DTU Library

Degradation prediction of PEMFCs using stacked echo state network based on genetic algorithm optimization

Deng, Zhihua; Chen, Qihong; Zhang, Liyan; Zhou, Keliang; Zong, Yi; Liu, Hao; Li, Jishen; Ma, Longhua

Published in:
IEEE Transactions on Transportation Electrification

Link to article, DOI:
10.1109/TTE.2021.3111906

Publication date:
2022

Document Version
Peer reviewed version

Citation (APA):
Deng, Z., Chen, Q., Zhang, L., Zhou, K., Zong, Y., Liu, H., Li, J., & Ma, L. (2022). Degradation prediction of PEMFCs using stacked echo state network based on genetic algorithm optimization. *IEEE Transactions on Transportation Electrification*, *8*(1), 1454 - 1466. https://doi.org/10.1109/TTE.2021.3111906

Degradation prediction of PEMFCs using stacked echo state network based on genetic algorithm optimization *

Zhihua Deng[1], Qihong Chen[1 ,†](Member, IEEE), Liyan Zhang[1 ,‡](Member, IEEE), Keliang Zhou[1](Senior Member, IEEE), Yi Zong[2](Member, IEEE), Hao Liu[3](Member, IEEE), Jishen Li[1], Longhua Ma[4](Senior Member, IEEE)

[1]School of Automation, Wuhan University of Technology, Wuhan, Hubei, 430070, China

[2]Center for Electric Power and Energy, Technical University of Denmark, Risø, Roskilde, 4000, Denmark

[3]State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, Zhejiang, 310027, China

[4]School of information and engineering, NingboTech University, Ningbo, Zhejiang 315100, China

Abstract: Durability is considered as one of the main technical obstacles to the large-scale commercialization of proton exchange membrane fuel cells (PEMFCs), which can be effectively improved through prognostics prediction techniques. This paper proposes a stacked echo state network (ESN) based on the genetic algorithm (GA) to predict the future degradation trend of PEMFCs. By alternately using the projection layer and the encoding layer, the proposed method can make full use of the temporal kernel property of the ESN to encode the multi-scale and multi-level dynamics of the stack voltage, thereby obtaining more robust generalization performance and higher accuracy than existing methods. Specifically, a stack voltage time series of PEMFCs is projected into the high-dimensional echo state space of the reservoir. Then, an auto-encoder projects the echo state representation into the low-dimensional feature space. After that, the genetic algorithm is utilized to optimize the hyperparameters of the developed model. Based on two open-source datasets of PEMFCs with different accelerated test conditions, this paper systematically tested the proposed degradation prediction methods based on different model structures. Test results demonstrate that the proposed method is superior to traditional prediction methods in terms of accuracy and generalization performance.

Keywords: Proton exchange membrane fuel cells, prognostics prediction, stacked echo state network, genetic algorithm, projection-encoding

*The work is supported by National Natural Science Foundation of China (62173264), Ningbo Science and Technology Plan Project (2019B10116), the "SPGC" (640170582) and "PRESS" (807300026B) projects granted by the Danish Energy Technology Development and Demonstration program (EUDP) and the Danish Agency for Science & Higher Education, respectively. The work is also supported by the Fundamental Research Funds for the Central Universities (2020YB034) and China Scholarship Council (202006950031).

†Corresponding author, E-mail: chenqh@whut.edu.cn (Q. Chen).

‡Corresponding author, E-mail: zlywhut@whut.edu.cn (L. Zhang).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TTE.2021.3111906, IEEE Transactions on Transportation Electrification

1. Introduction

Due to the advantages of little impact by environmental changes, high energy conversion efficiency, zero pollution, low noise, short refueling time, and long cruising range, the proton exchange membrane fuel cells (PEMFCs) system is considered one of the most promising energy sources [1, 2]. It has been widely used in vehicles, stationary power generation, portable equipment, etc. [3, 4]. However, the poor durability of PEMFCs under complex dynamic operating conditions is one of the main technical obstacles to the large-scale commercialization of PEMFCs systems. The existing studies have shown that in addition to material improvement and structural optimization, the use of prognostics and health management (PHM) techniques can also improve the durability of PEMFCs. The main principle of PHM is to take effective maintenance measures in advance to avoid further deterioration by accurately estimating the current state of health and predicting the future state of degradation for PEMFCs [5, 6]. This paper focuses on predicting the future degradation trend of PEMFCs.

The degradation of PEMFCs is a complex nonlinear process involving multi-physics, multi-scale, and multi-component. Various factors in design, production, and application will affect the degradation path of PEMFCs. Therefore, it is a challenge to accurately predict the future degradation trend of PEMFCs. The existing PEMFCs degradation trend prediction methods [7, 8] are mainly divided into three categories, namely model-based methods [9, 10], data-based methods [11, 12], and hybrid methods [7, 13-15]. For model-based methods, most of the research oriented to online applications achieves the prediction of degradation trends by establishing semi-mechanical or empirical models and combining filtering algorithms, such as the method based on semi-mechanical models and unscented kalman filter [16], the method based on extended kalman filter [17], the method based on empirical models and particle filters [9], etc.

Compared with the model-based method and hybrid method, the data-based method has the advantages of small computational burden, simple deployment, strong generality, and does not require an in-depth understanding of the degradation mechanism. In recent years, it has been widely used in PHM related technologies of various complex systems. The existing data-based methods used to predict the degradation trend of PEMFCs include: stacked long short-term memory (LSTM) method [18], bi-directional LSTM recurrent neural network (RNN) method [19], LSTM RNN method [20], wavelet analysis and nonlinear autoregressive exogenous neural network method [21], and particle filter and RNN fusion method [22], etc. In particular, many deep learning-based methods have achieved impressive high performance. An attention-based RNN model was proposed to improve the prog-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TTE.2021.3111906, IEEE Transactions on Transportation Electrification

nostics of PHM [23], which enables a more accurate prediction of the stack voltage degradation of PEMFCs based on the original long-term dynamic loading cycle durability test data. Similar works have been developed by Ma *et al.*, in [24]. The remaining useful life (RUL) prediction based on deep learning (DL) algorithm was proposed to build the degradation model, which can effectively and accurately forecast the degradation of PEMFCs system. At the same time, some studies have also tried to use automatic hyper-parameter optimization to further improve the prediction accuracy and the automation of the method. A novel degradation model of fuel cells based on grey neural network was proposed by Chen *et al.*, in [25], which was optimized by particle swarm optimization algorithm. Chen *et al*,.[5] proposed a hybrid method to establish the degradation model of PEMFCs based on the wavelet analysis, extreme learning machine and genetic algorithm. Subsequently, the authors also proposed a novel degradation model of PEMFCs based on the backpropagation neural network and evolutionary algorithm [26]. Similar works in [18] proposed a stacked LSTM prognostic model based on the differential evolution algorithm, which can improve the prediction accuracy. Similar works also in [21] proposed a novel wavelet neural network method to establish the degradation model of fuel cells. The parameters of built model were optimized by cuckoo search algorithm. The results indicated that the accuracy of the optimized degradation model was better than the manual parameters model.

Although most of the above methods have realized the accurate prediction of the future aging trend of PEMFCs, the sensitivity analysis of the number of layers and training data lengths to degradation prediction accuracy of the DL model mentioned above are rarely involved. Thus, these problems motivate us to conduct the research of this paper.

Among all the methods mentioned above, DL and RNN methods have been proved to have a powerfully ability to predict the RUL of fuel cells [20, 23-24]. Nevertheless, the lengthy process of training deep RNN remains a practical problem, which limits its practical application. To overcome its shortcomings, the structure of echo state network (ESN) was first proposed by Jaeger *et al.*, in [27], which was worth mentioning. ESN is a typical RNN with a fixed state transition structure (the reservoir) that has exhibited excellent performance in time series prediction problems. Subsequently, Hua *et al.*, [28] proposed a multi-input ESN method to improve the accuracy of RUL prediction, which can obtain the higher prediction accuracy than single-input ESN method under both static operating conditions and dynamic conditions. Similar works also in [29] proposed an improved multi-input ESN method to predict the RUL degradation of PEMFCs. Due to the most time series contains a multi-scale and multi-level structure, a single-hidden-layer reservoir computing (RC) model cannot extract effective degradation features from the input time series [30]. Therefore, it is an important challenge to develop a new ESN structure to further improve the prediction accuracy of fuel cells life.

Besides, the stack output voltage is chosen as the aging state of PEMFC in most literature, which is easily measured by a sensor. For example, the aging of catalyst will cause a larger activation overvoltage, which results in lower output voltage [31]. In addition, the stack voltage time series usually contains the multi-scale and multi-level information of the RUL degradation, which changes with operating conditions. Thus, the stack output voltage is regarded as training input and an evaluation index for the aging state of PEMFCs in this paper, which can be considered as a time series prediction problem, and model it using machine learning methods.

