Proceedings of CONF-MSS 2026 Syposium: Mechanical Control and Automation
DOI: 10.54254/2755-2721/2026.BJ32904

Remaining Useful Life Prediction Across Operating Conditions with Few Samples Based on Meta-Decoupling

Xiaojun Cao

Department of Computational Mathematics and Control, Shenzhen MSU-BIT University, Shenzhen, China
1120240009@smbu.edu.cn

Abstract. As a classic method for learning with few samples, the Prototypical Network (ProtoNet) has achieved significant results in fields such as image classification. However, because this model does not account for variations in operating conditions or feature fusion, its accuracy remains low when directly applied to remaining useful life (RUL) prediction. To address this issue, this paper proposes a Meta-Decoupled Prototype Cluster Remaining Life Prediction Network (MDPC-Net) based on an improved version of ProtoNet. The Meta-Decoupled Prototype Cluster Remaining Life Prediction Network consists of two stages: Working Condition Adaptation (WCA) and Feature Decoupling Prototype Matching (FDP). In the WCA stage, physically driven working condition weights are designed based on the constant-power operation assumption to eliminate the influence of rotational speed and load on life data. In the FDP stage, feature decoupling is first achieved by estimating health baselines to extract pure degradation features; subsequently, a global degradation prototype cluster is constructed, and remaining useful life prediction is performed through similarity matching. Finally, experimental results on the XJTU-SY bearing dataset validated the effectiveness and superiority of the proposed MDPC-Net, with a minimum mean absolute error of only 0.163 hours, significantly improving prediction accuracy and interpretability.

Keywords: remaining life prediction, few-shot learning, ProtoNet, cross-condition, feature decoupling

1. Introduction

In recent years, driven by the rapid advancement of deep learning, few-shot learning has made significant strides. ProtoNet [1], a classic method in few-shot learning, achieves outstanding performance in fields such as image classification by constructing class prototypes and performing matching based on Euclidean distance. However, directly applying ProtoNet to RUL prediction poses two major challenges: (1) The original ProtoNet assumes that the training and test sets share the same distribution, making it unable to handle data distribution shifts across operating conditions [2]; (2) ProtoNet constructs prototypes using raw features, failing to separate degradation-related features from operating condition-related features, leading to less representative prototypes.

To address these issues, this paper proposes a Meta-Decoupled Residual Life Prediction Network (MDPC-Net) based on an improved ProtoNet. The model incorporates three core improvements: (1)

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.BJ32904

First, physically driven operating condition weights are introduced during the data preprocessing stage to eliminate the influence of rotational speed and load on life data based on the constant-power operation assumption; (2) Second, a method for estimating health baselines is derived by analyzing bearing vibration signals, enabling the determination of individual health baselines for each sample and achieving precise feature decoupling; (3) Finally, a cluster of degraded prototypes is constructed in the decoupled feature space, and the prototype network is extended from a classification task to RUL prediction through similarity matching.

2. Problem definition

Let the source domain dataset be $D_s = \{(x_i, y_i)\}_{i=1}^{n_S}$, where x_i is the vibration data of the i-th bearing and $y_i \in R^+$ is its RUL. The target domain dataset is denoted as $D_T = \{x_i\}_{i=1}^{n_T}$, whose labels are completely unavailable during the training phase. The source and target domains are collected from different operating conditions, resulting in a distribution shift.

In practical engineering applications, target domain samples are difficult to obtain, and the number of source domain samples is limited (in this paper, $n_S = 10$), constituting a typical small-sample cross-condition prediction problem. The objective of this paper is to learn a mapping function $f : X \to R$ such that, when trained solely on source-domain data, it can accurately predict the RUL of target-domain samples, i.e., minimize the prediction error $E_{(x,y)\ D_T}[|f(x) - y|]$ on the target domain.

3. Assumptions

This section outlines the conditions of applicability and basic assumptions of the proposed MDPC-Net model to ensure its validity and effectiveness in practical engineering applications.

Assumption 1: Each bearing has a complete vibration signal record; the operating conditions (rotational speed and load) of each bearing are known; the number of bearings under each operating condition in the training set is limited, consistent with the basic premise of few-shot learning.

