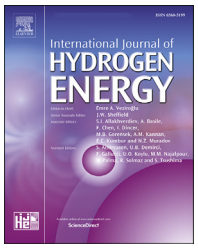
INTERNATIONAL JOURNAL OF HYDROGEN ENERGY 47 (2022) 33466–33478

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/he



Degradation prediction of proton exchange membrane fuel cell based on Bi-LSTM-GRU and ESN fusion prognostic framework

Songyang Li, Weiling Luan, Chang Wang, Ying Chen, Zixian Zhuang*

Key Laboratory of Power Battery Systems and Safety (CPCIF), Key Laboratory of Pressure Systems and Safety (MOE), School of Mechanical and Power Engineering, East China University of Science and Technology, Shanghai 200237, China

HIGHLIGHTS

- Bi-LSTM and Bi-GRU are deployed to extract degradation characteristics of PEMFC.
- The introduction of ESN reduces the risk of overfitting.
- Both the short-term degradation prediction and RUL estimation can be achieved.
- Only the first fifth of the dataset is used as the training dataset.
- The prediction accuracy is improved compared with the traditional model.

ARTICLE INFO

Article history:
Received 29 May 2022
Received in revised form
16 July 2022
Accepted 27 July 2022
Available online 25 August 2022

Keywords:
PEMFC
Prognostic
Remaining useful life
Bi-LSTM-GRU
Deep learning

ABSTRACT

The durability of proton exchange membrane fuel cell (PEMFC) is one of the technical challenges restricting its commercial applications. To enhance the reliability and durability of PEMFC, a fusion prognostic framework is proposed based on bi-direction long short-term memory (Bi-LSTM), bi-direction gated recurrent unit (Bi-GRU) and echo state network (ESN), which can achieve short-term degradation prediction and remaining useful life (RUL) estimation of PEMFC with fewer training datasets. For short-term prediction, using the first 200 h of voltage degradation data for training can achieve an acceptable and accurate prediction, with the root mean square error (RMSE), mean absolute error (MAE) and coefficient of determination (R^2) of 0.0235, 0.0195 and 0.9822, respectively. Compared with traditional machine learning methods, the proposed fusion prognostic framework shows a better predictive performance. In addition, a 100-step-sliding-windows method based on the fusion prognostic framework was implemented for RUL estimation. The results show that the percentage error (E_r) is only 1.22% with the first 200 h of training data. The proposed method has great significance for online testing and health management of PEMFC.

* *Corresponding author.*
E-mail address: luan@ecust.edu.cn (W. Luan).

https://doi.org/10.1016/j.ijhydene.2022.07.230

Introduction

PEMFC is considered to be one of the most promising clean alternative power supply for automotive power sources with high energy conversion efficiency, high power density, low operating temperature and less operating noise [1–4]. However, the large-scale commercialization of PEMFC has been restricted by its short lifespan [1,5–8]. Therefore, Prognostic and Health Management (PHM) technology is needed to predict the health status of PEMFC throughout the whole life cycle to extend its lifespan and enhance its performance.

In general, there are three approaches for PEMFC prognostic: model-based methods, data-driven methods and hybrid methods. Model-based methods rely on physical modeling of PEMFC degradation behavior (also called failure physical model) to build an empirical model for the description of the declining trajectory of PEMFC [9–15]. However, a general modeling approach for the description of all the degradation mechanisms of PEMFC does not exist yet [16]. Linear or nonlinear estimation formula developed in the literature [10,12,13] ignore some of the degradation mechanisms of PEMFC. To simplify the model, some models such as empirical degradation models combined with particle filter (PF) [11], Unscented Kalman filter (UKF) [14,17], Extended Kalman filter (EKF) [15,18] have been proposed. Unfortunately, the methods also rely on the models of PEMFC [16].

Recently, with the rapid development of semiconductor materials and digital technology, data-driven methods have gradually become a research hotspot in fuel cell degradation prediction. Compared with the model-based methods, the data-driven methods can be more flexible and applicable to the prediction of the performance degradation trend of PEMFC, regardless of the implementation mechanism and propagation rules. Nevertheless, a large amount of data is often needed for training, which is time-consuming and costly. Adaptive Neuro Fuzzy Inference Systems (ANFIS) with different fuzzy inference system creation strategies have been applied in Refs. [19,20]. Online estimation of PEMFC can be implemented with a constraint-based summation wavelet extreme learning machine (SW-ELM) integrated model [21]. Some scholars used modified support vector machine (SVM) [22–24] or relevance vector machine (RVM) [25,26] to achieve high-precision prediction. However, these works are all focused on the long-term prognostic of PEMFC. Some deep learning models such as recurrent neural network (RNN) [27–31], ESN [16,32], convolutional neural network (CNN) [33] provide a promising solution to deal with the nonlinearity and temporal problems. Ma et al. [27,28] proposed grid long short-term memory (G-LSTM) for the degradation prediction of PEMFC with almost 60% of the training data. Bi-LSTM model is proposed in Ref. [29], in which the first 550 h and 692 h operation data are used for training. The attention mechanism is introduced in combination with LSTM to realize the short-term prediction and improve the prediction accuracy of PEMFC [29,30]. However, a lot of data is also used for training. Morando et al. [32] developed an ESN for PEMFC aging prediction. Although the method can overcome the shortcomings of the traditional RNN, the simple linear fitting seems to require a large amount of training data as support.

Considering the advantages and disadvantages of the model-based method and the data-driven method, some scholars combine the two methods, namely the hybrid method [34]. Pan et al. [35] proposed a novel hybrid approach using a combination of model-based adaptive EKF and data-driven NARX neural network, which the overall degradation trend is captured by empirical aging model and KF, and NARX is used to depict the detail degradation information. Chen et al. [36] achieved degradation prediction of PEMFC based on wavelet analysis and NARX. The original voltage of PEMFC is decomposed into several sub-waveforms by wavelet analysis. Then, NARX is used to predict the degradation of each sub-waveform, and the overall degradation prediction of PEMFC is obtained by integrating the degradation results of each sub-waveform. Xie et al. [37] combined filtering method with LSTM respectively to achieve RUL estimation and short-term degradation prediction. Cheng et al. [38] presented a hybrid prognostic method for PEMFC based on least square support vector machine (LSSVM) and regularized particle filter (RPF), which can not only provide the estimate of RUL, but describe the uncertainty of RUL by probability distribution. Liu et al. [39] introduced AutoML technology firstly, combined with semi-mechanistic degradation model and adaptive UKF method to predict the long-term degradation trend of voltage.

As shown by the above-mentioned research, most studies can only perform short-term prediction or RUL estimation and require a huge proportion of training data [37]. Therefore, a comprehensive prognosis method based on Bi-LSTM-GRU and ESN is proposed innovatively in this paper, which can predict the short-term degradation and estimate RUL based on the first quintile data in the training process. The short-term prediction and RUL estimation process are shown in Fig. 1. The use of Bi-LSTM can effectively avoid the problem of gradient explosion and disappearance, and be more suitable for capturing the information of adjacent time nodes [27–31,37]. The Bi-GRU tends to grasp overall time-series information [40]. ESN replaces the traditional flatten layer as the final output layer. In addition, ESN introduces a large number of sparse neurons to further reduce the risk of overfitting, which is a simple linear fitting without considering the excessive training time. To combine the advantages of the three networks, a better fusion prognostic framework can be realized. The main contributions of this paper are summarized as follows:

1) A fusion prognostic framework for PEMFC without regarding to physical properties is proposed.
2) Short-term degradation prediction and RUL estimation of PEMFC are realized with the fusion prognostic framework.
3) The accuracy of the short-term degradation prediction is improved by the fusion method compared with traditional machine learning methods.

