Letter

RGCNU: Recurrent Graph Convolutional Network With Uncertainty Estimation for Remaining Useful Life Prediction

Qiwu Zhu, Qingyu Xiong, Zhengyi Yang, and Yang Yu

Dear Editor,

This letter focuses on the problem of remaining useful life (RUL) prediction of equipment. Existing graph neural network (GCN)-based approaches merely provide the point estimation of RUL. However, the estimated RUL often varies widely due to the model parameters and the noise in data. It is important to know the uncertainty in predictions for reliable risk analysis and maintenance decision making. To map the relationship between noisy condition monitoring data and RUL with uncertainty, we propose a recurrent graph convolutional network with uncertainty estimation (RGCNU) for RUL prediction. In our approach, the correlation exploiting module captures the spatial-temporal correlations based on the learned graph structure. Furthermore, the fusion module associates the RUL prediction and data uncertainty to improve the robustness of the model to noisy data.

In the long-term operation process, the reliability of mechanical equipment will gradually decline [1]. RUL prediction plays an increasingly important role in improving equipment operation reliability. RUL prediction methods are usually composed of two major types: model-based methods and data-driven methods [2]. The main idea of model-based methods is to use physical mechanisms or statistical knowledge to describe the degradation process. They can include exponential model, Paris-Erdogan model, and Gamma process model. In practice, however, even if the same system under the same operating state may be a significant difference in degradation process. Also, additional prior knowledge and experience are often required for specific situations. Therefore, these limitations make the model-based method unsuitable for a large-scale application.

Data-driven methods attract more attention from scholars. Without requiring expert knowledge, machine learning-based algorithms can construct a mapping relationship between the monitoring data and the equipment degradation process to determine the working condition of equipment. Neural networks [3] have been widely used in equipment health monitoring and achieve good results in the past several years. For example, convolution neural network (CNN) can extract the temporal features from multi-sensors [4]. As another neural network structure, recurrent neural network (RNN) can capture long-term dependency on the degradation process and model sequential data.

Recently, GCN has gradually attracted researcher's attention due to its powerful node representation ability [5]. GCN-based approaches have also been used in RUL prediction. For example, Wang *et al.* [6] proposed a fixed GCN model for RUL prediction, in which CNN and RNN learn the spatial and sequence relations, respectively. Li *et al.* [7] presented a hierarchical GCN to fuse the representations from condition monitoring data for engine RUL prediction.

However, there are two drawbacks in existing GCN-based studies for RUL prediction. One is that these studies decouple the spatial-temporal correlations of condition monitoring data, leading to loss of some implicit information. The other is that these studies ignore the uncertainty modeling for noisy data, which can improve the model robustness and provide additional decision information.

To this end, a recurrent graph convolutional network with uncertainty estimation is proposed for RUL prediction in this letter. To our best knowledge, this letter is the first to apply GCN and uncertainty estimation simultaneously in RUL prediction. To discover the hidden correlation among sensors, a graph learning module first builds the adjacency matrix, which is input into correlations exploiting module. Then, the correlations exploiting module captures spatial-temporal dependencies in condition monitoring data. According to the learned representation, the fusion module associates the RUL prediction and the data uncertainty to improve the robustness of the network to noisy data. Besides, Monte Carlo dropout (MCDO) is used to estimate the model uncertainty of the network. Finally, the above two types of uncertainty are combined for uncertainty estimation of RUL prediction.

Our contributions are described as follows: 1) We propose a correlation exploiting module to capture the spatial-temporal correlations in the condition monitoring data; 2) We design a fusion module for associating the RUL prediction and the data uncertainty to improve the robustness of network to noisy data; 3) Extensive experiments on C-MAPSS show the effectiveness of the proposed RGCNU.

Preliminaries: Given the historical condition monitoring data recorded by sensors on the equipment, we focus on predicting the RUL and uncertainty of the new condition monitoring data.

