

# Data-based flooding fault diagnosis of proton exchange membrane fuel cell systems using LSTM networks

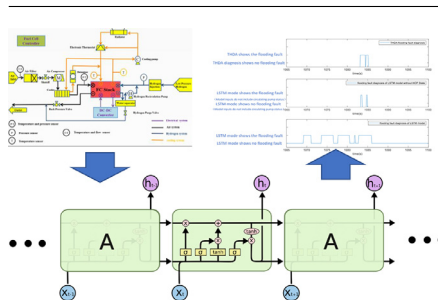
Xin Gu, Zhongjun Hou\*, Jun Cai

Shanghai Hydrogen Propulsion Technology Co., Ltd., Unit 10, BLDG 17, Innovation Park, Lane 56, Antuo Rd. Jiading, Shanghai PR China

## HIGHLIGHTS

- LSTM model for flood faulting diagnosis is proposed.
- Using auxiliary system information avoids the installation of a large number of sensors.
- The proposed method has better performance in diagnose/pre-diagnose.

## GRAPHICAL ABSTRACT



## ARTICLE INFO

### Article history:

Received 20 December 2020  
Received in revised form 8 February 2021  
Accepted 8 February 2021  
Available online 24 February 2021

### Keywords:

Proton exchange membrane fuel cell  
Flooding diagnosis  
Data-driven diagnosis  
Long Short-Term Memory network

## ABSTRACT

Flooding fault diagnosis is critical to the stable and efficient operation of fuel cells, while the on-board embedded controller has limited computing power and sensors, making it difficult to incorporate the complex gas-liquid two-phase flow models. Then in fuel cell system for cars, the neural network modeling is usually regarded as an appropriate tool for the on-line diagnosis of water status. Traditional neural network classifiers are not good at processing time series data, so in this paper, Long Short-Term Memory (LSTM) network model is developed and applied to the flooding fault diagnosis based on the embedded platform. Moreover, the fuel cell auxiliary system statuses are adopted as the inputs of the diagnosis network, which avoids installing a large number of sensors in the fuel cell system, and contributes to reduce the total system cost. The bench test on the 92 kW vehicle fuel cell system proved that this model can effectively diagnose/pre-diagnose the fuel cell flooding, and thus help optimize the water management under vehicle conditions.

## 1. Introduction

Proton Exchange Membrane fuel cell (PEMFC) technology has the advantages of fast start-up, high efficiency, low operating temperature, high power density, stable operation, and no environmental pollution. As environment and global warming issues become increasing severe, PEMFCs are regarded as one of the clean solutions for automotive powertrains [1,2].

For vehicular applications, performance, durability and reliability of PEMFC systems are the most challenging issues, which largely depends on the water and thermal management [3]. With the growing requirement of high power density, fuel cell stacks need to operate at higher current density, which means more liquid water is produced and is prone

to block the porous electrode and flow channels, causing the reactant gas starvation, and reduce the performance and durability [4-5]. Therefore, good water management is required to ensure the stable operation of fuel cell stack, by balancing the membrane humidifying and oxygen transport, and then the diagnosis of water flooding state is critical to realize the timely and delicate control of the fuel cell system.

Considerable efforts have been devoted to the water state estimation, and the main approaches usually include modeling and experiment, such as the analytical model [6], fuzzy logic [7-8] and neural network model [9], polarization curve analysis, electrochemical impedance spectroscopy (EIS) analysis, wavelet and fourier analysis, and etc.

Pressure drop model (most of the pressure drop models are also an analytical model) is widely used in water state analysis, as the pressure

\* Corresponding author.

E-mail addresses: [hong\\_zhongjun@shpt.com](mailto:hong_zhongjun@shpt.com), [zhjhou@hotmail.com](mailto:zhjhou@hotmail.com) (Z. Hou).

drop can reflect the internal water state of fuel cell stack in different conditions [10]. Liu et al. analyzed the relationship between stack water state and cathode/anode pressure drop and thought that cathode pressure drop can better reflect stack water state [11]. Banerjee et al. analyzed the temperature effect with the developed pressure-water model under both single and two phase flow [12–13]. Anderson R et al. investigated the relationship between current density effect using the pressure-water model [14]. Pei et al. suggested that pressure drop can reduce if increasing the inlet gas flow density [15]. However, due to the large amount of liquid water at the anode and cathode outlets, the noise of outlet pressure detection is significant, and the traditional pressure difference detection is difficult to accurately determine the water state of fuel cell stack under vehicle-mounted conditions. Moreover, traditional methods require massive inputs, which is uneconomic. Based on the discrete wavelet transform approach, Jonghoon et al. analysed the system output terminal voltage signal (OTVS) and then evaluate the healthiness of system [16], while this approach requires high-precision sensors and will increase the vehicle system cost.

Neural network model is another widely adopted approach [19]. Li et al. developed a pressure drop model and labelled the data. The model uses the cell voltages as the main input, and the SVM (Support Vector Machine, SVM) classifier is designed to diagnose the fuel cell water status and obtain high diagnostic accuracy [17–18]. However, the classification result of the SVM classifier is only related to the instantaneous state and cannot handle the time series, which leads to limited predictive diagnosis capabilities. In addition, in previous studies, the fuel cell stack water status diagnosis relied on the inputs from stack inlet and outlet parameters of hydrogen, air, and thermal parameters (such as inlet flow rate, outlet dew point temperature, etc.). Some parameters are difficult to monitor under vehicle conditions, for example, when fuel cell stack operates with anode hydrogen recycling, a large amount of liquid water circulates back from the outlet to the hydrogen inlet, while the flow sensor cannot accurately detect the hydrogen flow, or the sensor is expensive for vehicle system.

In recent years, as deep learning becoming overly popular, neural networks such as CNN (Convolutional Neural Network, CNN), DNN (feedforward neural network, DNN) and RNN (Recurrent Neural Network, RNN) are experiencing a fast development. The neuron output of RNN can be used as the input of next time step's neuron, overcoming the shortcoming of other networks that cannot model the changes in the time series [20]. So RNNs are widely used in time series processing, such as language processing, voice recognition and gesture recognition [21–22]. The Long Short-Term Memory (LSTM) network, as a special kind of RNN, is currently a neural network that is very suitable and has been widely used in the prediction of time series. Moreover, for on-board fuel cell systems, the status and signal of fuel cell auxiliary system in a period of time can reflect the stack flooding state. For example, when the opening frequency of the hydrogen purge valve is low, a period of operation in high power or the pressure disorder between hydrogen and air can easily cause the stack flooding.

Combining the above reasons, this paper uses the LSTM network to model the fuel cell water state, and compares it with the traditional "memoryless" neural network. The model is based on the water transport mechanism inside the stack, and the fuel cell system accessory status is adopted as the model inputs, which effectively reduces the use of sensors. After finishing the deployment of model on the embedded platform, experiments on a 92kw fuel cell system are conducted to validate the model.