This paper is organized as follows: Section 2 describes two experimental PEMFC stacks. Section 3 introduces the fusion prognostic framework. The experimental results and

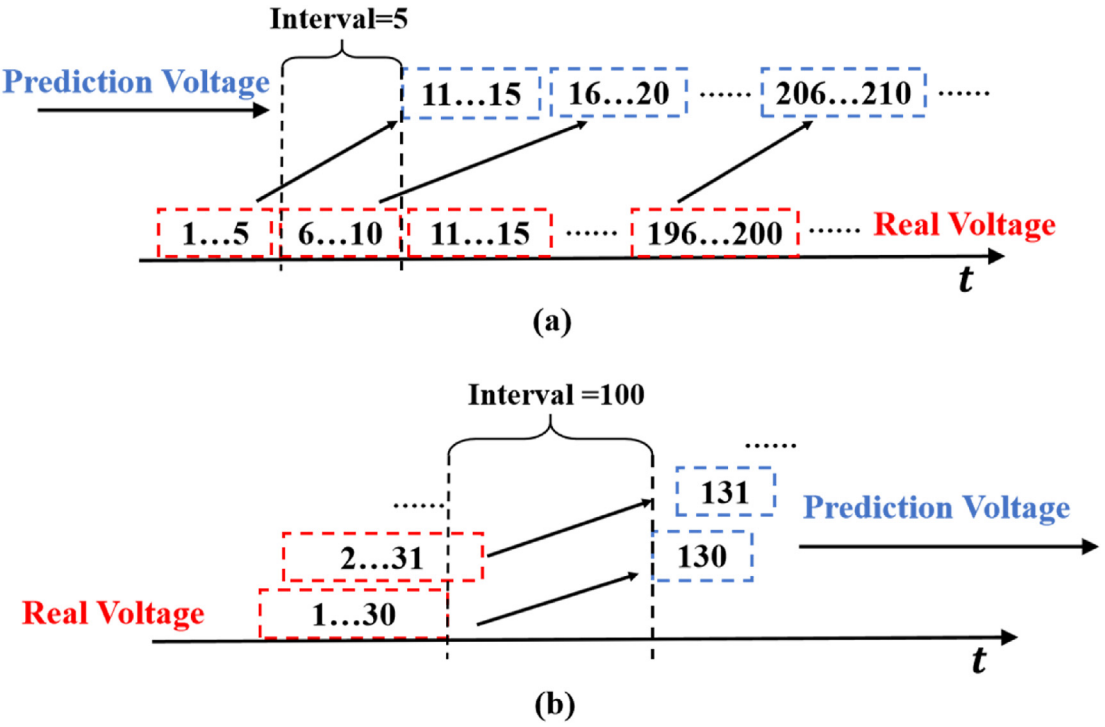


Fig. 1 – (a) Short-term degradation prediction; (b) RUL estimation.

discussion are shown in Section 4. The conclusion of this paper is given in Section 5.

Aging experimental implementation

The Bi-LSTM-GRU and ESN fusion prognostic methods are verified using data from the two experimental PEMFC stacks, which called FC1 and FC2 of the IEEE PHM 2014 Data Challenge [41]. The aging test of FC1 is carried out at a steady current of 70 A, and the FC2 experiment is carried out at a dynamic current of 70 A. The test is carried out on a test bench developed by the fuel cell laboratory (FR CNRS 3539). The test bench is suitable for PEMFC with power up to 1 KW. The test system consists of a fuel cell system, control system, electronic load and LabView interface, and the fuel cell system consists of a fuel cell stack and auxiliary systems. Gas humidification is achieved by two independent boilers upstream of the stack. Air and hydrogen pass through boilers before entering the stack when the hydrogen boiler is kept at room temperature. The temperature of the stack is controlled by the cooling water system. The control system receives commands from the LabView interface, and signals are obtained by sensors. The signal obtained by the sensor will be displayed in the LabView interface. The control signals are sent from the control system to the fuel cell system [41]. Table 1 lists the physical parameters involved in the stack.

The fuel cell stack is integrated by FCLAB, which is composed of five cells. PEMFC is realized by using an industrial membrane, diffusion layer, and processed flow distribution

Table 1 – Range of physical parameters controlled.

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

Table 2 – Aging parameters gathered during experiments.

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

plate [31]. The collected aging parameters are shown in Table 2. The output voltage of PEMFC is used as an indicator of stack degradation in this paper.

Fusion prognostic framework based on Bi-LSTM-GRU and ESN

This section begins with an introduction of Bi-LSTM and Bi-GRU. Then, a feature extraction method based on Bi-LSTM-GRU is proposed. The ESN output network realizing short-term degradation prediction and RUL estimation for PEMFC, a PSO that optimizes the hyperparameters of the feature extraction method will be introduced.

Bi-LSTM architecture

LSTM was proposed by Hochreiter et al. [42] in 1997. As a special RNN, it is capable of long-term memory. As shown in Fig. 2, an LSTM unit comprises three gated states and a cell state, where X_t is the input of the current step, h_t denotes the output of the current step and C_t is the cell state of the current step. X_{t-1}, h_{t-1} and C_{t-1} are the input, output and the cell state of the previous step, respectively.

The mathematical definition of the gate can be expressed as:

$$g(t) = \sigma(Wx + b) \tag{1}$$

As shown in Fig. 2, when the information X_t inputs into the LSTM unit, the input gate i_t determines the information which can be retained in the cell state. The output is a value of [0, 1], where 0 means completely forgotten, 1 means completely reserved. The mathematical expression is:

$$i_t = \sigma(W_i[X_t, h_{t-1}] + b_i) \tag{2}$$

The forgetting gate f_t determines which information passed from the previous moment needs to be discarded from C_t. The calculation of f_t is expressed as:

$$f_t = \sigma(W_f[X_t, h_{t-1}] + b_f) \tag{3}$$

The cell state needs to be updated. The updating process can be divided into two parts. The first part is to create a candidate value from a *tanh* layer, which contains the new information to be added. The second part is to update the information of C_{t-1}. Two parts of information are combined to update the unit. The so-called candidate value $\tilde{C}_t$ and the updated unit state C_t can be expressed as:

$$\tilde{C}_t = \tanh(W_c[X_t, h_{t-1}] + b_c) \tag{4}$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tilde{C}_t \tag{5}$$

The final output is obtained through the output gate. The mathematical expression of the output gate o_t and final output h_t can be written as:

$$o_t = \sigma(W_o[X_t, h_{t-1}] + b_o) \tag{6}$$

$$h_t = o_t \otimes \tanh(C_t) \tag{7}$$

where σ is the activation function, and the sigmod function represented here. W_x ($x = i, f, c, o$) and b_x ($x = i, f, c, o$) involved in the above formula are the weight matrix and deviation vector, respectively.

Schuster et al. [43] proposed the concept of bi-direction RNN. As shown in Fig. 3, the principle of Bi-LSTM is similar to that of bi-direction RNN. The network has two LSTM units, forward and reverse, to process time series. The sequence to be processed allows sliding in both directions, which improves the accuracy of the model. The forward layer and the Backward layer are connected to the output layer [29]. The forward layer is calculated from time 1 to t, and the state of the forward layer at each time is obtained and saved. A backward layer is calculated in the reverse direction from time t to 1. The final output is obtained by combining the hidden layer states of the forward layer and the backward layer at each moment to get. The mathematical expression is shown as follows:

$$h_t = f(W_1 x_t + W_2 h_{t-1}) \tag{8}$$

$$h'_t = f(W_3 x_t + W_4 h_{t+1}) \tag{9}$$

$$o_t = g(W_5 h_t + W_6 h'_t) \tag{10}$$

among them, h_t, h'_t represent the forward and backward layer state at time t, respectively.

