International Journal of Hydrogen Energy 58 (2024) 839–851

Contents lists available at ScienceDirect

International Journal of Hydrogen Energy

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

Accuracy improvement of fuel cell prognostics based on voltage prediction

Chang Liu [a], Jiabin Shen [b], Zhen Dong [a], Qiaohui He [a], Xiaowei Zhao [a,*]

[a] Intelligent Control and Smart Energy (ICSE) Research Group, School of Engineering, University of Warwick, Coventry, CV4 7AL, UK
[b] General Motors Canada Company, Oshawa, ON, L1J 0C5, Canada

ARTICLE INFO

Handling Editor: Dr M Mahdi Najafpour

Keywords:
Deep learning
Degradation prediction
Fuel cell
Prognostics
Voltage

ABSTRACT

Proton exchange membrane fuel cell (PEMFC) is a promising hydrogen technique with various application prospects. However, all the PEMFCs are subject to degradation resulting from mechanical and chemical aging. To tackle this challenge, accurately predicting fuel cell degradation is essential for its durability optimization. In this study, an enhanced data-driven prognostic framework is developed to accurately predict short-term and medium-term degradation using only fuel cell voltage as the input feature. Firstly, a local outlier factor (LOF) algorithm is adopted for automatic detection of outliers in raw data collected from actual sensing environments. Then, an advanced deep learning model, residual–CNN–LSTM-random attention, is proposed to optimize voltage prediction to better indicate future PEMFC degradation trend. The proposed work is validated by the IEEE PHM 2014 Data Challenge. Compared to state-of-the-art methods, the proposed framework provides superior prediction accuracies with high stability. For instance, the framework improves short-term prediction, achieving a root mean square error (RMSE) of 0.0021 and a mean absolute percentage error (MAPE) of 0.0323 at steady state when training stops at 600 h. For medium-term prediction, our method also attains better results with an RMSE of 0.0085 and a MAPE of 0.4237 under same working conditions. Additionally, the comparative analyses demonstrate a lower computational burden and higher suitability of proposed work for practical applications.

1. Introduction

1.1. Background and literature review

Renewable energy sources, including hydrogen, are considered as ideal alternatives to alleviate carbon emissions and meet the growing global energy demand [1–3]. The fuel cell is a technique that converts hydrogen and oxygen into electricity with exhaust water as the only by-product [4]. Among various fuel cell types, proton exchange membrane fuel cell (PEMFC) is the most promising one, with applications in transportation and stationary power generation [5,6]. In addition, the demands for PEMFC are increasing annually [6,7]. PEMFC offers various advantages [8,9], such as low operating temperature, high energy efficiency, minimal noise level, etc. However, the unsatisfactory performance and high cost still hinder its wider commercialization. For instance, the current durability of PEMFCs for automotive driving cycles is around 5000 h and aims to reach 7000 h within the next 16 years. This falls significantly short of the target value of 8000 h [10,11]. Durability is the most important criterion for evaluating the performance of PEMFCs and is strongly affected by the degradation. Due to material worn out and improper operations, fuel cells are unavoidably subjected to degradations. However, the PEMFC degradation is a complex process [12], which involves various components. Therefore, the full understanding of the degradation is still an open challenge and further efforts are needed to enhance the PEMFC performance. Hence, it is necessary to develop an effective degradation prognostic method to assess the aging state of the entire PEMFC and capture its degradation trend, which could facilitate service maintenance and operational optimization to efficiently maximize the durability of PEMFC.

PEMFC prognostics are generally categorized into model-based and data-driven methods. The model-based one relies on semi-empirical, empirical, or physical models to predict the degradation trend. An equivalent circuit model (ECM) is presented in Ref. [13] to estimate the degradation state of a single cell. In Ref. [14], the electrochemical impedance spectroscopy (EIS) technique is utilized to build an ECM for estimating fuel cell degradation. In addition, EIS is employed along with physical models or Kalman filters in Refs. [15–20] to capture the degradation behaviors. Although the EIS technique has proved its efficacy in extracting degradation features, its application is limited due to its high cost and time-consuming nature [21]. Empirical models, Kalman

* Corresponding author. ICSE, University of Warwick, Coventry, CV4 7AL, UK.
E-mail address: Xiaowei.Zhao@warwick.ac.uk (X. Zhao).

https://doi.org/10.1016/j.ijhydene.2024.01.238
Received 11 November 2023; Received in revised form 8 January 2024; Accepted 20 January 2024
Available online 30 January 2024

filters and particle filters are proposed in Refs. [22–26] for PEMFC degradation estimations. These methods have achieved sufficient accuracies in tracking degradation. However, they require vehicle data or prior knowledge of fuel cells to fit critical parameters. Besides, these parameters may not be universally applicable in all cases. A physical model based entirely on the working principles of PEMFC is applied for the degradation prediction in Ref. [27]. In addition, the work in Ref. [28] develops a dual numerical physical model to map complex operating conditions to fuel cell degradations. Moreover, a hybrid model is built in Ref. [29] to forecast PEMFC aging trend by integrating various physical principles. Overall, the model based prognostic methods are limited by the difficulty of developing an effective mathematical model considering all aspects of fuel cell degradation [21]. Therefore, precisely predicting degradation using model-based approaches is still a challenge.

Nowadays, data-driven methods have gained significant attentions in degradation prognostics. It is considered as a "black box" model that does not rely upon analytical models and expertise on the PEMFC degradation process. A wide range of data-driven techniques have been developed for fuel cell prognostics. An adaptive neuro-fuzzy inference system (ANFIS) [30] and a machine learning method [31] are presented to predict long-term degradation, i.e., the prediction interval exceeds 168 h. Likewise, the work of [32] adopts the wavelet decomposition method to assist a data-driven model to achieve the same objective. Nonetheless, these two studies only emphasized long-term prognostics and did not address the short-term degradation prediction with prediction intervals less than 24 h. A high-precise short-term prediction is useful, as the predicted values can be adopted to compare with the reference values to identify instantaneous faults at an early stage [33]. To enhance the short-term prognostics capabilities of ANFIS, fuzzy c-means and particle swarm optimization algorithms are integrated in Ref. [34] to assist in degradation predictions. Other conventional machine learning techniques, such as relevance vector machine [35], group method of data handling [36], and wavelet-extreme learning machine [37], are applied to address short-term prognostic challenges as well.

Deep learning methods, which have demonstrated superior predictive capabilities compared to conventional methods, have also been utilized in fuel cell prognostics. The backpropagation neural network (BPNN) [38] and sparse autoencoder neural network (SAE-NN) [39] have proven their capabilities in predicting PEMFC voltage one-time step ahead. In contrast, the recurrent neural network (RNN) [40] has shown better performance in processing time series data. Ensemble echo state network (E-ESN) [41], hybrid model [42] and stacked ESN [43] are then applied to fuel cell to solve the potential gradient explosion and vanishing issue of RNN. However, they require either prior projection of raw data or knowledge of load file. To solve these problems, several algorithms such as long short-term memory (LSTM) [44], ESN with moving weight matrix [45], support vector LSTM [46], deep belief network [47] and variable input length LSTM [48] are applied and performs well in PEMFC degradation prognostics. Nevertheless, they only focus on steady-state operating conditions and their performances for dynamic or quasi-dynamic conditions are not addressed.

To this end, more research works have been conducted on the topic of degradation prognostic using deep learning methods. The stacked-LSTM (S-LSTM) [49], Grey NN [50], Resevior-ESN [51] and LSTM with Gaussian process regression [52] are applied to enhance short-term prediction under the quasi-dynamic condition. Although these works properly capture the overall fuel cell degradation trend, their performance at the late stage of voltage degradation remains untested. The Grid-LSTM [53] and bi-directional LSTM [54] are then introduced to further strengthen the prediction capabilities under both steady-state and quasi-dynamic operating conditions. In addition, nonlinear autoregressive exogenous NN [55], mind evolutionary BPNN [56] and gated recurrent unit [57] are also utilized for the same purpose. These methods all require multiple features as input, where measuring multiple fuel cell parameters in some applications can be costly or even inaccessible. To achieve the degradation prognostic task using only fuel cell output voltage, a hybrid LSTM method [58] and cycle reservoir jump ESN [59] are applied. Unfortunately, these works have not yet tackled the medium-term prediction. The longer term predictions have been achieved by using the hybrid models RCLMA in Ref. [60], Bi-LSTM-ESN in Ref. [61] and transfer transformer in Ref. [62]. All literature based on data driven methods is summarized in Table 1.

1.2. Need for research

According to the comparative Table 1 and literature review, further research is needed to address the following research gaps:

1) Several studies [44–52] fall short of providing robust short-term predictions under different working conditions.
2) Multiple features are applied for degradation prognostic in works [44–57]. Using multiple features would increase the prediction complexity and limit multi-step ahead medium-term prediction, since multiple prediction targets are required for each time step. Therefore, degradation prognostics using a single feature is preferred.
3) The medium-term prediction is not addressed in works [34–59]. However, medium-term prediction is important for PEMFC, as it could generate future voltages between 24 and 168 h to guide maintenance and energy management for PEMFC [16].
4) It is worth noting that the outliers arising from the recovery phenomenon following each characterization of the experiment are pre-processed in the literature [38–62]. Nonetheless, outliers are manually selected rather than automatically detected, which limits the potential online applications of these methods.

Overall, the previous works [38–62] has fallen short in providing reliable short or medium-term predictions under both steady-state and quasi-dynamic conditions relying solely on a single voltage feature. The accuracy of both predictions is paramount for the effective health management of PEMFC systems.