## 2. Basic ideas of the model

### 2.1. The water transport inside fuel cell and selection of the input vectors for LSTM model

The water transport mechanism inside the fuel cell is shown in Fig. 1, which mainly includes thermal-osmotic drag (TOD), electro-osmotic

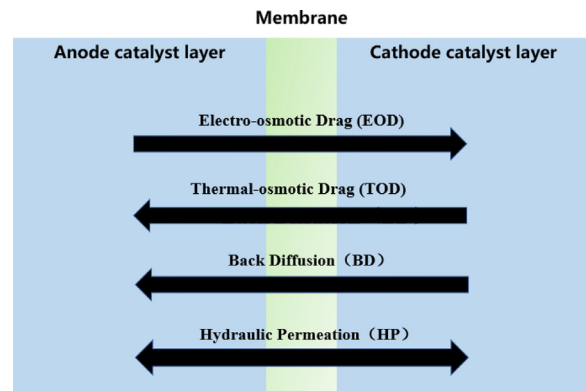


Fig. 1. Water transmission inside the fuel cell stack.

drag (EOD) with proton transport, back diffusion (BD) driven by water content difference and hydraulic permeation (HP) driven by pressure difference [23]. EOD and BD are dominated for water transport in PEMFCs, which significantly affect the humidity of membrane and thus the proton resistance [24].

During the operation of PEMFC, EOD occurs as the proton transports from the anode to cathode with combined water molecules [25], and the water transport amount induced by EOD is proportional to the current density. In high current density condition, EOD may cause the anode drying and cathode flooding, so the developed diagnostic model proposed in this study takes the stack current density as an important model input [26]. Zhiping et al. found that the EOD coefficient is closely related to temperature, when the temperature rises, the EOD coefficient increases significantly. Therefore, the stack coolant inlet and outlet temperatures as also adopted as the model inputs, and the cooling pump speed is selected as the input to reflect the state of coolant flow [27]. The excess water produced at the cathode may transport to the anode due to concentration gradients and pressure difference, this is the called BD and HP [28]. The water transport amount resulting from BD and HP depends on the water content gradient, membrane thickness, pressure difference and temperature. Therefore, the hydrogen and air inlet pressure are adopted as the model inputs [29–30]. Due to the Thermal-osmotic Drag phenomenon, water may flow from the colder area to hotter area [31–32], while it has little effect on the water state of the fuel cell stack and can be ignored.

Moreover, the state of the hydrogen purge valve, air flow rate, cell voltage, cathode inlet temperature are also incorporated as the model inputs. Visual experiments in previous studies have confirmed that the accumulation of liquid water in the flow channel experiences four stages: stray droplets, stable droplets, liquid film and slug [33]. Before the fuel cell is flooded, the liquid water in the flow channel continues to increase, and then causing the extra power consumption of the fuel cell system accessories, especially for the hydrogen recirculation pump, so the DC current and the speed of hydrogen recirculation pump are adopted as model inputs. As described above, with the increase of liquid water, channels will experience single-phase flow, droplet flow, film flow and slug flow, successively. In the flow pattern map, the superficial gas and liquid water velocities are related to the state of the two-phase flow [34], for example, at a higher superficial liquid velocity and a lower superficial gas velocity, slug flow more likely occurs. In vehicle applications, the anode superficial gas velocity relates with the speed of hydrogen recirculation pump, and the cathode superficial gas velocity relates to the air inlet flow rate. From the observations in the experiment, the liquid removal speed at the cathode and anode outlets is related to the inlet pressure and inlet flow, so we believe that qualitatively, the superficial liquid speed is related to the speed of hydrogen recirculation pump, anode inlet pressure, air flow rate, cathode inlet pressure.

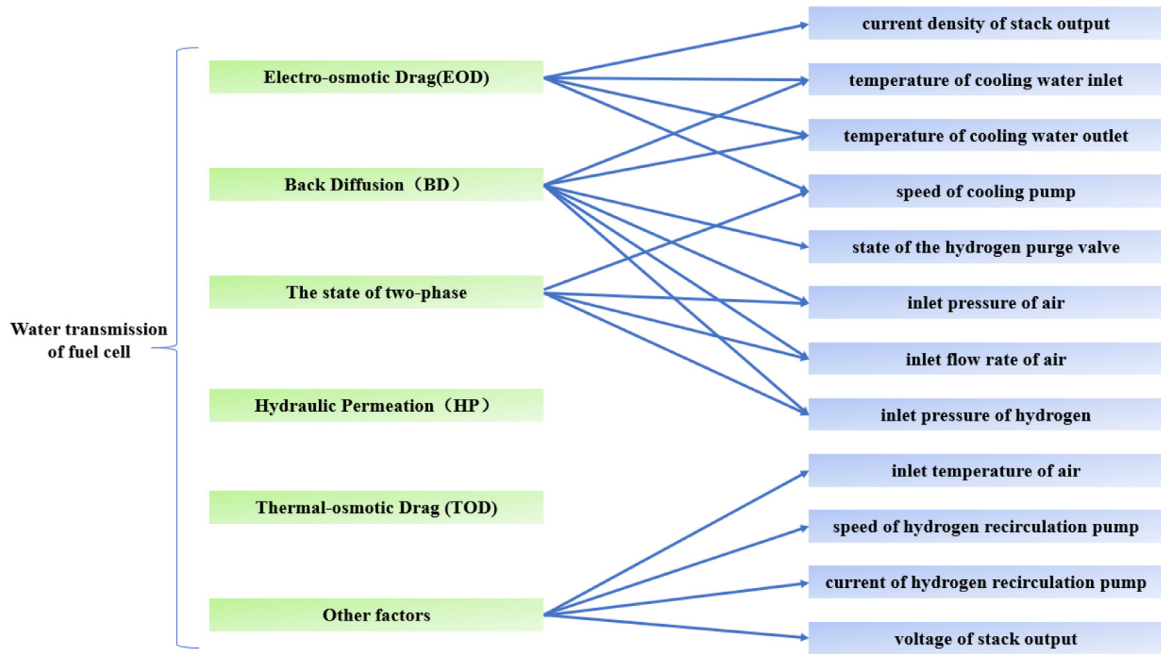


Fig. 2. Input vector of LSTM model.

In summary, this study developed a data-driven model, and takes the outputs of fuel cell system accessories as model inputs based on their relationship with the water status of fuel cell stack. All the model inputs can be measured with a few cheap sensors, which avoids installing more sensors and saves the total system cost that is vital to vehicle applications. The inputs of the LSTM model are shown in Fig. 2

## 2.2. Fuel cell power system in vehicles

In order to meet the on-board application of fuel cells, fuel cell stacks must be operated within the appropriate control of the fuel supply, power output, water and heat management, fault diagnosis and protection of the fuel cell, which is realized by fuel cell auxiliary systems.

As shown in Fig. 3, the fuel cell auxiliary system generally consists of the hydrogen, air, cooling, electrical and control subsystems. The electrical subsystem adjusts the stack output power through the DC-DC converter, and the hydrogen subsystem adjusts the hydrogen inlet pressure and flow rate through both the injector and hydrogen recirculation pump. The water separator separates the liquid water at the anode outlet, and avoids that excessive liquid water is recycled to the anode inlet. The air subsystem provides the required air inlet pressure and flow rate by adjusting the air compressor speed and opening ratio of the back pressure valve. The air filter prevents large particles in the air flow from entering the air compressor, which improves the system life. The cooling subsystem adjusts the speed of cooling pump and the opening ratio of electronic thermostat to control the system heat dissipation, and ensure that the fuel cell stack operates below 95 °C. The control subsystem coordinates the operation of all electronic control components with considering the economy, power and durability of the on-board fuel cell system.