Bi-GRU architecture

As a variant of LSTM, GRU has a simpler structure than LSTM with fewer internal weight parameters, which reduces overfitting to a certain extent. As shown in Fig. 4, GRU comprises only two gates: update gate and reset gate, where the update gate z_t corresponds to the input gate and forgetting gate in the LSTM unit. The memory cell state C_t of the network is not retained in the GRU unit. The reset gate r_t acts directly on the hidden state. The propagation formula of the network can be expressed as follows:

$$r_t = \sigma(W_{rx} \cdot X_t + W_{rh} \cdot h_{t-1} + b_r) \tag{11}$$

$$z_t = \sigma(W_{zx} \cdot X_t + W_{zh} \cdot h_{t-1} + b_z) \tag{12}$$

$$\tilde{h}_t = \tanh(W_{\tilde{h}h} \cdot r_t \odot h_{t-1} + W_{hx} \cdot x_t + b_{\tilde{h}}) \tag{13}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{14}$$

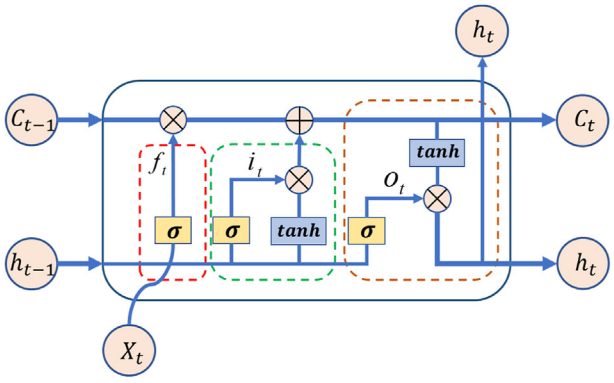


Fig. 2 – LSTM architecture flowchart.

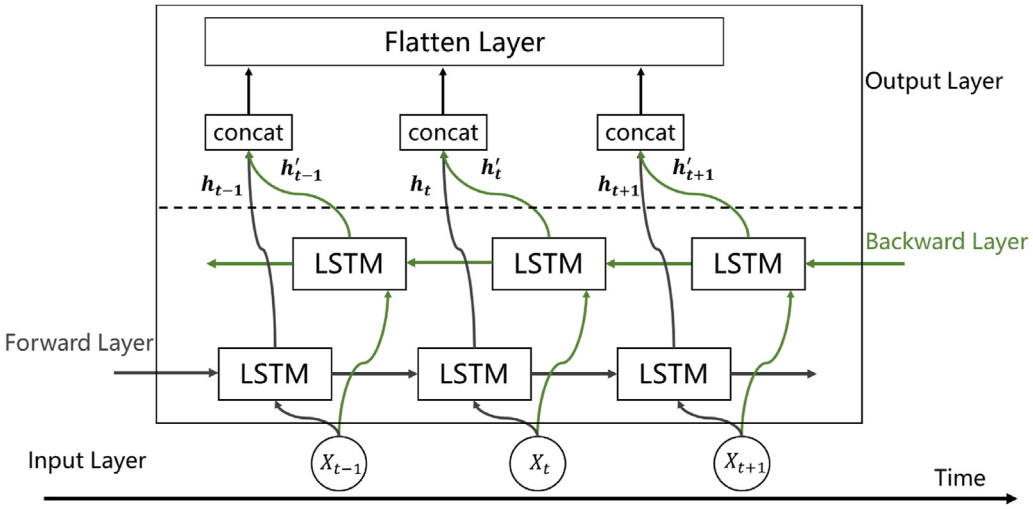


Fig. 3 – Bi-LSTM architecture flowchart.

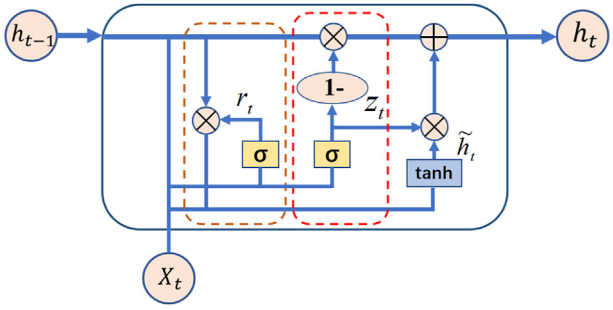


Fig. 4 – GRU architecture flowchart.

where ⊙ is Hadamard product, $\tilde{h}_t$ and h_t represent candidate hidden state and final output, respectively. The principle of the Bi-GRU is similar to that of Bi-LSTM, details of which can be found in Ref. [40].

ESN network framework

A large dynamic reservoir is introduced to replace the original hidden layer in ESN network framework. As shown in Fig. 5, the so-called reservoir is the middle part, which consists of a large number of sparse neurons. The basic idea of ESN is to generate a complex dynamic space that is constantly changing with input from the reservoir. When the state space is sufficiently complex, the output can be linearly expressed by the internal state. The reservoir has the following characteristics :

1) The connection of neurons in the reservoir is random. In other words, whether or not to establish associations between neurons is not artificially determined.
2) The input weight matrix W_{in} and recurrent weight matrix W_{res} remain constant once the structure is fixed. In the training phase, only the output weight W_{out} is changed. Therefore, ESN is optimized through the linear regression method, which greatly reduces the number of training calculations and avoids the occurrence of local optimization in optimization algorithms such as gradient descent, that is, over-fitting.

In discretized time, the status update of the reservoir can be expressed as:

$$\tilde{x}(n) = f(W_{in}u(n) + W_{res}x(n-1)) \tag{15}$$

$$x(n) = (1-\alpha)x(n-1) + \alpha\tilde{x}(n) \tag{16}$$

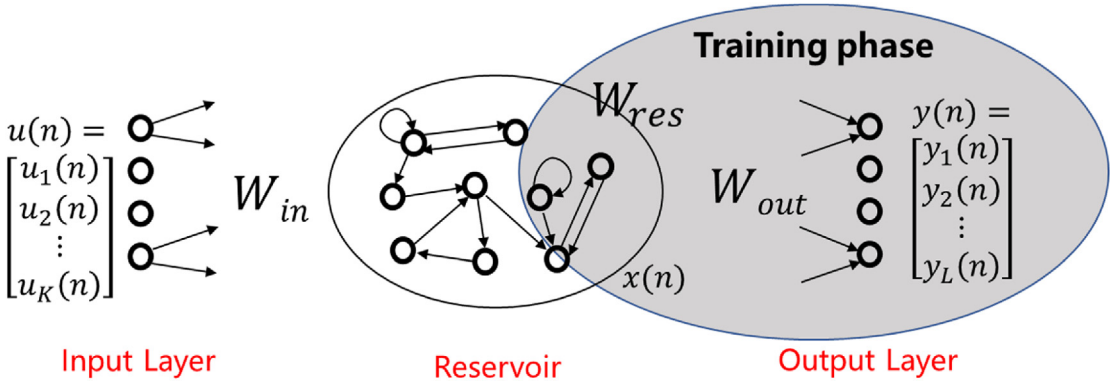


Fig. 5 – ESN architecture flowchart.

where $f(\cdot)$ is the activation function. $u(n)$ is the input vector, $\alpha \in [0,\ 1]$ is the leakage rate, $\tilde{x}(n)$ is the update state at time step n, and $x(n)$ is the state of the reservoir at time n.

The final output of ESN is represented as follows:

$$y(n) = W_{out}[u(n); x(n)] \tag{17}$$

where $y(n)$ is the output vector, [;] represents the vertical vector concatenation. This paper uses the most universal and stable method, that is, ridge regression to calculate the output weight matrix W_{out}:

$$W_{out} = Y^{target}X^T\left(XX^T + \beta I\right)^{-1} \tag{18}$$

where Y^{target} is the target output, X represents the output of reservoir, β and I are the regularization parameter and the identity matrix, respectively.

