Applied Energy Symposium 2021: Low carbon cities and urban energy systems
September 4-8, 2021, Matsue, Japan
Paper ID:29

Diagnosis and prediction of local gas starvation in proton exchange membrane fuel cells based on machine learning method

Huicui Chen[1*], Lebin Chu[1], Tong Zhang[1], Pucheng Pei[2]

1 School of Automotive Studies, Tongji University, Shanghai 201804, China

2 State Key Laboratory of Automotive Safety and Energy, Tsinghua University, Beijing 100084, China

ABSTRACT

The life and durability problems of proton exchange membrane fuel cell (PEMFC) have limited the commercialization process. The main reason for the degradation of life due to the frequent occurrence of local gas starvation in the dynamic process. The existing research is mainly carried out by experiment or simulation to diagnose local gas starvation, there is almost no research using machine learning methods to predict the local gas starvation through operating parameters. To solve this problem, a snake-shaped five-channel PEMFC model is established in this paper, and obtained source data through CFD simulation. Principal component analysis and k-means clustering algorithm are used to effectively define the local gas starvation state of each sample point, and complete sample labeling. Five operating parameters (temperature, pressure, humidity, gas stoichiometric ratio and current density), were used as model inputs. Three machine learning methods are chosen for training and prediction, and compare their accuracy. The prediction accuracy rate based on the extreme learning machine regression model is the highest, which is 93.49%, and have a fast prediction speed. It can quickly and accurately predict the local gas starvation state under a certain working condition, which has guiding significance for the optimization of operating parameters in the fuel cell control process.

Keywords: Proton exchange membrane fuel cells; Diagnosis and prediction of local gas starvation; Machine learning

1. INTRODUCTION

To solve the problem of energy shortage and environmental pollution, many countries around the world have strengthened the research and development of new energy vehicles. Proton exchange membrane fuel cell (PEMFC) uses hydrogen as fuel, has the advantages of high efficiency, zero emissions, rapid refueling and low operating temperature[1], its development prospects are widely recognized. When PEMFC is used as a fixed power source, the service life can reach 30,000 hours[2], while used as a vehicle power source, the service life is generally about 3,000 hours [3], which severely restricts the large-scale commercialization of fuel cells. Studies have shown that frequent local gas starvation during dynamic process are the main cause of fuel cell life degradation[4]. Therefore, the diagnosis and prediction of local gas starvation in PEMFC is of great significance to improve the life of fuel cells.

Many scholars have conducted research on the diagnosis of local gas starvation in fuel cells. At present, the methods used for the diagnosis mainly include: voltage monitoring method[5], current density difference method[6], air excess coefficient method[7] and visualization method[8].

In the process of fuel cell operation, the external characteristic parameters such as fuel cell output parameters and operating parameters are easily collected. The effective diagnosis of local gas starvation in fuel cells through external characteristic parameters is helpful to reduce the occurrence of starvation, and there are still deficiencies in this type of diagnosis method. Using a suitable method to predict the local gas starvation is of great significance for improving the life of the fuel cell. The prediction of the local gas starvation state through external characteristic parameters is a multivariate nonlinear prediction with tags. Using machine learning methods, a suitable regression model can be selected, to achieve rapid prediction of the local gas starvation state of the fuel cell under a certain working condition.

Based on the previous analysis, this paper established a three-dimensional fuel cell model, and obtained source data through simulation. The gas starved area ratio is used as a measure of the degree of local gas starvation. The local gas starvation state of the fuel cell is effectively diagnosed based on the principal component analysis method and the K-means clustering algorithm. Then each sample was marked according to the clustering results, and 5 operating parameters (temperature, pressure, humidity, gas stoichiometric ratio and current density) were selected, compared a variety of machine learning methods, and studied the method of predicting the local gas starvation state of fuel cells through operating parameters. Accurate prediction models can effectively and quickly predict the local gas starvation state of fuel cells under certain operating conditions, provide guidance for optimizing the control and design parameters of fuel cells, which are of great significance for improving the life and durability of fuel cells.

