China Instrument and Control Society
Instrumentation
INSTRUMENTATION, Vol. 11, No. 1, March 2024 https://doi.org/10.15878/j.instr.202300155

Article

Performance Degradation Prediction of Proton Exchange Membrane Fuel Cell Based on CEEMDAN-KPCA and DA-GRU Networks

Tingwei Zhao, Juan Wang*, Jiangxuan Che, Yingjie Bian and Tianyu Chen

Shaanxi Key Laboratory of Nanomaterials and Nanotechnology, School of Mechanical and Electrical Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China

* Corresponding author email: juanwang618@126.com

Abstract: In order to improve the performance degradation prediction accuracy of proton exchange membrane fuel cell (PEMFC), a fusion prediction method (CKDG) based on adaptive noise complete ensemble empirical mode decomposition (CEEMDAN), kernel principal component analysis (KPCA) and dual attention mechanism gated recurrent unit neural network (DA-GRU) was proposed. CEEMDAN and KPCA were used to extract the input feature data sequence, reduce the influence of random factors, and capture essential feature components to reduce the model complexity. The DA-GRU network helps to learn the feature mapping relationship of data in long time series and predict the changing trend of performance degradation data more accurately. The actual aging experimental data verify the performance of the CKDG method. The results show that under the steady-state condition of 20% training data prediction, the CKDA method can reduce the root mean square error (RMSE) by 52.7% and 34.6%, respectively, compared with the traditional LSTM and GRU neural networks. Compared with the simple DA-GRU network, RMSE is reduced by 15%, and the degree of over-fitting is reduced, which has higher accuracy. It also shows excellent prediction performance under the dynamic condition data set and has good universality.

Keywords: proton exchange membrane fuel cell; dual-attention gated recurrent unit; data-driven model; time series prediction

Citation: Tingwei Zhao, Juan Wang, Jiangxuan Che, Yingjie Bian and Tianyu Chen. "Performance Degradation Prediction of Proton Exchange Membrane Fuel Cell Based on CEEMDAN-KPCA and DA-GRU Networks." *Instrumentation* 11, no. 1 (March 2024). https://doi.org/10.15878/j.instr.202300155

0 Introduction

Proton exchange membrane fuel cell (PEMFC) is considered to be one of the most promising green power sources due to its high energy conversion efficiency, high power density, and clean reaction products[1-2]. However, due to the influence of material degradation and complex operating conditions[3], the performance of PEMFC will decline and even lead to failure. Therefore, accurate prediction of PEMFC performance degradation is significant for prognostic and health management (PHM) technology and related optimal operational decisions to extend its service life and improve its performance.

At present, there are three types of PEMFC performance degradation prediction methods: model-based methods, data-driven methods, and hybrid methods. The model-based method relies on the physical modeling of the internal degradation behavior of PEMFC to establish an empirical model of PEMFC performance degradation trajectory[4-6]. Bressel M et al.introduced the aging factors of ohmic resistance and limiting current density, constructed a linear attenuation model of the fuel cell, used aging factors as health indicators, used an extended Kalman filter to estimate aging factors in real-time, and predicted the remaining service life of fuel cell[7]. Chen et al.constructed a linear empirical degradation model and applied an unscented Kalman filter to predict the remaining useful life of postal fuel cell electric vehicles[8]. However, the PEMFC system has multi-physical and multi-scale

complex characteristics, often showing nonlinear solid characteristics, and its degradation mechanism is not completely clear[9]. Currently, there is no general modeling method to describe all the degradation mechanisms of PEMFC[10]. However, the hybrid method is still more dependent on the construction of the fuel cell model in essence, and the parameters of the introduced filtering algorithm are susceptible to noise.

Compared with the model-based method, which is highly dependent on model complexity and parameter selection, the data-driven prediction method uses aging data to establish a PEMFC degradation model, which avoids the complex internal mechanism relationship of PEMFC and has greater flexibility and excellent nonlinear fitting ability[11]. Data-driven methods are also widely used in the remaining life prediction of lithium-ion batteries[12]. PEMFC has more complex operating conditions than lithium-ion batteries, and the decay process involves strong time dependence. The design and training of prediction models pay more attention to capturing dynamic characteristics in time series. For example, a recurrent neural network (RNN) contains a recursive hidden layer suitable for time series prediction processing. Various RNN architectures have been studied in the literature[13]. Zuo et al.proposed a recurrent neural network (RNN) model based on attention processing data and proved that the attention mechanism can further improve the performance of the model[14]. Ma et al. proposed a grid-based long short-term memory (G-LSTM) network, which has the potential to deal with complex spatio-temporal sequence data, but the model contains a large number of parameters that are prone to overfitting[15]. Zhang et al.proposed a Bi-directional Gated Recurrent Unit (Bi-GRU) recurrent neural network, adjusting the model's hyperparameters to obtain the optimal hyperparameter combination, and the model's universality has room for improvement[16]. The GRU network uses update gates and reset gates to save and filter the previous information. Compared with LSTM, it simplifies the iterative structure, can screen helpful information more quickly and avoid gradient explosion and gradient disappearance problems, and has more efficient performance in the prediction application of PEMFC multi-factor and multi-data input[17]. However, most of the PEMFC performance degradation prediction performance based on the GRU method is not ideal, which may be due to the following reasons: 1) Directly using nonlinear and noisy original aging data for prediction is easily affected by noise and spikes in the original data. 2) The ability to extract implicit aging features is limited, and it is difficult to accurately learn features when dealing with long-term sequence data with complex information. 3) There are over-fitting or under-fitting problems in the training process.

Because of the above problems, this paper proposes a fusion prediction method (CKDG) of CEEMDAN, KPCA, and DA-GRU. The CEEMDAN is used to decompose the characteristic data affecting the output voltage. The KPCA algorithm is used to extract the high-weight data from the component data, and the DA-GRU network is used to extract the weight selectively. The ability to completely capture long-term time-dependent relationships improves the prediction performance of PEMFC voltage decay data. The main contributions of this paper are as follows:

1) A fusion prediction method that can simplify the feature data set and accurately extracting the feature weight is proposed, which provides a new perspective for PEMFC performance degradation prediction.

2) The CKDG method is used to predict the performance degradation of PEMFC, and the feasibility and universality of the CKDG method are verified on different data sets.

3) Compared with other mainstream neural network methods, CKDG can provide more accurate prediction performance with less training data.

1 Experiment and Dataset Analysis

The CKDG fusion prediction method is validated using two aging data sets from the experimental PEMFC stack. The PEMFC aging experiment is implemented by the Federation for Fuel Cell Research (FCLAB). The PEMFC comprises an industrial membrane, a diffusion layer, and a machined port plate and consists of five single cells with an area of 100 cm^2.