To further improve the accuracy of degradation trend prediction without loss of simplicity, this paper proposes a stacked ESN based on the genetic algorithm (GA) for the first time. The most notable feature of the proposed method is its multiple projection-encoding-based stacked architectures. More specifically, when the stack voltage input time series is projected into the echo state space of the reservoir, the subsequent encoding layer receives the echo state of the previous reservoir as input and encodes the high-dimensional echo state representations into a lower-dimensional feature space. Then, these encoded representations are again projected into the high-dimensional state space of the next reservoir through random connections. Using this multiple projection-encoding method, the stacked ESN can make full use of the temporal kernel property of each reservoir to capture the multi-scale and multi-level dynamics of the stacked voltage time series instead of directly stacking multiple reservoirs in a completely random way. After that, based on the stacked ESN architecture, the GA is utilized to optimize the hyper-parameters in the constructed model, which can further improve the degradation trend prediction accuracy of PEMFCs. The contributions of this paper are as follows:

(1) A stacked ESN based on projection-encoding between reservoirs is proposed to build the degradation model of PEMFCs, which takes advantage of the merits of both DL and RC, and bridges the gap between them. Because the advantages of a deep structure are used to capture the multi-scale and multi-level dynamic of the stack voltage. In addition, RC is an effective method to construct recurrent networks that require less training of the network and pursues conciseness and effectiveness.

(2) Unsupervised encoding the echo states layer by layer and using the parameters optimization algorithm, the proposed method realizes the use of the advantages of the deep structure to capture the multi-scale and multi-level dynamics of the stack voltage, while obtaining more robust generalization performance and higher accuracy than existing methods.

(3) An evolutionary algorithm is proposed to optimize the reservoir hyper-parameters of stacked ESN, which improves the accuracy and automation of degradation trend prediction for PEMFCs.

(4) Based on the training data of different lengths and different dataset sources, the proposed methods of varying model structures have been tested and compared. Test results show that the proposed method can guarantee the accuracy of the prediction even with less training data.

The rest contents of this paper are organized as follows. Problem formulation is illustrated in Section 2. Section 3 presents the principle of stacked ESN based on the GA. The experiment results are given and further analyzed in Section 4, and it is followed by a conclusion and further study in Section 5.

2.Problem formulation

A PEMFCs system mainly consists of air supply subsystem, hydrogen subsystem, PEMFCs stack, cooling water subsystem, gas humidification, electronic load, and control subsystem [8]. The schematic of PEMFCs system is illustrated in Figure 1. The air and hydrogen are humidified and transported to the cathode and anode of the stack, respectively. To avoid the fuel cells stack "starvation" phenomenon, the supply of gas is regulated by the pressure and flow controller. And to maintain the internal temperature of the stack within an acceptable range, an additional cooling water subsystem is often added to the system. Also, in order to effectively prevent hydrogen leakage and control the pressure difference between the cathode and anode of the stack, hydrogen circulating pump and back pressure valve components are added to the system.

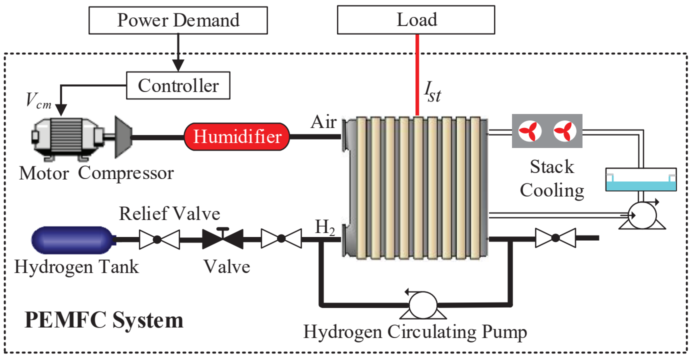

Figure 1: Schematic of PEMFCs system.

There are many factors that affect the PEMFCs degradation performance, operating conditions (open-circuit/idling, dynamic load, startup-shutdown and constant load) are the most key factors that accelerate the aging of the materials and components of fuel cells stack [29]. Among them, the constant load is represented by a constant current of 70A in this paper. Dynamic load conditions are the most serious impact on the degradation performance of fuel cells, and account for the highest proportion of performance degradation factors. Firstly, dynamic load conditions mainly bring the change of dynamic heat/humidity, and even operating pressure, and consequently, the mechanical degradation of components will be accelerated. Secondly, potential voltage cycling accelerates degradation of the electrochemical surface area and the growth of Pt particles. Thirdly, dynamic load increases the change of gas starvation due to the combined effects of demand dynamics, internal differences, and lagging gas supply in local reactant supply. Understanding the degradation mechanisms under different conditions is of great significance to extend the RUL of fuel cells. This paper focuses on the two harsh conditions for PEMFCs: constant load conditions and dynamic load conditions.

For most degradation prediction methods, there are some problems such as the prediction accuracy of the degradation model and the complexity of the degradation model. To solve these problems, a stacked ESN based on the GA method is proposed to compromise the accuracy of the degradation

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TTE.2021.3111906, IEEE Transactions on Transportation Electrification

model and the complexity of the degradation model. An accurate and efficient degradation model of fuel cells plays an important role in ensuring the safety and reliability of the system and extending its service life.

3.Stacked ESN modeling based on GA

In this section, a novel stacked ESN method is proposed to develop accurate degradation model of fuel cells. As illustrated in Figure 2, the implementation processes of the proposed degradation prediction can be divided into five steps: (1) Collect two datasets from the fuel cells aging test bench. (2) Preprocess the stack voltage time series data. (3) Train the stacked ESN model and optimize its hyper-parameters using GA. (4) Calculate the output weight matrix of stacked ESN based on the GA model. (5) Predict new data. The detailed content of the stacked ESN based on the GA method is described as follows.

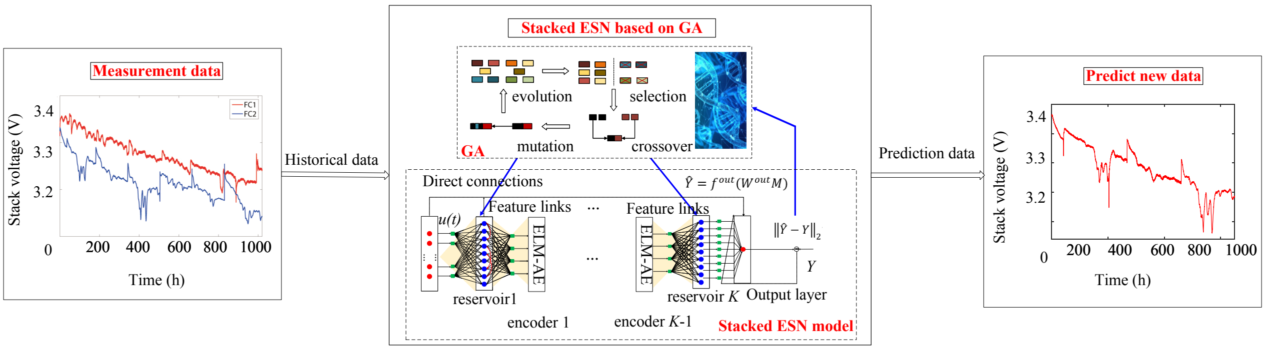

Figure 2: Implementation framework of the stacked ESN based on the GA.

3.1.The stacked ESN model

According to the analysis of the aforementioned, the stack voltage of fuel cells is regarded as the input time series (U) in this paper. Thus, the degradation prediction of fuel cells can be considered as a time series prediction problem, and model it using a stacked ESN method. ESN has a strong ability (a "reservoir" of dynamics) to deal with complex time series problems [32]. The input weights and recurrent weights of ESN are randomly initialized and fixed during the training phase and only the outputs weights are calculated. It can effective avoid the laborious process of gradient-descent RNN training, yet achieve excellent performance in time series prediction. Therefore, in this paper a stacked ESN method is proposed to predict the degradation trends of fuel cells. By using projection layers and encoding layers alternately and using parameters optimization algorithm, this method can not only learn the multi-scale and multi-level dynamics of the stack voltage measurement data, and also provide robust generalization degradation performance, which make it accurately predict the RUL of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TTE.2021.3111906, IEEE Transactions on Transportation Electrification

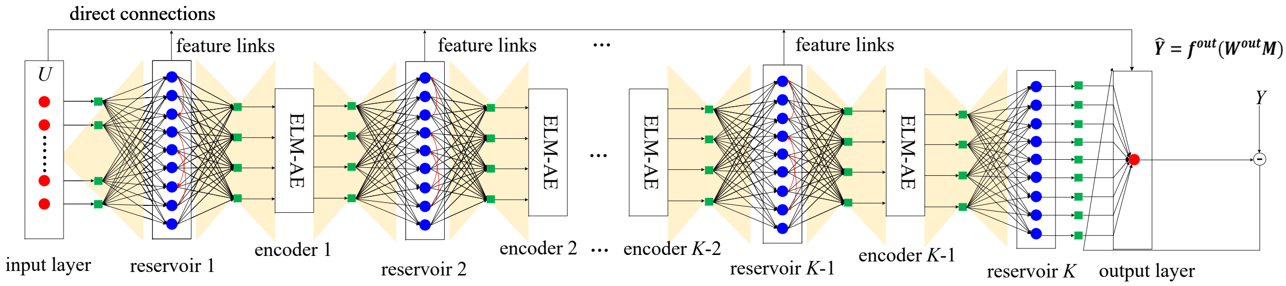

Figure 3: The structure diagram of stacked ESN based on ELM-AE method.

fuel cell. The formulation details of the stacked ESN method can be described as follows.

