



HAL
open science

A novel long short-term memory networks-based data-driven prognostic strategy for proton exchange membrane fuel cells

Chu Wang, Zhongliang Li, Rachid Outbib, Manfeng Dou, Dongdong Zhao

► To cite this version:

Chu Wang, Zhongliang Li, Rachid Outbib, Manfeng Dou, Dongdong Zhao. A novel long short-term memory networks-based data-driven prognostic strategy for proton exchange membrane fuel cells. *International Journal of Hydrogen Energy*, 2022, 47 (18), pp.10395-10408. 10.1016/j.ijhydene.2022.01.121 . hal-03581357

HAL Id: hal-03581357

<https://hal.science/hal-03581357>

Submitted on 19 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Novel Long Short-term Memory Networks-based Data-driven Prognostic Strategy for Proton Exchange Membrane Fuel Cells

Chu Wang^{a, b, *}, Zhongliang Li^{a, **}, Rachid Outbib^a, Manfeng Dou^b and Dongdong Zhao^b

^a LIS Lab (UMR CNRS 7020), Aix-Marseille University, 13397 Marseille, France

^b School of Automation, Northwestern Polytechnical University, Xi'an 710072, China

* Corresponding author: Chu Wang (e-mail: chu.wang@etu.univ-amu.fr)

** Corresponding author: Zhongliang Li (e-mail: zhongliang.li@lis-lab.fr)

Abstract

Proton exchange membrane fuel cell (PEMFC) long-term prognostic facilitates reducing the time/cost of the durability tests and is a critical starting point for control/maintenance suggestions. Long short-term memory (LSTM) recurrent neural networks have excellent time series processing capabilities and are proved to be useful for the short-term prognostic of PEMFC. However, LSTM prognostic models usually suffer from accumulated errors and model recognition uncertainties, which make it difficult to break the historical degradation data limitations, resulting in unsatisfactory long-term prediction performance. To tackle the problem, this paper proposes a novel model named navigation sequence driven LSTM (NSD-LSTM) for long-term prognostic. In the strategy, a navigation sequence is firstly generated by using an autoregressive integrated moving average model with exogenous variables. The sequence is then fed iteratively into LSTM in the implementation stage to achieve long-term prediction. The proposed strategy is evaluated using the aging experimental data of two types of PEMFC under different operating conditions. The long-term prognostic performance of the proposed model and other two state-of-the-art prognostic models, namely, nonlinear autoregressive exogenous and echo state network, are evaluated through comparison experiments. The simulation and experimental results show that the proposed prognostic strategy has better long-term degradation trend prediction consistency and remaining useful life estimation robustness.

Keywords:

Long-term prognostic;

Navigation sequence driven long short-term memory networks;

Proton exchange membrane fuel cell;

Robustness;

Remaining useful life

1 Introduction

Fuel cell (FC) is considered as an energy conversion system that has received widespread attention. Compared with traditional fossil fuels, the large-scale use of hydrogen, one of the most commonly used fuel of FC, has the potential to greatly reduce environmental pollution. At present, among different types, proton exchange membrane fuel cell (PEMFC) is one of the most widely used ones [1]. In particular, PEMFC can convert chemical energy into electrical energy without generating greenhouse gases during work, and also has a suitable operating temperature range and response characteristics. This makes PEMFC suitable for the use of transportation represented by electric vehicles [2]. However, for PEMFC used in electric vehicles, many factors affect its degradation in the process of energy conversion [5]. At the same time, electric vehicles often have to face different operating conditions and loads, which results in the remaining useful life (RUL) of PEMFC often failing to meet the demand [3,4]. Nowadays, and technically speaking, the target is to double the average fuel cell lifetime estimated presently about 13,000 hours [5]. In order to better understand the degradation mechanisms of PEMFC, and shorten the time/cost of durability tests, it is essential to estimate the degradation trend and RUL as accurately as possible.

In recent years, three types of PEMFC prognostic strategies [6-20,24-33] have been proposed: model-based, data-based, and hybrid methods based on model and data fusion. Among them, the model-based strategy describes the degradation processes by constructing a physical model [6-8], which is beneficial to the prediction accuracy in the condition that the physical degradation model is sufficiently accurate. The authors in [6], proposed a multi-agent prediction method based on particle filter (PF) for three parallel FC stacks. In this work, only a simple degradation model is considered, and the RULs of three FC stacks were discussed without analyzing the real degradation trend. In [7], the authors proposed a reconstructed prognostic model for FC city buses. The FC voltage deviation model was built by analyzing different operating conditions for the city bus. However, only the short-term prediction (less than one hour) was analyzed in this work, and apparently, RUL was not estimated. The authors of [8] proposed a prognostic model based on the adaptive unscented Kalman filter for PEMFC. This model estimated the health state and the RUL by improving the initial parameter setting of conventional unscented Kalman filter. The same defect as in [6] is that the prediction of the degradation trend was not considered.

It has been widely known that the FC degradation process is complicated and complex. It is usually difficult to establish an analytical degradation model with sufficient fidelity. Data-based methods and hybrid methods have therefore been attracting wider attention compared with model-based ones. In literature [9], Javed et al. proposed a Summation Wavelet-Extreme Learning Machine (SW-ELM) algorithm, which performed long-term prognostic and RUL estimation. It is worth mentioning that the authors discussed the

robustness and applicability of the proposed method, which has become a problem faced by most prognostic strategies, including the one proposed in the same paper. Literatures [10,11,12] have used the echo state network (ESN) framework for prognostic. Among them, Li et al. [11,12] conducted a long-term aging experiment with a compact fuel cell system and proposed a prognostic strategy combining the Linear Parameter Varying (LPV) model and ESN. This method extracts the health indicator (HI) of the fuel cell under dynamic operating conditions and further predicts the degradation trend and RUL. In [13], Jouin et al. proposed a hybrid method based on a particle filtering framework for behavior prediction and RUL construction. As in literature [11,12], this method pays more attention to the long-term change trend rather than to the capture of local details.

Inspired by the recent remarkable advances in deep neural network (DNN), the special attention has been put on the development of data-based prognostic tools by configuring and training different DNN structures. Among different DNN structures, long short-term memory (LSTM), a recurrent neural network (RNN) paradigm, has been considered as a potentially effective tool to handle time series prediction problems [14-20,24-33]. Several recent studies have been proposed to explore the LSTM application in fuel cell prognostics. In these studies, the LSTM is deployed for either single-step or multi-step prediction (realized by implementing single-step prediction iteratively). Meanwhile, the degradation trend can be effectively predicted only for a short horizon of a few hours [14-18,30-33]. To achieve RUL estimation, multi-step prediction with long prediction horizon is required, particularly for the prognostics of aging degradation [19,20,24-29]. However, several issues arise as the prediction horizon is extended.

- The prediction error is accumulated when well-trained LSTM models are fed with the output in previous-step prediction iteratively.
- The FC degradation behavior is disturbed and influenced by the time-varying operations and recoverable faults, which poses difficulties for tracing the intrinsic degradation trend.
- The prognostic models built only on the historical data usually suffer from the model epistemic uncertainty. The predicted results are hardly beyond the numerical interval of the historical-based training set.

This work is devoted to tackling the afore-mentioned issues and make LSTM model more appropriate for long-term prognostic use. Basically, the above issues are due to the lack of reliable data or information in the prediction zone. As one potential solution, we consider that an approximate trend sequence, inherited from the training data, can be pre-generated for the prediction zone. The pre-generated sequence can then be used to guide the LSTM model in order to avoid the afore-mentioned undesired predictive behaviors. Specifically, in this work, an autoregressive integrated moving average model with exogenous variables (ARIMAX) is

proposed to generate a navigation sequence (NS) to guide LSTM for multi-step prediction. As a consequence, a novel navigation sequence driven LSTM (NSD-LSTM) prognostic strategy design is proposed. The strategy is then tested using different fuel cell long-term aging datasets, including those obtained in both static and dynamic operating conditions. The enhanced multi-step prediction capability and RUL estimation performance are then highlighted by comparing the proposed NSD-LSTM with two state-of-the-art data-driven prognostic models, namely, nonlinear autoregressive exogenous (NARX) [34] and ESN.

The organization of the remaining paper is as follows. In Section 2, the principles of LSTM and ARIMAX are briefly introduced. The prognostic strategy of NSD-LSTM is also presented in the same section. In Section 3, FC aging experiments and data preprocessing are introduced. Then, the proposed prognostic strategy is tested and compared with NARX and ESN models in Section 4. Finally, the work is concluded in Section 5.

2 NSD-LSTM based prognostic strategy

2.1 ARIMAX model

The ARIMAX model is one of the most commonly used time series analysis models [21,22], which can be used to explain the relationship between system variables. In literature [22], the author applies several artificial intelligence models and ARIMAX to predict time series. Compared with other state-of-the-art models, thanks to the addition of exogenous sequences (XS), the ARIMAX model is considered to be a useful method for predicting time series.