1.3. Contributions

To tackle the above issues, a structurally enhanced data-driven framework is proposed in this paper to improve the prognostic capability without knowing the working mechanism of PEMFC. The framework is capable of automatically detecting outliers, and handling both short-term and medium-term prognostics with a high precision using only the voltage as the input. The overall prognostic framework is divided into two steps. Firstly, a local outlier factor (LOF) algorithm is adopted herein to identify outliers and remove noise for raw PEMFC voltage data. The corresponding data completion and pre-processing techniques are then applied. Secondly, an innovative hybrid deep learning model is proposed to optimize the future voltage prediction of the fuel cell and reduce computational burden to better indicate its degradation trend. The efficacy of the proposed framework is validated on two fuel cell datasets of the IEEE Data Challenge [63]. The main advantages and contributions are summarized as follows.

1) Based on the LOF algorithm, the proposed framework can automatically and accurately detect outliers without any artificial manipulation, unlike other state-of-the-art frameworks [38–62]. The LOF is applied to the automatic outlier detection in fuel cell for the first time, which reduces the impact of noisy input and aids in the correct online fine-tuning of the data-driven framework.
2) A structurally enhanced hybrid deep learning model, residual-convolutional neural network-long short term memory with random self-attention (residual–CNN–LSTM-random attention), is proposed herein to improve the fuel cell degradation prognostic

Table 1
Comparisons of recent research works.

References	Prediction Accuracy	Consider quasi-dynamic/ dynamic state	Prior projection of raw data	Single feature	Consider Medium-term prediction	Automatic outlier detection	Applicability
[34–37]	Moderate	Yes	No	Yes	No	No	High
[38–43]	Moderate	No	Yes	Yes	No	No	High
[44–48]	Medium	No	No	No	No	No	High
[49–52]	Medium	Yes	No	No	No	No	High
[53–57]	Medium	Yes	No	No	No	No	High
[58,59]	Medium	Yes	No	Yes	No	No	High
[60–62]	High	Yes	No	Yes	Yes	No	Moderate
Proposed model	High	Yes	No	Yes	Yes	Yes	High

capability. It demonstrates high stability when utilizing the voltage as only input feature.

3) The proposed work illustrates its superior performance under steady-state and quasi-dynamic operating conditions through the practical aging data. The comparative results demonstrate the proposed model outperforms other state-of-the-art works [38–59,64] in both short-term and medium-term prediction accuracies. For instance, the proposed work improves the performance and achieves the root mean square error and mean absolute percentage error at 0.0021 and 0.0323, respectively, in short-term prediction.
4) In contrast with other methods [54,64], the proposed method achieves a lower time complexity along with better prediction accuracy, which demonstrates it stronger suitability for field applications.

The rest of the paper is organized as follows. Section II presents the datasets of two aging stacks. Section III describes the details of the prognostic framework. The experiment results, comparisons, and discussions are included in Section IV. Section V concludes the results and findings of this paper.

2. Fuel cell degradation data

In the IEEE PHM 2014 Data Challenge [63], two fuel cell stacks were selected for degradation tests, namely FC1 and FC2. Both stacks consist of five cells, producing a maximum total power of 1 kW. Each monolithic cell within these two stacks has a uniform active cross-sectional area of 100 cm^2, of which the nominal and maximal current densities are 0.7 A/cm^2 and 1 A/cm^2, respectively.

The aging experiments of FC1 and FC2 were conducted under two different load conditions. FC1 was operated under a steady current of 70 A, while FC2 was tested under a quasi-dynamic current of 70 A with a sinusoidal high-frequency ripple at a magnitude of 7 A and a frequency of 5 kHz. The FC1 was operated for 1154 h and a total of 143862 consecutive sets of data were recorded. In terms of FC2, a total of 127730 sets of data were recorded during its 1020 h operating period. This paper selects output voltage as the degradation indicator due to its consistency with the degeneration behaviors [49], and up samples it to an hourly frequency. The output voltages range from 3.19 V to 3.31 V and from 3.15 V to 3.32 V for FC1 and FC2, respectively.

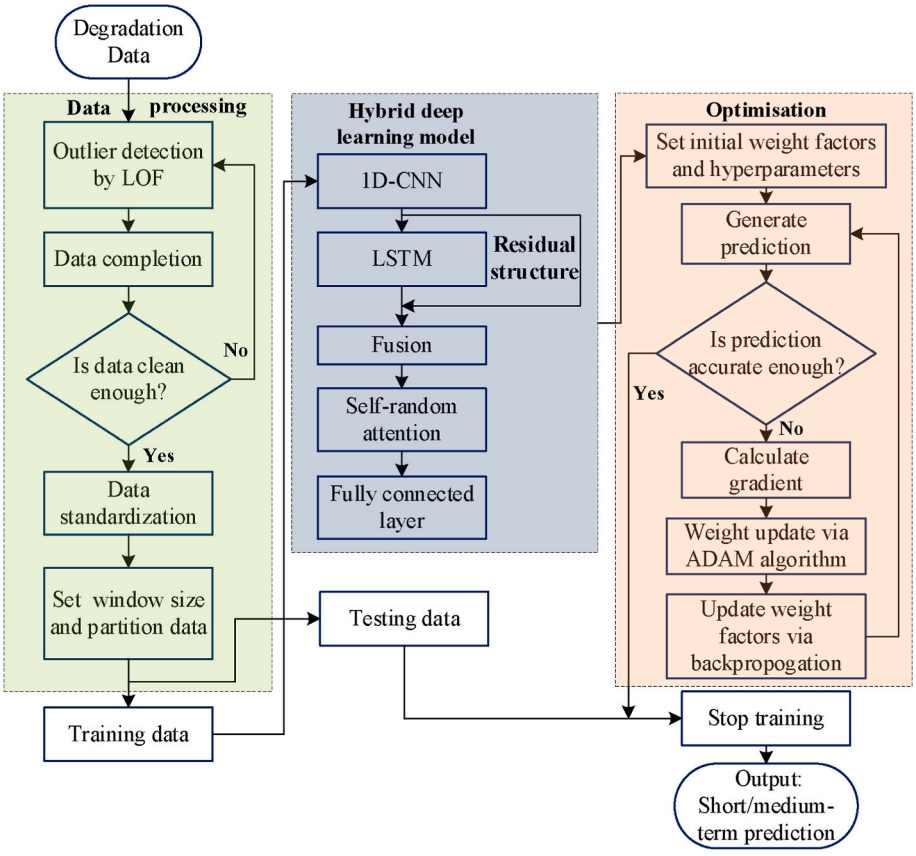

Fig. 1. Flow chart of the proposed data-driven prognostic framework.

3. Prognostic framework

This section reveals the working principles of the proposed prognostic framework. As shown in Fig. 1, the framework consists of the data pre-processing step, the proposed hybrid deep learning model, and the model optimization process. In the first stage of the framework, an outlier detection algorithm and a data completion method are applied to process the raw input data. Then, the pre-processed dataset is fed into the hybrid deep learning model, generating future voltage values to reveal the degradation trend. At the same time, the deep learning model is continuously optimized until it satisfies the training requirements.

3.1. Outlier detection and data completion

In practice, the raw data recorded by sensors is inevitably contaminated by outliers caused by the noises of the measuring devices or sudden changes in the operating conditions. In the aging experiments of FC1 and FC2, the outliers were caused by the recovery phenomenon after each characterization of polarization and EIS. Therefore, the raw voltage data exhibits sharp spikes followed by rapid declines, deviating from the normal operating values. As a result, feeding contaminated raw data into a data-driven model would affect its training process and lead to inaccurate predictions.

To tackle impacts brought by outliers, this paper proposes the use of the LOF algorithm [65] to pre-process the raw data of fuel cells. The working principle of LOF is illustrated in Fig. 2 and the Euclidean distance is defined as

$$d_E(x_i, x_j) = \sqrt[n]{\sum_{c=1}^{n} (x_{ic} - x_{jc})^2} \tag{1}$$

where d_E is the Euclidean distance introduced to measure distances between different voltage sampling points, and n represents the total dimension of each selected point. For example, the Euclidean distance between points x_i and x_j is calculated based on the voltage magnitude v and time step t. The same calculation principle applies to all other sampling points. Based on the Euclidean distance, the k-distance d_k of the point x_i is defined as

$$d_k(x_i) = d_E(x_i, x_j) \tag{2}$$

where k is the hyperparameter of the algorithm. Eq. (2) is applicable when two prerequisites are met. Firstly, the voltage sample space must contain at least k points x_j' satisfying the condition $d_E(x_i,x_j) \geq d_E(x_i,x_j')$, where x_j' represents the unified symbol of all possible points. Secondly, it is necessary to ensure that the number of data points in the voltage sample space for which $d_E(x_i,x_j) > d_E(x_i,x_j')$ is less than k. As illustrated in the blue area of Fig. 2, when k is equal to 7, 7 points meet the criteria of $d_E(x_i,x_j) \geq d_E(x_i,x_j')$ and 6 points meet the criteria of $d_E(x_i,x_j) > d_E(x_i,x_j')$, respectively. After obtaining valid $d_k(x_i)$, the k-distance domain can be written as

$$D_k(x_i) = \left\{ points\ x_j' \middle| d_k(x_i) \geq d_E\left(x_i, x_j'\right) \right\} \tag{3}$$

where $D_k(x_i)$ denotes the k-distance domain of voltage point x_i, encompassing all possible points x_j' fulfilling the requisite condition of Eq. (3). As shown in example Fig. 2, the points within the blue area form $D_k(x_i)$, while other k-distance domains within the voltage sample space are denoted by the remaining colorful areas.