The developed model in this study runs in the system controller of the control subsystem, and the information from above auxiliary systems is used as inputs for flooding fault diagnosis, which can effectively reduce the installation of sensors, as well as the system complexity and cost.

## 2.3. Calculation and training of LSTM network

Ordinary recurrent neural networks will experience gradient disappearance or gradient explosion during training, which makes it difficult

to deal with long-distance dependence in practical applications. In response to the above problems, the LSTM eliminates the above defects, and thus becoming the most popular RNN due to its advantage in processing the time series data.

As shown in Fig. 4, the basic structure of LSTM is the block. Each block is connected in a chain and receives the input vectors at different times. The key to LSTM is the block state vector, which is defined as C in this study. Vector C evolves along the entire chain, making the "memory" of the network easy to flow along it. The dimension of the vector C is called the scale of the model hidden layer neuron, and the larger this dimension, the better model expression ability, while the calculation amount and training data will increase accordingly.

Each block uses the forget gate and input gate to control the block state vector C. The calculation method of forget gate is described as Eq. (1), where  $W_f$  is the weight matrix of forget gate,  $X_t$  is the input vector at the current time step,  $h_{t-1}$  is the output vector at the previous time step, and  $b_f$  is the bias item. The forget gate determines how much of the block state  $C_{t-1}$  from the previous time step is retained to the current time step, and the sigmoid function is usually used as the activation function.

$$f_t = f(W_f * \begin{bmatrix} h_{t-1} \\ X_t \end{bmatrix} + b_f) \quad (1)$$

The calculation of input gate is given in Eq. (2), where  $W_i$  is the input gate weight matrix, which determines how much of the input  $X_t$  of the network at the current time step is saved in the block state  $C_t$ .

$$i_t = f(W_i * \begin{bmatrix} h_{t-1} \\ X_t \end{bmatrix} + b_i) \quad (2)$$

The current state vector  $C_t$  is calculated as Eq. (3) through the state at the previous time step and the input at the current time step.

$$\begin{aligned} \tilde{C}_t &= f(W_c * \begin{bmatrix} h_{t-1} \\ X_t \end{bmatrix} + b_c) \\ C_t &= f_t \circ C_{t-1} + i_t \circ \tilde{C}_t \end{aligned} \quad (3)$$

The output at the current time step is calculated through the output gate.

$$\begin{aligned} O_t &= f(W_o * \begin{bmatrix} h_{t-1} \\ X_t \end{bmatrix} + b_o) \\ h_t &= O_t \circ \tanh(O_t) \end{aligned} \quad (4)$$

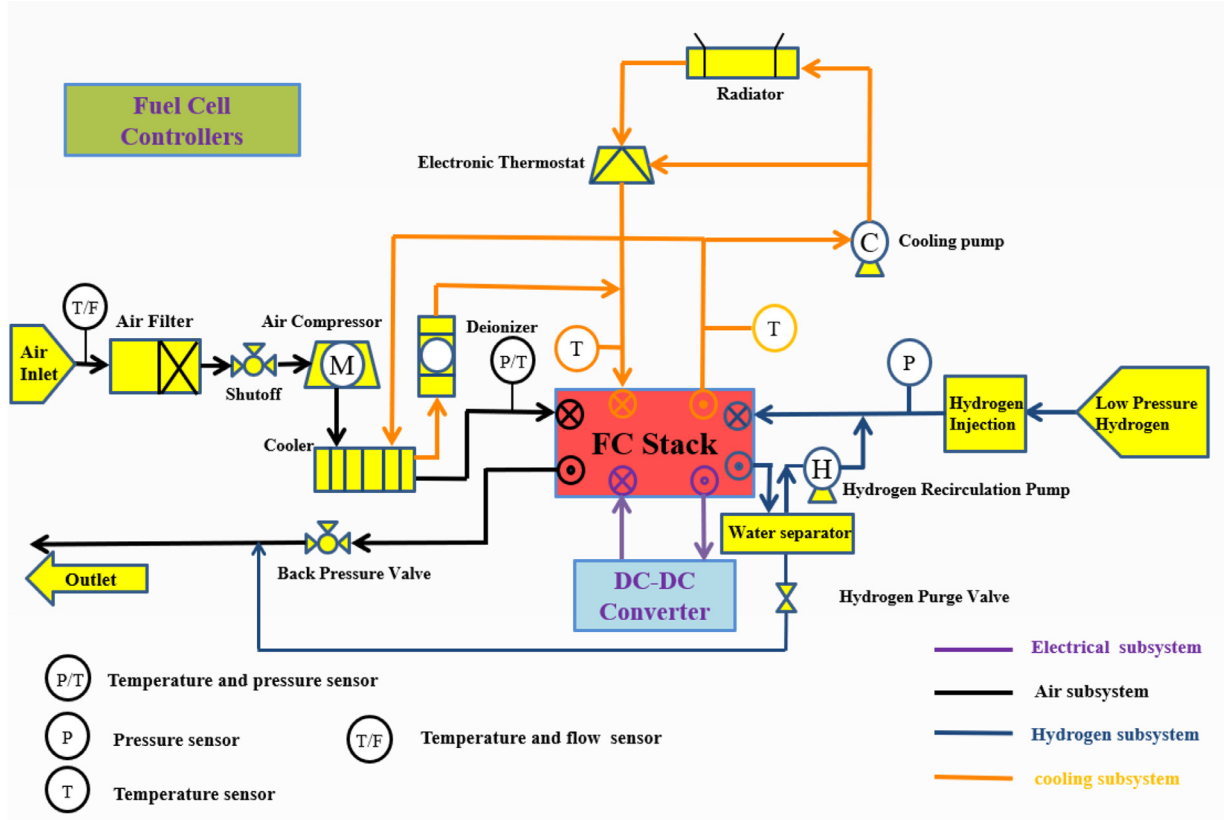


Fig. 3. Schematic of the used PEMFC system.