The structure diagram of the stacked ESN model is illustrated in Figure 3. The T-length time series inputs are represented as $U = [u(1), u(2), \cdots, u(T)]$ and the prediction results are denoted as the vector $\hat{Y} = [\hat{y}(1), \hat{y}(2), \cdots, \hat{y}(T)]$. For the i-th reservoir (i=1, $\ldots$, K), its high-dimensional states can be expressed as

$$z^{(i)}(t+1) = W^{res(i)} x_{res}^{(i)}(t) + W^{in(i)} x_{in}^{(i)}(t+1) \tag{1}$$

$$x_{res}^{(i)}(t{+}1){=}\,(1{-}\gamma)\, x_{res}^{(i)}(t) + \gamma f(z^{(i)}(t+1)) \tag{2}$$

where $x_{res}^{(i)}(t)$ and $x_{in}^{(i)}(t+1)$ are the updated states in i-th layer reservoir at time step t and the input of i-th reservoir at time step $(t+1)$, $W^{res(i)}$ and $W^{in(i)}$ denote the hidden-to-hidden weights and the input-to-hidden weights of i-th reservoir, respectively. $f(\cdot)$ and γ are the nonlinear activation function ($tanh(\cdot)$) and the leak rate of the reservoir, respectively. When i is greater than one, the inputs of i-th reservoir are the output of the (i-1)-th encoder, i.e. $x_{in}^{(i)}(t+1) = x_{enc}^{(i-1)}(t+1)$. On the contrary, we have $x_{in}^{(1)}(t+1) = u(t+1)$. For simplicity, an operator $\mathcal{F}_i$ is used to express the high-dimensional projection Eq. (1) and the update step Eq. (2), which can be expressed as

$$x_{res}^{(i)}(t+1) = \mathcal{F}_i(x_{res}^{(i)}(t), x_{in}^{(i)}(t+1)) \tag{3}$$

The states of the previous reservoir are known, to keep the computational merits of RC, the encoder $\mathcal{T}$ should have low learning cost. The dimensionality reduction technique of ELM-AE [32] has the advantages of less training parameters, fast learning speed and good generalization performance. It is used for simplifying the training of traditional auto-encoders. The key idea is to achieve the hidden random features $\mathrm{H} \in {}^{M\times N}$ by using random weights $W^0 \in {}^{M\times D}$ and bias $b^0 \in {}^{M\times D}$, which can be expressed as

$$\mathrm{H} = g(W^0 X + b^0) \tag{4}$$

in which $X \in {}^{D\times N}$ and g are the inputs and the activation function, respectively. Then the dimension reduction mappings $W^* \in {}^{M\times D}$ can be obtained by

$$W^* {=} \arg\max_{W} \|W\mathrm{H}{-}X\|_2 + \lambda\|W\|_2 \tag{5}$$

in which λ is the regularization parameter. The above equation can be solved by the pseudo-inverse method. Then, the reduced data can be represented by $\mathrm{H}_{enc}=(W^{*})^{T}X$. Therefore, the encoding procedure of the j-th encoder (j=1, ..., K-1) can be defined as

$$x_{enc}^{(j)}(t+1) = \mathcal{T}\left(x_{res}^{(j)}(t+1)\right) \tag{6}$$

Further, we can instantiate $\mathcal{T}(\cdot)$ in Eq. (6) as

$$\mathcal{T}\left(x_{res}^{(j)}(t+1)\right) = f_{enc}\left(W^{enc(j)}x_{res}^{(j)}(t+1)\right) \tag{7}$$

in which $f_{enc}(\cdot)$ denotes the activation function of the encoder, which is the identity function.

According to the above derivation, the state representations of the last reservoir can be formulated as

$$x_{res}^{(K)}(t+1) = \mathcal{F}_K \circ \mathcal{H}_{K-1} \circ \cdots \circ \mathcal{H}_1(u(t+1)) \tag{8}$$

in which $\mathcal{H}_j = \mathcal{T}_j \circ \mathcal{F}_j$ and the symbol $\circ$ is a composition operator.

Unlike the traditional ESNs, the stacked ESN incorporates additional middle-layer encoded features into the last output layer. Thus, the outputs of the whole system at time step (t+1) can be obtained as

$$\hat{y}(t+1) = f^{out}(W^{out}M(t+1)) \tag{9}$$

where $M(t+1)$ can be represented as

$$M(t+1) = \left[\underbrace{x_{res}^{(K)}(t+1)^T}_{\mathrm{A}}, \underbrace{u(t+1)^T}_{\mathrm{B}}, \underbrace{\left\{x_{enc}^{(1,\ldots,K-1)}(t+1)^T\right\}}_{\mathrm{C}}\right]^T \tag{10}$$

in which A, B and C are the echo states of the last reservoir, the input along with direct connections, and the multi-scale representations along with the feature links, respectively. Thus, the Eq. (9) can be rewritten as

$$\hat{Y} = f^{out}(W^{out}M) \tag{11}$$

in which f^{out} denotes the element-wised output activation function and the columns of $\hat{Y}$ and M range over 1, ..., T.

The parameters W^{out}can be represented by using a standard squared error loss function

$$E(W^{out}) \propto \left\|\hat{Y} - Y\right\|_2^2 \tag{12}$$

in which Y is the real data of the stack output voltage. Thus, Eq.(12) is a regression problem on the parameters W^{out}. As the time series is usually a high-dimensional form, this problem always is over-determined, and the ridge-regression with Tikhonov regularization [33] is regarded as the most universal and stable way to compute the output weight matrix.

$$\hat{W}^{out} = YM^T(MM^T + \beta I)^{-1} \tag{13}$$

where I and β are the identity matrix and a small regularization parameter (such as 10^{-4}).

According to the mentioned above, the proposed prediction method can preserve sufficient input features in the states through using projection layers and encoding layers alternately, especially for predicting long-term dependent measurement data. In addition, an analysis of stacked ESN's computational complexity is illustrated in Appendix A. Furthermore, to promote the convergence rapidity of the optimizing the hyper-parameters in each reservoirs and to further improve the accuracy of the degradation prediction of fuel cells, a GA algorithm will be proposed in the next part.

3.2.Stacked-ESN fuel cells life prediction model based on GA

As the depth of stacked RC models increases, it becomes especially important to optimize the reservoir hyper-parameters of stacked ESN, since it indirectly affects the prediction accuracy of the degradation trends of fuel cells. Generally, the manual adjustment of parameters is unsuitable for such the deep reservoir networks structure. In addition, the parameters of stacked ESN are randomly generated.

Table 1: The parameter settings of GA.

Parameters	Value
Population size	20
Maximum number of generations	100
Crossover probability	0.6
Mutation probability	0.1
Generational gap	10

Especially, when the randomly generated parameter value is zero, some neurons in the hidden layer are ineffective, which reduces the accuracy of the stacked ESN model. The reservoir hyper-parameters of stacked ESN should be optimized to obtain the better prediction performance. The GA can simplify the calculation procedure of the traditional method. Therefore, the GA is introduced into the stacked ESN, which is used for optimizing the reservoir hyper-parameters of the stacked ESN. Compared with traditional ESN method, the proposed method only needs to optimize the key hyper-parameters in each reservoir: the input scaling (IS), the spectral radius (SR), and the leak rate (γ), which can improve the prediction accuracy of the degradation model of PEMFCs. GA is proposed to optimize three hyper-parameters in each reservoir for the improvement of convergence rapidity. A flow chart of the GA is illustrated in Figure 4. The major implementation steps of the proposed stacked ESN based on the GA are described as follows.

Step 1: Establish the degradation model of stacked ESN. According to the characteristics of the PEMFCs aging data, the topology of the degradation model of stacked ESN includes the number of neurons in the input layer, the number of neurons in the hidden layer, and the number of neurons in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TTE.2021.3111906, IEEE Transactions on Transportation Electrification

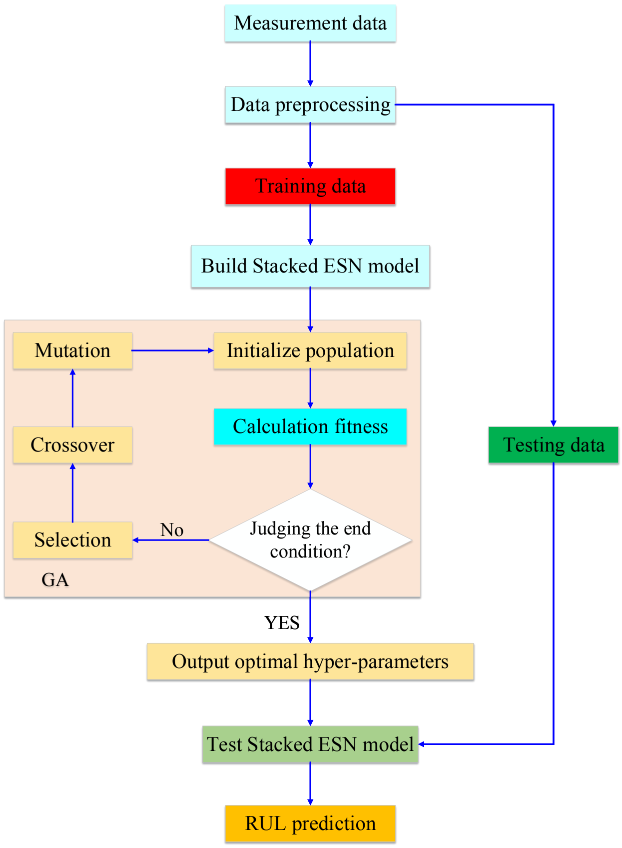

Figure 4: Schematic diagram of stacked ESN based on the GA.

the output layer.