We consider an undirected graph $G=(V,E,A)$ where V is a finite set of $|V|=N$ nodes, E is a set of edges and fully connects different nodes, and $A\in\mathbb{R}^{N\times N}$ is the adjacency matrix of the whole graph. From a graph-based perspective, all the sensors are considered as nodes in the graph, and the relationships between sensors are described using the adjacency matrix A. Let $x_t\in\mathbb{R}^N$ is the measured values of all sensors at the time t, where x_{ti} is the measured value of i-th sensor. $X^j=\{x_1^j,x_2^j,\ldots,x_F^j\}\in\mathbb{R}^{N\times F}$ is the j-th sequence of F time steps of condition monitoring data, y_j is the corresponding RUL label. With the dataset $\mathbf{D}=\{(X^1,y_1),\ldots,(X^j,y_j),\ldots\}$, the task of this letter is to learn a mapping graph model $f(\cdot)$ from dataset $\mathbf{D}$ to accurately predict the RUL and estimate the corresponding uncertainty.

Method: RGCNU is proposed to address the problem of spatial-temporal correlations exploiting and uncertainty estimation in RUL prediction. As shown in Fig. 1, the proposed framework mainly consists of three modules: the graph learning module, correlations exploiting module, and fusion module. To discover the hidden correlation among sensors, a graph learning module builds the adjacency matrix, which is input into the next module. The correlations exploiting module aims to capture spatial-temporal dependencies in condition monitoring data. The fusion module performs the RUL prediction and data uncertainty estimation based on the learned representation. The framework is described in detail as follows.

Graph learning module: The graph learning module adaptively builds an adjacency matrix to discover the hidden correlation among sensors. To build up the graph, the learning process is described as follows.

For the feature matrix $X\in\mathbb{R}^{N\times F}$, the intermediate variables A_1, A_2 can be expressed as

$$A_1=\tanh(\alpha X\Theta_1) \tag{1}$$

$$A_2=\tanh(\alpha X\Theta_2) \tag{2}$$

where $\Theta_1,\Theta_2\in\mathbb{R}^{F\times N}$ represent learnable parameters, tanh is the activation function, and α is the hyper-parameter.

Corresponding author: Qingyu Xiong.

Citation: Q. W. Zhu, Q. Y. Xiong, Z. Y. Yang, and Y. Yu, "RGCNU: Recurrent graph convolutional network with uncertainty estimation for remaining useful life prediction," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 7, pp. 1640–1642, Jul. 2023.

Q. W. Zhu, Q. Y. Xiong, and Z. Y. Yang are with the School of Big Data &Software Engineering, Chongqing University, Chongqing 400044, and also with the Key Laboratory of Dependable Service Computing in Cyber Physical Society, Ministry of Education, Chongqing 400044, China (e-mail: qiwuzhu@163.com; cquxqy@163.com; zyyang@cqu.edu.cn).

Y. Yu is with the School of Biology and Engineering, Guizhou Medical University, Guiyang 550025, China (e-mail: cquyy@cqu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2023.123369

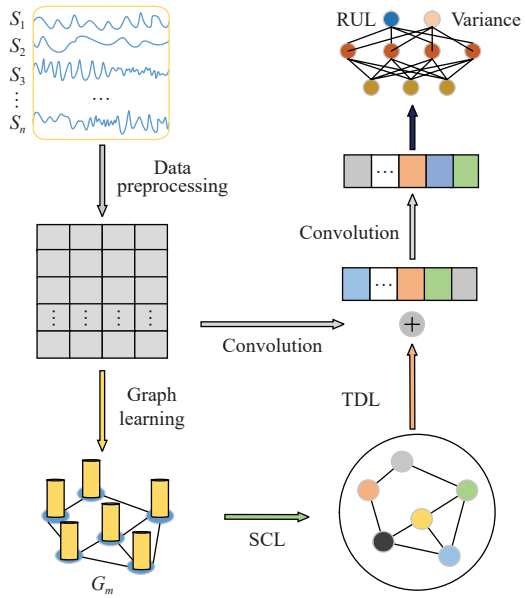

Fig. 1. The framework of proposed approach.

