Online inverse solution for deep learning-based prognostics

Li, Tianzhi; Moradi, M.; Xiao, Ming ; Wang, Lihui

DOI
10.21741/9781644903513-14

Publication date
2025

Document Version
Final published version

Published in
Structural Health Monitoring - The 10th Asia-Pacific Workshop on Structural Health Monitoring, 10APWSHM 2024

Citation (APA)
Li, T., Moradi, M., Xiao, M., & Wang, L. (2025). Online inverse solution for deep learning-based prognostics. In S. Xue, K. Ikago, L. Xie, & M. Cao (Eds.), *Structural Health Monitoring - The 10th Asia-Pacific Workshop on Structural Health Monitoring, 10APWSHM 2024: Materials Research Proceedings* (Vol. 50, pp. 119-126). Article 14 (Materials Research Proceedings; Vol. 50). Materials Research Forum LLC. https://doi.org/10.21741/9781644903513-14

Materials Research Proceedings 50 (2025) 119-126 https://doi.org/10.21741/9781644903513-14

Online inverse solution for deep learning-based prognostics

Tianzhi Li[1,a *], Morteza Moradi[3,b], Ming Xiao[2,c], Lihui Wang[1,d]

[1]Department of Production Engineering, KTH Royal Institute of Technology, 10044 Stockholm, Sweden

[2]Division of Information Science and Engineering, KTH Royal Institute of Technology, 10044 Stockholm, Sweden

[3]Center of Excellence in Artificial Intelligence for structures, prognostics & health management, Aerospace Engineering Faculty, Delft University of Technology, Delft, 2629 HS, the Netherlands

[a]tianzhil@kth.se, [b]M.Moradi-1@tudelft.nl, [c]mingx@kth.se, [d]lihuiw@kth.se

Keywords: Structural Health Monitoring, Remaining Useful Life Prediction, Inverse Solution, State and Parameter Estimation, Multiple Local Particle Filter

Abstract. Data-driven prognostic models have been extensively utilized in current structural health monitoring (SHM) practices. They are designed to provide the health indicator (HI) - a representation of the system's current health state - through sensor data. To enhance performance, online learning is often used to take care of uncertainties that arise from the run-to-failure process. The inverse solution, though demonstrated in online uncertainty quantification applications, remains unexplored in the context of online data-driven prognostics. Therefore, this work proposes a generic inverse solution for a deep prognostic model to online address uncertainties. The proposed method is tested using the open-access XJTU-SY bearing datasets, showcasing its capacity to online enhance the performance of a given model.

Introduction

Structural health monitoring (SHM) has been extensively investigated across various engineering application scenarios, such as in metals [1], composites [2], or rotating machinery [3-5]. SHM encompasses four key levels: damage detection [6], isolation or localization [7, 8], identification [9], and prognostics [1], efficiently ensuring the integrity and safety of engineering structures. Regarding prognostics, it often consists of three main sequential steps [1]: defining the damage state or health indicator (HI) to assess potential failure, building the prognostic models to calculate the RUL, and (iii) refining these models to improve the RUL prediction accuracy.

The first task typically involves characterizing the structure's health status by either a physics-based damage state or a data-driven HI. The definition of a physics-based state varies: for metal, it can be crack length [1, 10-12] or crack shape [13]; for composites, matrix cracking density [14], delamination length [15], or delamination shape [16, 17]; and for gears, crack length, pitting level, or wear depth [18]. On the other hand, a data-driven HI is often obtained by online structural health monitoring (SHM) data, such as acceleration [19, 20], strain [21], guided waves [22], or acoustic emission [23]. Deep models are often designed to provide advanced HI through sensor data, involving two steps. First, a function that involves the service time and end-of-life (EOL), such as the square ratio of the current time to EOL, is designed as a HI label simulator. Second, a deep model is trained by simulated labels and sensor data [22, 23]. During the testing phase, the model output DHI can directly produce EOL or RUL, typically without future HI projections [24, 25]. Therefore, this DHI construction model is often considered the prognostic model, though it does not describe the HI evolution with time.