Assuming that k input variable sequences ($\{x_{1t}\}, \{x_{2t}\}, \dots, \{x_{kt}\}$), including the XS and the historical HI sequence, are stationary, where $t \in \mathbb{R}^l$ corresponds to l time steps. The regression model between the output variable sequence $\{y_t\}$ and the input variable sequences is as follows:

$$y_t = \mu + \sum_{i=1}^k [\theta_i(B)/\phi_i(B)] B^{L_i} x_{it} + \varepsilon_t \quad (1)$$

where μ is the mean vector of $\{y_t\}$, and $\{\varepsilon_t\}$ represents regression residual sequence. $\{x_{it}\}$ is the i -th ($i = 1, 2, \dots, k$) input variable sequence. L_i represents the i -th lag degree. B is backshift operator such as $B^{L_i} x_{it} = x_{it-L_i}$. $\theta_i(B) = \theta_0^i - \sum_{j=1}^q \theta_j^i B^j$ denotes q -order moving average polynomial of $\{x_{it}\}$, and $\phi_i(B) = \phi_0^i - \sum_{j=1}^p \phi_j^i B^j$ denotes p -order autoregressive polynomial of $\{x_{it}\}$.

Since both $\{y_t\}$ and $\{x_{it}\}$ are stationary sequences, the $\{\varepsilon_t\}$ is also stationary:

$$\varepsilon_t = y_t - \left\{ \mu + \sum_{i=1}^k [\theta_i(B)/\phi_i(B)] B^{L_i} x_{it} \right\} \quad (2)$$

Assuming that $\{\varepsilon_t\}$ is a non-white noise sequence, the autoregressive moving average model is then used to extract the information. The final fitted model is:

$$\begin{cases} y_t = \mu + \sum_{i=1}^k [\theta_i(B)/\phi_i(B)] B^{L_i} x_{it} + \varepsilon_t \\ \varepsilon_t = [\theta(B)/\phi(B)] a_t \end{cases} \quad (3)$$

where $\theta(B) = 1 - \sum_{j=1}^q \theta_j B^j$ is the q -order moving average operator, and $\phi(B) = 1 - \sum_{j=1}^p \phi_j B^j$ is the p -order autoregressive operator. $\{a_t\}$ is a zero-mean white noise sequence [21,22]. The specific process of generating NS using the ARIMAX model will be described in detail in Section 2.3.

2.2 LSTM structure

The aging degradation process of FC generally lasts from 1000-15000 hours. LSTM can selectively store the intrinsic information of long-term data and capture the long time-scale correlation between time series data. In fact, as in [14-19,24-33], the LSTM-based prognostic framework is demonstrated to be capable of predicting the short-term PEMFC degradation behavior. Specifically, the basic network configuration of the LSTM contains three layers, namely the input layer, the single LSTM layer, and the output layer, as in Fig. 1 (a).

To eliminate the risk of gradient disappearance or explosion in RNN, the LSTM creates a path that allows the gradient to flow continuously for a long time through an ingenious controllable self-circulation [15]. Specifically, in a single LSTM layer, h hidden units are connected sequentially to form a chain, which allows the cell state ($c_t \in \mathbb{R}^h$) and the hidden state ($h_t \in \mathbb{R}^h$) to be continuously passed, as in Fig. 1 (b). The $x_t \in \mathbb{R}^d$ (the superscript d refers to the number of input features) from the input layer enters the hidden unit along with the previous unit outputs (c_{t-1} and h_{t-1}). The new outputs (c_t and h_t) are generated with the collaboration of forget gate ($f_t \in \mathbb{R}^h$), input gate ($i_t \in \mathbb{R}^h$), output gate ($o_t \in \mathbb{R}^h$), and the internal state unit ($\tilde{c}_t \in \mathbb{R}^h$) [18].

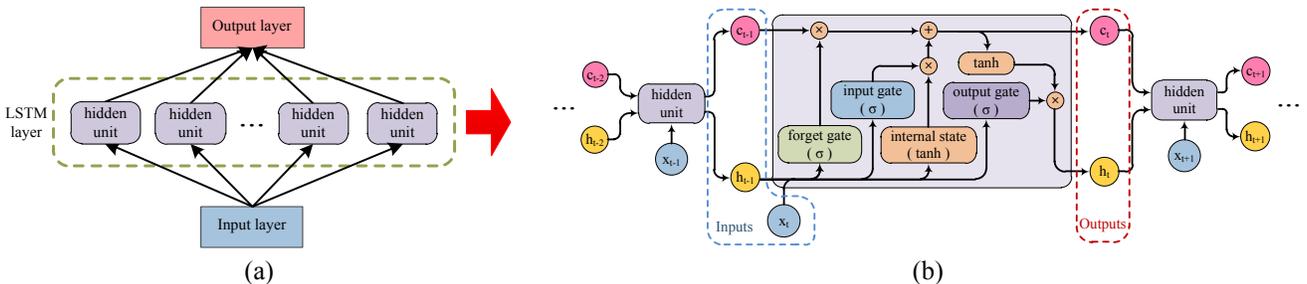


Fig. 1 - Illustration of LSTM model:
(a) LSTM (single-layer) network configuration;
(b) hidden unit chain and individual unit structure.

Equations (4) to (9) show the functional equations of LSTM

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (4)$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (5)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (6)$$

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (7)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \quad (8)$$

$$h_t = o_t \circ \tanh(c_t) \quad (9)$$

where σ is sigmoid function. \tanh is hyperbolic tangent function. The operator “ \circ ” denotes the Hadamard (element-wise) product [15]. The matrices $W_q \in \mathbb{R}^{h \times d}$ and $U_q \in \mathbb{R}^{h \times h}$ contain, respectively, the weights of the input and recurrent connections, where the subscript q can be either the input gate i , output gate o , the forget gate f or the cell state c . Similarly, the vectors $b_q \in \mathbb{R}^h$, such as b_f , b_i , b_o and b_c , are the bias of each component [14].

2.3 NSD-LSTM model architecture

FC operation involves coupled multi-physics processes. The aging degradation of FC is time-varying, highly uncertain and correlated to these processes. This leads to the prognostic, supposed to predict the intrinsic FC degradation evolution, is disturbed. Moreover, the prognostic model built on the historical data of the identical FC also suffers from the epistemic uncertainty as the predicted data evolve within an unseen quantity interval. In order to handle the disturbances from both the system operation/degradation and model uncertainties, the LSTM based prognostic model should be trained so as to trace the intrinsic degradation evolution. With this in mind, this paper proposes a novel prognostic strategy based on NSD-LSTM, which takes into account the characteristics of LSTM and ARIMAX. The overall NSD-LSTM based prognostic strategy is shown in Fig. 2. The strategy is divided into three parts: dataset preprocessing, offline phase and online phase.

2.3.1 Dataset preprocessing

The upper left frame of Fig. 2 shows the dataset preprocessing phase of the strategy. The purpose of this phase is to adjust the HI sequence to make it suitable for prognostic. The HI can be the stack voltage sequence from the aging experiment, or the virtual nominal stack voltage mentioned in the literature [12].

The first step is to select the split point. The split point can divide the HI sequence into a training set and a test set. Further, the test set can be used to calculate the real RUL. Subsequently, it is necessary to select the failure threshold (FT), and determine the actual end of life (EOL) according to FT. Then the actual remaining useful life (RUL) is calculated as

$$RUL = EOL - T_{SP} \quad (10)$$

where T_{SP} is the operating time at the split point. It is worth mentioning that the test set is considered unknown in both offline and online phases, in order to avoid data leakage, as in practical conditions.

The second step is to standardize the training set using:

$$z_{std} = (z - \bar{z})/SD \quad (11)$$

where $\{z\}$ is the raw training set sequence, \bar{z} is the average, SD is the standard deviation, and $\{z_{std}\}$ is the standardized training set sequence.

