

Article

Life Prediction Based on D-S ELM for PEMFC

Xuexia Zhang * , Zixuan Yu and Weirong Chen

School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China; Yuuzixuan@163.com (Z.Y.); wrchen@swjtu.edu.cn (W.C.)

* Correspondence: zhangxx@home.swjtu.edu.cn; Tel.: +86-28-8760-3332

Received: 15 August 2019; Accepted: 28 September 2019; Published: 30 September 2019



Abstract: The proton exchange membrane fuel cell (PEMFC) is an extremely clean and efficient power generation device. However, its limited lifespan has restricted the large-scale commercial development of PEMFCs. Life prediction is a promising solution for the further life extension of PEMFCs. In this paper, D-S ELM(DWT-SaDE ELM), define as, an enhanced extreme learning machine (ELM) optimized by discrete wavelet transform (DWT) and self-adaptive differential evolutionary algorithm (SaDE), is proposed to predict the remaining useful life (RUL) of PEMFCs. In D-S ELM, DWT is employed to extract available features from multi-input data with stochastic noise. Then, SaDE explores the optimal parameter configuration for the ELM neural network. Moreover, the influence of training data sizes on the prediction results is discussed. Simulations show that D-S ELM has obvious advantages in prediction accuracy. Furthermore, the superiority of D-S ELM in small sample applicability, prediction speed and robustness make it more suitable for the online prediction of PEMFCs.

Keywords: Proton exchange membrane fuel cell (PEMFC); remaining useful life (RUL); extreme learning machine (ELM); discrete wavelet transform (DWT); self-adaptive differential evolutionary (SaDE)

1. Introduction

Energy issue is the hottest topic in the 21st century. Hydrogen energy is attracting more and more attention due to its high efficiency, cleanliness and sustainability. As a prospective power supply device, the proton exchange membrane fuel cell (PEMFC) has been widely applied in many fields such as substation, cars, buses, locomotives, trams and portable devices [1–3]. However, the performance of each single cell will be affected by catalyst poisoning, membrane drying, water flooding, current spillover and other problems, resulting in temporary or permanent voltage degradation. In that case, the remaining useful life of fuel cells is greatly shortened, which greatly hinders the large-scale application of fuel cell technology in the industrial field. Prognostics and health management (PHM) [4], especially the life prediction of PEMFC, is a challenging issue to be addressed.

At present, life prediction methods can be mainly divided into three categories including model-driven methods, data-driven methods and hybrid-driven methods.

Model-driven methods are realized by an analysis model based on non-linear physical phenomena. Due to the lack of an adaptive aging model and the disturbance caused by the recovery phenomenon, Marine Jouin et al. [5] adopted a particle filter (PF) algorithm to predict power degradation based on an empirical model. Taejin Kim et al. [6] proposed a method based on an equivalent circuit model (ECM) and the training electrochemical impedance spectroscopy (EIS) data to predict the remaining useful life of PEMFC. Elodie Lechartier et al. [7] proposed a behavioral model of PEMFC consisting of the static model based on the Butler-Volmer equation and the dynamic model based on the electrical equivalent, respectively, to describe the physical phenomena of PEMFC. Xinfeng Zhang et al. [8] proposed an

empirical model by estimating the degradation rate of output voltage. The effect of load distribution on voltage degradation can be analyzed by the characteristic values, which are obtained by analyzing the aging process of PEMFC.

Considering the difficulty and inaccuracy of physical modeling, data-driven methods have been developed rapidly. However, a large amount of historical data is required during prediction. Mona Ibrahim et al. [9] proposed a data driven method based on an auto-regressive integrated moving average model (ARIMA) using discrete wavelet transform (DWT). The experimental results show that the prediction error is less than 3%. R.E. Silva et al. [10] proposed a scheme to divide the non-periodic degenerate voltage into normal degradation and external disturbance. Only the normal degradation part can be used to train an adaptive neuro-fuzzy inference system (ANFIS) and the error caused by external disturbances is eliminated. Hao Liu et al. [11] employed wavelet analysis to decompose an original voltage curve into sub-waveforms. Therefore, it contributes to reducing the disturbances from the original data. Simon Morando et al. [12,13] applied an echo state network (ESN) to predict the aging process of PEMFC. ESN can fast forecast lifetime through modeling a linear regression neural network to express a complex nonlinear system. Rui Ma et al. [14] presented a grid long short-term memory (G-LSTM) neural network which can achieve an accurate voltage degradation curve when solving complex prediction problems. Kamran Javed et al. [15,16] explored a summation wavelet-extreme learning machine (SW-ELM) by optimizing the initialized parameters including random weights and bias, which can improve the prediction accuracy and robustness of remaining useful life (RUL) prediction.

Hybrid-driven methods combine the advantages of model-driven and data-driven methods. Nevertheless, complex calculation is a disadvantage of hybrid-driven methods. Zhou Daming et al. [17] introduced a hybrid-driven method. An empirical degradation model and nonlinear auto-regressive neural network (NARNN) are fused by weight factors which are inversely proportional to the prediction residuals. A moving window algorithm is applied to increase the prediction accuracy. Yujie Cheng et al. [18,19] presented a hybrid method based on least square support vector machine (LSSVM) and regularized particle filter (RPF). LSSVM has good prediction performance in processing large sample data and RPF improves the accuracy by solving the problem of particle diversity loss. The proposed hybrid algorithm uses LSSVM to get a preliminary prediction, then the uncertainty of the prediction results is quantified by RPF.

From the above review, both model-driven methods and hybrid-driven methods require accurate physical models. Nevertheless, the complicated physical phenomena in PEMFCs are difficult to precisely model. During the operation process of PEMFCs, a large amount of data can be recorded, thus, the data-driven prediction method has great advantages. However, the traditional neural network algorithm needs a large number of iterations in training and testing processes, which will take a long time for prediction. Previous data-driven algorithms (e.g., ESN and LSTM etc.) focus on improving the prediction accuracy. These algorithms with complex network structures and long prediction time cannot be applied on-line. ELM can solve this problem due to its fast training and testing speed. As a consequence, this paper proposes D-S ELM, which improves the prediction accuracy of extreme learning machine (ELM) while retaining the superiority of speed to predict the RUL of PEMFCs, which lays a foundation for its online real-time prediction.

2. Experimental Data

2.1. Raw Experimental Data

The experimental data used in this paper are provided by FCLAB (Fuel Cell Lab) in 2014 IEEE Fuel Cell Data Challenge Competition [20]. In the test bench shown in Figure 1, the fuel cell stack consists of five single cells with an effective area of 100 cm². The control range of experimental physical parameters is shown in Table 1.