The constant current of the fuel cell working under the static current load (FC1) is 70 A, the nominal current of the fuel cell working under the dynamic current load (FC2) is 70 A, and it has a 7 A oscillating current with a frequency of 5 kHz. The experimental platform and load settings are shown in Fig.1. The experiment includes two aging data sets: electrochemical impedance spectroscopy

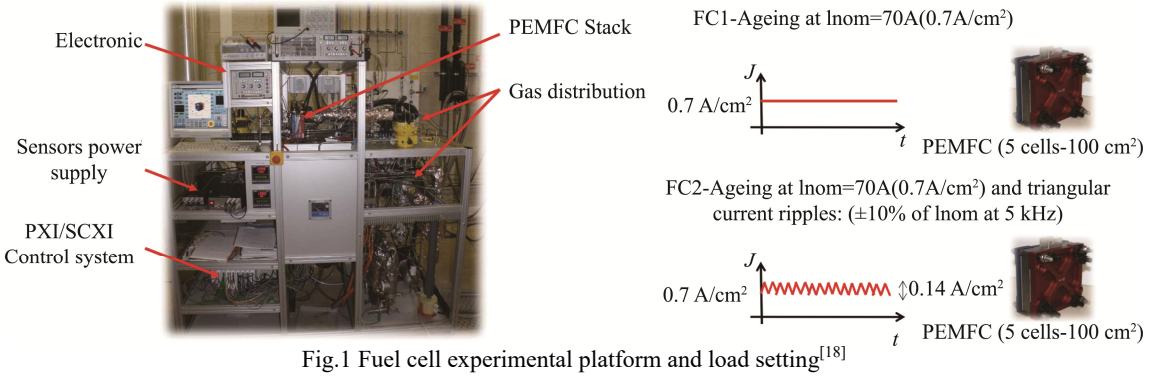

Fig.1 Fuel cell experimental platform and load setting[18]

(EIS) data and polarization data[18]. The physical parameters involved in the experiment can be measured and controlled by LabVIEW® 2014 and the self-made interface in the national instrument control system, respectively. Table 1 lists the physical parameters involved in the test. In addition, part of the voltage recovery phenomenon is caused by polarization curve tests and EIS measurements. The collected aging parameters are shown in Table 2. In this paper, the degradation trend is reflected by the change of stack voltage drop with time.

Table 1 Range of physical parameters controlled

Parameter	Control range
Cooling temperature	20℃-80℃
Cooling flow	0-10 l/min
Gas temperature	20℃-80℃
Gas humidification	0-100%RH
Air flow	0-100 l/min
H_2 flow	0-30 l/min
Gas pressure	0-2bars
Fuel Cell current	0-300A

Table 2 Aging parameters gathered during experiments

Parameter	Physical meaning
Time	Time Ageing time (h)
U1 to U5;	Single cells and stack voltage (V)
I; J	Current (A) and current density (A/cm^2)
TinAIR; ToutAIR	Inlet and outlet temperatures of Air (°C)
TinWAT; ToutWAT	Inlet and outlet temp. of cooling Water(°C)
TinH_2; ToutH_2	Inlet and outlet temperatures of H_2(°C)
PinAIR; PoutAIR	Inlet and outlet Pressure of Air (mbara)
PinH_2; PoutH_2	Inlet and outlet Pressure of H_2 (mbara)
DinH_2; DoutH_2	Inlet and outlet flow rate of H_2 (l/mn)
DWAT	Flow rate of cooling water (l/mn)
DinAIR; DoutAIR	Inlet and outlet flow rate of Air (l/mn)
HrAIRFC	Inlet Hygrometry (Air)-estimated (%)

Because the voltage data of the aging test contains certains noise and spikes, the large number of experimental data points may increase the complexity of the model and the calculation time. In order to effectively obtain the degradation trend, the Gaussian weighted moving average filter is used to smooth the original voltage data. Filter out high-frequency disturbances and retain useful low-frequency trends in the time series[19]. Then, FC1 and FC2, two new voltage decay data sets with an interval of 1 h, are obtained, denoted as FC1, as shown in Fig.2. The two new voltage aging data sequences contain 1154 and 1020 sets of voltage data, respectively. At the same time, the data integration of 23 characteristic environmental parameters of the stack voltage is divided into a series of 1h intervals, and a unified complete data set is formed with the new voltage aging data sequence. In addition, it can be seen in Fig.2 that the voltage degradation of the FC1 test is much more regular and stable than that of FC2.

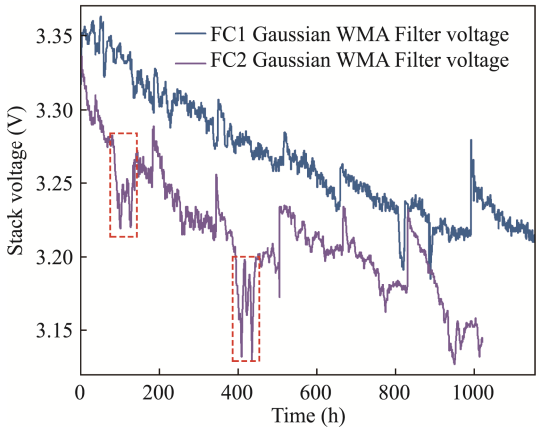

Fig.2 Degradation data of stack voltage

2 Methodology

2.1 Prediction Framework

In this work, the stack voltage data is used to display the performance of the PEMFC system, and the CKDG method is implemented, as shown in Fig.3.

In the framework of this method, the characteristic data affecting the output voltage of PEMFC, such as temperature, flow rate, humidity, current density, and superimposed voltage, are first decomposed into n Intrinsic Mode Function(IMF) sequences by CEEMDAN and the nth IMF sequence is also called residual sequence. Then, it is input into the KPCA algorithm to complete the high-weight data extraction. The Gaussian kernel is used as the kernel function to extract the IMF component sequence with a contribution more significant than 90% and reintegrate it into a new stack voltage-related data set. Finally, the DA-GRU network is used to extract the feature data weights at each time selectively, and the long-term time dependence is completely captured. The multi-input single-output mode obtains the final stack voltage prediction result.