In addition, the reachable k-distance is calculated as

$$r_k\left(x_i, x_j'\right) = \max\left(d_k\left(x_j'\right), d_E\left(x_i, x_j'\right)\right) \tag{4}$$

where $r_k(x_i,x_j')$ is the reachable k-distance from point x_j' to point x_i, which is determined by the maximum value between the Euclidean distance $d_E(x_i,x_j')$ and the k-distance of x_j'. For instance, the $r_k(x_i,x_j')$ in Fig. 2 is represented by the $d_E(x_i,x_j')$. Similarly, the reachable k-distance from other points x_j' in $D_k(x_i)$ to x_i are also calculated. Subsequently, they are combined with $D_k(x_i)$ to evaluate the local reachable density (LRD) for point x_i, written as

$$\rho(x_i) = \frac{|D_k(x_i)|}{\sum_{x_j' \in D_k(x_i)} r_k\left(x_i, x_j'\right)} \tag{5}$$

where $\rho(x_i)$ represents the reciprocal of the average reachable k-distance from each point in $D_k(x_i)$ to x_i. The high density indicates a greater likelihood that x_i and its adjacent points belong to the same cluster. Finally, the LOF is quantified through the density $\rho(x_i)$ as

$$F(x_i) = \frac{\sum_{x_j' \in D_k(x_i)} \rho\left(x_j'\right)}{|D_k(x_i)| \bullet \rho(x_i)} \tag{6}$$

The LOF $F(x_i)$ stands for the density of x_i relative to its adjacent points. The LOFs of other sampling points are obtained in the same way. If the value of LOF is significantly greater than 1, it indicates that the corresponding point is likely to be an outlier. In contrast, points with LOF values less than 1 and close to 1 are considered as density points and normal points relative to their adjacent points, respectively.

Subsequently, the identified outliers are replaced using local mean values. By considering the magnitude and trend of adjacent points, the local mean values rectify abnormal voltage values to reasonable levels. As a result, a complete and outlier-free voltage time series dataset is obtained.

Furthermore, this dataset requires data standardization before being fed into the prediction model. Since the model is applied to fuel cells under two different operating conditions, voltages with different numerical ranges are used for model training. To avoid the gradient saturation caused by such variations during training and ensure the weight factors perform equally for different inputs, it is crucial to standardize inputs and outputs. The mathematical illustration is summarized as

$$x' = \frac{x - \mu}{\sigma} \tag{7}$$

where μ, σ, x, and x' stand for the mean value, standard deviation, pre-processed data and standardized data of the voltage, respectively. Meanwhile, the minute noises are filtered out through the standardization process. In the following sections, to evaluate the effectiveness of the model, the predicted voltage values are reversely standardized.

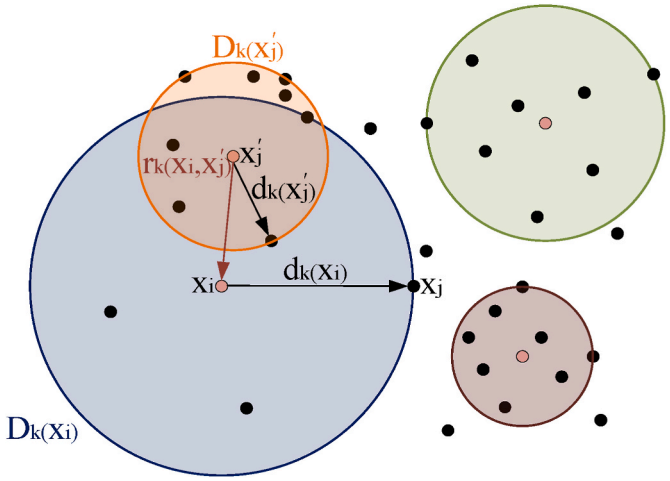

Fig. 2. Working principle of LOF algorithm.

3.2. The proposed hybrid deep learning model

The workflow of the proposed hybrid deep learning model for voltage prediction is depicted in Fig.1. The hybrid model consists of one-dimensional convolutional neural network (1D-CNN), residual LSTM, self-random attention, and fully connected blocks. The working principles of each block and hybrid model are presented as follows.

3.2.1. CNN block

The CNN block [66] has shown remarkable performances in numerous fields, especially in image classification and object detection. In this paper, the 1D-CNN block is employed in the initial stage to extract features and capture dynamic patterns from the voltage series. It is shown in Fig. 3, the 1D-CNN block is a multilayer perceptron comprising a convolutional layer, a max pooling layer, and a flatten layer.

Firstly, the convolutional layer is responsible for identifying changes and extracting features from each local input slice. To retain the size of the input after convolution remains unchanged, a zero-padding mechanism is applied at both ends of the input. Then, the kernel in the convolutional layer functions as a sliding window, successively extracting local representative features from each time series slice. These extracted features are combined to form the new global time series feature, which contains less redundant information. By convolving inputs from different perspectives using kernels of varying weights, a global feature set is ultimately constructed.

In the second step, the max pooling layer is applied to down sample each global feature into a representative form, reducing the computational complexity of the subsequent structures. Finally, the flatten layer is utilized to reshape the output generated by the max pooling layer. This step ensures that the output of 1D-CNN is properly aligned with the input dimension of the subsequent block.

3.2.2. Vanilla LSTM block

The LSTM [67] is a type of RNN architecture specifically designed to handle time series data with long-term and short-term dependencies. This paper employs an LSTM block, as shown in Fig. 4, to develop the residual LSTM in hybrid model.

LSTM operates on a multi-gated memory mechanism, which can efficiently capture and integrate temporal information from past and current time steps. The working principles of these gates are explained as

$$I_t = \sigma(w_I[h_{t-1}, x_t] + b_I) \tag{8}$$

$$F_t = \sigma(w_F[h_{t-1}, x_t] + b_F) \tag{9}$$

$$O_t = \sigma(w_O[h_{t-1}, x_t] + b_O) \tag{10}$$

where the input gate, forget gate and output gate are denoted as I_t, F_t and Q_t, respectively. Moreover, h_{t-1}, x_t, w, and b indicate the preceding hidden state, temporal information of current step, weight, and noise matrices for relevant gates, respectively. The sigmoid function σ is responsible for activating each nonlinear transformation in three gates. As depicted in Fig. 4, each gate performs a unique function. Their detailed functions are explained as follows.

$$\widetilde{M_t} = tanh(w_M[h_{t-1}, x_t] + b_M) \tag{11}$$

$$M_t = F_t M_{t-1} + I_t \widetilde{M_t} \tag{12}$$

$$h_t = O_t \tanh(M_t) \tag{13}$$

where $\widetilde{M_t}$ refers to the transient memory state that nonlinearly stores the merged information from h_{t-1} and x_t. As stated in Eq. (12), the input gate selects essential information from $\widetilde{M_t}$, while the forget gate discards insignificant information from the preceding memory state M_{t-1}. As a result, these two gates are used to form the current memory state M_t. Then, the output gate collaborates with the *tanh* function to choose the preserved hidden state h_t for the next time step, as demonstrated in Eq. (13). Finally, multiple LSTM blocks are connected sequentially to improve their capability in capturing the underlying information flow over time series.

3.2.3. Self-random attention block

The attention mechanism is a weight assignment method, ranking each element within the input sequence according to its intrinsic relevance. The subsequent blocks then prioritize the elements with higher importance. As a result, it improves the overall performance of the prediction model. However, calculating the weight factors for each element of a long sequence introduces a computational burden, impacting the prediction accuracy and real-time applications of the model. To reduce the computational complexity and enhance the performance of the self-attention block [68], a self-random attention block is proposed in this study, as illustrated in Fig. 5. The following equations

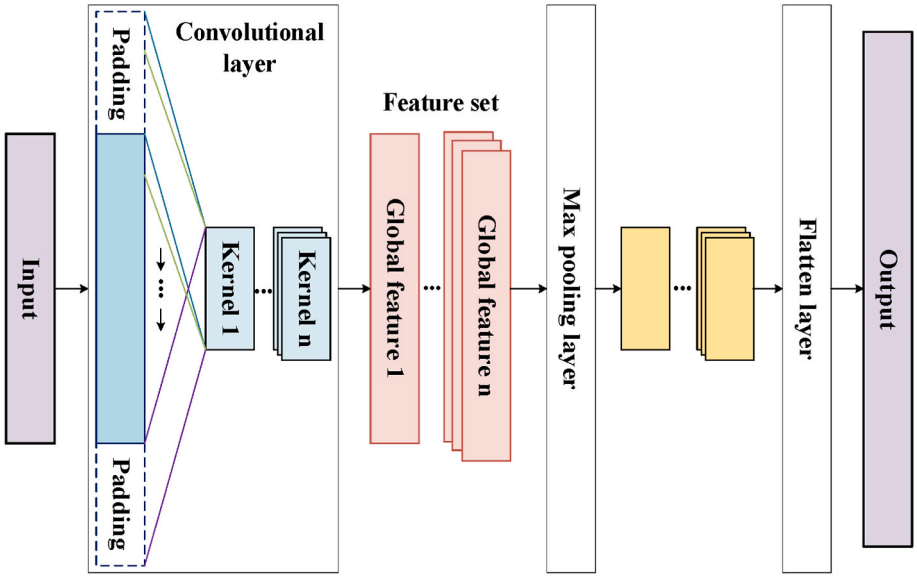

Fig. 3. Flow chart of 1D-CNN.

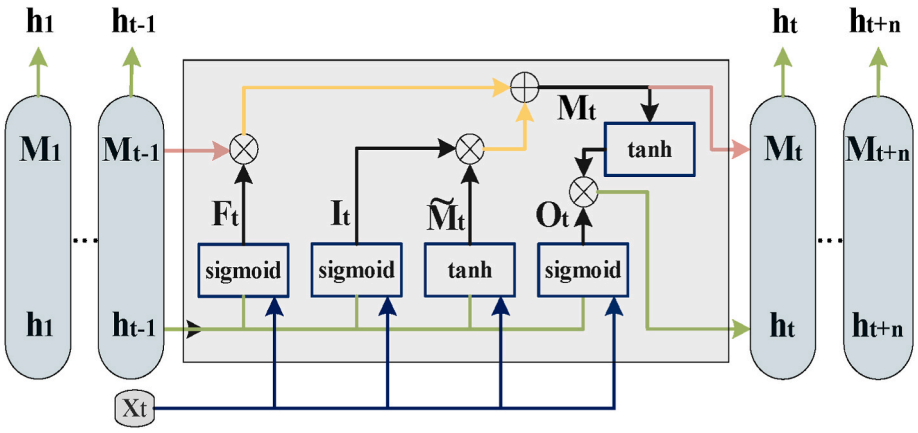

Fig. 4. Flow chart of vanilla LSTM.

