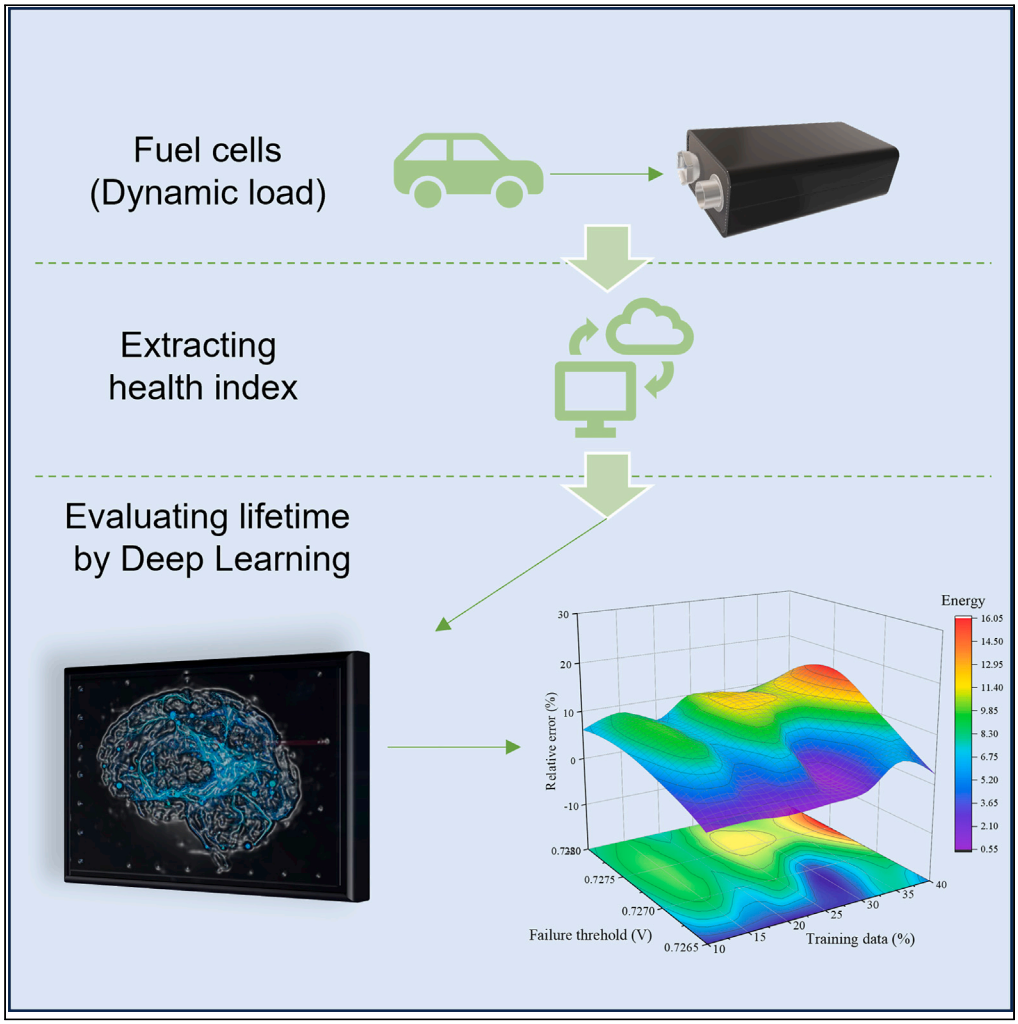
A health index-based approach for fuel cell lifetime estimation

iScience

CellPress
OPEN ACCESS

Article

A health index-based approach for fuel cell lifetime estimation



Hangyu Wu, Ruiming Zhang, Wenchao Zhu, ..., Bingxin Guo, Changzhi Li, Rui Xiong

zhuwenchao@whut.edu.cn

Highlights

A method for extracting health indicators for fuel cells under dynamic variable loads

A data-driven method based on Bayesian theory adjusts model parameters

Utilization of effective features for model training further enhances prediction accuracy

Adaptability of the prediction method is validated through comparisons of three datasets

Wu et al., iScience *27*, 110979
November 15, 2024
https://doi.org/10.1016/j.isci.2024.110979

iScience

CellPress
OPEN ACCESS

Article

A health index-based approach for fuel cell lifetime estimation

Hangyu Wu,[1] Ruiming Zhang,[2] Wenchao Zhu,[3,4,7,*] Changjun Xie,[1,4] Yang Li,[5] Yang Yang,[1,4] Bingxin Guo,[4] Changzhi Li,[4] and Rui Xiong[6]

SUMMARY

Efficient health indicators (HI) and prediction methods are crucial for assessing the remaining useful life (RUL) of fuel cells. However, obtaining HI under dynamic conditions with frequently changing loads is highly challenging. Therefore, this study proposes a prediction framework based on dynamic conditions. A method combining complete ensemble empirical mode decomposition with adaptive noise, power spectral density, and energy analysis (CPE) is proposed to extract HI under dynamic conditions from the perspectives of frequency and energy. Furthermore, the time convolution network with adaptive Bayesian optimization (AB-TCN) is introduced to address parameter optimization and prediction challenges. Effective feature parameters of the data are identified using random forest and used to train the AB-TCN. Results show that the extracted HI can effectively determine the end-of-life. The AB-TCN achieves accurate RUL estimation with a prediction error of only 6.825% and shows strong adaptability to various prediction tasks.

INTRODUCTION

Background and literature review

In recent years, escalating global environmental challenges have compelled people to seek sustainable, clean, and efficient energy solutions.[1] In this context, Proton Exchange Membrane Fuel Cell (PEMFC) has gained significant attention as a favored energy conversion technology, offering numerous advantages such as high energy conversion efficiency, low carbon emissions, and silent operation.[2] However, PEMFC faces a range of challenges for commercialization, including limited lifespan and high costs.[3] Prolonged operation of PEMFC leads to component aging and damage, resulting in a decline in fuel cell performance.[4] To achieve reliable operation and extend their lifespan, effective prognostics and health management (PHM) strategies are imperative.[5]

PHM is a monitoring, diagnostic, and predictive technology aimed at enhancing equipment and system reliability and maintenance efficiency.[6] Among these strategies, accurate prediction plays a critical role.[7] In the field of fuel cells, prediction methods can be classified into model-based methods and data-driven methods.[8] Nevertheless, most mechanistic models of fuel cells struggle to adapt to dynamic conditions, prompting the increasing popularity of data-driven methods.[9] Data-driven methods employ historical data and statistical techniques to establish models and predict system behavior. By analyzing and learning patterns and trends within the data, these methods facilitate prediction and decision-making.[10] Commonly used algorithms include Artificial Neural Network (ANN), Support Vector Machine (SVM), and Recurrent Neural Network (RNN).[11,12] Among these, RNN has found widespread applications due to its potent memory capabilities, with Long Short-Term Memory (LSTM) being the most commonly employed.[13,14]

However, RNN is susceptible to issues of gradient vanishing and exploding, affecting the parameter updates during backpropagation and model convergence. Furthermore, RNNs may encounter information leakage during prediction, as they rely on historical observations to update internal states, potentially revealing sensitive information.[15] These limitations have prompted researchers to explore deep learning techniques, such as convolutional neural network (CNN), for time series prediction.[16] For instance, Xie et al.[17] combined CNN with LSTM to predict vehicle trajectories, achieving improved accuracy through spatial dilation. Lu et al.[18] proposed a CNN model with bidirectional LSTM and attention mechanism for predicting stock closing prices, achieving enhanced performance. Zhang et al.[19] proposed a hybrid prediction model that combines temporal convolutional network (TCN) for feature extraction and transfer learning. TCN is employed to extract features from raw voltage data, and then Kernel Principal Component Analysis (KPCA) is applied to select HI. Following this, an LSTM with transfer

[1]Hubei Key Laboratory of Advanced Technology for Automotive Components, Wuhan University of Technology, Wuhan 430070, China
[2]Guangdong Hydrogen Energy Institute of Wuhan University of Technology, Foshan 528000, China
[3]Hubei Provincial Key Laboratory of Fuel Cells, Wuhan 430070, China
[4]School of Automation, Wuhan University of Technology, Wuhan 430070, China
[5]The Department of Electrical Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden
[6]National Engineering Research Center of Electric Vehicles, School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China
[7]Lead contact
*Correspondence: zhuwenchao@whut.edu.cn
https://doi.org/10.1016/j.isci.2024.110979

learning capabilities is utilized to estimate the RUL. While CNN is being applied to fuel cell aging prediction, most studies used CNN for feature extraction to support other RNN-based models, with limited analysis of the predictive characteristics of CNN.[20]

After selecting the prediction method, researchers determine the health indicator (HI) that characterizes the degradation behavior of the fuel cell.[21] HI is used to assess the state and performance of the fuel cell system, providing information about its aging degree and remaining useful life (RUL).[22] Many studies analyze two datasets obtained under steady-state and quasi-dynamic conditions, using the output voltage as the HI to assess future performance and RUL of the fuel cell.[14–16] Additionally, some research also extracts the HI based on the output voltage. For instance, by combining the output voltage with a semi-empirical voltage degradation model under steady-state conditions, Liu et al.[23] extracted aging indicators associated with total reaction resistance and limiting current, which were subsequently utilized as HI for RUL estimation. Wang et al.[24] proposed a method based on the Hilbert-Huang Transform to extract an HI under dynamic operating conditions. In this method, the empirical mode decomposition (EMD) is used for decomposition within the Hilbert-Huang Transform, and a frequency threshold is set to filter the HI. Chen et al.[25] proposed a hybrid model combining EMD, Gray Relational Analysis (GRA), and RNN for predicting the RUL of lithium-ion batteries. EMD is used for data denoising, and the processed data is then evaluated for RUL.

The extraction of HI under dynamic conditions with frequent load changes poses numerous challenges. Firstly, under dynamic loads, the output voltage exhibits transient fluctuations during load transitions. Consequently, the output voltage cannot directly reflect the aging trend of the stack due to load changes. Secondly, some parameters in the semi-empirical voltage degradation model require identification based on polarization curve data during the initial stage. Polarization curves describe the static characteristics of fuel cells and are not suitable for dynamic loads. Therefore, when extracting HI under dynamic loads, certain prerequisites are established. For instance, Wang et al.[26] utilized an electrochemical mechanism model to extract HI and employed multiple cycles of data to identify an HI value, thereby reducing identification distortion through data simplification. Similarly, Xie et al.[27] extracted HI from voltage data below 70A within each 24-h measurement window. Ma et al.[28] categorized the data under the same current level for dynamic loads, applied aging parameters to a mechanistic model at specific current levels to obtain virtual steady-state voltages, and regarded them as HI. The above-mentioned efforts demonstrate that HI can only reflect the general aging status of fuel cells but may not effectively determine the specific performance of fuel cells in particular instances. Thus, under dynamic loads, the output voltage remains the most direct indicator of fuel cell performance, while the extracted HI can be focused on determining the end-of-life (EoL) point to predict the RUL.

