An Artificial Intelligence Solution for Predicting Short-Term Degradation Behaviors of Proton Exchange Membrane Fuel Cell

Zijun Yang [1,†], Bowen Wang [1,†], Xia Sheng [2], Yupeng Wang [2], Qiang Ren [3], Shaoqing He [4,*], Jin Xuan [5,*] and Kui Jiao [1,*]

1   State Key Laboratory of Engines, Tianjin University, 135 Yaguan Road, Tianjin 300350, China; yangzijun@tju.edu.cn (Z.Y.); wangbw94@tju.edu.cn (B.W.)
2   Powertrain Department, General R&D Institute, FAW, No. 1 Xinhongqi Street, Changchun 130013, China; shengxia@faw.com.cn (X.S.); wangyupeng@faw.com.cn (Y.W.)
3   GAC Automotive Research & Development Center, Guangzhou 510000, China; renqiang@gacrnd.com
4   China Automotive Technology & Research Center Co., Ltd., Tianjin 300300, China
5   Department of Chemical Engineering, Loughborough University, Loughborough LE11 3TU, UK
*   Correspondence: heshaoqing@catarc.ac.cn (S.H.); j.xuan@lboro.ac.uk (J.X.); kjiao@tju.edu.cn (K.J.)
†   Equal contribution.

Abstract: The dead-ended anode (DEA) and anode recirculation operations are commonly used to improve the hydrogen utilization of automotive proton exchange membrane (PEM) fuel cells. The cell performance will decline over time due to the nitrogen crossover and liquid water accumulation in the anode. Highly efficient prediction of the short-term degradation behaviors of the PEM fuel cell has great significance. In this paper, we propose a data-driven degradation prediction method based on multivariate polynomial regression (MPR) and artificial neural network (ANN). This method first predicts the initial value of cell performance, and then the cell performance variations over time are predicted to describe the degradation behaviors of the PEM fuel cell. Two cases of degradation data, the PEM fuel cell in the DEA and anode recirculation modes, are employed to train the model and demonstrate the validation of the proposed method. The results show that the mean relative errors predicted by the proposed method are much smaller than those by only using the ANN or MPR. The predictive performance of the two-hidden-layer ANN is significantly better than that of the one-hidden-layer ANN. The performance curves predicted by using the sigmoid activation function are smoother and more realistic than that by using rectified linear unit (ReLU) activation function.

Keywords: PEM fuel cell; artificial neural network; short-term degradation behavior; dead-ended anode; anode recirculation

1. Introduction

Proton exchange membrane (PEM) fuel cell is a high efficiency and low emission electro-chemical device which can be applied in many fields including automobiles, distributed generation and military [1]. In recent years, the PEM fuel cell technology has made great progress in increasing performance, reducing costs and improving durability. However, the output performance and durability of PEM fuel cells are still insufficient for wide commercialization at this stage [2,3]. There exist many performance degradation processes in the PEM fuel cell operation that generally include two categories, recoverable performance degradation in the short term, such as the cell in the cold start, dead-ended anode (DEA), and anode recirculation operation [4–7], and unrecoverable performance degradation in the long term [8,9]. The mechanisms causing performance degradation are various and complicated in these degradation processes, but it is widely recognized that analysis and prediction of the degradation behaviors are critical [10,11].

Developing the physical model is the main method in fuel cell studies [1,11]. For the physical model, the governing equations with strong physical meaning are solved, and

thus the physical model generally has high prediction accuracy on the cell performance and other physical parameters. The DEA and anode recirculation modes are commonly used to improve the hydrogen utilization, especially for the automotive PEM fuel cell [12–14]. Nitrogen crossover and liquid water accumulation in the anode will occur that lead to the performance degradation. To improve the durability and operation stability of the PEM fuel cell, it is necessary to understand the degradation process under various working conditions. Compared to the experimental test, predicting the performance degradation process of the PEM fuel cell by modeling could reduce experimental time and expense. Peng et al. [15] developed a three-dimensional single-channel computational fluid dynamics (CFD) model of the PEM fuel cell with the DEA, and processes of the performance decay under various operating conditions are predicted. Rizvandi and Yesilyurt [16] developed a pseudo-three-dimensional transient model to analyze the dynamic and localized characteristics of the PEM fuel cell with the DEA. The numerical results of the performance decay were compared to the experiment results to validate the model. Xu et al. [17] developed a control-oriented model of the PEM fuel cell with the DEA operation considering the internal transport mechanisms. The working processes under various purge intervals and purge durations were simulated and analyzed, and some critical indicators, such as pressure and nitrogen content, were monitored during the simulation. Furthermore, the calculated performance decay in the DEA operation, and performance recovery in the purge duration were validated by comparing to the experiment results. Luo et al. [18] and Zhou et al. [19] developed a PEM fuel cell stack cold start model to investigate the cold start characteristics. The failure processes of cold start were well simulated and compared with experimental data. Three start modes, constant power, current and voltage, were analyzed, and the optimized cold start strategies were proposed based on the simulation results. Sultan et al. [20] used a bonobo optimizer (BO) algorithm to identify the unknown parameters of the PEM fuel cell model, and the simulation results could well match the experimental results. The BO algorithm also showed better performance than other optimization algorithms. Wang et al. [21] developed a quasi-two-dimensional transient model of PEM fuel cells with anode recirculation. The transient characteristics of the cell with anode recirculation was investigated, including the performance improvement in the initial period due to self-humidification and performance decay due to nitrogen crossover.

Besides physical models, many studies developed data-driven models to predict the transient behaviors of PEM fuel cells [22–25]. The data-driven model constructs the input-output relationship from the experimental dataset, and does not require much prior knowledge. Furthermore, after the model is well-trained, conducting the data-driven model does not require complex calculation, and the computational efficiency is generally much higher than that of the physical model [26,27]. Xie et al. [28] proposed a fusion prognostic approach based on particle filter and long short-term memory (LSTM) network. This approach could effectively predict the remaining useful life and short-term degradation of fuel cells. Zuo et al. [29] reported an attention-based recurrent neural network (RNN) model that could accurately predict the output voltage degradation of PEM fuel cells based on original long-term dynamic loading cycle durability test data. The attention-based LSTM and attention-based gated recurrent unit (GRU) showed the higher prediction accuracy compared to LSTM and GRU. Vichard et al. [30] presented the results of a long-term durability test performed on an open-cathode fuel cell system operating during 5000 h under specific operating conditions. A degradation model based on echo state network (ESN) was then used to predict the performance evolution. The results showed that the model could achieve accurate performance prediction for more than 2000 h. Chen et al. [31] proposed a new grey neural network model (GNNM) method, which combined the particle swarm optimization (PSO) and the moving window method to predict the degradation of PEM fuel cell under different operating conditions. The influence of load current, inlet temperature, input hydrogen pressure and inlet relative humidity are considered. Three PEM fuel cell aging experiments were used to demonstrate that the proposed method can accurately predict the degradation for PEM fuel cell on different applications.