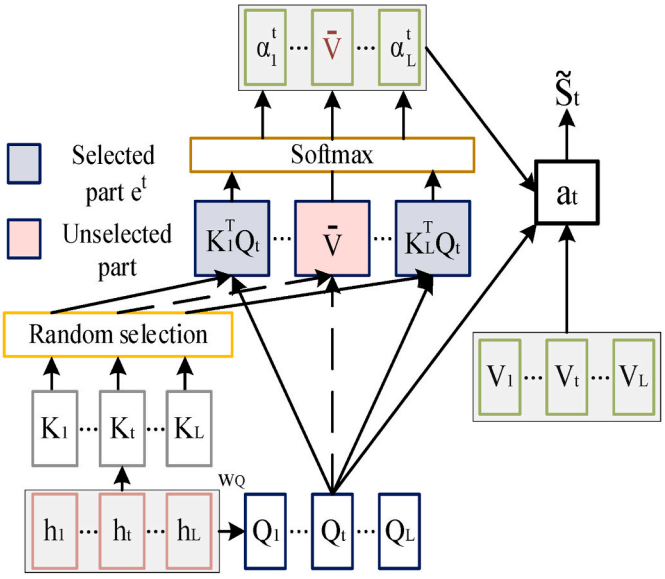

Fig. 5. Flow chart of self-random attention.

illustrate the working principles of self-random attention.

$$Q = w_Q h + b_Q \tag{14}$$

$$K = I_K h \tag{15}$$

$$V = I_V h \tag{16}$$

Firstly, the query vector Q is obtained by linearly transforming the hidden state h using the weight matrix w_Q and the bias matrix b_Q. It is noteworthy that the weight matrix w_Q is designed to be shared at each step, effectively reducing the time complexity of the block. Meanwhile, the key vector K and the value vector V are intentionally set to be the same as the hidden state to simplify the calculation. These processes are shown in Eqs. (15) and (16), where I indicates the identity matrix. The subsequent explanations will take Q_t, i.e., Q at time step t, as an example.

$$s(Q_t, K_1) = K_1^T Q_t \tag{17}$$

$$e^t = \begin{bmatrix} K_1^T Q_t & \cdots & K_m^T Q_t \end{bmatrix} \quad \text{subject to } m < L \tag{18}$$

Secondly, the inner product method is chosen, as described in Eq. (17), to compute the similarity score $s(Q_t,K_1)$ between Q_t and the first key element K_1. Similarly, the remaining similarity scores are obtained using the same method. Afterward, the random selection method is applied in selecting the components in Eq. (18), and its workflow is illustrated in Fig. 5. m and L denote the selecting length and input length, respectively. The random selection process

reduces the computational cost by excluding non-essential details since most similarity scores are close to 0. Thus, e^t contains only the scores between Q_t and the randomly selected m key elements.

$$\alpha^t = softmax(e^t) \tag{19}$$

$$a_t = \sum_{i=1}^{m} \alpha_i^t V_i + (L-m)\overline{V} \tag{20}$$

Then, a softmax function is utilized to compute the attention score set α^t. By using normalization, the valuable information in e^t is further enhanced. Thereafter, we calculate the attention value α_t for Q_t. As explained in Eq. (20), α_t is designed to include two parts. The first part is the product sum of randomly selected elements, whereas the second part is the mean of the value vector $\overline{V}$. As illustrated in Fig.6, replacing unselected elements with the mean value can enrich the information diversity, such as the information interaction between Q_t and $\overline{V}$.

$$\widetilde{s}_t = tanh\left(w_s[a_t, Q_t] + b_s\right) \tag{21}$$

where, the $\widetilde{s}_t$ is the output of the attention block. Finally, $\widetilde{s}_t$ is calculated based on attention value, query vector, weight matrix $\widetilde{s}_t$ and bias b_s.

3.2.4. Proposed residual–CNN–LSTM-random attention model

To improve the accuracy of fuel cell voltage prediction, the hybrid deep learning model is proposed. As shown in Fig. 6, the model is designed by fusing 1D-CNN, residual LSTM, random self-attention, and fully connected blocks.

Firstly, the 1D-CNN is employed to extract features from the input voltage sequence. Within the input aging voltage, short-term consecutive steps exhibit strong temporal correlations and repetitive patterns, introducing noises to the prediction model. To eliminate adverse effects of the input, the 1D-CNN operates on local time slices to extract representative features and filter out unnecessary variations. As a result, the extracted local features enhance the dynamic trend of the input without modifying its original pattern, which helps the model to accurately learn the fuel cell degradation.

Then, the residual LSTM is utilized to process these features from a time-series perspective, where it discovers internal long and short-term correlations between features extracted at different time steps and generates deep-level features. Simultaneously, the residual structure enables the concurrent transmission of initial features extracted by the 1D-CNN to interact with other features within the deep-layered block, thereby enhancing the representativeness of the fused features. It also

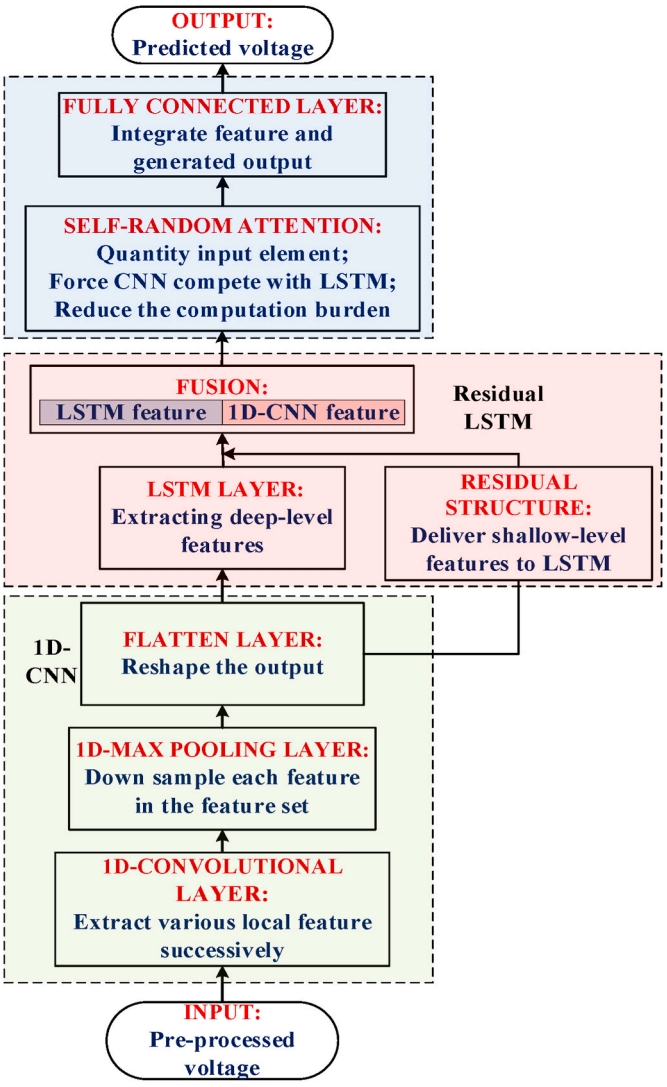

Fig. 6. Flow chart of the proposed deep learning method.

mitigates the risks associated with potential preliminary information loss and performance degradation that may arise during the deep-level feature extraction process of LSTM. However, simultaneous inputs from two sources, especially the flattened output of 1D-CNN, increase the computational complexity of the subsequent block. In addition, there may be an overlap between certain parts of the two data sources, which may impede the accurate training of the prediction model [69].

To address these issues, self-random attention is then proposed to assign weights to different elements in the integrated CNN and LSTM sequence. Consequently, the two blocks strive to extract representative information and compete to dominate the predicted results. Therefore, the information overlaps and less important features are eliminated. Moreover, the overall computational burden is reduced by simplifying matrix similarity calculations between query and key vectors, as illustrated in Fig. 5. By discarding some possible repetitive features during the random selection process, the impact of noise interference is also minimized. As a result, the combination of residual LSTM and self-random attention enhances the prediction capability of the hybrid model.

Subsequently, the fully connected block integrates all the received information. Each block of the hybrid model is updated continuously until the training requirements are met. Finally, the optimized hybrid deep learning model generates the predictive voltage. To sum up, the combination of these blocks would improve the prediction accuracy of the model.

4. Results and discussions

The prognostic framework proposed in this study is implemented using Python 3.69, where the TensorFlow 1.14 package is used to build the hybrid deep learning model. To verify the effectiveness of the prognostic framework in detecting outliers and predicting voltage, the IEEE PHM 2014 Data Challenge dataset is utilized. Firstly, the hybrid deep learning model is trained using historical data with different training stopping points (TSPs). For instance, the first 500 h of data are set as training set when TSP is 500. Then, the remaining data is used to validate the short-term and medium-term prediction accuracy of the model. The impact of different TSPs on prediction capability is also analyzed. Furthermore, stability and comparative analyses are conducted. The corresponding results prove the superior predicting accuracy and the high stability of the proposed prognostic framework.