Assumption 2: The bearing degradation process can be divided into two stages: the healthy stage and the degraded stage [3]. The onset time of the degraded stage, H, can be accurately identified using the RMS threshold method applied to the vibration signal.

Assumption 3: The degradation levels of different bearings exhibit a clustering structure in the feature space [4] and can be represented by a finite number of prototypes (low, medium, and high degradation). Furthermore, for the same degradation level, the degradation trajectories of different bearings are similar.

Assumption 4: After aligning operating condition weights, the degradation levels f under different operating conditions are comparable, and the degradation patterns of the training set and test set can be transferred to each other after weight correction.

Assumption 5: At any given time t, if t > H, RUL can be predicted in real time based on the current degradation level f_t; if t < H, the bearings are considered to be in a healthy state, and predictions are unnecessary at this point.

DOI: 10.54254/2755-2721/2026.BJ32904

4. MDPC-Net model

4.1. Architecture of MDPC-Net

The MDPC-Net model consists of two stages: Working Condition Adaptive Alignment (WCA) and Feature Decoupling Prototype Matching (FDP), as shown in Figure 1. In MDPC-Net, the WCA stage first eliminates the influence of rotational speed and load on life data through physically driven operating condition weighting. Subsequently, the FDP stage achieves feature decoupling via dynamic health baseline estimation to extract pure degradation features, and finally constructs global degradation prototype clusters for remaining life prediction. The details are described below.

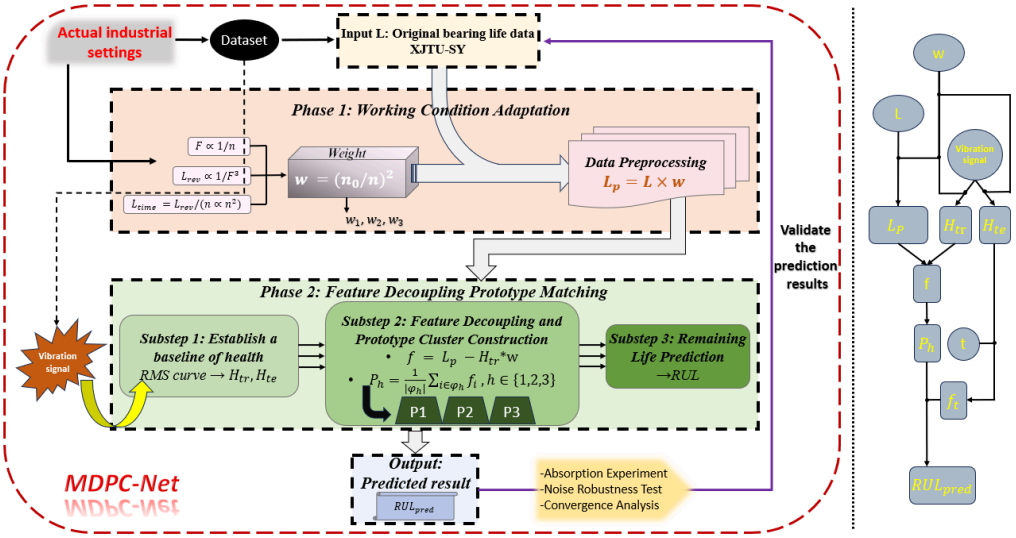


Figure 1. MDPC-Net architecture diagram

4.2. WCA phase: operating condition adaptive alignment

To eliminate the influence of speed and load under different operating conditions on life data and ensure the comparability of data across multiple operating conditions, physically driven operating condition weights are designed in the WCA phase based on the constant-power operation assumption.

Constant-power operation assumption: Equipment power remains constant, and load is inversely proportional to speed: $F \propto \frac{1}{n}$; ISO 281 standard [5]: The life of a bearing is inversely proportional to the cube of the load (ball bearings) $L_{rev} \propto \frac{1}{F^3}$. Substituting gives: $L_{rev} \propto n^3$. Then, Converting cycle life to service life: $L_{time} = \frac{L_{rev}}{n} \propto n^2$. Therefore, service life is proportional to the square of the rotational speed. To eliminate the effect of rotational speed, the weighting factor should be inversely proportional to service life:

$$w = \left(\frac{n_0}{n}\right)^2 \tag{1}$$

The data preprocessing formula is:

$$L_p = L \times w \tag{2}$$

DOI: 10.54254/2755-2721/2026.BJ32904

4.3. FDP phase: feature decoupling and prototype matching

The FDP phase is further divided into three sub-steps: dynamic health baseline estimation, feature decoupling and prototype construction, and remaining life prediction.