Based on the literature review, machine learning was mainly used for the long-term degradation process prediction of PEM fuel cells. The cell performance degradation process in the initial period (generally hundreds of hours) was measured, and the data was used to predict the degradation process in the future time through machine learning. However, to the best of authors' knowledge, the study of machine learning modeling to predict the full-process degradation process of the PEM fuel cell at different operating conditions still lacks. In this study, we propose a data-based degradation prediction method to simulate the degradation behaviors of the PEM fuel cell in the DEA and anode recirculation modes. We firstly used artificial neural network (ANN) to predict the degradation behaviors of the PEM fuel cell, but the results were not satisfactory. Considering that the initial value and change evolution of the PEM fuel cell performance both change under different operating conditions, we try to prediction them separately to improve the accuracy by the multivariate polynomial regression (MPR) method and ANN, respectively. Some previous works also reported the reasonable combination of algorithms can achieve a higher accuracy [32]. Therefore, we propose a combined the MPR method and ANN approach, M-ANN. The main idea is using the MPR method to predict the initial value under different conditions, and using ANN to predict the variation value over time. The workflow of the proposed method, M-ANN, is shown in Figure 1 for predicting the PEM fuel cell degradation process in the DEA mode. We use a verified physical model to simulate the degradation characteristics of the PEM fuel cell in the DEA and anode recirculation modes, and the simulation data is used instead of experimental data. The potential of using experimental data for the proposed modeling method in the future is also discussed.
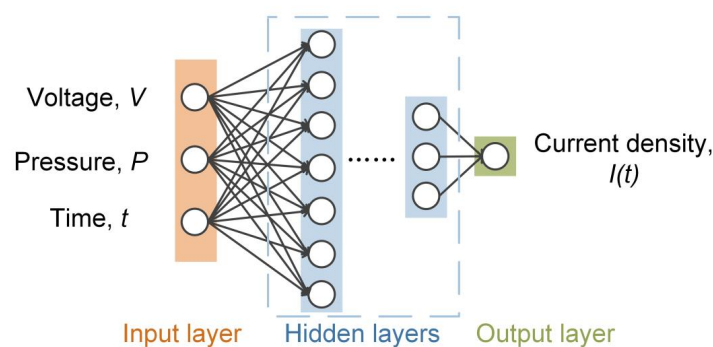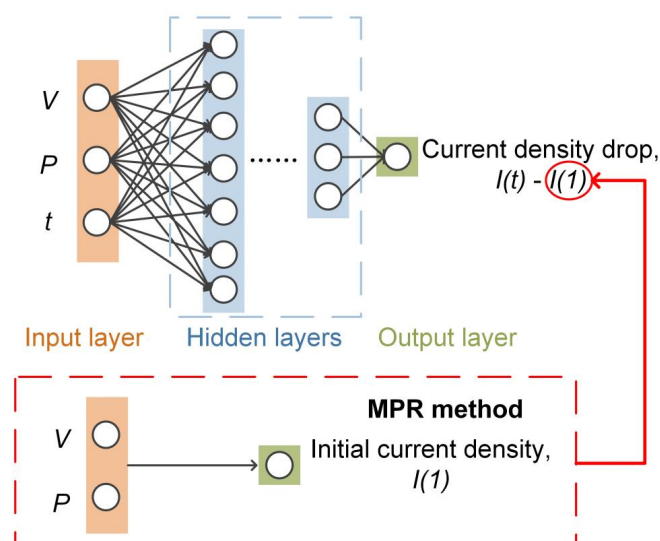

Figure 1. Workflow of the ANN-based and M-ANN based models for predicting the PEM fuel cell degradation process in the DEA mode.

2. Methods

2.1. Data Acquisition

In this study, a mechanism model is employed to simulate the degradation processes of the PEM fuel cell with the DEA mode and anode recirculation mode [21]. The quasi-two-dimensional transient model fully considers the transport process of the electrode and channels, including gas and liquid transport, water phase change, heat transfer, nitrogen crossover and electrochemical reactions. Evolutions of performance change simulated by the model are utilized as the data. The parameters of the simulated PEM fuel cell are listed in Table 1. The governing equations of the employed model are briefly introduced as follows, and more details of transport coefficient and source term could found in Ref. [21].

Table 1. PEM fuel cell parameters.

Parameter	Value
Temperature, T_0	353.15 K
Cathode inlet velocity, u_c	$10.0\ \mathrm{m \cdot s^{-1}}$
Cathode relative humidity, RH_c	1.0
Thickness of the GDL; MPL; CL; membrane, δ_{GDL}; δ_{MPL}; δ_{CL}; δ_m	300.0 μm; 40.0 μm; 10.0 μm; 25.4 μm
Length; Width; Depth of the channel, L; W; d	10 cm; 1 mm; 1 mm
Activation area, A_{act}	$2.0 \times 10^{-4}\ \mathrm{m^2}$
Contact area between the GDL and channel, A_{CH}	$1.0 \times 10^{-4}\ \mathrm{m^2}$
Inlet area, A_{in}	$1.0 \times 10^{-6}\ \mathrm{m^2}$
Porosity of the GDL; MPL; CL ε_{GDL}; ε_{MPL}; ε_{CL}	0.6; 0.4; 0.3
Contact angle of the GDL; MPL; CL, θ_{GDL}; θ_{MPL}; θ_{CL}	110°; 120°; 95°
Permeability of the GDL; MPL; CL; membrane, K_{GDL}; K_{MPL}; K_{CL}; K_m	$1.0 \times 10^{-11}\ \mathrm{m^2}$; $1.0 \times 10^{-12}\ \mathrm{m^2}$; $1.0 \times 10^{-13}\ \mathrm{m^2}$; $2.0 \times 10^{-20}\ \mathrm{m^2}$
Ionomer fraction of the anode and cathode CL, ω_a; ω_c	0.25, 0.25
Transfer coefficient, α;	0.5
Reference exchange current density of anode and cathode, $i_{0,a}^{ref}$; $i_{0,c}^{ref}$	$2.0 \times 10^3\ \mathrm{A \cdot m^{-2}}$; $1.0 \times 10^{-5}\ \mathrm{A \cdot m^{-2}}$
Reference reactant concentration of anode and cathode, $c_{H_2}^{ref}$; $c_{O_2}^{ref}$	$41\ \mathrm{mol \cdot m^{-3}}$; $41\ \mathrm{mol \cdot m^{-3}}$

One-dimensional mass transport along the channel is calculated. For the DEA mode, the liquid water and nitrogen accumulation in the anode channel must be involved. Nitrogen will be accumulated at the dead end of the anode channel firstly, and meanwhile, the nitrogen concentration gradient along the channel is also increasing over time. Although convection is generally dominated for gas species transport in channels, the species diffusion in the dead-ended channel still cannot be ignored. Gas species concentration (hydrogen, oxygen, water vapor and nitrogen) and liquid fraction in channels are calculated:

$$\left(u_{in}^{k-1,[n]} - u_{out}^{k,[n]} \right) A_{in} c_g = \sum_i \phi_{i,CH}^{[n]} A_{in} + \left(\sum_i \phi_{i,CH\text{-}GDL}^{[n]} + S_{i,CH} \right) \frac{A_{act}}{N} \tag{1}$$

$$\frac{\left(c_i^{k,[n]} - c_i^{k-1,[n]} \right) A_{in} L / N}{dt} = \left(J_{i,in}^{[n]} - J_{i,out}^{[n]} + \phi_{i,CH}^{[n]} \right) A_{in} - \left(\phi_{i,CH\text{-}GDL}^{[n]} + S_{i,CH} \right) A_{act} / N \tag{2}$$

$$\frac{\rho_{lq}}{M_w} \frac{\left(s_{CH}^{k,[n]} - s_{CH}^{k-1,[n]} \right) d}{dt} = J_{lq,in}^{[n]} - J_{lq,out}^{[n]} - \phi_{lq,CH\text{-}GDL}^{[n]} - S_{lq,CH} \tag{3}$$

where u $(\mathrm{m \cdot s^{-1}})$ is the velocity; A $(\mathrm{m^2})$ is the area; c $(\mathrm{mol \cdot m^{-3}})$ is the molar concentration; ϕ $(\mathrm{mol \cdot m^{-2} \cdot s^{-1}})$ is the diffusion flux; S $(\mathrm{mol \cdot m^{-2} \cdot s^{-1}})$ is the source term; N is the number

of control volumes along the channel; L (m) is the channel length; J (mol·m^{-2}·s^{-1}) is the convection flux along the channel; ρ (kg·m^{-3}) is the density; M (kg·mol^{-1}) is the molar weight; s is the liquid volume fraction; d (m) is the channel height. The superscripts and subscript: in and out represent the inlet and outlet of the control volume or channel, respectively; $k-1$ and k represent the time step; [n] represents the control volume along the channel; i represents gas species; CH represents the channel; CH-GDL represents the interface between the channel and gas diffusion layer (GDL); act represents the activation; lq represents liquid water; w represents water. For the anode recirculation mode, liquid water in the channel could be quickly blown away and removed by the liquid-vapor separator. We assume that the liquid fraction is fixed as zero in the anode channel, and Equation (3) can be ignored.

Mass and heat transfer in the electrodes are diffusion-dominated along the through-plane direction. Gas species, liquid saturation, membrane water content and temperature are calculated:

$$\frac{\delta_p \varepsilon_p (1-s)\left(c_{i,p}^k - c_{i,p}^{k-1}\right)}{dt} = \phi_{i,p} + S_{i,p} \tag{4}$$

$$\frac{\rho_{lq}}{M_w}\frac{\delta_p \varepsilon_p \left(s_p^k - s_p^{k-1}\right)}{dt} = \phi_{lq,p} + S_{lq,p} \tag{5}$$

$$\frac{\rho_m}{EW}\frac{\delta_p \omega \left(\lambda_p^k - \lambda_p^{k-1}\right)}{dt} = \phi_{mw,p} + S_{mw,p} \tag{6}$$

$$\frac{(\rho C_p)_p^{eff}\delta_p\left(T_p^k - T_p^{k-1}\right)}{dt} = \phi_{T,p} + S_{T,p} \tag{7}$$

where δ (m) is the thickness; ε is the porosity; EW (kg·mol^{-1}) is the equivalent weight of the membrane; ω is the ionomer volume fraction of the catalyst layer (CL); λ is the membrane water content; C_p is the specific heat capacity (J·mol^{-1}·K^{-1}); T (K) is the temperature. The superscripts and subscript: p represents the components of the electrode, including GDLs, micro-porous layers (MPLs) and CLs of the anode and cathode; m represents the membrane; mw represents the membrane water; T represents energy; eff represents the effective coefficient.