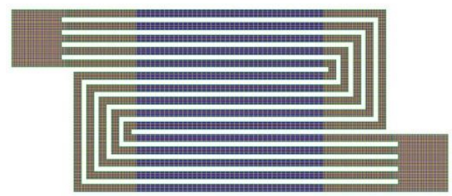
2. MODEL DEVELOPMENT

First, a snake-shaped five-channel PEMFC model is established in this paper, which can not only accurately describe the internal mass transfer and heat transfer process of the fuel cell, and observe macroscopic physical quantities such as voltage, but also obtain the relevant gas distribution under different operating conditions, so as to calculate the proportion of gas-starved area and measure the degree of local gas starvation. Operating parameters are randomly generated to ensure the uniformity of machine learning data.

2.1 Geometry and mesh

The PEMFC geometric model can be divided into three parts, including the anode, the proton exchange membrane, and the cathode. Both the anode and the cathode are composed of a plate, a flow channel, a gas diffusion layer and a catalyst layer, and are separated by a proton exchange membrane. Based on the assumptions of the model and the basic equations followed in [9], a snake-shaped five-channel PEMFC model is established, as shown in Figure 1(a). The geometric parameters are shown in Table 1. This paper used a hexahedral grid to divide the simulation model, divide and combine the grids of the flow channel, proton exchange membrane, catalytic layer, gas diffusion layer, electrode plate and gas inlet and outlet positions, as shown in Figure 1(b). By setting appropriate parameters, subsequent simulation calculations can be converged in a short time.

(a)



(b)

Fig 1 Geometry and mesh

Table 1 Geometric parameters

Ridge width	Channel width	Channel thickness	Land width
0.8mm	1.2mm	0.8mm	32mm
Membrane thickness	GDL thickness	Catalyst Layer thickness	Graphite Plates thickness
0.036mm	0.200mm	0.012mm	1.2mm

2.2 Simulation conditions and model parameters

In this paper, the independent variables are temperature, pressure, humidity, gas stoichiometric ratio, and current density, the physical quantities related to the operating state of the fuel cell are used as dependent variables. First, the above operating parameters are randomly combined within the range of reasonable working conditions， to obtain the original random working condition sample points. Then, the three-dimensional modeling simulation is carried out under the working condition of each sample point, other physical quantities of the fuel cell are obtained. The simulation sample points can cover as many working conditions as possible, so that the diagnosis conclusions obtained by the clustering model are more effective and accurate.

The range and resolution of each parameter of operating conditions are defined as follows:

Temperature: 30~90°C, resolution is 1°C;

Cathode inlet pressure: 0.3~3 bar, resolution is 0.1 bar;

Cathode humidity: 0.1-1, resolution is 0.1;

Cathode gas stoichiometric: 1.2-4.0, resolution is 0.1;

Current density:1-4A/cm^2, resolution is 0.1A/cm^2.

Using the random number function in Matlab, the random number combination working conditions are uniformly generated according to the value of the resolution within the above range, so that the data in the test sample covers as many working conditions as possible.

In order to ensure the accuracy of the subsequent clustering algorithm model and balance the time and resources required for the PEMFC simulation experiment, the sample size was finally set to 500 groups. Among them, 10 groups of working conditions are set as forced gas starvation under the initial simulation conditions, that is, the amount of oxygen introduced is lower than the calculated amount of oxygen that should be introduced. Figure 2 shows the distribution of temperature and gas stoichiometric ratio in 500 combined operating conditions randomly generated. It can be seen that within the range of values, the combined operating conditions can cover most of the possible operating conditions.

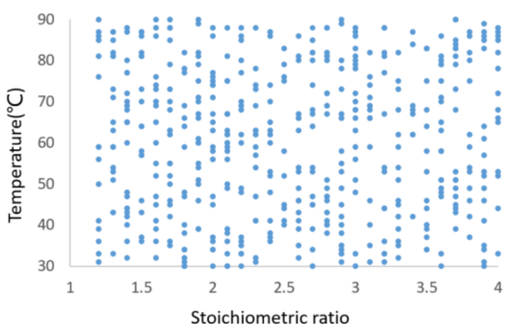


Fig 2 the distribution of temperature and stoichiometric ratio

3. METHOD DESCRIPTION

In this paper, the gas-starved area ratio was selected as a measure of the degree of gas starvation[10]. Then, the local gas starvation is effectively diagnosed based on the principal component analysis method and the K-means clustering algorithm. According to the clustering results, complete sample labeling for each sample point, compare three machine learning methods, and choose one method that can quickly predict the state of local gas starvation through operating parameters.