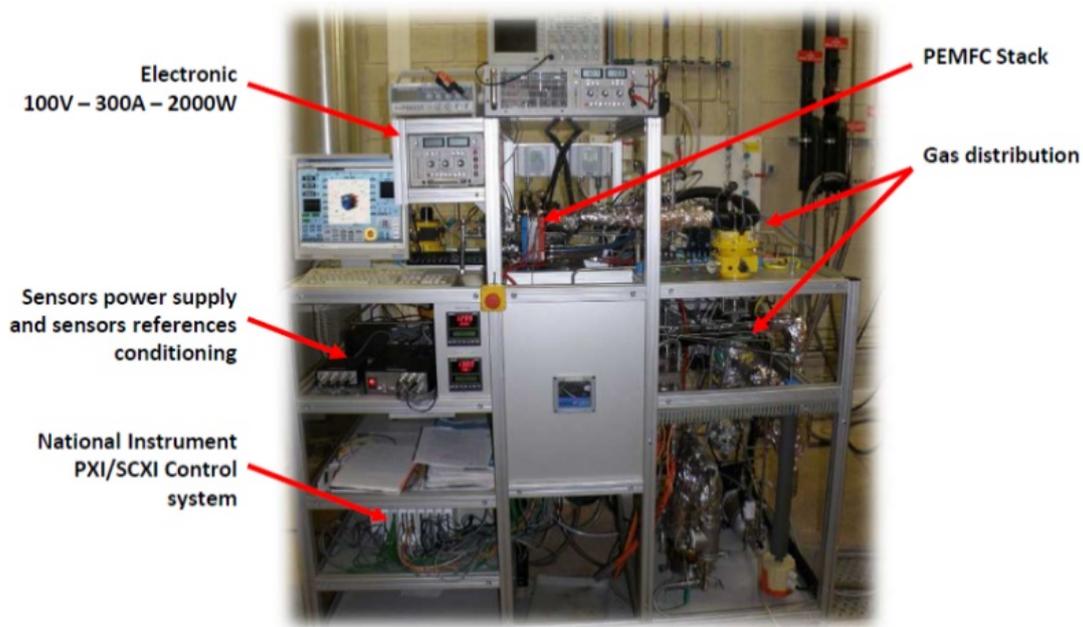


Figure 1. Proton exchange membrane fuel cell (PEMFC) test bench provided by the 2014 IEEE Fuel Cell Data Challenge Competition [20].

Table 1. Control range of experimental physical parameters provided by the 2014 IEEE Fuel Cell Data Challenge Competition [20].

| Physical Parameters | Control Range |
|---------------------------|---------------|
| Cooling water temperature | 20 °C–80 °C |
| Cooling water flowrate | 0–10 L/min |
| Gas temperature | 20 °C–80 °C |
| Gas humidity | 0%–100% |
| Air flowrate | 0–100 L/min |
| Hydrogen flowrate | 0–100 L/min |
| Gas pressure | 0–2 bars |
| Fuel cell current | 0–300 A |

In this paper, RUL prediction is considered under a constant current condition, in this case, FC1 data set which is observed under constant current stable environment $J = 0.7\text{A}/\text{cm}^2$ is chosen as raw experimental data.

To accurately predict the RUL of PEMFC, the selection of indicators characterizing the aging process is the primary problem. The curves of candidate indicators in the FC1 data set are shown in Figure 2. We can see that the candidate indicators (e.g., outlet hydrogen temperature, outlet water temperature and outlet hydrogen pressure) cannot reflect the aging phenomena of PEMFC from the curves of these indicators. Compared with the above indicators, the output voltage has a distinct downward trend, which can be used as an aging indicator.

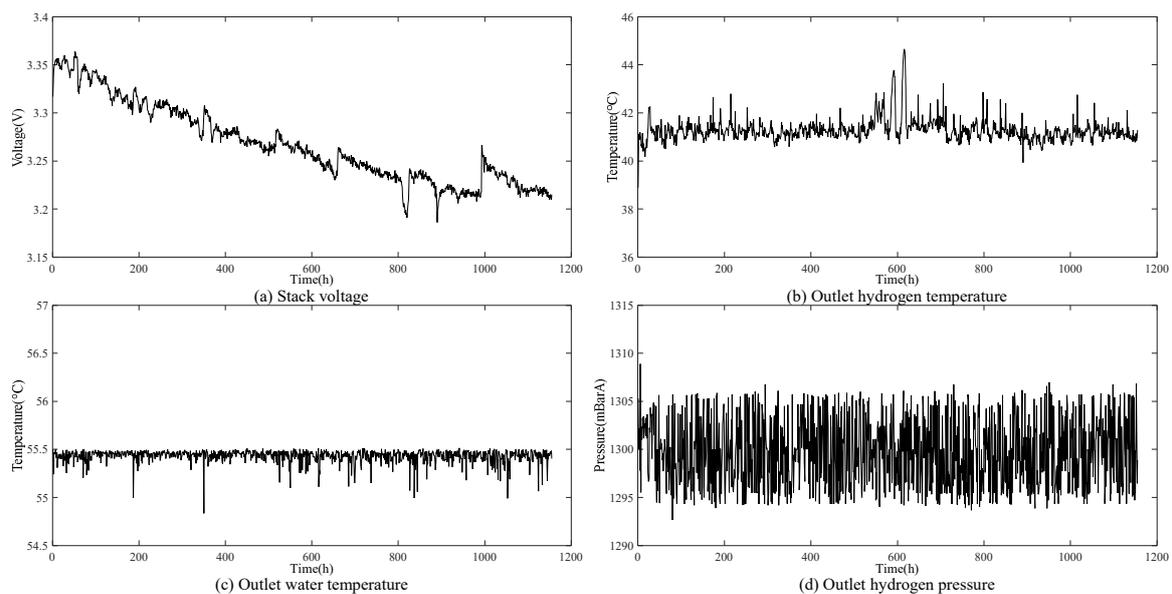


Figure 2. FC1 data curves.

The characteristics of experimental data can be summarized as follows:

- (1) The size of monitoring data is large. For better analysis, data preprocessing is applied.
- (2) The data set only includes the operation parameters (e.g., inlet hydrogen pressure) and the output parameters (e.g., voltage), but not the internal aging parameters (e.g., catalyst active surface area) of PEMFC. Considering the difficulty of establishing an accurate physical model, the data-driven method is a promising solution.
- (3) The data set has the characteristic of non-linear and non-Gauss disturbance, further data tracking in time series contributes to ensuring the accuracy of prediction.

2.2. Data Preprocessing

In this paper, the data preprocessing method includes data reduction and normalization. The purpose of data reduction is to simplify the original huge data without loss of data properties. The specific operation is to extract data from the original data with a step size of 100. To unify multiple physical quantities, data normalization is applied. The operation parameters and output parameters are unified in the range of [0,1].

3. D-S ELM Methodology

3.1. Discrete Wavelet Transform

Discrete wavelet transform (DWT) can filter out noise and retain useful information of the time series [21]. Basic wavelet translation and scaling operations give the wavelet transform good local characteristics in both time and frequency domains, which is an important feature of the wavelet transform. Wavelet translation can obtain the time-domain information of time series and wavelet scaling can obtain the frequency-domain information of time series. The main process of DWT is decomposition and reconstruction. The decomposition process is to first determine a basic wavelet, decompose the time series into N-layer sub-wavelets and then get the approximate coefficients and detail coefficients. Figure 3 gives a brief introduction of decomposition hierarchically.

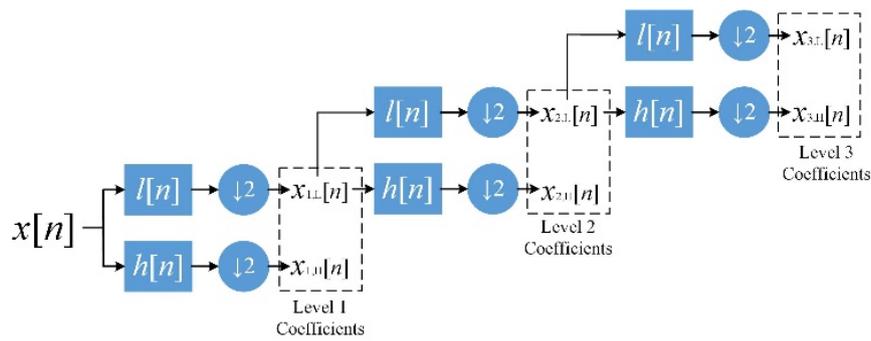


Figure 3. Decomposition process of discrete wavelet transform (DWT).