The hyperparameters of ESN include the input matrix W_{in}, the recurrent weight matrix W_{res}, the input size K, the output size L, the number of neurons in the reservoir N_{res}, the leakage rate α, the spectral radius ρ, and the regularization parameter β. The matrix is obtained from a normal distribution centered on zero. The spectral radius is used to achieve scaling of the circulant matrix and maintain better performance [44]. The general rule is that the more N_{res}, the better output performance of ESN can be, although it will take longer to fit. Therefore, there is a trade-off between the "number of neurons-fitting time". In this paper, the proper values of key parameters are taken from Ref. [16], as shown in Table 3.

PSO optimization model hyperparameters

Optimization of the relevant parameters of the Bi-LSTM-GRU model is implemented based on PSO optimization. The PSO algorithm, which originated from the study of bird predation behavior, utilizes collaboration and information sharing between individuals in a group to find the optimal solution. The individual optimal position and the global optimal position of the particle are obtained from calculating the fitness value of the objective function, by continuously updating the speed and position of the particle. In this paper, the training error is taken as the fitness value. The mathematical formula for updating speed and position is as follows:

$$v_{ij}(t+1) = wv_{ij}(t) + c_1 r_{ij}(t)\left(p_{ij}(t) - X_{ij}(t)\right) + c_2 \widehat{r_{ij}}(t)\left(p_{gj}(t) - X_{ij}(t)\right) \tag{19}$$

$$X_{ij}(t+1) = X_{ij}(t) + v_{ij}(t+1) \tag{20}$$

Table 3 – Key parameters of ESN.

Parameters	Values
Reservoir neurons N_{res}	400
Leakage rate α	[0.3, 0.5]
Spectral radius ρ	[0.4, 1.0]
Regularization parameter β	8×10^{-2}
Input weight matrix W_{in}	[-0.5, 0.5]
Recurrent weight matrix W_{res}	[-0.5, 0.5]

where i is the i-th particle, j represents the dimensional space, w indicates the inertia weight factor, v and X are the particle velocity and current position, respectively. $c_x(x = 1,\ 2)$ are learning factors, r_{ij} and $\widehat{r_{ij}}$ are independent random numbers between [0, 1], $p_{gj}(t)$ is the historical optimal position of the particle at time t. In order to ensure the global search ability and convergence ability, the weight update adopts a linearly decreasing strategy:

$$w = w_{max}\frac{t(w_{max} - w_{min})}{t_{max}} \tag{21}$$

In this paper, according to the standard PSO algorithm [45], $w_{max} = 0.9, w_{min} = 0.4$, and t represent the current number of iterations.

Prognostic implementation based on the fusion approach

The implementation of Bi-LSTM-GRU and ESN fusion prognostic framework is shown in Fig. 6. The feature extraction model based on Bi-LSTM-GRU learns the forward and reverse sequence information of the data at the same time. The depth of the feature parameters is extended by the connection of hidden vectors. In addition, reconstructed hidden vectors have more diversified features with both the memory of time and the memory of depth. The hidden vectors input into the ESN is subjected to simple linear fitting to obtain the final predicting value.

The results are executed on an Intel Core processor i5-10400CPU 2.9 GHz and NVIDIA GeForce GTX 1660s GPU with 6 GB memory. The Bi-LSTM-GRU is based on Pytorch, while ESN framework is conducted based on NumPy. The specific implementation steps are summarized as follows:

1) The aging raw data (see section 4.1 for processing details) is processed. FC1 is divided into the training set and the verification set, and FC2 is used to verify the universality of the fusion model.
2) A feature extraction method is adopted, based on deep learning, which uses the PSO algorithm to determine the hyperparameters of the prediction model. See section 4.1 for specific optimization details.
3) Adam optimization is applied in this paper. The choice of loss function will be introduced in section 4.2.
4) ESN is constructed, which determines the key parameters of the model, such as the leakage rate and the number of neurons in the reservoir.
5) Denormalization is used to the output of the fusion framework, which makes it physically meaningful and compares it with the real voltage to judge the predictive ability of the framework.

Results and discussion

In this section, the Bi-LSTM-GRU and ESN fusion prognostic framework is validated by IEEE challenge data. The prediction performance of this framework in short-term prediction and RUL estimation are discussed respectively.

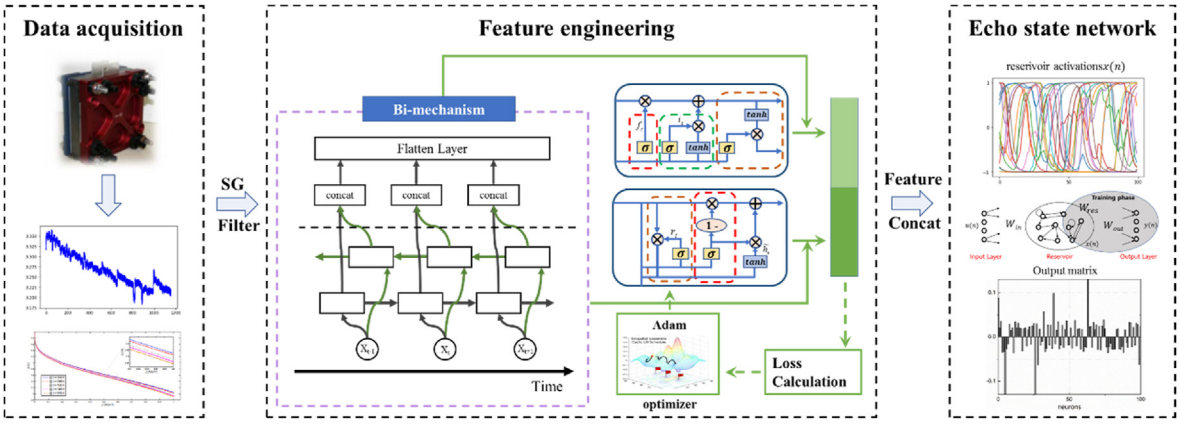


Fig. 6 – Implementation framework of the fusion prognostic method.

Data processing and model parameter determination

Since the original dataset contains a large number of noise and peaks, the SG filter algorithm is used to smooth the data stream. SG algorithm is a polynomial smoothing algorithm based on the principle of least squares, also known as convolution smoothing. Least square method is applied to fit a hypothetical higher order polynomial to enable weighted filtering of a section of data. The filter window width is selected as 51, the polynomial order is 3. As shown in Fig. 7, the smoothed data not only retain the main trend of the original information but effectively remove noise and peaks. There is some reversible degradation inside the PEMFC. Voltage recovery occurs when the operation is stopped for a short period of time or when cyclic voltammetry scan occurs, that is, the sudden rise in voltage in Fig. 7. Since the smoothed voltage has 24-dimensional attributes [31], it is necessary to normalize the data and map the data to [0, 1], which can converge quickly during training.

A two-layer hidden layer structure is adopted in the Bi-LSTM-GRU model. The number of hidden neurons in the first and second layers are recorded as h_i^1, $h_i^2 (i = lstm, gru)$, $h_i^1 \in [64, 200]$ and $h_i^2 \in [16, 100]$. The batch size is recorded as $h_0 \in [400, 600]$, learning rate is recorded as $\alpha \in [0.0001, 0.001]$. These parameters are adjusted to form the corresponding particle structure $[h_0, h_{lstm}^1, h_{gru}^1, h_{lstm}^2, h_{gru}^2, \alpha]$, where the RMSE is taken as fitness value. The fitness value is calculated according to the new position, which guides the updating of individual optimal position and global optimal position of the particles. When the number of iterations reaches the maximum, the optimal particle is obtained (h_0, h_1 and h_2 are all integers, which means rounding the result). The number of PSO iterations in this paper is 100, and the optimal parameters of Bi-LSTM-GRU model are shown in Table 4.