Research gap and contributions

In the field of time series prediction, RNN has consistently been the preferred choice for most deep learning practitioners. However, CNN has shown superiority over typical recurrent networks such as LSTM on various tasks and datasets, demonstrating longer effective memory.[15,29] Furthermore, many existing data-driven methods focus on prediction accuracy but overlook prediction reliability.[27] Recent research findings have demonstrated that temporal convolutional network (TCN), a CNN variant with causal convolutions and residual connections, effectively eliminates information leakage from the future to the past and addresses the issues of gradient vanishing and exploding.[15] However, the challenge lies in the need to set appropriate hyperparameters for individual TCNs, which has hindered its widespread adoption in time series prediction tasks.

Therefore, this study introduces a predicting framework under dynamic load. This framework proposes a CPE algorithm combining complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), power spectral density (PSD), and energy analysis (EA) to process hierarchical dynamic data and generate a simple HI that can reflect the EoL time point. In terms of prediction methodology, an adaptive Bayesian-optimized TCN (AB-TCN) algorithm is proposed, which combines Bayesian optimization theory with TCN to tune appropriate model hyperparameters. Furthermore, AB-TCN, trained with effective parameters, achieves more accurate RUL predictions and demonstrates adaptability across multiple datasets. The contributions of this study are as follows.

(1) AB-TCN exhibits strong data adaptability and delivers accurate prediction performance even with limited training data.
(2) Utilizing CPE for HI extraction accurately characterizes degradation trend and EoL time point.
(3) The adoption of effective feature training can effectively enhance prediction accuracy.

The rest of this article is as follows. Section results outlines the data sources and the process of HI extraction. Section discussion validates the estimated RUL results and assesses the adaptability of AB-TCN. The conclusion is elaborated in Section conclusion. Section limitations of the study discusses the study's limitations and potential directions for future research. Section STAR Methods introduces the fuel cell prediction management framework.

RESULTS

The proposed fuel cell prognostics framework based on AB-TCN is illustrated in Figure 1. First, the dynamic variable load is categorized based on current levels to explore the relationship between isolated quasi-dynamic load voltage data and the original dataset. The quasi-dynamic load data are decomposed using CPE to extract the HI representing the reaction trend. Subsequently, the AB-TCN is employed for RUL prediction. Moreover, the adaptability of AB-TCN is assessed based on two different types of datasets.

The objective of prognostics is to predict the RUL. However, there is no unified standard for the failure threshold (FT) under dynamic loads.[5] In certain cases, different FT definitions lead to significantly different RUL estimations.[30] For fuel cells used in different applications, a single FT may not be the most optimal approach. Figure 2 illustrates the definitions of relevant terms related to RUL.

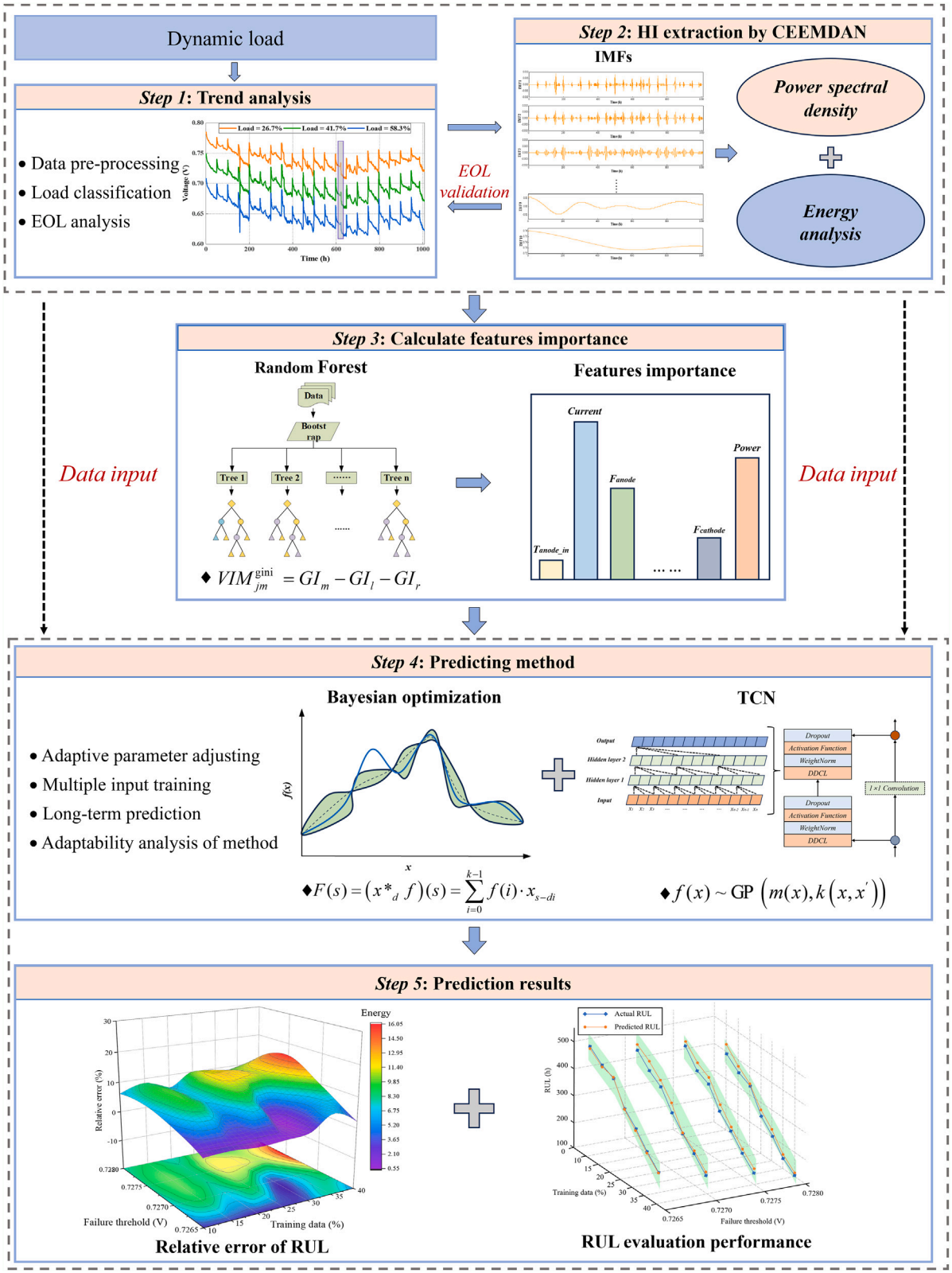


Figure 1. Hybrid Forecasting framework

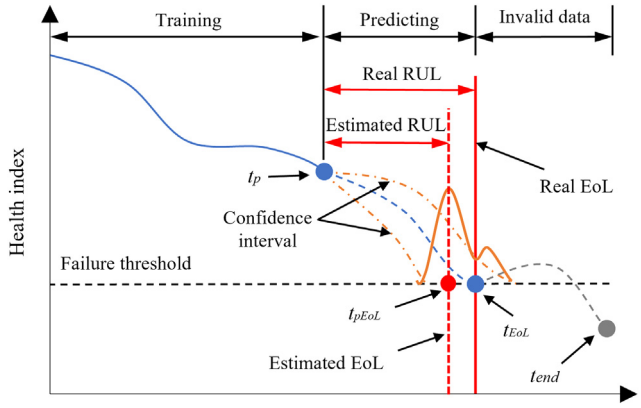


Figure 2. The remaining useful life estimation

As shown in Figure 2, the predicted trend is further utilized to estimate the RUL. The prediction starting point corresponds to time t_p. Given the FT of the HI, the intersection time between the FT and each predicted trend represents the EoL. The estimated RUL calculated at time t_p for the prediction is:

$$RUL_{estimate} = t_{pEOL} - t_p \quad \text{(Equation 1)}$$

In this process, multiple predictions are made at the expected starting point t_p, and the peak of the prediction distribution is selected as the estimated EoL. t_{pEoL} represents the estimated time of the EoL. $RUL_{estimate}$ indicates the estimated RUL.

Figure 3 illustrates the predictive scheme adopted in this study. The predictive scheme employs an iterative prediction approach. Taking the initialization of predictions with 20% of the training set as an example, the first training process involves modeling with 20% of the data. In the initial prediction, AB-TCN forecasts future data points at a distance of 20% *L with a step size of m. Subsequently, the second prediction is made based on actual data at 20% *L + m, continuing to forecast future data points with a step size of m. This iterative prediction process is repeated multiple times, ultimately generating all prediction results.

Aging test under dynamic condition

This study utilized dynamic load cycling (FC-DLC) experimental data specific to commercial PEMFC for vehicles.[31] Table 1 presents detailed information about one FC-DLC cycle of the fuel cell, involving nine different load currents ranging from 0% to 100%.[26]

As shown in Figure 4, each cycle consists of 35 testing steps, with a total duration of 1181 s per cycle. Figure 5 illustrates the overall data of voltage variation with current loads. However, from the total data, it is challenging to clearly discern the aging trend of the PEMFC, making it unsuitable as a dataset for studying the RUL.

End-of-life validation

Under dynamic load operating conditions, the voltage response process of the PEMFC exhibits transient phenomena, leading to significant fluctuations in the raw data.[32] Additionally, operating the fuel cell in the Ohmic loss region represents high efficiency.[26] Taking all factors into

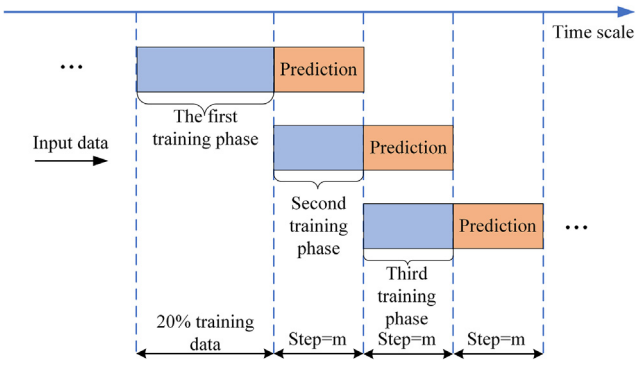


Figure 3. Iterative prediction process diagram

Table 1. Detailed information about FC-DLC

Load (%)	0	5	12.5	26.7	29.2	41.7	58.3	83.3	100
Duration (s)	81	409	52	140	100	162	143	50	44

consideration, three loads are selected for classification near the Ohmic loss region, with the longest working time (Load = 26.7%, 41.7%, 58.3%).[33] The acquired quasi-dynamic load data are smoothed using a moving average method and uniformly sampled at 3076 points with equal time intervals. The preprocessed quasi-dynamic load data are shown in Figure 6.