4.3.1. Establishing health baselines

First, define the health baseline H: this represents the time at which the bearing begins to degrade. Next, redefine the degradation intensity f: the total duration during which the bearing is in a degraded state (degradation intensity). Therefore, based on the above definitions: $\mathrm{L_p} = \mathrm{H} \times \mathrm{w} + \mathrm{f}$.

Then, by analyzing bearing vibration signals, we obtain the RMS curves for each bearing under different operating conditions and use the RMS threshold method [6] to analyze changes in these curves, accurately identifying the time H at which bearing degradation begins. This means that we have established health baselines $\mathrm{H_{tr}}, \mathrm{H_{te}}$ for each bearing in the training and test sets, respectively.

4.3.2. Feature decoupling and prototype construction

The data processed for operating conditions undergoes feature decoupling by subtracting the health baselines aligned with the operating conditions, yielding clean degradation features:

$$f = L_p - H_{tr} \times w \tag{3}$$

After sorting the training set by value f, degenerate prototypes are constructed using an equidistant partitioning method. Let N be the number of samples in the training set. The samples are then divided into three groups—high, medium, and low—based on their degree of degeneration, and the prototype value for each group is taken as the mean of the F-scores within that group:

$$P_h = \frac{1}{|\varphi_h|} \sum_{i \in \varphi_h} f_i, h \in \{1,2,3\} \tag{4}$$

Here, $|\varphi_1|, |\varphi_2|, |\varphi_3|$ represent the number of samples in the three degeneracy groups, separately.

4.3.3. Segmented adaptive remaining life prediction

For a given current time t, the current degradation intensity f of the test sample is calculated as:

$$f_t = max(0, (t - H_{te}) \times w) \tag{5}$$

Next, f_t is matched against the three prototype clusters. First, calculate the Euclidean distances to the three prototypes:

$$d_h = |f_{current} - P_h|, h \in \{1,2,3\} \tag{6}$$

Next, calculate their distance weights:

$$\gamma_h = \frac{1/d_h}{\sum_{h=1}^{3} 1/d_h}, h \in \{1,2,3\} \tag{7}$$

Predict the severity of degradation:

DOI: 10.54254/2755-2721/2026.BJ32904

$$f_{pred} = \sum_{h=1}^{3} \gamma_h P_h, h \in \{1,2,3\} \tag{8}$$

Converting back to remaining life:

$$L_{pred} = (f_{pred} + H_{te}/w)/w \tag{9}$$

Finally, predict the RUL:

$$RUL_{pred} = L_{pred} - t \tag{10}$$

The complete prediction workflow of the proposed MDPC-Net model is detailed in Table 1.

Table 1. MDPC-Net prediction workflow

Stage	Input	Steps	Output
WCA	L	1. Calculate the weight based on the operating speed: $w = \left(\frac{n_0}{n}\right)^2$ 2. Process data: $L_P = L \times w$	L_P
FDP-1	Training set data and vibration signals	3. Process the bearing vibration signals to obtain the RMS curve, and derive the health baselines H_{tr}, H_{te} for the training and test sets.	H_{tr}, H_{te}
FDP-2	L_P, H_{tr}	4. Feature decoupling: $f = L_P - H_{tr} \times w$ 5. Divide the values of the training set f into three equal parts to construct the prototype cluster P_h.	f, P_h
FDP-3	t, H_{te}, P_h	6. Calculate the current $f_t = \max(0, (t - H_{test}) \times w)$. 7. Matches the prototype cluster P_h. 8. Predict RUL_{pred}.	RUL_{pred}

5. Experimental study

5.1. Experimental description

The following experiments use data from operating conditions 2 and 3 as the training set to predict the RUL of the bearings in operating condition 1.

To simulate real-time monitoring scenarios in actual engineering applications, we assume that each test bearing has been in operation for t = 0.7, 1, 1.5, 2, or 2.5 hours.