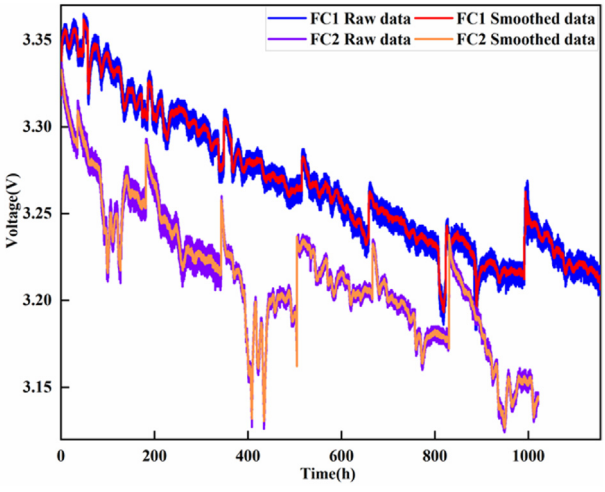


Fig. 7 – Degradation data of stack voltage.

Evaluation index of prediction accuracy

In order to evaluate the accuracy of the prediction results, it is important to give a series of evaluation indicators. Three statistical criteria standards of RMSE, MAE and R^2 are selected to evaluate the performance of short-term prediction. The smaller RMSE and MAE are, the higher accuracy of the model. The larger value of R^2 indicates the better prediction accuracy. When the prediction result is consistent with the real label, the calculation formula of $R^2 = 1$. The formulas for calculating these loss functions are shown as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{1}^{N}(Y(t) - \hat{Y}(t))^2} \tag{22}$$

$$MAE = \frac{1}{N}\sum_{1}^{N}|Y(t) - \hat{Y}(t)| \tag{23}$$

Table 4 – Optimal parameters for feature extraction framework.

	Bi-LSTM	Bi-GRU
h_0	488	488
h_1	154	140
h_2	65	96
α	0.0002	0.0002

$$R^2 = 1 - \frac{\sum_1^N (Y(t) - \widehat{Y}(t))^2}{\sum_1^N (Y(t) - \overline{Y}(t))^2} \quad (24)$$

where $Y(t)$ is the actual measured voltage, $\widehat{Y}(t)$ is the predicted voltage value, $\overline{Y}(t)$ indicates the mean value of the measured voltage, and N is the number of the measured voltage.

As for RUL estimation, the E_r between the actual RUL (RUL_{Act}) and the predicted RUL (RUL_{Prdt}) is generally used to determine the accuracy of the RUL estimation [29]:

$$E_r = \frac{|RUL_{Act} - RUL_{Prdt}|}{RUL_{Act}} \quad (25)$$

Short-term degradation prediction

In this section, the short-term degradation prediction is implemented based on Bi-LSTM-GRU and ESN fusion prognostic framework according to the aging data FC1. The prediction results are compared with the degradation prediction results of SVR, LSTM and GRU models, as shown in Fig. 8(a). When the training data is [0, 200 h], the fusion prognostic framework is more effective than other models. The fusion model and LSTM quickly converged in the training phase. A total of 200 epochs are trained, and the fusion model begins to converge in the 10th epoch. The LSTM method can predict the details of fuel cell degradation and has the best performance in the training phase. RMSE, MAE, and R^2 are 0.0097, 0.0038, 0.9981, respectively, whereas the prediction error increases substantially as prediction time increases. However, in the verification stage, LSTM does not predict well, which indicates the occurrence of over-fitting. For GRU, large fluctuations occurred in data-fitting with instability during the training period. R^2 of GRU shows 0.7125. However, the RMSE, MAE, R^2 of GRU in the verification stage are 0.0408, 0.0331, 0.9665, respectively, which are better than LSTM and indicate the long-term predictive characteristics of this model. In the whole life of PEMFC, only the fusion model can fit the measured data. The RMSE, MAE and R^2 of the four methods are shown in Table 5 the lowest RMSE and MAE in the verification stage are achieved by the fusion method of 0.0235 and 0.0195, and the R^2 shows the highest of 0.9822. In contrast, the RMSE and MAE of SVR method are obtained the highest, which are 0.0989 and 0.1063. The results show that the proposed fusion method achieves the best prediction accuracy, whereas the SVR method shows the worst prediction accuracy. It is worth mentioning that we achieve similar accurate prediction effect as those reported in other studies [29–31,36] with only the first 200 h of data for training, which is half the training sample.

In order to compare the performance of the four methods systematically, the prediction results of the four methods under different training phases are analyzed. The voltage degradation prediction results when the training phases are [0, 300 h], [0, 400 h] and [0, 500 h], are shown in Fig. 8(b)–(d). The results show that all models achieved better accuracy with the increase of training data, while Bi-LSTM-GRU and ESN fusion model maintained high-precision prediction results in the overall process, indicating good stability. It is necessary to mention that the prediction result of the fusion model seems not appear in Fig. 8(d), because of the prediction result is almost completely close to the real output voltage, which is covered by the output results of GRU and LSTM. To compare the estimation accuracy of four methods more precisely, the RMSE and MAE bar diagrams are shown in Fig. 9. It can be seen that the fusion model shows better predictive ability at any phase.

To further verify the universality of the fusion model, the degradation voltage data of FC2 under dynamic operation is used for verification. The error of the fusion model is summarized in Table 6. The prediction results of the Bi-LSTM-GRU and ESN fusion model in different training phases are shown in Fig. 10. It can be seen that the results of the fusion method can fit the measured data well with 200 h training data. The RMSE, MAE and R^2 of the fusion model are 0.0197, 0.0155, 0.9818, respectively, which seems to show a better result than FC1. The results can be explained by the obvious degradation trend of the variable load data in the early stage, where the fusion model can accurately extract these unexplainable characteristics. With the increase of training data, the fusion model still maintains high predictive results. It seems to prove that the fusion model still has the accurate and stable short-term predictive ability under different operating conditions.

RUL estimation

The short-term degradation prediction performance of the fusion method has already been verified in the preceding section. In the last part, we present RUL estimation with fusion model. When the total voltage of the stack drops to 96.5% of the initial voltage, PEMFC is defined as failure [31]. Therefore, for FC1, when T = 809 h, the total voltage of the stack (U_{tot}) is 3.203 V, which is defined as the end of life (EOL) of the stack. As for FC2, it becomes failure when T = 395 h (U_{tot} = 3.182 V). As shown in Fig. 1(b), a 100-step ahead prediction with the sliding window method is applied to obtain the degradation trend. When the predicted voltage is lower than 3.203 V (FC1) or 3.182 V (FC2), a fault is reported.

The percentage error E_r is used to evaluate the RUL estimation. The prediction results under different training phases are given in Fig. 11. For FC1, an acceptable prediction performance can be obtained with [0, 200 h] for training. Although there are large forecast fluctuations, the fusion model can track the long-term nonlinear trend in voltage. Generally, it is not possible to achieve high accuracy for tracking the short-term trend change of the voltage, but a relatively accurate RUL estimation can be performed. With the increase of training data, the fusion model can predict the short-term trend change of voltage more stably. It is interesting to note that error increases with the increase of training data increases. The reason is due to the choice of E_r as the RUL metric. An increase in training data means an increase in the operation time of PEMFC and a decrease of RUL. Therefore, E_r increase inevitably. The E_r of fusion model are summarized in Table 7. It can be found that the errors are all controlled below 3%.