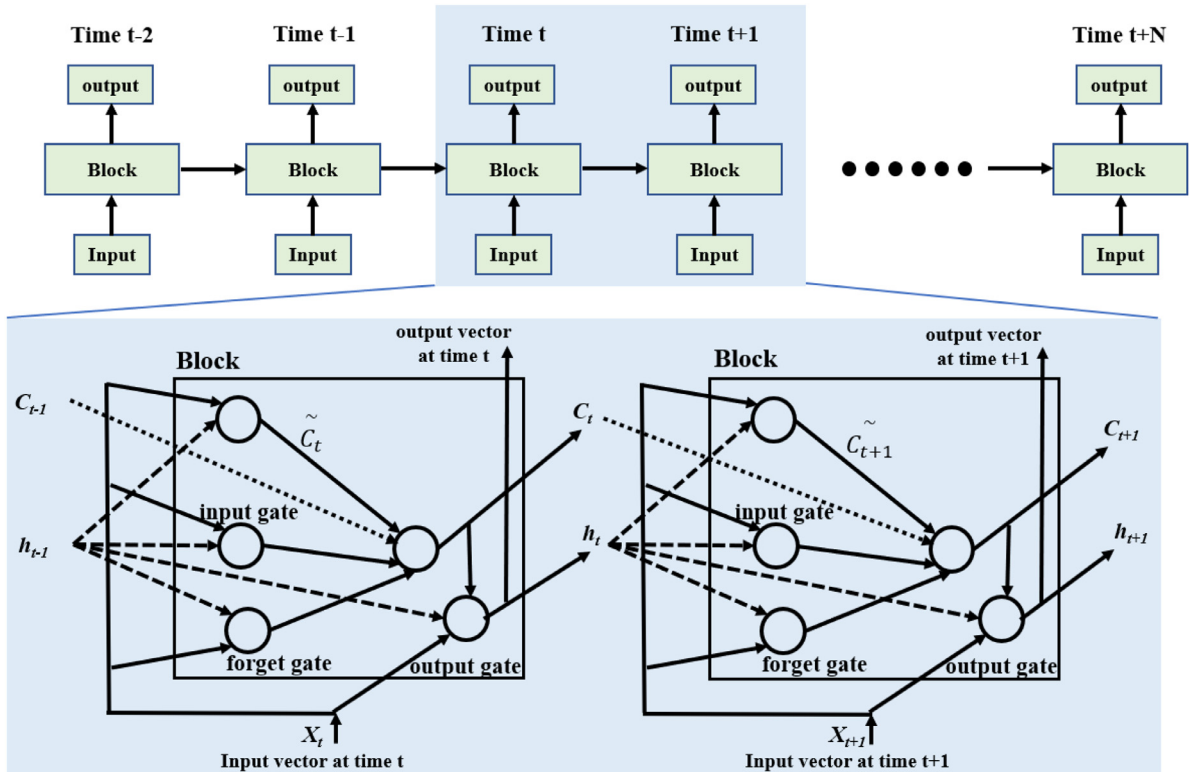


Fig. 4. LSTM network architecture.

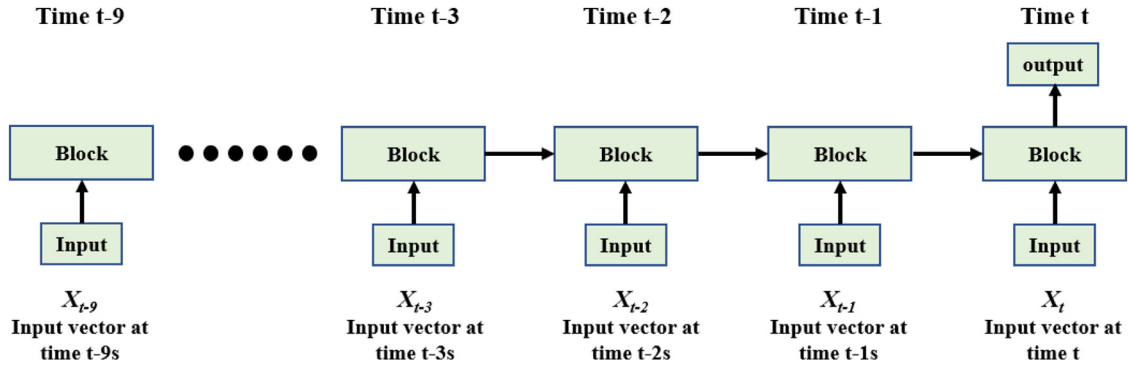


Fig. 5. Water fault diagnosis model based on LSTM.

**Table 1**  
Input vectors of LSTM model.

	Components of the model input vector X	unit	Signal acquisition method under vehicle operation
1	inlet pressure of hydrogen	Kpa	AD sampling of controller
2	inlet pressure of air	Kpa	AD sampling of controller
3	inlet flow rate of air	g/s	AD sampling of controller
4	t emperature of cooling water inlet	°C	AD sampling of controller
5	t emperature of cooling water outlet	°C	AD sampling of controller
6	speed of cooling pump	r/min	automotive CAN bus
7	state of the hydrogen purge valve	no unit	digital sampling of controller
8	current density of stack output	A	automotive CAN bus
9	voltage of stack output	V	automotive CAN bus
10	speed of hydrogen recirculation pump	r/min	automotive CAN bus
11	current of hydrogen recirculation pump	A	automotive CAN bus



Fig. 6. THDA equipment.

The training of LSTM uses the back propagation algorithm, which includes three main steps [35]:

- Calculate the values of the five vectors  $f_t, i_t, C_t, O_t$  and  $h_t$ ;
- Do the reverse calculation for the error value of each neuron. Like the recurrent neural network, the back propagation of the error term also includes two directions: one is the back propagation evolved as time, and the another is the propagation to the upper layer;
- Calculate the gradient of each weight according to the corresponding error term.

### 3. Methodology

#### 3.1. Model establishment

In view of the high cost of obtaining the flooding fault data from fuel cell stack and avoiding the underfitting due to insufficient fault data, the model simply classifies the stack water status into two categories, namely, normal status and flooding status.

As shown in Fig. 5, the water state in fuel cell stack is a long-term accumulation effect, so the input vectors of the model should be the time

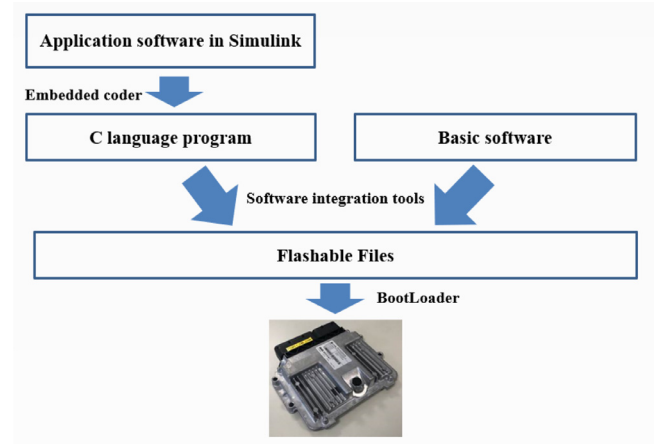


Fig. 7. Fuel cell system test bench architecture.

sequence, and the length is 9 s in this study. The definition of the input vectors is introduced in Table 1, where the state of the hydrogen purge valve is unitless (1 means the state is on, 0 means the state is off). The reasons for choosing these input vectors are explained in Section 2.2. When the model starts running, the current time step is pushed back by nine seconds, and the vector sampling is used as the inputs of the model at every time step. As the controller runs, rolling estimates are performed.

The structure of each block in the model is shown in Section 2.2. The activation functions in each block all use the sigmoid function, and the dimension of hidden layer neurons is 15.  $h_t$  is the output of the model, and the softmax classifier is used to classify the outputs of the LSTM model to give the final results. The output vector of the model is a one-dimensional vector, and the value of 0 means no flooding fault, and the value of 1 means there is the flooding fault.





Fig. 8. P390 full-power fuel cell system of SHPT.

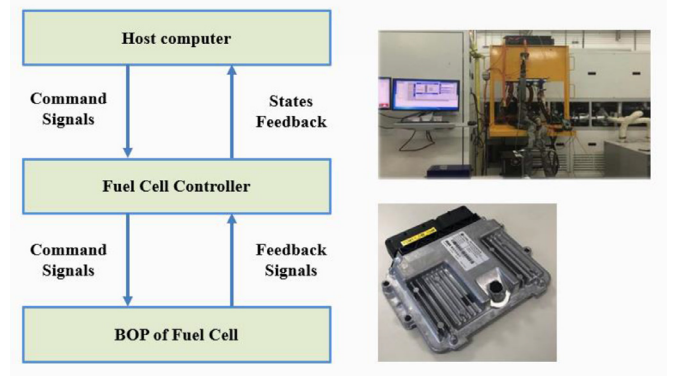


Fig. 9. Fuel cell system test bench architecture.