Step 2: Initialize population. All triple hyper-parameters of each reservoir are concatenated into a vector, which is coded with the binary form to produce the initial population.

Step 3: Calculation fitness. To calculate the individual fitness value of the current population, the training data are applied to train the stacked ESN model.

Step 4: Judge the end condition. If the individual reaches the optimization algebra or reaches the set stop threshold, stop the iteration and jump to step 6; otherwise, perform step 5.

Step 5: Update the population. The selection, crossover, and mutation operations are performed according to the current individual fitness value, and then the new progeny population is generated. Take the progeny population as the current population, jump to step 2.

Step 6: Export the optimal IS, SR, and γ of the key hyper-parameters in each reservoir.

Step 7: Build the degradation model with the optimal parameters based on stacked ESN.

4.Experiment and validations

The experimental data comes from IEEE PHM 2014 DATA CHALLENGE [34], which is used to verify the different models. The aging experimental test platform for PEMFCs system is 1.0 kW electrical power, and the stack contains 5 cells and the active area of each cell is 100 cm^2, which is integrated at Federation for Fuel Cell Research (FCLAB). Some measurable and controllable physical parameters in the test bench like the temperature, pressure, flow, current, etc. are presented in Table 2. All the experiments are conducted under the MATLAB R2015b, and the computer with 2.3 GHz Intel Core i5 processor.

Table 2: Range of physical parameters controlled of the aging experimental test bench.

Parameters	Control range	Unit
Cooling temperature	20-80	°C
Cooling flow	0-10	L/min
Gas temperature	20-80	°C
Gas humidification	0-100%	RH
Air flow	0-100	L/min
H_2 flow	0-30	L/min
Gas pressure	0-2	Bar
Fuel cell current	0-300	A

The stack output voltage is regarded as a health management indicator of a PEMFCs system in this paper. The aging experimental test conducted two long-term durability tests for more than 1000 h under constant load and dynamic load operating conditions. In the first aging test, the constant load current of 70 A (FC1) was imposed on the aging test. In the second aging test, a triangular ripple current of 7 A with 5 kHz (FC2) was superimposed to the constant current of 70 A. The aging experimental test data contain lots of noise and large spikes (durability test shut down after restart) that would a very adverse effect on the accurate prediction of degradation trend of PEMFCs. The aging sampled data of FC1 with in red and FC2 with in blue are smoothed by gaussian-weighted moving average filter, which is shown in Figure 5. It can be seen from Figure 5 that the stack voltage of FC1 and FC2 decrease as time increased, which indicates the degradation phenomena of the PEMFCs stack.

4.1.Evaluation metric

Two quantitative indicators are applied to evaluate the prediction performance of different models such as root mean squared error (RMSE) and mean average percentage error (MAPE), which are

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TTE.2021.3111906, IEEE Transactions on Transportation Electrification

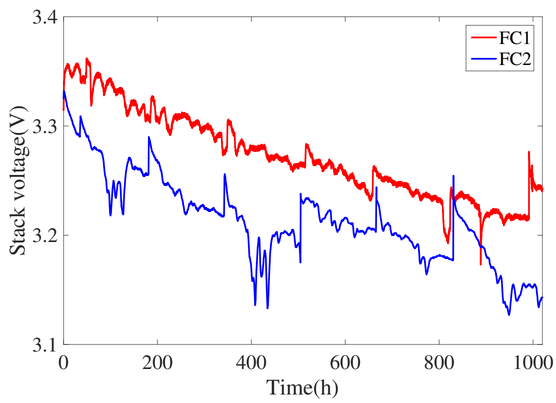

Figure 5: Degradation voltage curves of FC1 and FC2.

respectively formulated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t(n) - \hat{Y}_t(n)\right)^2} \tag{14}$$

$$\text{MAPE} = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{Y_t(n) - \hat{Y}_t(n)}{Y_t(n)}\right| \tag{15}$$

in which $Y_t(n)$ and $\hat{Y}_t(n)$ are the real voltage and the prediction voltage of the model during the RUL time (n data points), respectively. T denotes the number of points. The lower RMSE, and MAPE values, the better their corresponding degradation performances.

4.2.Effects of hyper-parameters

For all the weights of stacked ESN, $W^{res(i)}$, $W^{in(i)}$, and $W^{enc(j)}$ are given randomly, only W^{out} needs be trained. According to the output Eq. (11), the output weights depend on the last reservoir state $x_{res}^{(K)}(t+1)$, the input $u(1+t)$ and the $(K-1)$ reservoir states $x_{enc}^{(1,\ldots,K-1)}(t+1)$. The purpose of training output weights is to minimize the network output error, such that the prediction output can approximate to the desired output as much as possible.

Hyper-parameters optimization is seldom discussed in the existing methods. Since the optimization of the built model parameters indirectly affects the prediction accuracy of degradation performance for fuel cells, it is very important for optimize the hyper-parameters such as IS, SR and γ. IS is adopted to scale the randomly-generated input matrix W^{in}, SR means the spectral radius of W^{res}, and γ is the leak rate of the reservoir.

In this paper, GA is used to optimize the hyper-parameters for the improvement of convergence rapidity and the prediction accuracy of stacked ESN model. The triple hyper-parameters of each reservoir are concatenated into a vector which is viewed as a population in GA. The search space for IS limited to the range [0.00001,1], which is assigned as a small value of 0.00001 to avoid IS becoming

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TTE.2021.3111906, IEEE Transactions on Transportation Electrification

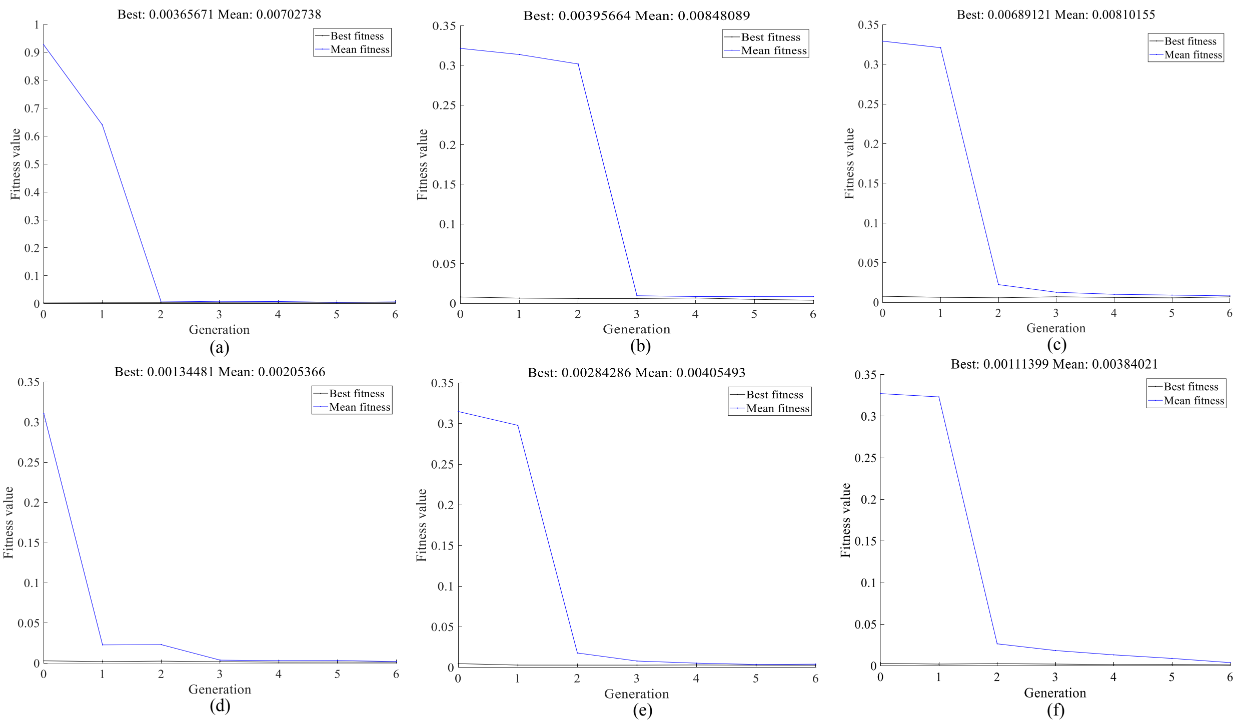

Figure 6: The process of GA optimization parameter under FC1 and FC2. (a) 2 layers under FC1. (b) 3 layers under FC1. (c) 4 layers under FC1. (d) 2 layers under FC2. (e) 3 layers under FC2. (f) 4 layers under FC2.

Table 3: The reservoir parameters optimization results.