3.1 *Diagnosis of local gas starvation based on principal component analysis and K-means clustering algorithm*

The operating parameters of fuel cell will have an impact on the degree of gas starvation. In this paper, the data obtained by CFD simulation is used as the original data, and physical quantities that can directly reflect the operating state of PEMFC are selected as characteristic parameters. A diagnosis method combining principal component analysis method and clustering method is proposed to effectively diagnose and divide the local gas starvation conditions.

Selected 12 characteristic parameters include: temperature, pressure, air stoichiometric, air intake flow rate, air intake relative humidity, current density, output voltage, oxygen molar concentration at the oxygen outlet, oxygen molar concentration at the oxygen inlet, the molar concentration of water at the oxygen outlet, the molar concentration of water at the oxygen inlet, and the ratio of starvation.

To reduce invalid and repetitive calculations, it is necessary to reduce the dimensionality of the characteristic parameters, and select the characteristic parameters which best characterize the operating conditions of the PEMFC. In this paper, principal component analysis (PCA) is used to transform multiple indicators into a few comprehensive indicators, namely principal components. When the cumulative contribution rate of the principal components can reach 90%, it can be considered that this part can represent most of the information in the original sample and can be used to characterize the overall source data. Finally, this paper selected the five principal components with the largest contribution rate as the input value for model training.

The K-means clustering algorithm uses a certain distance between the sample data point and the prototype as the optimized objective function to classify the sample data. The initial cluster center and initial classification number k value are selected; then calculate the Euclidean squared distance of all source data samples, indicating the "degree of intimacy" between each sample. The Euclidean squared distance is:

$$d_{ij} = \sum_{k=1}^{P} \left(x_{ik} - x_{jk}\right)^2 \tag{1}$$

Where d_{ij} represents the distance between the ith and jth samples; x_{ik} represents the kth variable in the ith sample data point; and x_{jk} is similar as x_{ik}.

In the traditional K-means algorithm, the selection of the initial clustering center is often random, which will have a greater impact on the subsequent iterative calculation results, and the clustering results lack a certain interpretability for the diagnosis of local gas starvation. Therefore, this paper optimized the selection process of the initial clustering center, the sample point with forced gas starvation was selected as the cluster center. The proportions of gas-starved areas of the sample points after clustering were compared, and the results of K-means clustering were further analyzed.

3.2 Prediction of local gas starvation based on machine learning methods

The above-mentioned diagnosis method based on principal component analysis and K-means clustering algorithm can accurately define local gas starvation state, but cannot make rapid predictions in actual operating conditions. A suitable machine learning method can realize rapid prediction of local gas starvation state through several operating parameters.

Based on the clustering results of 500 sample points in Section 3.1, the local gas starvation attribute of each sample point is defined as the "label". The 500 sample points are randomly divided into training set and test set, and three machine learning methods are used to predict local gas starvation state.

The process steps of the machine learning model are shown in Figure 3:

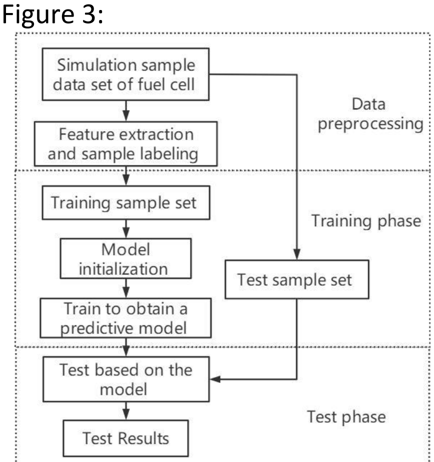


Fig 3 The process steps of the machine learning model

3.2.1 Prediction of local gas starvation based on support vector machine regression model

Using the support vector machine regression model to predict local gas starvation, it is necessary to find a separating hyperplane in a high-dimensional space. Based on the method of maximizing the interval, the hyperplane and the support vector that maximize the "distance" of the sample data point set are obtained. According to the obtained hyperplane and support vector, the regression model is used to classify the sample, and the prediction of the local gas starvation state is realized. For the training data set T:

$$T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\} \quad (2)$$

Where $x_i \in R^n$, $y_i \in \{+1, -1\}$, $i = 1,2,\ldots,N$, x_i represents the i^th^ feature vector, and y_i represents two different markers. In this paper, it represents local gas starvation and non-local gas starvation. The expression of the hyperplane is as follows:

$$\boldsymbol{\omega} \cdot \boldsymbol{x} + b = 0 \quad (3)$$

3.2.2 Prediction of local gas starvation based on extreme learning machine regression model

To predict the state of local gas starvation, the extreme learning machine model firstly initialize the number of neurons in the hidden layer, and randomly generate the values of ω and b. ω represents the connection weight between the input layer and the hidden layer, b represents the threshold of the hidden layer neuron; determine the activation function G of the hidden layer; calculate the output matrix H of the hidden layer; according to the output matrix H and the network output matrix T , The weight β of the output layer is calculated, and the trained extreme learning machine model is obtained.