### 3.2. Data labeling and model training

The training set is obtained from the actual experimental data of fuel cell stack system. As shown in Fig. 6, data labeling is conducted by AVL's THDA (Harmonic Distortion Factor Analysis, THDA) detection equipment. During the operation of fuel cell stack, the THDA device superimposes a sinusoidal AC current component, which has an amplitude of 2A in the stack output through the parallel circuit, and uses 24bitA/D

to collect voltage/current signals. Based on the fundamental wave and harmonic analysis of current and voltage signals, it is possible to diagnose in real time whether the stack is flooded. The diagnosis result of the THDA device is used as the criterion for judging whether the model is correct in the following experiments.

In order to eliminate the underfitting phenomenon, which is caused by the imbalance between the fault data and the normal operation data, this study uses the sampling of fault data segment at different starting moments to obtain multiple fault data sets. At the same time, the data

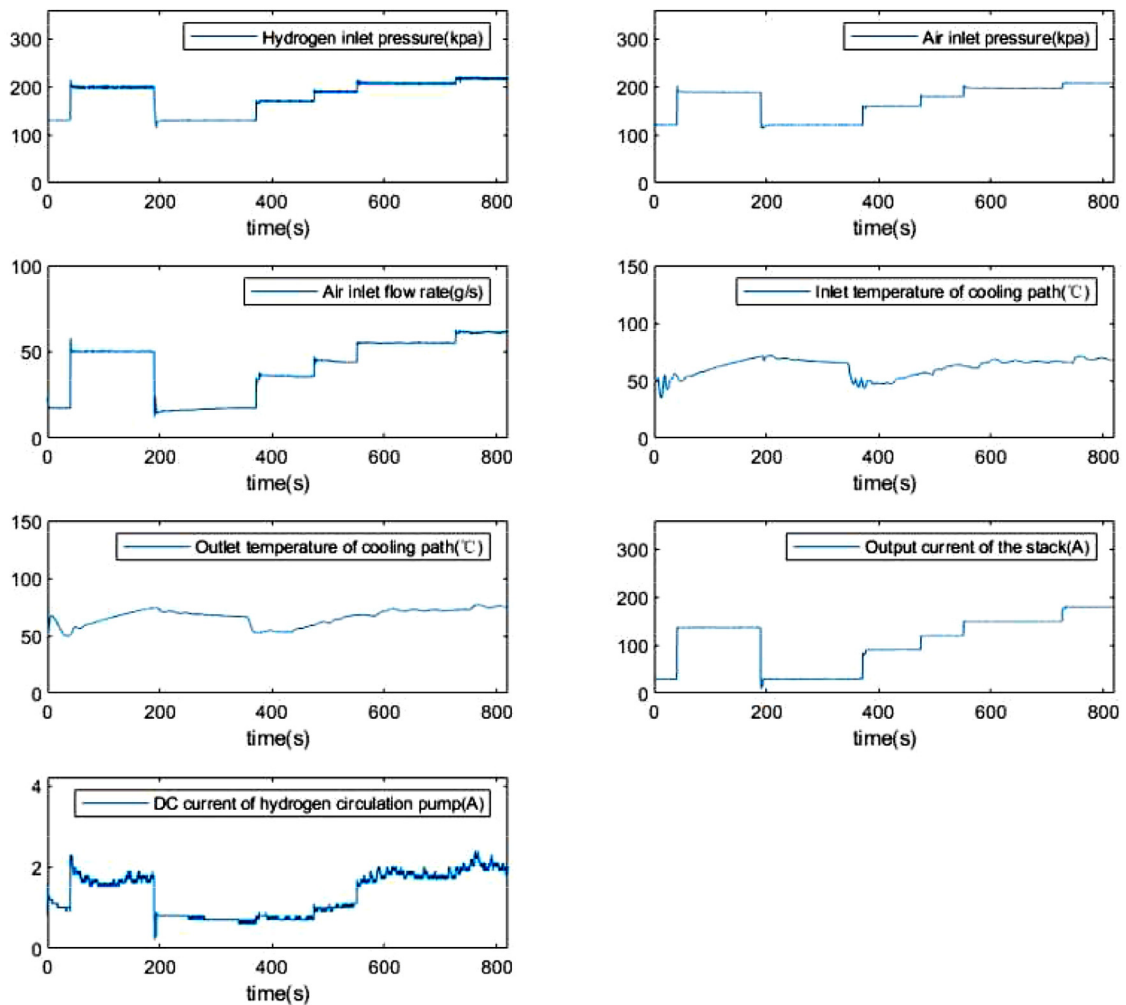


Fig. 10. System parameters under low current density operation.

under the failure label is tripled to balance the failure and non-fault data. The above method makes the data under different labels reach a balance in quantity, avoiding the occurrence of over-fitting and under-fitting.

The training method of the model is shown in Section 2.3, and the training is implemented using MATLAB's deep learning toolbox. The relevant parameters are set as follows:

- 1 The solver is: 'adam';
- 2 The gradient threshold is 1, and the maximum epochs are 800;
- 3 The sequence length is designated as 'longest';
- 4 The computing hardware is 'GPU';

### 3.3. Embedded realization of control software

As shown in Fig. 7, the algorithm proposed in this study is programmed in the application software layer. The application software is implemented in Simulink tool and converted into C code through the embedded coder. The generated C code and the basic software are compiled to generate a refreshable file. The refreshable file is flashed into the fuel cell system controller through BootLoader, and the controller is used on the system to conduct trial control.

In order to ensure the online operation of the model on the vehicle embedded platform, the calculation is simplified in the establishment of the model and the embedded realization. The dimensions of hidden layer neurons and the numbers of blocks are reduced during the embedded implementation. In addition, the execution cycle of the program at the application software is 1 s to reduce the computational burden of the controller.

## 4. Experiments and database

### 4.1. PEMFC system and test bench

As shown in Fig. 8, the fuel cell system used in the experiment is a P390 full-power fuel cell system from SHPT (Shanghai Hydrogen Propul-

sion Technology Co. Ltd). The maximum system power is 92 kW and the maximum system efficiency is 60%. This fuel cell system could rapidly start up from  $-30\text{ }^{\circ}\text{C}$  in 30 s without external heating.

The experimental test setup is given in Fig. 9. The host computer is responsible for the experimental conditions settings and help calculate the power output instructions of the fuel cell. The power output instruction is sent to the FCU (Fuel Cell Control Unit) through the CAN bus. The FCU is the control core of the fuel cell system based on the built-in control programs and diagnostic algorithms. FCU controls the auxiliary parts of the system through drive interface, communication subsystem and sampling interfaces. Data were collected from auxiliary parts for system diagnosis and feedback to the host computer.

### 4.2. Experimental database

The control software of fuel cell system utilized in the experiment tests is modified from the commercial control software of P390 fuel cell system. The control software of P390 fuel cell system receives the power instructions and control the feed gasses, heat dissipation and fuel cell output parameters by the actuators in a closed loop. The water flooding diagnosis model based on the LSTM network proposed in this study was compiled into the software used in the experimental tests. the diagnosis work based on the conventional SVM model and the LSTM network model without circulating pump information are also implemented as control group.