Given the uncertainties arising from factors such as complex structure degradation and environmental influences [26], when the same model is applied to several identical specimens

Materials Research Proceedings 50 (2025) 119-126 https://doi.org/10.21741/9781644903513-14

undergoing identical run-to-failure tests, significant variations in prognostic performance will occur. Therefore, online updating is often necessary to take care of the uncertainties arising from the degradation process. Regarding the HI-based model, since the EOL cannot be obtained during the run-to-failure process, it is rarely possible to acquire the true label for each data stream in real-time. As a result, updating this deep model with the latest data represents a typical unsupervised online learning problem, which only received a few investigations, i.e., online transfer learning [27-29] and online incremental learning [30].

These investigations [27-30] have provided successful online prognostic solutions, and they all consider data-driven prognostics as a mapping problem - progressing sequentially from sensor measurement, through the prognostic model, and finally resulting in the HI or RUL prediction. On the other hand, inverse solution has been extensively demonstrated in online uncertainty quantification [16], with the potential to be applied to a deep prognostic model for addressing degradation uncertainties. This, however, has not been explored.

In this context, this work proposes the a generic inverse solution for a given deep prognostic model to online address uncertainties. First, a prognostic model is constructed using a user-defined modeling strategy. Next, a state-space model is developed by incorporating the prognostic model and prior information. Finally, state and parameter estimation is performed to generate the HI posterior. The proposed method is tested using the open-access XJTU-SY bearing datasets, showcasing its capacity to enhance the performance of a given prognostic model.

The rest of this paper is organized as follows: Section 2 introduces the proposed method. Section 3 provides the results of the proposed method applied to the XJTU-SY bearing datasets, respectively. Finally, Section 4 concludes this paper.

Proposed method

The proposed inverse solution with both the offline and online phases. The offline phase involves the development of four models: prognostic model (PM), feature extraction model (FEM), measurement model (MM), and state space model (SSM). The online phase focuses on the state and parameter estimation with the latest measurement from the FEM model.

This section introduces the development of four models: training the PM, defining the FEM, training the MM, and formulating the SSM. To train the first model, a function needs to generate the HI labels. For example, the HI can be defined by a linear function as the ratio of current time to EOL:

$$x_k = \frac{t_k}{EOL} \tag{1}$$

where x, k, and t denote the HI, time step, and service time, respectively. Then, given the simulated labels, a data-driven model as PM can be built to link the sensor data $\boldsymbol{u}$ and the HI x as:

$$x_k = f\left(\boldsymbol{u}_k\right) \tag{2}$$

Then, a FEM model is firstly split from the PM excluding certain last layers, as follows:

$$\boldsymbol{Y}_k = f_s\left(\boldsymbol{u}_k\right) \tag{3}$$

In this study, FEM is considered as PM without only the last layer, i.e., the HI output layer. As a result, the FEM's output $\boldsymbol{Y}$ is a vector of the neurons of PM last hidden layer. By using the HI labels as input and the FEM's output as output, the MM can be constructed as follows:

$$\boldsymbol{Y}_k = g\left(x_k\right) \tag{4}$$

Then, by combining Eqs. (1) and (4), the SSM can be formulated as:

$$\begin{cases} \boldsymbol{X}_k = \begin{bmatrix} \boldsymbol{\theta}_k \\ EOL_k \\ x_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{\theta}_{k-1} + \boldsymbol{\omega}_{\theta,k} \\ EOL_{k-1} + \omega_{e,k} \\ t_k / EOL_k + \omega_{x,k} \end{bmatrix} \\ \boldsymbol{Y}_k = g\left(x_k, \boldsymbol{\theta}_k\right) + \boldsymbol{v}_k \end{cases} \tag{5}$$

where $\boldsymbol{\theta}$ is a vector of the model parameters within the measurement equation, $\boldsymbol{\omega}_\theta$, $\boldsymbol{\omega}_e$, and ω are the process noises for the model parameters, EOL and HI, respectively, and v is the measurement noise. Moreover, certain prior information can be included, such as:

$$k \le EOL_k \le EOL_{max} \tag{6}$$

which means EOL should always lie within the range between the current service time and the maximum EOL. The EOL should be adjusted to the nearest boundary when falling out of the range.