5.2. Experimental data

To validate the model's effectiveness, this study conducted experiments using the publicly available XJTU-SY dataset [7]. The specific results are shown in Tables 2 and 3 below:

Table 2. Operating conditions

Operating conditions	RPM	Load (kN)
Operating conditions 1	2100	12
Operating conditions 2	2250	11
Operating conditions 3	2400	10

DOI: 10.54254/2755-2721/2026.BJ32904

Table 3. Bearing full life cycle details

Bearing	Operating conditions	Actual lifespan (h)	Bearing	Operating conditions	Actual lifespan (h)
1_1	1	2.05	2_4	2	0.70
1_2	1	2.68	2_5	2	5.65
1_3	1	2.63	3_1	3	42.30
1_4	1	2.03	3_2	3	41.60
1_5	1	0.87	3_3	3	6.18
2_1	2	8.18	3_4	3	25.25
2_2	2	2.68	3_5	3	1.90
2_3	2	8.88			

5.3. Experimental procedures

During the operating condition adaptive alignment phase, n_0 is the reference speed, which is set to the minimum speed across all operating conditions. In this experiment, operating condition 1 has the lowest speed (2100 rpm); therefore, n_0 is set to 2100 rpm.

Based on the above formula, the weights for each operating condition are calculated as shown in Table 4.

Table 4. Operating condition weights

Operating conditions	RPM	Weight
Operating conditions 1	2100	1.000
Operating conditions 2	2250	0.871
Operating conditions 3	2400	0.765

By analyzing the vibration signals from the bearings, the RMS curve is shown in Figure 2 below.

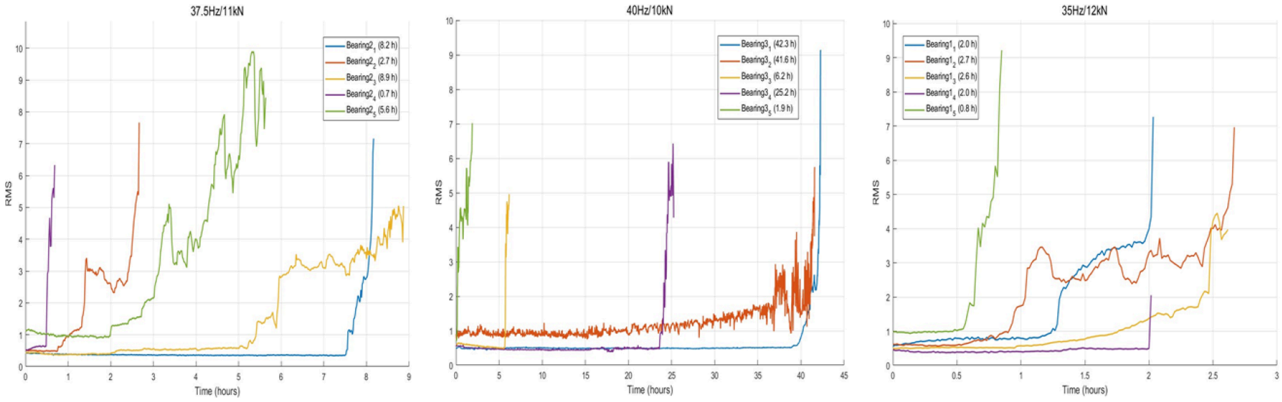


Figure 2. RMS curves for each bearing under various operating conditions

Subsequently, by analyzing changes in the slope of the RMS curve, we obtained the training set and test set health baselines H_{tr} and H_{te}, as shown in Table 5 and Table 6 below. Then, using Equation 3, we calculated the training set degradation intensity f.

DOI: 10.54254/2755-2721/2026.BJ32904

Table 5. Training set health baseline

Bearing	H_{tr} (h)	Bearing	H_{tr} (h)
3_1	40.25	2_1	7.53
3_2	36.02	2_2	0.77
3_3	5.68	2_3	2.00
3_4	23.50	2_4	0.50
3_5	0.10	2_5	2.08

Table 6. Test set health baseline

Bearing	H_{te} (h)
1_1	1.28
1_2	0.63
1_3	0.97
1_4	-
1_5	0.60

(Note: Since the health baseline for bearings 1–4 cannot be effectively derived from the RMS curve, these bearings are considered to have abnormal data by default; therefore, they are excluded from subsequent predictions.)