The adjacency matrix A is computed as

$$A = \mathrm{ReLU}\left(\tanh\left(\alpha\left(A_1 A_2^T - A_2 A_1^T\right)\right)\right) \tag{3}$$

where ReLU is the activation function.

Correlation exploiting module: The correlation exploiting module aims to capture the spatial-temporal correlation and consists of spatial correlation learning (SCL) and temporal dependency learning (TDL).

To fully benefit from high-order neighbors' information, a two-layer fully-connected GCN is employed to combine the features from itself and different orders of neighbors for learning spatial correlation. Specifically, the spatial features of the sensors network at each time slice are captured.

The condition monitoring data is regarded as the multi-sensors time-series data. In the SCL, GCN represents the mutual spatial correlations among all sensors, and the dropout layer is followed. In TDL, one-dimensional convolution is first appended to combine the internal feature of each sensor, and then LSTM is adopted in the time dimension for temporal dependency.

Fusion module: The fusion module integrates the residual information from X and the representation obtained from TDL, aiming to expedite the model training and mitigate the model overfitting problem. Besides, a CNN with 1×1 kernel size is used for the preprocessed input X to guarantee that the residual information has the same dimensions as the representation obtained from TDL. To associate the predicted RUL and the estimated data uncertainty, a CNN and fully connected layers are then adopted to obtain the RUL and variance, respectively. The expression of the fusion module is expressed as

$$\mathrm{Mid} = \Phi_2 * (\mathrm{ReLU}(\Phi_1 * X + H)) \tag{4}$$

$$\hat{y} = FC_1(Flat(\mathrm{Mid})) \tag{5}$$

$$\hat{\sigma} = FC_2(Flat(\mathrm{Mid})) \tag{6}$$

where H denotes the output of correlation exploiting module, Φ_1 and Φ_2 denote the convolution kernel's parameters, $Flat$ is Flatten operator, FC_1 and FC_2 mean the fully connected layers with different weights, $\hat{y}$ and $\hat{\sigma}$ are the predicted RUL and variance, respectively.

Loss function: The cost function is reformulated as

$$\mathcal{L}(\theta) = (1-\lambda)\times\left(\frac{1}{M}\sum_{i=1}^{M}\frac{1}{2}\exp(-s_i)\|\boldsymbol{y}_i - \hat{\boldsymbol{y}}_i\|^2 + \frac{1}{2}s_i\right) + \lambda\times\left(\sum_{i=1}^{M}\exp^{2s_i}\right) \tag{7}$$

where $s_i := \log\hat{\sigma}_i^2$ is the log variance, $\lambda \in (0,1)$ is a tuning parameter.

Uncertainty estimation of RUL: To obtain the uncertainty estimation of RUL, the model uncertainty of RUL also needs to be estimated. Through using dropout at the test stage to sample from the approximate posterior [8], the predicted RUL and the corresponding uncertainty of a test sample can be expressed as

$$\mathrm{Mean}(y) \approx \frac{1}{L}\sum_{l=1}^{L}\hat{y}_l \tag{8}$$

$$\mathrm{Var}(y) \approx \frac{1}{L}\sum_{l=1}^{L}\hat{y}_l^2 - \left(\frac{1}{L}\sum_{l=1}^{L}\hat{y}_l\right)^2 + \frac{1}{L}\sum_{l=1}^{L}\hat{\sigma}_l^2 \tag{9}$$

where $\left\{\hat{y}_l, \hat{\sigma}_l^2\right\}_{l=1}^{L}$ is a set of sampled outputs after L stochastic forward passes.