2.2 CEEMDAN

Adaptive noise complete ensemble empirical mode decomposition CEEMDAN algorithm adaptively adjusts the noise coefficient based on the EMD algorithm to generate Gaussian noise with different signal-to-noise ratios. It introduces the signal to be decomposed, which effectively avoids modal aliasing and almost no reconstruction error[20]. The operation data of PEMFC usually have nonlinear and non-stationary characteristics, which may contain noise introduced by measurement errors, environmental changes, and other factors. CEEMDAN can effectively process nonlinear and noisy feature data and extract more obvious feature data. The specific steps are as follows:

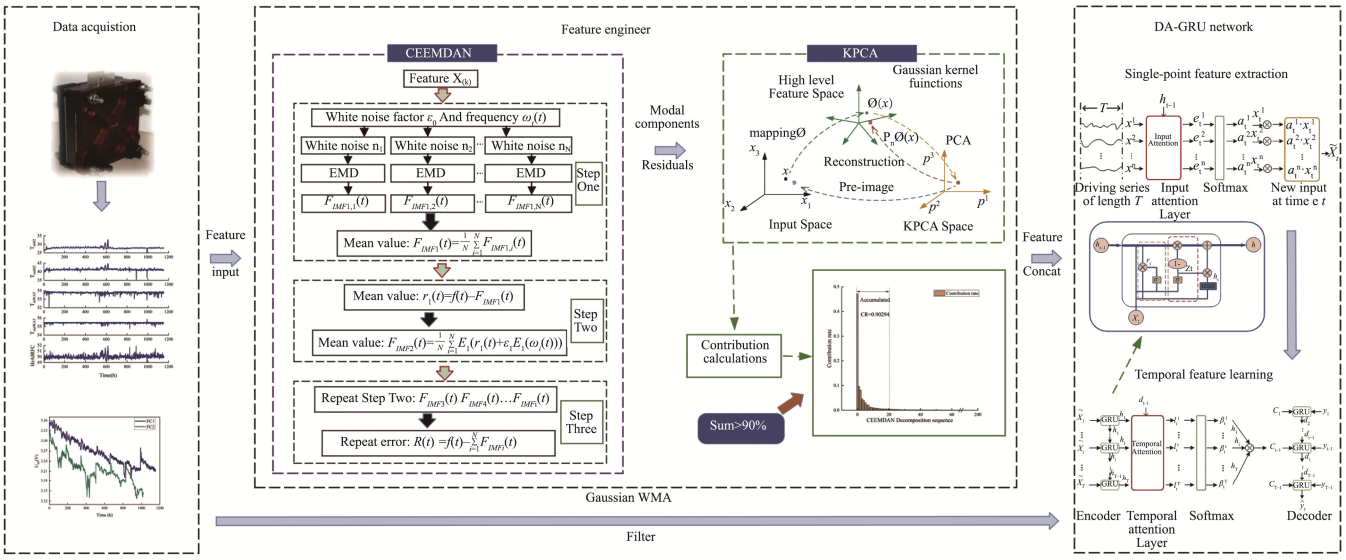

Fig.3 Framework of CKDG data prediction method

Step 1: The white noise sequence with a K times mean value of 0 is added to the decomposition sequence $f(t)$, and the signal sequence $f_i(t)$ containing noise is obtained. $F_{IMF1,i}(t)(i=1, 2, ..., N)$ is obtained after N times EMD decomposition of this sequence, and the arithmetic average value is taken to obtain the first modal component $F_{IMF1}(t)$:

$$\begin{cases} f_i(t) = f(t) + \varepsilon_0 \omega_i(t) \\ F_{IMF1}(t) = \dfrac{1}{N}\sum_{i=1}^{N} F_{IMF1,i}(t) \end{cases} \tag{1}$$

where ε_0 is the noise coefficient, $\omega_i(t)$ is the white noise sequence which obeys the standard normal distribution added by the i th decomposition.

Step 2: Calculate the first residual component : $r_1(t) = f(t) - F_{IMF1}(t)$, define $E_j(\cdot)$ as the jth modal component after EMD decomposition of the sequence, then decompose $r_1(t) + \varepsilon_1 E_1(\omega_i(t))$ to obtain:

$$F_{IMF2}(t) = \frac{1}{N}\sum_{i=1}^{N} E_1\left(r_1(t) + \varepsilon_1 E_1\left(\omega_i(t)\right)\right) \tag{2}$$

Step 3: Repeat step 2 to obtain the remaining modal components and residual components until the number of extreme points of the obtained residual components is less than or equal to 2, the decomposition is stopped, and the intrinsic modal component F_{IMFK} and the final residual component $R(t)$ are obtained. The final residual component can be expressed as:

$$R(t) = f(t) - \sum_{i=1}^{K} F_{IMFi}(t) \tag{3}$$

2.3 KPCA

The feature data set of PEMFC contains multiple feature components after CEEMDAN processing, and low correlation components need to be removed to reduce the difficulty of model processing.The KPCA method is a classical nonlinear dimensionality reduction method based on the kernel, which can retain the global characteristics of data. The basic principle of the algorithm is to use the nonlinear function $\phi(x)$ to map the original data x_i to the high-dimensional space K to reduce the original data's nonlinearity first. Then, PCA is performed on the data in the high-dimensional space[21]. The primary calculation process of the KPCA method is to solve the eigenvalues and eigenvectors of the covariance matrix in high-dimensional space K:

The covariance matrix of the mapped data in the high-dimensional space K is as follows:

$$Cov = \frac{1}{N}\sum_{i=1}^{N} \phi(x_i)(x_i)^T \tag{4}$$

By converting Eq. (4), the characteristic equation can be obtained as follows:

$$\lambda \cdot w = Cov \cdot w \tag{5}$$

Where λ is the eigenvalue and w is the corresponding eigenvector. Converting Eq. (5), we can get:

$$\lambda \cdot \left[\phi(x_j), w\right] = \left[\phi(x_j), Cov \cdot w\right] \tag{6}$$

In Eq. (6), w can be replaced by $w = \sum_{i=1}^{N} v_i \phi(x_i)$, $v_i \in v = (v_1, v_2, \ldots, v_n)$, so that we get:

$$\begin{aligned} &\lambda \sum_{i=1}^{N} v_i \left[\phi(x_j),\ \phi(x_i)\right] = \\ &\frac{1}{N}\sum_{i=1}^{N} v_i \left[\phi(x_j),\ \sum_{k=1}^{N}\phi(x_k)\right]\left[\phi(x_k),\ \phi(x_i)\right] \end{aligned} \tag{7}$$

The kernel function $Ker(x_i, x_j) = \left[\phi(x_i),\ \phi(x_j)\right]$ is introduced, and the projection of the mapping data in the high-dimensional space K is calculated to be $Proj_i = \sum_{i=1}^{N} v_i^j Ker(x_i, x_j)$. The kernel function used in the KPCA method is the Gaussian kernel function $KG(x_i, x_j) = \exp\left(\dfrac{-\|x_i - x_j\|^2}{2\sigma^2}\right)$, where σ is the bandwidth of the Gaussian kernel function.