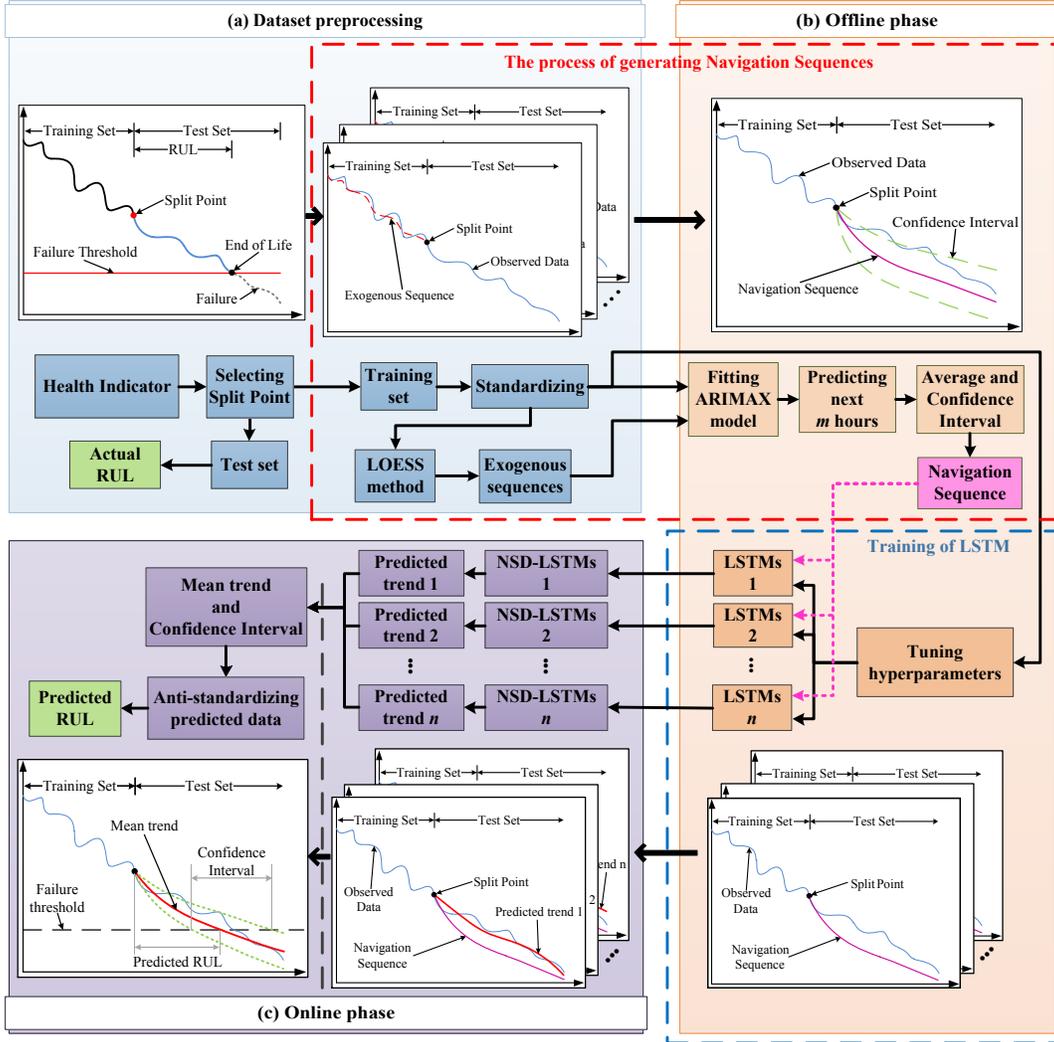


Fig. 2 - The prognostic process based on NSD-LSTM.

- (a) Dataset preprocessing: datasets segmentation, standardization, exogenous sequences generation;
- (b) Offline phase: navigation sequence generation, training of LSTM;
- (c) Online phase: long-term degradation trends prediction, RUL prediction.

The third step is to generate multiple XS for the ARIMAX model. The XS are introduced to better retain the main trends in the raw training set while filtering out anomalous and meaningless data fluctuations. Specifically, by adjusting the smoothing coefficient in the locally weighted scatterplot smoothing (LOESS) method [14], multiple XS with different smoothing levels can be obtained. These different smoothing levels facilitate diversified exogenous information for ARIMAX to prevent the impact by uncertainty disturbances.

Moreover, each XS can eliminate some fluctuations mainly due to the system operation while retaining the long-term degradation trend.

2.3.2 Offline phase

The offline phase is displayed on the right frame of Fig. 2. This phase contains two parts: generating NS and training LSTM.

1) Generating NS

The NS generation process is shown in the red dashed box in Fig. 2. The training set and XS generated in the preprocessing phase will be used to fit the ARIMAX model. Specifically, a series of ARIMAX models are fitted using multiple (here n) different XS. Based on these models n candidate sequences for NS can be predicted. Then, at the k -th predicted instance after the split point, the average predicted value and the confidence interval (CI) can be calculated respectively, as

$$\bar{\hat{y}}_k = \left(\sum_{i=1}^n \hat{y}_k^i \right) / n \quad (12)$$

$$CI_k = \bar{\hat{y}}_k \pm 1.96SD_k \quad (13)$$

Where \hat{y}_k^i is the i -th ($i \in \{1, 2, \dots, n\}$) predicted value, $\bar{\hat{y}}_k$ represents the average predicted value of the k -th time step, and SD_k denotes the identified standard deviation. The constant 1.96 means 95% CI. The CI_k are the confidence interval bounds. The average predicted value sequence $\{\bar{\hat{y}}\}$ is considered as the NS.

2) Training LSTM

In the training LSTM stage, Adam is utilized as an optimizer to facilitate hyperparameters tuning, because it is computationally efficient and has little memory requirements [23]. This drives the model to converge to the minimum generalization error in a stable and rapid manner. Meanwhile, Adam is proven to be suitable for training LSTM [14,18]. Different settings of hyperparameters and initialized weight matrices can bring diversity to the LSTM model. This variation of model is not fully controllable for the prediction phase. Any well-trained LSTM model can be considered to perform effectively in the training set, yet there is often a significant difference in its performance in the test set. Setting the multiple initialization weight matrixes of hidden units randomly, multiple LSTM can be trained to generate multiple prognostic models, such as LSTM 1, LSTM 2, ..., LSTM n shown in the blue dashed box in Fig. 2. The multiple LSTM setting benefits to provide diverse prediction results, which will be used further to improve the model epistemic uncertainty.

2.3.3 Online phase

As shown in the right frame in Fig. 2, the HI trend prediction and the RUL estimation are completed in the online phase.

1) Predicting degradation trends

Multi-step prediction is achieved by deploying well-trained LSTM iteratively for single-step prediction. Specifically, in order to break the historical numerical interval and mitigate the cumulative single-step prediction error, the NSD-LSTM model uses NS rather than the previous step prediction value as model input. Corresponding to multiple well-trained LSTM models in the offline phase, a series of NSD-LSTM models (NSD-LSTM 1, NSD-LSTM 2, ... , NSD-LSTM n) will be formed in the online phase. These NSD-LSTM models will give n different HI degradation prediction results. Assuming that the n predicted degradation trends follow the standard Gaussian distribution, which can be tested by hypothesis testing, the mean trend composed of the average of these predicted values is regarded as the final predicted HI degradation trend, and the CI is configured as a 95% probability interval using (13).

2) RUL estimation

The estimated end of life (\widehat{EOL}) is considered to be the operating time when the mean trend first-time reaches the FT. The estimated remaining useful life (\widehat{RUL}) is calculated according to equation:

$$\widehat{RUL} = \widehat{EOL} - T_{SP} \quad (14)$$

Additionally, the computing cost of the NSD-LSTM model will be discussed in Section 4.

3 Long-term FC aging tests and prognostic data sets

3.1 FC aging experiments

In view of the FC lifetime data scarcity limitations and the operating conditions influences. In this paper, two long-term FC aging experiments are conducted under constant and dynamic-load conditions, respectively.

3.1.1 Constant-load aging experiment

The first FC aging experiment in constant-load operating condition is completed on a 1kW Proton Motor 200 (PM200) fuel cell. In this experiment, the PM200 FC stack with 96 cells is tested under the operating conditions shown in Table 1 [15].

Table 1 - Constant-load aging experiment operating condition

Operating mode	Constant-load
Air supply	Air blower & filter
Cooling system	Deionized-water/glycol
Fuel gas supply	99.99% dry H ₂ @1.5 bar
Number of cells	96
Operating hours	10430 h
Stack temperature	58 °C
Current density	0.64 A/cm ²

3.1.2 Dynamic-load aging experiment

In the dynamic-load aging experiment, the concerned FC stack is of an open cathode and a dead-end anode structure. Some operating conditions are shown in Table 2. A 24V DC fan realizes the functions of air supply and temperature adjustment. The pressure of hydrogen as fuel is fixed at about 1.35 bar. In addition, FC is self-humidifying and a purge lasting for 0.5 seconds is activated every 30 seconds. In order to get closer to the dynamic operating conditions of FC in electric vehicle applications, a programmable electronic DC load is used to set the output current in the experiment [12].

Table 2 - Dynamic-load aging experiment operating condition

Operating mode	Dynamic-load
FC type	Open cathode/Dead-end anode
Active surface	33.63 cm ²
Number of cells	15
Nominal output power	73.5 W
Operating temperature	corresponding to current
Maximum temperature	75 °C
Maximum current	13.45 A
Lowest permitted stack voltage	7.5 V
Pressure interval at hydrogen inlet	0.10 to 0.40 bar

3.2 Datasets for prognostic

The observed/virtual average cell voltages are specified as HI sequences to verify the reliability and generalization of the prognostic strategy. Through the data set preprocessing method mentioned in Section 2, the HI sequences are processed to be suitable for prognostic. Fig. 3 (a) and (b) respectively show the constant-load and dynamic-load datasets obtained after preprocessing.

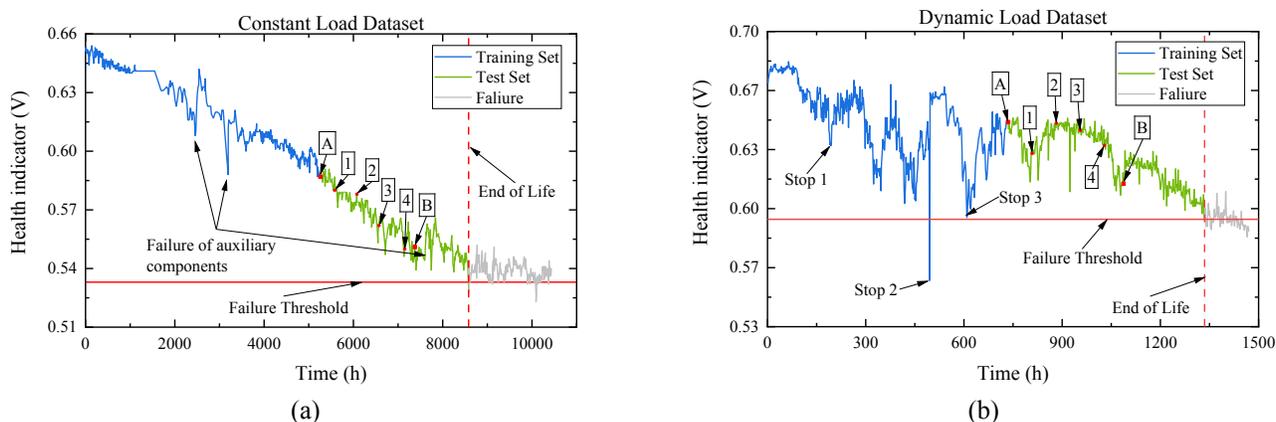


Fig. 3 - Fuel cell aging experimental datasets:
 (a) average cell voltage in the constant-load,
 (b) virtual nominal average cell voltage in the dynamic-load.