For a given discrete input signal $x[n]$, a low-pass filter $l[n]$, a high-pass filter $h[n]$ and a down-sampling filter $\downarrow Q$ (set $Q = 2$), $x_{\alpha,L}[n]$ and $x_{\alpha,H}[n]$ are the approximate coefficients and the detail coefficients in α^{th} layer.

$$x_{\alpha,L}[n] = \sum_{k=0}^{N-1} x_{\alpha-1,L}[2n-k]l[k] \tag{1}$$

$$x_{\alpha,H}[n] = \sum_{k=0}^{N-1} x_{\alpha-1,L}[2n-k]h[k] \tag{2}$$

The reconstruction process reconstructs the original time series by the approximate coefficients and detail coefficients. For further understanding, Figure 4 shows the decomposition and reconstruction process of DWT, meanwhile, the approximate coefficients and detail coefficients of the original signal can be obtained. After decomposition, the original signal, whose time domain length is L , is divided into approximate coefficients and detail coefficients whose length is $L/2$. In this paper, the experimental data provided by FCLAB is decomposed by using one level wavelet transform. The basic wavelet is "db3" wavelet.

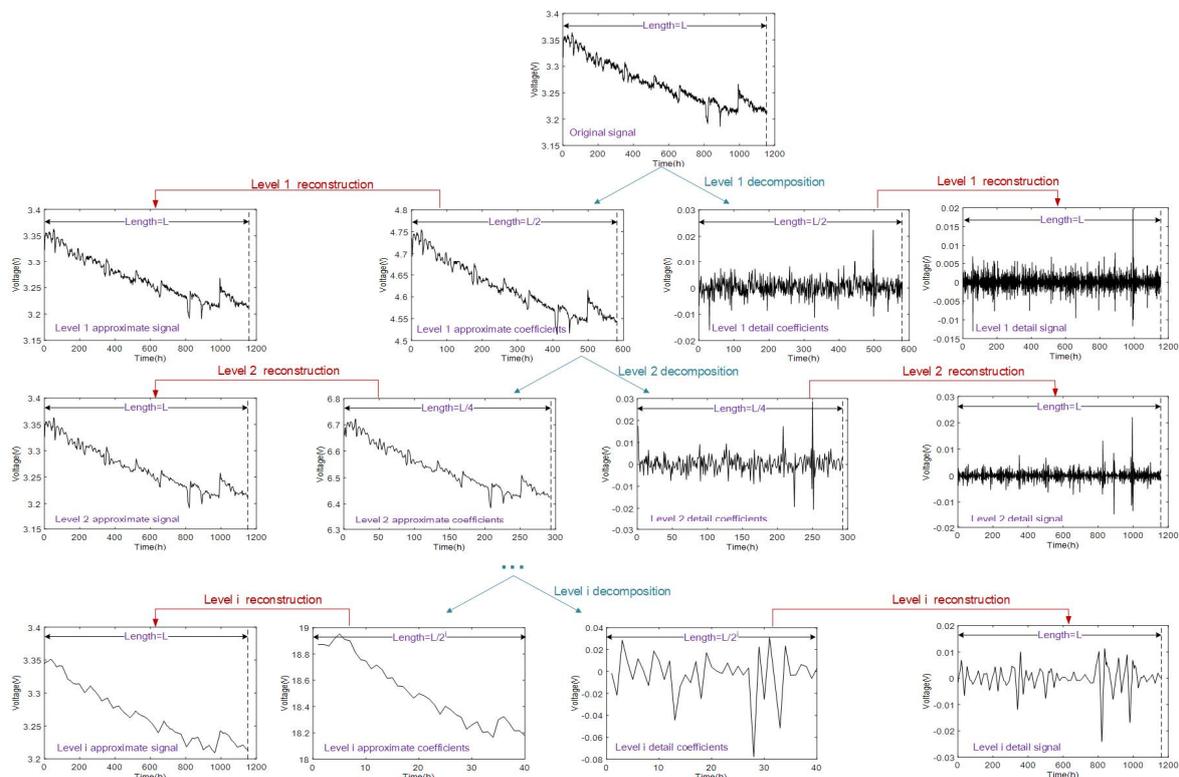


Figure 4. Decomposition and reconstruction process of DWT.

3.2. Self-Adaptive Differential Evolutionary Algorithm

Since the output weights of ELM [22] are calculated based on the random input weights w_i and stochastic hidden layer bias b_i , a set of non-optimal or unnecessary input weights and hidden layer bias exist. In some applications, ELM may need more hidden neurons than the traditional algorithm [23]. Therefore, the algorithm is expected to get more compact and respond faster to the neural network in practical applications.

Self-adaptive differential evolutionary algorithm [24] (SaDE) is a methodology available to optimize the neural network of ELM. According to the above analysis, the optimized objectives are the two randomly selected parameters including the input weight w_i and hidden layer bias b_i . First, the randomly selected parameters are used as the initial population. Then the fitness of each individual in the population is calculated. In this paper, the root mean square error (RMSE) was chosen as the fitness function. When the fitness calculation of all individuals was completed, the process of mutation, crossover and selection was carried out and the individuals with better fitness will be retained for the next generation.

3.3. D-S ELM

This paper proposes D-S ELM, which is optimized by discrete wavelet transform (DWT) and a self-adaptive differential evolutionary algorithm (SaDE). The framework of D-S ELM is shown in Figure 5.

