Loading (kN)*				
II	2250	11	Outer Race	2_2	2_4	2_5

B. Inflection Point Identification

Using the proposed inflection model, inflection points for selected experiments are found as shown in Fig. 4.

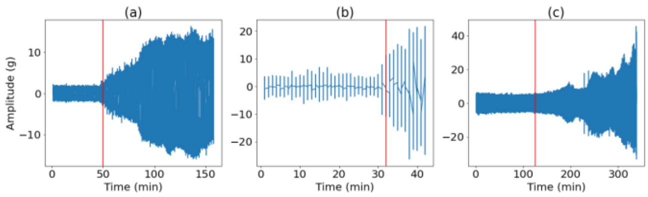

Fig. 4. Plots of the vibration signals of bearings in (a) 2_2, (b) 2_4, and (c) 2_5. The proposed inflection point by the model is denoted in red.

C. EOL Mel-spectrograms Similarity

Considering only the experiments with the "outer race" fault element under Condition II, the cosine similarity between the EOL spectrograms of the bearings is listed in Table II. From the case study in this work, EOL spectrograms exhibit a relatively high degree of cosine similarity on a band-by-band basis, attaining an average magnitude of 0.807.

Dissimilar bands with cosine similarity of less than 0.5 tend to be around low-frequency bands near the bottom of the spectrogram. Labelling the bands based on pixels of the spectrogram, the bands exhibiting similarity are across the top, while some of the highest similarities are often sandwiched between the dissimilar bands near the neighborhood of Band 240. It has been attributed that the dissimilarity is primarily due to the differences in amplitude, as the shapes of the EOL spectrograms are largely similar, despite minor shifts of bands.

TABLE II. COSINE SIMILARITY BETWEEN PAIRS OF BEARINGS

Bearing Pairs	Cosine Similarity		
	Minimum	*Maximum*	*Mean*
2_2 vs. 2_4	0.2449 *(Band 227)*	0.9985 *(Band 253)*	0.8120
2_2 vs. 2_5	0.2767 *(Band 254)*	0.9951 *(Band 222)*	0.8182
2_4 vs. 2_5	0.2591 *(Band 252)*	0.9997 *(Band 239)*	0.7911

D. VHI Threshold

This section presents the findings of a generic EOL spectrogram used to determine the EOL threshold at later stages of analysis.

The bearings are divided into 3 combinations of training and testing bearing sets, shown in Table III. EOL Mel-spectrograms of the training bearing dataset are combined through arithmetic mean. The combined Mel-spectrogram specific to the working condition and failure type is treated as the 'golden' EOL representation for the corresponding testing bearing set. A sampled pair of training and testing spectrograms of Combination II are shown in Fig. 5.

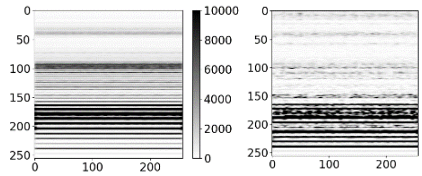

Fig. 5. *(left)* An average of the EOL spectrograms of Bearings 2_2 and 2_5. *(right)* The actual EOL spectrogram of Bearing 2_4.

TABLE III. BEARING SET II

Combination	Bearing(s) In Scope	
	Training	*Testing*
I	2_4, 2_5	2_2
II	2_2, 2_5	2_4
III	2_2, 2_4	2_5

Through application of the trained encoder-decoder model to the 'golden' representation, important EOL features can be extracted. The log-scaled values of these features are selected as the threshold of the testing set correspondingly.

E. RUL Prediction

This work aims to predict the RUL of the candidate bearings. Root Mean Squared Error (RMSE) is used as the primary performance evaluation criterion. To compare to other works, cumulative relative accuracy (CRA) is also used. In particular, CRA is given as:

$$CRA = \sum_{k=1}^{K} w_k \left(1 - \frac{|RUL(T_k) - \widehat{RUL}(T_k)|}{RUL(T_k)}\right) \tag{9}$$

where w_k is the normalized weight factor of time step k, given by $k/\sum_{k=1}^{K} k$, $RUL(T_k)$ is the actual RUL observed at time T_k, and $\widehat{RUL}(T_k)$ is the predicted RUL by the model.

Due to the lack of suitable literature employing unsupervised VHI approach in RUL prediction, the performance result of a hybrid prognostics approach by Wang *et al.* employing a physical Health Indicator is used to benchmark the proposed approach [16]. Table IV displays the benchmark results, with the best score highlighted in **bold**.