Nitrogen crossover from the anode to the cathode is calculated:

$$J_{N_2,cro} = K_{N_2}\frac{p_{N_2,c} - p_{N_2,a}}{\delta_m} \tag{8}$$

where K (mol·m^{-1}·s^{-1}·Pa^{-1}) is the permeability; p (Pa) is the pressure. The superscripts and subscript: cro represents crossover; a and c represent the anode and cathode, respectively.

The electro-chemical model represents the relationship between the current density and output voltage:

$$V = E_{rev} - RI - \frac{RT}{n\alpha F}\ln\left(\frac{I}{I_0}\right) \tag{9}$$

where V (V) is the output voltage; E (V) is the potential; R (Ω· m^2, 8.314 J·mol^{-1}·K^{-1}) is the area resistance, and universal gas constant; I (A·m^{-2}) is the current density; n is the number of electron transfer; α is the transfer coefficient; F (95,486 C·mol^{-1}) is Faraday constant. The superscripts and subscript: rev represents reversible; 0 represents the reference value.

The numerical process is done by the in-house C++ code. The simulation results of performance transients in the DEA and anode recirculation modes were compared with experimental results, and a good agreement was achieved that could be found in our previous work [14,21]. The employed model could well reflect the degradation process of the PEM fuel cell in the DEA and anode recirculation modes, and thus the simulated data can be used as the substitute for experimental data to develop the data-driven model.

2.2. ANN Model Development

When the hidden layer in the multilayer perceptron (MLP) has an appropriate activation function and a sufficient number of hidden layers and neurons, the MLP can estimate any function with any accuracy. The ANN used in this work is implemented by an MLP. The ANN architecture can be described as follows: the first layer of input variable feeding network is represented as the input layer, the final layer is called the output layer, and all layers between the input layer and output layer are called the hidden layer (as shown in Figure 1). All neural networks have an input layer and an output layer, but the number of hidden layers may vary.

In this study, two different ANN structures are explored and their performances are tested. They are ANN with one-hidden-layer and ANN with two-hidden-layer. The initial number of neurons in each hidden layer is less than 20 [33]. The exact number of neurons, the learning rate and the number of training epochs in the current ANN are tested by a hyperparameter selection procedure. The data set has 12 groups of complete time-series data, of which 8 groups are randomly divided into train set, 2 groups into validation set, and 2 groups into the test set. The train set is used to train the ANN model, the validation set is used to select hyperparameter and select the best model, and the test set is used to check the generalization ability of the selected ANN model. The number of neurons in the hidden layer is 3–20, the learning rate is 0.001, 0.004, 0.007, 0.01 and 0.04, and the number of training epochs is within 200. The code automatically saves the model when we get a better model during validation. The validation of the forecasted model is based on the mean and maximum relative errors. Learning and building the ANN is done using the PyTorch library of the interpreted programming language Python.

In DEA mode, PEM fuel cell works in the constant voltage mode, and the current density changes over time. In this study, 12 processes of the PEM fuel cell in the DEA mode with different operating conditions (voltage and inlet pressure) are simulated (as shown in Table 2). The duration for the DEA mode of each process is 200 s. The sampling time is 1 s. Therefore, the input layer of ANN has three neurons. When in DEA mode, they receive the following three parameters.

1. Voltage;
2. Inlet pressure;
3. Time.

Table 2. Operating conditions of the twelve processes in the DEA mode.

	Train Set					
Voltage, V	0.7	0.6	0.6	0.7	0.5	0.4
Inlet pressure, bar	1.6	1.4	1.3	1.0	1.2	1.8
	Train Set		**Validation Set**		**Test Set**	
Voltage, V	0.4	0.5	0.5	0.6	0.6	0.7
Inlet pressure, bar	1.3	2.0	1.6	1.6	1.1	1.3

In anode recirculation mode, PEM fuel cell works in the constant current mode, and the voltage changes over time. In this study, 12 processes of the PEM fuel cell in the anode recirculation mode with different operating conditions (current density and anode stoichiometry) are simulated (as shown in Table 3). The duration for the anode recirculation mode of each process is 600 s. The sampling time is 1 s. When in anode recirculation mode, they receive the following three parameters.

1. Current density;
2. Anode stoichiometry;
3. Time.

Table 3. Operating conditions of the twelve processes in the anode recirculation mode.

	Train Set					
Current density, A·cm^{-2}	1.6	1.5	0.4	1.2	1.2	1.1
Anode stoichiometry	1.3	1.5	2.0	1.5	2.4	1.8
	Train Set		Validation Set		Test Set	
Current density, A·cm^{-2}	0.8	1.0	0.8	1.3	1	0.6
Anode stoichiometry	1.2	3.0	2.4	1.4	1.6	1.8

The input values are normalized before input. This is conducive to the improvement of the training speed and numerical performance of the ANN, so that the error in the training process can be smoothly reduced, and the accuracy of gradient estimation can be improved.

2.3. Multivariate Polynomial Regression (MPR) Method

MPR method is a regression analysis method to study the polynomial relationship between a dependent variable (target) and several independent variables (input). The greatest advantage of the MPR method is that it can approach the measured points by adding the high-order term of the independent variable until the results are satisfactory. In this paper, the first to fifth order equations are considered for the MPR method. The regression equations of the following two independent variables are used to predict the initial values:

First order equation:

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 \tag{10}$$

Second order equation:

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_1^2 + \alpha_4 x_1 x_2 + \alpha_5 x_2^2 \tag{11}$$

When in DEA mode, where y is the predicted value of the initial current density, and x_1 and x_2 are the normalized voltage and inlet pressure, respectively. When in anode recirculation mode, where y is the predicted value of the initial voltage, and x_1 and x_2 are the normalized current density and Anode stoichiometry, respectively. α_i (i = 0, 1, 2, 3, 4, 5) is the corresponding regression coefficient which is calculated using statistical methods. After comparing the accuracy, we choose the third order equation of the MPR method for modeling. The development of the MPR model is done by using the scikit-learn library of the interpreted programming language Python.

3. Results and Discussion

3.1. Performance Decay Prediction of the DEA Mode

In this section, the performance decay prediction of the PEM fuel cell in the DEA mode is evaluated. Firstly, the prediction performance of the ANN and M-ANN is compared to select the better modeling method. Then, the MPR method is employed as the benchmark model to demonstrate the necessity of ANN in modeling. Lastly, the effect of hidden layer numbers and activation functions on the prediction performance is discussed. We introduce two errors to evaluate the prediction performance, mean relative error δ_{mean} and maximum relative error δ_{max}:

$$\delta_{\text{mean}} = \frac{1}{n} \sum_{i=1}^{n} \frac{|I_i - \hat{I}_i|}{I_i} \times 100\% \tag{12}$$

$$\delta_{\text{max}} = \max\left\{ \frac{|I_i - \hat{I}_i|}{I_i} \right\} \times 100\% \tag{13}$$

where n is the number of data points; I_i (A·cm^{-2}) is the true value of current density; $\hat{I}_i$ (A·cm^{-2}) is the predictive value of current density.