4.1. Performance evaluation criteria

Three quantitative criteria have been selected to evaluate the performance of the proposed framework, including the accuracy (*AC*), the root mean square error (*RMSE*), and the mean absolute percentage error (*MAPE*).

$$AC = \frac{N_d}{N_o} \tag{22}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(x_t - \hat{x}_t)^2} \tag{23}$$

$$MAPE = \frac{1}{N}\left(\sum_{t=1}^{N}\frac{|x_t - \hat{x}_t|}{|x_t|}\right) \tag{24}$$

where N_o, N_d, and N represent the total number of outliers, the number of outliers correctly identified by the LOF algorithm, and the total number of predicted voltage points, respectively. Besides, x_t denotes the actual voltage, while $\hat{x}_t$ stands for the predicted voltage. Firstly, *AC* is used to quantify the effectiveness of the outlier detection algorithm in the proposed framework. The high value of *AC* indicates the strong ability of the detection algorithm. Secondly, the performance of the framework in predicting voltages is evaluated using *RMSE* and *MAPE*, where the lower values reflect higher prediction accuracy.

4.2. Performance of data pre-processing

The characterization of polarization and EIS was conducted weekly. These processes may cause the recovery phenomena of voltage degradation due to sudden stops and changes in operating conditions. As a result, the voltages in the vicinity of the event points exhibit abrupt ripples, deviating from adjacent voltage points and being identified as outliers.

The events occurred at 48, 185, 348, 515, 658, 823, 991 h and 35, 182, 343, 515, 666, 830, 1016 h for FC1 and FC2 respectively. In FC1, the voltage and its trend are relatively smooth, and thus the recovery phenomenon has a minor impact. Only the outliers with large fluctuations should be adjusted, as they have higher impacts on the predictions. Hence, the first task of data pre-processing is to automatically detect outliers near these events for FC2. Under the quasi-dynamic working condition, the presented LOF algorithm locates voltage outliers at 36, 182, 343, 506, 667 and 831 h. Compared with the experimental data, these identified points match all the actual recovery event timings, resulting in an *AC* of 1. Then, a data completion is applied to complete the dataset. Fig. 7 displays the raw and pre-processed data of FC2 as well as the raw data of FC1. Afterward, the pre-processed dataset is used to train and validate the performance of the proposed hybrid deep learning model.

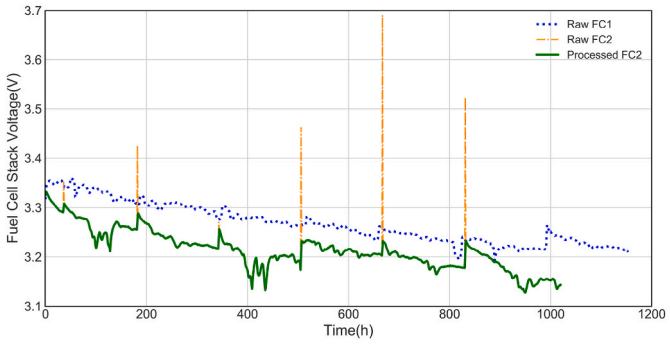

Fig. 7. FC1 and FC2 of the IEEE PHM 2014 data challenge.

4.3. Short-term degradation prediction results

The objective of short-term degradation prediction is to accurately predict short-term voltages. Meanwhile, the prediction model needs to maintain stable performances across different TSPs. Furthermore, the proposed model utilizes a look-back time step of 10 h. To evaluate the short-term prediction performances of the proposed deep learning method, it is compared with other state-of-the-art prognostic models, including LSTM, S-LSTM, BILSTM-Attention, and CNN-LSTM-Attention. The relevant results are analyzed in this section.

Firstly, FC1 of the IEEE PHM 2014 Data Challenge dataset is utilized to train and test prediction models. As shown in Fig. 8(a), all models except for the LSTM method capture the overall degradation trend of voltage under steady operating conditions with TSP equaling 500. With the increase in TSP, the LSTM and other methods all fit the degradation curves well, as illustrated in Fig. 8(b)–(c). The results reveal the influence of different TSPs on the prediction stability of the LSTM. Furthermore, Table 2 summarizes all relevant errors and comparison results obtained under steady operating conditions. It is worth noting that most methods perform the best in the case of TSP equals to 600, and the proposed hybrid deep learning model outperforms others in all TSP cases. For instance, under the condition where TSP equals 600, the RMSE and MAPE values of the proposed model are 0.0021 and 0.0323, respectively. Moreover, the proposed deep learning model achieves high-precision prediction without overfitting, as demonstrated by the training and test curves in Fig. 8(d).

To evaluate the short-term prediction performances of the proposed model under the quasi-dynamic operating condition, a quantitative analysis is conducted. FC2 of the IEEE PHM 2014 Data Challenge dataset is employed to fit and verify different prediction models. As depicted in Fig. 9(a)-(c), although the overall performances of LSTM and S-LSTM models are satisfactory, they cannot accurately capture the late dynamic changes, i.e., behaviors after 900 h. In contrast, other models all fit the overall degradation curve of FC2 well. Amongst them, the proposed model exhibits the best predictive performance without overfitting, as demonstrated by Fig. 9(d). The performances of BILSTM-Attention, CNN-LSTM-Attention, and the proposed methods are attributed to the presence of the attention block. The attention block selectively focuses on elements in the sequence, providing the model with a strong ability to capture dynamic variations. Moreover, Table 2 also presents the comparative results under the quasi-dynamic operating condition. Most models exhibit their strongest performances at TSP equaling 600 and their weakest performances at TSP equaling 500. For instance, the proposed model performs the best under the condition where TSP equals 600, with *RMSE* and *MAPE* values of 0.0023 and 0.0361, respectively.

Several conclusions can be drawn from the results under two operating conditions. Firstly, LSTM and S-LSTM show similar performance in both operating cases. This suggests that merely stacking LSTM blocks to increase the model depth does not effectively improve predictive ability. Secondly, the results of CNN-LSTM-Attention and BILSTM-Attention indicate that combining the CNN block leads to better performance than stacking the LSTM block. Thirdly, the proposed method outperforms CNN-LSTM-Attention under both operating cases. The comparative results prove that the residual structure and self-random attention mechanism are effective in enhancing the prediction accuracy.

To verify the stability of the proposed method, relevant quantitative analyses are conducted. The proposed method provides similar performances under different TSPs. To further illustrate its stability, we take the example of TSP equaling 600. Initially, the percentage error curves of the proposed model under both operating conditions are plotted. As depicted in Fig. 10(a)-(b), the errors of test data remain stable within a small range without any abrupt fluctuations, which demonstrates the high stability of the model. Additionally, the prediction confidence

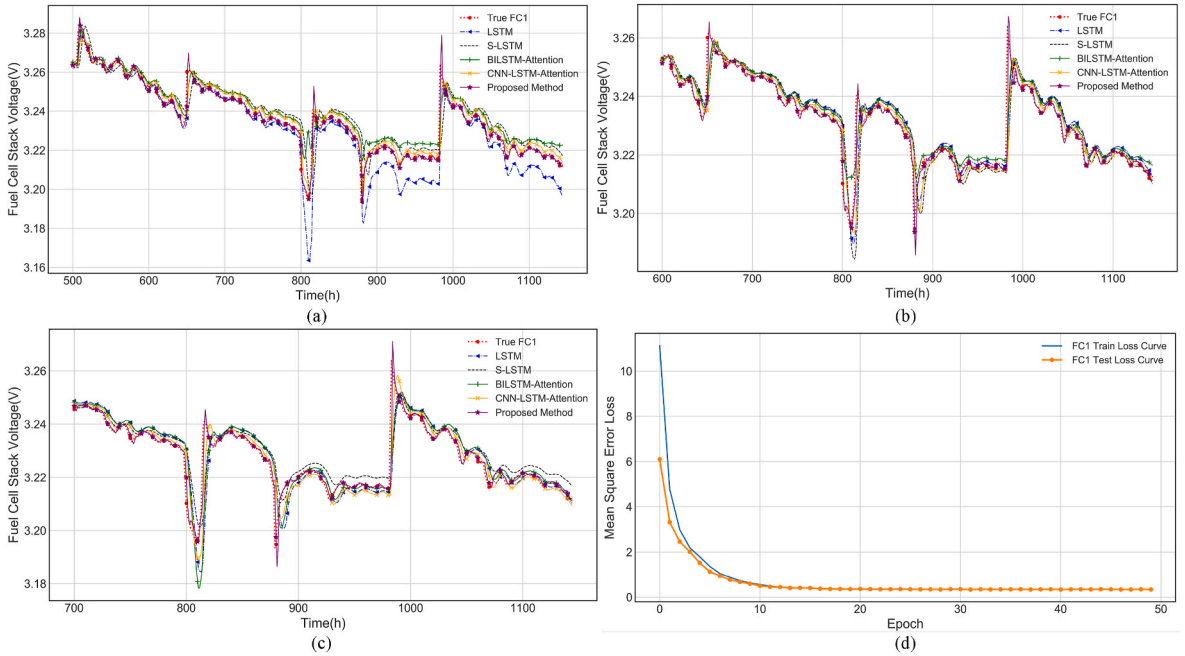

Fig. 8. Performances of different methods on FC1 (a) TSP = 500 (b) TSP = 600 (c) TSP = 700 (d) Training and test curves when TSP = 600.

Table 2
Comparison results of short-term prediction for various methods under different conditions.