In practical applications, both water flooding and membrane dehydration issues may lead to same result, such the decrement of fuel cell output voltage. The main reason for the performance degradation is difficult to identify based on the same phenomenon with simple methods. Therefore, larger air flow rate and lower power output are given for the fuel cell stack to dry the membrane, in order to test whether the diagnosis will be misdiagnosed.

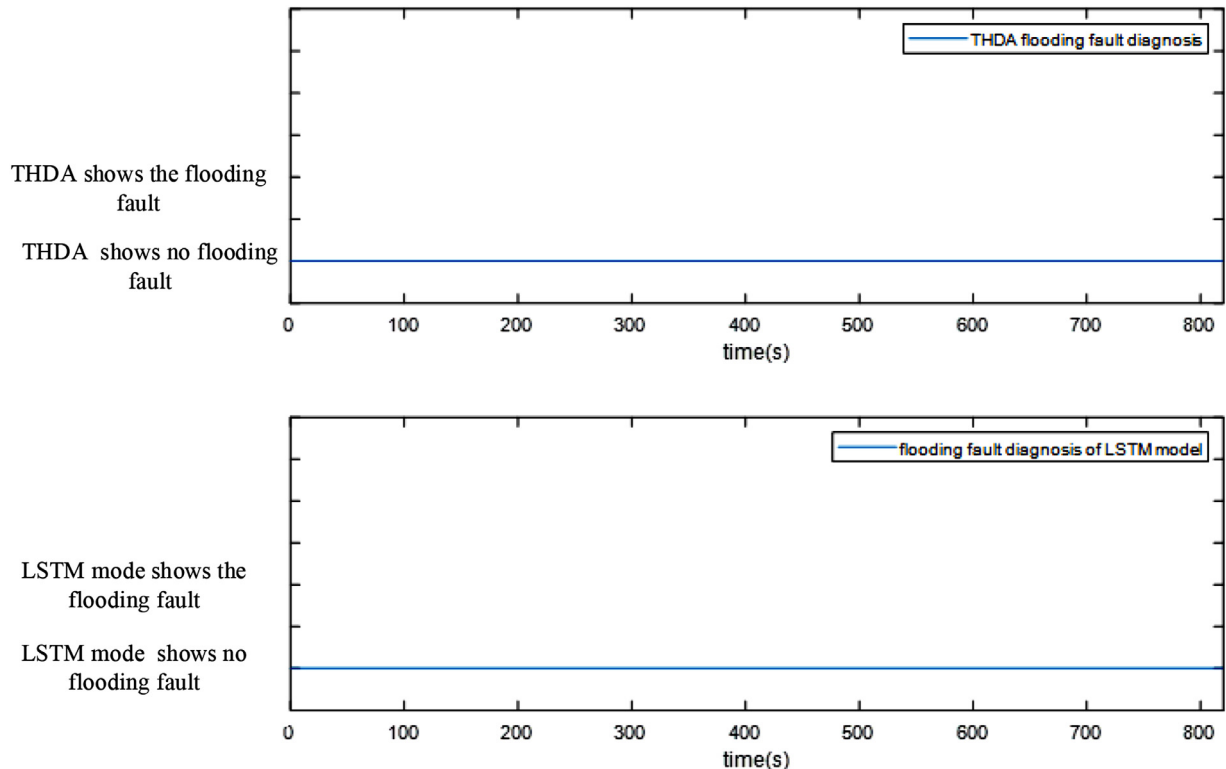


Fig. 11. Under low current density operation, THDA device diagnosis results and LSTM diagnosis network output results.

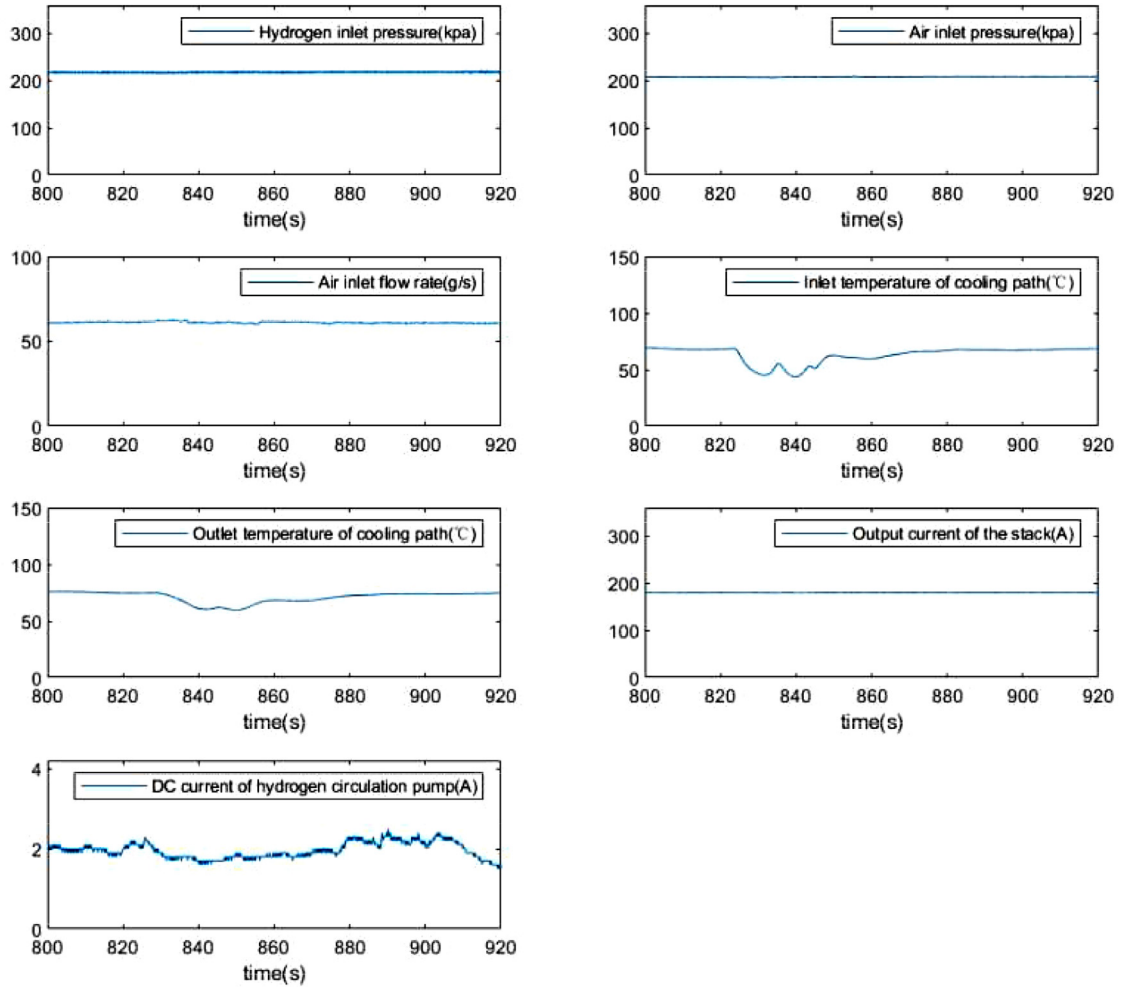


Fig. 12. Operating data in fault1.