3.1.1. Comparison of the ANN and M-ANN

In the DEA mode, the PEM fuel cell performance firstly increases due to the self-humidification effect, and then gradually decreases due to the nitrogen crossover and accumulation in the anode, as shown for the data in the 12 processes. More physical explanation could be found in our previous work [21]. Comparison of one-hidden-layer ANN and one-hidden-layer M-ANN predictive values and true values of current density change evolution of the PEM fuel cell in the DEA mode are shown in Figures 2 and 3, respectively. Comparison of two-hidden-layer ANN and two-hidden-layer M-ANN predictive values and true values of current density change evolution of the PEM fuel cell in the DEA mode are shown in Figures 4 and 5. (a)–(h), (i,j) and (k,l) represent the train set, validation set, and test set, respectively. The trend of current density firstly increasing and then decreasing is predicted by the ANN-based model for all the 12 processes. It is worth mentioning that it takes around 4 hours' calculation for the small workstation to obtain the best one-hidden-layer ANN model through grid search, while it takes around 70 h to obtain the best two-hidden-layer ANN model. Prediction performance, including the mean and maximum relative errors of different modeling methods for the DEA mode is listed in Table 4.
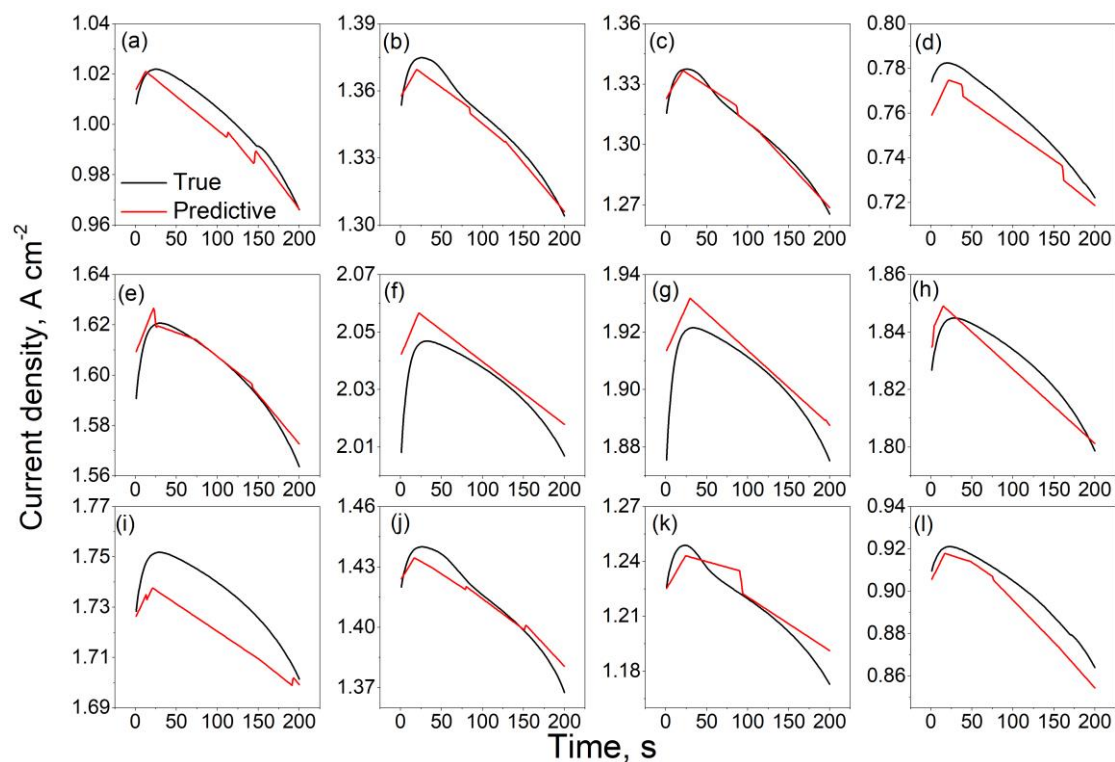

Figure 2. One-hidden-layer ANN with ReLU activation function. Comparison of predictive values and true values of current density change evolution for the PEM fuel cell with the DEA mode. (**a**–**l**) represent the 12 transient processes. (**a**–**h**), (**i**,**j**) and (**k**,**l**) are the train set, validation set, and test set, respectively.
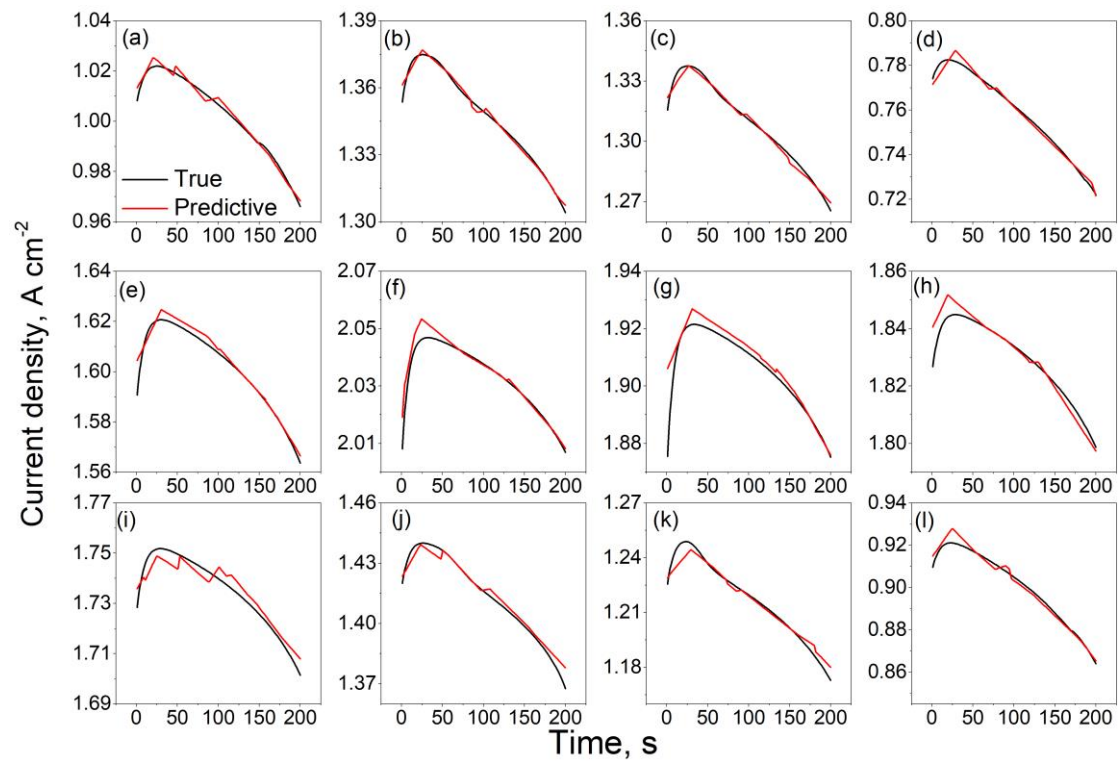
Figure 3. One-hidden-layer M-ANN with ReLU activation function. Comparison of predictive values and true values of current density change evolution for the PEM fuel cell with the DEA mode. (**a**–**l**) represent the 12 transient processes. (**a**–**h**), (**i**,**j**) and (**k**,**l**) are the train set, validation set, and test set, respectively.