Datasets	Layers	Parameters	Parameters of the n^{th} layer			
			1^{st} layer	2^{nd} layer	3^{rd} layer	4^{th} layer
FC1	2	IS	0.0420	0.5586		
		SR	0.8804	0.4794		
		γ	0.8660	0.8056		
	3	IS	0.5266	0.2463	0.8105	
		SR	0.4724	0.3006	0.2469	
		γ	0.8979	0.9754	0.5287	
	4	IS	0.3460	0.8878	0.8787	0.3873
		SR	0.8824	0.4607	0.3840	0.6360
		γ	0.5124	0.6602	0.9483	0.3851
FC2	2	IS	0.2024	0.6880		
		SR	0.3418	0.6012		
		γ	0.9407	0.8037		
	3	IS	0.5543	0.7584	0.5662	
		SR	0.6315	0.1770	0.4951	
		γ	0.7052	0.7637	0.5065	
	4	IS	0.4218	0.8235	0.5853	0.4733
		SR	0.9340	0.4387	0.8909	0.7572
		γ	0.7431	0.6797	0.8143	1.0000

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TTE.2021.3111906, IEEE Transactions on Transportation Electrification

zero. SR and γ are limited to (0,1). SR is set to be smaller than 1 to ensure the echo state property in each reservoir [34]. In the case of the proposed stacked ESN based on the GA algorithm, the prediction error is selected as the fitness value of population (the smaller the loss, the higher the fitness). The process of GA optimization parameters under FC1 and FC2 are respectively illustrated in Figure 6. In Figure 6 (c), the maximum number of iterations for the optimization parameters of 4 layers is only 5, indicating the GA has a fast convergence speed. Figure 6 (d) also shows that the best convergence is only 3 with 2 layers under FC2. The reservoir parameters optimization results of 2 layers, 3 layers, and 4 layers are listed in Table 3. The results show that the GA has strong convergence ability when optimizing hyper-parameters problems. In conclusion, the proposed prediction method can play a very important guiding role in system fault diagnosis and PHM.

4.3.Prognostic under constant current load conditions

In the constant current load condition testing task, the datasets come from a 1154 h duration test on the FC1 stack. No matter what the data-driven prognostics methods applied, the whole prognostics process of PEMFCs should include two phases, which are training process and prediction process. The data between 0 h and 550 h is used for training, the data between 551 h and 1154 h is used for the prediction model.

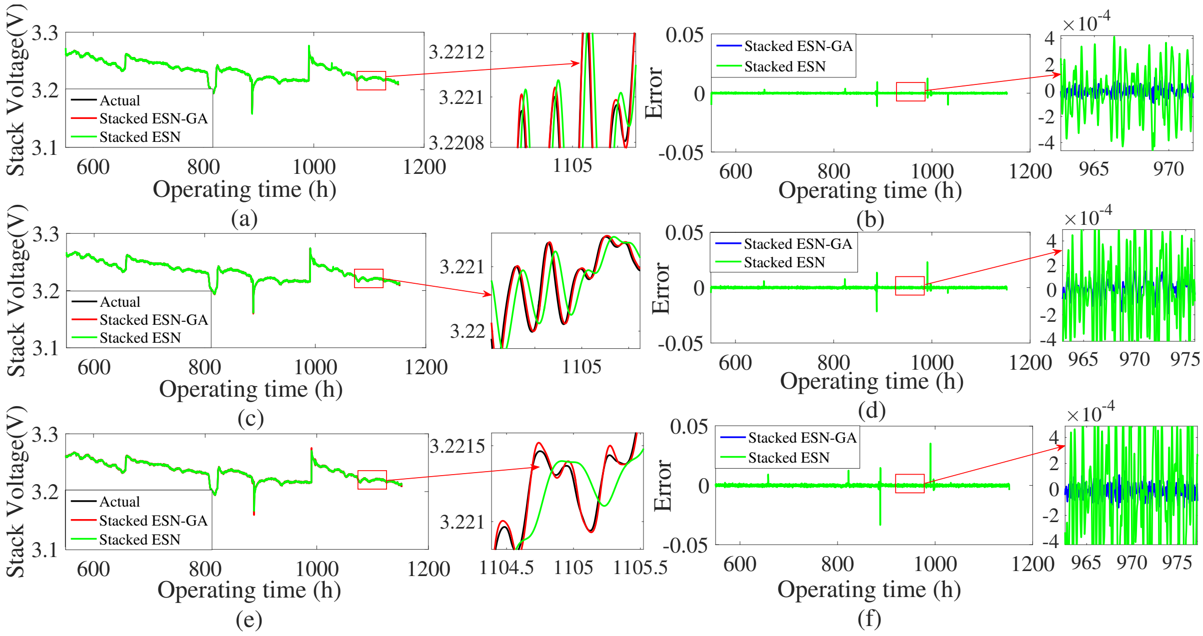

Figure 7: Prediction and error results of stacked ESN model before and after GA optimization for different layers under FC1 (1154 h). (a) The prediction results of 2 layers. (b) The error results of 2 layers. (c) The prediction results of 3 layers. (d) The error results of 3 layers. (e) The prediction results of 4 layers. (f) The error results of 4 layers.

Recently, a DL method is applied to automatically extract the degradation feature. However, the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TTE.2021.3111906, IEEE Transactions on Transportation Electrification

training of the DL method requires a large amount of aging data, which will take a long time and poses a challenge for its applications such as energy management of electric vehicles and the distributed energy generations, etc. In addition, as the number of ESN layers increases, the prediction speed of degradation trend of PEMFCs will inevitably decrease, which is difficult to apply in actual automotive systems. Therefore, a compromise solution is considered to choose 2, 3, and 4 layers to build the degradation prediction model for FC1 and FC2.

The degradation prediction results of FC1 can be obtained by the proposed stacked ESN based on the GA under different layers. As shown in Figure 7 (a), the green is the prediction phase and its length is 604 h, respectively. Figure 7 (a), (c) and (e) show the predicted results of the stacked ESN model before and after GA optimization for 2 layers, 3 layers and 4 layers, respectively. It can be clearly seen from local zoomed in Figure 7 (a), (c) and (e) that the proposed method has a high prediction accuracy. Figures 7 (b), (d) and (f) show the prediction error of stacked ESN model before and after GA optimization with different layers, which can accurately predict the general degradation trend of the real voltage in the prediction phase under different layers.

Table 4: Comparison results of degradation prediction error under FC1.

Methods	Layers	Evaluation metric	
		RMSE	MAPE
Stacked ESN based on the GA	2	**4.4711e-05**	**1.1051e-05**
	3	9.2535e-05	2.2579e-05
	4	6.1152e-05	1.5064e-06
Stacked ESN	2	2.7565e-04	5.7584e-05
	3	3.7690e-04	9.1256e-05
	4	5.1074e-04	1.2562e-04
Hua *et al.* [28]	SISO-ESN	0.01435	0.004010

The error results of the stacked ESN model before and after GA optimization with different layers are summarized in Table 4. From Table 4, the error results of the proposed stacked ESN based on the GA method for 2 layers are the smallest. The best result has been bolded in black. The RSME and MAPE values of 2 layers are 4.4711e-05 and 1.1051e-05, respectively. The RMSE value of the stacked ESN based on the GA model for 2 layers is 16.2 times that of stacked ESN model. The prediction errors of the other layers are also small, which are within an acceptable range. In addition, it can be found that the prediction accuracy of the proposed stacked ESN based on the GA outperforms the stacked ESN and traditional prediction method [28] in terms of RMSE and MAPE.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TTE.2021.3111906, IEEE Transactions on Transportation Electrification

4.4.Prognostic under dynamic operation

To verify the prediction results of the different degradation models under dynamic operation, the data comes from a 1020 h duration test on the FC2. Since the accelerated aging test time of FC2 is limited and its voltage degradation rate is significantly faster than the first aging test for FC1. Similar with the FC1, a compromise solution is considered to choose 2, 3, and 4 layers to build the RUL prediction model for FC2.

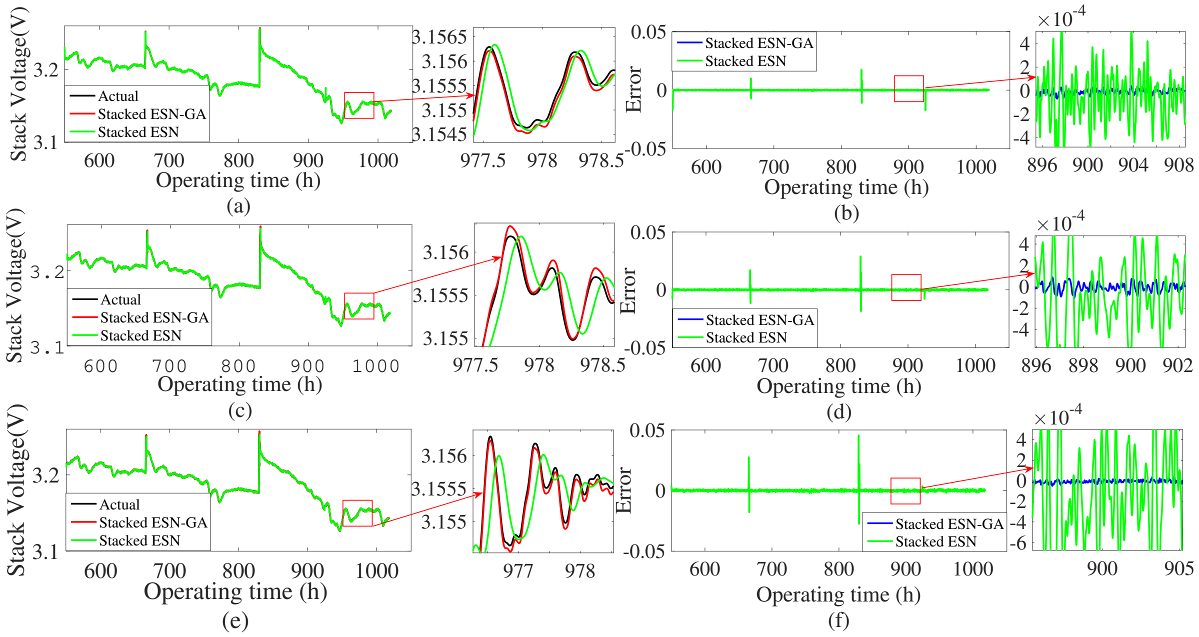

Figure 8: Prediction and error results of stacked ESN model before and after GA optimization for different layers under FC2 (1020 h). (a) The prediction results of 2 layers. (b) The error results of 2 layers. (c) The prediction results of 3 layers. (d) The error results of 3 layers. (e) The prediction results of 4 layers. (f) The error results of 4 layers.