#### 4.3. Experiments and experimental results

While the experiment is being implemented, AVL THDA diagnostic equipment is used to help diagnose the system. The diagnosis result of the THDA device is regarded as the standard result in order to verify the diagnosis accuracy of the LSTM model.

Moreover, if the data without hydrogen recirculation pump is used as the inputs of LSTM, the diagnosis accuracy of the network could be estimated. Taking SVM for an example, conventional memoryless network and LSTM network is analyzed and compared in the following sections.

### 5. Results

#### 5.1. Diagnostic results during trouble-free operation

As shown in Fig. 10, the system starts to operate at a low current density. The hydrogen pressure, air pressure and air flow rate are controlled by the FCU to increase with current. With the operation of fuel cell stack, the coolant temperature continues to rise, thus the stack continuing to heat up. As displayed in Fig. 11, the stack produces less water, and the THDA diagnostic equipment shows that no flooding fault accurs, in which 0 means no flooding fault has occurred and 1 means flooding fault has occurred). Meanwhile, the output of the LSTM diagnostic model proposed in this paper is 0, which indicates that no flooding

fault has occurred. The output result of the diagnostic model shows a good agreement with the result of the THDA device.

#### 5.2. Comparative analysis of LSTM network and SVM network

To compare and analyze the diagnostic accuracy of LSTM network and conventional feedforward network, the same training data was used to construct the SVM network with gaussian kernel function. For comparison, simulation work based on the LSTM network are also carried out.

Figs. 12 and 13 shows the data from 800 to 920 s. In the 827 s of system operation, the output of the LSTM diagnostic model changed from 0 to 1, indicating that the system was about to be flooded. During the 832 s of system operation, the THDA diagnostic equipment reported a flooding fault, and the output of the SVM network diagnostic model remains 1, representing the system flooding fault. Experimental results show that the LSTM model can classify the state for a period of time. Therefore, it works on the diagnosis of flooding faults. However, the conventional feedforward network, such as SVM network, has no ability to process time series. It only identifies the fault state, but could not predict the possible flooding issue to be happened.

In addition, in terms of system parameters, there is a significant increase in the power consumption of the hydrogen recirculation pump at 820 s, while the speed of the hydrogen recirculation pump does not change. Therefore, the change in the state of the hydrogen recirculation pump is an important feature of the flooding fault.



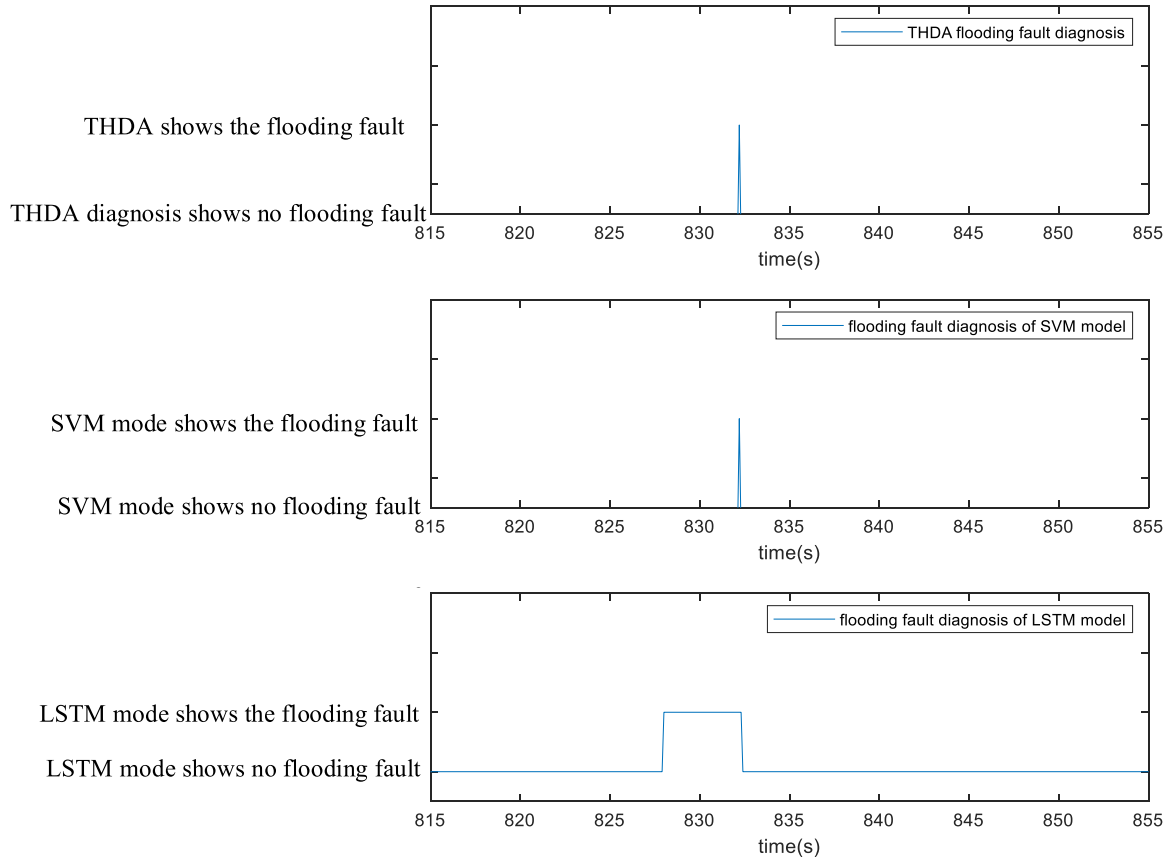


Fig. 13. Comparison of THDA equipment diagnosis results and LSTM diagnosis results and SVM network diagnosis results.

### 5.3. The importance of circulating pump information for flooding diagnosis

In order to estimate the identification improvement of the diagnosis network by the state of the fuel cell auxiliary system, the LSTM network trained based on the data without the hydrogen recirculation pump, and the original LSTM are compared to evaluate the diagnosis accuracy.

Fig. 14 and Fig. 15 gives the data from 1000 to 1200 s. At 1069 s, the output of the LSTM diagnostic model proposed in this paper continuously fluctuate between 0 and 1, indicating that the system is about to be flooded. From 1083 to 1085.2 s, the low high frequency resistance showed by THDA diagnostic equipment manifests that the system has been already flooded. About 0.5 s after the fault occurred, the output of the LSTM network without the hydrogen recirculation pump status changes from 0 to 1.

The result shows that the LSTM diagnosis model proposed in this paper could pre-diagnose the flooding issue. The LSTM network diagnosis that removes the hydrogen recirculation pump information shows a diagnosis lag, which emphasizes the important role of accessory information, especially the hydrogen recirculation pump data.

### 5.4. Membrane drying failure

When the stack membrane is dry, the high membrane ohmic resistance will lead to performance degradation. On the other hand, the water flooding will also result in serious mass transfer issues, thus lower fuel cell output voltage. Membrane drying and flooding are difficult to accurately identify in the diagnosis algorithm. In order to verify the ef-

fectiveness of the flooding diagnosis algorithm proposed in this paper, a large gas flow is used in the experiment test to make the membrane dry and to check out the accuracy of the model.

As shown in Fig. 16, after 8 s of operation, the input power of the stack is 3.5kw, and the air flow rate is increased to 25 g/s. At the same time, the hydrogen inlet pressure, air inlet pressure and cooling water temperature are adjusted to maintain the low power output of the stack. The high-frequency impedance rise sharply due to the low power and large air flow rate.