The SSM Eq. (5) is developed based on the PM Eq. (2), while it can provide more accurate prognostic performance, as it incorporates additional prior knowledge, and takes care of the uncertainties arising from the degradation process.

The online phase includes the feature extraction and the state and parameter estimation. The latest sensor data should be processed through the aforementioned FEM to extract specific features, which serve as measurements in the state space model. Finally, a state and parameter estimation algorithm will be used to provide the HI posterior.

Given that Eq. (5) is often high-dimensional and nonlinear, because of the parameter vector and the prognostic and measurement models, respectively. A high-dimensional system identification method, i.e., multiple local particle filter [16], is used in this study for online state and parameter estimation.

Testing with XJTU-SY Bearing Datasets

The XJTU-SY bearing datasets [19] are used in this study. Figure 1 depicts one bearing under the run-to-failure test. Two PCB 352C33 accelerometers are placed on the bearing housing to collect the vertical and horizontal accelerations at 60-second intervals. The sampling frequency and duration are 25.6 kHz and 1.28 seconds, respectively. Table 1 presents one operating condition from the datasets, involving tests on five bearings. For safety reasons, testing for each bearing is stopped once the vibration amplitude exceeds 20 g. Consequently, the time when the vibration amplitude exceeds the threshold is considered the EOL. Due to the uncertainties inherent in the degradation process, bearings may have different failure modes, including inner race wear, outer race wear, and outer race fractures, which result in varying EOLs. One can refer to [19] for further experimental details.

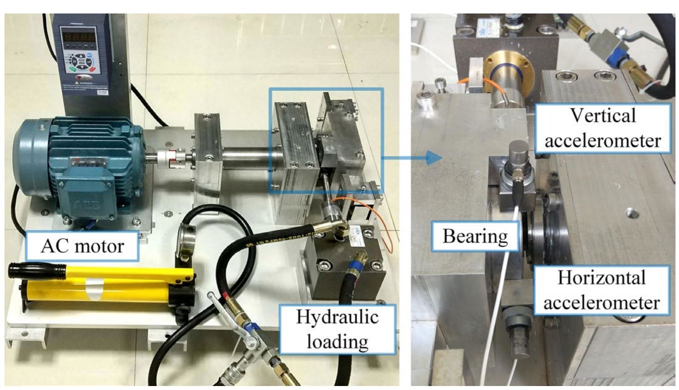

Figure 1: XJTU-SY Bearing run-to-failure test.

Table 1: Operating condition.

Condition	Rotating speed [rpm]	Radial force [kN]	Bearing specimen	EOL [Minute]
I	2100	12	1_1, 1_2, 1_3, 1_4, 1_5	123, 161, 158, 122, 52

Cross-validation will be performed for each bearing under Condition I. Specifically, one bearing will be designated for testing, while the remaining four bearings will be utilized for modeling. Table 2 lists the prognostic and measurement models used for each testing scenario. For each, two CNN-based PMs are separately used for testing, and their layouts are defined as 'Layout 1' and 'Layout 2', respectively. Then, the FEM, MM, and SSM have to be developed based on each PM.

Table 2: Prognostic and measurement models used for XJTU-SY datasets.

	Input	Output	Layout
Prognostic model 1 (PM1)	Raw data	HI	CNN 1
Prognostic model 2 (PM2)	Raw data	HI	CNN 2
Measurement model 1 (MM1)	HI	FEM 1 output	MLP
Measurement model 2 (MM2)	HI	FEM 2 output	MLP

The performance of the proposed method is first tested by applying PM1 to the bearings under Condition I. Three separate routines are compared:

- PM1: PM1 is used to provide HI prediction results.
- PM1 (P): The PM1 results are modified by the prior knowledge. Specifically, at each time step, the HI is set as the ratio of service time to the maximum EOL when it is lower than the ratio, and it is set as one when it is above the value of one.
- New: The proposed method is developed based on the given PM.