3.2.1 Constant-load dataset

The constant-load dataset is the average cell voltage, which comes from the actual stack voltage measured in the PM200 aging experiment. In the dataset, the highest voltage is 0.654 V and the lowest voltage is 0.523

V. The average initial voltage is 0.650 V, and the failure threshold is set to 0.533 V. The actual EOL corresponding to this failure threshold is 8583.3 h. As shown in Fig. 3 (a), there are three obvious voltage fluctuations, which are caused by the failures of auxiliary components such as the air blower or the cooling system.

In the experiment to verify the prognostic strategy, a series of split points are selected. From point **A** (5261.1 h) to point **B** (7378.2 h), 21 split points are set at equal intervals. Among them, between points **A** and **B**, four points are selected for comparison experiment: point **1** (5580.5 h), point **2** (6078.3 h), point **3** (6573.2 h), and point **4** (7151.6 h).

3.2.2 Dynamic-load dataset

The dynamic-load dataset is about a HI extracted using the same method proposed in the literature [12]. The HI can be considered as the nominal average cell voltage of the concerned stack. In the dynamic-load dataset, the highest voltage is 0.683 V and the lowest voltage is 0.559 V. The average initial voltage is 0.676 V, and the failure threshold is set to 0.595 V. The actual EOL corresponding to this failure threshold is 1335 h. As shown in Fig. 3 (b), there are three abnormal stops. Among them, the most obvious voltage fluctuation is near to stop 2 which is related to an unexpected system stop of hydrogen depletion.

In the experiment to verify the prognostic strategy, a series of split points are selected. From point **A** (735 h) to point **B** (1087.5 h), 25 split points are set at equal intervals. Among them, between points **A** and **B**, four points are selected for comparison experiment: point **1** (808.5 h), point **2** (882 h), point **3** (955.5 h), and point **4** (1029 h).

4 Evaluation of prognostic strategy

4.1 Evaluation criteria

This paper sets two evaluation criteria: the root-mean-square error (RMSE) and the relative error (RE). The RMSE evaluates the single-step prediction performance, and the RE is used in the online phase to evaluate the accuracy of the estimated RUL from the prognostic strategy. The mathematical description of RMSE and RE are given by:

$$RMSE = \sqrt{\left[\sum_{j=1}^m (y_j - \hat{y}_j)^2 \right] / m} \quad (15)$$

$$RE = (|RUL - \widehat{RUL}| / RUL) \times 100\% \quad (16)$$

where, y_j is the actual stack voltage observation value in the training set; \hat{y}_j is the predicted voltage value by the well-trained LSTM model in the offline phase; m is the sample number under evaluation.

4.2 Prognostic experiment and discussion

The programs used in the FC prognostic experiment are developed in a Matlab R2019b environment. Even with the help of Adam, some hyperparameters configuration and tuning still rely on manual adjustment. The following are part of the hyperparameters configured and tuned in this paper: the number of hidden units is set to 200 to ensure prediction accuracy and regulate model complexity; the dynamic learning rate is set to 0.001 initially and becomes smaller by 80% after 100 epochs to avoid training loss oscillation, overfitting, or slow convergence; the maximum training epoch is set to 150 based on the data characteristics of the training sets; the gradient threshold is set to 1 for avoiding gradient explosion.

4.2.1 Single-step prediction performance comparison test between LSTM and ARIMAX

In this paper, long-term/multi-step prediction is achieved by deploying single-step prediction iteratively. Therefore, it is essential to evaluate the credibility of the single-step prediction. Specifically, this evaluation uses the test set as input to evaluate the well-trained LSTM model in the NSD-LSTM. Moreover, in order to more comprehensively evaluate the single-step prediction performance, a comparison test between the ARIMAX model and the well-trained LSTM model is utilized. The results of the tests performed at point [A](#) of the two different load datasets are shown in Table 3 and Fig. 4. The results show that the ARIMAX model tends to be more trend-holding, while the LSTM is better at describing degradation behavior in detail. The LSTM predicted values match the test set well, and the RMSEs (as in Table 3) are slightly better than ARIMAX under different loads. However, it is undeniable that the error of LSTM gradually increases as the prediction continues, especially in the dataset with dynamic-load. In addition, the errors of LSTM are almost always positive when the errors become progressively larger, which confirms that the LSTM is limited by the historical data. Not to be neglected, the above single-step prediction results depend on future observations (test set). In case of observation scarcity, single-step predictions cannot provide long-term RUL estimates due to the prediction length limitation. Meanwhile, prediction results that reflect positive errors tend to provide lifetime valuations larger than the actual RUL, which further leads to the risk of fatal failures before maintenance. In summary, LSTM outperforms ARIMAX in single-step prediction, but still relies on multi-step prediction to achieve long-term prognostics.

Table 3 - Comparison experiments of single-step prediction performance

Dataset	Model	RMSE	Error (range)	Note
Constant-load	ARIMAX	0.004702	-0.0158 to 0.0135	Fig. 4 (a)
	LSTM	0.002574	-0.0044 to 0.0105	Fig. 4 (b)
Dynamic-load	ARIMAX	0.006095	-0.0129 to 0.0334	Fig. 4 (c)
	LSTM	0.005577	-0.0072 to 0.0135	Fig. 4 (d)

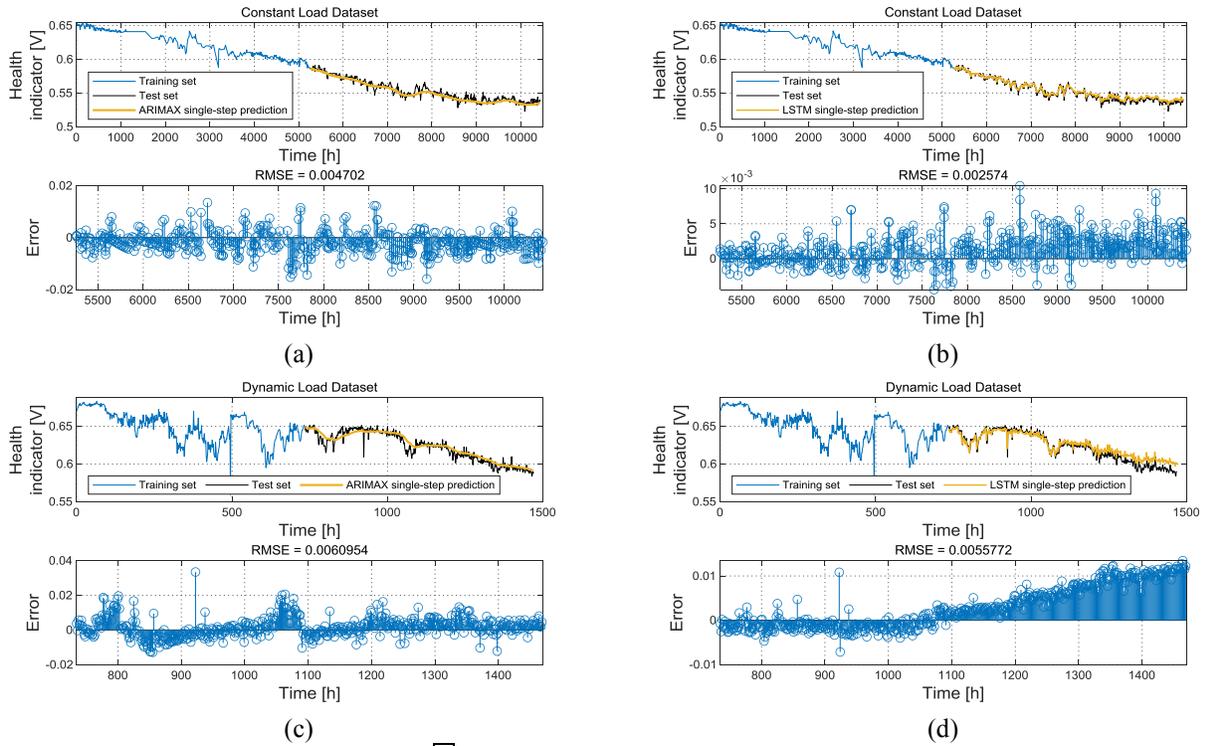


Fig. 4 - Single-step prediction at split point \square : in constant-load dataset, (a) ARIMAX model; (b) LSTM model; in dynamic-load dataset, (c) ARIMAX model; (d) LSTM model.

Remark: (1) Section 4.2.1 is devoted to discuss the single-step prediction capability of well-trained LSTM models in NSD-LSTM. (2) In the actual deployment of NSD-LSTM, the single-step prediction performance is considered to be fully optimized through the training phase. (3) In practice, highly accurate single-step prediction is more meaningful in real-time controlling fuel cell applications.

4.2.2 Evaluating the multi-step prediction performance

Multi-step prediction provides a longer prediction range and is the prerequisite for estimating RUL. To properly evaluate the multi-step prediction capability of the NSD-LSTM, 20 trend predictions are performed at each split point between points \square and \square . The mean trend is calculated as the multi-step prediction result and then compared with the ARIMAX and LSTM, as in Fig. 5.