In the purple-shaded region, three quasi-dynamic load data experience a voltage drop phenomenon around 622 h and reach their minimum values around 640 h. Wang et al.[26] investigated the same dataset, and they identified system dynamics using an electrochemical mechanism model and extracted HI. The HI also reached their highest or lowest values around 640 h. This indicates that analyzing quasi-dynamic load data can provide insights into the RUL of dynamic variable loads.

Due to the extensive amount of testing data in this dynamic load, it is not feasible to clearly present the prediction results of the algorithm.[31] Therefore, the analysis is performed on the first 16 FC-DLC data, which is defined as dataset 1. As the three quasi-dynamic loads exhibit similar variation trends, the load of 26.7% is selected as dataset 2 for further analysis.

Extraction result of health indicators

To further extract HI from dataset 2, this section would utilize CPE for decomposition. To reduce uncertainty during the decomposition process, the original data are subjected to 20 repetitions of decomposition, and the average values are taken. The resulting IMF sequences are illustrated in Figure 7.

After CPE decomposition, ten IMF sequences are obtained. Each sequence represents the aging condition at different time scales and exhibits distinct trends and ranges. The high-frequency components of these IMF sequences are believed to have a potential relationship with stack operating conditions.[1] To further analyze the IMF sequences that characterize irreversible aging trends, the PSD and EA are utilized to calculate the contribution of each IMF to the overall data. By applying specific criteria, the aging trend can be identified and filtered out.

Figure 8 shows the main flowchart of the CPE for extracting HI. After decomposing dataset 2, the resulting IMF sequences are subjected to collaborative screening using EA and PSD. IMF sequences with an EA value less than 0.1 are removed. When using PSD for screening, IMF sequences with a frequency f greater than 5×10^{-5} Hz and a PSD value greater than 0.1 are also removed. Table 2 shows the EA values of each IMF as a percentage of the original data.

After the EA calculation, only IMF10 has an energy ratio greater than 0.05. Additionally, the maximum value of the PSD for IMF10 is 21.87, which is much greater than 0.1. Therefore, IMF10 is designated as the HI trend of dataset 2. Through the analysis in Section 3.2, it is found that the extracted quasi-dynamic load (dataset 2) effectively reflects the aging trend of the dynamic variable load. The trend of the HI extracted through CPE is shown in Figure 9. Dataset 2 reflects a certain performance degradation trend of PEMFC, but it also includes voltage recovery phenomena caused by start-stop operations and voltage fluctuations during operation. These transient effects, induced by human operations and changes in operating conditions, do not directly reflect the long-term aging state of the fuel cell. As marked in red, the HI accurately identifies the EoL point. Therefore, the study of RUL for dynamic variable loads can be simplified by analyzing the changes in the HI.

Furthermore, the voltage drop phenomenon observed at 622h is different from the voltage drop after start-stop operations, and it reaches the lowest voltage value at 640h. Therefore, in this study, the HI at 622h is set as the time warning line for dynamic variable load, and the actual voltage value (0.7265V) at 622h is defined as the failure threshold.

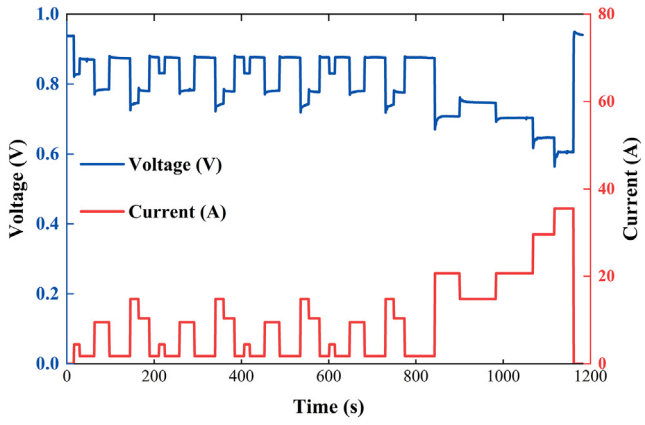


Figure 4. Current load and voltage values for a single cycle

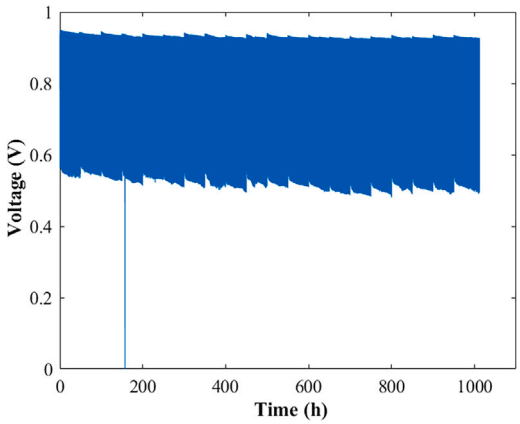


Figure 5. Total dynamic voltage for all cycles

DISCUSSION

Evaluation criteria of prediction performance

To determine the predicting performance, this article evaluates and compares the forecasting performance of each method based on multiple indicators: root-mean-square Error (RMSE), mean absolute percentage error (MAPE), coefficient of determination (R^2), and relative error (RE).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad \text{(Equation 2)}$$

$$MAPE = \frac{100}{N}\cdot\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad \text{(Equation 3)}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2} \quad \text{(Equation 4)}$$

$$RE = \frac{|t_{RUL} - t_{\overline{RUL}}|}{t_{RUL}}\cdot 100\% \quad \text{(Equation 5)}$$

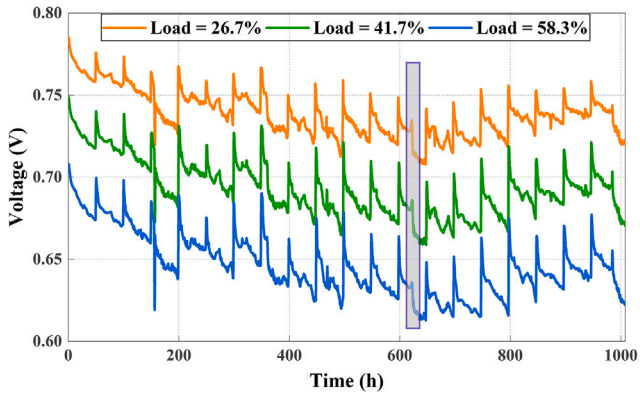


Figure 6. The quasi-dynamic load

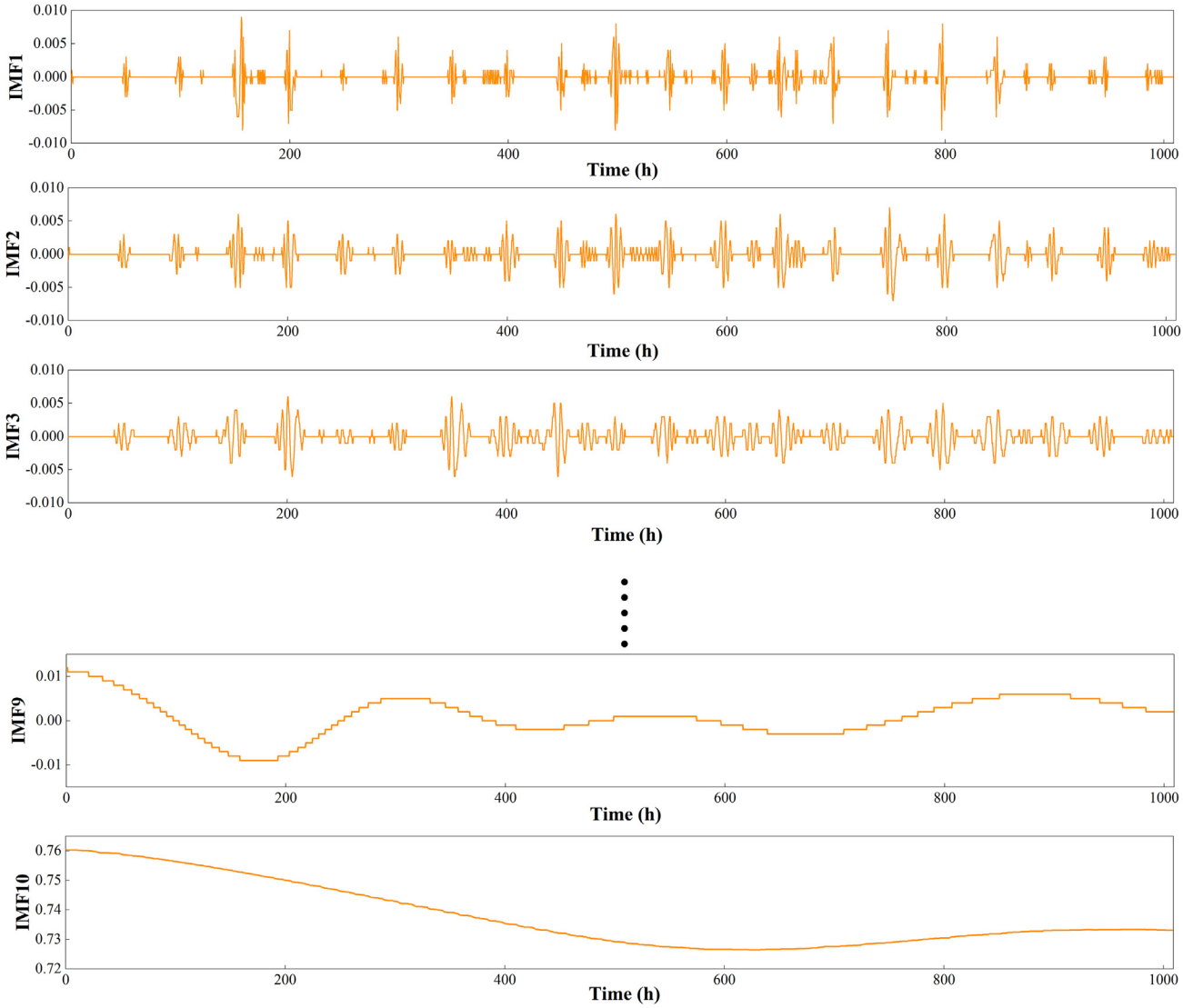


Figure 7. The results of IMF sequences

where N is the number of test data, y_i is the real value, and $\hat{y}_i$ is the predicted value, t_{RUL} is the real RUL at time t and $t_{\overline{RUL}}$ is the estimated RUL at time t.