Figure 4. Two-hidden-layer ANN with ReLU activation function. Comparison of predictive values and true values of current density change evolution for the PEM fuel cell with the DEA mode. (**a**–**l**) represent the 12 transient processes. (**a**–**h**), (**i**,**j**) and (**k**,**l**) are the train set, validation set, and test set, respectively.
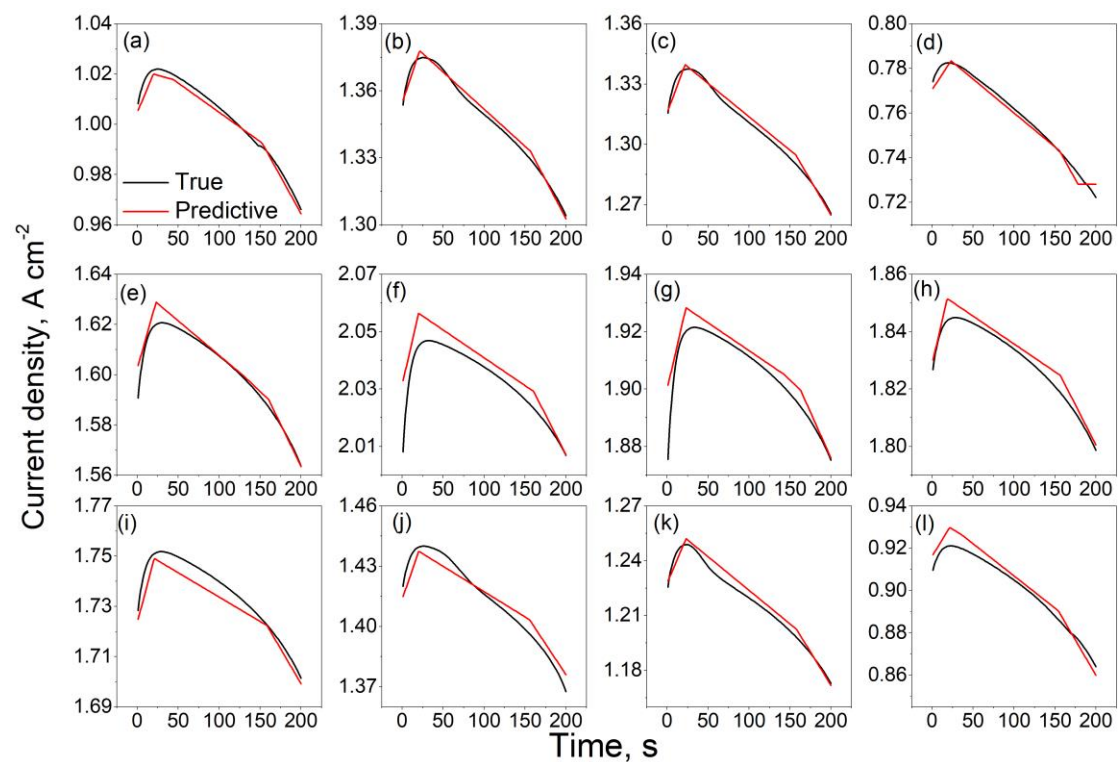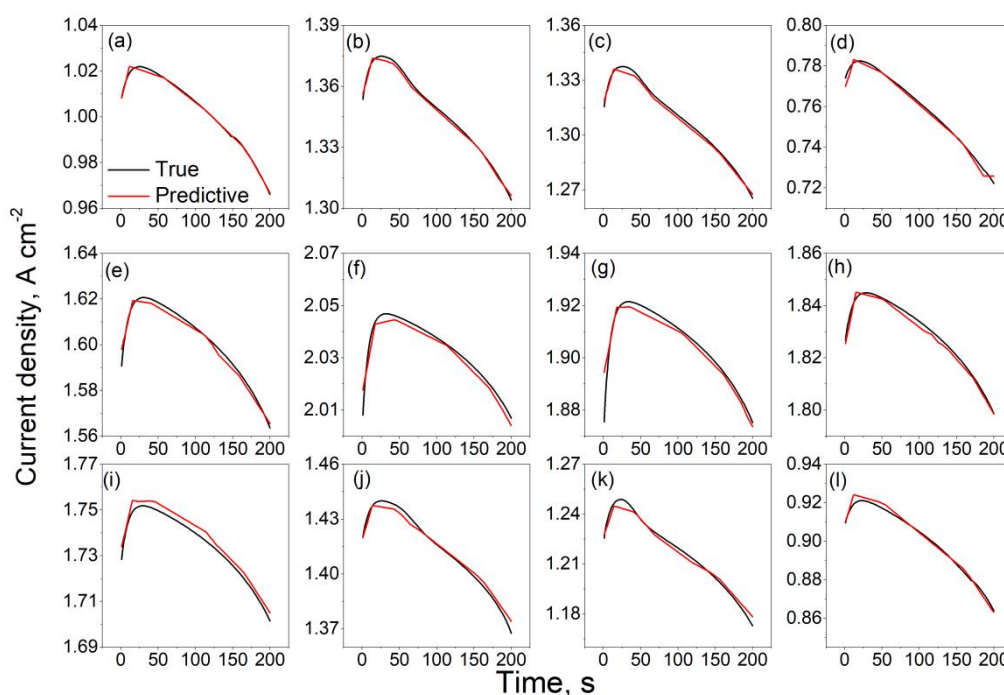
Figure 5. Two-hidden-layer M-ANN with ReLU activation function. Comparison of predictive values and true values of current density change evolution for the PEM fuel cell with the DEA mode. (**a**–**l**) represent the 12 transient processes. (**a**–**h**), (**i**,**j**) and (**k**,**l**) are the train set, validation set, and test set, respectively.

Table 4. Prediction performance (mean and maximum relative error) of different modeling methods in the train, validation and test set for the DEA mode.

	Mean/Maximum Relative Error, %		
	Train	**Validation**	**Test**
one-hidden-layer ANN (ReLU)	0.405/2.037	0.625/1.122	0.725/1.563
one-hidden-layer M-ANN (ReLU)	0.119/1.632	0.180/0.759	0.198/0.803
two-hidden-layer ANN (ReLU)	0.186/1.381	0.291/0.665	0.341/0.932
two-hidden-layer M-ANN (ReLU)	0.094/1.010	0.146/0.484	0.158/0.534
MPR-MPR	0.228/2.210	0.264/1.263	0.348/1.510
two-hidden-layer M-ANN (sigmoid)	0.074/0.822	0.122/0.444	0.157/0.617

The mean and maximum relative errors by one-hidden-layer ANN are 0.725% and 1.563% that are much larger than that by one-hidden-layer M-ANN of 0.198% and 0.803%. The mean and maximum relative errors by two-hidden-layer ANN are 0.341% and 0.932% that are much larger than that by two-hidden-layer M-ANN of 0.158% and 0.534%. Although the prediction degradation process of M-ANN includes the error of the MPR model to predict the initial current density and the error of the ANN model to predict the current density change. However, from the results of M-ANN, this is tolerable. It can be seen that the prediction curves by ANN show more obvious error compared with the true curves than those by M-ANN. The reason why M-ANN can obtain better prediction results than ANN may be as follows:

1. The MPR method solves the problem of initial current density prediction under various working conditions. This is because each operating condition has a different starting point for its current density due to different initial conditions, which is the reason for the poor effectiveness of ANN in predicting the overall degraded current density. ANN can also be used to predict initial current density at various operating conditions. the initial current density dataset has only 12 samples. It is well known that the ANN model needs a large amount of data to be trained to get better results.

Therefore, when the amount of data is small, it is a wise choice to use the MPR method to predict results.

2. The ANN solves the problem of predicting current density change. After eliminating the influence of the initial point, the current density changes under various operating conditions are similar. However, this change is relatively complex, not only rise or fall, but also the magnitude and time of rise and fall are uncertain. Therefore, with a large number of samples, a relatively complex ANN is more effective in learning the current density change. Figure 6 shows the result of MPR method predicting current density change. MPR-MPR means to use MPR to predict the initial current density and then MPR to predict the current density change. Compared with Figures 5 and 6, ANN is better at predicting the current density change than MPR.
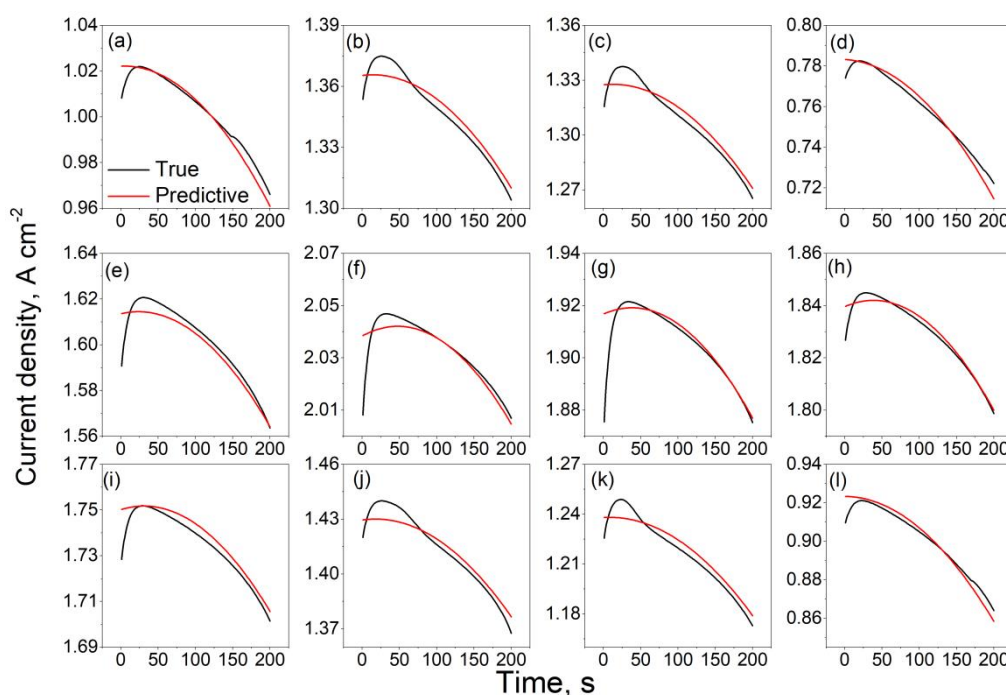

Figure 6. MPR-MPR. Comparison of predictive values and true values of current density change evolution for the PEM fuel cell with the DEA mode. (**a–l**) represent the 12 transient processes. (**a–h**), (**i,j**) and (**k,l**) are the train set, validation set, and test set, respectively.

Overall, this study decomposes the predictive degradation process into two different types of problems, and then gives the two problems to two algorithms to solve, which achieve the complementary advantages between algorithms, resulting in better performance than a single algorithm to solve a single problem.