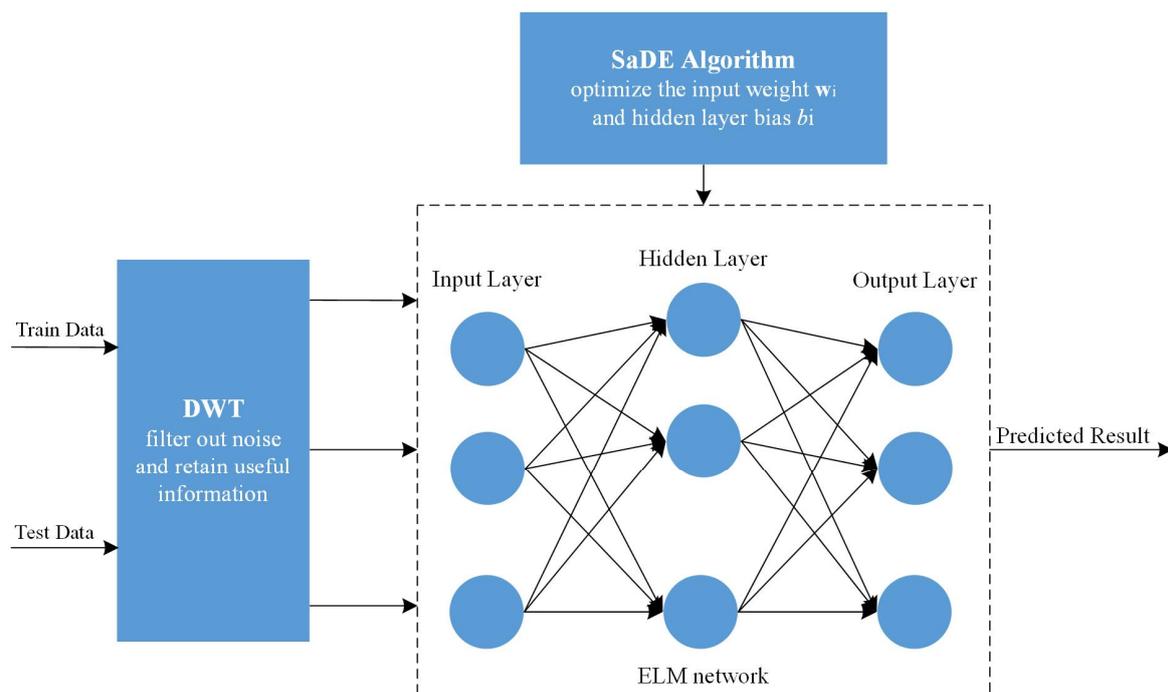


Figure 5. Remaining useful life (RUL) prediction framework based on D-S ELM.

The main concept and steps of D-S ELM are summarized as follows.

Step 1. Discrete wavelet transform

Use discrete wavelet transform (DWT) to transform the training data and test data.

Step 2. ELM initialization

Set the number of hidden layer nodes N and the activation function $g(x)$. Original ELM can be approximated to N zero-error samples, which is expressed as the following equations.

$$\sum_{i=1}^N \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{p}_j, j = 1, \dots, N \quad (3)$$

Equation (3) can be written succinctly.

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{P} \quad (4)$$

In Equation (3), \mathbf{w}_i and $b_i (i = 1, \dots, N)$ is the input weight and the bias of hidden layer. $\boldsymbol{\beta}$ is the output weight, and \mathbf{x}_j and $\mathbf{p}_j (j = 1, \dots, N)$ is the training input and training output. In Equation (4), \mathbf{H} is the output matrix of hidden layer and \mathbf{P} is the training output matrix.

Step 3. SaDE initialization

Set and the maximum number of iterations G_{max} . The population $\boldsymbol{\theta}_{k,G}$ with NP individuals is initialized.

$$\boldsymbol{\theta}_{k,G} = \left[\mathbf{w}_{1,(k,G)}^T, \dots, \mathbf{w}_{N,(k,G)}^T, b_{1,(k,G)}^T, \dots, b_{N,(k,G)}^T \right] \quad (5)$$

where \mathbf{w}_i and $b_i (i = 1, \dots, N)$ are randomly selected. G is the current generation and $k = 1, 2, \dots, NP$.

Step 4. Calculations of hidden layer output matrix H .

$$\mathbf{H}(\mathbf{w}_1, \dots, \mathbf{w}_N, b_1, \dots, b_N, \mathbf{x}_1, \dots, \mathbf{x}_N) = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \dots & g(\mathbf{w}_N \cdot \mathbf{x}_1 + b_N) \\ \vdots & \dots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \dots & g(\mathbf{w}_N \cdot \mathbf{x}_N + b_N) \end{bmatrix}_{N \times N} \quad (6)$$

According to Equation (3) and Equation (4), the i^{th} column of matrix \mathbf{H} is the i^{th} hidden neuron output vector response to training input $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$.

Step 5. Calculations of output weights $\boldsymbol{\beta}$ and RMSE

$$\boldsymbol{\beta}_{k,G} = \mathbf{H}_{k,G}^+ \mathbf{P} \quad (7)$$

$$\text{RMSE}_{k,G} = \sqrt{\frac{\sum_{j=1}^N \|\sum_{i=1}^L \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) - \mathbf{p}_j\|_2^2}{N}} \quad (8)$$

Step 6. Mutation

$$\mathbf{v}_{i,G} = \boldsymbol{\theta}_{r_1^i,G} + F \cdot (\boldsymbol{\theta}_{r_2^i,G} - \boldsymbol{\theta}_{r_3^i,G}) \quad (9)$$

$$\lambda = e^{1 - \frac{G_{max}}{1 + G_{max} - G}} \quad (10)$$

$$F = F_0 \cdot 2^\lambda \quad (11)$$

$\mathbf{v}_{i,G}$ is a new generated mutant population, where r_1^i, r_2^i, r_3^i are randomly generated in the population. Parameter F is an adaptive scaling factor which is usually chosen between $[0,2]$. In this paper, F is set as a value that changes with the number of iterations.

Step 7. Crossover

$$u_{i,G}^j = \begin{cases} \mathbf{v}_{i,G}^j & \text{if } \text{rand}(0,1) \leq CR \\ \boldsymbol{\theta}_{i,G}^j & \text{Otherwise} \end{cases} \quad (12)$$

Crossover process can increase the diversity of the population. CR is the crossover probability which is usually chosen between $[0,1]$.

Step 8. Selection

$$\theta_{k,G+1} = \begin{cases} u_{k,G+1} & \text{if } RMSE_{u_{k,G+1}} < RMSE_{\theta_{k,G}} \\ \theta_{k,G} & \text{Otherwise} \end{cases} \quad (13)$$

The individuals with lower fitness value will be retained for the next generation. Repeat step 4 to step 8 until $G = G_{max}$.

Step 9. Calculations of test results

Input the test data, calculate the predicted results according to β .

4. Life Prediction Based on D-S ELM

4.1. Selection of Input and Output Layer Data

From the analysis in Section 2, only the output stack voltage of PEMFC has an obvious downward trend compared with other indicators in the original data set, which is suitable for predicting the remaining useful life of PEMFC. Therefore, the output voltage was chosen as the data of output layer network. The number of input network layers is determined by the number of factors which affect the prediction results. After data analysis, outlet water temperature, air pressure and hydrogen pressure were used as data for the input layer network. These three parameters directly affect the performance of PEMFC and the RUL of the fuel cell. Compared with BP and deep learning algorithms, in this paper, the proposed algorithm used fewer historical data to predict the current state of the PEMFC system, and can also reflect the real-time changes of the state of the PEMFC system over time.

4.2. Selection of the Number of Hidden Layer Nodes

The choice of the number of hidden layer nodes is a sensitive issue [25], which has a great influence on the prediction performance of a neural network. In order to make the D-S ELM network have a better generalization performance, the number of hidden layer nodes determined should be as few as possible to make the network more compact. When determining the number of hidden layer nodes, the following preconditions should be satisfied:

The number of hidden layer nodes should be less than that of the training samples, otherwise the network will have no practical value and low generalization ability. Similarly, the number of input layer neurons should be less than that of output layer neurons. In this paper, the number of hidden layer nodes is determined to be 3, which has best performance in simulations.

4.3. Selection of Activation Function

Activation function plays an important role in the neural network because of the ability of nonlinearity capture, which can introduce non-linear factors to the neuron model. In that case, the neural network can approximate non-linear function arbitrarily. The input data is non-linearly transformed by activation function, and then is transmitted to the next neuron as input layer data. Here enumerate some common activation functions, such as sigmoid function, tanh function and hardlim function. In this paper, sigmoid function was used as the activation function. The sigmoid function is generally used as the threshold function of artificial neural network, which maps variables to $[0,1]$.