The parameter estimation results using the new approach are shown in Figures 2 (a) - (e). As the estimation process progresses, the spread of parameter samples decreases and converges around final values, indicating satisfactory convergence. The HI predictions using both the PM1 and the new approach are presented in Figures 2 (f) - (j). The new approach yields significantly smoother and more accurate results, highlighting its superiority in improving prognostic performance. Then, the results from the above three routines are evaluated by three performance metrics: root-mean-square error (RMSE), mean absolute percentage error (MAPE), and mean absolute error (MAE). Results from the three above routines are given in Table 3. The proposed method consistently achieves lower RMSE, MAPE, and MAE values, demonstrating its ability to enhance the prognostic performance of the given model PM1.

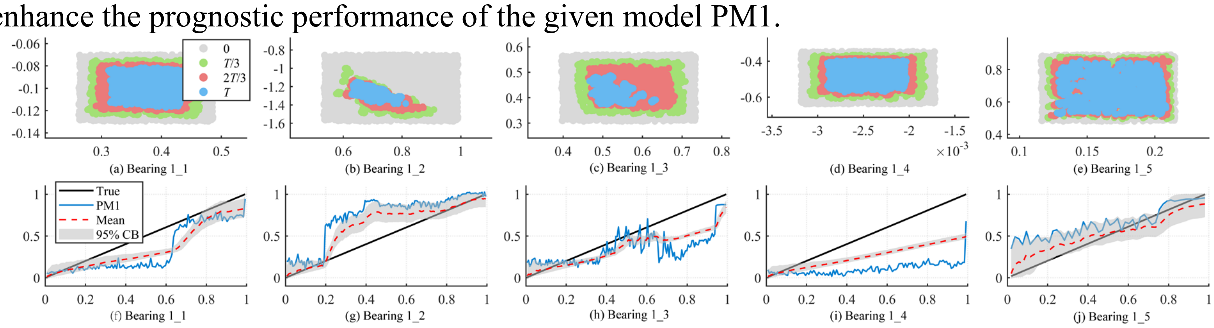

Figure 2: Results of PM1 and new methods for bearings under Condition I.

Note: (i) 'T' means the number of total time steps, (ii) the x and y axes for parameter estimation are two model parameters, while those for HI prediction are 'Service time / True EOL', and 'HI', respectively, and (iii) 'Mean' and '95% confidence boundary (CB)' are derived from the results of new method.

Table 3: Prognostic performances of PM1 and new methods for bearings under Condition I.

Bearing		1_1	1_2	1_3	1_4	1_5
RMSE	PM1	0.216	0.285	0.256	0.466	0.204
	PM1 (P)	0.173	0.284	0.188	0.301	0.204
	New	**0.172***	**0.171***	**0.186***	**0.291***	**0.097***
MAPE	PM1	46.6	112.7	82.6	86.3	130.2
	PM1 (P)	41.0	112.6	77.0	61.7	130.2
	New	**31.3***	**37.4***	**33.6***	**45.2***	**42.6***
MAE	PM1	0.174	0.237	0.196	0.402	0.163
	PM1 (P)	**0.147***	0.235	0.155	0.261	0.163
	New	**0.147***	**0.125***	**0.153***	**0.247***	**0.088***

Note: The symbol '*' denotes 'best performance among the three routines'.

The same validation is conducted on the five bearings using PM2. Three routines are included: PM2, PM2 (P), and the new method. The estimation and prognostic results are given in Figure 3 and Table 4. Although a different PM can yield slightly different prognostic performance for the same bearing, the proposed method can always have the capacity to improve the performance of the given model.