Through the feature mapping of activation function G, the output value of the extreme learning machine can be obtained:

$$f_L(x) = \sum_{i=1}^{L} \beta_i G(\omega_i * x_i + b_i) \quad (4)$$

L is the total number of samples. The diagnosis and prediction of the local gas starvation state is realized through the output value $f_L(x)$.

3.2.3 Prediction of local gas starvation based on based on decision tree classifier

To predict the local gas starvation, this paper selected the CART decision tree with Gini coefficient as the characteristic criterion to establish the prediction model. For the binary classification problem, the expression of the Gini coefficient of the probability distribution is as follows:

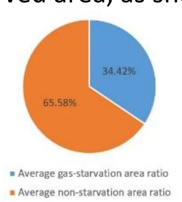
$$Gini(p) = 2p(1-p) \quad (5)$$

In the prediction process, the CART decision tree started from the root node, calculated the Gini coefficient of temperature, pressure, gas stoichiometric ratio, humidity and current density. Used the smallest feature attribute of Gini as the division feature of the leaf node, according to the value of the division feature, create a new leaf node, and continuously call the above division method to the new leaf node to establish a CART decision tree model. The CART decision tree classification algorithm may overfit the test data set, leading to the low accuracy prediction result. Therefore, the CART tree needed to be pruned and optimized according to its complexity, to balance the fit of the sample data set and the complexity of the model.

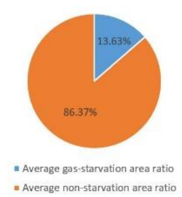
4. RESULTS AND DISCUSSION

4.1 Diagnosis results based on principal component analysis and K-means clustering algorithm

The data of 500 random operating points are obtained through simulation and processed by the principal component analysis method. Each sample is characterized by the principal components of 5 main classifications. According to the contribution of different principal components to distinguishing the distance between sample points, the corresponding weight coefficients are applied, and the most representative parameters of different sample points are obtained after superposition. From the 10 sample points of forced gas starvation, the sample point with the farthest representative parameter is selected as the initial clustering center of one of them for clustering. Divide all sample points into two types, calculate the ratio of the gas-starved area, as shown in Figure 4.



(a)



(b)

Fig 4 The average gas-starved area ratio diagram of the two types of working conditions: (a) The first type of working condition (b) The second type of working condition

Analyzing the clustering results, the average gas-starved area ratio in Type 1 reaches 34.42%, and the average gas-starved area ratio in Type 2 is only 13.63%. The 10 sample points that are forced to be gas starvation are all in Type 1. Based on the actual situation, it can be considered that type one characterizes the working condition of local gas starvation for PEMFC, and type two characterizes non-local gas starvation. Among them, type one contains a total of 228 sample points, type two contains 272 sample points. The average gas-starved area ratio of 500 sample points in original is 23.11%, and the average gas-starved area ratio of sample points in type 1 after classification by the K-means clustering algorithm is 34.42%. It can be considered that the algorithm can effectively diagnose the local gas starvation conditions of the PEMFC.

4.2 Prediction results based on machine learning methods

The clustering result is used as the "label" of the sample point, that is, it is divided into a local gas starvation sample point or a non-local gas starvation sample point. The 500 sample points are randomly divided into a training set and a test set, of which the training set accounts for 80%, thereby completing the collection and labeling of the original sample data set. Choose different machine learning methods to train the data set and get the results.

The input of machine learning model is five operating parameters (temperature, pressure, humidity, gas stoichiometric ratio and current density), the 100 sample points in the test set will get a predicted local gas starvation state, and the prediction accuracy of machine learning can be obtained by comparing the prediction results with the clustering results of each sample point.

The prediction results of fuel cell local gas starvation based on support vector machine algorithm, extreme learning machine algorithm and decision tree algorithm are shown in Table2-4, respectively.