Prediction and error results of stacked ESN model before and after GA optimization for different layers under FC2 (1020 h) are illustrated in Figure 8. The data between 0 h and 550 h is used for training, the rest of the data is used for the prediction with the green color. Figures 8 (a), (c) and (e) show the prediction results of stacked ESN model before and after GA optimization for 2 layers, 3 layers and 4 layers, respectively. Figures 8 (b), (d) and (f) show the error results of stacked ESN model before and after GA optimization for 2 layers, 3 layers and 4 layers, respectively. It can be seen from local zoomed in Figure 8 that the proposed prediction method can also accurately predict the general degradation trend of the real voltage in the prediction phase under dynamic conditions. The results in Figure 8 are highly consistent with the results in Figure 7. This is because, by using projection layers and encoding layers alternately, the proposed prediction method can provide much more robust generalization performance than the standard prediction method, and also fully take advantage of the temporal kernel property of ESN [35] to encode the multi-scale dynamics of the voltage time series.

The RMSE and MAPE assessment results are summarized in Table 5. The best result has been bolded in black. The prediction error of the 3 layers proposed method is the smallest. The RSME and MAPE values of the prediction of 3 layers proposed stacked ESN based on the GA are 1.1027e-04 and 3.0978e-05, respectively. The RMSE value of stacked ESN model after GA optimization for 3 layers is 3 times that of stacked ESN model. In addition, the prediction error results of the other layers methods are also within an acceptable range. It is obvious from the above results that the stacked ESN model after GA optimization can accurately predict the actual degradation voltage signal no matter in constant current load conditions (FC1) and dynamic conditions (FC2). It can be seen from Table 4 and Table 5 that the stacked ESN model based on the GA can significantly improve the prediction accuracy. These results also reveal that the number of different layers has an impact on the prediction accuracy. Compared with the traditional method, the results also indicate that the proposed method has a strong robustness to accurately predict the degradation trends under the variable extreme operating conditions. This is because a distinctive feature of the proposed method is its multiple projection-encoding based stacked architecture. The stacked ESN alternates between

Table 5: Comparison results of degradation prediction error under FC2.

Methods	Layers	Evaluation metric	
		RMSE	MAPE
Stacked ESN based on the GA	2	1.8081e-04	5.2952e-05
	3	**1.1027e-04**	**3.0978e-05**
	4	1.4130e-04	4.1660e-05
Stacked ESN	2	3.6550e-04	5.8857e-05
	3	3.8123e-04	9.2243e-05
	4	5.6994e-04	1.4431e-04
Hua *et al.* [28]	SISO-ESN	0.033538	0.009735

a projection layer and an encoding layer to connect the reservoirs. More specifically, when the stack voltage time series is projected into the echo-state space of a reservoir, a subsequent encoding layer receives the echo states of the previous reservoir as input and encodes the high-dimensional echo-state representations into a lower-dimensional feature space (e.g., by ELM-AE). Then these encoded representations are once again projected into the high-dimensional state space of the following reservoir by random connections. By using this multiple projection-encoding method, the stacked ESN can fully take advantage of the temporal kernel property of each reservoir to represent the multi-scale dynamics of the stack voltage time series, rather than directly stacking multiple reservoirs in an entirely random way. The experimental results illustrated that the stacked ESN based on the GA outperforms both the stacked ESN and traditional ESN models on the stack voltage time series prediction by capturing the rich multi-scale dynamics of the data.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TTE.2021.3111906, IEEE Transactions on Transportation Electrification

4.5.Prognostic under different training lengths

In order to further analyze the effectiveness of different training data lengths on the prediction accuracy of the degradation trends of PEMFCs. The training data is divided into 600 h and 700 h and 800 h to verify different prediction models, respectively. Figure 9 shows the prediction results under constant

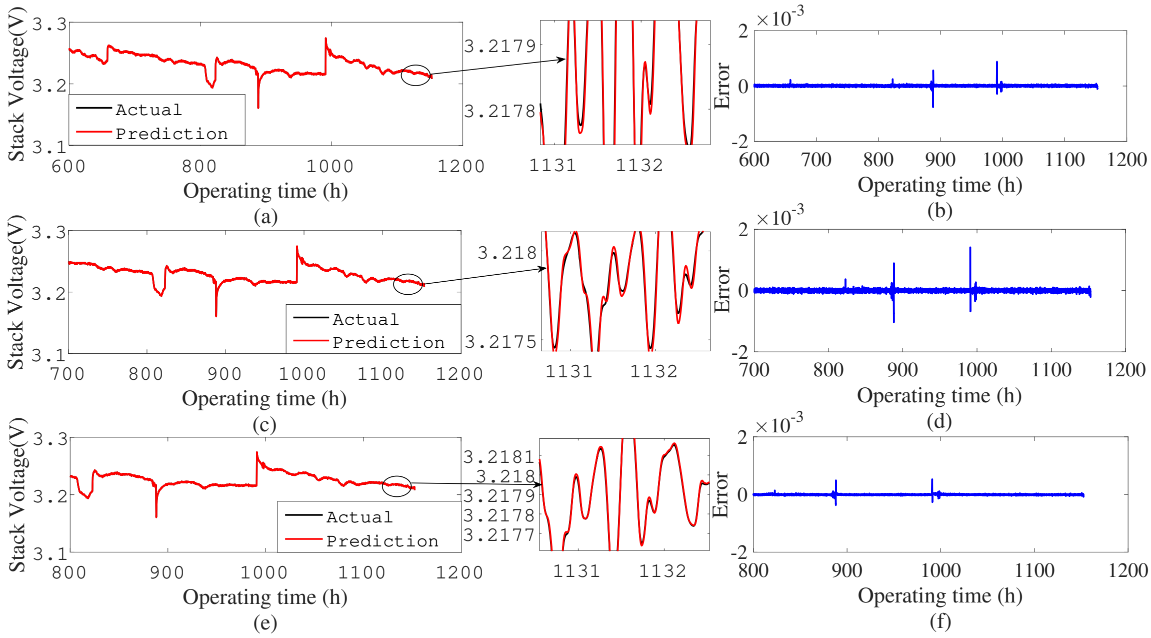

Figure 9: Stacked ESN based on the GA method for FC1 with various training length. (a) The prediction results of 600 h training length. (b) The prediction error results of 600 h training length. (c) The prediction results of 700 h training length. (d) The prediction error results of 700 h training length. (e) The prediction results of 800 h training length. (f) The prediction error results of 800 h training length.

current load. The results clearly show that the proposed model can accurately predict the voltage degradation trend under different training lengths. Among them, when the training length reaches 80%, and the prediction accuracy can achieve the highest. The RMSE and MAPE assessment results for different methods are given in Table 6. Despite training length varies, the proposed method is better than other methods with smaller RMSE and MAPE in fuel cell degradation prediction under constant current load.

In order to further verify the prediction results of different training lengths of the proposed method under dynamic operating conditions, the results of stacked ESN based on the GA method for FC2 with various training lengths are illustrated in Figure 10. Figure 10 (a), (c) and (e) present the predicting performance under 60%, 70% and 80% of the overall training degradation data, respectively. The training length is 70%, and the prediction accuracy is the highest. The prediction and error analysis for different methods under dynamic load are also given in Table 6. In a word, the degradation prediction of PEMFCs with varying training lengths, the proposed method is significantly better than

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TTE.2021.3111906, IEEE Transactions on Transportation Electrification