Models	Steady (FC1)						Quasi-dynamic (FC2)					
	TSP = 500		TSP = 600		TSP = 700		TSP = 500		TSP = 600		TSP = 700	
	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
LSTM	0.0085	0.1688	0.0057	0.1034	0.0071	0.1217	0.0143	0.2421	0.0079	0.1514	0.0118	0.2282
S-LSTM	0.0078	0.1758	0.0071	0.1237	0.0063	0.1211	0.0074	0.1537	0.0070	0.1400	0.0079	0.1713
BILSTM-Attention	0.0064	0.1525	0.0049	0.0934	0.0066	0.1136	0.0048	0.0927	0.0044	0.0875	0.0049	0.0995
CNN-LSTM-Attention	0.0046	0.0991	0.0058	0.0880	0.0053	0.0873	0.0038	0.0807	0.0034	0.0625	0.0033	0.0543
Proposed model	0.0025	0.0356	0.0021	0.0323	0.0027	0.0329	0.0025	0.0395	0.0023	0.0361	0.0026	0.0437

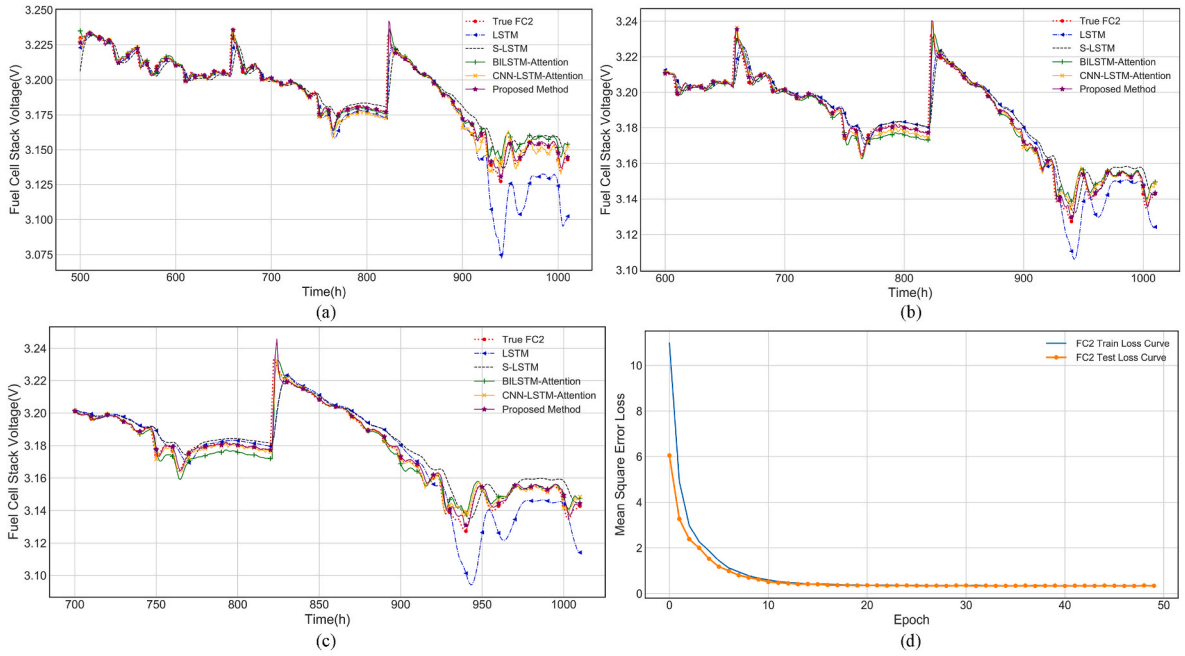

Fig. 9. Performances of different methods on FC2 (a) TSP = 500 (b) TSP = 600 (c) TSP = 700. (d) Training and test curves when TSP = 600.

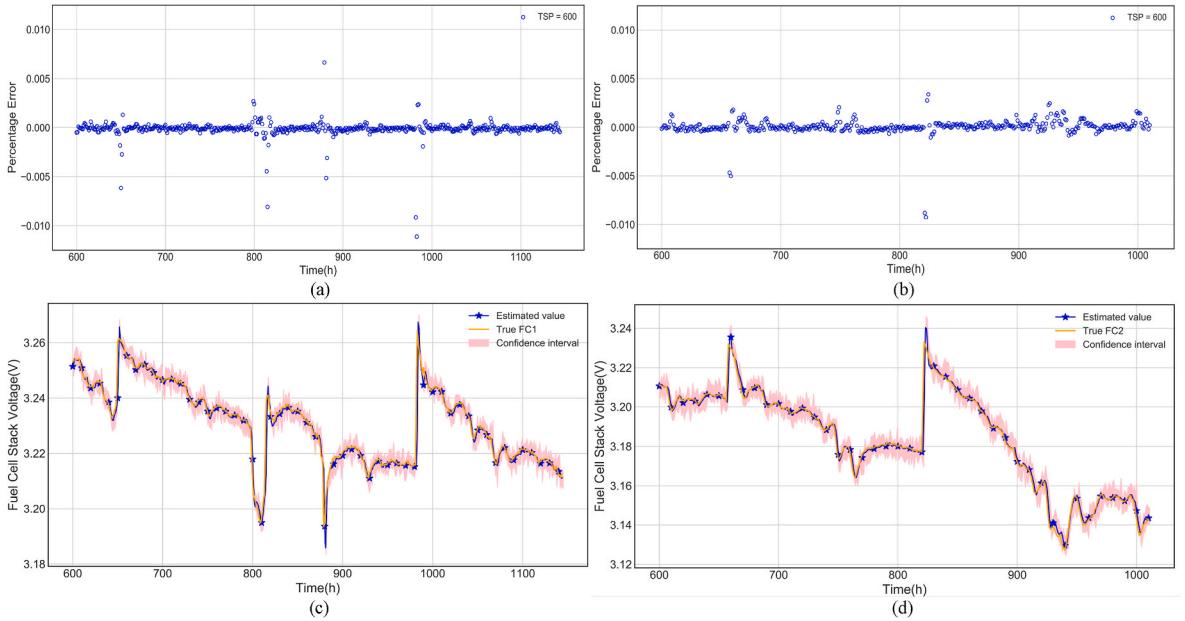

Fig. 10. Short-term prediction stability analysis when TSP = 600 (a) Estimation percentage error of FC1 (b) Estimation percentage error of FC2 (c) Prediction interval of FC1 (d) Prediction interval of FC2.

intervals of the proposed method are illustrated in Fig. 10(c)–(d). The prediction confidence interval is generated by adjusting the model's hyperparameters and fusing its lowest and highest performance limits. The small prediction confidence intervals for both operating conditions, as depicted in Fig. 10(c)–(d), provide further evidence supporting the high stability of the proposed method. To sum up, the proposed model demonstrates superior short-term prediction capability with a high level of stability.

4.4. Medium-term degradation prediction results

This section evaluates the performance of the proposed deep learning method on medium-term prediction. In this study, the medium-term prediction is realized by applying iterative short-term predictions. Therefore, accurate short-term prediction is a prerequisite for a satisfactory medium-term prediction. As a result, we only compare BILSTM-Attention, CNN-LSTM-Attention, and the proposed method, which perform the best in short-term predictions. For medium-term prediction, the look-back timestep is also set at 10 h.

Fig. 11(a) shows the prediction curves of different methods under steady operating conditions with TSP equaling 600. As illustrated in the figure, the proposed method outperforms the others in capturing the trend of the degradation curve. As demonstrated in Fig. 11(b), BILSTM-Attention and CNN-LSTM-Attention models cannot accurately capture the variations in the degradation curve and exhibit certain oscillations under the quasi-dynamic case. In contrast, the proposed method outperforms other methods. Despite some differences between the actual and predicted values, the proposed model fits the overall trend well.

Furthermore, Table 3 presents the comparison results of medium-term predictions. It is worth noting that an increase in TSP leads to a decrease in both the *RMSE* and *MAPE* values in FC1, while only *MAPE* values decrease in FC2. The results also indicate that all models perform better under the steady operating condition compared to the results under the quasi-dynamic operating condition. Additionally, the proposed model shows the best performance with the minimum RMSE and MAPE values in all cases. This further emphasizes the importance of the residual structure and the self-random attention mechanism in enhancing prediction precisions.

Moreover, the low prediction confidence intervals presented in Fig. 11(c)–(d) verify the high stability of the proposed model in medium-term prediction. In conclusion, the proposed model demonstrates higher medium-term prediction accuracies than others, while simultaneously maintaining a high level of stability.

4.5. Working zone

The study [70] defines the working zone of a PEMFC as the zone where the proposed method can provide reliable performance under certain working conditions. Identifying the working zone is crucial, as it can be used to evaluate the robustness of the proposed method for different application scenarios. In our research, the combinations of prediction intervals and operating states are applied. Besides, the RMSE is the criteria that used to select the working zone.

In terms of the short-term prediction [21], prediction with a RMSE lower than 0.012 is preferable. As illustrated in Table 4, all the RMSMs are lower than 0.012 for both working conditions, where the RMSEs for quasi-dynamic state are slightly higher than those for steady state. Therefore, the proposed work performs well for all the combinations within the working zone.

For medium-term prediction [21], it is important to have a robust model with an RMSE lower than 0.04. Table 4 shows that all predictive performances under steady state fulfil the criteria. In contrast, only the predictive performances under quasi-dynamic state with a prediction interval less than 80 h meet the requirement. To sum up, all short-term and medium-term predictions that meet the accuracy criteria form the working zone, which can be used to guide the proposed work for various applications.

