applied sciences

MDPI

Article

Statistical Analysis on Random Matrices of Echo State Network in PEMFC System's Lifetime Prediction

Zhiguang Hua [1,2,3,*], Zhixue Zheng [3,4], Marie-Cécile Péra [2,3] and Fei Gao [2,3]

[1] School of Automation, Northwestern Polytechnical University, Road Dongxiang, Xi'an 710129, China
[2] FEMTO-ST Institute, Université Bourgogne Franche-Comté, UTBM, CNRS, Rue Ernest Thierry Mieg, 90010 Belfort, France; marie-cecile.pera@univ-fcomte.fr (M.-C.P.); fei.gao@utbm.fr (F.G.)
[3] FCLAB, Université Bourgogne Franche-Comté, UTBM, CNRS, Rue Ernest Thierry Mieg, 90010 Belfort, France; zhixue.zheng@univ-lorraine.fr
[4] LMOPS Laboratory, Université de Lorraine & CentraleSupélec, 2 Rue Edouard Belin, 57070 Metz, France
* Correspondence: zhiguang.hua@utbm.fr

Abstract: The data-driven method of echo state network (ESN) has been successfully used in the proton exchange membrane fuel cell (PEMFC) system's lifetime prediction area. Nevertheless, the uncertainties of the randomly generated input and internal weight matrices in ESN have not been reported yet. In view of this, an ensemble ESN structure is proposed in this paper to analyze the effects of random matrices from a statistical point of view. For each ESN, the particle swarm optimization (PSO) method is utilized to optimize the hyperparameters of the leaking rate, spectral radius, and regularization coefficient. The statistical results of each ensemble ESN are analyzed from 100 repeated tests whose weight matrices are generated randomly. The mean value of the ensemble ESN and a confidence interval (CI) of 95% are given during the long-term lifetime prediction. The effects of two different distribution shapes, i.e., uniform distribution and Gaussian distribution, are fully compared. Finally, the effects of the ensemble structure and two different distribution shapes are tested by three experimental datasets under steady-state, quasi-dynamic, and full dynamic operating conditions.

Keywords: proton exchange membrane fuel cell; degradation; prognostic; data-driven; echo state network; reservoir computing; random matrices

check for updates

Citation: Hua, Z.; Zheng, Z.; Péra, M.-C.; Gao, F. Statistical Analysis on Random Matrices of Echo State Network in PEMFC System's Lifetime Prediction. *Appl. Sci.* **2022**, *12*, 3421. https://doi.org/10.3390/app12073421

Academic Editors: Jasna Jankovic and Nada Zamel

Received: 25 February 2022
Accepted: 25 March 2022
Published: 28 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Energy and environment are the focal points attracting the attention of the international community due to the depletion of fossil fuels and the increasing emissions of carbon dioxide [1,2]. Hydrogen (H_2) is considered to be a promising energy carrier of sustainable clean energy in the 21st century. The proton exchange membrane fuel cells (PEMFC) system can transform the chemical energy of H_2 into electricity directly, and the by-products during the electrochemical process are heat and water [3,4]. It is a promising power conversion device because of its satisfactory characteristics except for the high cost and limited lifetime. Predicting the lifetime of the PEMFC system can help the users to take action in advance to extend its working life.

Model-based and data-driven methods are two typical classifications of existing lifetime prediction methods [5,6]. Data-driven methods, which are mainly based on machine learning (ML) theory, construct the nonlinear degradation relations [7,8]. One of the patterns of ML, the recurrent neural network (RNN), has shown great power in nonlinear time series prediction. Nevertheless, the RNN has some weaknesses such as slow convergence rate and the existence of bifurcation. To overcome the weaknesses of RNN, the echo state network (ESN) was proposed by Prof. Jaeger et al. in [9]. There are two characteristics of ESN: one is a large, randomly generated reservoir that is used to replace the hidden layers of RNN; the other is the weight matrices (input and internal) that are generated randomly. The ESN has a faster convergence rate and lower computational training cost than the

traditional RNN. Nevertheless, the hyperparameters of ESN (leaking rate, spectral radius, regularization coefficient, reservoir neuron number, reservoir connectivity, etc.) need to be carefully tuned from human expertise [10].

For the PEMFC system's lifetime prediction application, the ESN has first been applied in [11] to estimate the steady-state cell voltage by the single-step ahead prediction pattern. Four kinds of reservoirs are compared in [12], and the delay line reservoir (DLR) and the simple cycle reservoir (SCR) have the best precision values. A cycle reservoir with a jump (CRJ) model has been proposed in [13], and this structure is validated on the steady-state datasets. The CRJ model can speed up the linear fitting process and improve prediction accuracy. In ref. [14], the steady-state cell voltages are divided into the approximation part and detail part by wavelet transform, and then double ESNs with different dynamic characteristics are utilized to deal with the voltage signal separately. This double ESN structure could improve prediction accuracy. The hyperparameters of ESN in [11–14] are set by the trial-and-error method. Inspired by the separation concept in [14], the wavelet transform and multiple ESNs are proposed in [15] to deal with the multi-timescale degradation features under the dynamic operating conditions in the long-term prediction horizon. In ref. [15], the hyperparameters of the spectral radius and leaking rate are optimized by the grid-search method, and the prediction accuracy has been improved. To improve the prediction efficiency, the original degradation data prediction is replaced by the shortened coefficients in the discrete wavelet transform process [16]. The prediction data points can be shortened by a factor of 8, and the hyperparameters of ESN (spectral radius, leaking rate, and regularization coefficient) are optimized by the genetic algorithms (GA) method. In ref. [17], the analysis of variance (ANOVA) method is applied to analyze the four hyperparameters of the reservoir neuron number, reservoir connectivity, spectral radius, and leaking rate under the steady-state dataset. Results show that the spectral radius has the biggest effects on the prediction performance. Additionally, based on the ANOVA method, the effects of three hyperparameters (i.e., leaking rate, spectral radius, and regularization coefficient) have been analyzed under the dynamic operating conditions [18]. Results show that the interaction of the spectral radius and regularization parameter has the biggest effect, and the effects of the spectral radius rank second.

Based on the ANOVA analysis results of [17], a multi-reservoir ESN structure is proposed in [19], and each reservoir has a different spectral radius. This multi-reservoir ESN can avoid the optimization of the spectral radius, and its effectiveness is validated under the steady-state cell voltage prediction condition. As an extension of the work in [19], the ESN and Markov chains are combined to predict the cell voltage under the dynamic micro-combined heat and power (μ-CHP) load profile [20]. The RUL prediction by the combined structure can be realized with a mean relative accuracy below 17%. An ensemble ESN network has recently been proposed to avoid hyperparameter optimization and improve the adaptiveness of lifetime prediction [21]. A dynamic health indicator of "virtual steady-state voltage (VSV)" is proposed based on linear parameter varying (LPV) models. Then, a number of ESNs with different combinations of hyperparameters (leaking rate and spectral radius) are used for the VSV prediction. Finally, the long-term RUL prediction and a confidence interval of 95% confidence level are given in detail. In the past two years, the multi-input ESN structure is proposed to improve the prediction accuracy of the RUL [18,22–24]. The effects of the operating parameters (i.e., current, temperature, and pressure) are fully analyzed under the steady-state and dynamic operating conditions in [18,22] separately. The multi-input ESN structure could improve the prediction accuracy regarding the health indicator and the scheduled current as the double inputs are more meaningful in practice. In ref. [23], the ambient temperature and stack voltage are used as the double inputs of ESN to improve the RUL prediction accuracy. The normalized root mean square error is equal to 0.098 with a prediction over 2000 h. In ref. [24], the input parameters of ESN are evaluated and selected by the least absolute shrinkage and selection operator (LASSO), and then the degradation-related parameters are used for the

RUL prediction. Results show that the ESN with eight input parameters is the optimal choice. Both [23] and [24] are validated by the steady-state datasets.