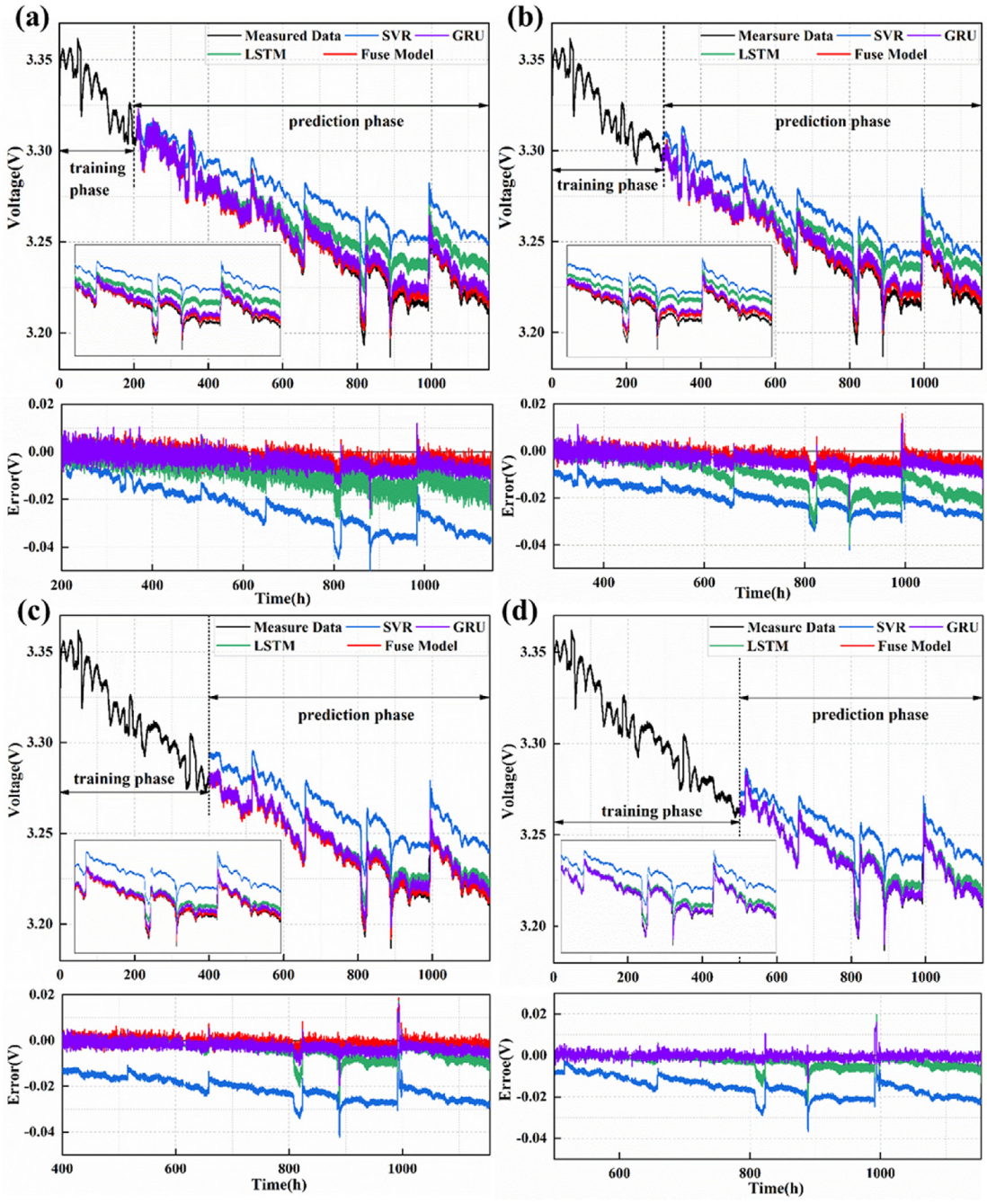


Fig. 8 – Short-term degradation prediction results of the four methods. (a) 200 h training phase; (b) 300 h training phase; (c) 400 h training phase; (d) 500 h training phase

Table 5 – Comparison of results of four models under 200 h for FC1.

	Training phase				Testing phase			
	SVR	LSTM	GRU	Fuse model	SVR	LSTM	GRU	Fuse model
RMSE	0.0478	0.0097	0.0461	0.0204	0.0989	0.0482	0.0408	0.0235
MAE	0.0408	0.0038	0.0364	0.0161	0.1063	0.0648	0.0331	0.0195
R^2	0.9331	0.9981	0.7125	0.9829	0.7946	0.8694	0.9665	0.9822

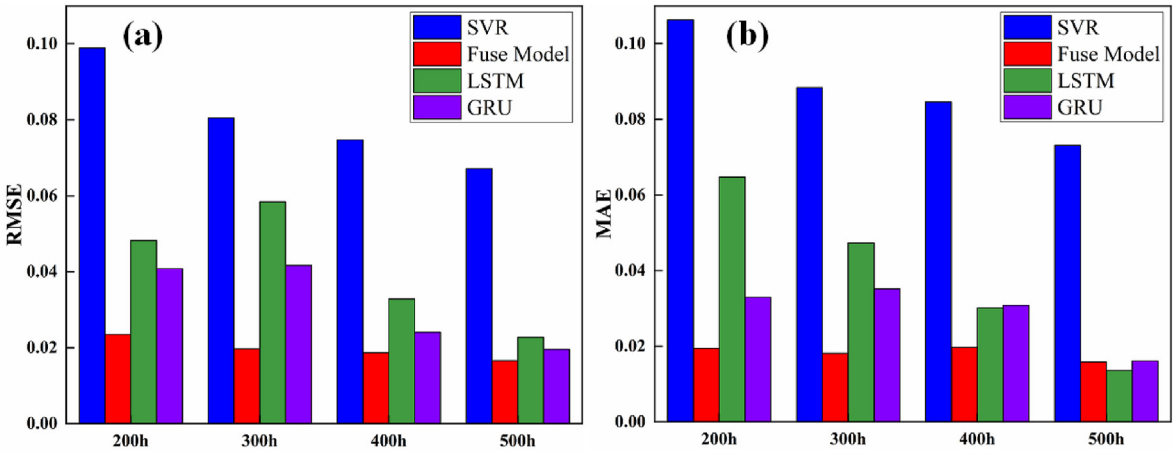


Fig. 9 – RMSE and MAE results with four methods.

Table 6 – The results of prediction of fusion prognostic method for FC2.

Training phase	200 h	300 h	400 h	500 h
RMSE	0.0197	0.0161	0.0148	0.0117
MAE	0.0155	0.0126	0.0116	0.0083
R^2	0.9818	0.9856	0.9857	0.9966

As for FC2, we can perform accurate RUL estimation using only the previous 200 h operating data, and the error is also controlled below 3%, which is 2.98%. However, due to the rapid degradation of the stack with variable load conditions, the stack reached EOL at 395 h, only the first 200 h is used for training, as shown in Fig. 11(e).