Firstly, as in Fig. 5(a) and (d), various XS are generated in the data preprocessing phase by setting five different smoothing levels. Subsequently, in the offline phase, NS is generated by using ARIMAX model. The generated NS can maintain the historical potential trends in the HI, which facilitates to guide the NSD-LSTM during long-term degradation prediction.

Secondly, the HI degradation trends (as the colorful lines in Fig. 5 (b), (e) and (g)) predicted by NSD-LSTM remain consistent across the different load data sets. Meanwhile, the predicted trends are slightly lower than the test set in most cases. This helps to avoid estimating larger RULs and is expected to provide further maintenance recommendations with redundancy.

Thirdly, the prediction results of LSTM eventually converge to a horizontal line as the prediction time increases, as in Fig.5 (c), (f) and (h). This is reminiscent of the prevalent positive errors in the single-step predictions mentioned in Section 4.2.1 (Fig. 4 (b) and (d)). It confirms that LSTM is difficult to break the historical numerical interval, as well as the lack of long-term prediction ability. Further, the prediction results of ARIMAX retain partially historical trend information, but are generally much lower than the test set. This can cause the estimated RUL to be too small, triggering maintenance prematurely and increasing the cost. In contrast, NSD-LSTM gives more appropriate predictions.

Another interesting feature of the proposed NSD-LSTM model is that both short-term voltage fluctuations and long-term voltage evolution in the historical data can be captured and reflected in the prediction results. As in Fig. 5 (g) and (h), the predicted degradation trends appear to be quite different from the previous ones. The predicted trends show a clear and rapid decline at the beginning of the prediction horizon, and then gradually restored a quasi-linear decline. Among them, the obvious rapid decline in the early stage of the prediction is due to the fact that the relatively abnormal voltage drop before the split point is captured by the NSD-LSTM model. The abnormal voltage drop may indicate that the fuel cell is undergoing changes in operating conditions or encountering the failure of auxiliary equipment. These factors will lead to the changes in the HI degradation trend in the future, so it is meaningful to consider such abnormal voltage drops in the long-term prognostics. Subsequently, the predicted trend returns to a state similar to a quasi-linear decline, and continues to maintain the global downward trend contained in the training set. Therefore, the prediction based on NSD-LSTM model not only considers the abnormal HI changes that may affect the trend evolution, but also adjusts the prediction results in a relatively short time to retain the global degradation trend shown in the training set. Such prediction property highlights the robustness of NSD-LSTM against the disturbances from system operations and recoverable faults.

Furthermore, on a regular personal computer, the CPU is used for model training and prediction. As in Table 4, the computing costs of the three models are compared. In the training/fitting (offline) phase, the computing cost of NSD-LSTM is slightly higher than that of LSTM, and ARIMAX performs the best. In the offline phase, there is usually ample time for model training and hyperparameter tuning, so this is acceptable. More interestingly, the computing cost of the all models is almost as well during prediction. It indicates that NSD-LSTM has the ability to implement online prediction.

Table 4 - Comparison of computing costs.

Dataset	Split point	Training/fitting time (s)			Predicting time (s)		
		ARIMAX	LSTM	NSD-LSTM	ARIMAX	LSTM	NSD-LSTM
Constant-load	3	3.85	57.27	64.45	1.53	1.37	1.44
Dynamic-load	1	3.75	36.70	43.56	1.71	1.57	1.59
	3	3.76	43.83	50.47	1.75	1.33	1.26

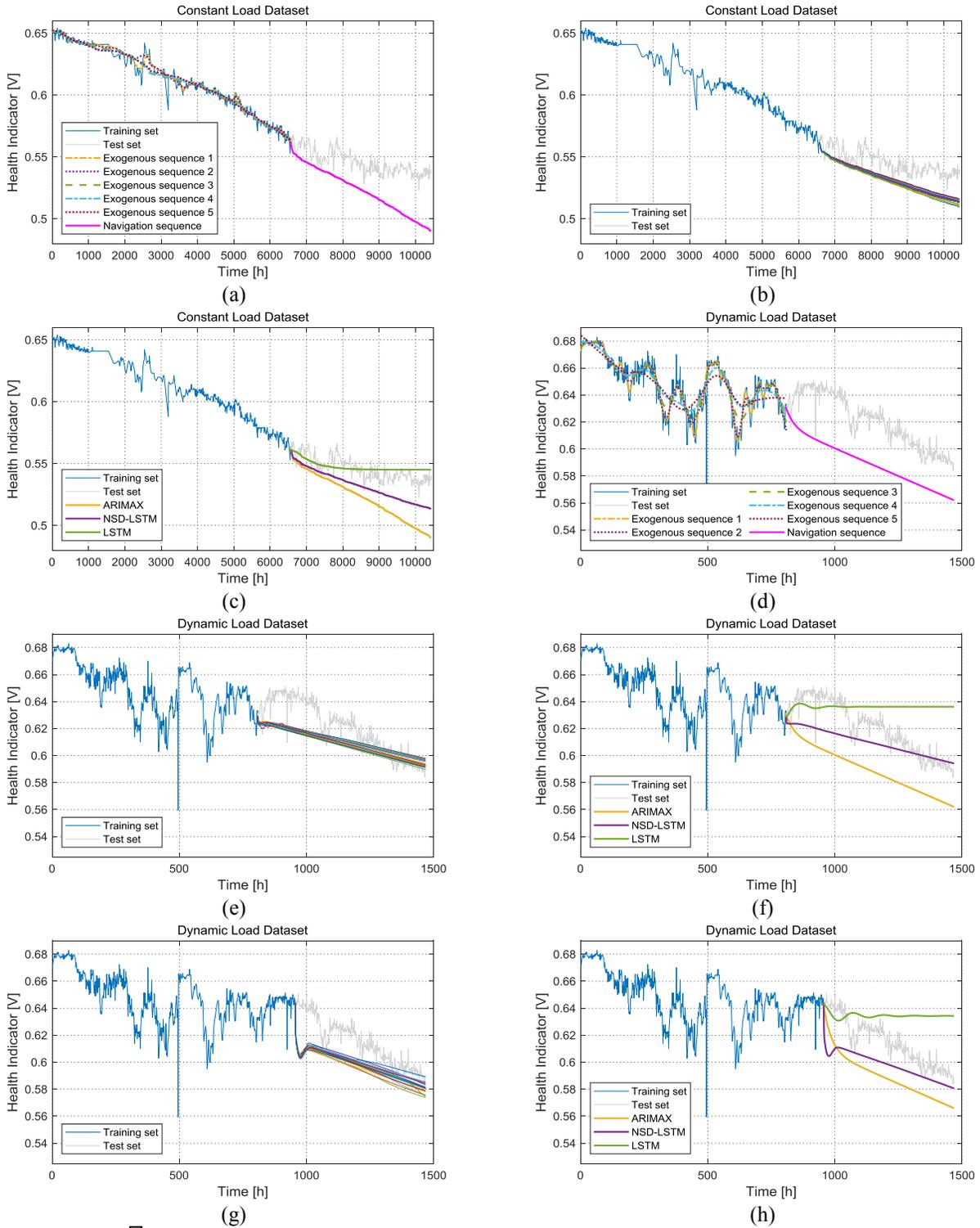


Fig. 5 - At split point $\textcircled{3}$ in constant-load dataset: (a) 5 different exogenous sequences (XS), and navigation sequence (NS) ; (b) 20 predicted degradation trends (the colorful lines); (c) multi-step prediction results comparison.

At split point $\textcircled{1}$ in dynamic-load dataset: (d) 5 different exogenous sequences (XS), and navigation sequence (NS) ; (e) 20 predicted degradation trends (the colorful lines); (f) multi-step prediction results comparison.

At split point $\textcircled{3}$ in dynamic-load dataset: (g) 20 predicted degradation trends (the colorful lines); (h) multi-step prediction results comparison.

Remark: The term “robustness” mentioned in this paper mainly refers to an ability of the proposed

prognostic method to counteract abnormal fluctuations in health indicators and to keep the predicted results in line with the inherent degradation trend of the fuel cell.

4.2.3 Evaluating RUL estimation

Thanks to the HI degradation trends predicted by the NSD-LSTM model, RUL can be estimated before the fuel cell failure. In order to evaluate the estimated RUL, multiple tests are performed between the split points **A** and **B** for both datasets.