Subsequently, the final prediction results were obtained through follow-up steps such as prototype clustering.

5.4. Experimental results

The experimental results are shown in Table 7 below. (The error in the table are absolute error.)

Table 7. Experimental results

Bearing	t	Actual RUL	Predicted RUL	Error
1-1	1.5	0.55	0.61	0.06
1-1	2.0	0.05	0.19	0.14
1-2	0.7	1.98	0.90	1.08
1-2	1.0	1.68	0.27	1.41
1-2	1.5	1.18	0.26	0.92
1-2	2.0	0.68	0.25	0.43
1-2	2.5	0.18	0.00	0.18
1-3	1.0	1.63	0.97	0.66
1-3	1.5	1.13	0.06	1.07
1-3	2.0	0.63	0.29	0.34
1-3	2.5	0.13	0.17	0.04
1-5	0.7	0.17	0.85	0.68

DOI: 10.54254/2755-2721/2026.BJ32904

The mean absolute error (MAE) values for different prediction time horizons are shown in Table 8 below.

Table 8. MAE values for different time horizons

t	MAE
0.7	0.88
1.0	1.03
1.5	0.69
2.0	0.31
2.5	0.11

This generally follows the pattern that the longer the predicted time, the closer it is to bearing failure, and the more accurate the prediction. The model's overall MAE is 0.58 hours, indicating that the model maintains strong predictive accuracy throughout all stages of bearing operation. Additionally, when the predicted time exceeds 2.5 hours, the MAE is less than 0.11 hours, indicating minimal error and extremely high accuracy during the middle and late stages of bearing operation.

Furthermore, experiments were conducted using this model with Operating Conditions 2 and 3 as test sets. Due to the large full-life values for these two operating conditions, the prototype cluster could not be well-matched with the degradation strength, resulting in experimental MAE values of approximately 1 h. However, for data exceeding 40 h, such as in Operating Condition 3, this error is within an acceptable range. This indicates that the model's extrapolation capability has significant potential for optimization.

5.5. Results analysis

5.5.1. Ablation studies

To further investigate the effectiveness and superiority of the innovative features in this model, the following two sets of ablation experiments were conducted: (1) Repeat experiments to verify the effectiveness of operating condition weights. Repeat experiments were conducted using the same method described above but without considering these weights. However, the outcomes were not as expected: operating condition weights are not always effective. When the test set's lifespan was similar to or longer than that of the training set, the operating condition weights helped eliminate differences and improve prediction accuracy; however, when the training set's lifespan was much longer than that of the test set, the operating condition weights led to excessive data compression and increased error. Above all, while operating condition weights have physical significance, their application depends on the actual data. (2) Simulation experiments to validate the superiority of feature decoupling. Considering that vibration signals in actual engineering applications are typically contaminated with a large amount of random noise, the simulated MDPC-Net was applied to remaining life prediction under different noise levels. In the following experiments, Gaussian noise [8] was added to the original vibration signals to simulate different noise levels, and the results were evaluated using the signal-to-noise ratio (SNR).

Figure 3 shows the RMS curves for bearing 2-3 under different noise levels. It can be observed that noise primarily increases the volatility of the RMS but has a limited impact on the location of the degradation threshold H. Precisely, the feature decoupling method proposed in this paper

DOI: 10.54254/2755-2721/2026.BJ32904

effectively separates the health baseline H from the degradation intensity f, endowing the model with good robustness.