The trust area (TA) establishes a range for a predicted value with a specific confidence level, within which there is a probability of encompassing the actual unknown value. The formula is as follows:

$$RUL_{estimate} - t_{pEOL} \cdot \alpha_{low} \leq TA \leq RUL_{estimate} + t_{pEOL} \cdot \alpha_{up} \quad \text{(Equation 6)}$$

where α_{low} and α_{up} are the lower and upper limits of the TA, respectively. When the predicted RUL has a bias, a smaller value compared to the actual RUL indicates a more sufficient time for control and maintenance. Therefore, upper and lower bounds are set at 10% and 5%, respectively. If the predicted RUL falls within the trust area TA, it is considered to be a reliable prognosis.

Feature contribution analysis

In this section, Random Forest is employed to calculate the importance of different datasets separately, and the selected effective features are input into AB-TCN for model training. To mitigate randomness in the process of assessing contributions under quasi-dynamic loads, calculations are performed on the raw data of dataset 2(I3), I5, and I6, respectively. Voltage is set as the label, and all feature parameters are input into a random forest to compute their importance scores relative to the label. The importance score for each feature is presented in Table 3.

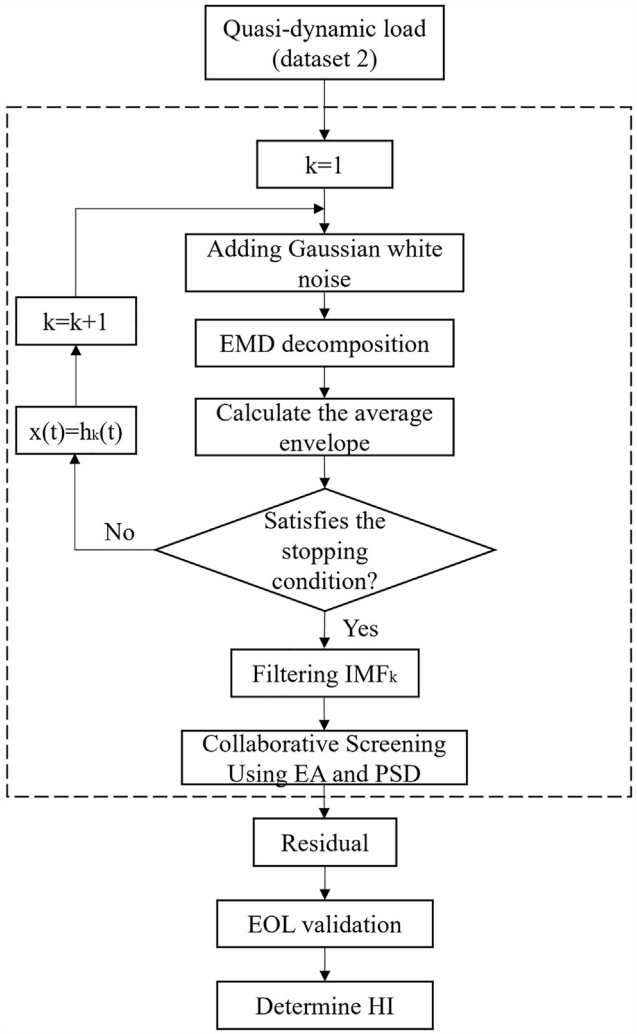


Figure 8. Flowchart of CPE

According to the analysis, features with scores greater than 0.01 are considered effective features. For dataset 1, the effective features are Current (0.04) and Power (0.95). For dataset 2, the effective feature is Power (0.99). As for the HI, the effective feature is Power (0.99). It is evident that the effective features for dataset 1 differ from those under quasi-dynamic loads (I3, I5, I6).

This distinction arises from the fact that dataset 1 is obtained under a continuously varying current, whereas dataset 2, extracted during constant operating condition classification, contains time intervals within the same current level. Consequently, these two datasets exhibit significant structural differences. Additionally, Zhu et al.,[4] when analyzing feature importance in other quasi-dynamics similar to dataset 2, observed lower importance of current. This may be attributed to the fact that, under pseudo-dynamic conditions, the current under constant load settings fluctuate only within the specified numerical range, rendering it insufficient to qualify as a high-contributing feature influencing voltage. Furthermore, the feature contributions under the three loads of I3, I5, and I6 are essentially consistent, providing additional confirmation of the accuracy of power as an effective feature under quasi-dynamic loads.

Table 2. Energy ratio values

	IMF1	IMF2	IMF3	IMF4	IMF5
EA	2.71E-6	3.84E-6	4.79E-6	4.08E-6	1.21E-5
	IMF6	IMF7	IMF8	IMF9	IMF10
EA	2.23E-5	3.81E-5	2.67E-5	3.93E-5	0.99

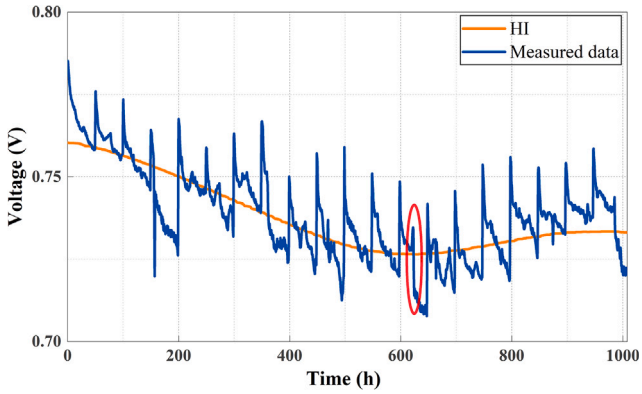


Figure 9. Dataset 2 and HI

Long-term prediction

Remaining useful life estimation results

Based on the HI obtained from Section 3.3, this section will validate the RUL estimation performance of AB-TCN through long-term predicting, with a step size of 100 steps. Additionally, Section 3.3 established 622 h as the actual EoL time point for the fuel cell. Considering that the data from 0 to 622 h roughly accounts for about 60% of the HI, the training set proportions of 10%, 15%, 20%, 25%, 30%, 35%, and 40% are set for model training.

During various prediction tasks using AB-TCN, Bayesian optimization is applied during the training phase. We conducted 200 iterations for optimization, which took only 37.5s. This approach not only significantly saves time compared to manual parameter tuning but also ensures the accuracy of parameters. Table 4 presents the optimal hyperparameters obtained through this optimization process.

In fuel cell prognostics, a slight deviation in the estimated RUL, which is slightly smaller than the actual RUL, is generally considered more acceptable.[34] Conversely, an overestimation of RUL may lead to delayed maintenance planning and potentially premature occurrence of critical failures. Therefore, the RUL estimation by AB-TCN is deemed reasonable. In contrast, the RNN series of neural networks (BIGRU, BIGRU-Att) struggle to accurately predict voltage. Both BIGRU and BIGRU-Att consistently produce predicted values that exceed the FT value, plateauing around 460h, making it impossible to calculate the estimated RUL (Figure 10). This limitation is likely attributed to the limited size of the training set, hindering BIGRU and BIGRU-Att from forming effective models. Despite the improvement observed in the 40% training set, they still failed to reach the end-of-life point during the prediction process.

The results demonstrate that AB-TCN can accurately predict multiple future voltage values under different training set proportions. Moreover, the estimated RUL is consistently smaller than the true RUL, ensuring sufficient time to address the faults.

Table 3. Importance score

	Importance value				
Feature	dataset 1	dataset 2(I3)	I5	I6	HI
Current	**4.39e-02**	9.72e-05	4.37e-05	2.37e-05	1.23e-05
Power	**9.52e-01**	**9.95e-01**	**9.99e-01**	**9.99e-01**	**9.99e-01**
Pressure anode inlet	1.48e-06	2.56e-03	6.78e-06	4.88e-04	9.38e-06
Pressure anode outlet	3.56e-06	1.80e-03	7.35e-06	2.02e-04	2.92e-05
Pressure cathode inlet	1.08e-05	2.92e-05	4.37e-06	1.70e-05	4.54e-05
Pressure cathode outlet	9.22e-06	4.86e-05	5.96e-06	3.68e-06	5.70e-05
Temp anode inlet	1.74e-05	8.17e-05	1.09e-05	7.91e-06	2.51e-04
Temp anode outlet	1.79e-05	6.03e-05	1.16e-05	1.13e-05	1.09e-05
Temp cathode inlet	1.18e-05	7.88e-05	1.17e-05	7.08e-06	1.82e-05
Temp cathode outlet	5.47e-05	6.67e-05	1.29e-05	1.02e-05	3.56e-05
Temp anode dewpoint water	2.30e-05	7.76e-05	1.15e-05	2.98e-05	2.05e-05
Temp cathode dewpoint water	2.29e-05	5.82e-05	1.17e-05	1.90e-05	1.37e-05
Total anode stack flow	3.27e-05	1.11e-04	1.12e-05	7.42e-06	5.08e-05
Total cathode stack flow	1.04e-03	4.02e-06	2.41e-07	3.80e-07	6.11e-06
Temp anode endplate	2.69e-03	1.90e-06	1.17e-06	0.00	2.12e-05

Table 4. Parameters after Bayesian optimization

	ksize	*levels*	*nhid*	*dropout*
Optimal value	5	7	75	0.21

To appropriately evaluate the RUL estimation performance of AB-TCN under different failure thresholds, three additional failure thresholds are set, ranging from 0.7265V to 0.7280V with an interval of 0.0005V. Figure 11 displays the RUL evaluation performance under different thresholds. In all test points, the predicted RUL falls within the trust area, indicating that even with a small training set, the result of RUL predictions is considerable. Furthermore, Figure 12 visually illustrates the variation of RE in long-term predictions across different dimensions. It is evident that regardless of changes in training set length or feature types (FT), the relative error of RUL consistently remains stable. The average of RE is only 6.825%.

Long-term prediction under different step lengths

To further validate the significance of effective features in model training, this section will analyze the prediction results under different step lengths. The comparison is conducted with the same number of model training iterations to ensure fairness.