2.4 DA-GRU Network

The traditional GRU network has gradient disappearance and explosion, and the traditional encoder-decoder network cannot capture long-term dependencies[22].In order to improve this limitation, a dual attention mechanism is introduced to comprehensively select the parameters that affect the output voltage of PEMFC. Based on Seq2Seq, an attention mechanism is introduced in both the encoder and decoder stages to form a DA-GRU network to achieve more targeted feature selection and enhance the neural network's grasp of temporal dependencies.

In the encoder, the feature attention base on GRU (FA-GRU) is used to selectively determine the importance weight of the relevant driving sequence when predicting the current target sequence. In the decoder, we use the Temporal Attention base on GRU (TA-GRU) to select the temporal importance weight of the hidden state of the relevant encoder. Using these two attention mechanisms, DA-GRU can accurately select relevant input features, capture the long-term time dependence of the time series, and finally output the predicted value of the voltage series for the future. Fig.4 shows the visual description of the model, where:

(a) Encoder:

In the overall framework, the encoder can be regarded as an RNN. For the time series prediction problem, given the input sequence $X=(x_1, x_2, ..., x_T)$, $x_t \in R^n$ (where n is the number of driving sequences), the encoder can be represented by a map from x_t to h_t as follows:

$$h_t = f_1\left(h_{t-1}, x_t\right) \tag{8}$$

Where $h_t \in R^m$ is the hidden size of the encoder at time t, m is the size of the hidden state, f_1 is a nonlinear function, and the GRU structure is selected as f_1 to capture the long-term dependence of the hidden state. As shown in Fig.5, each GRU structure has a cell state with a state of s_t at time t, which is retained in the hidden state of the GRU and will not be completely exposed to the outside. The access to the cell state is controlled by two sigmoid gates: r_t (the reset gate) and z_t (the update gate). The update of the GRU structure can be summarized as follows (9):

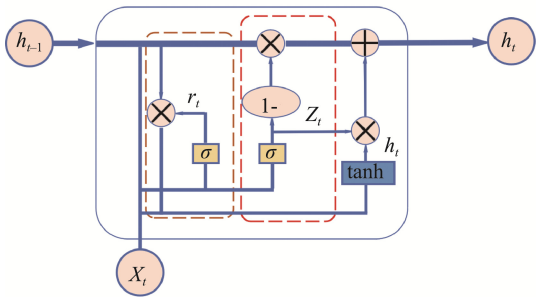

Fig.5 GRU Architecture Flowchart

$$\begin{cases} r_t = \sigma\left(W_r \cdot \left[h_{t-1}, x_t\right]\right) \\ z_t = \sigma\left(W_z \cdot \left[h_{t-1}, x_t\right]\right) \\ \tilde{h}_t = \tanh\left(W_h \cdot \left[r_t * h_{t-1}, x_t\right]\right) \\ h_t = \left(1 - z_t\right) * h_{t-1} + z_t * \tilde{h}_t \end{cases} \tag{9}$$

Among them, x_t is the input of the neuron, h_{t-1} is the state of the hidden layer at the previous moment, and W_r, W_z and W_h are the weight matrices of the reset gate r_t, the update gate z_t and the hidden layer, respectively.

Based on the given kth input drive sequence, $x^k = \left(x_1^k, x_2^k, \cdots, x_T^k\right) \in R^T$, referring to the hidden state h_{t-1} and the unit state s_{t-1} at the previous moment, the feature attention mechanism is constructed in the GRU unit of the encoder:

$$e_t^k = v_e^T \tanh\left(W_e h_{t-1} + U_e x^k\right) \tag{10}$$

$$\alpha_t^k = \frac{\exp\left(e_t^k\right)}{\sum_{i=1}^{n} \exp\left(e_t^i\right)} \tag{11}$$

Among them, $v_e \in R^T$, $W_e \in R^{T \times m}$ and $U_e \in R^{T \times T}$ are the parameters that need to be learned. α_t^k is the importance weight of the kth input feature at time t. The softmax function is applied to e_t^k to ensure that the sum of ownership weights is 1. Using these attention weights, the driver sequences related to the predicted target sequence can be selectively extracted.

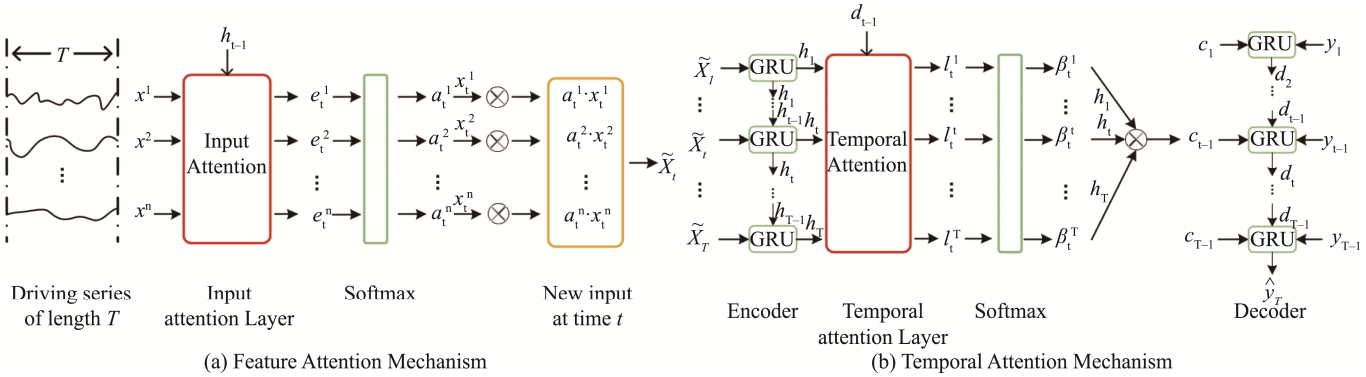

Fig.4 DA-GRU Architecture Flowchart

$$\tilde{X}_t = \left(\alpha_t^1 x_t^1,\ \alpha_t^2 x_t^2,\ \cdots,\ \alpha_t^n x_t^n\right)^T \tag{12}$$

Then the hidden state update at time t can be expressed as: $h_t = f_1(h_{t-1},\ \tilde{x}_t)$, and the weight matrix is updated according to the GRU, replacing x_t with the newly calculated $\tilde{x}_t$.

(b) Decoder:

The decoder introduces temporal attention based on the GRU network and uses it to quantify the influence of hidden states at different historical moments on the current state. The influence of hidden states at high weight moments is highlighted through weight distribution, and the importance weights of hidden states of related encoders in all time steps are selectively determined, which provides more accurate information for long-time series prediction and improves prediction accuracy[23].