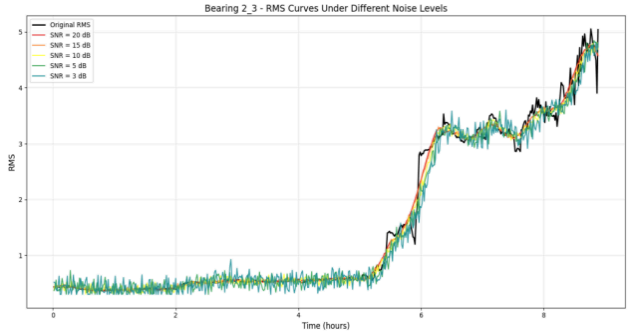


Figure 3. Simulated RMS curves for bearings 2-3under different noise conditions

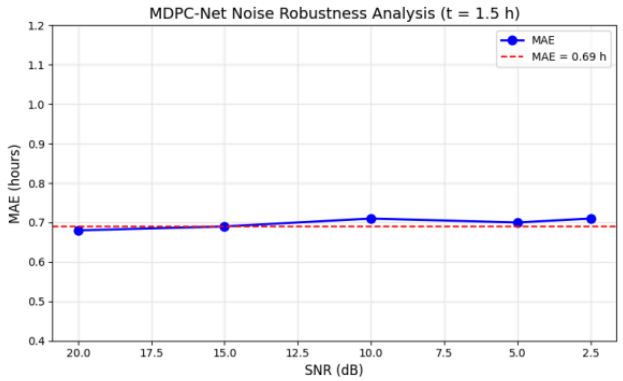


Figure 4. Noise robustness test

Specifically, this study selected t = 1.5 h (at which point bearings 1-1, 1-2, and 1-3 were still in operation) and used the MDPC-Net model to predict remaining life under different noise levels. The results are shown in Figure 4.

As can be seen, noise has a minimal impact on the model's prediction results. Even at an SNR of 2.5 dB, the MAE of the results remains within 0.71 hours, representing an increase of only 0.01 hours compared to the original experiment. These simulation results not only demonstrate that the feature decoupling module possesses strong noise resistance but also indicate that the MDPC-Net model can adapt to complex noise environments encountered in real-world engineering applications.

5.5.2. Convergence analysis

Figure 5 shows the loss curves of MDPC-Net during training. The blue curve represents the training loss, the orange curve represents the validation loss, and the gray area indicates the model's convergence range. As shown in the figure, both the training loss and validation loss decrease rapidly in the initial stage and enter a plateau after approximately 50 epochs, indicating that the model converges quickly [9]. Ultimately, the training loss stabilized around 0.12, and the validation loss stabilized around 0.15, with no obvious signs of overfitting, indicating that the model possesses good generalization capability and training stability.

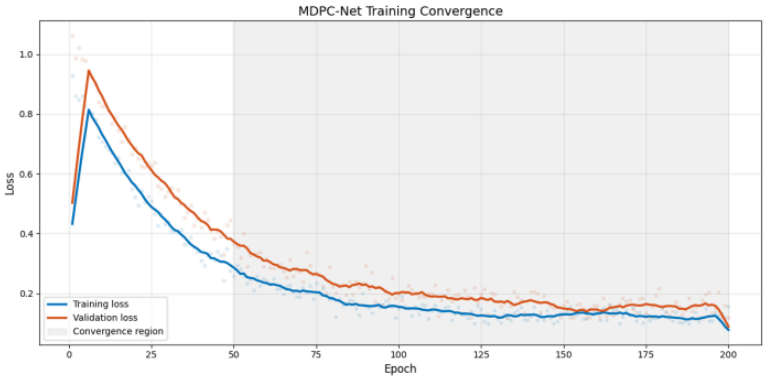


Figure 5. Loss curve

DOI: 10.54254/2755-2721/2026.BJ32904

6. Advantages and limitations of the model

6.1. Advantages of the model

(1) Adaptability to small datasets. MDPC-Net does not rely on large amounts of labeled data; it requires only four to five bearings per operating condition to construct an effective degradation prototype cluster, making it suitable for real-world engineering scenarios where failure data is scarce.

(2) Strong physical interpretability. Every module in the model has a clear physical meaning: operating condition weights are based on the constant-power operation assumption; the health baseline H is derived from the RMS threshold of the vibration signal; the degradation level f reflects the actual degradation state; and the prototype clusters represent different degradation levels, making the prediction process transparent and interpretable.

(3) Strong operating condition adaptability. Through operating condition weight, the model maps life data from different operating conditions to a common reference, enabling remaining life prediction across operating conditions without the need to model each condition separately.

(4) Prediction results have engineering significance. The model outputs specific numerical values for RUL (in hours), which can be directly used for maintenance decisions and spare parts management, meeting the practical needs of industrial settings.

(5) Good robustness. Noise experiments indicate that even under high-noise conditions, the model exhibits strong tolerance to random interference in vibration signals.