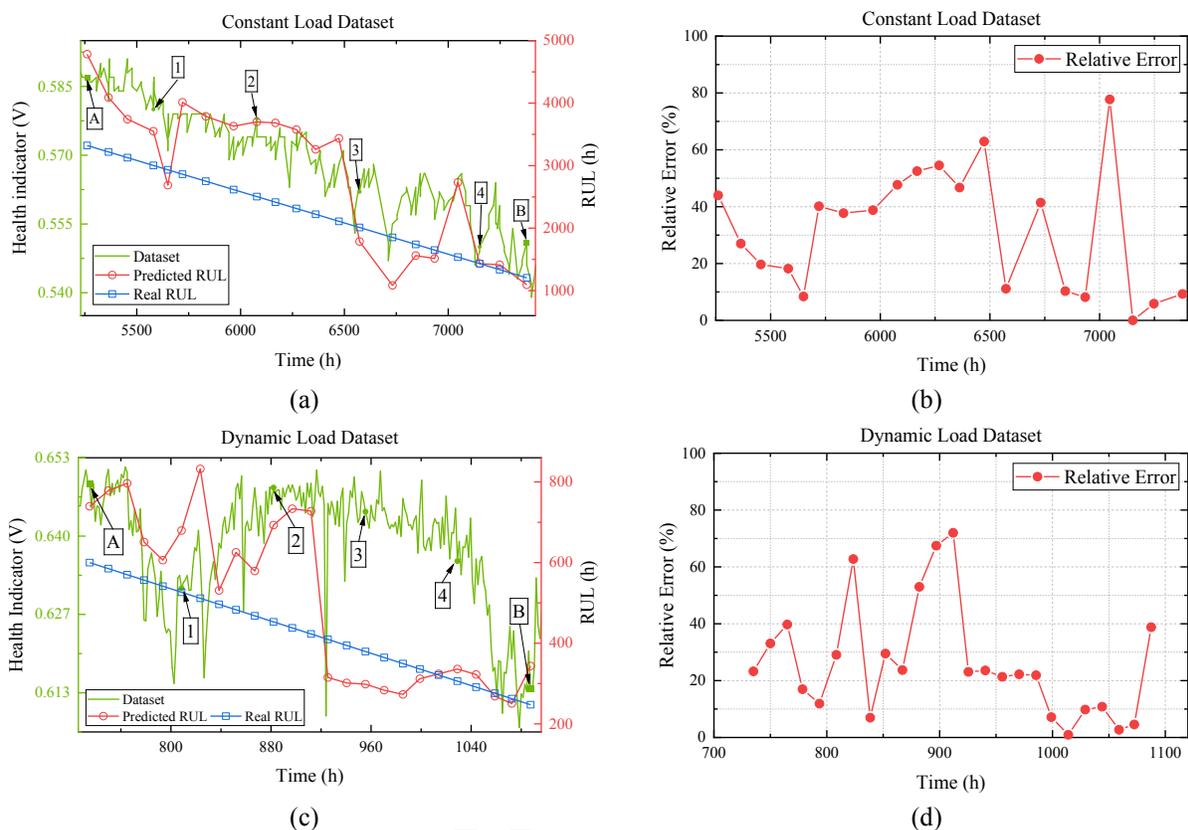


Fig. 6 - RUL estimation results from split point **A** to **B**:
in the constant-load dataset, (a) 21 estimated RULs versus real RULs and dataset; (b) 21 relative errors.
in the dynamic-load dataset, (c) 25 estimated RULs versus real RULs and dataset; (d) 25 relative errors.

Fig. 6 (a) and (b) show the results of 21 RUL estimations and REs for the constant-load dataset. Among them, Fig. 6 (a) has a dual y-axis of which the left y-axis (green) represents the health indicator (V), corresponding to the dataset curve (green line) and the right y-axis (red) represents RUL (h), corresponding to the estimated RULs (red line) and the real RULs (blue line). The estimated RULs from the split point **A** to point **B** as a whole exhibit a phenomenon that fluctuates around the real RULs. It can be observed that the predicted fluctuations of RULs are consistent with those of the dataset. Specifically, the predicted RULs from split point **A** to split point **1** gradually approach the real RULs. This is because the training set near the split point shows an increasingly steep downward trend, which makes the predicted RULs tend from overestimation to the real RULs. The estimated RUL drops significantly after split point **1**. This is because

NSD-LSTM takes into account the abnormal drop in HI near the split point. Then the estimated RUL near split point [2] tends to maintain the same trend of the dataset. It is worth mentioning that there are several short-term HI drops between points [2] and [3] whose effects are not obviously reflected in the RUL estimation. The prognostic model tends to maintain the prediction consistency during this period. As there is a significant HI drop in the short period before point [3], the estimated RUL also drops significantly. Between split point [3] and split point [4], there are obvious fluctuations in the dataset, and the estimated RULs also exhibit similar fluctuations. Between split point [4] and split point [B], the estimated RULs are very close to the real RULs. In Fig. 6 (b), the REs of the estimated RULs are shown. The maximum RE is 77.73%, the minimum is 0.01%, and the average is 32.28%. In 21 tests, 17 REs are less than 50%, and this is considered to be acceptable for long-term prognostic. It should be noted that from split point [A] to split point [B], the real RUL ranges from 3322.2 h to 1205.1 h, which means that NSD-LSTM need to perform RUL estimation for thousands of hours when the test set is unknown. In summary, for the constant-load dataset, the estimated RULs fluctuate near the real RULs, and the fluctuation of the estimated RULs is consistent with the evolution trend of the dataset.

Fig. 6 (c) and (d) shows the 25 RUL estimation results and the REs for the dynamic-load dataset. The axis significance is the same as in Fig. 6 (a) and (b). Between split point [A] and split point [2], the estimated RULs are generally higher than the real RULs; and between split point [2] and split point [B], the estimated RULs fluctuate near the real RULs. The estimated RULs are generally consistent with the dataset in terms of fluctuation trends. For instance, the estimated RULs rise slightly after point [A], which is due to the HI increase related to the recovery from an abnormal stop of the fuel cell before point [A] (“stop 3” in Fig. 3 (b)). Following that as far as split point [1], the dataset exhibits a significant HI drop and a recovery behavior. The estimated RULs also show a lagged similar evolution trend. It is worth noting that there are several “sags” in the dataset, such as the part between split point [A] and split point [3]. These “sags” can be understood as some recoverable faults or abnormal operating conditions [12]. From the results, the NSD-LSTM model can comprehensively consider the long-term and short-term degradation behaviors. In Fig. 6 (d), the REs of the 25 estimated RULs are shown. The maximum value of REs is 72%, the minimum value is 0.9%, and the average value is 26.3%. In 25 tests, 22 REs are lower than 40%, which is considered to be acceptable for the prognostics under dynamic operating conditions. It should be noted that the real RUL range from split point [A] to split point [B] is 600 h to 247.5 h, which means that when the test set is unknown, the NSD-LSTM model needs to perform hundreds of hours of long-term RUL estimation. In summary, for the dynamic-load dataset, the estimated RULs can track the changes in the degradation behavior, and provide consistent and robust RUL estimates in the long-term prognostics.

4.2.4 Comparison experiments

In order to further evaluate the consistency and robustness of the proposed prognostic strategy, comparison experiments are designed at split points $\boxed{1}$, $\boxed{2}$, $\boxed{3}$ and $\boxed{4}$ of the dynamic-load dataset. It should be pointed out that in these comparison experiments, the NARX model uses the same NS of the NSD-LSTM model as the exogenous variable. This will help reveal the guiding role of NS and highlight the LSTM performance. In addition, the ESN model in the comparison experiments is from [11,12].

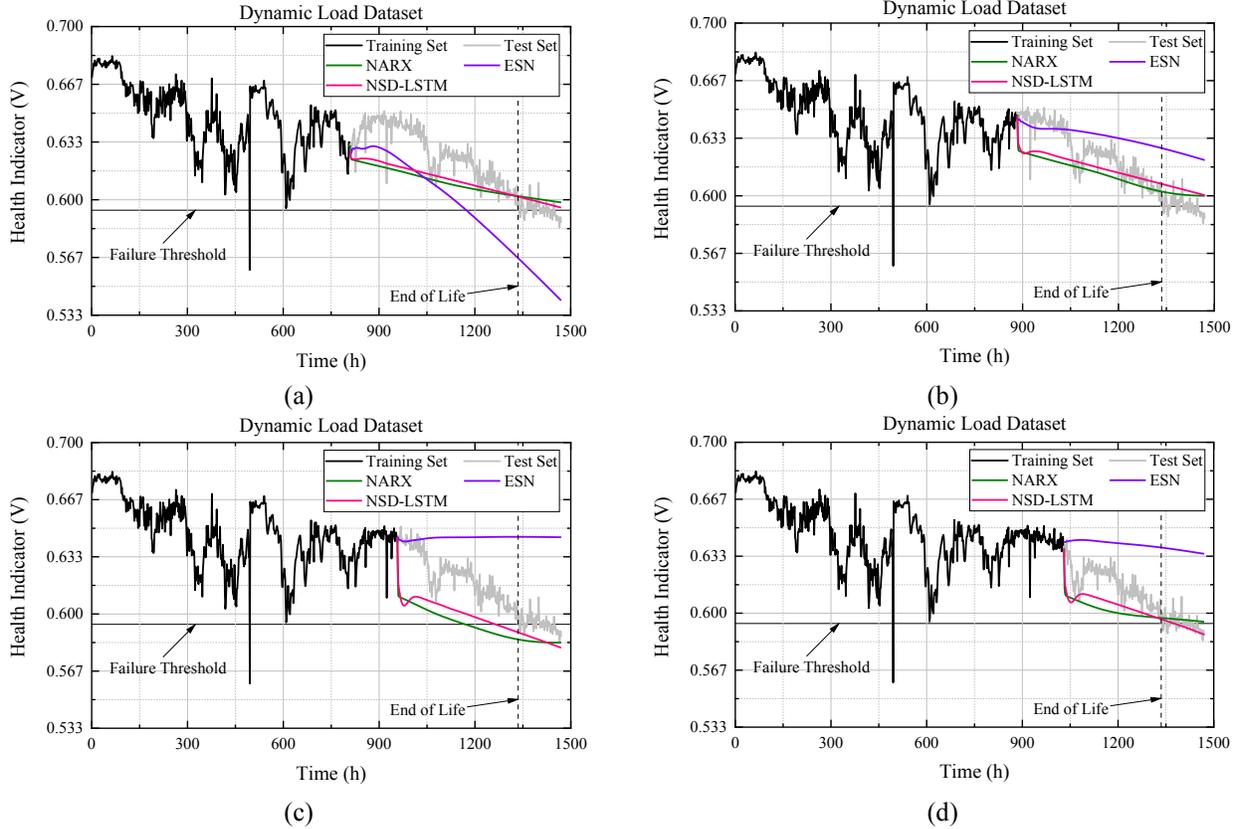


Fig. 7 - Comparison experiment of degradation trend predictions in the dynamic-load dataset, among them, (a), (b), (c) and (d) represent the results of split points $\boxed{1}$, $\boxed{2}$, $\boxed{3}$, and $\boxed{4}$, respectively.