According to the previous hidden state $d_{t-1} \in R^p$ of the decoder, the temporal attention mechanism weight coefficient l_t^i of the hidden state of each encoder at time t is calculated:

$$l_t^i = v_d^T \tanh\left(W_d d_{t-1} + U_d h_i\right),\ 1 \leqslant i \leqslant T \tag{13}$$

Where v_d and W_d is the weight parameter of the multi-layer perceptron, and U_d is the bias parameter of the multi-layer perceptron.

The obtained weight coefficient l_t^i of the temporal attention mechanism is normalized by the softmax function of Equation (14), and the influence of different hidden states $h_i (i = 1,\ 2,\ \cdots,\ T)$ on the current hidden state is quantified, which is recorded as β_t^i.

$$\beta_t^i = \frac{\exp\left(l_t^i\right)}{\sum_{j=1}^{T} \exp\left(l_t^j\right)} \tag{14}$$

Where β_t^i denotes the predictive importance weight of the hidden state of the i th encoder. Since the hidden state h_i of each encoder is a nonlinear mapping of the input data that changes over time, the attention mechanism calculates the context vector c_t as the weighted sum of all hidden state $\{h_1,\ h_2,\ ...,\ h_T\}$ of the encoder:

$$c_t = \sum_{i=1}^{T} \beta_t^i h_i \tag{15}$$

Where c_t changes with the change of time step. Similar to the encoder, they are combined with the history value $(y_1,\ y_2,\ \cdots,\ y_{T-1})$ of the target sequence to obtain:

$$\tilde{y}_{t-1} = \tilde{w}^T \left[y_{t-1}; c_{t-1}\right] + \tilde{b} \tag{16}$$

Where $\left[y_{t-1};\ c_{t-1}\right] \in R^{m+1}$ is the concatenation matrix of the context vector c_{t-1} calculated in decoder inputs y_{t-1} and expression (15), and parameters $\tilde{w}^T \in R^{m+1}$ and $\tilde{b} \in R$ map the concatenation matrix to the size of the decoder input. The calculated $\tilde{y}_{t-1}$ in (16) can be used to update the hidden state of the decoder at time t:

$$d_t = f_2\left(d_{t-1},\ \tilde{y}_{t-1}\right) \tag{17}$$

The GRU is selected as the nonlinear function f_2. In this kind of NARX model, the purpose is to use the DA-GRU network to approximate the function F and use the historical values of the target sequence and the driving sequence to obtain the current output $\hat{y}_T$ of the target sequence. $\hat{y}_T$ can be expressed by Eq. (18):

$$\begin{aligned} \hat{y}_T &= F\left(y_1,\ \cdots,\ y_{T-1},\ x_1,\ \cdots,\ x_T\right) \\ &= v_y^T \left(W_y \left[d_T;\ c_T\right] + b_w\right) + b_v \end{aligned} \tag{18}$$

Among them, $\left[d_T; c_T\right] \in R^{p+m}$ is the concatenation matrix of the decoder hidden state and the context vector. Parameters $W_y \in R^{p\times(p+m)}$ and $b_w \in R^p$ map the concatenation matrix to the decoder hidden state. The linear function with parameters $v_y \in R^p$ and $b \in R$ gives the final prediction result of the network.

3 Performance Index Test

In this section, the analysis is mainly divided into three steps: 1) The CEEMDAN and KPCA algorithms are used to optimize the feature data sequences of the two data sets, and the data complexity is reduced after decomposition and dimensionality reduction. 2) Verify the CKDG method under the FC1 data set, compare the prediction performance of the other three methods, and compare the model prediction ability under different training set lengths. 3) Based on the FC2 dataset, the CKDG method is used to predict the length of different training sets, and the universality of the model framework is verified.

Program in Windows 11 64bit computer with AMD Ryzen 7 5800H CPU 3.2 GHz and NVIDIA GeForce GTX 3060 Laptop GPU with 6 GB memory. The CEEMDAN-KPCA is based on MATLAB R2022a, while CKDG framework is conducted based on Python3.9.16, Pytorch1.12.0 environment.

In order to evaluate the performance of the CKDG method proposed in this paper, different performance indicators are used to evaluate the prediction effect of CKDG from the aspect of prediction accuracy. Including root mean square error RMSE, mean absolute percentage error MAPE, mean absolute error MAE, coefficient of determination R2. These indicators are defined as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \hat{x}_i\right)^2} \tag{19}$$

$$MAPE = \left(\frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i - \hat{x}_i}{x_i}\right|\right) * 100 \tag{20}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \hat{x}_i\right| \tag{21}$$

$$R^2 = 1 - \frac{\sum_{i}\left(\hat{x}_i - x_i\right)^2}{\sum_{i=1}\left(\overline{x}_i - x_i\right)^2} \tag{22}$$

Among them, x_i - the actual value of the output voltage of the fuel cell stack ; $\hat{x}_i$ - the predicted value of the output voltage of the fuel cell stack ; $\overline{x}_i$ -The mean value of the output voltage of the fuel cell stack. The smaller the first three indicators, the higher the prediction accuracy. The closer R2 is to 1, indicating that the higher the fitting degree of prediction, the better the prediction accuracy.

3.1 Case I:FC1/FC2 Data Set Optimization

Based on FC1 and FC2 data sets, CEEMDAN is used to expand the data. CEEMDAN modal decomposition is performed on 23 sets of characteristic data sequences except the heap voltage sequence, which is decomposed into n intrinsic mode function (IMF) sequences. The nth IMF sequence is also called the residual sequence. The decomposition results are shown in Table 3. There are 240 groups of samples after the decomposition of FC1 feature data, including modal components (217 groups) and residuals (23 groups). There are 231 groups of samples after the decomposition of FC2 feature data, including modal components (208 groups) and residuals (23 groups).

In order to effectively obtain the data sequence with a high contribution to the prediction sequence, the kernel principal component analysis method is used for correlation analysis to eliminate the low contribution data. The characteristic modal components and residual sequences processed by CEEMDAN are introduced into KPCA. The KPCA method is used to reduce the dimension of multi-dimensional data affecting load forecasting, and the data sequences with a contribution rate > 90% are extracted. As shown in Fig.6, the cumulative contribution rate of 20 data of FC1 is 90.29%, and the cumulative contribution rate of 25 data of FC2 is 90.14%.