4.4. Simulation Results Based on Two Algorithms

After the parameters of ELM network are set up, the number of input training data should also be determined to obtain the best prediction results. As mentioned earlier, one of the advantages of D-S ELM over traditional algorithms is that it needs fewer training data. The influence of training data size

on prediction accuracy is studied and discussed in Section 4.4.3. Therefore, to study the prediction accuracy improvements of the proposed algorithms, the prediction curves and comparison results given below are in the case of training data size $n = 10$.

One of the simulation results is shown in Figure 6. The error curves of two prediction methods are shown in Figure 7. As can be seen from the Figure 6, the predicted voltage curve based on ELM fluctuates greatly, which limits the on-line prediction of RUL. On the basis of ELM, D-S ELM is more stable and closer to the expected output and the improvement of prediction accuracy can also be seen from Figure 7.

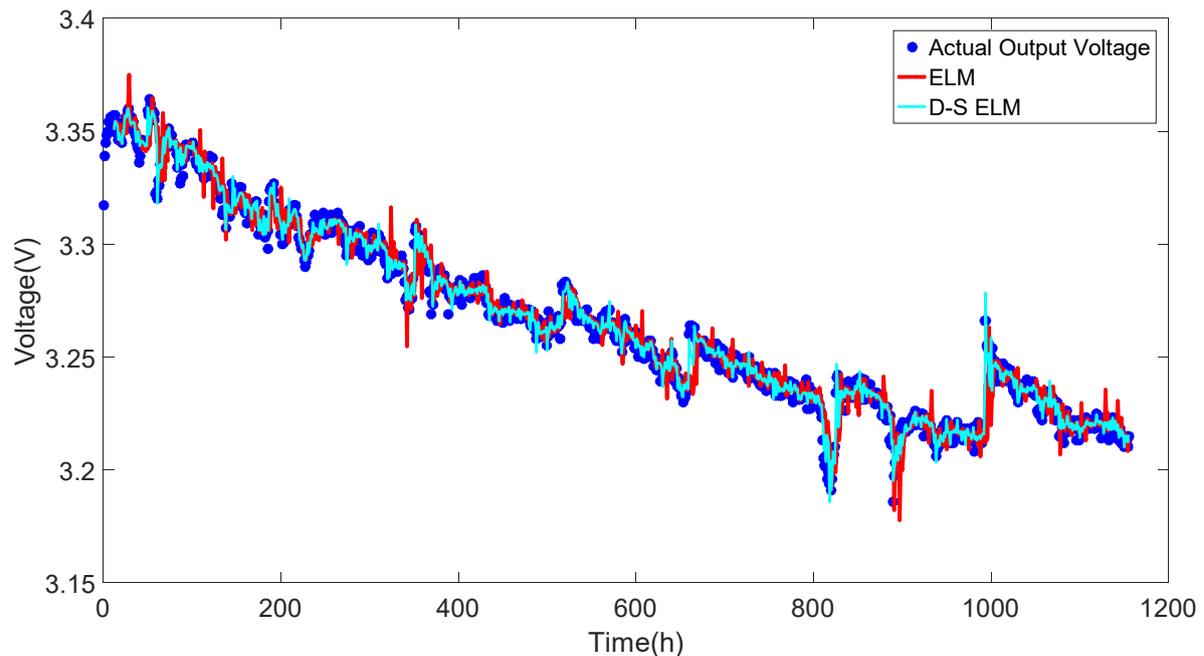


Figure 6. Prediction results of extreme learning machine (ELM) and D-S ELM (the size of training sample is set as 10).

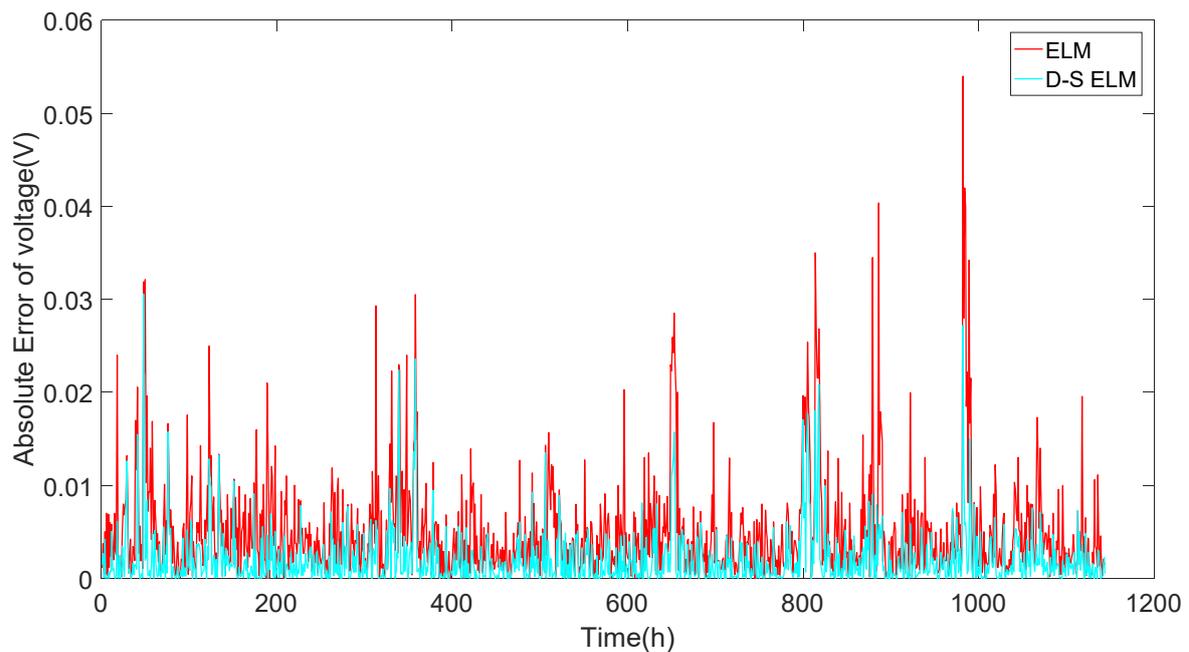


Figure 7. Error curves of ELM and D-S ELM (the size of training sample is set as 10).

4.4.1. Research on Robustness of D-S ELM

In addition, to study the robustness of the algorithm, the accuracy of life-end-point prediction was considered. Because data-driven algorithms have the characteristic of randomness, each simulation is different, which leads to the inaccuracy of prediction. In that case, use the two algorithms mentioned above to make 100 simulation predictions and obtain the end-of-life probability distribution.

As a promising energy converter, PEMFC is expected to ensure its power supply capability for a certain time in practical applications. Due to the aging phenomenon, the voltage drops and the power output is insufficient, which leads to the issue of RUL prediction. The data set used in this paper was under constant current $J = 0.7A/cm^2$, and voltage drop was used as an indicator of life end. In order to verify the accuracy and universality of the proposed algorithm, and according to the definition of 2014 IEEE Fuel Cell Data Challenge Competition [20], this paper considered a variety of voltage losses as the life end point—3.5%, 4%, 4.5% of initial voltage.

As can be seen from Figure 8, when the total output voltage reaches the failure threshold, the fuel cell will reach the end-of-life stage. The initial voltage of PEMFC test bench was 3.359 V, and the voltage failure threshold was set to 3.2418 V, 3.225 V, 3.2082 V, respectively. The actual end-of-life point was 627 h, 809 h, 811 h.