Experiments:

Datasets: A series of experiments are conducted on C-MAPSS [9] to evaluate the effectiveness of the proposed approach.

Evaluation metrics: Two metrics, root mean square error (RMSE) and score function, are used to evaluate the performance of the proposed approach.

Result analysis: Fig. 2 shows the predicted RUL, the corresponding uncertainty, and the true RUL of four units in four datasets. When the engine units are in the degradation stage, it can be found that the uncertainty and the error between the true RUL and the predicted RUL will decrease. The reason is that degradation information is enhanced while the engine unit is near failure. That shows our approach can capture the information for better prediction and uncertainty estimation.

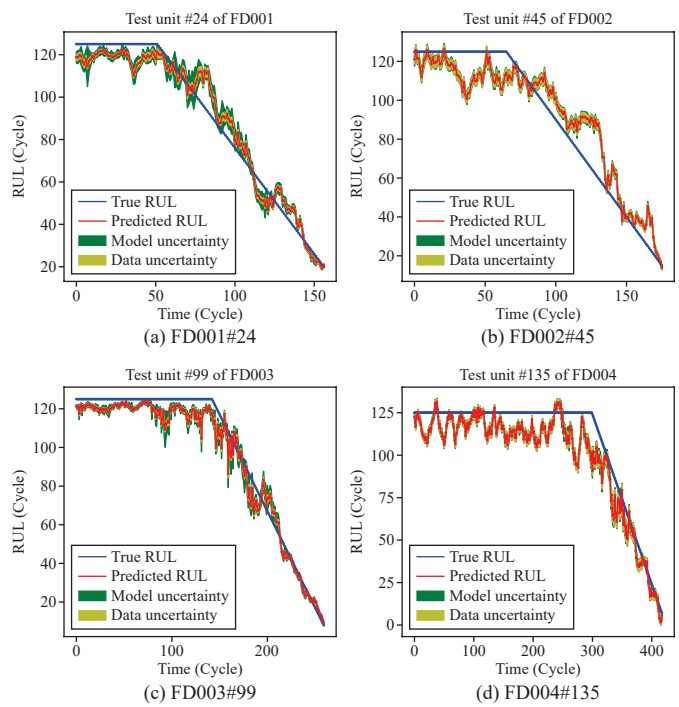

Fig. 2. Four examples of test units in different datasets.

Performance comparison: To verify the effectiveness, the proposed approach is compared with other approaches without and with uncertainty estimation. Table 1 summarizes the RMSE and score of different approaches with the same settings on four subsets. It can be found that the two metrics results on FD001 and FD004 are the smallest and largest among the four subsets, respectively. The reason is that FD001 has the fewest operating settings and fault modes. Among the existing approaches, the GCN approaches have shown their advantage in RUL prediction. These approaches can achieve better performance and lower error, the reason is that these approaches have a strong ability to learn the spatial correlation from

Table 1. The Performance of Different Approaches

Approaches	FD001		FD002		FD003		FD004	
Without	RMSE	Score	RMSE	Score	RMSE	Score	RMSE	Score
DCNN [10]	12.61	273.7	22.36	10412	12.64	284.1	23.31	12466
HALSTM [11]	14.53	322.44	N/A	N/A	N/A	N/A	27.08	5649.14
Autoencoder [12]	13.58	228	19.59	2650	19.16	1727	22.15	2901
AGCNN [13]	12.42	225.51	19.43	**1492**	13.39	**227**	21.50	3392
Transformer [14]	**11.27**	N/A	22.81	N/A	**11.42**	N/A	24.86	N/A
GCN [6]	12.76	266	N/A	N/A	12.07	278	N/A	N/A
HAGCN [7]	11.93	**222.3**	N/A	N/A	11.53	240.3	N/A	N/A
GAT [15]	13.21	303.18	N/A	N/A	15.36	507.52	N/A	N/A
With	RMSE	Score	RMSE	Score	RMSE	Score	RMSE	Score
LSTMBS [16]	14.89	481.1	26.86	7982	15.11	493.4	27.11	5200
Bayesian [17]	12.19	267.21	**18.49**	2007.81	12.07	409.39	**19.41**	**2415.71**
BDL [18]	18.6	2774.1	22.9	7734.7	27.9	19990.6	28.1	53295.6
RGCNU	**11.18**	**173.59**	**16.22**	**1148.16**	**11.52**	**225.03**	**19.11**	**2215.9**
imp	0.79%	21.91%	12.27%	23.04%	–	0.86%	1.54%	8.27%