As in Fig. 7, the degradation trends predicted by the ESN at each of the four split points are different. Especially at split points $\boxed{3}$ and $\boxed{4}$, the trends predicted by ESN are close to the horizontal line. Compared with the ESN model, the trends predicted by the NARX model have better consistency at the four split points. However, as the prediction time increases, the trends predicted by NARX model gradually become flat. This phenomenon is more pronounced at split points $\boxed{3}$ and $\boxed{4}$. Compared with the previous two models, the prediction results of the NSD-LSTM model show the best consistency and are closest to the test set. Especially at split points $\boxed{3}$ and $\boxed{4}$, the predicted trends of NARX and NSD-LSTM seems to be similar in the early stage of prediction, and both show a rapid decline thanks to the guidance of the NS. Then, the well-trained LSTM in NSD-LSTM provides more “long-term memory” information about the degradation trend for the prediction process, which enables NSD-LSTM to maintain stable and consistent trend prediction

performance.

Table 5 - Comparison experiments in the constant-load

Split point	Operating time (h)	Actual	RUL (h)			RE (%)		
			NARX	ESN	NSD-LSTM	NARX	ESN	NSD-LSTM
1	5580.5	3002.8	2369.8	2778.8	3549.9	21.08	7.46	18.22
2	6078.3	2505.0	4645.9	2024.5	3699.8	85.47	19.18	47.70
3	6573.2	2010.1	3271.9	1462.3	1786.2	62.77	27.25	11.14
4	7151.6	1431.7	1252.6	1052.6	1431.9	12.51	26.48	0.01

In the comparison experiment on the constant-load dataset, as shown in Table 5 and Fig. 8 (a) and (b), the consistency of the NARX model is the worst, the consistency of the NSD-LSTM model is better. It is of interest that the ESN model performs slightly better than the NSD-LSTM model in terms of consistency, but unfortunately, the prediction accuracy of the ESN model does not decrease as expected as the operating time increases. Specifically, at split point 1, the RE performances of the NARX and NSD-LSTM model are similar, while ESN model performs better. At split point 2, the prediction performance of all three models decreases, with the NARX model being the worst. At split point 3 and split point 4, both the NARX model and the NSD-LSTM model show a trend of gradually improving prediction accuracy, with the NSD-LSTM being more significant. Comparatively, the ESN model performs poorly.

Table 6 - Comparison experiments in the dynamic-load

Split Point	Operating time (h)	Actual	RUL (h)			RE (%)		
			NARX	ESN	NSD-LSTM	NARX	ESN	NSD-LSTM
1	808.5	526.5	798.0	361.5	649.5	51.57	31.34	23.36
2	882.0	453.0	1381.5	1042.5	777.0	204.97	130.13	71.52
3	955.5	379.5	202.5	2637.0	246.0	46.64	594.86	35.18
4	1029.0	306.0	457.5	1561.5	348.0	49.51	410.29	13.73

In the comparison experiment on the dynamic-load dataset, as shown in Table 6 and Fig. 8 (c) and (d), the consistency of the ESN model is the worst, the consistency of the NARX model is slightly better, and the consistency of the NSD-LSTM model is the best. Specifically, at split point 1, the performances of the three models are similar. At split point 2, the RUL estimated by the ESN model has increased significantly, and the RE has exceeded 130%. What is more serious is that at split point 3 and split point 4, the REs of the ESN model have exceeded 590% and 410%, respectively. In contrast, the RULs estimated by the NARX and NSD-LSTM models fluctuate close to the real RULs. The NARX model performs worse than the NSD-LSTM model at all four split points. Especially at split point 2, the RE of NARX exceeds 200%. In summary, compared with the other two models, the NSD-LSTM model is capable of reproducing the short-term and long-term degradation behaviors and provides a more consistent prediction of the degradation trends. The RULs estimated by the NSD-LSTM model are more reliable, with enhanced robustness.

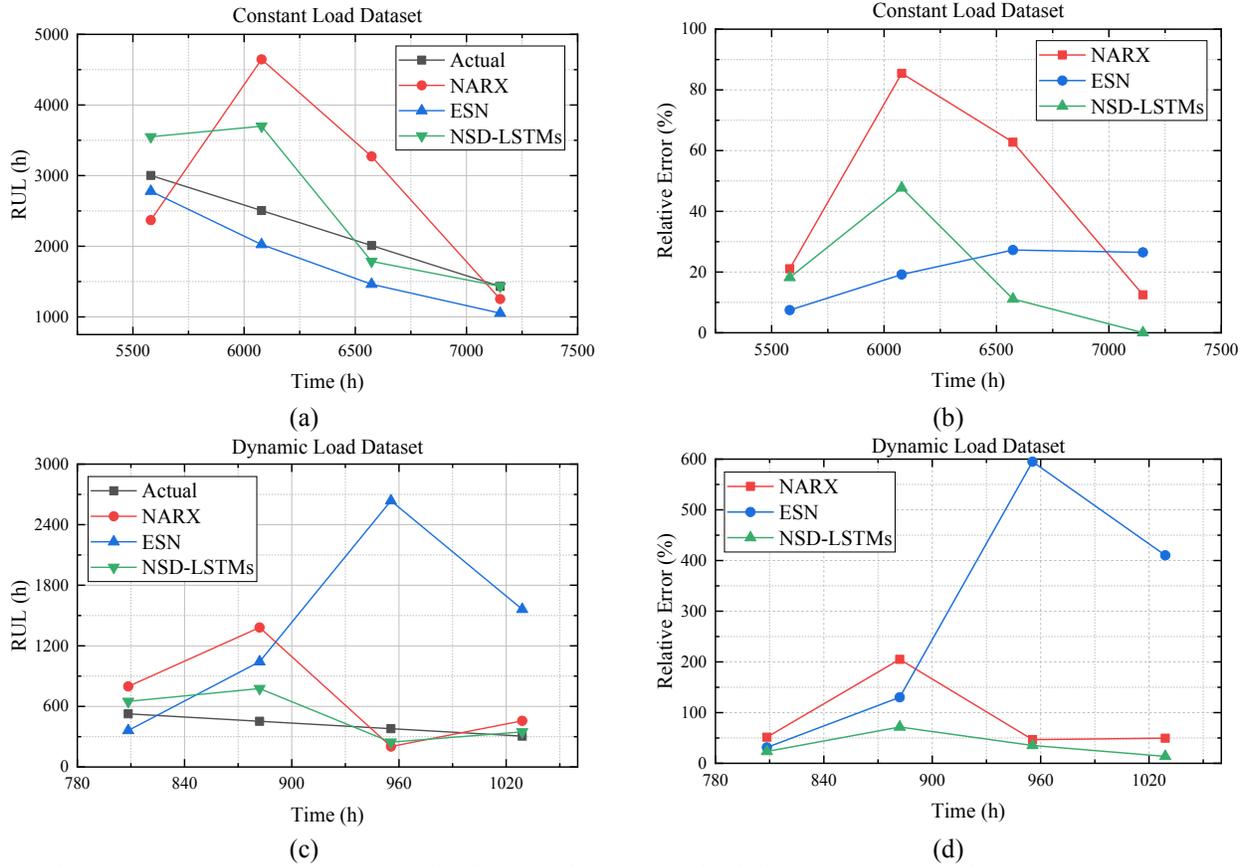


Fig. 8 - Comparison experiment of RUL estimations: in the constant-load dataset, (a) estimated RULs; (b) relative errors; in the dynamic-load dataset, (c) estimated RULs; (d) relative errors.

4.2.5 Discussion

In this paper, the fuel cell prognostics are discussed and the proposed NSD-LSTM model shows a satisfactory prediction performance. Especially, it outperforms other state-of-the-art prognostic methods in some cases, which can be considered to arise from the following factors.

1) In comparison experiments, the ESN model effectively identifies a part of the degradation trend, especially 100-200 hours before the prediction starting point. However, due to the effect of system disturbances/recoverable faults, a short-term trend may differ significantly from the inherent degradation trend. This results in poor prediction performance, especially in the dynamic load dataset.

2) In the comparison experiments, the NARX model predicts similar degradation trends as the NSD-LSTM model. This is due to the fact that both use the same NS, which also illustrates the availability and consistency of NS in guiding predictions.

(3) The prediction accuracy of the NARX model is slightly lower, which is due to the influence of the historical data interval on its prediction range. Especially, the phenomenon is more significant in Fig. 7(d). Along with the prediction time (step), the degradation trend predicted by NARX gradually converges to the

horizontal line, which leads to the predicted RUL abnormally higher. In contrast, the NSD-LSTM can “remember” and “inherit” the historical degradation trend to break the numerical interval limitation.