TABLE IV. PERFORMANCE EVALUATION FOR CANDIDATE BEARINGS

Set	Candidate Bearing	Hybrid Prognostics Approach	This Work		
		CRA	*CRA*	*RMSE*	*MAE*
II	2_2	**0.6521**	0.3751	47.3933	37.2657
	2_4	0.6276	**0.7640**	6.6151	4.0158
	2_5	0.6328	**0.6684**	9.1851	7.2738

[1] The complete result of this paper can be accessed at https://github.com/seanlau-flair/unsupervised-remaining-useful-life-prediction-for-bearings-with-virtual-health-index.git

Table V further shows the results of the approach suggested by this work, compared against other state-of-the-art approaches, including using Relevance Vector Machine (RVM), Deep Belief Network (DBN), Particle Filtering (PF), and Extended Kalman Filtering (EKF). For each candidate bearing, the best performing score is highlighted in **bold**, the second best in *italic*. It is noteworthy that among candidate bearings under Condition II, the proposed approach has consistently outperformed state-of-the-art approaches, and is even better than the Hybrid Prognostics Approach in the case with Bearings 2_4 and 2_5.

TABLE V. PERFORMANCE EVALUATION FOR CANDIDATE BEARINGS IN COMPARISON TO OTHER STATE-OF-THE-ART APPROACHES

Bearing	RVM	DBN	PF	EKF	This Work
	CRA	*CRA*	*CRA*	*CRA*	*CRA*
2_2	0.1789	-0.1977	0.2634	**0.4314**	*0.3751*
2_4	-0.0693	*0.5316*	0.4633	0.5004	**0.7640**
2_5	0.2563	0.0671	0.1833	*0.4815*	**0.6684**

Fig. 6 demonstrates the evolution of the RUL prediction values among the candidate bearings after the inflection point has been detected. The predicted values generally fluctuate right after the inflection point but converge toward the true values near the EOL.

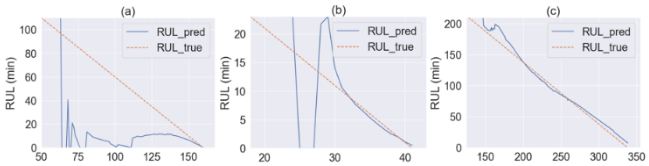

Fig. 6. RUL predictions of (a) Bearing 2_2, (b) Bearing 2_4, and (c) Bearing 2_5, by the proposed approach. Note that predictions only kick in once past the detected inflection point.

V. CONCLUSION

This paper has proposed a promising method to predict the RUL of a bearing using limited EOL data. This unsupervised method provides a viable alternative to RUL prediction by ridding the need to secure a large run-to-failure dataset for training and any domain knowledge of POF.

The proposed method only requires several samples of the EOL data without any long-term trends. For a bearing of interest, its raw vibration signal is consistently monitored until an inflection point is detected by means of statistical methods. The appearance of an inflection point indicates the transition into the degradation state, from which features can be drawn and selected from accumulated healthy data. A benchmark of EOL threshold for the bearing of interest is then proposed. This benchmark is generated from a generic EOL spectrogram, which is derived from a limited amount of EOL training data from other bearings. With the determined threshold, a VHI Prediction Model is developed to estimate the RUL of the said bearing.

This work has attained desirable results despite the challenges of working with limited EOL data. The proposed approach is capable of achieving results comparable to other state-of-the-art methods, at times exceeding such benchmarks. This work has faced difficulties in determining the appropriate EOL threshold for a candidate bearing. Future works can include (i) better generalization methods to determine a generic EOL spectrogram, (ii) better choice of features and smoothing for transition between features, and (iii) exploration of other failure cases and failure components to ensure extensibility. Such efforts can add flexibility to the approach to boost the reliability and the application proposed by this work.

ACKNOWLEDGMENT

This work was supported by the InnoHK funding launched by Innovation and Technology Commission, Hong Kong SAR.

REFERENCES

[1] W. Teng, X. Zhang, Y. Liu, A. Kusiak, and Z. Ma, “Prognosis of the Remaining Useful Life of Bearings in a Wind Turbine Gearbox,” *Energies*, vol. 10, no. 1, p. 32, Dec. 2016, doi: 10.3390/en10010032.