In the current state-of-art, the developments of ESN for the PEMFC system's lifetime prediction are focused on the ESN structure (reservoir forms, ensemble ESN, multi-input ESN, etc.) and hyperparameter optimization. In the practical implementation process of ESN, the input and internal weight matrices ($\mathbf{W}_{in}$ and $\mathbf{W}$) remain unchanged after they have been drawn from a random distribution. Nevertheless, the effects of the distribution shape of $\mathbf{W}_{in}$ and $\mathbf{W}$ on the lifetime prediction accuracy have not been reported yet. There is abundant evidence that the ensemble structure can improve the prediction robustness and stability such as the ensemble neural network model [25], particle filtering [26], and extreme learning machine [27]. The ensemble ESN is proposed in [21], and the effects of the leaking rate and spectral radius have been fully analyzed. Nevertheless, the statistical property of the ESN, e.g., the uncertainty caused by the random matrices, has not been analyzed yet. Based on these two motivations, the contributions of this paper are summarized as follows:

1. The effects of two distribution shapes (uniform and Gaussian distribution) of $\mathbf{W}_{in}$ and $\mathbf{W}$ on the prediction accuracy are explored;
2. The metaheuristic technique of particle swarm optimization (PSO) is utilized to optimize the hyperparameters of the leaking rate, spectral radius, and regularization coefficient;
3. The uncertainty of the ensemble ESN caused by the random matrices is statistically analyzed under three different operating conditions.

2. Mathematical Backgrounds

2.1. Echo State Network

Similar to the structure of RNN, there are three parts in the ESN: input layer, reservoir, and output layer. The different parts are connected by the weight matrix elements of input weight matrix $\mathbf{W}_{in} \in \mathbb{R}^{N\times K}$, internal weight matrix $\mathbf{W} \in \mathbb{R}^{N\times N}$, feedback weight matrix $\mathbf{W}_{fb} \in \mathbb{R}^{N\times L}$, and output weight matrix $\mathbf{W}_{out} \in \mathbb{R}^{L\times(N+K)}$. K is the dimension of the input signal, N is the number of neurons in the reservoir, and L is the dimension of the output signal. The spectral radius ρ is the maximum eigenvalue of $\mathbf{W}$. The $\mathbf{W}_{fb}$ is set to be $\mathbf{0}$ in this task. At step n, the neuron update in the reservoir is determined by

$$\tilde{\mathbf{x}}(n) = f(\mathbf{W}_{in}\cdot\mathbf{u}(n) + \mathbf{W}\cdot\mathbf{x}(n)) \tag{1}$$

where $\tilde{\mathbf{x}}(n)$ is the neuron update, $\mathbf{u}(n)$ is the input signal at step n, $\mathbf{x}(n)$ is the neuron state at step n, and f is the activation function. Then, the neuron state at step $(n+1)$ can be expressed as

$$\mathbf{x}(n+1) = (1-\alpha)\cdot\mathbf{x}(n) + \alpha\cdot\tilde{\mathbf{x}}(n) \tag{2}$$

where α is the leaking rate of the reservoir. The ESN output is represented by

$$\mathbf{y}(n+1) = \mathbf{W}_{out}\cdot(\mathbf{x}(n+1);\mathbf{u}(n)) \tag{3}$$

where $\mathbf{y}(n+1)$ is the output signal. In the training part, the $\mathbf{W}_{out}$ is calculated to minimize the root mean square error (RMSE) between the predicted value $\mathrm{y}^{pre}(n)$ and target value $\mathrm{y}^{tar}(n)$

$$\mathbf{W}_{out} = \mathrm{argmin}\sqrt{\frac{1}{m}\sum_{n=1}^{m}\left(\mathbf{y}^{pre}(n)-\mathbf{y}^{tar}(n)\right)^2} \tag{4}$$

where m is the total data points in the training part. When the ridge regression is utilized, the $\mathbf{W}_{out}$ can be calculated by linear regression

$$\mathbf{W}_{out} = \mathbf{Y}^{tar}\mathbf{X}^{\mathrm{T}}\left(\mathbf{X}\mathbf{X}^{\mathrm{T}}+\gamma\mathbf{I}\right)^{-1} \tag{5}$$

where $\mathbf{X}$ represents the neuron state and inputs ($\mathbf{x}(n+1)$;$\mathbf{u}(n)$), $\mathbf{Y}^{tar}$ is the target value in the training part, γ is the regularization coefficient, and $\mathbf{I}$ is the identity matrix. In this task, the reservoir neuron number N is 400, the bounds of leaking rate α, spectral radius ρ, and the regularization coefficient γ are (0, 1), (0, 2), and (1×10^{-3}, 9×10^{-3}), respectively. The implementation time of ESN without hyperparameter optimization takes about 20 s in Matlab 2018, Intel-i7 CPU (2.6 GHz), and 16 G RAM environment.

The $\mathbf{W}_{in}$ and $\mathbf{W}$ are generated randomly and sparsely. For simplicity, they are usually generated in the same distribution form. Based on the practical guide to applying ESN [28], the uniform distribution and Gaussian distribution are two popular distribution shapes. The uniform distribution is first performed for its continuity of values and boundedness. The distribution interval is $[-0.5, 0.5]$, and the variance $D(x)$ is calculated as $1/12$. When the Gaussian distribution has the same variance as uniform distribution, the weight matrix follows the normal distribution N (0, $1/12$) [29]. The probability density functions of uniform distribution and Gaussian distribution are shown in Figure 1.

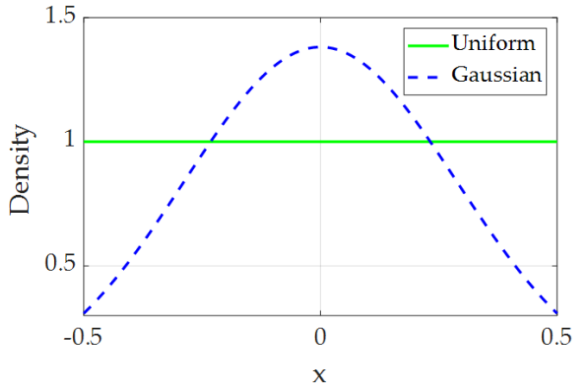


Figure 1. The probability density function of uniform distribution and Gaussian distribution.

2.2. Particle Swarm Optimization

The hyperparameters of the ESN need to be carefully tuned, including the leaking rate α, spectral radius ρ, regularization coefficient γ, reservoir neuron number N, reservoir connectivity c, etc. Some straightforward hyperparameter setting methods such as trial-and-error [18], grid-search [15], and stochastic gradient descent [10] are computationally complex. The effectiveness of some metaheuristic methods such as PSO [30], GA [31], and big bang–big crunch (BB–BC) [20] have been proven in different tasks. Optimizing all hyperparameters is time-consuming, so three key hyperparameters of the ESN (α, ρ, and γ), which have the biggest effects on the prediction results, are selected to be optimized [16].