5 Conclusion

This paper proposes a robust long-term prognostic strategy, i.e., NSD-LSTM, for fuel cells. In the strategy, the navigation sequence generated by the ARIMAX model is fed into the well-trained LSTM to reinforce the performance of long-term degradation trend prediction and RUL estimation. From the comparison experiment results, the following conclusions can be made:

- Compared with the existing LSTM-based prognostic models, the NSD-LSTM model achieves credible degradation prediction with longer prediction horizons. Based on the long-term prediction, RUL estimation is further realized.
- Thanks to the guidance and improvement of navigation sequence, the NSD-LSTM gains the ability to break the numerical interval of historical-based training set, and the predicted long-term degradation trends are consistent.
- Both the long-term and the short-term degradation behaviors are considered comprehensively in the fuel cell prognostic. Due to this property, better degradation prediction and RUL estimation results are obtained compared to several state-of-the-art methods.

Although the above work has been able to accurately predict some of the health indicators, it is still not fully apply this result to diverse practical operating conditions of fuel cells. In future work, efforts will be made to explore the effects of load conditions on various degradation processes. Meanwhile, the sensitivity of the prognostic model will be investigated under different practical operating conditions. Furthermore, the combination of the prognostic model with the fuel cell degradation mechanism model will be pursued to improve the prediction performance.

Acknowledgement

This work was supported in part by the China Scholarship Council (CSC) under Grant [grant number 201906290107].

References

- [1] Das V, Padmanaban S, Venkitusamy K, Selvamuthukumar R, Blaabjerg F, Siano P. Recent advances and challenges of fuel cell based power system architectures and control – A review. *Renew Sust Energ Rev* 2017;73:10-8. <https://doi.org/10.1016/j.rser.2017.01.148>.
- [2] David H, Jones S, Lewis J. The Fuel Cell Industry Review 2020. E4tech 2020. <https://fuelcellindustryreview.com>.

- [3] Lucia U. Overview on fuel cells. *Renew Sust Energ Rev* 2014;30:164-9. <https://doi.org/10.1016/j.rser.2013.09.025>.
- [4] Sharaf OZ, Orhan MF. An overview of fuel cell technology: Fundamentals and applications. *Renew Sust Energ Rev* 2014;32:810-53. <https://doi.org/10.1016/j.rser.2014.01.012>.
- [5] Eudy L, Post M B. Fuel Cell Buses in U.S. Transit Fleets: Current Status 2018. National Renewable Energy Lab. (NREL), Golden, CO (United States); 2019. <https://doi.org/10.2172/1489893>.
- [6] Liu J, Zio E. Prognostics of a multistack PEMFC system with multiagent modeling. *Energy Sci Eng* 2019;7(1):76-87. <https://doi.org/10.1002/ese3.254>.
- [7] 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 Manage* 2018;156:723-32. <https://doi.org/10.1016/j.enconman.2017.11.069>.
- [8] Liu H, Chen J, Zhu CY, Su HY, Hou M. Prognostics of Proton Exchange Membrane Fuel Cells Using A Model-based Method. *IFAC-PapersOnLine* 2017;50(1):4757-62. <https://doi.org/10.1016/j.ifacol.2017.08.947>.
- [9] 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 Aug;324:745-57. <https://doi.org/10.1016/j.jpowsour.2016.05.092>.
- [10] Hua Z, Zheng Z, Gao F, Péra MC. Remaining useful life prediction of PEMFC systems based on the multi-input echo state network. *Appl Energ* 2020 Mar;265:114791. <https://doi.org/10.1016/j.apenergy.2020.114791>.
- [11] Li Z, Jemei S, Gouriveau R, Hissel D, Zerhouni N. Remaining useful life estimation for PEMFC in dynamic operating conditions. in 2016 IEEE Vehicle Power and Propulsion Conference 2016. <https://doi.org/10.1109/VPPC.2016.7791762>.
- [12] Li Z, Zheng Z, Outbib R. Adaptive Prognostic of Fuel Cells by Implementing Ensemble Echo State Networks in Time-Varying Model Space. *IEEE T Ind Electron* 2020;67(1):379-89. <https://doi.org/10.1109/TIE.2019.2893827>.
- [13] Jouin M, Gouriveau R, Hissel D, Péra MC, Zerhouni N. Prognostics of PEM fuel cell in a particle filtering framework. *Int J Hydrogen Energ* 2014 Jan;39(1):481-94. <https://doi.org/10.1016/j.ijhydene.2013.10.054>.
- [14] 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 Energ* 2019;44(11):5470-80. <https://doi.org/10.1016/j.ijhydene.2018.10.042>.
- [15] 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 Energ* 2018;231:102-15. <https://doi.org/10.1016/j.apenergy.2018.09.111>.
- [16] 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 Transportation and Electrification Conference and Expo 2018:102-7. <https://doi.org/10.1109/ITEC.2018.8449962>.
- [17] Wang FK, Mamo T, Cheng XB. 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 Jun;461. <https://doi.org/10.1016/j.jpowsour.2020.228170>.
- [18] Wang FK, Cheng XB, Hsiao KC. Stacked long short-term memory model for proton exchange membrane fuel cell systems degradation. *J Power Sources* 2020 Feb;448. <https://doi.org/10.1016/j.jpowsour.2019.227591>.

- [19] 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 T Ind Appl* 2019 Jul-Aug;55(4):4321-31. <https://doi.org/10.1109/TIA.2019.2911846>.
- [20] Yang CC, Li ZH, Liang B, Lu WN, Wang XQ, Liu HD. A Particle Filter and Long Short Term Memory Fusion Algorithm for Failure Prognostic of Proton Exchange Membrane Fuel Cells. in 29th Chinese Control and Decision Conference 2017:5646-51. <https://doi.org/10.1109/CCDC.2017.7978172>.
- [21] Box GE, Jenkins GM, Reinsel GC, Ljung GM. *Time Series Analysis: Forecasting and Control*, 5th Edition. John Wiley & Sons, 2015.
- [22] Jalalkamali A, Moradi M, Moradi N. Application of several artificial intelligence models and ARIMAX model for forecasting drought using the Standardized Precipitation Index. *Int J Environ Sci Te* 2014;12(4):1201-10. <https://doi.org/10.1007/s13762-014-0717-6>.
- [23] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint 2014;arXiv:1412.6980. <https://arxiv.org/abs/1412.6980>.
- [24] Wang C, Li Z, Outbib R, Zhao D, Dou M. Proton Exchange Membrane Fuel Cells Prognostic Strategy Based on Navigation Sequence Driven Long Short-term Memory Networks. In *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, Singapore, 2020:3969-74. <https://doi.org/10.1109/IECON43393.2020.9255373>.
- [25] Wang C, Li Z, Outbib R, Dou M, Zhao D. Symbolic Deep Learning Based Prognostics for Dynamic Operating Proton Exchange Membrane Fuel Cells. *Appl Energ* 2022;305. <https://doi.org/10.1016/j.apenergy.2021.117918>.
- [26] Zheng L, Hou Y, Zhang T, Pan X. Performance Prediction of Fuel Cells Using Long Short - Term Memory Recurrent Neural Network. *Int J Energ Res* 2021;45(6):9141-61. <https://doi.org/10.1002/er.6443>.
- [27] Zhang Z, Wang YX, He H, Sun F. A Short- and Long-Term Prognostic Associating with Remaining Useful Life Estimation for Proton Exchange Membrane Fuel Cell. *Appl Energ* 2021;304. <https://doi.org/10.1016/j.apenergy.2021.117841>.
- [28] 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 Jul;113:149-65. <https://doi.org/10.1016/j.isatra.2020.06.005>.
- [29] Ma R, Xie R, Xu L, Huangfu Y, Li Y. A Hybrid Prognostic Method for Pemfc with Aging Parameter Prediction. *IEEE T Transp Electr* 2021;7(4):2318-31. <https://doi.org/10.1109/TTE.2021.3075531>.
- [30] Zuo B, Cheng J, Zhang Z. Degradation Prediction Model for Proton Exchange Membrane Fuel Cells Based on Long Short-Term Memory Neural Network and Savitzky-Golay Filter. *Int J Hydrogen Energ* 2021;46(29):15928-37. <https://doi.org/10.1016/j.ijhydene.2021.02.069>.
- [31] Zuo J, Lv H, 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 Energ* 2021;281. <https://doi.org/10.1016/j.apenergy.2020.115937>.
- [32] Zhou S, Zhao P, Fan L. A Hybrid Method for Estimation of Pemfc State of Health. in the 40th Chinese Control Conference, Shanghai, China, 2021 July 26-28. <https://doi.org/10.23919/CCC52363.2021.9549420>.
- [33] Wang FK, Huang CY, Mamo T, Cheng XB. Ensemble Model for the Degradation Prediction of Proton Exchange Membrane Fuel Cell Stacks. *Qual Reliab Eng Int* 2020;37(1):34-46. <https://doi.org/10.1002/qre.2718>.
- [34] Gomathi K, Karthik M, Usha S. An Intelligent Parametric Modeling and Identification of a 5 kw Ballard Pem Fuel Cell System Based on Dynamic Recurrent Networks with Delayed Context Units. *Int J Hydrogen Energ* 2021;46(29):15912-27. <https://doi.org/10.1016/j.ijhydene.2021.02.065>.