Table 3 FC1/ FC2 characteristic data subscales

Parameter	IMF of FC1	Residuals of FC1	IMF of FC2	Residuals of FC2
U1	9	1	9	1
U2	9	1	9	1
U3	9	1	9	1
U4	9	1	9	1
U5	9	1	9	1
J	10	1	9	1
I	10	1	9	1
$TinH_2$	9	1	9	1
$ToutH_2$	9	1	9	1
TinAIR	11	1	9	1
ToutAIR	9	1	9	1
TinWAT	10	1	9	1
ToutWAT	10	1	9	1
PinAIR	9	1	10	1
PoutAIR	10	1	9	1
$PoutH_2$	9	1	9	1
$PinH_2$	9	1	9	1
$DinH_2$	9	1	9	1
$DoutH_2$	10	1	10	1
DinAIR	10	1	9	1
DoutAIR	9	1	8	1
DWAT	10	1	9	1
HrAIRFC	9	1	9	1
SUM	217	23	208	23

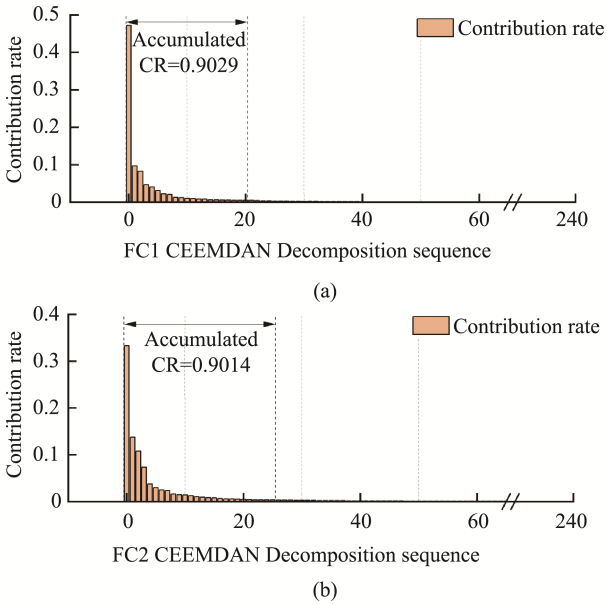

Fig.6 Contribution value distribution. (a) FC1; (b) FC2

3.2 Case II: FC1 Prediction Results

In this section, based on the FC1 data set processed by the CEEMDAN-KPCA method, the CKDG fusion prediction method and the LSTM, GRU, and DA-GRU-only methods are used to achieve short-term voltage decay degradation prediction, and 200 epochs are trained. The prediction characteristics of the four methods, when the training data set accounts for 20%, are analyzed in detail, and the prediction accuracy of the four methods is compared. The accuracy data are shown in Table 4. At the same time, in order to compare the performance of the four methods more systematically, the voltage degradation prediction results of the four methods when the proportion of training data set is 20%, 30%, 40%, and 50% are analyzed, as shown in Fig.7(a)-(d). The results show that with the increase in the proportion of training data, the prediction accuracy of all models is improved, and the CKDG fusion method maintains the best prediction accuracy in the whole process, which thoroughly verifies the stability and accuracy of its prediction.

When the training data set accounts for 20%, the LSTM method can predict the details of partial degradation of the fuel cell and performs well in the training phase. RMSE, MAPE, and R2 are 0.0032, 0.0760, and 0.9674, respectively. Fig.7(a) shows that the prediction effect of LSTM in the verification phase is could be better, and there is an over-fitting situation. RMSE and MAPE are 0.0108 and 0.2869, respectively. The prediction error increases significantly with the increase in prediction time. The data fitting degree of GRU during training is poor; R2 is 0.8060, while the RMSE, MAPE and, R2 of GRU in the verification stage are 0.0078, 0.1889 and 0.9284, respectively, which are better than LSTM, indicating that the model has better long-term prediction characteristics. The DA-GRU model has the best accuracy in the training stage, and its RMSE,

MAPE, and R2 are 0.0013, 0.0338, and 0.9944, respectively. However, the prediction accuracy in the prediction verification stage is much lower than that in the training stage, indicating that there is still a specific over-fitting situation.

In the whole life cycle of PEMFC, only the CKDG fusion prediction method can accurately fit the measured data. The RMSE and MAPE in the verification stage of the fusion method are optimal, which are 0.0051 and 0.1083, respectively, and the R2 is 0.9792. The accuracy is the highest, and the degree of fitting is good. It is worth mentioning that CKDG only uses 20% of the training data to obtain an accurate prediction effect similar to that of other studies under 40% of the training data. In order to compare the prediction accuracy of the four methods more intuitively, the data histograms of RMSE and MAPE are shown in Fig.8. It can be seen that the fusion model has better prediction accuracy at any stage.

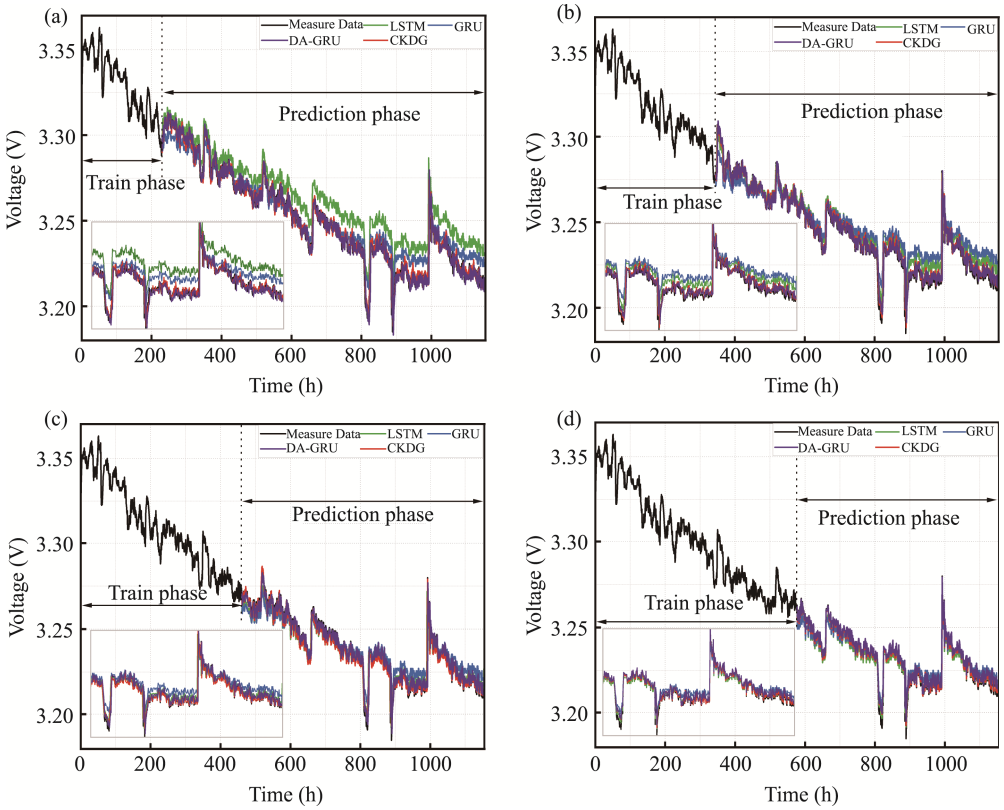

Fig.7 Short-term degradation prediction results of the four methods. (a) 20 % Training Set Training Phase; (b) 30 % Training Set Training Phase; (c) 40 % Training Set Training Phase; (d) 50 % Training Set Training Phase