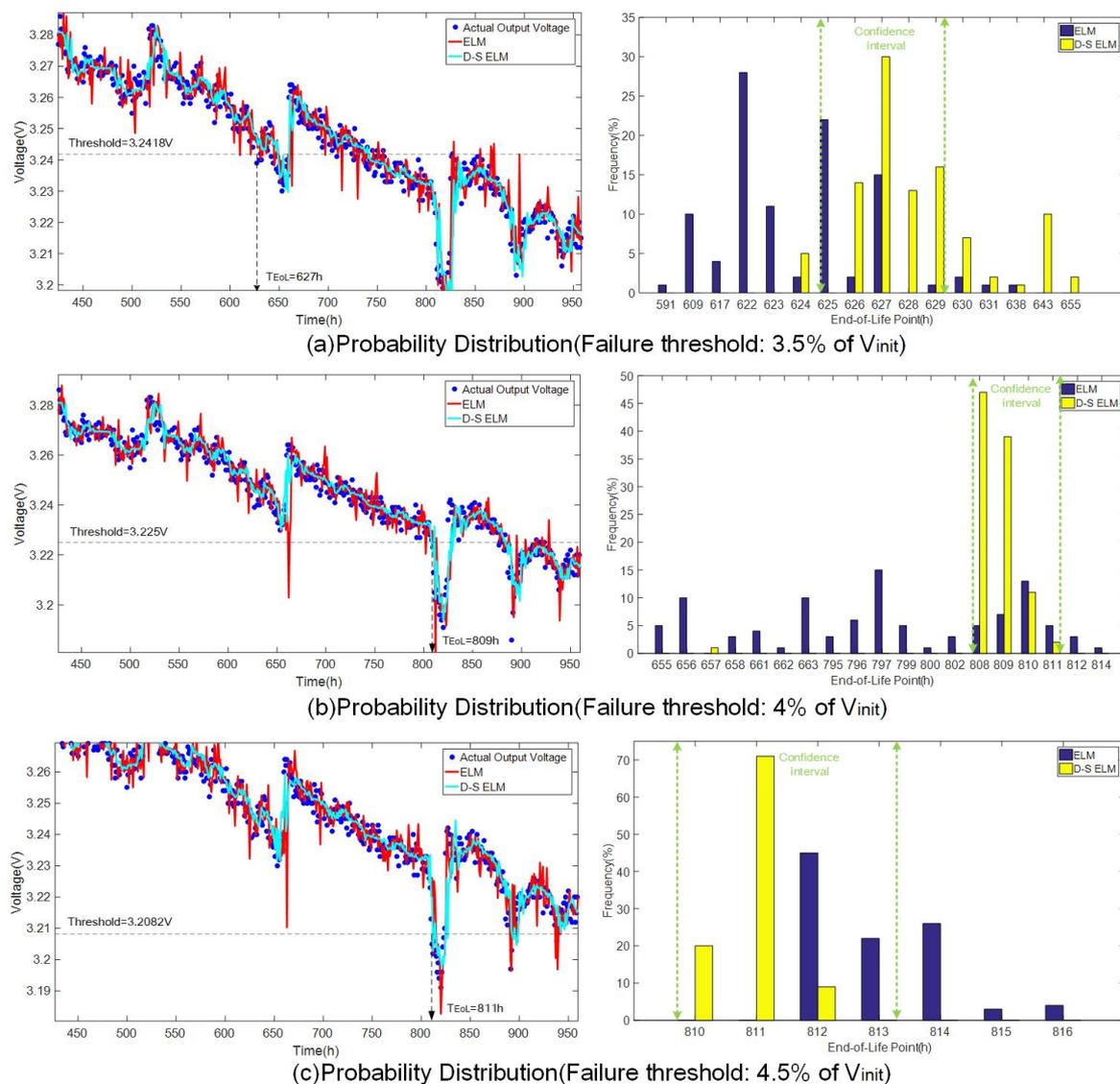


Figure 8. Probability distribution of end-of-life point (the size of training sample is set as 10).

After simulations, the predicted results were compared with the actual voltage curve to find the earliest point which reach the failure threshold. The probability distribution of end-of-life point is shown in Figure 8.

Failure threshold is set to 3.5% of initial voltage

The actual end-of-life point is 627 h. Set $(627 \text{ h} \pm 2 \text{ h})$ as the confidence interval. The probability of the predicted life end point based on ELM in the confidence interval is 39%, while the probability of D-S ELM in the confidence interval reaches 73%. As can be seen from Figure 8a, the fluctuation of ELM between 600 hours and 640 hours is large, leading to the predicted point ahead of the actual point, and the prediction time point is scattered.

Failure threshold is set to 4% of initial voltage

The actual end-of-life point is 809 h. Set $(809 \text{ h} \pm 2 \text{ h})$ as the confidence interval. The probability of the predicted life end point based on ELM in the confidence interval is 30%, the probability of D-S ELM in the confidence interval reaches 99%. This phenomenon happens because the set failure threshold is very close to the first trough of the output voltage curve, at a position of about 660 h. When the set failure threshold is confusing, the instability of ELM will lead to inaccurate prediction results.

Failure threshold is set to 4.5% of initial voltage

The actual end-of-life point is 811 h. Set $(811 \text{ h} \pm 2 \text{ h})$ as the confidence interval. The probability of the predicted life end point based on ELM in the confidence interval is 67%, and the probability of D-S ELM in the confidence interval reaches 100%. As for the voltage threshold set in Figure 8c, the predicted time points of ELM and D-S ELM are concentrated due to the significant voltage drop, and the accuracy is improved accordingly.

The accuracy of the algorithm is different for different voltage failure thresholds. However, according to three probability distribution figures, D-S ELM closely matched the expected voltage curve, and the prediction results of the proposed D-S ELM were more accurate than those of ELM. According to the above analysis, compared with ELM, the predicted end-of-life point of D-S ELM is more robust.

4.4.2. Comparison of Predicted Results

For a better comparison of final predicted results and further study, the following experiments are proposed on the basis of 4% voltage loss.

To measure the accuracy of prediction results, RMSE and ME are defined as follows.

$$\text{ME} = \frac{\sum_{j=1}^N \sum_{i=1}^K \beta_i g(w_i \cdot x_j + b_i) - \mathbf{p}_j}{N} \quad (14)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{j=1}^N \|\sum_{i=1}^K \beta_i g(w_i \cdot x_j + b_i) - \mathbf{p}_j\|_2^2}{N}} \quad (15)$$

Two algorithms are compared according to the predicted end-of-life point, deviation, root mean square error, mean error and predicted time of the algorithm. The comparison results are shown in Table 2. The actual end-of-life point is 809 h. Due to the fluctuation of each simulation, the average end-of-life deviations were obtained from the previous 100 runs of simulations. The average end-of-life point of ELM and D-S ELM is 756 h, 807.16 h, and the average time deviation is 53 h and 1.84 h, respectively.

As can be seen from the Table 2, the proposed D-S ELM has the better accuracy, with the prediction results achieving a RMSE of 0.0036 and a ME of 0.0020. Although the application of self-adaptive differential evolutionary algorithm increased the prediction time, the training time and testing time is still much smaller than that of traditional algorithms, which is enough for online prediction.

Table 2. Comparison of prediction results of two algorithms (the size of training sample is set as 10).