4.6. Time complexity analysis

Fine-tuning is necessary for the degradation prognostic applications, which require high prediction accuracy. Previous results demonstrate that models with the attention block, particularly the proposed method, outperform other models in prediction accuracy. In addition, the time

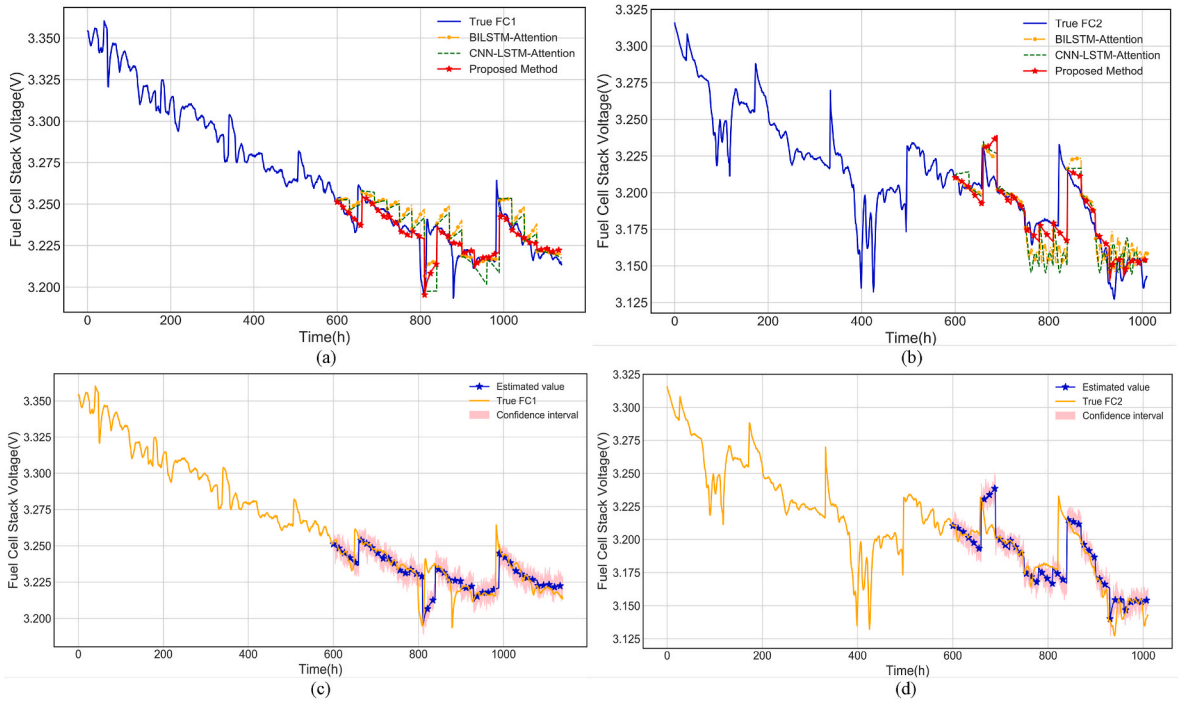

Fig. 11. Medium-term prediction performances of different methods when TSP = 600 (a) Estimation curve of FC1 (b) Estimation curve of FC2 (c) Prediction interval of FC1 (d) Prediction interval of FC2.

Table 3
Comparison results of medium-term prediction for various methods under different conditions.

Models	Steady (FC1)				Quasi-dynamic (FC2)			
	TSP = 600		TSP = 700		TSP = 600		TSP = 700	
	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
BILSTM-Attention	0.0127	0.5285	0.0123	0.4119	0.0189	0.9255	0.0229	0.8188
CNN-LSTM-Attention	0.0139	0.5647	0.0126	0.4133	0.0199	0.9567	0.0227	0.7993
Proposed model	0.0085	0.4237	0.0082	0.3410	0.0123	0.7957	0.0147	0.7319

Table 4
Comparison results of different prediction intervals.

Prediction performance (TSP = 600)	Steady (FC1)						Quasi-dynamic (FC2)					
	Prediction interval (hour)						Prediction interval (hour)					
	1	2	4	8	16	24	1	2	4	8	16	24
RMSE of short-term prediction	0.0021	0.0032	0.0039	0.0043	0.0071	0.0079	0.0023	0.0028	0.0049	0.0071	0.0095	0.0113
	Prediction interval (hour)						Prediction interval (hour)					
	30	40	50	80	120	160	30	40	50	80	120	160
RMSE of medium-term prediction	0.0081	0.0085	0.0093	0.0097	0.0103	0.0122	0.0119	0.0123	0.0149	0.0397	0.0851	0.0951

Table 5
Time complexity comparison results of various methods.

Models	BILSTM-Attention	CNN-LSTM-Attention	Proposed model
Structure:			
First layer	LSTM: Process elements sequentially with medium time complexity [71].	1D-CNN: Process elements parallelly with low time complexity [71].	1D-CNN: Process elements parallelly with low time complexity.
Second layer	LSTM: Process elements sequentially with medium time complexity.	LSTM: Process elements sequentially with medium time complexity.	LSTM: Process elements sequentially with medium time complexity.
Third layer	Self-attention: High time complexity.	Self-attention: High time complexity.	Self-random attention: Simplify similarity calculation with medium time complexity.
Experiment:			
Time Complexity per epoch	1 s	Less than 1 s	Less than 1 s
Number of epochs	Around 100	Around 100	Around 70
Overall time complexity	Moderate	Medium	Low
Accuracy	Moderately Limited	Medium	High
Application	Limited applicability	Medium applicability	High applicability

complexity is another crucial factor for field applications. This section compares the time complexity of the proposed method with BILSTM-Attention and CNN-LSTM-Attention models.

As illustrated in Table 5, the time complexity of each layer of the three models is analyzed. Then, a comparative experiment is conducted to analyze the time complexity of the entire models in the application. The models are tested on FC1 and FC2 using the hardware environment of the i7-6500U CPU. The results clearly show that the training time per epoch of BILSTM-Attention, CNN-LSTM-Attention, and the proposed model are 1 s, less than 1 s, and less than 1 s, respectively. Besides, the proposed method requires a minimum number of epochs to achieve convergence. As a result, the proposed model outperforms others in terms of time complexity. Therefore, our model with higher prediction accuracy and lower time complexity is the preferred option for field applications.

5. conclusion

This paper proposes a data-driven prognostic framework for fuel cell degradation prediction, which consists of a data pre-processing step and a hybrid deep learning model. Firstly, the data pre-processing step is responsible for detecting and handling the outliers. Then, the hybrid deep learning model is proposed based on 1D-CNN, residual LSTM, self-random attention, and fully connected blocks. It extracts representative information, processes features, evaluates the importance of received elements as well as predicts voltages. The experimental results demonstrate the improved short-term and medium-term predicting accuracies achieved by this prognostic method. Meanwhile, the method has a high level of stability and low time complexity, illustrating its suitability for field implementation. The results indicate the benefits of the method for the development of PEMFC health management system.

In future research, certain aspects require further improvements before applying in the field. Firstly, the performance of the proposed framework has not been tested under more dynamic working conditions. In such cases, the degradation mechanism becomes more complex, leading to a less stable prediction of the degradation trend. To this end, we plan to conduct more degradation experiments and enhance the structure of the proposed framework to improve its predictive capabilities under various working conditions. In addition, the online learning of the prognostic framework will be explored to enhance its real-time application. The quick training response of the framework enables the implementation. Finally, the proposed work is planned to be integrated into health management system to optimize the durability of PEMFC for future research.

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.

References

[1] Yang Z, Ghadamyari M, Khorramdel H, Alizadeh SMS, Pirouzi S, Milani M, et al. Robust multi-objective optimal design of islanded hybrid system with renewable and diesel sources/stationary and mobile energy storage systems. Renew Sustain Energy Rev 2021;148:111295.

[2] Mehrpooya M, Ghadimi N, Marefati M, Ghorbanian SA. Numerical investigation of a new combined energy system includes parabolic dish solar collector, Stirling engine and thermoelectric device. Int J Energy Res 2021;45:16436–55.

[3] Ye H, Jin G, Fei W, Ghadimi N. High step-up interleaved dc/dc converter with high efficiency. Energy Sources Part A; 2019.

[4] Jarry T, Jaafar A, Turpin C, Lacressonniere F, Bru E, Rallieres O, et al. Impact of high frequency current ripples on the degradation of high-temperature PEM fuel cells (HT-PEMFC). Int J Hydrogen Energy 2023;48(54):20734–42.

[5] Park D, Ham S, Sohn Y, Choi Y, Kim M. Mass transfer characteristics according to flow field and gas diffusion layer of a PEMFC metallic bipolar plate for stationary applications. Int J Hydrogen Energy 2023;48(1):304–17.

[6] Yue M, Jemei S, Zerhouni N, Gouriveau R. Proton exchange membrane fuel cell system prognostics and decision-making: current status and perspectives. Renew Energy 2021;179:2277–94.

[7] Liu J, Chen C, Liu Z, Jermsittiparsert K, Ghadimi N. An IGDT-based risk-involved optimal bidding strategy for hydrogen storage-based intelligent parking lot of electric vehicles. J Energy Storage 2020;27:101057.

[8] Kannan R, Sundharajan V. A novel MPPT controller based PEMFC system for electric vehicle applications with interleaved SEPIC converter. Int J Hydrogen Energy 2023;48(38):14391–405.

[9] Dehghani M, Ghiasi M, Niknam T, Kavousi-Fard A, Shasadeghi M, Ghadimi N, et al. Blockchain-based securing of data exchange in a power transmission system considering congestion management and social welfare. Sustain Times 2021;13(1):90.

[10] Kongkanand A, Masten D, Wang C, Chen J, Spendelow J, James B, et al. Automotive fuel cell targets and status. In: Tech rep. Washington DC, United States: U.S. DOE; Aug 2020, 20005.

[11] Sahajpal K, Sahajpal K, Kumar V. Accurate long-term prognostics of proton exchange membrane fuel cells using recurrent and convolutional neural networks. Int J Hydrogen Energy 2023;48(78):30532–55.

[12] Guida M, Postiglione F, Pulcini G. A random-effects model for long-term degradation analysis of solid oxide fuel cells. Reliab Eng Syst Saf 2015;140:88–98.

[13] Lee J, Lee J-H, Choi W, Park K-W, Sun H-Y, Oh J-H. Development of a method to estimate the lifespan of proton exchange membrane fuel cell using electrochemical impedance spectroscopy. J Power Sources 2010;195:6001–7.

[14] Dhirde AM, Dale NV, Salehfar H, Mann MD, Han TH. Equivalent electric circuit modeling and performance analysis of a PEM fuel cell stack using impedance spectroscopy. IEEE Trans Energy Convers 2010;25:778–86.