The PSO is a satisfactory global search method that is easier to be implemented, has fewer parameters, is computationally less expensive, and can converge more quickly [32]. A population of particles is used for searching, and each particle indicates a candidate solution to the problem. The velocity v_i and position x_i updating of the particle at time step k are shown as

$$\begin{cases} v_i^{k+1} = wv_i^k + c_1r_1(p_{ib} - x_i^k) + c_2r_2(g_{ib} - x_i^k) \\ x_i^{k+1} = x_i^k + v_i^{k+1} \end{cases} \quad (6)$$

where p_{ib} and g_{ib} are the best personal and global positions, w is the inertial weight, c_1 and c_2 are the learning factors, and r_1 and r_2 are the random variables from 0 to 1. The objective function is to minimize the RMSE between the predicted value $\mathrm{y}^{pre}(n)$ and target value $\mathrm{y}^{tar}(n)$, as shown in Equation (4). In this case study, the swarm size is 30, the maximum number of iterations is 600, the function tolerance is 1×10^{-6}, both learning factors are 1.49, and the inertial weight is 0.8.

2.3. Implementation of Ensemble ESN

The implementation process of the ensemble ESN is shown in Figure 2. There are four steps in total:

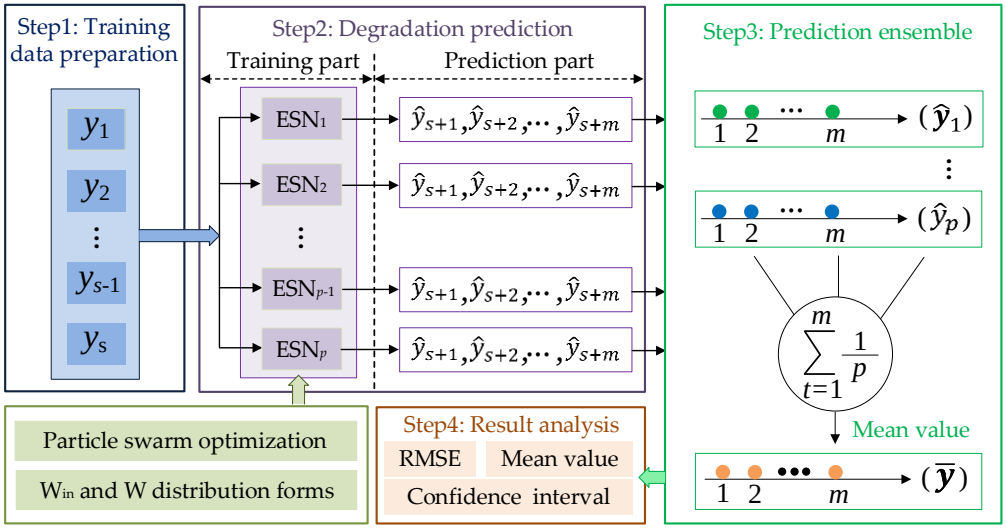


Figure 2. The implementation process of ensemble ESN.

Step 1 is the training data preparation. Some pretreatment procedures are necessary, such as filtering, resampling, and normalization after the data (voltage, current, etc.) acquisition. In this task, the moving average filtering method is utilized, and the length of the moving window is empirically set to 31. The resampling time interval between two adjacent points is 0.5 h. The data are normalized to the same interval [0, 1] by Equation (7)

$$z' = \frac{z - z_{\min}}{z_{\max} - z_{\min}} \tag{7}$$

where z' is the normalized value, z is the original value, and $z_{\max}$ and $z_{\min}$ are the maximum and minimum values, respectively. The health indicators can be extracted, i.e., voltage for steady-state and quasi-dynamic operating conditions and the relative power loss rate (RPLR) for the dynamic operating condition.

Step 2 is the degradation prediction based on individual ESNs. The randomly generated procedure of $\mathbf{W}_{in}$ and $\mathbf{W}$ are similar in the ESN, so the distribution shapes of these two matrices are the same for convenience. For each ESN_i ($1 \le i \le p$), the uniform distribution and Gaussian distribution shapes of these two matrices are explored separately. The hyperparameters of the ESN are optimized by PSO in the training part, and the optimized parameter combination is used in the prediction part. Thus, each ESN would have two long-term prediction results under two random matrices distribution shapes. The total number of ensemble ESNs is set to be 100 (p = 100) for statistical analysis.

Step 3 is the ensemble of prediction results. For a certain distribution shape, 100 times prediction results are assembled to form a single new model to statistically analyze the randomness. The weight of each ESN is taken to be the same, i.e., each ESN has the same contribution during the RUL prediction. Thus, the mean value of 100 ESNs at each time is calculated as the final prediction value.

Step 4 is the result analysis. Each ESN could realize an individual prediction, and its RMSE between the predicted signal and target signal can be calculated by

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{t=1}^{m}\left(\mathbf{y}^{pre}(t) - \mathbf{y}^{tar}(t)\right)^2} \tag{8}$$

At a certain time step t ($1 \leq t \leq m$), the mean and the variance of the predicted value at time step ($s + t$) can be represented as

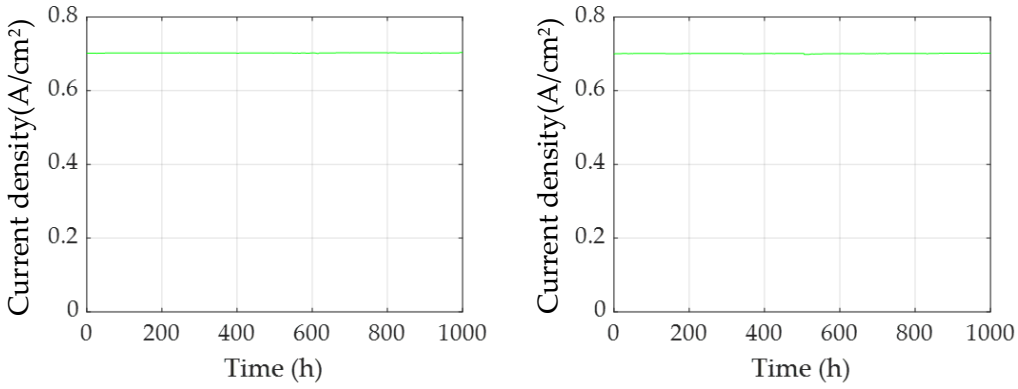
$$\begin{cases} \overline{y}(s+t) = \frac{1}{p}\sum_{i=1}^{p} \hat{y}_i(s+t) \\ D(s+t) = \frac{1}{p-1}\sum_{i=1}^{p} (\overline{y}(s+t) - \hat{y}_i(s+t))^2 \end{cases} \tag{9}$$

where p is the total number of ensemble ESNs, $\hat{y}_i(s+t)$ is the predicted value, $\overline{y}(s+t)$ is mean value, and $D(s+t)$ is the variance. Assuming the predicted value follows the standard Gaussian distribution, the 95% confidence interval (CI) of the results at time step ($s + t$) can be expressed as

$$\mathrm{CI}(s+t) = [\overline{y}(s+t) - 1.96\sqrt{D(s+t)}, \overline{y}(s+t) + 1.96\sqrt{D(s+t)}] \tag{10}$$

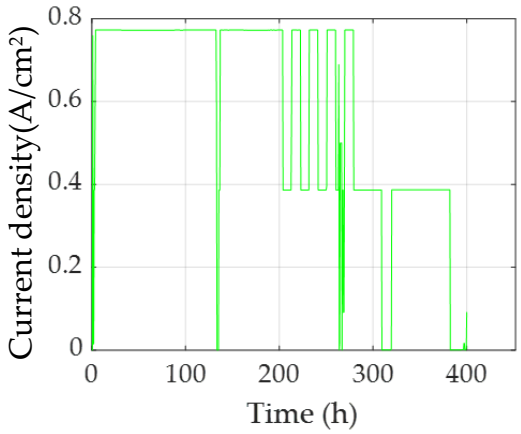
3. Experimental Results

Three degradation datasets under different operation conditions are used for exploring the prediction effects of two distribution shapes of weight matrices, testing the performance improvements of the ensemble ESN, and statistically analyzing the uncertainty of random matrices. The steady-state and quasi-dynamic datasets come from the open data of the 2014 Prognostic and Health Management Data Challenge [22]. Two PEMFC stacks are used for the 1000 h duration test under the steady-state and quasi-dynamic operating conditions. There are 5 cells for both stacks, and their active area is 100 cm^2, the operating temperature is 80 °C, and the back pressure is 0.2 MPa. For the steady-state test, the constant current is 70 A. For the quasi-dynamic test, the current ripples (7 A) at different frequencies (from 50 mHz to 10 kHz) are superimposed to the constant current (70 A) for the electrochemical impedance spectroscopy (EIS) measurement. The third dataset comes from a PEMFC stack for 382 hours' duration test under the dynamic μ-CHP load profile. The current has a cycle between 170 A and 85 A. The stack consists of 8 cells, and the active area is 220 cm^2, the operating temperature is 80 °C, and the back pressure is 0.15 MPa [18]. The load profiles of the three tests are shown in Figure 3.