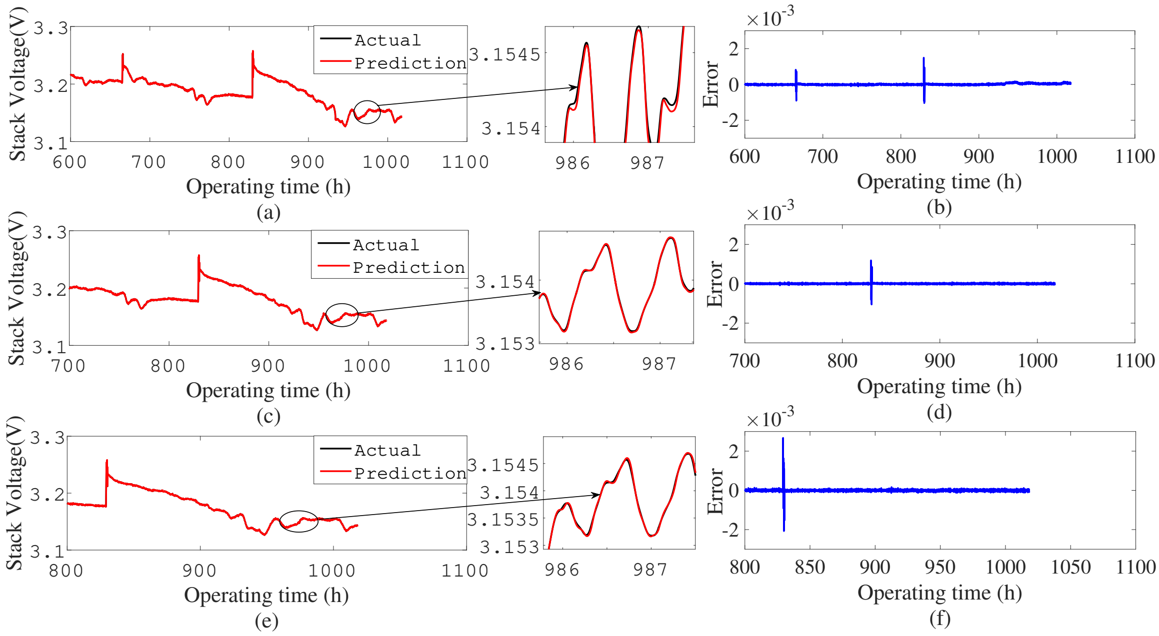

Figure 10: Stacked ESN based on the GA method for FC2 with various training length. (a) The prediction results of 600 h training length. (b) The prediction error results of 600 h training length. (c) The prediction results of 700 h training length. (d) The prediction error results of 700 h training length. (e) The prediction results of 800 h training length. (f) The prediction error results of 800 h training length.

other methods in terms of RMSE and MAPE.

Table 6: Different training lengths on the accuracy of life prediction under FC1 and FC2.

Datasets	Methods	Training Length	Evaluation metric	
			RMSE	MAPE
FC1	Stacked ESN based on the GA (2 layers)	60%	**1.9468e-05**	**4.7683e-06**
		70%	**3.3681e-05**	**8.3055e-06**
		80%	**1.4046e-05**	**3.4779e-06**
	LSTM RNN [36]	60%	0.0091	0.0027
		70%	0.0046	0.0016
		80%	0.0059	0.0021
	ARIMA [36	60%	0.0111	0.0030
		70%	0.0084	0.0022
		80%	0.0073	0.0019
	Fusion [36]	60%	0.0044	0.0011
		70%	0.0041	0.0010
		80%	0.0039	0.0010
FC2	Stacked ESN based on the GA (3 layers)	60%	**7.2543e-05**	**2.0558e-05**
		70%	**2.2239e-05**	**5.7112e-06**
		80%	**3.8044e-05**	**9.7321e-06**
	LSTM RNN [36]	60%	0.0058	0.0017
		70%	0.0054	0.0014
		80%	0.0062	0.0015
	ARIMA [36]	60%	0.0206	0.0042
		70%	0.0219	0.0042
		80%	0.0248	0.0063
	Fusion [36]	60%	0.0165	0.0036
		70%	0.0178	0.0034
		80%	0.0206	0.0051

5.Conclusion

The stack voltage degradation prediction of the PEMFCs plays an important role in its applications, such as energy management of electric vehicles and the distributed energy generations, etc. To improve the prediction performance under different operating conditions, this paper proposed a stacked ESN based on the GA to implement the degradation prediction for the PEMFCs system. As the improvement structure of standard ESN, the stacked ESN has the advantages of low computational complexity and high life prediction accuracy. The effectively and feasibility of the stacked ESN based on the GA model with different levels were verified under both constant current load conditions and dynamic operation conditions. Furthermore, the effect of different training lengths on the prediction accuracy of the models were also verified. Experimental results indicated that the proposed stacked ESN based on the GA has the ability to perform robust prediction and capture the rich multi-scale dynamic in each reservoir. Since RC was an efficient method to construct recurrent networks that required less training network and pursued conciseness and effectiveness, this was the difference from deep learning methods. Thus, there was a gap between the advantages and disadvantages of two methods. Therefore, the proposed stacked ESN based on the GA algorithm can provide a novel perspective towards bridging this gap between RC and deep learning. In future work, the proposed prediction method will be further applied to system diagnosis and PHM of electric vehicles system.

CRediT authorship contribution statement

Zhihua Deng: Investigation, Methodology, Writing-original draft, Writing-review & editing. **Qihong Chen:** Funding acquisition, Writing-review & Manuscript revision. **Liyan Zhang:** Supervision, Writing-review & Manuscript revision. **Keliang Zhou:** Manuscript revision. **Yi Zong:** Manuscript revision & Funding acquisition. **Hao Liu:** Manuscript revision. **Jishen Li:** Manuscript revision. **Longhua Ma:** Manuscript revision & Funding acquisition.

Declaration of competing interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Appendix A. An analysis of stacked ESN's computational complexity

Assuming a stacked ESN has K reservoirs and $K-1$ ELM-AE, in which all reservoirs' sizes are fixed by N, and the reduced dimensionality is $M(M<N)$. Given T-length D-dimensional input sequences, the computational complexity of stacked ESN is analyzed as follows.

The complexity at the steps of high-dimensional projection (1) and update (2) in i-th reservoir can be expressed as

$$\begin{cases} C_{res(i)} = \mathcal{O}\left(2\alpha TN^2 + 2TND\right), i = 1 \\ C_{res(i)} = \mathcal{O}\left(2\alpha TN^2 + 2TNM\right), i = 2, 3, \ldots, K. \end{cases} \tag{A.1}$$

in which the sparsity α is small (such as 0.01).

Besides, the complexity of encoding j-th states with ELM-AE mentioned before can be calculated by [32]

$$C_{enc(j)} = \mathcal{O}\left(NTM\right) \tag{A.2}$$

After updating echo states of all the layers at all the time stamps, the last reservoirs states, inputs and all the middle-layer-encoded features are collected into a matrix M with the size of $(N+(K-1)M+D)T$ which is full row rank. Solving the regression problem in (13) has the complexity

$$C_{regression} = \mathcal{O}((T + (P/3)))P^2 \tag{A.3}$$

in which $P = N + (K-1)M + D$. Since the dimension of a reservoir usually is much larger than the sizes and inputs of encoders. In this way, $C_{regression}$ can be approximately rewritten as $\mathcal{O}\left(TN^2 + N^3\right)$. Further, if N is much less than T (high dimension property of the input time series), then we can have $N \leq T$ and $C_{regression} = \mathcal{O}\left(TN^2\right)$.

Finally, the computational complexity of stacked ESN can be expressed as

$$C_{DeePr-ESN} = \sum_{i=1}^{K} C_{res(i)} + \sum_{j=1}^{K-1} C_{enc(j)} + C_{regression} \tag{A.4}$$

That is

$$C_{DeePr-ESN} \approx \mathcal{O}\left(2\alpha TKN^2 + 2TND + (K-1)\,2TNM + (K-1)\,TNM + TN^2\right) \approx \mathcal{O}(TN^2) \tag{A.5}$$

As can be seen from the above analysis, with efficient unsupervised encoders (e.g., ELM-AE) and the assumption of high dimension property of the input time series, the computational complexity of stacked ESN is $\mathcal{O}(TN^2)$. It is the training complexity of stacked ESN and the run-time burden is much smaller. In addition, a conventional single-reservoir ESN' s computational complexity can be computed by

$$C_{ESN} = C_{res} + C_{regression} \approx \mathcal{O}(2\alpha TN^2 + 2TND + TN^2) \approx \mathcal{O}(TN^2) \quad \text{(A.6)}$$

Therefore, the proposed stacked ESN can achieve equivalent computational performance to a standard ESN (single-reservoir), which shows that the proposed method remains the high computational efficiency of traditional reservoir-computing networks.

References

[1] Q. Li, T. Wang, S. Li, W. Chen, H. Liu, E. Breaz, et al, "Online extremum seeking-based optimized energy management strategy for hybrid electric tram considering fuel cell degradation," *Applied Energy*, vol. 285, pp. 116505-116516, 2021. https://doi.org/10.1016/j.apenergy.2021.116505.

[2] Z. Deng, Q. Chen, L. Zhang, Z. Fu, K. Zhou, "Weighted fusion control for proton exchange membrane fuel cell system," *International Journal of Hydrogen Energy*, vol. 45, no. 30, pp. 15327-15335, 2020. https://doi.org/10.1016/j.ijhydene.2020.01.137.

[3] S. Meraghni, L. Terrissa, M. Yue, J. Ma, S. Jemei, N. Zerhouni, "A data-driven digital-twin prognostics method for proton exchange membrane fuel cell remaining useful life prediction," *International Journal of Hydrogen Energy*, vol. 46, no. 2, pp. 2555-2564, 2021. https://doi.org/10.1016/j.ijhydene.2020.10.108.

[4] Z. Deng, Q. Chen, L. Zhang, Y. Zong, K. Zhou, Z. Fu, "Control oriented data driven linear parameter varying model for proton exchange membrane fuel cell systems," *Applied Energy*, vol. 277, pp. 115540-115548, 2020. https://doi.org/10.1016/j.apenergy.2020.115540.