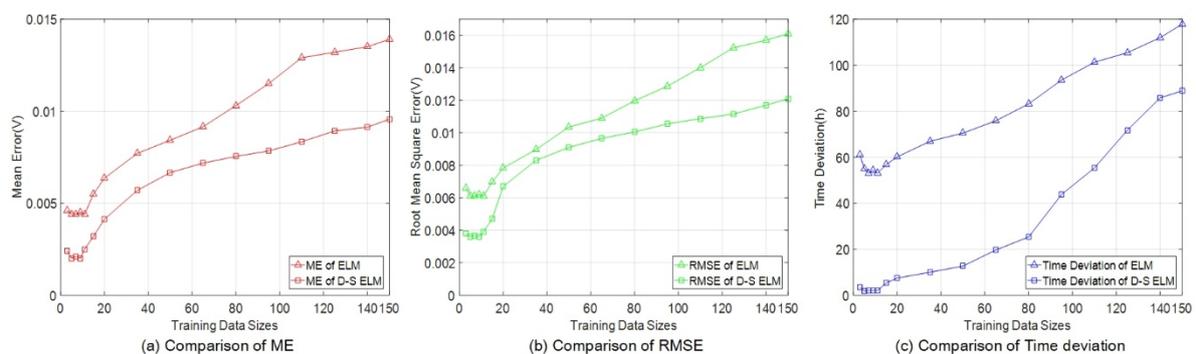
| Prediction Performance Indicators | ELM | D-S ELM |
|-----------------------------------|---------------|---------|
| End-of-life Point (h) | 756 | 807.16 |
| Deviation (h) | 53 | 1.84 |
| ME (V) | 0.0044 | 0.0020 |
| RMSE (V) | 0.0061 | 0.0036 |
| Training Time (s) | ≤ 0.0156 | 0.0781 |
| Testing Time (s) | ≤ 0.0156 | 0.0313 |

4.4.3. Influence of Training Data Size

Online RUL prediction requires high accuracy and fast speed. For that reason, it requires algorithms to use as little experimental data as possible to get results quickly and accurately. In the actual operation of the PEMFC system, the changes of environmental factors or human factors will lead to real-time voltage fluctuation. The data far from the current time cannot represent the current state of the system and has a negative impact on the mathematical model of the neural network. This phenomenon has also been confirmed in Reference [17].

In this paper, to get the most accurate predictive output, constantly changing data is used to update the mathematical model of the neural network in ELM and D-S ELM. The data sizes are set from 3 to 150 to find the appropriate training data size for PEMFC. The comparisons of ME, RMSE and the time deviation of two algorithms are shown in Figure 9. In order to see the influence of data sizes on prediction results more intuitively, five groups of predicted voltage output curves of D-S ELM and prediction result comparisons are shown in Figure 10.

As the data size increases from 3 to 150, the prediction error and time deviation of both ELM and D-S ELM increases. The ME of D-S ELM increases from 0.0020 to 0.00956, the RMSE increases from 0.0036 to 0.0121, and the time deviation increases from 1.84 h to 88.79 h. When the size of the training sample is set between [5,10], the time deviation and prediction error is smallest. As can be seen from Figure 10, with the data size increasing, the predicted voltage gradually deviates from the expected voltage. When the training data size is set to 150, only a roughly decline curve can be obtained, and the RUL deviation reaches 88.79 h. This phenomenon also confirms the assumption mentioned earlier. Compared with traditional BP algorithm, D-S ELM has little requirement for the number of data samples, which is more suitable for a small sample for online RUL prediction.

**Figure 9.** Comparisons of different training data sizes.

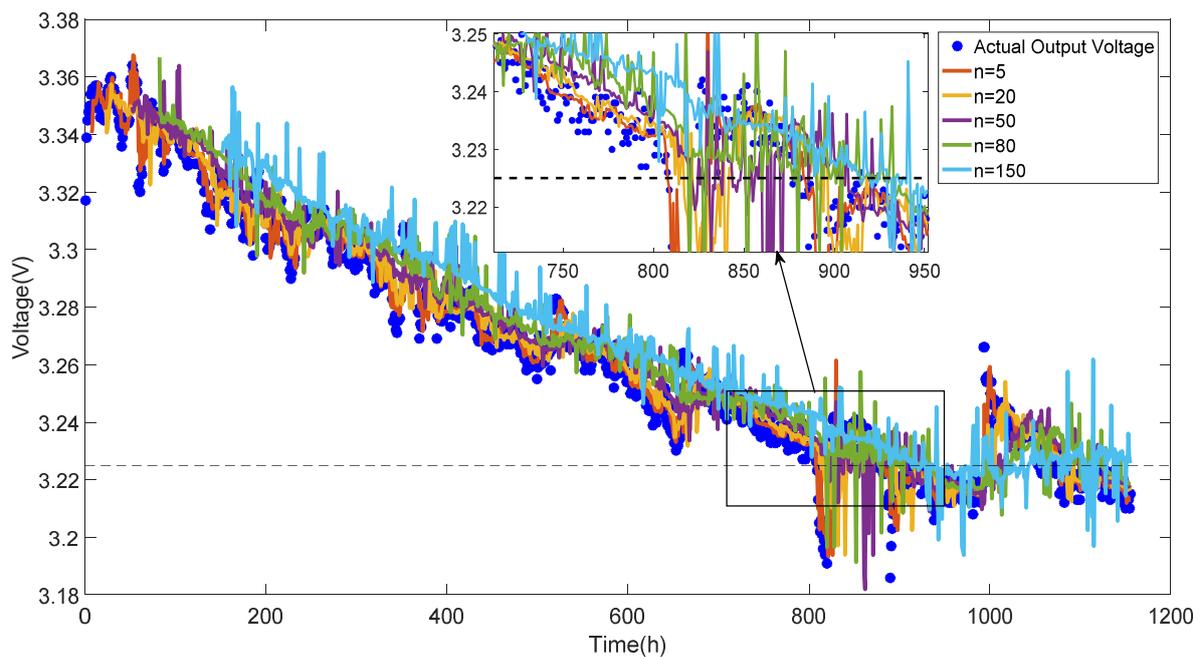


Figure 10. Prediction results of different training data sizes.

5. Conclusions

This paper proposes D-S ELM to predict the remaining useful life of PEMFC. In the proposed algorithm, the discrete wavelet transform and self-adaptive differential evolutionary algorithm are applied to optimize ELM. The proposed D-S ELM is compared with the original ELM, using the data set provided by FCLAB. The results show that D-S ELM has best performance.

The main advantages of the proposed algorithm can be summarized as follows:

- (1) The proposed D-S ELM algorithm has strong robustness. In order to study the uncertainties of data-driven algorithm, 100 simulations provide the RUL probability distribution of different algorithms. D-S ELM has best stability and smallest RUL deviation.
- (2) The proposed algorithm has little requirement for the size of training samples which is suitable for small sample prediction.
- (3) D-S ELM overcomes the shortcomings of a traditional neural network, such as a large amount of calculation and slow convergence speed.
- (4) D-S ELM improves the prediction accuracy of ELM, with the predicted voltage achieving an RMSE of 0.0036 and a ME of 0.0020.

The above research provides a basis for data-driven life prediction of PEMFC. However, the data set used in this research is under constant current conditions, while in practical applications, PEMFC usually works under variable loading conditions. Therefore, the next step of this study is to realize the RUL prediction of PEMFC with a variable load, which is a huge challenge but has a certain practical significance for prolonging the life of PEMFC and improving its reliability.