[2] Y. Hu, H. Li, P. Shi, Z. Chai, K. Wang, X. Xie, and Z. Chen, “A prediction method for the real-time remaining useful life of wind turbine bearings based on the Wiener process,” *Renewable Energy*, vol 127, pp. 452-460, Nov 2018.

[3] L. Cao *et al.*, “Prediction of Remaining Useful Life of Wind Turbine Bearings under Non-Stationary Operating Conditions,” *Energies*, vol. 11, no. 12, p. 3318, Nov. 2018, doi: 10.3390/en11123318.

[4] M. Pagitsch, G. Jacobs, and D. Bosse, “Remaining Useful Life Determination for Wind Turbines,” *Journal of Physics: Conference Series,* vol. 1452, issue 1, Jan. 2020, doi: 10.1088/1742-6596/1452/1/012052

[5] G. Hong and D. Suh, "Supervised-Learning-Based Intelligent Fault Diagnosis for Mechanical Equipment," in *IEEE Access*, vol. 9, pp. 116147-116162, 2021, doi: 10.1109/ACCESS.2021.3104189.

[6] D. K. B. Kulevome, H. Wang and X. Wang, "Deep Neural Network based Classification of Rolling Element Bearings and Health Degradation Through Comprehensive Vibration Signal Analysis," in *Journal of Systems Engineering and Electronics*, vol. 33, no. 1, pp. 233-246, February 2022, doi: 10.23919/JSEE.2022.000023.

[7] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, and D. Siegel, “Prognostics and health management design for rotary machinery systems – Reviews, methodology and applications,” *Mechanical Systems and Signal Processing*, vol 42, Issues 1-2, pp. 314-334, Jan 2014.

[8] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Morello, N. Zerhouni, and C. Varnier, “PRONOSTIA: An Experimental Platform for Bearings Accelerated Degradation Tests,” in *IEEE Prognostics and Health Management*, pp. 1-8, Denver, Colorado, Jun 2012

[9] X. Liu, P. Song, C. Yang, C. Hao and W. Peng, "Prognostics and Health Management of Bearings Based on Logarithmic Linear Recursive Least-Squares and Recursive Maximum Likelihood Estimation," in *IEEE Transactions on Industrial Electronics*, vol. 65, no. 2, pp. 1549-1558, Feb. 2018, doi: 10.1109/TIE.2017.2733469.

[10] M. Hervé de Beaulieu, M. Shekhar Jha, H. Garnier, and F. Cerbah, “Unsupervised Prognostics based on Deep Virtual Health Index Prediction”, *PHME_CONF*, vol. 7, no. 1, pp. 193–199, Jun. 2022.

[11] X. Liu and F. Machida, "Failure Threshold Setting for Wiener-process-based Remaining Useful Life Estimation," *2021 Global Reliability and Prognostics and Health Management (PHM-Nanjing)*, 2021, pp. 1-5, doi: 10.1109/PHM-Nanjing52125.2021.9613018.

[12] Y. Wu, M. Yuan, S. Dong, L. Lin, and Y. Liu, “Remaining useful life estimation of engineered systems using vanilla LSTM neural networks,” in *Neurocomputing*, vol. 275, pp. 167-179, Jan 2018.

[13] W. Yu, I. Y. Kim, and C. Mechefske, “An improved similarity-based prognostic algorithm for RUL estimation using an RNN autoencoder scheme,” in *Reliability Engineering & System Safety*, vol. 199, Jul 2020.

[14] D. O’shaughnessy, *Speech Communications: Human and Machine*. Reading, Massachusetts: Addison-Wesley Publishing Company, 1987.

[15] M. S. Bartlett, “Properties of Sufficiency and Statistical Tests,” in *Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences*, vol. 160, no. 901, pp. 268-282, May 1937.

[16] B. Wang, Y. Lei, N. Li, and N. Li, “A hybrid prognostics approach for estimating remaining useful life of rolling element bearings,” IEEE Transactions on Reliability, vol 69, no. 1, pp. 401-412, 2020.

[17] Z. Zhong, Y. Zhao, A. Yang, H. Zhang, and Z. Zhang, “Prediction of Remaining Service Life of Rolling Bearings Based on Convolutional and Bidirectional Long- and Short-Term Memory Neural Networks,” *Lubricants*, vol. 10, no. 8, p. 170, Jul. 2022, doi: 10.3390/lubricants10080170.