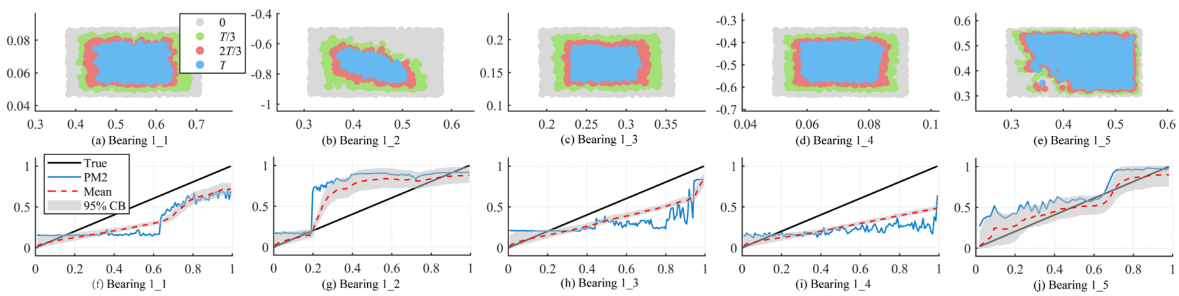

Figure 3: Results of PM2 and new methods for bearings under Condition I.

Note: (i) 'T' means the number of total time steps, (ii) the x and y axes for parameter estimation are two model parameters, while those for HI prediction are 'Service time / True EOL', and 'HI', respectively, and (iii) 'Mean' and '95% confidence boundary (CB)' are derived from the results of new method.

Table 4: Prognostic performances of PM2 and new methods for bearings under Condition I.

Bearing		1_1	1_2	1_3	1_4	1_5
RMSE	PM2	0.244	0.301	0.285	0.379	0.184
	PM2 (P)	0.213	0.301	0.206	0.299	0.184
	New	**0.212***	**0.211***	**0.198***	**0.295***	**0.068***
MAPE	PM2	75.0	108.0	97.9	86.4	107.6
	PM2 (P)	71.0	108.0	90.6	79.0	107.6
	New	**38.7***	**44.2***	**34.4***	**47.5***	**23.0***
MAE	PM2	0.214	0.240	0.235	0.314	0.154
	PM2 (P)	0.192	0.240	0.182	0.258	0.154
	New	**0.188***	**0.160***	**0.168***	**0.250***	**0.059***

Note: The symbol '*' denotes 'best performance among the three routines'.

Conclusion

Data-driven prognostic models must be continuously updated to account for uncertainties in the degradation process. This work introduces a novel inverse solution to address these uncertainties in real time. Based on prognostic results from the open-access XJTU-SY bearing datasets, the following conclusions can be drawn: Incorporating prior knowledge, such as the maximum EOL of certain specimens, can significantly enhance prognostic performance, even with simple online adjustments. For further improvements, it is essential to employ prognostic and measurement models to construct a state-space model, integrating prior knowledge for real-time state and parameter estimation. The effectiveness of the proposed method has been shown to rely on the proper utilization of prior information and measurement models. This work has only adopted a very simple prior, i.e., a constant maximum EOL. One may consider leveraging more advanced priors, such as the degradation-related physical constraints, or incorporating a physics-informed neural network into the proposed method.

References

[1] T. Li, Particle filter-based fatigue damage prognosis using prognostic-aided model updating, Mechanical Systems and Signal Processing, 211 (2024) 111244. https://doi.org/10.1016/j.ymssp.2024.111244

[2] L. Lomazzi, R. Junges, M. Giglio, F. Cadini, Unsupervised data-driven method for damage localization using guided waves, Mechanical Systems and Signal Processing, 208 (2024) 111038. https://doi.org/10.1016/j.ymssp.2023.111038

[3] J. Tian, D. Han, H.R. Karimi, Y. Zhang, P. Shi, A universal multi-source domain adaptation method with unsupervised clustering for mechanical fault diagnosis under incomplete data, Neural Networks, (2024) 106167. https://doi.org/10.1016/j.neunet.2024.106167