Table 5 presents comparative results for step lengths of 10 and 50. AB-TCN denotes a model trained with single-input voltage, while AB-TCN(m) signifies a model trained with both voltage and power as inputs.

When the step length is set to 10, AB-TCN(m) demonstrates slightly lower predictive performance than AB-TCN for the case with a 30% training dataset. In contrast, AB-TCN(m) consistently maintains higher accuracy than AB-TCN in other scenarios. Additionally, both AB-TCN and AB-TCN(m) exhibit a common trend in prediction accuracy: an initial increase followed by a decrease. The highest errors for both models occur when the training set is at 20%. As the training set size continues to increase, the prediction errors decrease.

These results indicate that incorporating power as an effective feature for model training significantly enhances the predictive performance of AB-TCN. To further elaborate on the benefits of effective features in model training, the subsequent sections will comprehensively analyze the predictive performance of AB-TCN(m) from multiple perspectives.

Adaptability analysis of adaptive Bayesian-temporal convolutional network

In this section, the predictive performance of AB-TCN will be further validated through single-step forecasting. We will analyze the adaptability of the AB-TCN model to the data from multiple perspectives: *a.* Different types of datasets (dataset 1 and dataset 2). *b.* Different training set lengths. *c.* Model training with and without other effective features. *d.* Comparison with other algorithms. Specifically, the following sections primarily focus on the performance analysis under dynamic conditions (dataset 1), while the performance analysis under pseudo-dynamic conditions (dataset 2) is addressed in method comparison under dataset 2.

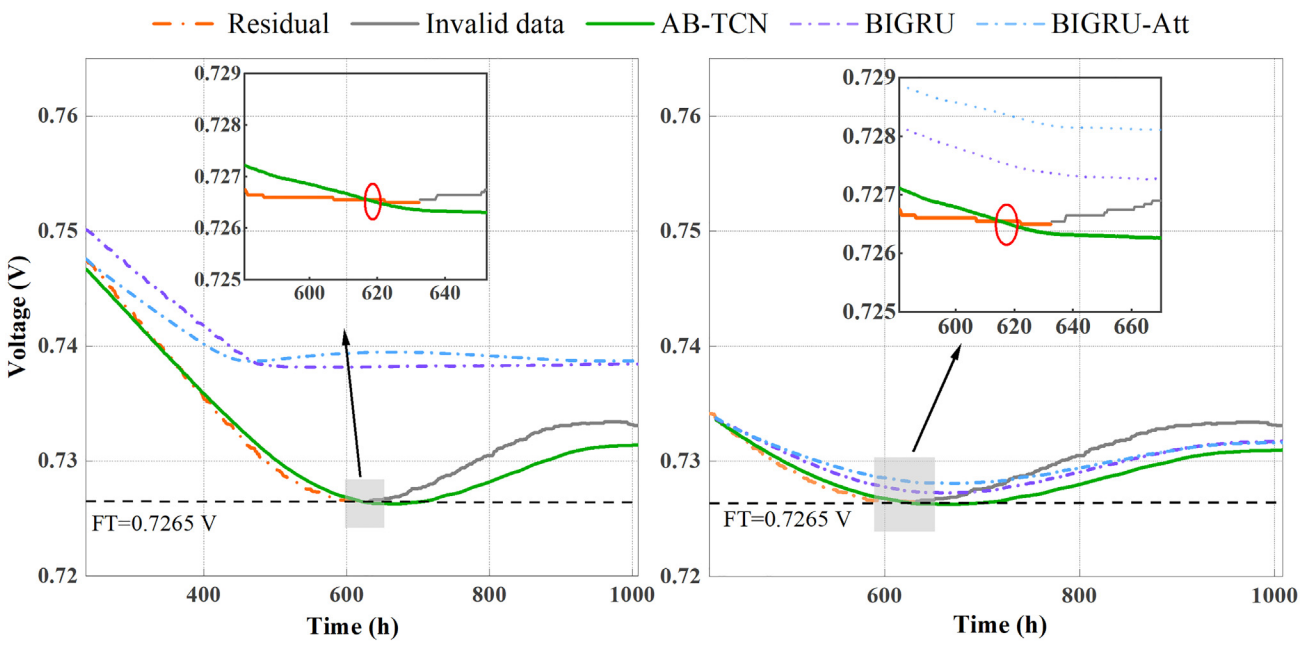


Figure 10. The long-term prediction results (20% training data and 40% training data)

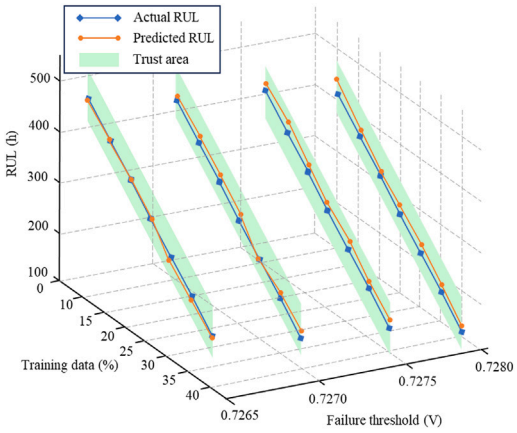


Figure 11. RUL evaluation performance

Figure 13 illustrates the single-step prediction results of AB-TCN under dataset 1 with 50% training data. AB-TCN represents only voltage as the training input, and AB-TCN(m) denotes that all effective features are included in the model training.

Compared to the regular TCN, the AB-TCN achieved a reduction in RMSE of 30.7% and a decrease in MAPE of 19.4%. Moreover, the AB-TCN(m) demonstrated even more remarkable improvements over the AB-TCN. Specifically, the AB-TCN(m) exhibited an increase in RMSE by 59.1% and an elevation in MAPE by 53.6% compared to the AB-TCN model.

As shown in enlarged view of Figure 13, the TCN (orange curve) exhibits noticeable fluctuations during the prediction process, possibly due to insufficiently optimized model parameters. AB-TCN (purple curve) demonstrates lower and fewer amplitude fluctuations compared to TCN. Notably, AB-TCN(m) accurately predicts the transient voltage fluctuations at the load switching inflection point. During a stable load operation, both AB-TCN and TCN show less stable prediction performance, while AB-TCN(m) overcomes this drawback. This is attributed to the inclusion of relevant features associated with voltage in the neural network training, enhancing the model's robustness in the prediction process.

Table 6 compares the prediction performance of AB-TCN and other algorithms under different training lengths to analyze the impact on model performance.

In the comparison results for dataset 1, AB-TCN(m) achieves the best performance across all training lengths, followed by AB-TCN. Moreover, AB-TCN(m) exhibits smaller fluctuations in RMSE and MAPE, with ranges of only 2.8%–8.3% and 4.7%–30.6%, respectively, using the 50% training set as the baseline. Notably, the R^2 of AB-TCN(m) remains consistently at 0.998. For dataset 2 validation, AB-TCN(m) continues to demonstrate the best predictive performance. With a 50% training set, AB-TCN(m) achieves the lowest RMSE and MAPE, reaching only 0.0013 and 0.107, respectively. Additionally, except for R^2 of 0.892 at the 30% training set, the R^2 of AB-TCN(m) is not below 0.920 for other cases.

These results demonstrate that AB-TCN can achieve a more stable model with a smaller training dataset. Incorporating effective features during the training process allows the AB-TCN model to better adapt to the target data, resulting in more accurate predictions. Notably, the underperformance of BILSTM-Att in dataset 1 improves significantly in dataset 2 (method comparison under dataset 2). This discrepancy is attributed to BILSTM-Att's inability to cope with frequent switches in dynamic load data, leading to decreased predictive accuracy.

To further validate the superiority of AB-TCN, this section compares it with other methods from past literature that also use the same dynamic dataset. Before the comparison, the dataset is standardized according to the standards used in the selected literature to eliminate potential biases caused by inconsistent experimental setups. Table 7 lists the error results of each method. Zuo et al.[31] proposed a GRU

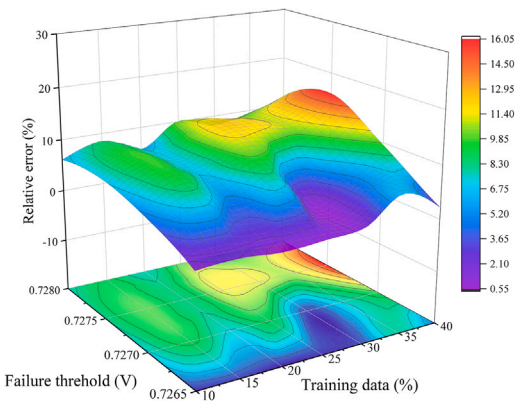


Figure 12. Relative error of RUL

Table 5. Predictions with the different step sizes

Train data (%)	Metric	AB-TCN	AB-TCN(m)
Step = 10			
10	RMSE	0.00090	**0.00059**
	MAPE	0.1090	**0.0731**
20	RMSE	0.00146	**0.00126**
	MAPE	0.1887	**0.1558**
30	RMSE	**0.00078**	0.00080
	MAPE	**0.0843**	0.0985
40	RMSE	0.00075	**0.00059**
	MAPE	0.0880	**0.0631**
Step = 50			
10	RMSE	0.00100	**0.00079**
	MAPE	0.1032	**0.0920**
20	RMSE	0.00324	**0.00246**
	MAPE	0.4068	**0.3026**
30	RMSE	0.00281	**0.00150**
	MAPE	0.3516	**0.1613**
40	RMSE	0.00129	**0.00106**
	MAPE	0.1434	**0.1275**

network with the attention mechanism to predict the voltage variation trend of fuel cells. He et al.[35] proposed a method combining auto-encoder (AE) and LSTM, and compared it with the Polarization model-LSTM. Compared to AE-LSTM, AB-TCN achieves similar performance in the MAPE. In terms of RMSE, AB-TCN reduces the error by 6.25% compared to AE-LSTM. In general, the error index of AB-TCN is smaller than that of other methods, which represents a higher prediction accuracy.

Conclusion

This article introduces a prognostic framework for fuel cells under dynamic variable loads. CPE is employed to extract a straightforward HI from the graded dynamic data. The utilized AB-TCN algorithm in the prediction process demonstrates robust data adaptability and accurate RUL estimation. To validate the proposed prognostic framework and the advantages of AB-TCN, extensive performance studies are conducted, leading to the following conclusions.