3.1.2. Comparison of Hidden Layer Numbers and Activation Functions

Comparison of one-hidden-layer ANN and that of two-hidden-layer ANN predictive values and true values of current density change evolution for the PEM fuel cell with the DEA mode are shown in Figures 2 and 4. The mean and maximum relative errors by one-hidden-layer ANN are 0.725% and 1.563% that are much larger than that by two-hidden-layer ANN of 0.341% and 0.932%. Comparison of one-hidden-layer M-ANN and that of two-hidden-layer M-ANN predictive values and true values of current density change evolution for the PEM fuel cell with the DEA mode are shown in Figures 3 and 5. The mean and maximum relative errors by one-hidden-layer M-ANN are 0.198% and 0.803% that are slightly larger than that by two-hidden-layer M-ANN of 0.158% and 0.534%. From the above two groups of comparison, it is obvious that the performance of two-hidden-layer is better than that of one-hidden-layer, which is a predictable result. Because

the two-hidden-layer neural network is more complex than the one-hidden-layer, it can learn deeper knowledge from the data, which leads to superior prediction results.

Figure 7 shows the comparison between the predictive values of two-hidden-layer M-ANN with sigmoid activation function and the true values of current density evolution in DEA mode. The mean and maximum relative errors by two-hidden-layer M-ANN with sigmoid activation function are 0.157% and 0.617%. Comparing Figures 5 and 7, it can be seen that the prediction curves obtained by using sigmoid activation function are smoother than those obtained by using rectified linear unit (ReLU) activation function.
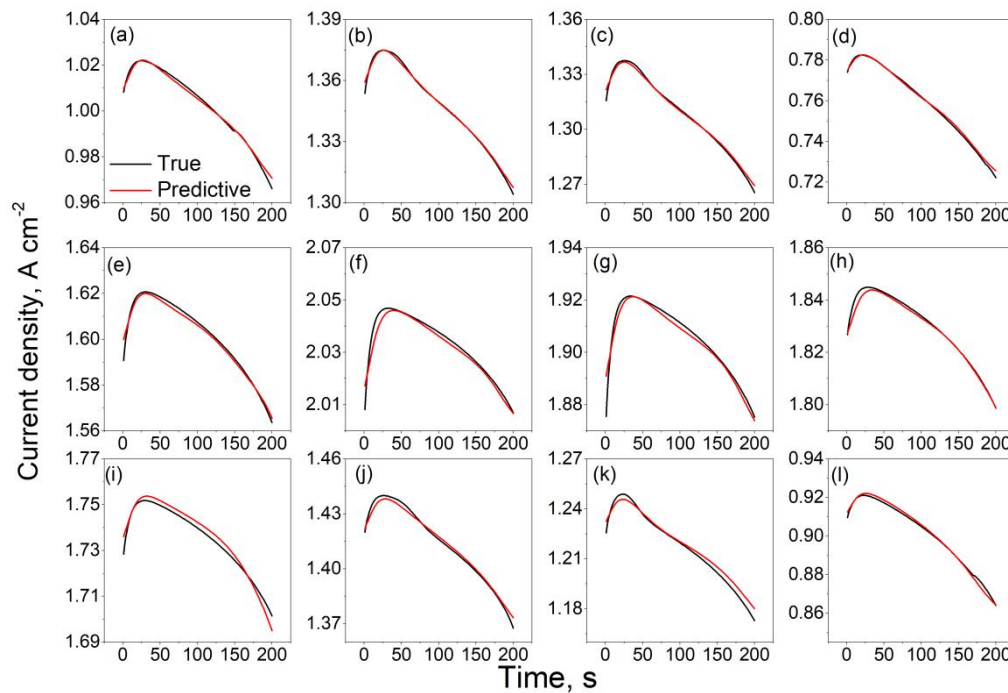

Figure 7. Two-hidden-layer M-ANN with sigmoid activation function. Comparison of predictive values and true values of current density change evolution for the PEM fuel cell with the DEA mode. (**a**–**l**) represent the 12 transient processes. (**a**–**h**), (**i**,**j**) and (**k**,**l**) are the train set, validation set, and test set, respectively.

Sigmoid activation function:

$$S(x) = \frac{1}{1 + e^{-x}} \tag{14}$$

ReLU activation function:

$$ReLU(x) = \max(0, x) \tag{15}$$

As can be seen from Figure 8, the ReLU activation function outputs 0 at $x < 0$ and x at $x > 0$. The sudden change of output at $x = 0$ is the reason for the unsmooth time-series output. The sigmoid activation function is smooth in the whole range, so its time-series output is also smooth. ReLU activation function has widely replaced sigmoid activation function in the field of deep learning for three reasons:

1. The sigmoid activation function is computationally intensive while the ReLU activation function is much less so when the backpropagation algorithm is solving for the gradient.
2. For deep neural networks with sigmoid activation function, the vanishing gradient problem can easily occur when the backpropagation algorithm is solving for the gradient.

3. The ReLU activation function can make the output of some neurons zero, which will cause the sparsity of the neural network, reduces the interdependence of parameters and alleviates the overfitting problem.
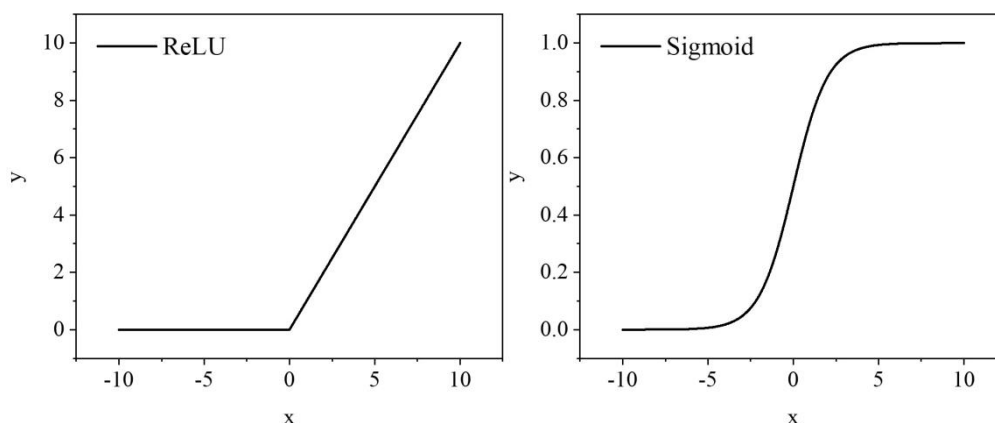

Figure 8. ReLU activation function and sigmoid activation function.

It can be seen from the results in this paper that in the case of a small number of neurons, the ReLU activation function will make the time-series results unsmooth. Obviously, a smooth prediction curve is more representative of the reality. In this respect, the sigmoid activation function has an advantage over the ReLU activation function. Although from the perspective of mean relative error, there is almost no difference in performance between the two in this study.

3.2. Performance Decay Prediction of the Anode Recirculation Mode

Different from DEA mode, the independent variable in anode recirculation mode is changed to current density, anode stoichiometry and time, and the dependent variable is changed to voltage. The operating conditions of the 12 processes in anode recirculation mode are shown in Table 3. Comparison of two-hidden-layer M-ANN with ReLU activation function and that of two-hidden-layer M-ANN with sigmoid activation function predictive values and true values of voltage change evolution for the PEM fuel cell with the anode recirculation mode are shown in Figures 9 and 10. The mean and maximum relative errors by two-hidden-layer M-ANN with ReLU activation function are 0.143% and 0.458%. The mean and maximum relative errors by two-hidden-layer M-ANN with sigmoid activation function are 0.155% and 0.359%.