the raw data. The improvements ($imp = \frac{r_e - r_p}{r_e}$) are calculated, where r_e is the best result of compared approaches and r_p is our result. The performances are improved on four subsets except for the RMSE of FD003. Those results indicate the proposed approach can achieve better prediction performance compared with other approaches. For example, two metrics of the proposed approach are reduced by 12.27% and 23% compared with existing results on FD002, respectively.

Ablation study: The influence of TDL and fusion module is studied by conducting several experiments on four subsets. To achieve this goal, the part of uncertainty output is removed to build a model, denoted by RGCN. Similarly, the TDL in RGCN is replaced with CNN to build another model, denoted by GCNN. Table 2 shows the performances on four subsets. It can be noted that RGCN achieves better performance than GCNN in all of the two metrics. This shows that the correlations exploiting module can better capture the spatial-temporal correlations. As for RGCN and RGCNU, RGCNU also achieves better performance overall. These results indicate that the introduced uncertainty estimation can improve the robustness of the model and make the model more reliable in RUL prediction.

Table 2. The Performance of Different Ablation Studies

Approaches	FD001		FD002		FD003		FD004	
	RMSE	Score	RMSE	Score	RMSE	Score	RMSE	Score
GCNN	12.88	244.57	18.00	1775.51	12.98	389.06	21.98	3579.88
RGCN	11.93	205.19	17.05	1186.10	12.51	288.04	19.87	2225.99
RGCNU	**11.18**	**173.59**	**16.22**	**1148.16**	**11.52**	**225.03**	**19.11**	**2215.9**

Conclusion: This letter proposes a recurrent graph convolutional network with uncertainty estimation for RUL prediction. In our approach, a graph learning module firstly builds the adjacency matrix to discover the hidden correlation among sensors. Then, the correlation exploiting module captures the spatial-temporal correlations based on the learned graph structure. Furthermore, the fusion module associates the RUL prediction and the data uncertainty to improve the robustness of the model to noisy data. Besides, MCDO estimates the model uncertainty of the network. Finally, the above two types of uncertainty are combined for uncertainty estimation of RUL prediction. Sufficient experimental results on the C-MAPSS indicate that the proposed approach achieves satisfying results compared to the existing approaches. Besides, ablation studies are conducted to show the effectiveness of several schemes.

Acknowledgments: This work was supported by the Major Special Program of Chongqing Science & Technology Commission (CSTC 2019jscx-zdztzxX0031), Graduate Scientific Research and Innovation Foundation of Chongqing (CYB21068, CYS22128).

Supplementary material: The supplementary material of this letter can be found in links https://drive.google.com/file/d/1hlCGD2JyEzucMqTAzY6TO3H50CpY2Op/view?usp=share_link.

References

[1] F. Li, T. Zheng, N. He, and Q. Cao, "Data-driven hybrid neural fuzzy network and ARX modeling approach to practical industrial process identification," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 9, pp. 1702–1705, 2022.

[2] J. Zhu, N. Chen, and W. Peng, "Estimation of bearing remaining useful life based on multiscale convolutional neural network," *IEEE Trans. Industrial Electronics*, vol. 66, no. 4, pp. 3208–3216, 2018.

[3] R. Jiao, K. Peng, and J. Dong, "Remaining useful life prediction for a roller in a hot strip mill based on deep recurrent neural networks," *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 7, pp. 1345–1354, 2021.