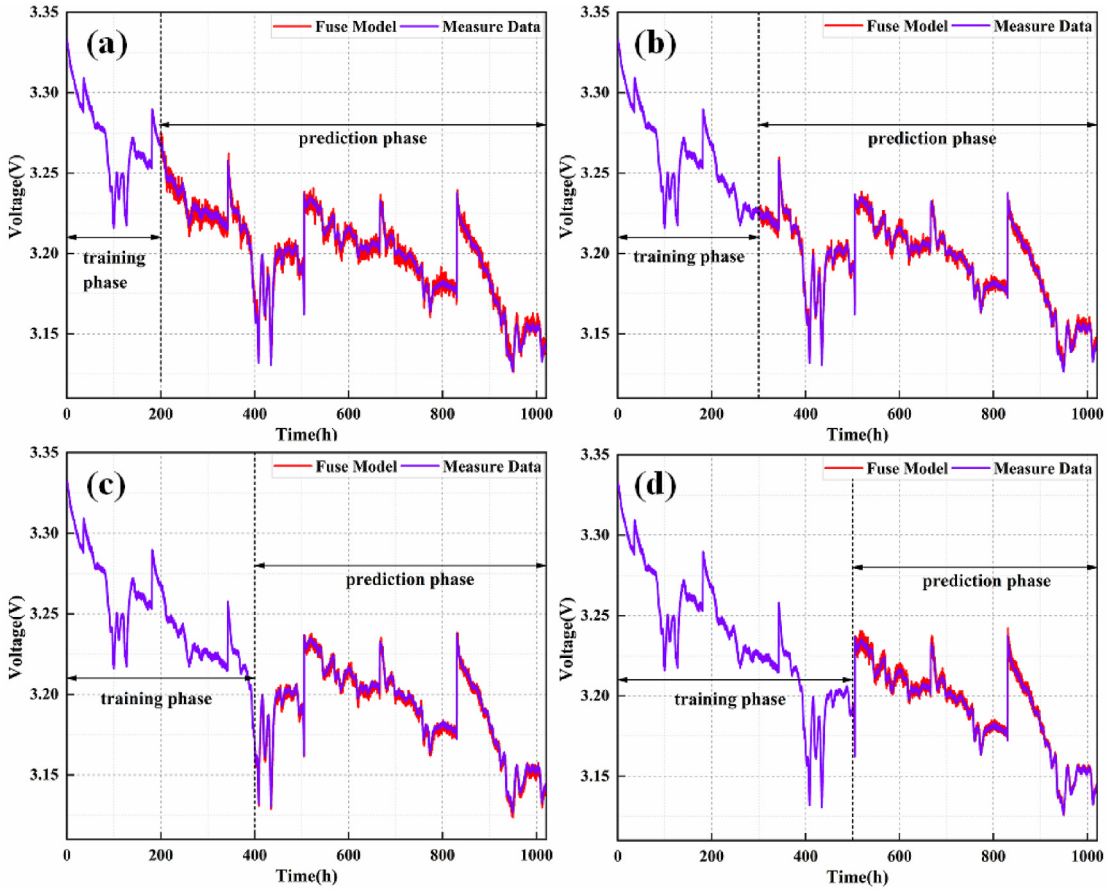


Fig. 10 – Short-term degradation prediction results of fusion prognostic method for FC2 (a)200 h training phase; (b) 300 h training phase; (c) 400 h training phase; (d) 500 h training phase.

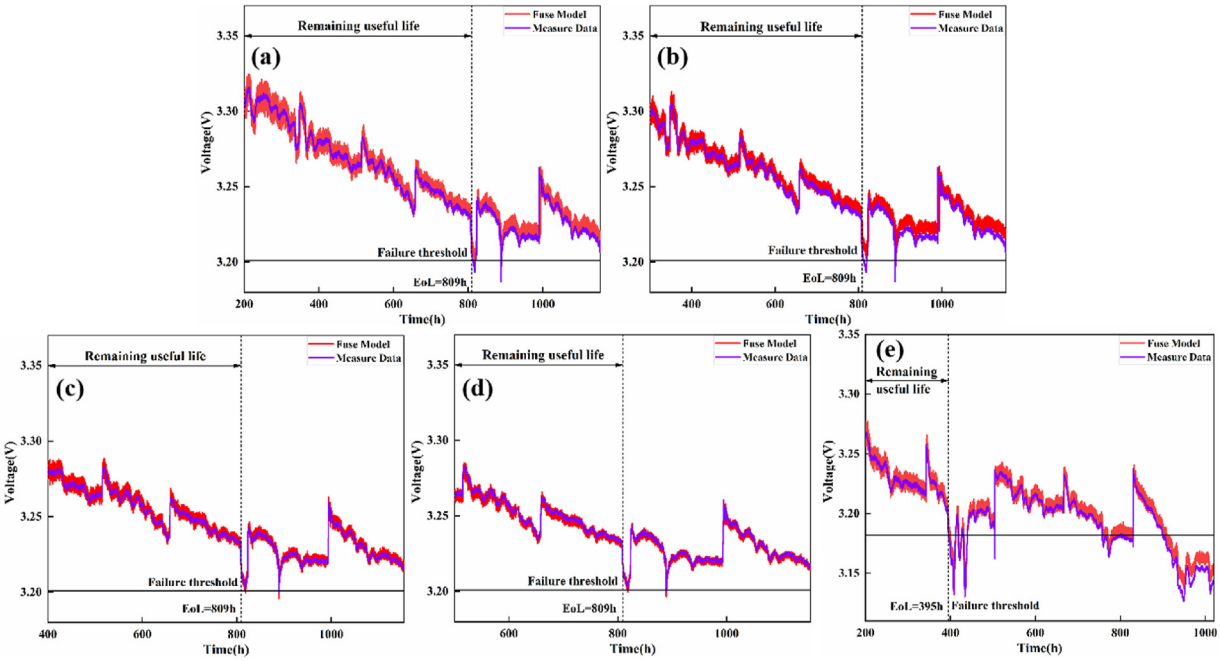


Fig. 11 – RUL estimation of fusion prognostic method with different training phases. (a) 200 h FC1; (b)300 h FC1; (c)400 h FC1; (d)500 h FC1; (e)200 h FC2.

Table 7 – The results of RUL estimation with different training phases.

Dataset	EOL	Training phase	Actual RUL	Prediction RUL	E_r
FC1	3.203 V	200 h	609 h	616.43 h	1.22%
		300 h	509 h	518.46 h	1.86%
		400 h	409 h	417.01 h	1.96%
		500 h	309 h	317.01 h	2.59%
FC2	3.182 V	200 h	195 h	189.19 h	2.98%

Conclusion

Aiming at the voltage degradation of the PEMFC stack during operation, a comprehensive fusion prognostic framework based on Bi-LSTM-GRU combined with ESN is proposed in this paper. The short-term degradation prediction and RUL estimation can be realized by this framework. The SG algorithm is used to smooth the voltage data, which can effectively remove noise and spikes while retaining the original trend. To get the best performance, PSO algorithm is used to obtain the optimal hyperparameters of the feature extraction model based on Bi-LSTM-GRU.

The short-term prediction results show that the proposed fusion model has better prediction results compared with other data-driven methods, whether under steady-state conditions or dynamic conditions. A better prediction result was achieved because of the stronger ability to extract deep features of fusion model under less training data. The RMSE, MAE, R^2 of fusion model are 0.0235, 0.0195 and 0.9822, respectively, under the first 200 h of data training. In addition, RUL estimation is conducted with the fusion model. In different training phases, the percentage error E_r is less than 3%. The relatively accurate RUL was also obtained by training with only the first 200 h of voltage data. In summary, the proposed fusion model can achieve short-term prediction and estimate the RUL under the training condition using only the first quintile of data. Although the effectiveness of this method has been proven, the complex real operating conditions have not been considered. The prediction of PEMFC under start-stop, idle speed and large load variation can be focused in the future.

Declaration of competing interest

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 paper.

Acknowledgment

This work was supported by Shanghai Automobile Industry Science and Technology Development Foundation (2111), Fundamental Research Funds for the Central Universities (JKG01211523).