Changing the independent and dependent variables is to explore whether the M-ANN can continue to perform well when different independent variables are used. The comparison between Tables 3 and 5 shows that M-ANN has strong robustness. Compared with DEA mode, the performance change evolution in anode recirculation mode is simpler, with the slope of rising or fall remaining almost constant, and the voltage changes little. However, the mean relative error in anode recirculation mode is similar to that of DEA mode. In this case, we adjust the number of training epochs to within 2000. However, it does not yield a superior model. The most likely reason is that the prediction performance of two-hidden-layer M-ANN has reached the limit. Perhaps when using a three-hidden-layer M-ANN model, the prediction results will be better. However, the two-hidden-layer M-ANN is more complex than the three-hidden-layer M-ANN, so it needs more data to train the model to get better results.
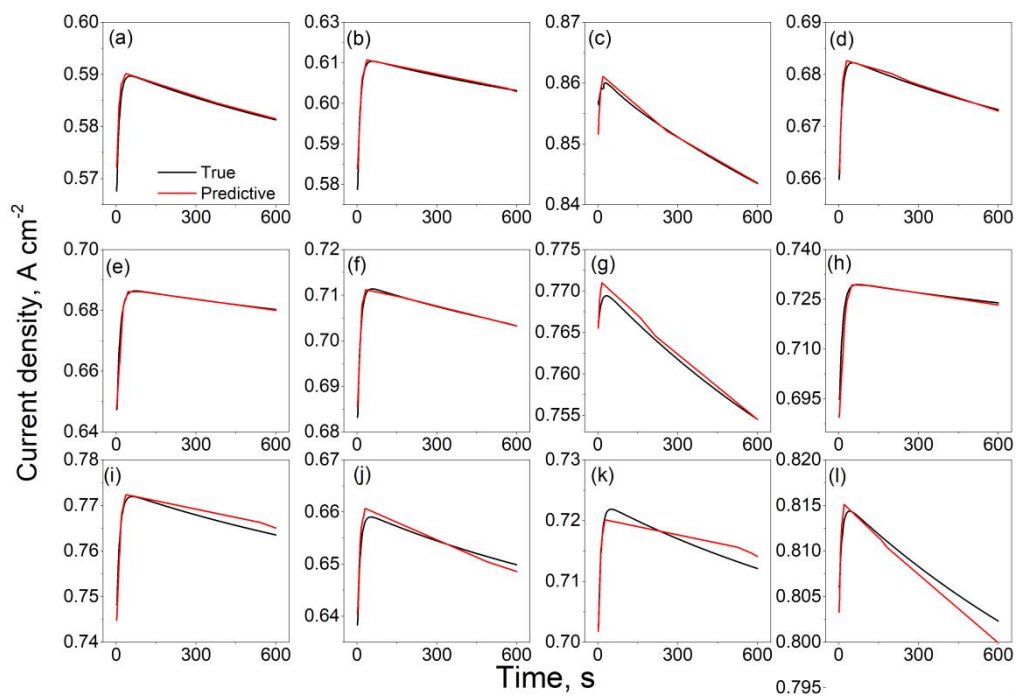
Figure 9. Two-hidden-layer M-ANN with ReLU activation function. Comparison of predictive values and true values of current density change evolution for the PEM fuel cell with the anode recirculation mode. (**a–l**) represent the 12 transient processes. (**a–h**), (**i,j**) and (**k,l**) are the train set, validation set, and test set, respectively.
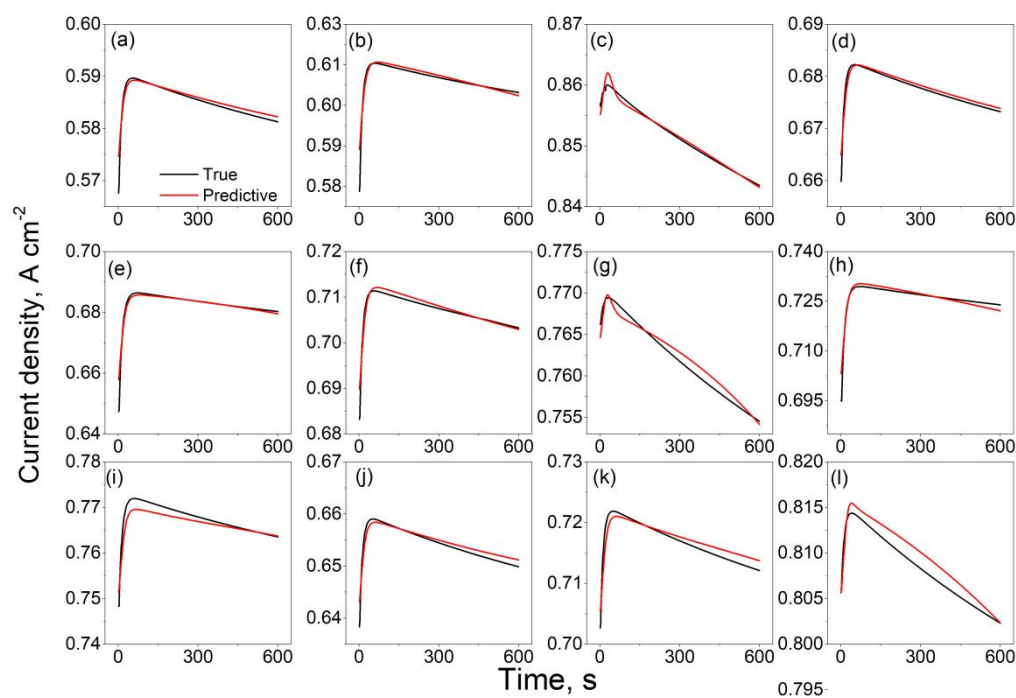

Figure 10. Two-hidden-layer M-ANN with sigmoid activation function. Comparison of predictive values and true values of current density change evolution for the PEM fuel cell with the anode recirculation mode. (**a–l**) represent the 12 transient processes. (**a–h**), (**i,j**) and (**k,l**) are the train set, validation set, and test set, respectively.

Table 5. Prediction performance (mean and maximum relative error) of different modeling methods in the train, validation and test set for the anode recirculation mode.

	Mean/Maximum Relative Error, %		
	Train	**Validation**	**Test**
two-hidden-layer M-ANN (ReLU)	0.082/1.768	0.128/0.768	0.143/0.458
two-hidden-layer M-ANN (sigmoid)	0.051/1.351	0.139/0.518	0.155/0.359

4. Conclusions

In this study, we propose the M-ANN method to simulate the short-term degradation behaviors of the PEM fuel cell. Two cases of degradation data, the PEM fuel cell in the DEA and anode recirculation modes, are employed to train the model. The results show that the predictive performance by M-ANN is significantly better than that only by ANN or MPR. Firstly, the MPR method solves the problem of initial values prediction under various working conditions, and this problem cannot be easily solved by ANN because the sample size is extremely small. Secondly, the ANN solves the problem of predicting the variation value of the cell performance in the degradation process, and meanwhile the prediction effect of the MPR is much worse than that of the ANN. Then, we investigate the effect of the hidden-layer number on the predictive performance of M-ANN, and it is found that the performance of two-hidden-layer is better than that of one-hidden-layer. In addition, we also compare the effects of different activation functions, ReLU and sigmoid, on the predictive performance of M-ANN. Although each of them does not show an obvious advantage over the other on the errors, the predictive curves by using the sigmoid activation function are smoother than that by using ReLU activation function. Obviously, the smooth prediction curve is more realistic. Generally speaking, the M-ANN model shows good predictive performance in both two cases, which proves that M-ANN has good universality on the short-term degradation or dynamic behaviors of the PEM fuel cell.

Author Contributions: Conceptualization: B.W. and K.J.; Methodology: Z.Y. and B.W.; Investigation: B.W., X.S. and Y.W.; Writing—original draft preparation: Z.Y. and B.W.; Writing—review and editing: Q.R., S.H., J.X. and K.J.; Supervision: S.H., J.X. and K.J.; Project administration: K.J.; Funding acquisition: K.J. All authors have read and agreed to the published version of the manuscript.

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

Nomenclature

T	temperature, K
u	flow rate, $\text{m} \cdot \text{s}^{-1}$
RH	relative humidity
GDL	gas diffusion layer
MPL	micro-porous layer
CL	catalyst layer
L	length of the channel, cm
W	width of the channel, mm
d	depth of the channel, mm
A	area, m^2
K	permeability, $\text{mol} \cdot \text{m}^{-1} \cdot \text{s}^{-1} \cdot \text{Pa}^{-1}$
I	current density, $\text{A} \cdot \text{m}^{-2}$

c	reactant concentration, $mol \cdot m^{-3}$
S	source term, $mol \cdot m^{-2} \cdot s^{-1}$
N	the number of control volumes along the channel
J	the convection flux along the channel, $mol \cdot m^{-2} \cdot s^{-1}$
M	molar weight, $kg \cdot mol^{-1}$
s	liquid volume fraction
EW	equivalent weight of the membrane, $kg \cdot mol^{-1}$
C_p	specific heat capacity, $J \cdot kg^{-1} \cdot K^{-1}$
P	pressure, Pa
V	voltage, V
E	potential, V
R	area resistance, $\Omega \cdot m^2$; universal gas constant, $8.314\ J \cdot K^{-1} \cdot mol^{-1}$
n	the number of electron transfer
F	Faraday constant, $96{,}485\ C \cdot mol^{-1}$
ReLU	rectified linear unit

Greek Letter

δ	thickness, m
ε	porosity
θ	contact angle, $\circ$
ω	ionomer volume fraction of the catalyst layer
α	transfer coefficient
ϕ	diffusion flux, $mol \cdot m^{-2} \cdot s^{-1}$
ρ	density, $kg \cdot m^{-3}$
λ	membrane water content

Subscript and Superscript

0	reference value
a	anode
c	cathode
GDL	gas diffusion layer
MPL	micro-porous layer
CL	catalyst layer
m	membrane
act	activation
CH	channel
in	flow inlet; inlet surface of the control volume
out	flow outlet
ref	reference
k	time step
n	control volume along the channel
i	gas species
CH-GDL	interface between the channel and gas diffusion layer
lq	liquid water
w	water
p	components of the electrode
m	membrane
mw	membrane water
T	Energy
eff	effective coefficient
cro	crossover
rev	reversible