(**a**) (**b**)

Figure 3. *Cont.*

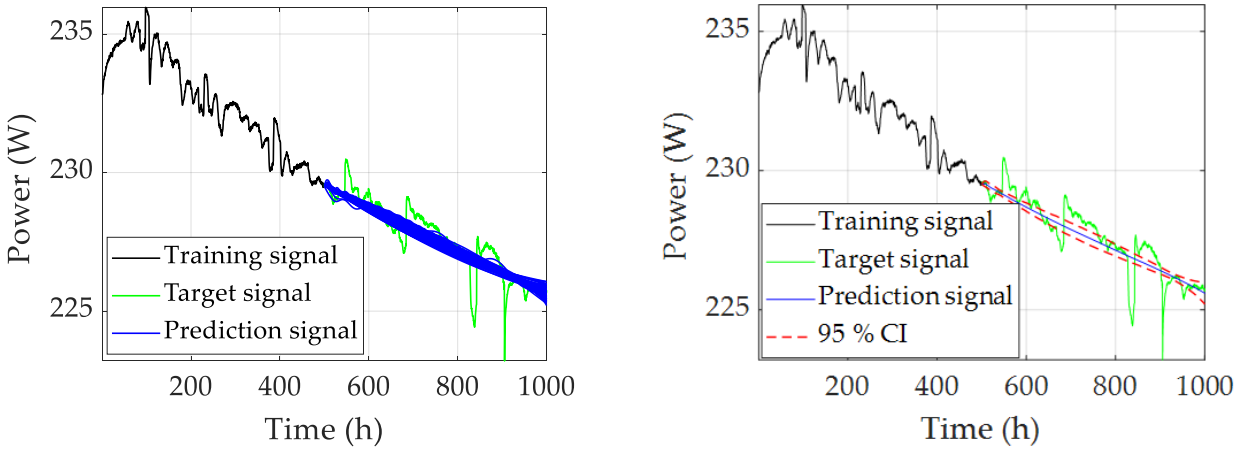


(**c**)

Figure 3. Load profiles of three tests: (**a**) steady-state operating condition; (**b**) quasi-dynamic operating condition; (**c**) dynamic operating condition.

3.1. Steady-State Operating Condition

In the steady-state operating condition, the stack output power is used as the degradation health indicator. Half of the data (0–500 h) is used for training the ESN, and the rest of the data (501–1000 h) is used for prediction. During the prediction procedure, 100 individual ESNs have 100 groups of prediction results. Then, the mean value, variance, and 95% CI of the predicted results can be calculated at each time step. When $\mathbf{W}_{in}$ and $\mathbf{W}$ follow the uniform distribution, the RUL prediction results are shown in Figure 4. Taking the first 10 groups of ESNs as an example, their hyperparameters and RMSE are shown in Table 1.



(**a**) (**b**)

Figure 4. Steady-state RUL results under the uniform distribution: (**a**) Prediction results of 100 individual ESNs (blue lines); (**b**) Mean prediction results (blue line) of ensemble ESN with 95% CI.

Table 1. Steady-state results of ESN_1-ESN_{10} when $\mathbf{W}_{in}$ and $\mathbf{W}$ follow the uniform distribution.

Test Group		ESN_1	ESN_2	ESN_3	ESN_4	ESN_5	ESN_6	ESN_7	ESN_8	ESN_9	ESN_{10}
Hyper-para.	α	0.80000	0.24477	0.80000	0.80000	0.10000	0.61598	0.55070	0.28000	0.10000	0.10000
	ρ	0.38125	0.69227	0.79748	0.93218	0.20756	0.57188	0.56348	0.72581	0.18374	0.17359
	γ	0.00100	0.00900	0.00900	0.00900	0.00164	0.00363	0.00100	0.00900	0.00136	0.00340
RMSE		0.58050	0.59487	0.59647	0.60032	0.59554	0.57550	0.57331	0.57092	0.60588	0.63377

After changing the distribution shapes of $\mathbf{W}_{in}$ and $\mathbf{W}$ to Gaussian distribution, the RUL prediction results under the steady-state operating condition are shown in Figure 5. The hyperparameters and RMSE of the first 10 groups of ESNs are shown in Table 2. To analyze the effects of the two matrices' distribution shapes of $\mathbf{W}_{in}$ and $\mathbf{W}$ on the prediction accuracy, the RMSE of 100 ESNs is shown in Figure 6.

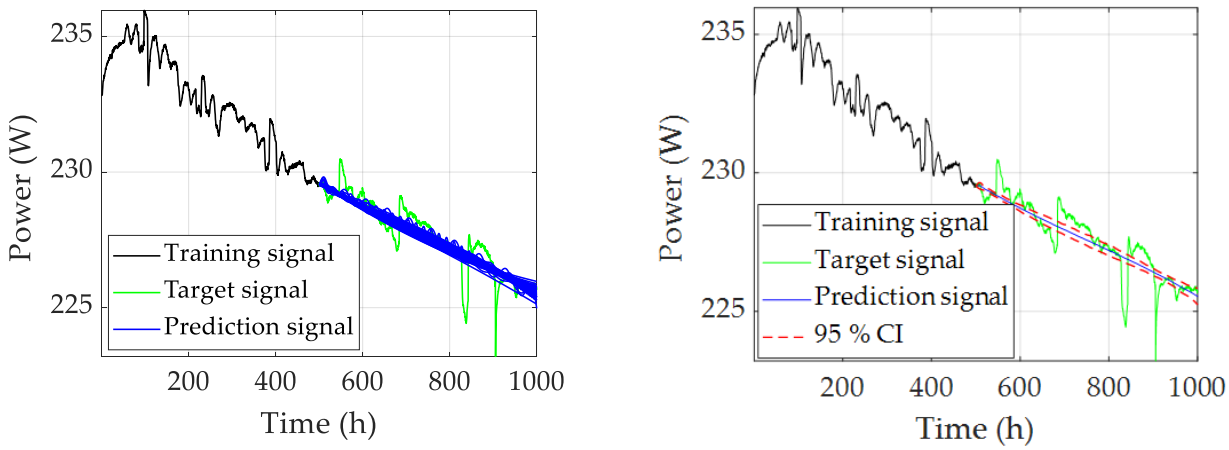


Figure 5. Steady-state RUL results under the Gaussian distribution: (**a**) Prediction results of 100 individual ESNs (blue lines); (**b**) Mean prediction results (blue line) of ensemble ESN with 95% CI.

Table 2. Steady-state results of ESN_1-ESN_{10} when $\mathbf{W}_{in}$ and $\mathbf{W}$ follow the Gaussian distribution.