[4] R. Jin, M. Wu, K. Wu, K. Gao, Z. Chen, and X. Li, "Position encoding based convolutional neural networks for machine remaining useful life prediction," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 8, pp. 1427–1439, 2022.

[5] C. Zhang, H.-Y. Zhou, Q. Qiu, Z. Jian, D. Zhu, C. Cheng, L. He, G. Liu, X. Wen, and R. Hu, "Augmented multi-component recurrent graph convolutional network for traffic flow forecasting," *ISPRS Int. J. Geo-Information*, vol. 11, no. 2, p. 88, 2022.

[6] M. Wang, Y. Li, Y. Zhang, and L. Jia, "Spatio-temporal graph convolutional neural network for remaining useful life estimation of aircraft engines," *Aerospace Systems*, vol. 4, no. 1, pp. 29–36, 2021.

[7] T. Li, Z. Zhao, C. Sun, R. Yan, and X. Chen, "Hierarchical attention graph convolutional network to fuse multi-sensor signals for remaining useful life prediction," *Reliability Engineering & System Safety*, vol. 215, p. 107878, 2021.

[8] A. Kendall and Y. Gal, "What uncertainties do we need in bayesian deep learning for computer vision?" in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5574–5584.

[9] A. Saxena, K. Goebel, D. Simon, and N. Eklund, "Damage propagation modeling for aircraft engine run-to-failure simulation," in *Proc. IEEE International Conf. Prognostics Health Management*, 2008, pp. 1–9.

[10] X. Li, Q. Ding, and J.-Q. Sun, "Remaining useful life estimation in prognostics using deep convolution neural networks," *Reliability Engineering & System Safety*, vol. 172, pp. 1–11, 2018.

[11] Z. Chen, M. Wu, R. Zhao, F. Guretno, R. Yan, and X. Li, "Machine remaining useful life prediction via an attention-based deep learning approach," *IEEE Trans. Industrial Electronics*, vol. 68, no. 3, pp. 2521–2531, 2020.

[12] W. Yu, I. Y. Kim, and C. Mechefske, "An improved similarity-based prognostic algorithm for RUL estimation using an RNN autoencoder scheme," *Reliability Engineering & System Safety*, vol. 199, p. 106926, 2020.

[13] H. Liu, Z. Liu, W. Jia, and X. Lin, "Remaining useful life prediction using a novel feature-attention-based end-to-end approach," *IEEE Trans. Industrial Informatics*, vol. 17, no. 2, pp. 1197–1207, 2020.

[14] Y. Mo, Q. Wu, X. Li, and B. Huang, "Remaining useful life estimation via transformer encoder enhanced by a gated convolutional unit," *J. Intelligent Manufacturing*, vol. 32, no. 7, pp. 1997–2006, 2021.

[15] T. Li, Z. Zhou, S. Li, C. Sun, R. Yan, and X. Chen, "The emerging graph neural networks for intelligent fault diagnostics and prognostics: A guideline and a benchmark study," *Mechanical Systems and Signal Processing*, vol. 168, p. 108653, 2022.

[16] Y. Liao, L. Zhang, and C. Liu, "Uncertainty prediction of remaining useful life using long short-term memory network based on bootstrap method," in *Proc. IEEE International Conf. Prognostics Health Management*, 2018, pp. 1–8.

[17] M. Kim and K. Liu, "A Bayesian deep learning framework for interval estimation of remaining useful life in complex systems by incorporating general degradation characteristics," *IISE Transactions*, vol. 53, no. 3, pp. 326–340, 2020.

[18] Y.-H. Lin and G.-H. Li, "A Bayesian deep learning framework for RUL prediction incorporating uncertainty quantification and calibration," *IEEE Trans. Industrial Informatics*, vol. 18, no. 10, pp. 7274–7284, 2022.