REFERENCES

[1] Mehta V, Cooper J. Review and analysis of PEM fuel cell design and manufacturing. J Power Sources 2003;114(1):32–53.
[2] Connolly D, Lund H, Mathiesen B, Leahy M. A review of computer tools for analysing the integration of renewable energy into various energy systems. Appl Energy 2010;87(4):1059–82.
[3] Salameh M. Can renewable and unconventional energy sources bridge the global energy gap in the 21st century? Appl Energy 2003;75(1–2):33–42.
[4] Li Q, Meng X, Gao F, Zhang G, Chen W. Approximate cost-optimal energy management of hydrogen electric multiple unit trains using double Q-learning algorithm. IEEE Trans Ind Electron 2022;69(9):9099–110.
[5] Ade N, Wilhite B, Goyette H, Mannan M. Intensifying vehicular proton exchange membrane fuel cells for safer and durable, design and operation. Int J Hydrogen Energy 2020;45(7):5039–54.
[6] Li Q, Meng X, Gao F, Zhang G, Chen W, Rajashekara K. Reinforcement learning energy management for fuel cell hybrid system: a review. IEEE Trans Ind Electron 2022:3148568.
[7] Messing M, Kjeang E. Empirical modeling of cathode electrode durability in polymer electrolyte fuel cells. J Power Sources 2020;451:227750.
[8] Li Q, Yin L, Yang H, Wang T, Qiu Y, Chen W. Multi-objective optimization and data-driven constraint adaptive predictive control for efficient and stable operation of PEMFC system. IEEE Trans Ind Electron 2020;68(12):12418–29.
[9] Hu X, Xu L, Lin X, Pecht M. Battery lifetime prognostics. Joule 2020;4(2):310–46.
[10] Polverino P, Pianese C. Model-based prognostic algorithm for online RUL estimation of PEMFCs. In: Conference on control and fault-tolerant systems. SysTol); 2016. p. 599–604.
[11] Jouin M, Gouriveau R, Hissel D, Péra M, Zerhouni N. Prognostics of PEM fuel cell in a particle filtering framework. Int J Hydrogen Energy 2014;39(1):481–94.
[12] Chen H, Pei P, Song M. Lifetime prediction and the economic lifetime of proton exchange membrane fuel cells. Appl Energy 2015;142:154–63.
[13] Pei P, Chen D, Wu Z, Ren P. Nonlinear methods for evaluating and online predicting the lifetime of fuel cells. Appl Energy 2019;254:113730.
[14] Chen K, Laghrouche S, Djerdir A. Fuel cell health prognosis using unscented kalman filter: postal fuel cell electric vehicles case study. Int J Hydrogen Energy 2019;44(3):1930–9.
[15] Bressel M, Hilairet M, Hissel D, Bouamama B. Remaining useful life prediction and uncertainty quantification of proton exchange membrane fuel cell under variable load. IEEE Trans Ind Electron 2016;63(4):2569–77.
[16] Hua Z, Zheng Z, Péra M, Gao F. Remaining useful life prediction of PEMFC systems based on the multi–input echo state network. Appl Energy 2020;265:114791.
[17] Pisu P, Zhang X. An unscented kalman filter based approach for the health–monitoring and prognostics of a electrolyte membrane fuel cell polymer. In: 2012 Annual conference of prognostics and health. Management Society; 2012. p. 1–9.
[18] Li H, Ravey A, N'Diaye A, Djerdir A. Online adaptive equivalent consumption minimization strategy for fuel cell hybrid electric vehicle considering power sources degradation. Energy Convers Manag 2019;192:133–49.
[19] Liu H, Chen J, Hissel D, Su H. Short-term prognostics of PEM fuel cells: a comparative and improvement study. IEEE Trans Ind Electron 2018;66(8):6077–86.
[20] Silva R, Gouriveau R, Jemei S, Hissel D, Boulon L, Agbossou K, et al. Proton exchange membrane fuel cell degradation prediction based on Adaptive Neuro–Fuzzy Inference Systems. Int J Hydrogen Energy 2014;39:11128–44.
[21] Javed K, Gouriveau R, Zerhouni N, Hissel D. Prognostics of proton exchange membrane fuel cells stack using an ensemble of constraints based connectionist networks. J Power Sources 2016;324:745–57.
[22] Chen K, Laghrouche S, Djerdir A. Remaining useful life prediction for fuel cell based on support vector regression and grey wolf optimizer algorithm. IEEE Trans Energy Convers 2022;37(2):778–87.
[23] Li C, Zhu X, Cao G, Sui S, Hu M. Identification of the Hammerstein model of a PEMFC stack based on least squares support vector machines. J Power Sources 2008;175:303–16.
[24] Zhong Z, Zhu X, Cao G, Shi J. A hybrid multi–variable experimental model for a PEMFC. J Power Sources 2007;164:746–51.
[25] Wu Y, Breaz E, Gao F, Miraoui A. A modified relevance vector machine for PEM fuel–cell stack aging prediction. IEEE Trans Ind Appl 2016;52:2573–81.
[26] Wu Y, Breaz E, Gao F, Paire D, Miraoui A. Nonlinear performance degradation prediction of proton exchange membrane fuel cells using relevance vector machine. IEEE Trans Energy Convers 2016;31:1570–82.
[27] Ma R, Yang T, Breaz E, Li Z, Briois P, Gao F. Data-driven proton exchange membrane fuel cell degradation predication through deep learning method. Appl Energy 2018;231:102–15.
[28] Ma R, Breaz E, Liu C, Bai H, Briois P, Gao F. Data-driven prognostics for PEM fuel cell degradation by long short-term memory network. In: IEEE transportation electrification conference and expo. ITEC); 2018. p. 102–7.
[29] Wang F, Mamo T, Cheng X. Bi–directional long short–term memory recurrent neural network with attention for stack voltage degradation from proton exchange membrane fuel cells. J Power Sources 2020;461:228170.
[30] Zuo J, Lv Hong, Zhou D, Xue Q, Jin L, Zhou W, et al. Deep learning based prognostic framework towards proton exchange membrane fuel cell for automotive application. Appl Energy 2021;281:115937.
[31] Liu J, Li Q, Chen W, Yan Y, Qiu Y, Cao T. Remaining useful life prediction of PEMFC based on long short-term memory recurrent neural networks. Int J Hydrogen Energy 2019;44(11):5470–80.
[32] Morando S, Jemei S, Hissel D, Gouriveau R, Zerhouni N. Proton exchange membrane fuel cell ageing forecasting algorithm based on echo state network. Int J Hydrogen Energy 2017;42(2):1472–80.
[33] Li J, Li X, He D. A directed acyclic graph network combined with CNN and LSTM for remaining useful life prediction. IEEE Access 2019;7:75464–75.
[34] Liu H, Chen J, Hissel D, Lu J, Hou M, Shao Z. Prognostics methods and degradation indexes of proton exchange membrane fuel cells: a review. Renew Sustain Energy Rev 2020;123:109721.
[35] Pan R, Yang D, Wang Y, Chen Z. Performance degradation prediction of proton exchange membrane fuel cell using a hybrid prognostic approach. Int J Hydrogen Energy 2020;45(55):30994–1008.
[36] Chen K, Laghrouche S, Djerdir A. Prognosis of fuel cell degradation under different applications using wavelet analysis and nonlinear autoregressive exogenous neural network. Renew Energy 2021;179(10):802–14.
[37] Xie R, Ma R, Pu S, Xu L, Zhao D, Huang Y. Prognostic for fuel cell based on particle filter and recurrent neural network fusion structure. Energy and AI 2020;2:100017.

[38] Cheng Y, Zerhouni N, Lu C. A hybrid remaining useful life prognostic method for proton exchange membrane fuel cell. Int J Hydrogen Energy 2018;43(27):12314–27.
[39] Liu H, Chen J, Hissel D, Su H. Remaining useful life estimation for proton exchange membrane fuel cells using a hybrid method. Appl Energy 2019;237:910–9.
[40] Du B, He Y, An B, Zhang C. Remaining useful performance estimation for complex analog circuit based on maximal information coefficient and bidirectional gate recurrent Unit. IEEE Access 2020;8:102449–66.
[41] FCLAB Research. IEEE PHM 2014 data challenge details for participants. 2014.
[42] Hochreiter S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. Int J Uncertain Fuzziness Knowledge-Based Syst 1998;6(2):107–16.
[43] Schuster M, Paliwal K. Bidirectional recurrent neural networks. IEEE Trans Signal Process 1997;45(11):2673–81.
[44] Lukoševičius M. A practical guide to applying echo state networks. Neural Network: Tricks Trade 2012;7700:659–86.
[45] Shi Y, Eberhart R. A modified particle swarm optimizer. IEEE World Congress On Computational Intelligence; 1998. p. 69–73.