Table 4 Comparison of prediction results of four methods under 20 % training set of FC1

	Training phase				Testing phase			
	LSTM	GRU	DA-GRU	CKDG	LSTM	GRU	DA-GRU	CKDG
RMSE	0.0032	0.0078	0.0013	0.0026	0.0108	0.0078	0.0061	0.0051
MAPE	0.0760	0.1936	0.0338	0.0681	0.2869	0.1889	0.1245	0.1083
MAE	0.0025	0.0065	0.0011	0.0022	0.0092	0.0061	0.0040	0.0035
R2	0.9674	0.8060	0.9944	0.9774	0.8537	0.9242	0.9539	0.9792

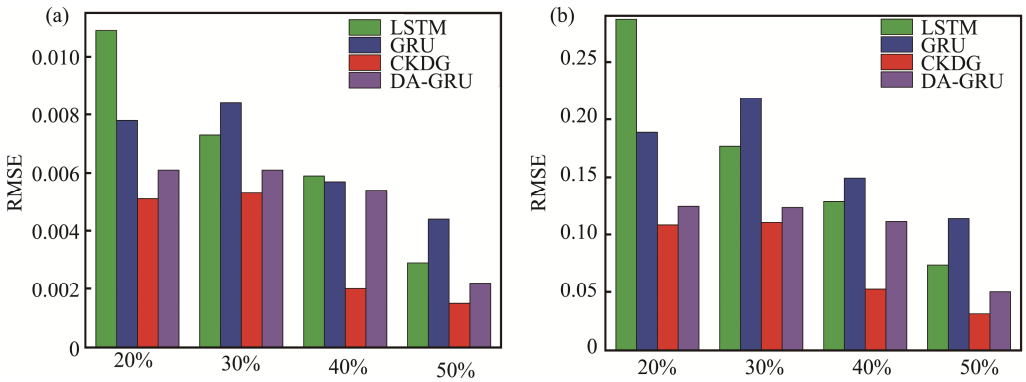

Fig.8 RMSE and MAPE results with four methods

3.3 Case III: FC2 Prediction Results

In order to further verify the generalization ability of the fusion model, based on the FC2 data set processed by CEEMDAN-KPCA, the degradation voltage data under dynamic operation are verified under different training set proportions. The error summary of the fusion model is shown in Table 5. The prediction results of the CKDG fusion prediction method in different training stages are shown in Fig.9. It can be seen that the results of CKDG fit the measured data and the data under the proportion of the first 20% training set.

The RMSE of the fusion model under the FC2 data set is 0.0039, which is 23.5% lower than that of FC1, and the MAPE is 0.0928, which is 14.3% lower than that of FC1. The R2 is approximate, the fitting degree is good, and the prediction accuracy is higher than that of FC1. This can be explained by the fact that the variable load data of fuel cells under dynamic load have a significant degradation trend in the early stage, and CKDG can accurately extract these unexplained features. With the increase of training data, the fusion model still maintains a high prediction effect, which proves that the fusion model still has accurate and stable short-term prediction ability under different working conditions. In addition, we can observe that the voltage spikes marked by red boxes in Fig.2 may reduce the model's prediction accuracy when using the normalization method for the original data because large data fluctuations will impact model learning.

Table 5 The results of prediction of CKDG method for FC2

Training phase	20%	30%	40%	50%
RMSE	0.0039	0.0034	0.0028	0.0019
MAPE	0.0928	0.0852	0.0828	0.0463
MAE	0.0030	0.0027	0.0026	0.0015
R2	0.9809	0.9832	0.9883	0.9951

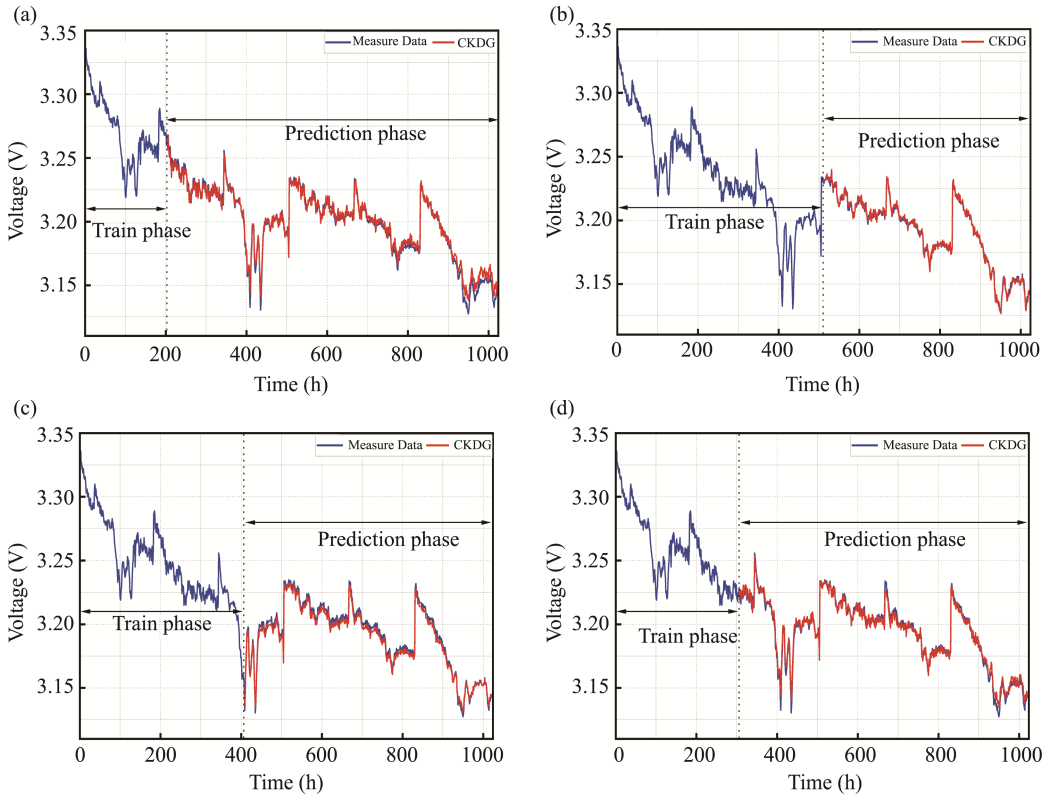

Fig.9 Short-term degradation prediction results of CKDG method for FC2. (a)20% Training Phase; (b) 30% Training Phase; (c) 40% Training Phase; (d) 50% Training Phase