Test Group		ESN_1	ESN_2	ESN_3	ESN_4	ESN_5	ESN_6	ESN_7	ESN_8	ESN_9	ESN_{10}
Hyper-para.	α	0.90000	0.90000	0.90000	0.85000	0.39040	0.90000	0.90000	0.71572	0.90000	0.90000
	ρ	0.19560	0.33440	0.44590	0.10000	0.86250	0.10000	0.31760	0.10000	1.41060	0.10000
	γ	0.00100	0.00320	0.00650	0.00660	0.00350	0.00110	0.00100	0.00100	0.00100	0.00120
RMSE		0.56508	0.57730	0.57151	0.56970	0.56199	0.56915	0.56801	0.58255	0.57192	0.56398

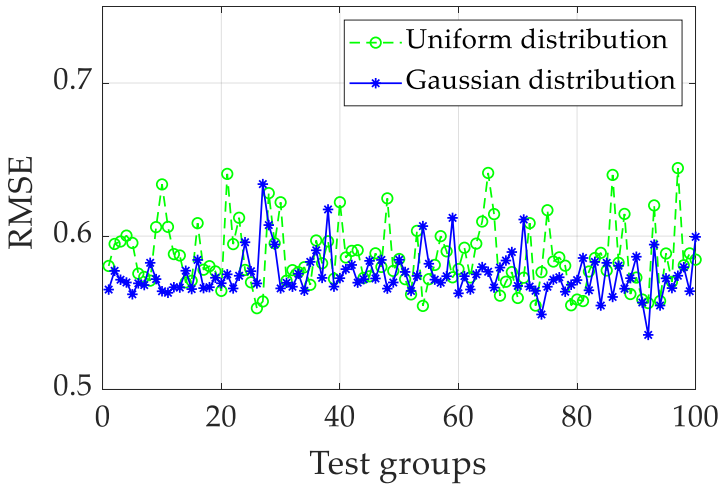


Figure 6. RMSE of 100 groups of ESNs under the steady-state operating condition.

The average RMSE under the Gaussian distribution (0.57488) is lower than the average RMSE under the uniform distribution (0.58602) for the steady-state test. The mean prediction results of the ensemble ESN are calculated, i.e., 0.57841 for uniform distribution and 0.56966 for Gaussian distribution.

3.2. Quasi-Dynamic Operating Condition

In the quasi-dynamic operating condition, the stack output power is used as the degradation-related health indicator. Similar to the steady-state operating condition, half of the data (0–500 h) is used for training the ESN, and the rest of the data (501–1000 h) is used for prediction. When $\mathbf{W}_{in}$ and $\mathbf{W}$ follow the uniform distribution, the RUL prediction results are shown in Figure 7. Taking the first 10 groups of ESNs as an example, their hyperparameters and RMSE are shown in Table 3.

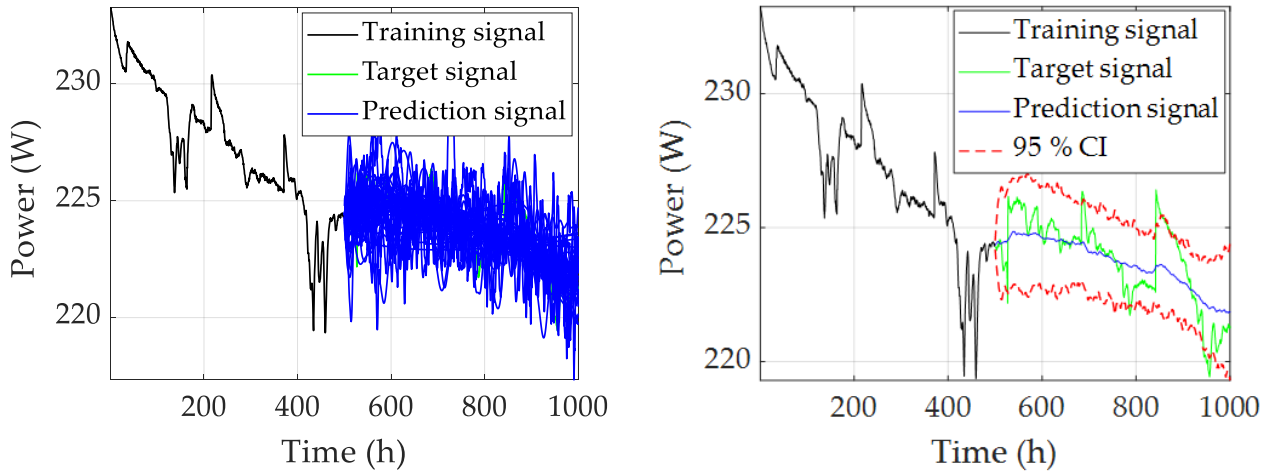


Figure 7. Quasi-dynamic RUL results under the uniform distribution: (**a**) Prediction results of 100 individual ESNs (blue lines); (**b**) Mean prediction results (blue line) of ensemble ESN with 95% CI.

Table 3. Quasi-dynamic results of ESN_1-ESN_{10} when $\mathbf{W}_{in}$ and $\mathbf{W}$ follow the uniform distribution.

Test Group		ESN_1	ESN_2	ESN_3	ESN_4	ESN_5	ESN_6	ESN_7	ESN_8	ESN_9	ESN_{10}
Hyper-para.	α	0.29901	0.10970	0.10820	0.13740	0.10000	0.34770	0.10000	0.41590	0.10800	0.12280
	ρ	1.90000	1.80230	1.74110	1.90000	1.62237	1.55670	1.90000	1.42740	1.42460	1.61780
	γ	0.00299	0.00710	0.00860	0.00150	0.00900	0.00110	0.00900	0.00890	0.00900	0.00100
RMSE		1.07394	1.01290	1.02710	1.10380	1.41801	1.03080	1.88830	1.28040	1.09540	1.49140

After changing the distribution shapes of $\mathbf{W}_{in}$ and $\mathbf{W}$ to Gaussian distribution, the RUL prediction results under the quasi-dynamic operating condition are shown in Figure 8. The hyperparameters and RMSE of the first 10 groups of ESNs are shown in Table 4.

Table 4. Quasi-dynamic results of ESN_1-ESN_{10} when $\mathbf{W}_{in}$ and $\mathbf{W}$ follow the Gaussian distribution.

Test Group		ESN_1	ESN_2	ESN_3	ESN_4	ESN_5	ESN_6	ESN_7	ESN_8	ESN_9	ESN_{10}
Hyper-para.	α	0.87000	0.35830	0.38420	0.10110	0.87420	0.54400	0.84620	0.82820	0.10000	0.10900
	ρ	1.36500	1.80570	1.58110	1.51860	1.47940	1.90000	1.83080	1.43060	1.76060	1.86000
	γ	0.00790	0.00830	0.00900	0.00900	0.00330	0.00900	0.00880	0.00100	0.00100	0.00580
RMSE		1.03890	0.93920	1.28170	0.89560	0.96380	1.37308	1.05020	0.97690	1.13970	0.90080