[5] K. Chen, S. Laghrouche, S. Djerdir, "Degradation model of proton exchange membrane fuel cell based on a novel hybrid method," *Applied Energy*, vol. 252, pp. 113439-113448, 2019. https://doi.org/10.1016/j.apenergy.2019.113439.

[6] H. Liu, J. Chen, D. Hissel, J. Lu, M. Hou, Z. Shao, "Prognostics methods and degradation indexes of proton exchange membrane fuel cells: A review," *Renewable and Sustainable Energy Reviews*, vol. 123, pp. 109721-109743, 2020. https://doi.org/10.1016/j.rser.2020.109721.

[7] H. Liu, J. Chen, D. Hissel, H. Su, "Remaining useful life estimation for proton exchange membrane fuel cells using a hybrid method," *Applied Energy*, vol. 237, pp. 910-919, 2019. https://doi.org/10.1016/j.apenergy.2019.01.023.

[8] J. Liu, Q. Li, Y. Han, G. Zhang, W. Chen, "PEMFC residual life prediction using sparse autoencoder-based deep neural network," *IEEE Transactions on Transportation Electrification*, vol. 5, no.4, pp. 1279-1294, 2019. https://doi.org/10.1109/TTE.2019.2946065

[9] M. Ou, R. Zhang, Z. Shao, B. Li, D. Yang, P. Ming, et al, "A novel approach based on semi-empirical model for degradation prediction of fuel cells," *Journal of Power Sources*, vol.488, pp. 229435-229445, 2021. https://doi.org/10.1016/j.jpowsour.2020.229435.

[10] X. Zhang, D. Yang, M. Luo, Z. Dong, "Load profile based empirical model for the lifetime prediction of an automotive PEM fuel cell," *International Journal of Hydrogen Energy*, vol. 42, no.16, pp.11868-11878, 2017. https://doi.org/10.1016/j.ijhydene.2017.02.146.

[11] M. Ibrahim, N. Steiner, S. Jemei, D. Hissel, "Wavelet-based approach for online fuel cell remaining useful lifetime prediction," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 8, pp. 5057-5068, 2016. https://doi.org/10.1109/TIE.2016.2547358.

[12] H. Liu, Z. Liu, W. Jia, X. Lin, "Remaining useful life prediction using a novel feature-attention-based end-to-end approach," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 1197-1207, 2021. https://doi.org/10.1109/TII.2020.2983760.

[13] Y. Cheng, N. Zerhouni, C. Lu, "A hybrid remaining useful life prognostic method for proton exchange membrane fuel cell," *International Journal of Hydrogen Energy*, vol. 43, no. 27, pp. 12314-12327, 2018. https://doi.org/10.1016/j.ijhydene.2018.04.160.

[14] J. Ma, X. Liu, X. Zou, M. Yue, P. Shang, L. Kang, et al, "Degradation prognosis for proton exchange membrane fuel cell based on hybrid transfer learning and intercell differences," *ISA Transactions*, 2020. https://doi.org/10.1016/j.isatra.2020.06.005.

[15] D. Zhou, F. Gao, E. Breaz, A. Ravey, A. Miraoui, "Degradation prediction of PEM fuel cell using a moving window-based hybrid prognostic approach," *Energy*, vol. 138, pp. 1175-1186, 2017. https://doi.org/10.1016/j.energy.2017.07.096.

[16] K. Chen, S. Laghrouche, A. Djerdir, "Fuel cell health prognosis using unscented kalman filter: postal fuel cell electric vehicles case study," *International Journal of Hydrogen Energy*, vol. 44, no. 3, pp. 1930-1939, 2019. https://doi.org/10.1016/j.ijhydene.2018.11.100.

[17] M. Bressel, M. Hilairet, D. Hissel, B. Bouamama, "Extended kalman filter for prognostic of proton exchange membrane fuel cell," *Applied Energy*, vol. 164, pp. 220-227, 2016. https://doi.org/10.1016/j.apenergy.2015.11.071.

[18] F. Wang, X. Cheng, K. Hsiao, "Stacked long short-term memory model for proton exchange membrane fuel cell systems degradation," *Journal of Power Sources,* vol. 448, pp. 227591-227598, 2020. https://doi.org/10.1016/j.jpowsour.2019.227591.

[19] F. Wang, T. Mamo, X. Cheng, "Bi-directional long short-term memory recurrent neural network with attention for stack voltage degradation from proton exchange membrane fuel cells," *Journal of Power Sources,* vol. 461, pp. 228170-228183, 2020. https://doi.org/10.1016/j.jpowsour.2020.228170.

[20] J. Liu, Q. Li, W. Chen, Y. Yan, Y. Qiu, T. Cao, "Remaining useful life prediction of PEMFC based on long short-term memory recurrent neural networks," *International Journal of Hydrogen Energy*, vol. 44, no.11, pp. 5470-5480, 2019. https://doi.org/10.1016/j.ijhydene.2018.10.042.

[21] K. Chen, S. Laghrouche, A. Djerdir, "Prognosis of fuel cell degradation under different applications using wavelet analysis and nonlinear autoregressive exogenous neural network," *Renewable Energy,* vol. 179, pp. 802-814, 2021. https://doi.org/10.1016/j.renene.2021.07.097.

[22] R. Xie, R. Ma, S. Pu, L. Xu, D. Zhao, Y. Huangfu, "Prognostic for fuel cell based on particle filter and recurrent neural network fusion structure," *Energy and AI,* vol. 2, pp.100017-100029, 2020. https://doi.org/10.1016/j.egyai.2020.100017.

[23] J. Zuo, H. Lv, D. Zhou, Q. Xue, L. Jin, W. Zhou, et al, "Deep learning based prognostic framework towards proton exchange membrane fuel cell for automotive application," *Applied Energy*, vol. 281, pp. 115937-115950, 2021. https://doi.org/10.1016/j.apenergy.2020.115937.

[24] R. Ma, T. Yang, E. Breaz, Z. Li, P. Briois, F. Gao, "Data-driven proton exchange membrane fuel cell degradation predication through deep learning method," *Applied Energy*, vol. 231, pp. 102-115, 2018. https://doi.org/10.1016/j.apenergy.2018.09.111.

[25] K. Chen, S. Laghrouche, A. Djerdir, "Degradation prediction of proton exchange membrane fuel cell based on grey neural network model and particle swarm optimization," *Energy Conversion and Management*, vol. 195, pp. 810-818, 2019. https://doi.org/10.1016/j.enconman.2019.05.045.

[26] K. Chen, S. Laghrouche, A. Djerdir, "Aging prognosis model of proton exchange membrane fuel cell in different operating conditions," *International Journal of Hydrogen Energy*, vol. 45, no. 20, pp.11761-11772, 2020. https://doi.org/10.1016/j.ijhydene.2020.02.085.

[27] S. Morando, S. Jemei, D. Hissel, R. Gouriveau, N. Zerhouni, "Proton exchange membrane fuel cell aging forecasting algorithm based on echo state network," *International Journal of Hydrogen Energy* , vol. 42, no. 2, pp. 1472-1480, 2017. https://doi.org/10.1016/j.ijhydene.2016.05.286.

[28] Z. Hua, Z. Zheng, M. Péra, F. Gao, "Remaining useful life prediction of PEMFC systems based on the multi-input echo state network," *Applied Energy*, vol. 265, pp. 114791-114804, 2020. https://doi.org/10.1016/j.apenergy.2020.114791.

[29] Z. Hua, Z. Zheng, E. Pahon, M. Péra, F. Gao, "Remaining useful life prediction of PEMFC systems under dynamic operating conditions," *Energy Conversion and Management*, vol. 231, pp. 113825-113840, 2021. https://doi.org/10.1016/j.enconman.2021.113825.

[30] M. Lukoševičius, H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Computer Science Review*, vol. 3, no. 3, pp. 127-149, 2009. https://doi.org/10.1016/j.cosrev.2009.03.005.

[31] P. Ren, P. Pei, Y. Li, Z. Wu, D. Chen, S. Huang, "Degradation mechanism of proton exchange membrane fuel cell under typical automotive operating conditions," *Progress in Energy and Combustion Science*, vol. 80, pp. 100859-100902, 2020. https://doi.org/10.1016/j.pecs.2020.100859.

[32] Q. Ma, L. Shen, G. Cottrell, "DeePr-ESN: A deep projection-encoding echo-state network," *Information Sciences*, vol. 511, pp. 152-171, 2020. https://doi.org/10.1016/j.ins.2019.09.049.

[33] A. Tikhonov, V. Arsenin, "Solutions of ill-posed problems," *Mathematics of Computation*, vol. 32, no. 144. pp. 1320-1322, 1977. https://doi.org/10.2307/2006360.

[34] R. Gouriveau, M. Hilairet, D. Hissel, S. Jemei, M. Jouin, E. Lechartier, et al, "IEEE PHM 2014 data challenge: outline, experiments, scoring of results, winners," *IEEE 2014 PHM Challenge,* 2014.

[35] C. Gallicchio, A. Micheli, "Echo state property of deep reservoir computing networks," *Cognitive Computation*, vol. 6, pp. 1-14, 2017. https://doi.org/10.1007/s12559-017-9461-9.

[36] R. Ma, Z. Li, E. Breaz, C. Liu, H. Bai, P. Briois, et al, "Data-fusion prognostics of proton exchange membrane fuel cell degradation," *IEEE Transactions on Industry Applications*, vol. 1, pp.4321-4332, 2019.