4 Conclusion

This paper proposes a CKDG fusion prediction method combining CEEMDAN, KPCA data processing algorithm, and DA-GRU. The Gaussian weighted moving average filter is used to smooth the output voltage, effectively remove the noise, and maintain the original trend. CEEMDAN processes the feature data combined with the KPCA algorithm, extracting the high-contribution feature

data component. The DA-GRU network is used to achieve accurate voltage decay prediction. Based on the actual fuel cell aging experimental data set, the effectiveness of the proposed method is verified. The results show that the proposed CKDG method has a better prediction effect than other data-driven methods. Under training the first 20% of the data, the RMSE, MAE, and R2 of the CKDG method are 0.0051, 0.1083 and 0.9792, respectively, and a good prediction effect can be obtained. In addition, the universality of the method is verified based on the dynamic condition data set, which can effectively help study the battery's remaining life. Although the effectiveness of this method has been proved, the amount of aging test data of PEMFC under start-stop, idle speed, and significant load changes is still tiny, and the universal performance of the prediction algorithm is still the focus of future research.

Author Contributions:

Tingwei Zhao: Data curation, Investigation, Methodology, Software, Writing – original draft. Juan Wang: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing. Jiangxuan Che: Supervision, Data curation, Writing – review & editing. Yingjie Bian: Conceptualization, Validation. Tianyu Chen: Formal analysis, Visualization.

Funding Information:

This research was funded by Shaanxi Province Key Industrial Chain Project (2023-ZDLGY-24); Industrialization Project of Shaanxi Provincial Education Department (21JC018);Shaanxi Province Key Research and Development Program (2021ZDLGY13-02);the Open Foundation of State Key Laboratory for Advanced Metals and Materials (2022-Z01).

Data Availability:

The authors declare that the main data supporting the findings of this study are available within the paper and its Supplementary Information files.

Conflict of Interest:

The authors declare no competing interests.

Dates:

Received 17 July 2023; Accepted 17 December 2023; Published online 31 March 2024

References

[1] Li, S. (2022). Degradation prediction of proton exchange membrane fuel cell based on Bi-LSTM-GRU and ESN fusion prognostic framework. *International Journal of Hydrogen Energy*, 47(78), PP. 33466-33478.

[2] Li, Q. (2022). Approximate Cost-Optimal Energy Management of Hydrogen Electric Multiple Unit Trains Using Double Q-Learning Algorithm. *IEEE Transactions on Industrial Electronics*, 69(9), PP. 9099-9110.

[3] Ma, R. (2019). Data-Fusion Prognostics of Proton Exchange Membrane Fuel Cell Degradation. *IEEE Transactions on Industry Applications*, 55(4), PP. 4321-4331.

[4] Hu, X. (2020). Battery Lifetime Prognostics. *Joule*, 4(2), PP. 310-346.

[5] Pei, P. (2019). Nonlinear methods for evaluating and online predicting the lifetime of fuel cells. *Applied Energy*, 254, PP. 113730.

[6] Bressel, M. (2016). Remaining Useful Life Prediction and Uncertainty Quantification of Proton Exchange Membrane Fuel Cell Under Variable Load. *IEEE Transactions on Industrial Electronics*, 63(4), PP. 2569-2577.

[7] Bressel, M. (2016). Extended Kalman Filter for prognostic of Proton Exchange Membrane Fuel Cell. *Applied Energy*, 164, PP. 220-227.

[8] Chen, K. (2019). Fuel cell health prognosis using Unscented Kalman Filter: Postal fuel cell electric vehicles case study. *International Journal of Hydrogen Energy*, 44(3), PP. 1930-1939.

[9] Sun, B. (2023). Short-term performance degradation prediction of a commercial vehicle fuel cell system based on CNN and LSTM hybrid neural network. *International Journal of Hydrogen Energy*, 48(23), PP. 8613-8628.

[10] Hua, Z. (2020). Remaining useful life prediction of PEMFC systems based on the multi-input echo state network. *Applied Energy*, 265, PP. 114791.

[11] Chen, K. (2019). Degradation prediction of proton exchange membrane fuel cell based on grey neural network model and particle swarm optimization. *Energy Conversion and Management*, 195, PP. 810-818.

[12] Zou, Y. (2023). Advancements in Artificial Neural Networks for health management of energy storage lithium-ion batteries: A comprehensive review. *Journal of Energy Storage*, 73, PP. 109069.

[13] Benaggoune, K. (2022). A data-driven method for multi-step-ahead prediction and long-term prognostics of proton exchange membrane fuel cell. *Applied Energy*, 313, PP. 118835.

[14] Zuo, J. (2021). Deep learning based prognostic framework towards proton exchange membrane fuel cell for automotive application. *Applied Energy*, 281, PP. 115937.

[15] Ma, R. (2018). Data-driven proton exchange membrane fuel cell degradation predication through deep learning method. *Applied Energy*, 231, PP. 102-115.

[16] Zhang, R. (2022). Bi-directional gated recurrent unit recurrent

neural networks for failure prognosis of proton exchange membrane fuel cells. International Journal of Hydrogen Energy, 47(77), PP. 33027-33038.

[17] Liu, H. (2017). Data-based short-term prognostics for proton exchange membrane fuel cells. International Journal of Hydrogen Energy, 42(32), PP. 20791-20808.

[18] Research, F. (2014). IEEE PHM 2014 data challenge details for participants.

[19] Zhang, Z. (2021). A short- and long-term prognostic associating with remaining useful life estimation for proton exchange membrane fuel cell. *Applied Energy*, 304, PP. 117841.

[20] Zhang, W. (2017). A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. *Energy Conversion and Management*, 136, PP. 439-451.

[21] Liu, X. (2023). A novel power transformer fault diagnosis method based on data augmentation for KPCA and deep residual network. *Energy Reports*, 9, PP. 620-627.

[22] Liu, L. (2021). An enhanced encoder–decoder framework for bearing remaining useful life prediction. *Measurement*, 170, PP. 108753.

[23] Cao, L. (2023). A parallel GRU with dual-stage attention mechanism model integrating uncertainty quantification for probabilistic RUL prediction of wind turbine bearings. *Reliability Engineering & System Safety*, 235, PP. 109197.