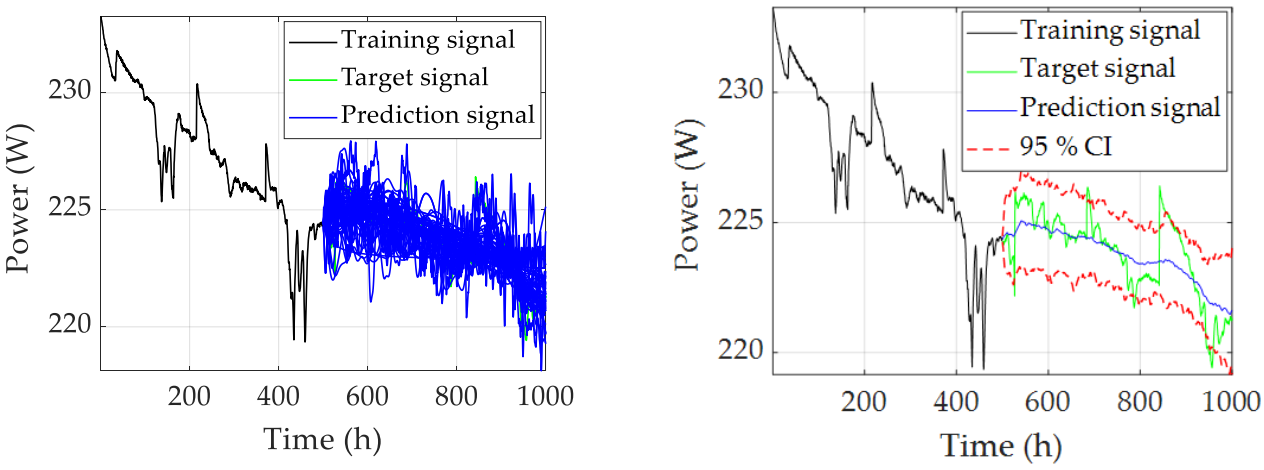


Figure 8. Quasi-dynamic RUL results under the Gaussian distribution: (**a**) Prediction results of 100 individual ESNs (blue lines); (**b**) Mean prediction results (blue line) of ensemble ESN with 95% CI.

To analyze the effects of the two matrices' distribution shapes of $\mathbf{W}_{in}$ and $\mathbf{W}$ on the prediction accuracy, the RMSE of 100 groups of ESNs is shown in Figure 9. The average RMSE under the Gaussian distribution (1.19634) is lower than the average RMSE under the uniform distribution (1.27562) for the quasi-dynamic test. The mean prediction results of the ensemble ESN are calculated, i.e., 0.95657 for uniform distribution and 0.90618 for Gaussian distribution.

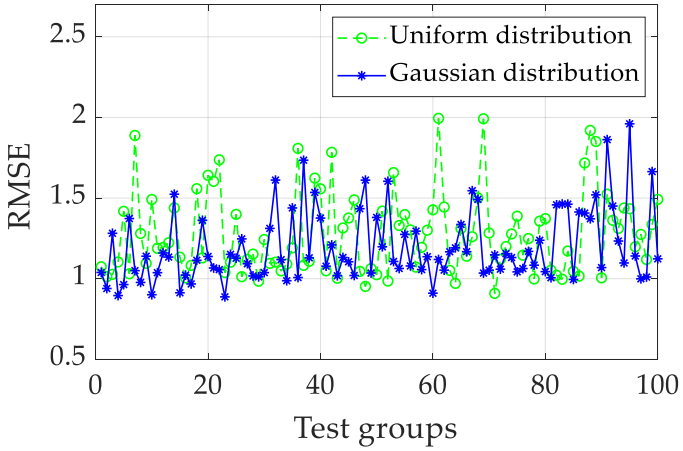


Figure 9. RMSE of 100 groups of ESNs under the quasi-dynamic operating condition.

3.3. Dynamic Operating Condition

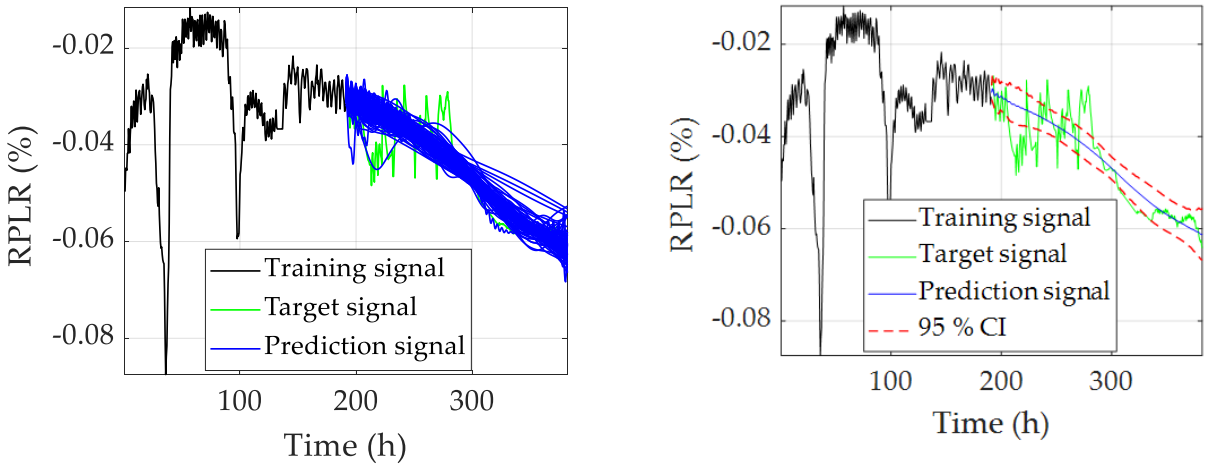
In the dynamic operating condition, the RPLR is used as the degradation health indicator, which has been described in detail in our previous work [18]. First, the polarization curve of the PEMFC stack at the beginning of life (BoL) is measured. The BoL power at different currents can be calculated to form a table. Secondly, the power at time step t can be calculated by the measured current (I_t) and voltage (U_t). Finally, the RPLR at time step t is calculated by

$$\text{RPLR} = (P_t - P_0)/P_t \tag{11}$$

where P_t is the power at time step t and P_0 is the BoL power, which is determined by the look-up table method. The P_t and P_0 are at the same current level.

Half of the data (0–191 h) is used for training the ESN, and the rest of the data (192–382 h) is used for prediction. When $\mathbf{W}_{in}$ and $\mathbf{W}$ follow the uniform distribution, the

RUL prediction results are shown in Figure 10. Taking the first 10 groups of ESNs as an example, their hyperparameters and RMSE are shown in Table 5.



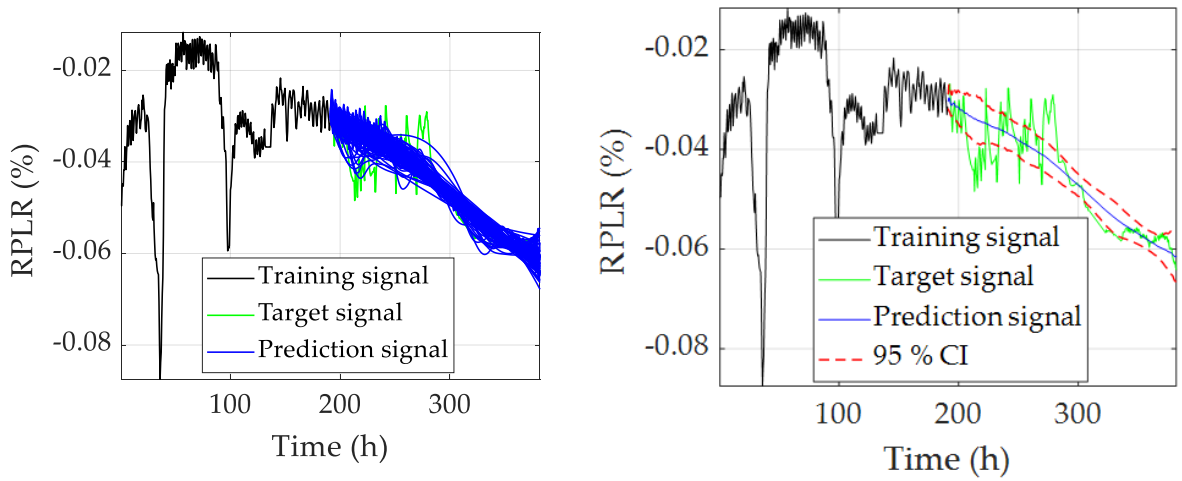
(**a**) (**b**)

Figure 10. Dynamic RUL results under the uniform distribution: (**a**) Prediction results of 100 individual ESNs (blue lines); (**b**) Mean prediction results (blue line) of ensemble ESN with 95% CI.