[15] Lechartier E, Laffly E, P_era M-C, Gouriveau R, Hissel D, Zerhouni N. Proton exchange membrane fuel cell behavioral model suitable for prognostics. Int J Hydrogen Energy 2015;40(26):8384–97.

[16] Onanena R, Oukhellou L, Candusso D, Harel F, Hissel D, Aknin P. Fuel cells static and dynamic characterizations as tools for the estimation of their ageing time. Int J Hydrogen Energy 2011;36(2):1730–9.

[17] Lechartier E, Laffly E, Péra M-C, Gouriveau R, Hissel D Zerhouni N. Proton exchange membrane fuel cell behavioral model suitable for prognostics. Int J Hydrogen Energy 2015;40(26):8384–97.

[18] Kim T, Kim H, Ha J, Kim K, Youn J, Jung J, et al. A degenerated equivalent circuit model and hybrid prediction for state-of-health (SOH) of PEM fuel cell. In: 2014 int. Conf. Progn. Heal. Manag. PHM 2014; 2015. p. 1–7.

[19] 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.

[20] Zhang D, Cadet C, Yousfi-Steiner N, Bérenguer C. Proton exchange membrane fuel cell remaining useful life prognostics considering degradation recovery phenomena. Proc Inst Mech Eng Part O J Risk Reliability 2018;232(4):415–24.

[21] Hua Z, Zheng Z, Pahon E, Péra M-C, Gao F. A review on lifetime prediction of proton exchange membrane fuel cells system. J Power Sources 2022;529:231–56.

[22] 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.

[23] 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;44:1930–9.

[24] 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.

[25] Hu Z, Xu L, Li J, Ouyang M, Song Z, Huang H. A reconstructed fuel cell life-prediction model for a fuel cell hybrid city bus. Energy Convers Manag 2018;156:723–32.

[26] Kimotho JK, Meyer T, Sextro W. PEM fuel cell prognostics using particle filter with model parameter adaptation. In: 2014 int. Conf. Progn. Heal. Manag. PHM 2014; 2015. p. 1–6.

[27] Zhang X, Pisu P. An unscented Kalman filter based approach for the health-monitoring and prognostics of a polymer electrolyte membrane fuel cell. In: 2012 int. Conf. Progn. Heal. Manag. PHM. vol. 3; 2012. p. 1–9.

[28] Robin C, Gerard M, Franco AA, Schott P. Multi-scale coupling between two dynamical models for PEMFC aging prediction. Int J Hydrogen Energy 2013;38(11):4675–88.

[29] Chandesris M, Vincent R, Guetaz L, Roch J-S, Thoby D, Quinaud M. Membrane degradation in PEM fuel cells: from experimental results to semi-empirical degradation laws. Int J Hydrogen Energy 2017;42(12):8139–49.

[30] Silva RE, Gouriveau R, Jemeï 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(21):11128–44.

[31] Wang C, Li Z, Outbib R, Dou M, Zhao D. A novel long short-term memory networks-based data-driven prognostic strategy for proton exchange membrane fuel cells. Int J Hydrogen Energy 2022;47(18):10395–408.

[32] Ibrahim M, Steiner NY, Jemei S, Hissel D. Wavelet-based approach for online fuel cell remaining useful lifetime prediction. IEEE Trans Ind Electron 2016;63(8):5057–68.

[33] Ma R, Yang T, Breaz E, Li Z, Brios P, Gao F. Data-driven proton exchange membrane fuel cell degradation predication through deep learning method. Appl Energy 2018;231:102–15.

[34] Liu H, Chen J, Hissel D, Su H. Short-term prognostics of PEM fuel cells: a comparative and improvement study. IEEE Trans Ind Electron 2019;66(8):6077–86.

[35] 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(4):1570–82.

[36] Liu H, Chen J, Hou M, Shao Z, Su H. Data-based short-term prognostics for proton exchange membrane fuel cells. Int J Hydrogen Energy 2017;42:20791–808.

[37] Javed K, Gouriveau R, Zerhouni N, Hissel D. PEM fuel cell prognostics under variable load: a data-driven ensemble with new incremental learning. In: 2016 int. Conf. Control, decis. Inf. Technol; 2016. p. 252–7.

[38] He K, Zhang C, He Q, Wu Q, Jackson L, Mao L. Effectiveness of PEMFC historical state and operating mode in PEMFC prognosis. Int J Hydrogen Energy 2020;45(56):32355–66.

[39] Liu J, Li Q, Han Y, Zhang G, Meng X, Yu J, et al. PEMFC residual life prediction using sparse autoencoder-based deep neural network. IEEE Trans Transp Electrificat 2019;5(4):1279–93.

[40] Gers FA, Schraudolph NN, Schmidhuber J. Learning precise timing with LSTM recurrent networks. J Mach Learn Res 2002;3:115–43.

[41] Mezzi R, Yousfi-Steiner N, Péra MC, Hissel D, Larger L. An Echo State Network for fuel cell lifetime prediction under a dynamic micro-cogeneration load profile. Appl Energy 2021;283:116297.

[42] 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.

[43] Deng Z, Chen Q, Zhang L, Zhou K, Zong Y, Liu H, et al. Degradation prediction of PEMFCs using stacked echo state network based on genetic algorithm optimization. IEEE Trans Transp Electrificat 2022;8(1):1454–66.

[44] 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.

[45] Hua Z, Zheng Z, Péra M, Gao F. Data-driven prognostics for PEMFC systems by different echo state network prediction structures. In: Proc. ITEC IEEE transp. electrif. Conf. Expo; 2020. p. 495–500.

[46] Liu Z, Xu S, Zhao H, Wang Y. Durability estimation and short-term voltage degradation forecasting of vehicle PEMFC system: development and evaluation of machine learning models. Appl Energy 2022;326:119975.

[47] Xie Y, Zou J, Li Z, Gao F, Peng C. A novel deep belief network and extreme learning machine based performance degradation prediction method for proton exchange membrane fuel cell. IEEE Access 2020;8:176661–75.

[48] Ma J, Liu X, Zou X, Yue M, Shang P, Kang L, et al. Degradation prognosis for proton exchange membrane fuel cell based on hybrid transfer learning and intercell differences. ISA Trans 2021;113:149–65.

[49] Wang F-K, Cheng X-B, Hsiao K-C. Stacked long short-term memory model for proton exchange membrane fuel cell systems degradation. J Power Sources 2020;448:227591.

[50] Chen K, Laghrouche S, Djerdir A. Degradation prediction of proton exchange membrane fuel cell based on grey neural network model and particle swarm optimization. Energy Convers Manag 2019;195:810–8.

[51] Zhang S, Chen T, Xiao F, Zhang R. Degradation prediction model of PEMFC based on multi-reservoir echo state network with mini reservoir. Int J Hydrogen Energy 2022;47(94):40026–40.

[52] Wang T, Zhou H, Zhu C. A short-term and long-term prognostic method for PEM fuel cells based on Gaussian process regression. Energies 2022;15(13):4844.

[53] 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: 2018 IEEE transp. electrif. Conf. Expo; 2018. p. 102–7.

[54] Wang F-K, Mamo T, Cheng X-B. 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.

[55] 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;179:802–14.

[56] Chen K, Laghrouche S, Djerdir A. Aging prognosis model of proton exchange membrane fuel cell in different operating conditions. Int J Hydrogen Energy 2020;45(20):11761–72.

[57] Long B, Wu K, Li P, Li M. A novel remaining useful life prediction method for hydrogen fuel cells based on the gated recurrent unit neural network. Appl Sci 2022;12(1):432.

[58] Ma R, Li Z, Breaz E, Liu C, Bai H, Briois P, et al. Data-fusion prognostics of proton exchange membrane fuel cell degradation. IEEE Trans Ind Appl 2019;55(4):4321–31.
[59] Jin J, Chen Y, Xie C, Zhu W, Wu F. Remaining useful life prediction of PEMFC based on cycle reservoir with jump model. Int J Hydrogen Energy 2021;46(80):40001–13.
[60] Sun X, Xie M, Fu J, Zhou F, Liu J. An improved neural network model for predicting the remaining useful life of proton exchange membrane fuel cells. Int J Hydrogen Energy 2023;48(65):25499–511.
[61] Li S, Luan W, Wang C, Chen Y, Zhuang Z. Degradation prediction of proton exchange membrane fuel cell based on Bi-LSTM-GRU and ESN fusion prognostic framework. Int J Hydrogen Energy 2022;47(78):33466–78.
[62] Tang X, Qin X, Wei K, Xu S. A novel online degradation model for proton exchange membrane fuel cell based on online transfer learning. Int J Hydrogen Energy 2023;48(36):13617–32.
[63] Mao L, Jackso L. IEEE 2014 data challenge data. Loughborough University. Dataset; 2016.
[64] Yang Z, Yang D, Dyer C, He X, Smola A, Hovy E. Hierarchical attention networks for document classification. In: 2016 conf. Of the N. Am. Chap. Of the assoc. For comput. Linguist. Hum. Lang. Tech; 2016. p. 1480–9.
[65] Breunig MM, Kriegel H-P, Ng RT, Sander J. LOF: identifying density-based local outliers. ACM SIGMOID Record 2000;29(2):93–104.
[66] Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998;86(11):2278–324.
[67] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9(8):1735–80.
[68] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. Adv Neural Inf Process Syst 2017:6000–10.
[69] Cho K, Merriënboer BV, Bahdanau D, Bengio Y. On the properties of neural machine translation: encoder-decoder approaches. arXiv preprint; 2014. arXiv:1409.1259.
[70] Zou W, Froning D, Shi Y, Froning W. Working zone for a least-squares support vector machine for modeling polymer electrolyte fuel cell voltage. Appl Energy 2021;283(1):116191.
[71] Garg S, Krishnamurthi R. A CNN encoder decoder LSTM model for sustainable wind power predictive analytics. Sustai Comput Inform Syst 2023;38:100869.