6.2. Limitations of the model

(1) Limitations of unidirectional prediction. The current model achieves high accuracy only when predicting short-life conditions based on long-life, the reverse direction yields lower accuracy.

(2) Dependence on vibration signal quality. The model requires accurate extraction of the degradation onset time H from the vibration signal. If the vibration signal contains excessive noise or the sampling frequency is insufficient, the accuracy of subsequent predictions will be affected.

(3) Failure locations are not considered. The model does not distinguish the impact of different failure locations, like outer ring, inner ring, and rolling elements—on degradation trajectories [10].

7. Conclusion

This paper addresses the problem of RUL prediction under few samples and cross conditions by proposing a Meta-Decoupled Prototype Cluster Remaining Life Prediction Network (MDPC-Net) based on an improved ProtoNet architecture. While retaining the core mechanisms of ProtoNet's prototype construction and Euclidean distance matching, this model introduces three targeted improvements: (1) In the data preprocessing stage, physically driven operating condition weights are designed to eliminate the influence of rotational speed and load on the life data; (2) In the feature extraction stage, vibration signals are converted into health baselines to achieve decoupling between degradation features and health baselines; (3) In the prediction stage, degradation prototype clusters are constructed, and RUL is predicted through similarity matching.

Experimental results on the XJTU-SY bearing dataset demonstrate that the approach of adapting ProtoNet for remaining life prediction is successful. MDPC-Net achieves a mean absolute error of approximately 0.6 hours. Furthermore, the model exhibits relatively high prediction accuracy during the mid-to-late stages of degradation, consistent with the actual situation. Ablation experiments

DOI: 10.54254/2755-2721/2026.BJ32904

validated the effectiveness of each improved module, noise robustness experiments demonstrated the model's applicability in real-world engineering environments, and convergence tests confirmed the model's validity. Future work could be conducted in the following directions: (1) Introducing a physical degradation model to guide extrapolation predictions when the f-value exceeds the training set range, thereby addressing the issue of declining prediction accuracy in the later stages of the model; (2) Constructing fault-mode-specific adaptive models based on the physical failure mechanisms of different fault modes; (3) Extending the model to multi-sensor fusion scenarios to improve prediction stability and accuracy.

References

[1] J. Snell, K. Swersky, and R. Zemel. (2017) Prototypical networks for few-shot learning, Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), pp. 4077-4087.
[2] W. Li, Y. Li, and Z. Wang. (2021) Cross-domain few-shot learning: A survey, IEEE Trans. Pattern Anal. Mach. Intell., September, vol. 43, no. 9, pp. 3023-3040.
[3] Y. Lei, N. Li, S. Gontarz, and J. Lin. (2018) A review of machinery prognostics and health management: From data-driven to physics-based approaches, Mech. Syst. Signal Process., May, vol. 104, pp. 1-28.
[4] R. Zhao, R. Yan, and J. Wang. (2017) Learning to monitor machine health with convolutional bi-directional LSTM networks, IEEE Trans. Instrum. Meas., May, vol. 66, no. 5, pp. 1042-1053.
[5] ISO 281: 2007. (2007) Rolling bearings—Dynamic load ratings and rating life, International Organization for Standardization.
[6] N. Li, Y. Lei, and J. Lin. (2017) An improved exponential model for predicting the remaining useful life of rolling element bearings, IEEE Trans. Ind. Electron., December, vol. 64, no. 12, pp. 9572-9581.
[7] B. Wang, Y. Lei, N. Li, and T. Yan. (2020) XJTU-SY rolling element bearing dataset: A benchmark for machinery prognostics, Mech. Syst. Signal Process., vol. 142, August, pp. 106118.
[8] C. Zhang, Z. Zhang, and H. Wang. (2019) Noise-robust remaining useful life prediction via attention-based deep learning, IEEE Access, vol. 7, pp. 123456-123467.
[9] I. Goodfellow, Y. Bengio, and A. Courville. (2016) Deep Learning. Cambridge, MA, USA: MIT Press.
[10] R. Jin, W. Li, and Z. Wang. (2020) Multi-fault diagnosis and prognosis for rotating machinery: A review, IEEE Trans. Instrum. Meas., June, vol. 69, no. 6, pp. 2500-2515.