Abbreviation

DEA	dead-ended anode
PEM	proton exchange membrane
MPR	multivariate polynomial regression
ANN	artificial neural network
CFD	computational fluid dynamics
BO	bonobo optimizer
LSTM	long short-term memory

RNN	recurrent neural network
GRU	gated recurrent unit
ESN	echo state network
GNNM	grey neural network model
PSO	particle swarm optimization

References

1. Jiao, K.; Ni, M. Challenges and opportunities in modelling of proton exchange membrane fuel cells (PEMFC). *Int. J. Energy Res.* **2017**, *41*, 1793–1797. [CrossRef]
2. Wang, Y.; Seo, B.; Wang, B.; Zamel, N.; Jiao, K.; Adroher, X.C. Fundamentals, materials, and machine learning of polymer electrolyte membrane fuel cell technology. *Energy AI* **2020**, *1*, 100014. [CrossRef]
3. Wang, Y.; Yuan, H.; Martinez, A.; Hong, P.; Bockmiller, F.R. Polymer electrolyte membrane fuel cell and hydrogen station networks for automobiles: Status, technology, and perspectives. *Adv. Appl. Energy* **2021**, *2*, 100011. [CrossRef]
4. Luo, Y.; Jiao, K. Cold start of proton exchange membrane fuel cell. *Prog. Energy Combust. Sci.* **2018**, *64*, 29–61. [CrossRef]
5. Yang, Y.; Zhang, X.; Guo, L.; Liu, H. Local degradation in proton exchange membrane fuel cells with dead-ended anode. *J. Power Sources* **2020**, *477*, 229021. [CrossRef]
6. Wang, B.; Deng, H.; Jiao, K. Purge strategy optimization of proton exchange membrane fuel cell with anode recirculation. *Appl. Energy* **2018**, *225*, 1–13. [CrossRef]
7. Rabbani, A.; Rokni, M.; Hosseinzadeh, E.; Mortensen, H.H. The Start-Up Analysis of a PEM Fuel Cell System in Vehicles. *Int. J. Green Energy* **2014**, *11*, 91–111. [CrossRef]
8. Pei, P.; Chen, H. Main factors affecting the lifetime of Proton Exchange Membrane fuel cells in vehicle applications: A review. *Appl. Energy* **2014**, *125*, 60–75. [CrossRef]
9. He, L.; Zhan, Z.; Chen, H.; Jiang, P.; Yu, Y.; Yang, X.; Sun, Y.; Wan, X.; Liao, L.; Li, S.; et al. A quick evaluation method for the lifetime of the fuel cell MEA with the particle filter algorithm. *Int. J. Green Energy* **2021**, *18*, 1–14. [CrossRef]
10. Ren, P.; Pei, P.; Li, Y.; Wu, Z.; Chen, D.; Huang, S. Degradation mechanisms of proton exchange membrane fuel cell under typical automotive operating conditions. *Prog. Energy Combust. Sci.* **2020**, *80*, 100859. [CrossRef]
11. Jiao, K.; Li, X. Water transport in polymer electrolyte membrane fuel cells. *Prog. Energy Combust. Sci.* **2011**, *37*, 221–291. [CrossRef]
12. Kurnia, J.C.; Sasmito, A.P.; Shamim, T. Advances in proton exchange membrane fuel cell with dead-end anode operation: A review. *Appl. Energy* **2019**, *252*, 113416. [CrossRef]
13. Hou, J.; Yang, M.; Zhang, J. Active and passive fuel recirculation for solid oxide and proton exchange membrane fuel cells. *Renew. Energy* **2020**, *155*, 1355–1371. [CrossRef]
14. Wang, B.; Wu, K.; Xi, F.; Xuan, J.; Xie, X.; Wang, X.; Jiao, K. Numerical analysis of operating conditions effects on PEMFC with anode recirculation. *Energy* **2019**, *173*, 844–856. [CrossRef]
15. Peng, Y.; Mahyari, H.M.; Moshfegh, A.; Javadzadegan, A.; Toghraie, D.; Shams, M.; Rostami, S. A transient heat and mass transfer CFD simulation for proton exchange membrane fuel cells (PEMFC) with a dead-ended anode channel. *Int. Commun. Heat Mass Transf.* **2020**, *115*, 104638. [CrossRef]
16. Rizvandi, O.B.; Yesilyurt, S. A Transient Pseudo-3D model of the PEM Fuel Cell for the Analysis of Dead-Ended Anode and Anode Bleeding Operation Modes. *Electrochim. Acta.* **2019**, *324*, 134866. [CrossRef]
17. Xu, L.; Fang, C.; Li, J.; Ouyang, M.; Lehnet, W. Nonlinear dynamic mechanism modeling of a polymer electrolyte membrane fuel cell with dead-ended anode considering mass transport and actuator properties. *Appl. Energy* **2018**, *230*, 106–121. [CrossRef]
18. Luo, Y.; Jiao, K.; Jia, B. Elucidating the constant power, current and voltage cold start modes of proton exchange membrane fuel cell. *Int. J. Heat Mass Transf.* **2014**, *77*, 489–500. [CrossRef]
19. Zhou, Y.; Luo, Y.; Yu, S.; Jiao, K. Modeling of cold start processes and performance optimization for proton exchange membrane fuel cell stacks. *J. Power Sources* **2014**, *247*, 738–748. [CrossRef]
20. Sultan, H.M.; Menesy, A.S.; Kamel, S.; Tostado-Véliz, Z.; Jurado, F. Parameter identification of proton exchange membrane fuel cell stacks using bonobo optimizer. In Proceedings of the 2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe), Milan, Italy, 9–12 June 2020; pp. 1–7.
21. Wang, B.; Wu, K.; Yang, Z.; Jiao, K. A quasi-2D transient model of proton exchange membrane fuel cell with anode recirculation. *Energy Convers. Manag.* **2016**, *129*, 108–121. [CrossRef]
22. Kheirandish, A.; Motlagh, F.; Shafiabady, N.; Dahari, M. Dynamic modelling of PEM fuel cell of power electric bicycle system. *Int. J. Hydrogen Energy* **2016**, *41*, 9585–9594. [CrossRef]
23. Zou, W.; Froning, D.; Shi, Y.; Lehnert, W. A least-squares support vector machine method for modeling transient voltage in polymer electrolyte fuel cells. *Appl. Energy* **2020**, *271*, 115092. [CrossRef]
24. Gu, X.; Hou, Z.; Cai, J. Data-based flooding fault diagnosis of proton exchange membrane fuel cell systems using LSTM networks. *Energy AI* **2021**, *4*, 100056. [CrossRef]
25. Lin, T.; Hu, L.; Wisely, W.; Gu, X.; Litster, S.; Kara, L.B. Prediction of high frequency resistance in polymer electrolyte membrane fuel cells using long short term memory based model. *Energy AI* **2021**, *3*, 100045. [CrossRef]

26. Wang, B.; Xie, B.; Xuan, J.; Jiao, K. AI-based optimization of PEM fuel cell catalyst layers for maximum power density via data-driven surrogate modeling. *Energy Convers. Manag.* **2020**, *205*, 112460. [CrossRef]
27. Wang, B.; Zhang, G.; Wang, H.; Xuan, J.; Jiao, K. Multi-physics-resolved digital twin of proton exchange membrane fuel cells with a data-driven surrogate model. *Energy AI* **2020**, *1*, 100004. [CrossRef]
28. Xie, R.; Ma, R.; Pu, S.; Xu, L.; Zhao, D.; Huangfu, Y. Prognostic for fuel cell based on particle filter and recurrent neural network fusion structure. *Energy AI* **2020**, *2*, 100017. [CrossRef]
29. 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]
30. 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]
31. Chen, K.; Laghrouche, S.; Djerdir, A. Degradation prediction of proton exchange membrane fuel cell based on grey neural network model and particle swarm optimization. *Energy Convers. Manag.* **2019**, *195*, 810–818. [CrossRef]
32. Buzau, M.M.; Tejedor-Aguilera, J.; Cruz-Romero, P.; Gómez-Expósito, A. Hybrid deep neural networks for detection of non-technical losses in electricity smart meters. *IEEE Trans. Power Syst.* **2019**, *35*, 1254–1263. [CrossRef]
33. Elisseeff, A.; Paugam-Moisy, H. Size of multilayer networks for exact learning: Analytic approach. *Adv. Neural Inf. Process. Syst.* **1996**, *9*, 162–168.