Author Contributions: Conceptualization, X.Z.; Funding acquisition, X.Z. and W.C.; Investigation, X.Z.; Methodology, X.Z.; Software, Zi.Y.; Supervision, X.Z.; Writing—original draft, Z.Y.; Writing—review & editing, X.Z. and W.C.

Funding: This research was funded by the National Natural Science Foundation (No. 51607149) and Department of Science and Technology of Sichuan Province (No. 2019YJ0236), Foundation of Key Laboratory of Magnetic Suspension Technology and Maglev Vehicle, Ministry of Education.

Acknowledgments: This work is supported by the National Natural Science Foundation (No. 51607149), Department of Science and Technology of Sichuan Province (No. 2019YJ0236), Foundation of Key Laboratory of Magnetic Suspension Technology and Maglev Vehicle, Ministry of Education.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sharaf, O.Z.; Orhan, M.F. An overview of fuel cell technology: Fundamentals and applications. *Renew. Sustain. Energy Rev.* **2014**, *32*, 810–853. [[CrossRef](#)]
2. Jahnke, T.; Futter, G.; Latz, A.; Malkow, T. Performance and degradation of proton exchange membrane fuel cells: State of the art in modeling from atomistic to system scale. *J. Power Sources* **2016**, *304*, 207–233. [[CrossRef](#)]
3. Breaz, E.; Gao, F.; Miraoui, A.; Tirnovan, R. A short review of aging mechanism modeling of proton exchange membrane fuel cell in transportation applications. In Proceedings of the Conference of the IEEE Industrial Electronics Society, Dallas, TX, USA, 29 October–1 November 2014; pp. 3941–3947.
4. Jouin, M.; Gouriveau, R.; Hissel, D.; P'era, M.C.; Zerhouni, N. Prognostics and health management of PEMFC—state of the art and remaining challenges. *Int. J. Hydrogen Energy* **2013**, *38*, 15307–15317. [[CrossRef](#)]
5. Jouin, M.; Gouriveau, R.; Hissel, D.; P'era, M.C.; Zerhouni, N. Prognostics of PEM Fuel Cell in a Particle Filtering Framework. *Int. J. Hydrogen Energy* **2014**, *39*, 481–494. [[CrossRef](#)]
6. Kim, T.; Kim, H.; Ha, J.; Kim, K.; Youn, J.; Jung, J.; Youn, B.D. A degenerated equivalent circuit model and hybrid prediction for state-of-health of PEM fuel cell. In Proceedings of the International Conference on Prognostics and Health Management, Cheney, WA, USA, 22–25 June 2014.
7. Lechartier, E.; Laffly, E.; P'era, M.C.; Gouriveau, R. Proton exchange membrane fuel cell behavioral model suitable for prognostics. *Int. J. Hydrogen Energy* **2015**, *40*, 8384–8397. [[CrossRef](#)]
8. Zhang, X.F.; Yang, D.J.; Luo, M.H.; Dong, Z.M. Load profile based empirical model for the lifetime prediction of an automotive PEM fuel cell. *Int. J. Hydrogen Energy* **2017**, *42*, 11868–11878. [[CrossRef](#)]
9. Ibrahim, M.; Steiner, N.Y.; Jemei, S.; Hissel, D. Wavelet-based approach for online fuel cell remaining useful lifetime prediction. *IEEE Trans. Ind. Electron.* **2016**, *63*, 5057–5068. [[CrossRef](#)]
10. Silva, R.E.; Gouriveau, R.; Jemei, S.; Hissel, D.; Boulon, L.; Agbossou, K.; Steiner, N.Y. Proton exchange membrane fuel cell degradation prediction based on adaptive neuro-fuzzy inference systems. *Int. J. Hydrogen Energy* **2014**, *39*, 11128–11144. [[CrossRef](#)]
11. Liu, H.; Chen, J.; Hou, M.; Shao, Z.G.; Ye, S.H. Data-based short-term prognostics for proton exchange membrane fuel cells. *Int. J. Hydrogen Energy* **2017**, *42*, 20791–20808. [[CrossRef](#)]
12. Morando, S.; Jemei, S.; Gouriveau, R.; Zerhouni, N.; Hissel, D. Fuel cells remaining useful lifetime forecasting using echo state network. In Proceedings of the 39th Annu. Conference of the IEEE Industrial Electronics Society, Coimbra, Portugal, 27–30 October 2014; pp. 1632–1637.
13. Morando, S.; Jemei, S.; Hissel, D.; Gouriveau, R.; Zerhouni, N. Proton exchange membrane fuel cell ageing forecasting algorithm based on echo state network. *Int. J. Hydrogen Energy* **2017**, *42*, 1472–1480. [[CrossRef](#)]
14. Ma, R.; Yang, T.; Breaz, E.; Li, Z.L.; Briois, P.; Gao, F. Data-driven proton exchange membrane fuel cell degradation prediction through deep learning method. *Appl. Energy* **2018**, *231*, 102–115. [[CrossRef](#)]
15. Javed, K.; Gouriveau, R.; Zerhouni, N.; Hissel, D. Improving accuracy of long-term prognostics of PEMFC stack to estimate remaining useful life. In Proceedings of the IEEE International Conference on Industrial Technology, Seville, Spain, 17–19 March 2015.
16. Javed, K.; Gouriveau, R.; Zerhouni, N. SW-ELM: A summation wavelet extreme learning machine algorithm with a priori parameter initialization. *Neurocomputing* **2014**, *123*, 299–307. [[CrossRef](#)]
17. Zhou, D.M.; Gao, F.; Breaz, E.; Ravey, A.; Miraoui, A. Degradation prediction of PEM fuel cell using a moving window based hybrid prognostic Approach. *Energy* **2017**, *138*, 1175–1186. [[CrossRef](#)]
18. Cheng, Y.J.; Zerhouni, N.; Lu, C. A hybrid remaining useful life prognostic method for proton exchange membrane fuel cell. *Int. J. Hydrogen Energy* **2018**, *43*, 12314–12327. [[CrossRef](#)]
19. Cheng, Y.J.; Zerhouni, N.; Lu, C. A prognostic framework for PEMFC based on least squares support vector regression particle filter. In Proceedings of the IEEE Vehicle Power and Propulsion Conference, Belfort, France, 11–14 December 2017.
20. FCLAB Research. *IEEE PHM 2014 Data Challenge*; FC Lab: Belfort, France, 2014.
21. Vetterli, M.; Herley, C. Wavelets and filter banks—theory and design. *IEEE Trans. Signal Process.* **1992**, *40*, 2207–2232. [[CrossRef](#)]

22. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: A new learning machine of feedforward neural networks. *Neural Netw. Conf. Proc.* **2004**, *2*, 985–990.
23. Cao, J.W.; Lin, Z.P.; Huang, G.B. Self-adaptive evolutionary extreme learning machine. *Neural Process. Lett.* **2012**, *36*, 285–305. [[CrossRef](#)]
24. Qin, A.K.; Huang, V.L.; Suganthan, P.N. Differential evolution algorithm with strategy adaptation for global numerical optimization. *IEEE Trans. Evol. Comput.* **2009**, *13*, 398–417. [[CrossRef](#)]
25. Zhao, W.T.; Li, P.; Liu, Q.; Liu, D.; Liu, X.W. Selecting the optimal hidden layer of extreme learning machine using multiple kernel learning. *Ksii Trans. Internet Inf. Syst.* **2018**, *12*, 5765–5781.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).