Table 5. Dynamic results of ESN_1-ESN_{10} when $\mathbf{W}_{in}$ and $\mathbf{W}$ follow the uniform distribution.

Test Group		ESN_1	ESN_2	ESN_3	ESN_4	ESN_5	ESN_6	ESN_7	ESN_8	ESN_9	ESN_{10}
Hyper-para.	α	0.12483	0.25671	0.15844	0.11213	0.34245	0.23291	0.23923	0.14834	0.26963	0.36064
	ρ	1.43205	1.74959	1.35257	1.53465	1.18192	1.33707	1.90000	1.23153	1.37328	1.25512
	γ	0.00218	0.00524	0.00900	0.00900	0.00129	0.00405	0.00100	0.00900	0.00165	0.00253
RMSE		0.00476	0.00527	0.00483	0.00482	0.00578	0.00523	0.00508	0.00485	0.00479	0.00501

After changing the distribution shapes of $\mathbf{W}_{in}$ and $\mathbf{W}$ to Gaussian distribution, the RUL prediction results under the dynamic operating condition are shown in Figure 11. The hyperparameters and RMSE of the first 10 groups of ESNs are shown in Table 6.



(**a**) (**b**)

Figure 11. Dynamic RUL results under the Gaussian distribution: (**a**) Prediction results of 100 individual ESNs (blue lines); (**b**) Mean prediction results (blue line) of ensemble ESN with 95% CI.

Table 6. Dynamic results of ESN_1-ESN_{10} when $\mathbf{W}_{in}$ and $\mathbf{W}$ follow the Gaussian distribution.

Test Group		ESN_1	ESN_2	ESN_3	ESN_4	ESN_5	ESN_6	ESN_7	ESN_8	ESN_9	ESN_{10}
Hyper-para.	α	0.30376	0.47659	0.54372	0.50078	0.13709	0.57584	0.32100	0.28035	0.11564	0.39271
	ρ	1.64620	1.29902	1.55139	1.71862	1.38232	1.44123	1.43160	1.59497	1.50729	1.90000
	γ	0.00634	0.00900	0.00586	0.00900	0.00900	0.00638	0.00157	0.00625	0.00900	0.00900
RMSE		0.00511	0.00527	0.00525	0.00481	0.00413	0.00512	0.00509	0.00473	0.00569	0.00485

The RMSE of 100 groups of ESNs is shown in Figure 12. The average RMSE under the Gaussian distribution (0.00489) is lower than the average RMSE under the uniform distribution (0.00502) for the dynamic test. The mean prediction results of the ensemble ESN are calculated, i.e., 0.00473 for uniform distribution and 0.00464 for Gaussian distribution.

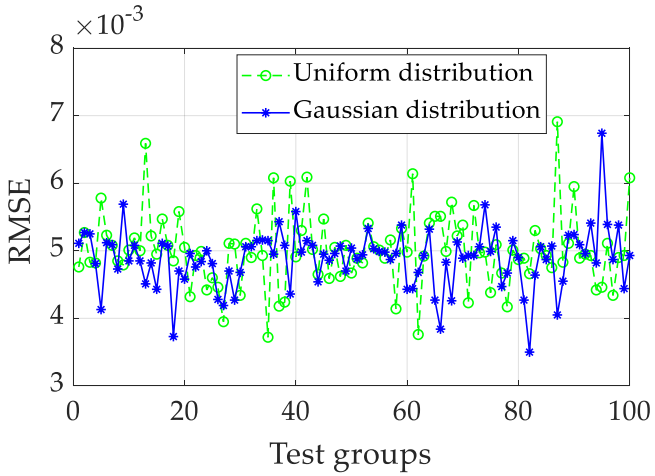


Figure 12. RMSE of 100 groups of ESNs under the dynamic operating condition.

Among these three different operating conditions, each random generation of $\mathbf{W}_{in}$ and $\mathbf{W}$ produces a different prediction result. The essential reason is that the randomness of the matrices changes the weight values of the connected neurons, and the random weight leads to different neuron behaviors, which are important for the nonlinear relationship construction ability of the reservoir. The numerous repeated tests also illustrate the uncertainty properties of random matrices. The ensemble ESN structure is utilized to statistically analyze the effects of random characteristics and the bounds of uncertainties are obtained. Taking into account these uncertainties, the robustness of the prediction can be further improved in practical applications.

4. Conclusions

The data-driven method of echo state network (ESN) is further explored in the long-term lifetime prediction of the PEMFC system. During the implementation process of the ESN, the input and the internal weight matrices ($\mathbf{W}_{in}$ and $\mathbf{W}$) are randomly generated. The traditional method keeps these two matrices unchanged once they are generated. Nevertheless, the random characteristics of these matrices will affect the performance of the ESN. To statistically analyze the effects of random characteristics, an ensemble ESN is proposed and tested under three different operating conditions. The conclusions are as follows:

1. The random characteristics of $\mathbf{W}_{in}$ and $\mathbf{W}$ affect the lifetime prediction results. The prediction results are presented in the statistical form by the ensemble computing technique, and this helps the user to analyze the uncertainties of the randomness. After the data analysis, a 95% confidence interval (CI) is given, which better qualifies the reliability of the result.

2. Based on the uniform and Gaussian distribution shapes of $\mathbf{W}_{in}$ and $\mathbf{W}$, the prediction performances are fully compared. To analyze the effects of random matrices, the PSO method is used to optimize the hyperparameters of the ESN. Based on the comparison results, the Gaussian distribution of $\mathbf{W}_{in}$ and $\mathbf{W}$ can decrease the prediction error slightly when compared to the uniform distribution. However, the effects of two different distribution shapes on the prediction results are rather insignificant.
3. Combining the ESN with other methods to improve the prediction accuracy and exploring the prognostic methods to realize the online lifetime prediction will be the focuses of our future work.

Author Contributions: Conceptualization, Z.H. and Z.Z.; Formal analysis, Z.H. and Z.Z.; Funding acquisition, M.-C.P. and F.G.; Methodology, Z.H. and Z.Z.; Project administration, M.-C.P. and F.G.; Software, Z.H. and Z.Z.; Supervision, Z.Z., M.-C.P. and F.G.; Validation, Z.H. and Z.Z.; Writing—original draft, Z.H.; Writing—review and editing, Z.H., Z.Z., M.-C.P. and F.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by the EIPHI Graduate school (contract "ANR-17-EURE-0002").