(1) By utilizing Bayesian theory to optimize model parameters, AB-TCN can adapt to different data types and prediction modes. Moreover, the predictive performance of AB-TCN is minimally affected by the length of the training dataset.
(2) Employing CPE for dynamic load data decomposition to extract HI, the obtained HI accurately reflects the genuine EoL time point of the original data. The mean relative error of RUL is only 6.825%.
(3) effective feature training models is validated in the adaptability analysis. In comparison to AB-TCN, AB-TCN(m) exhibits RMSE and MAPE increases of 10.3%–43.3% and 34.6%–49.5% in dataset 1, respectively. In dataset 2, RMSE and MAPE rise by 12.5%–35.0% and 19.0%–60%, respectively.

Limitations of the study

The proposed prognostic framework effectively determines the EoL time point, and the adopted prediction methodology yields predictions with high reliability. However, degradation behavior within fuel cell systems is exceedingly intricate. While accurately indicating the endpoint of the lifetime, HI should also be capable of characterizing major failures occurring during PEMFC operation. In future work, we will devote efforts to proposing HI that better reflects the aging mechanisms of PEMFC through model-based methods.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Wenchao Zhu (zhuwenchao@whut.edu.cn).

Materials availability

This study did not generate new unique reagents.

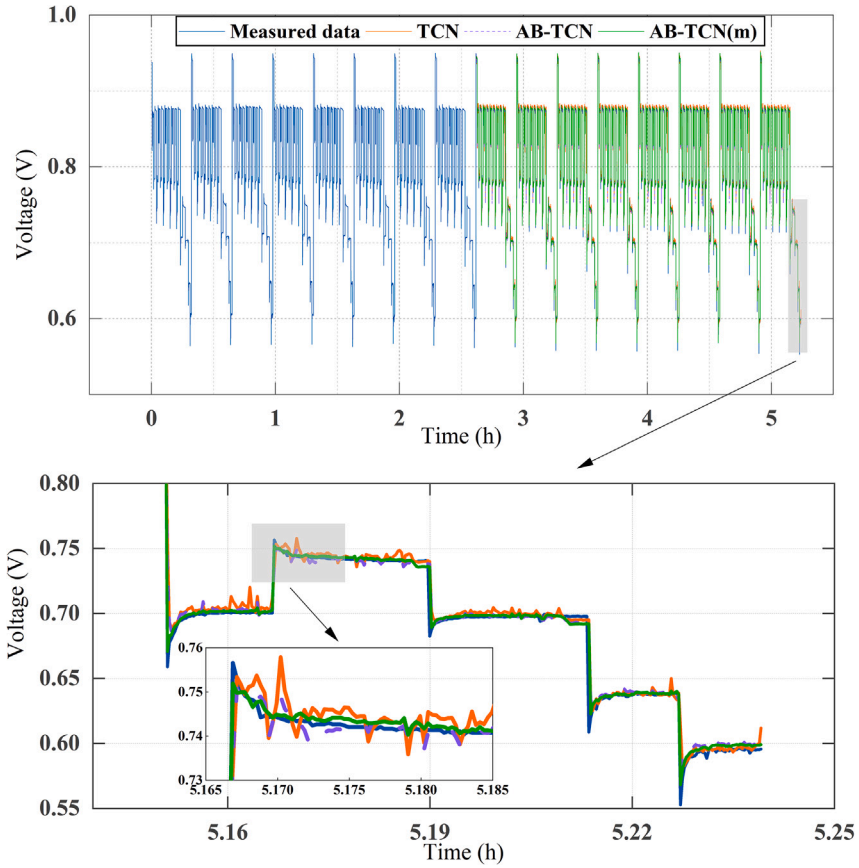


Figure 13. The single-step prediction results under dataset 1 (50% training data)

Data and code availability

- This article analyzes existing, publicly available data. These accession numbers for the datasets are listed in the key resources table.
- All original code has also been deposited at ScienceDB https://www.scidb.cn/s/nyAnQf.
- Any additional information required to reanalyze the data reported in this article is available from the lead contact upon request.

Table 6. Comparison of methods under dataset 1

Train data (%)	Metric	BILSTM	BILSTM-Att	TCN	AB-TCN	AB-TCN(m)
30	RMSE	0.0067	0.0188	0.0089	0.0057	**0.0038**
	MAPE	0.450	0.666	0.529	0.381	**0.205**
	R^2	0.994	0.952	0.989	0.996	**0.998**
40	RMSE	0.0056	0.0189	0.0076	0.0049	**0.0039**
	MAPE	0.344	0.721	0.458	0.311	**0.218**
	R^2	0.986	0.951	0.992	0.997	**0.998**
50	RMSE	0.0052	0.0188	0.0073	0.0049	**0.0036**
	MAPE	0.365	0.657	0.514	0.303	**0.193**
	R^2	0.996	0.951	0.992	0.997	**0.998**
60	RMSE	0.0047	0.0191	0.0065	0.0067	**0.0038**
	MAPE	0.296	0.685	0.485	0.364	**0.184**
	R^2	0.996	0.952	0.994	0.994	**0.998**
70	RMSE	0.0041	0.0187	0.0054	0.0039	**0.0035**
	MAPE	0.260	0.732	0.369	0.205	**0.134**
	R^2	0.998	0.953	0.996	0.998	**0.998**

Table 7. Comparison with different methods

	The training set proportion is 60%	
Method	RMSE	MAPE
AB-TCN	0.0075	0.7323
Attention-based GRU[31]	0.0155	–
AE-LSTM[35]	0.0080	0.7314
Polarization model-LSTM[35]	0.0134	1.4585

ACKNOWLEDGMENTS

This work was supported by the Key Research and Development Project of Guangdong Province (2020B0909040004), the Postdoctoral Fellowship Program of CPSF(GZC20232011) and the Fundamental Research Funds for the Central Universities(WUT: 104972024RSCbs0011).

AUTHOR CONTRIBUTIONS

Hangyu Wu: writing – original draft, data curation, software and implementation; Wenchao Zhu and Changjun Xie: writing – methodology and review and editing, conceptualization; Ruiming Zhang: funding acquisition and software; Yang Li and Yang Yang: supervision, project administration, formal analysis, and visualization; Bingxin Guo and Changzhi Li: experimental analysis and investigation.

DECLARATION OF INTERESTS

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

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- METHOD DETAILS
  - Method comparison under dataset 2
  - Random forest
  - CPE
  - Temporal convolutional network
  - Bayesian optimization
- QUANTIFICATION AND STATISTICAL ANALYSIS

Received: April 1, 2024
Revised: July 25, 2024
Accepted: September 13, 2024
Published: September 16, 2024

REFERENCES

1. Sun, B., Liu, X., Wang, J., Wei, X., Yuan, H., and Dai, H. (2023). Short-term performance degradation prediction of a commercial vehicle fuel cell system based on CNN and LSTM hybrid neural network. Int. J. Hydrogen Energy *48*, 8613–8628. https://doi.org/10.1016/j.ijhydene.2022.12.005.
2. Wan, W., Yang, Y., Li, Y., Xie, C., Song, J., Deng, Z., Tan, J., and Zhang, R. (2023). Operating conditions combination analysis method of optimal water management state for PEM fuel cell. Green Energy and Intelligent Transportation *2*, 100105. https://doi.org/10.1016/j.geits.2023.100105.
3. Wu, H., Wang, W., Li, Y., Zhu, W., Xie, C., and Gooi, H.B. (2024). Hybrid Physics-Based and Data-Driven Prognostic for PEM Fuel Cells Considering Voltage Recovery. IEEE Trans. Energy Convers. *39*, 601–612. https://doi.org/10.1109/TEC.2023.3311460.
4. Zhu, W., Guo, B., Li, Y., Yang, Y., Xie, C., Jin, J., and Gooi, H.B. (2023). Uncertainty quantification of proton-exchange-membrane fuel cells degradation prediction based on Bayesian-Gated Recurrent Unit. eTransportation *16*, 100230. https://doi.org/10.1016/j.etran.2023.100230.
5. Jouin, M., Bressel, M., Morando, S., Gouriveau, R., Hissel, D., Péra, M.-C., Zerhouni, N., Jemei, S., Hilairet, M., and Ould Bouamama, B. (2016). Estimating the end-of-life of PEM fuel cells: Guidelines and metrics. Appl. Energy *177*, 87–97. https://doi.org/10.1016/j.apenergy.2016.05.076.
6. Tian, Z., Wang, J., Al-Durra, A., Muyeen, S.M., Zhou, D., and Hua, S. (2023). A novel aging prediction method of fuel cell based on empirical mode decomposition and complexity threshold quantitative criterion. J. Power Sources *574*, 233120. https://doi.org/10.1016/j.jpowsour.2023.233120.
7. Peng, W., Wei, Z., Huang, C.-G., Feng, G., and Li, J. (2023). A Hybrid Health Prognostics Method For Proton Exchange Membrane Fuel Cells With Internal Health Recovery. IEEE Trans. Transp. Electrific. *9*, 4406–4417. https://doi.org/10.1109/TTE.2023.3243788.
8. Jin, J., Chen, Y., Xie, C., and Wu, F. (2023). Remaining useful life prediction of PEMFC based on the multi-input cycle reservoir with jump network. Int. J. Hydrogen Energy *48*, 12844–12860. https://doi.org/10.1016/j.ijhydene.2022.12.170.
9. Jin, J., Chen, Y., Xie, C., Zhu, W., and Wu, F. (2021). Remaining useful life prediction of PEMFC based on cycle reservoir with jump model. Int. J. Hydrogen Energy *46*, 40001–40013. https://doi.org/10.1016/j.ijhydene.2021.09.233.
10. Pan, R., Yang, D., Wang, Y., and Chen, Z. (2020). Performance degradation prediction of proton exchange membrane fuel cell using a hybrid prognostic approach. Int. J. Hydrogen Energy *45*, 30994–31008. https://doi.org/10.1016/j.ijhydene.2020.08.082.
11. Ma, R., Li, Z., Breaz, E., Liu, C., Bai, H., Briois, P., and Gao, F. (2019). Data-Fusion Prognostics of Proton Exchange Membrane

Fuel Cell Degradation. IEEE Trans. Ind. Appl. *55*, 4321–4331. https://doi.org/10.1109/TIA.2019.2911846.