[4] D. Yang, H.R. Karimi, L. Gelman, An explainable intelligence fault diagnosis framework for rotating machinery, Neurocomputing, 541 (2023) 126257. https://doi.org/10.1016/j.neucom.2023.126257

[5] R. Zhong, Y. Feng, P. Li, X. Wu, A. Guo, A. Zhang, C. Li, Uncertainty-aware nuclear power turbine vibration fault diagnosis method integrating machine learning and heuristic algorithm, IET Collaborative Intelligent Manufacturing, 6 (2024) e12108. https://doi.org/10.1049/cim2.12108

[6] C. Lai, P. Baraldi, E. Zio, Physics-Informed deep Autoencoder for fault detection in New-Design systems, Mechanical Systems and Signal Processing, 215 (2024) 111420. https://doi.org/10.1016/j.ymssp.2024.111420

[7] J. Tian, D. Han, M. Li, P. Shi, A multi-source information transfer learning method with subdomain adaptation for cross-domain fault diagnosis, Knowledge-Based Systems, 243 (2022) 108466. https://doi.org/10.1016/j.knosys.2022.108466

[8] R. Junges, Z. Rastin, L. Lomazzi, M. Giglio, F. Cadini, Convolutional autoencoders and CGANs for unsupervised structural damage localization, Mechanical Systems and Signal Processing, 220 (2024) 111645. https://doi.org/10.1016/j.ymssp.2024.111645

[9] X. Zhou, C. Sbarufatti, M. Giglio, L. Dong, A fuzzy-set-based joint distribution adaptation method for regression and its application to online damage quantification for structural digital twin, Mechanical Systems and Signal Processing, 191 (2023) 110164. https://doi.org/10.1016/j.ymssp.2023.110164

[10] J. Chen, S. Yuan, C. Sbarufatti, X. Jin, Dual crack growth prognosis by using a mixture proposal particle filter and on-line crack monitoring, Reliability Engineering & System Safety, 215 (2021) 107758. https://doi.org/10.1016/j.ress.2021.107758

[11] T. Li, L. Lomazzi, F. Cadini, C. Sbarufatti, J. Chen, S. Yuan, Numerical simulation-aided particle filter-based damage prognosis using Lamb waves, Mechanical Systems and Signal Processing, 178 (2022) 109326. https://doi.org/10.1016/j.ymssp.2022.109326

[12] T. Li, J. Chen, S. Yuan, F. Cadini, C. Sbarufatti, Particle filter-based damage prognosis using online feature fusion and selection, Mechanical Systems and Signal Processing, 203 (2023) 110713. https://doi.org/10.1016/j.ymssp.2023.110713

[13] X. Zhou, S. He, L. Dong, S.N. Atluri, Real-Time Prediction of Probabilistic Crack Growth with a Helicopter Component Digital Twin, AIAA Journal, 60 (2022) 2555-2567. https://doi.org/10.2514/1.J060890

[14] M. Chiachío, J. Chiachío, S. Sankararaman, K. Goebel, J. Andrews, A new algorithm for prognostics using Subset Simulation, Reliability Engineering & System Safety, 168 (2017) 189-199. https://doi.org/10.1016/j.ress.2017.05.042

[15] D. Cristiani, C. Sbarufatti, M. Giglio, Damage diagnosis and prognosis in composite double cantilever beam coupons by particle filtering and surrogate modelling, Structural Health Monitoring, (2020) 1475921720960067. https://doi.org/10.12783/shm2019/32281

[16] T. Li, C. Sbarufatti, F. Cadini, Multiple local particle filter for high-dimensional system identification, Mechanical Systems and Signal Processing, 209 (2024) 111060. https://doi.org/10.1016/j.ymssp.2023.111060