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Qiu, Y.; Li, Q.; Huang, L.; Sun, C.; Wang, T.; Chen, W. Adaptive uncertainty sets-based two-stage robust optimization for economic dispatch of microgrid with demand response. *IET Renew. Power Gener.* **2020**, *14*, 3608–3615. [CrossRef]
2. Pu, Y.; Li, Q.; Zou, X.; Li, R.; Li, L.; Chen, W.; Liu, H. Optimal sizing for an integrated energy system considering degradation and seasonal hydrogen storage. *Appl. Energy* **2021**, *302*, 117542. [CrossRef]
3. Li, Q.; Wang, T.; Li, S.; Chen, W.; Liu, H.; Breaz, E.; Gao, F. Online extremum seeking-based optimized energy management strategy for hybrid electric tram considering fuel cell degradation. *Appl. Energy* **2021**, *285*, 116505. [CrossRef]
4. Li, Q.; Yin, L.; Yang, H.; Wang, T.; Qiu, Y.; Chen, W. Multiobjective optimization and data-driven constraint adaptive predictive control for efficient and stable operation of PEMFC system. *IEEE Trans. Ind. Electron.* **2021**, *68*, 12418–12429. [CrossRef]
5. Wang, C.; Li, Z.; Outbib, R.; Dou, M.; Zhao, D. A novel long short-term memory networks-based data-driven prognostic strategy for proton exchange membrane fuel cells. *Int. J. Hydrogen Energy* **2022**, *47*, 10395–10408. [CrossRef]
6. Hua, Z.; Zheng, Z.; Pahon, E.; Péra, M.C.; Gao, F. A review on lifetime prediction of proton exchange membrane fuel cells system. *J. Power Sources* **2022**, *529*, 231256. [CrossRef]
7. Zuo, J.; Lv, H.; Zhou, D.; Xue, Q.; Jin, L.; Zhou, W.; Yang, D.; Zhang, C. Deep learning based prognostic framework towards proton exchange membrane fuel cell for automotive application. *Appl. Energy* **2021**, *281*, 115937. [CrossRef]
8. Yue, M.; Li, Z.; Roche, R.; Jemei, S.; Zerhouni, N. Degradation identification and prognostics of proton exchange membrane fuel cell under dynamic load. *Control Eng. Pract.* **2022**, *118*, 104959. [CrossRef]
9. Hadaeghi, F.; Jaeger, H. Computing optimal discrete readout weights in reservoir computing is NP-hard. *Neurocomputing* **2019**, *338*, 233–236. [CrossRef]
10. Thiede, L.A.; Parlitz, U. Gradient based hyperparameter optimization in echo state networks. *Neural Netw.* **2019**, *115*, 23–29. [CrossRef]
11. Morando, S.; Jemei, S.; Gouriveau, R.; Zerhouni, N.; Hissel, D. Fuel cells prognostics using echo state network. In Proceedings of the 39th Annual Conference of the IEEE Industrial Electronics Society (IECON), Vienna, Austria, 10–13 November 2013. [CrossRef]
12. Morando, S.; Jemei, S.; Hissel, D.; Gouriveau, R.; Zerhouni, N. Predicting the remaining useful lifetime of a proton exchange membrane fuel cell using an echo state network. In Proceedings of the International Discussion on Hydrogen Energy and Applications (IDHEA), Nantes, France, 12–14 May 2014.
13. Jin, J.; Chen, Y.; Wenchao, Z.; Xie, C.; Wu, F. Remaining useful life prediction of PEMFC based on cycle reservoir with jump model. *Int. J. Hydrogen Energy* **2021**, *46*, 40001–40013. [CrossRef]
14. 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]
15. Hua, Z.; Zheng, Z.; Pahon, E.; Péra, M.C.; Gao, F. Multi-timescale lifespan prediction for PEMFC systems under dynamic operating conditions. *IEEE Trans. Transp. Electrif.* **2022**, *8*, 345–355. [CrossRef]

16. Hua, Z.; Zheng, Z.; Pahon, E.; Péra, M.C.; Gao, F. Lifespan prediction for proton exchange membrane fuel cells based on wavelet transform and echo state network. *IEEE Trans. Transp. Electrif.* **2022**, *8*, 420–431. [CrossRef]
17. Morando, S.; Jemei, S.; Hissel, D.; Gouriveau, R.; Zerhouni, N. ANOVA method applied to proton exchange membrane fuel cell ageing forecasting using an echo state network. *Math. Comput. Simul.* **2017**, *131*, 283–294. [CrossRef]
18. Hua, Z.; Zheng, Z.; Pahon, E.; Péra, M.-C.; Gao, F. Remaining useful life prediction of PEMFC systems under dynamic operating conditions. *Energy Convers. Manag.* **2021**, *231*, 113825. [CrossRef]
19. Mezzi, R.; Morando, S.; Steiner, N.Y.; Péra, M.C.; Hissel, D.; Larger, L. Multi-reservoir echo state network for proton exchange membrane fuel cell remaining useful life prediction. In Proceedings of the 44th Annual Conference of the IEEE Industrial Electronics Society (IECON), Washington, DC, USA, 21–23 October 2018. [CrossRef]
20. Mezzi, R.; Yousfi-Steiner, N.; Péra, M.C.; Hissel, D.; Larger, L. An echo state network for fuel cell lifetime prediction under a dynamic micro-cogeneration load profile. *Appl. Energy* **2021**, *283*, 116297. [CrossRef]
21. Li, Z.; Zheng, Z.; Outbib, R. Adaptive prognostic of fuel cells by implementing ensemble echo state networks in time varying model space. *IEEE Trans. Ind. Electron.* **2019**, *67*, 379–389. [CrossRef]
22. Hua, Z.; Zheng, Z.; Péra, M.-C.; Gao, F. Remaining useful life prediction of PEMFC systems based on the multi-input echo state network. *Appl. Energy* **2020**, *265*, 114791. [CrossRef]
23. Vichard, L.; Harel, F.; Ravey, A.; Venet, P.; Hissel, D. Degradation prediction of PEM fuel cell based on artificial intelligence. *Int. J. Hydrogen Energy* **2020**, *45*, 14953–14963. [CrossRef]
24. He, K.; Mao, L.; Yu, J.; Huang, W.; He, Q.; Jackson, L. Long-term performance prediction of PEMFC based on LASSO-ESN. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–11. [CrossRef]
25. Napoli, G.; Ferraro, M.; Sergi, F.; Brunaccini, G.; Antonucci, V. Data driven models for a PEM fuel cell stack performance prediction. *Int. J. Hydrogen Energy* **2013**, *38*, 11628–11638. [CrossRef]
26. Zhang, D.; Baraldi, P.; Cadet, C.; Yousfi-Steiner, N.; Bérenguer, C.; Zio, E. An ensemble of models for integrating dependent sources of information for the prognosis of the remaining useful life of proton exchange membrane fuel cells. *Mech. Syst. Signal Process.* **2019**, *124*, 479–501. [CrossRef]
27. Javed, K.; Gouriveau, R.; Zerhouni, N.; Hissel, D. Prognostics of proton exchange membrane fuel cells stack using an ensemble of constraints based connectionist networks. *J. Power Sources* **2016**, *324*, 745–757. [CrossRef]
28. Lukoševičius, M. A practical guide to applying echo state networks. In *Neural Networks: Tricks of the Trade*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 659–686. [CrossRef]
29. Wu, Q.; Kudithipudi, D. An ensemble learning approach to the predictive stability of echo state networks. *J. Inform. Math. Sci.* **2018**, *10*, 181–199. [CrossRef]
30. Chouikhi, N.; Fdhila, R.; Ammar, B.; Rokbani, N.; Alimi, A.M. Single- and multi-objective particle swarm optimization of reservoir structure in echo state network. In Proceedings of the International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016. [CrossRef]
31. Zhong, S.; Xie, X.; Lin, L.; Wang, F. Genetic algorithm optimized double-reservoir echo state network for multi-regime time series prediction. *Neurocomputing* **2017**, *238*, 191–204. [CrossRef]
32. Wang, H.; Yan, X. Optimizing the echo state network with a binary particle swarm optimization algorithm. *Knowl. Based Syst.* **2015**, *86*, 182–193. [CrossRef]