Table 2 The prediction results of SVM

Label	Local gas starvation	Non-local gas starvation
Clustering result	46	54
Prediction result	43	57

Table 3 The prediction results of ELM

Label	Local gas starvation	Non-local gas starvation
Clustering result	46	54
Prediction result	45	55

Table 4 The prediction results of decision tree algorithm

Label	Local gas starvation	Non-local gas starvation
Clustering result	46	54
Prediction result	39	61

According to the clustering results, among the 100 sample points in the test set, there are 46 sample points with local gas starvation. By comparing the machine learning prediction results and clustering results of each sample point, it can be found that: the number of samples accurately predicted by the three machine learning methods are 41, 43, and 38, respectively. It can be seen that the prediction result based on the extreme learning machine model has the highest accuracy. Moreover, the algorithm does not need to adjust all parameters in iterations, so it has faster prediction speed.

5. CONCLUSION

This paper establishes a serpentine five-channel PEMFC simulation model, selects temperature, humidity, pressure, stoichiometric ratio and current density as independent variables, randomly generates 500 working condition sample points, and obtains source data from simulation.

Based on the principal component analysis method, using the idea of dimensionality reduction, the 12 working condition characteristic parameters are converted into 5 principal components. These 5 principal components can represent most of the information of the original data set, and the information contained in each other is not repeated; Based on the K-means clustering algorithm, the sample point with forced gas starvation is selected as the initial clustering center, and the sample data set is clustered into two categories. Through analysis, it can be proved that the clustering results can effectively define the local gas starvation conditions.

Based on the clustering results, each sample point was marked, 500 working condition sample points were randomly divided into training set and test set. Three machine learning regression models were constructed, namely model based on support vector machine regression, model based on extreme learning machine regression and model based on decision tree classifier to predict the local gas starvation state. Comparing the results of the three machine learning methods, it is concluded that the prediction accuracy rates of the three regression models are 89.13%, 93.49% and 82.61%, respectively. The extreme learning machine regression model has the highest prediction accuracy, and the prediction speed is fastest, which is applicable to quickly predict the local gas starvation state of PEMFCs.

ACKNOWLEDGEMENT

This research is sponsored by the National Key Research and Development Program (No.2017YFB01031 02), the National Natural Science Foundation of China (No.21805210), and the Prospective Study Funding of Nanchang Automotive Innovation Institute, Tongji University(No.TPDTC202010-17).

REFERENCE

[1] Bizon N, Manuel Lopez-Guede J, Kurt E , et al. Hydrogen economy of the fuel cell hybrid power system optimized by air flow control to mitigate the effect of the uncertainty about available renewable power and load dynamics[J]. Energy Conversion & Management, 2019, 179(JAN.): 152-165.

[2] Verhage A J L , Coolegem J F , Mulder M J J , et al. 30,000 h operation of a 70 kW stationary PEM fuel cell system using hydrogen from a chlorine factory[J]. International Journal of Hydrogen Energy, 2013, 38(11):4714-4724.

[3] Hao Liu, Jian Chen, et al. Short-Term Prognostics of PEM Fuel Cells: A Comparative and Improvement Study [J]. IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, 2019, 8(66):6077-6086.

[4] Chen H , Zhao X , Zhang T , et al. The reactant starvation of the proton exchange membrane fuel cells for vehicular applications: A review[J]. Energy Conversion and Management, 2019, 182:282-298.

[5] Li P, et al. Diagnostic method of gas starvation in dynamic process of proton exchange membrane fuel cell[J]. Chin High Technol Lett 2013;23(2):189–195.

[6] Hu Zunyan, Xu Liangfei, Li Jianqiu, Hu Junming, Xu Xin, Du Xiaoli, et al. A cell interaction phenomenon in a multi-cell stack under one cell suffering fuel starvation[J]. Energy Convers Manage 2018;174:465–474. ISSN 0196-8904.

[7] Liang D, et al. Behavior of a unit proton exchange membrane fuel cell in a stack under fuel starvation[J]. Power Sources 2011;196(13):5595–5598.

[8] Rahimi-Esbo M, et al. Design, manufacturing, assembling and testing of a transparent PEM fuel cell for investigation of water management and contact resistance at dead-end mode[J]. Hydrogen Energy 2017;42(16):11673–11688.

[9] Chen H, He Y, Zhang X, et al. A method to study the intake consistency of the dual-stack polymer electrolyte membrane fuel cell system under dynamic operating conditions[J]. Applied Energy, 2018, 231:1050-1058.