12. Liu, J., Li, Q., Chen, W., Yan, Y., Qiu, Y., and Cao, T. (2019). Remaining useful life prediction of PEMFC based on long short-term memory recurrent neural networks. Int. J. Hydrogen Energy *44*, 5470–5480. https://doi.org/10.1016/j.ijhydene.2018.10.042.
13. Gu, X., Hou, Z., and Cai, J. (2021). Data-based flooding fault diagnosis of proton exchange membrane fuel cell systems using LSTM networks. Energy and AI *4*, 100056. https://doi.org/10.1016/j.egyai.2021.100056.
14. Li, S., Luan, W., Wang, C., Chen, Y., and Zhuang, Z. (2022). Degradation prediction of proton exchange membrane fuel cell based on Bi-LSTM-GRU and ESN fusion prognostic framework. Int. J. Hydrogen Energy *47*, 33466–33478. https://doi.org/10.1016/j.ijhydene.2022.07.230.
15. Bai, S., Kolter, J.Z., and Koltun, V. (2018). An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. Preprint at arXiv. https://doi.org/10.48550/arXiv.1803.01271.
16. Wilberforce, T., Alaswad, A., A, G.–P., Xu, Y., Ma, X., and Panchev, C. (2023). Remaining useful life prediction for proton exchange membrane fuel cells using combined convolutional neural network and recurrent neural network. Int. J. Hydrogen Energy *48*, 291–303. https://doi.org/10.1016/j.ijhydene.2022.09.207.
17. Xie, G., Shangguan, A., Fei, R., Ji, W., Ma, W., and Hei, X. (2020). Motion trajectory prediction based on a CNN-LSTM sequential model. Sci. China Inf. Sci. *63*, 212207. https://doi.org/10.1007/s11432-019-2761-y.
18. Lu, W., Li, J., Wang, J., and Qin, L. (2021). A CNN-BiLSTM-AM method for stock price prediction. Neural Comput & Applic *33*, 4741–4753. https://doi.org/10.1007/s00521-020-05532-z.
19. Zhang, Y., Ma, R., Xie, R., Feng, Z., Liang, B., and Li, Y. (2024). A Degradation Prediction Method for PEM Fuel Cell Based on Deep Temporal Feature Extraction and Transfer Learning. IEEE Trans. Transp. Electrific. *10*, 203–212. https://doi.org/10.1109/TTE.2023.3262588.
20. Zha, W., Liu, Y., Wan, Y., Luo, R., Li, D., Yang, S., and Xu, Y. (2022). Forecasting monthly gas field production based on the CNN-LSTM model. Energy *260*, 124889. https://doi.org/10.1016/j.energy.2022.124889.
21. Liu, H., Chen, J., Hissel, D., Hou, M., and Shao, Z. (2019). A multi-scale hybrid degradation index for proton exchange membrane fuel cells. J. Power Sources *437*, 226916. https://doi.org/10.1016/j.jpowsour.2019.226916.
22. Ma, R., Yang, T., Breaz, E., Li, Z., Briois, P., and Gao, F. (2018). Data-driven proton exchange membrane fuel cell degradation predication through deep learning method. Appl. Energy *231*, 102–115. https://doi.org/10.1016/j.apenergy.2018.09.111.
23. Liu, H., Chen, J., Hissel, D., and Su, H. (2019). Remaining useful life estimation for proton exchange membrane fuel cells using a hybrid method. Appl. Energy *237*, 910–919. https://doi.org/10.1016/j.apenergy.2019.01.023.
24. Wang, C., Dou, M., Li, Z., Outbib, R., Zhao, D., Zuo, J., Wang, Y., Liang, B., and Wang, P. (2023). Data-driven prognostics based on time-frequency analysis and symbolic recurrent neural network for fuel cells under dynamic load. Reliab. Eng. Syst. Saf. *233*, 109123. https://doi.org/10.1016/j.ress.2023.109123.
25. Chen, J.C., Chen, T.-L., Liu, W.-J., Cheng, C.C., and Li, M.-G. (2021). Combining empirical mode decomposition and deep recurrent neural networks for predictive maintenance of lithium-ion battery. Adv. Eng. Inf. *50*, 101405. https://doi.org/10.1016/j.aei.2021.101405.
26. Wang, C., Dou, M., Li, Z., Outbib, R., Zhao, D., and Liang, B. (2022). A fusion prognostics strategy for fuel cells operating under dynamic conditions. eTransportation *12*, 100166. https://doi.org/10.1016/j.etran.2022.100166.
27. Xie, R., Li, C., Ma, R., Xu, L., and Zhou, X. (2023). An Explainable Data-Driven Framework for Fuel Cell Aging Prediction Under Dynamic Condition. IEEE Trans. Ind. Electron. *70*, 5960–5970. https://doi.org/10.1109/TIE.2022.3199935.
28. Ma, R., Xie, R., Xu, L., Huangfu, Y., and Li, Y. (2021). A Hybrid Prognostic Method for PEMFC With Aging Parameter Prediction. IEEE Trans. Transp. Electrific. *7*, 2318–2331. https://doi.org/10.1109/TTE.2021.3075531.
29. Benaggoune, K., Yue, M., Jemei, S., and Zerhouni, N. (2022). A data-driven method for multi-step-ahead prediction and long-term prognostics of proton exchange membrane fuel cell. Appl. Energy *313*, 118835. https://doi.org/10.1016/j.apenergy.2022.118835.
30. Zhao, J., and Li, X. (2019). A review of polymer electrolyte membrane fuel cell durability for vehicular applications: Degradation modes and experimental techniques. Energy Convers. Manag. *199*, 112022. https://doi.org/10.1016/j.enconman.2019.112022.
31. Zuo, J., Lv, H., Zhou, D., Xue, Q., Jin, L., Zhou, W., Yang, D., and Zhang, C. (2021). Deep learning based prognostic framework towards proton exchange membrane fuel cell for automotive application. Appl. Energy *281*, 115937. https://doi.org/10.1016/j.apenergy.2020.115937.
32. Chen, H., Pei, P., and Song, M. (2015). Lifetime prediction and the economic lifetime of Proton Exchange Membrane fuel cells. Appl. Energy *142*, 154–163. https://doi.org/10.1016/j.apenergy.2014.12.062.
33. BenChikha, K., Kandidayeni, M., Amamou, A., Kelouwani, S., Agbossou, K., and Abdelghani, A.B.B. (2022). Fuel Cell Ageing Prediction and Remaining Useful Life Forecasting. In 2022 IEEE Vehicle Power and Propulsion Conference (VPPC), Merced, CA, USA (IEEE), pp. 1–6. https://doi.org/10.1109/VPPC55846.2022.10003313.
34. Wang, C., Li, Z., Outbib, R., Dou, M., and Zhao, D. (2022). Symbolic deep learning based prognostics for dynamic operating proton exchange membrane fuel cells. Appl. Energy *305*, 117918. https://doi.org/10.1016/j.apenergy.2021.117918.
35. He, K., Liu, Z., Sun, Y., Mao, L., and Lu, S. (2022). Degradation prediction of proton exchange membrane fuel cell using auto-encoder based health indicator and long short-term memory network. Int. J. Hydrogen Energy *47*, 35055–35067. https://doi.org/10.1016/j.ijhydene.2022.08.092.
36. Li, C., Lin, W., Wu, H., Li, Y., Zhu, W., Xie, C., Gooi, H.B., Zhao, B., and Zhang, L. (2023). Performance degradation decomposition-ensemble prediction of PEMFC using CEEMDAN and dual data-driven model. Renew. Energy *215*, 118913. https://doi.org/10.1016/j.renene.2023.118913.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
PEMFC data	Zuo et al.[30]	https://doi.org/10.1016/j.apenergy.2020.115937
Software and algorithms		
Python	Python Software Foundation	https://www.python.org/
Origin 2021	OriginLab	https://www.originlab.com/
MATLAB R2021b	MathWorks	https://ww2.mathworks.cn/?s_tid=gn_logo
Temporal convolutional network	Zhang et al.[19]	https://doi.org/10.1109/TTE.2023.3262588
Random forest	Zhu et al.[4]	https://doi.org/10.1016/j.etran.2023.100230
AB-TCN	This paper	https://www.scidb.cn/s/nyAnQf
CPE	This paper	https://www.scidb.cn/s/nyAnQf
PEMFC data (Dataset 1)	Zuo et al.[31]	https://doi.org/10.1016/j.apenergy.2020.115937
Datasets 2 and 3	This paper	https://www.scidb.cn/en/s/QR3QNj

METHOD DETAILS

Method comparison under dataset 2

Figure A1 presents the prediction results under dataset 2. AB-TCN (purple curve) shows more accurate voltage predictions and closer proximity to the true voltage values compared to TCN. AB-TCN(m) trained with power as an effective feature outperforms other neural networks, which achieves the best performance. Compared to single-input AB-TCN, AB-TCN (m) exhibits an increase of 35.0% in RMSE and 19.2% in MAPE.

In Figure A1B, it is evident that the TCN (orange curve) without Bayesian optimization exhibits the poorest prediction results, with predictions deviating from the true voltage values between 625h and 650h. Furthermore, in the time interval between 620h and 625h, AB-TCN (m) accurately predicts voltage drops and achieves stable prediction trends in a shorter time.

The results from dataset 2 demonstrate that AB-TCN has the ability to accurately predict voltage under both dynamic and constant loads. Moreover, AB-TCN, trained with multiple effective features, further enhances the ability to accurately predict transient voltage fluctuations and stabilize predictions in a shorter time.

Table. A.1. Comparison of methods under dataset 2

Train data (%)	Metric	BILSTM	BILSTM-Att	TCN	AB-TCN	AB-TCN(m)
30	RMSE	0.0044	0.0043	0.0034	0.0032	**0.0028**
	MAPE	0.437	0.313	0.293	0.268	**0.217**
	R2	0.791	0.806	0.823	0.843	**0.892**
40	RMSE	0.0029	0.0025	0.0032	0.0027	**0.0021**
	MAPE	0.295	0.204	0.261	0.207	**0.153**
	R^2	0.902	0.926	0.824	0.884	**0.936**
50	RMSE	0.0021	0.0019	0.0026	0.0020	**0.0013**
	MAPE	0.168	0.129	0.167	0.132	**0.107**
	R^2	0.951	0.958	0.924	0.956	**0.980**
60	RMSE	0.0025	0.0027	0.0026	0.0022	**0.0018**
	MAPE	0.186	0.183	0.201	0.172	**0.139**
	R^2	0.938	0.926	0.911	0.937	**0.960**

(*Continued on next page*)

Continued

Train data (%)	Metric	BILSTM	BILSTM-Att	TCN	AB-TCN	AB-TCN(m)
70	RMSE	0.0023	0.0018	0.0022	0.0020	**0.0013**
	MAPE	0.132	0.103	0.199	0.160	**0.100**
	R^2	0.916	0.954	0.917	0.938	**0.975**

Table A.1 presents a comparative analysis for different training ratios under dataset 2.