[17] T. Li, F. Cadini, M. Chiachío, J. Chiachío, C. Sbarufatti, Particle filter-based delamination shape prediction in composites subjected to fatigue loading, Structural Health Monitoring, 22 (2023) 1844-1862. https://doi.org/10.1177/14759217221116041

[18] P. Kundu, A.K. Darpe, M.S. Kulkarni, A review on diagnostic and prognostic approaches for gears, Structural Health Monitoring, 20 (2020) 2853-2893. https://doi.org/10.1177/1475921720972926

[19] B. Wang, Y. Lei, N. Li, N. Li, A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings, IEEE Transactions on Reliability, 69 (2020) 401-412. https://doi.org/10.1109/TR.2018.2882682

[20] B. Wang, Y. Lei, T. Yan, N. Li, L. Guo, Recurrent convolutional neural network: A new framework for remaining useful life prediction of machinery, Neurocomputing, 379 (2020) 117-129. https://doi.org/10.1016/j.neucom.2019.10.064

[21] G. Galanopoulos, N. Eleftheroglou, D. Milanoski, A. Broer, D. Zarouchas, T. Loutas, A novel strain-based health indicator for the remaining useful life estimation of degrading composite structures, Composite Structures, 306 (2023) 116579. https://doi.org/10.1016/j.compstruct.2022.116579

[22] M. Moradi, F.C. Gul, D. Zarouchas, A novel machine learning model to design historical-independent health indicators for composite structures, Composites Part B: Engineering, 275 (2024) 111328. https://doi.org/10.1016/j.compositesb.2024.111328

[23] M. Moradi, A. Broer, J. Chiachío, R. Benedictus, T.H. Loutas, D. Zarouchas, Intelligent health indicator construction for prognostics of composite structures utilizing a semi-supervised deep neural network and SHM data, Engineering Applications of Artificial Intelligence, 117 (2023) 105502. https://doi.org/10.1016/j.engappai.2022.105502

[24] X. Liu, Y. Lei, N. Li, X. Si, X. Li, RUL prediction of machinery using convolutional-vector fusion network through multi-feature dynamic weighting, Mechanical Systems and Signal Processing, 185 (2023) 109788. https://doi.org/10.1016/j.ymssp.2022.109788

[25] Q. Ni, J.C. Ji, K. Feng, Data-Driven Prognostic Scheme for Bearings Based on a Novel Health Indicator and Gated Recurrent Unit Network, IEEE Transactions on Industrial Informatics, 19 (2023) 1301-1311. https://doi.org/10.1109/TII.2022.3169465

[26] R. Gao, L. Wang, R. Teti, D. Dornfeld, S. Kumara, M. Mori, M. Helu, Cloud-enabled prognosis for manufacturing, CIRP Annals, 64 (2015) 749-772. https://doi.org/10.1016/j.cirp.2015.05.011

[27] F. Zeng, Y. Li, Y. Jiang, G. Song, An online transfer learning-based remaining useful life prediction method of ball bearings, Measurement, 176 (2021) 109201. https://doi.org/10.1016/j.measurement.2021.109201

[28] J. Zhuang, Y. Cao, M. Jia, X. Zhao, Q. Peng, Remaining useful life prediction of bearings using multi-source adversarial online regression under online unknown conditions, Expert Systems with Applications, 227 (2023) 120276. https://doi.org/10.1016/j.eswa.2023.120276

[29] W. Mao, K. Liu, Y. Zhang, X. Liang, Z. Wang, Self-Supervised Deep Tensor Domain-Adversarial Regression Adaptation for Online Remaining Useful Life Prediction Across Machines, IEEE Transactions on Instrumentation and Measurement, 72 (2023) 1-16. https://doi.org/10.1109/TIM.2023.3265109

[30] C. Liu, L. Zhang, Y. Zheng, Z. Jiang, J. Zheng, C. Wu, Online industrial fault prognosis in dynamic environments via task-free continual learning, Neurocomputing, 598 (2024) 127930. https://doi.org/10.1016/j.neucom.2024.127930