AB-TCN(m), trained using effective features, maintains remarkably high predictive accuracy in dataset 2. The predictive performance of AB-TCN(m) is significantly improved compared to AB-TCN, with RMSE, MAPE, and R^2 values increasing by 12.5%–35.0%, 19.0%–60%, and 2.5%–5.9%, respectively. At a training ratio of 50%, AB-TCN(m) achieves the lowest RMSE and MAPE indicators, reaching only 0.0013 and 0.107, respectively. Moreover, except for a case where R^2 is 0.892 with a 30% training ratio, AB-TCN(m) maintains R^2 values no less than 0.920 across other scenarios.

The results demonstrate that other algorithms fail to achieve satisfactory predictive performance with fewer training samples. Enhanced predictive capabilities are achieved with longer training sets. This enhancement is likely attributed to the model's improved ability to capture data variability and trends with more training data, thus enhancing prediction accuracy. Notably, AB-TCN(m) overcomes these limitations, consistently maintaining stability in evaluation metrics across various training set lengths. In summary, these findings validate the significance of training models with effective features.

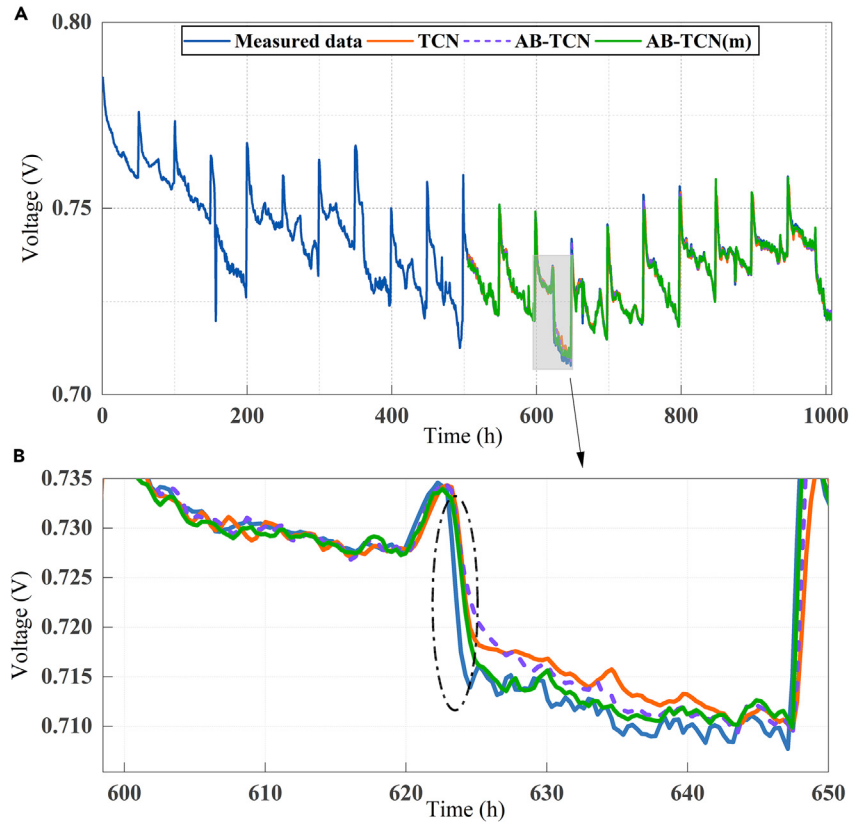


Figure A1. The single-step prediction results under dataset 2 (50% training data)

Random forest

Random forest is an ensemble learning algorithm widely used in machine learning and data mining for measuring feature importance. The main steps of random forest are outlined below.[4]

(1) Feature importance scores (*VIM*) and Gini coefficient (*GI*) are used as measures of contribution. First, denote the features as X_1, X_2, …, X_c. The average change in node impurity caused by splitting on the j-th feature in all decision trees of the random forest is represented as X_j.

$$GI_m = 1 - \sum_{k=1}^{|k|} p_{mk}^2 \quad \text{(Equation 7)}$$

where k represents the number of classes, p_{mk} is the score of the k-th class at node m. 2) The importance of feature X_j at node m can be calculated as follows:

$$VIM_{jm}^{gini} = GI_m - GI_l - GI_r \quad \text{(Equation 8)}$$

where GI_l and GI_r are the Gini indices of the corresponding nodes before and after the split.

(2) For the case where feature X_j is present in different nodes of decision tree l, and considering m belongs to the set M (m∈ M), the importance of feature X_j in the first tree i can be expressed as follows:

$$VIM_{ij}^{gini} = \sum_{m \in M} VIM_{jm}^{gini} \quad \text{(Equation 9)}$$

(3) By dividing the contribution of feature GI by the total number of features, we can obtain the contribution of feature X_j in the random forest.

$$VIM_j = \frac{\sum_{i=1}^{100} VIM_{ij}^{gini}}{\sum_{j=1}^{16}\sum_{i=1}^{100} VIM_{ij}^{gini}} \quad \text{(Equation 10)}$$

CPE

This part proposes a CPE algorithm for extracting HI under dynamic operating conditions. Based on CEEMDAN, PSD, and EA are incorporated to collaboratively screen HI. CEEMDAN is a method used for signal processing and analysis.[36] It is capable of decomposing non-stationary and non-linear signals into multiple intrinsic mode function (IMF) sequences and a residual component. In traditional EMD, due to the presence of mode mixing, the decomposed IMF may contain multiple frequency components. However, CEEMDAN addresses this issue by introducing adaptive noise, which enhances the stochasticity of the IMFs and effectively alleviates the problem of mode mixing. As a result, CEEMDAN can better separate the different frequency components of the signal.

EMD decomposition process

The EMD decomposition decomposes the signal $x(t)$ into a series of IMFs and a residual component $r(t)$. For each IMF, it can be calculated using the following equation:

$$x(t) = IMF_i + r(t) \quad \text{(Equation 11)}$$

Introduction of adaptive noise

For each IMF_i, after introducing adaptive noise, the IMF_i with noise can be calculated using the following equation:

$$IMF_i' = IMF_i + n_i(t) \quad \text{(Equation 12)}$$

where IMF_i' is the IMF with introduced noise, IMF_i is the original IMF, and $n_i(t)$ is the adaptively generated Gaussian white noise.

Stopping criteria

The stopping criteria are used to determine whether the decomposition has reached its termination condition. It can be determined by reaching a preset energy threshold or reaching the maximum number of decomposition layers N.

Adding PSD and EA calculations for each IMF to screen out low-frequency data reflecting the aging trend. PSD is a metric that describes the energy distribution of a signal in the frequency domain and can be expressed as

$$PSD(f) = \lim_{T\to\infty}[1/(2T)] * |X(f;T)|^2 \quad \text{(Equation 13)}$$

where *PSD*(*f*) represents the Power Spectral Density of the signal at frequency *f*. *X*(*f*;*T*) is the Fourier transform of the signal *x*(*i*) within the time window *T*.

The EA determines the relative energy proportions of signal components. Its equation can be expressed as follows

$$E(i) = \sum x(i)^2 \Big/ \sum x^2 \quad \text{(Equation 14)}$$

where *E*(*i*) represents the energy ratio of the *i*-th signal component, *x*(*i*) is the *i*-th signal component, and *x* is the original signal.

Temporal convolutional network

Temporal Convolutional Network (TCN) is a deep learning model designed for time-series data.[19] It uses dilated convolutions to capture longer-term temporal dependencies by expanding the receptive field. Causal convolutions prevent information leakage from the future during prediction. Residual connections mitigate information loss and vanishing gradients, aiding better information flow in the network.

The TCN is composed of a one-dimensional fully convolutional network (1-D FCN), causal convolutions, dilated convolutions, and residual modules. By stacking multiple layers of convolutional units, TCN can capture long-range dependencies within sequences, facilitating effective time series modeling and prediction. In contrast to RNNs, TCN offers advantages like parallel computation and stable gradient propagation, making it suitable for various time series tasks.

For a sequence $X = (x_1, x_2, \ldots, x_n)$ and introducing a filter *F*, the unfolded dilated convolution *F*(*s*) with dilation factor *d* at position x_n can be computed as follows:

$$F(s) = (x *_d f)(s) = \sum_{i=0}^{k-1} f(i) \cdot x_{s-di} \quad \text{(Equation 15)}$$

where $*d$ is the dilated convolution operator, *d* is the dilation factor, and *k* is the filter size. The *s*-*di* represents the causal direction in the past.

Bayesian optimization

Bayesian optimization is used to optimize four model parameters in TCN: *kernel size*, *levels*, *nhid*, and *dropout*. The main steps of Bayesian optimization are as follows.

(1) Adopt the Gaussian Process to model the underlying distribution of the objective function. The Gaussian Process regression model can be represented as follows:

$$f(x) \sim \mathcal{GP}(m(x), k(x, x')) \quad \text{(Equation 16)}$$

where *f*(*x*) is the objective function, *x* is the input variable, *m*(*x*) is the mean function of the Gaussian process, and *k*(*x*, *x*') is the covariance function of the Gaussian process.

(2) By utilizing prior information and existing observed data, we can estimate the latent distribution of the objective function. Bayesian inference is employed to obtain the posterior distribution:

$$p(f(x)|X, y) \quad \text{(Equation 17)}$$

where *X* is the observed values of the input variables, and *y* is the corresponding values of the objective function.

(3) To select the next point for evaluation, a sampling strategy is used to balance exploration and exploitation. One commonly used strategy is Expected Improvement (EI). The EI represents the expected improvement in the objective function value over the current observed minimum value *x**. It can be expressed as follows:

$$\mathrm{EI}(x) = \mathbb{E}[\max(f(x^*) - f(x), 0)] \quad \text{(Equation 18)}$$

By maximizing the Expected Improvement, the next point for evaluation can be selected. The Gaussian process model is updated based on the observed results, and the optimization process continues iteratively.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were conducted with Python 3.9 and MATLAB R2021b. We utilized a random forest algorithm to assess the significance of features, the results of which were presented in Table 3. Figures 6 and 7 displayed the long-term forecasting outcomes achieved by the AB-TCN model. Table 5 delineated the predictive performance of the AB-TCN across different time steps and training lengths. Figure 8 illustrated the forecasting performance of the AB-TCN on dataset1, with detailed predictive results tabulated in Table 6.