As shown in Fig. 17, since there is no flooding fault, the output of the modeling work remains zero. This indicate that this model has enough precision in the estimation of water flooding issues, which is considered difficult for the conventional vehicle sensors.

### 5.5. Analysis of results

There is a complete fault protection mechanism designed for the P390 fuel cell system. Once the flooding fault happens, the system will take protective measures to increase the anode drainage and alleviate the flooding fault. However, it will also sacrifice the hydrogen utilization rate. Therefore, the occurrence time of the flooding fault should be limited within a few seconds.

Experimental results show that the LSTM diagnostic network could effectively pre-diagnose flooding faults, help the control system to respond in time and avoid the occurrence of faults. Based on auxiliary system information, especially the circulating pump status the efficiency of the flooding fault diagnosis could be improved, which also helps reduce the sensor amount.

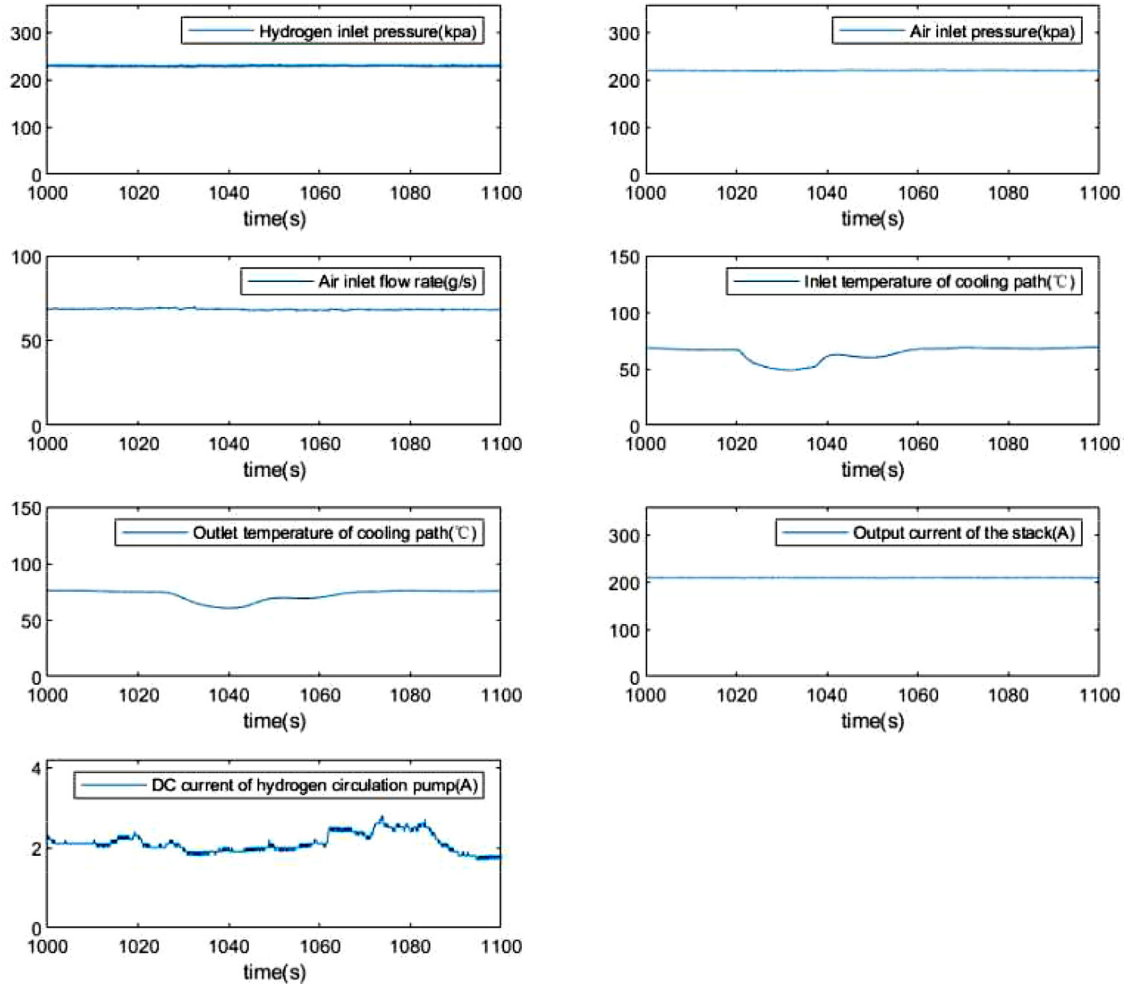


Fig. 14. Operating data in fault2.

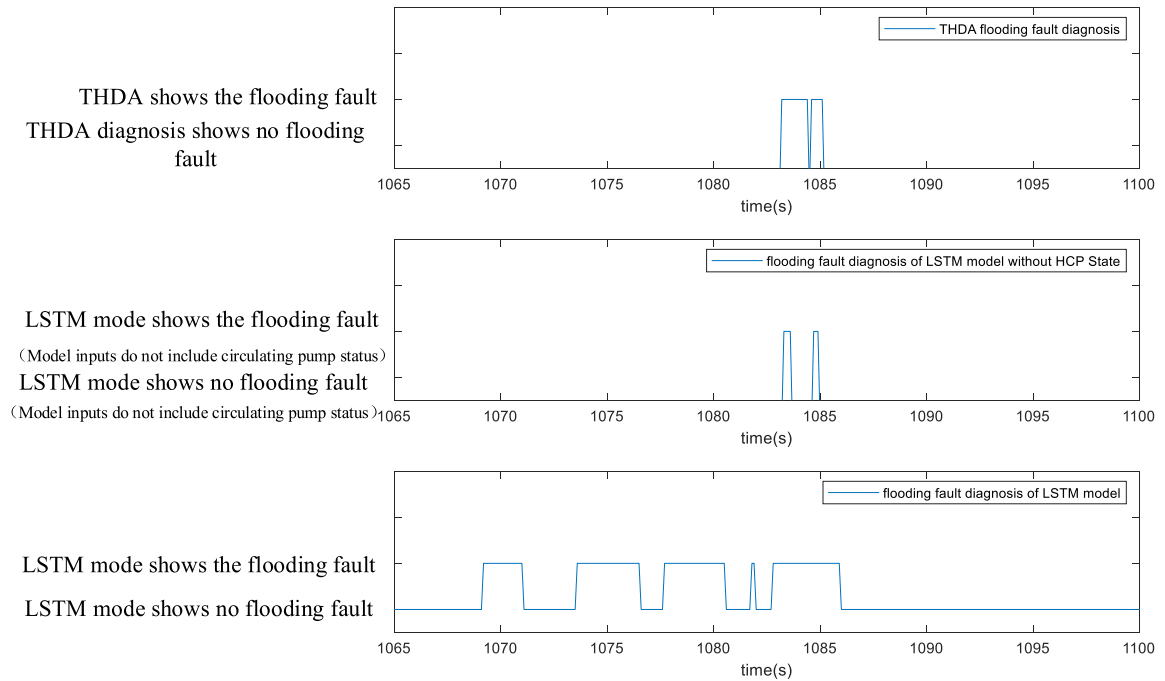


Fig. 15. Comparison of THDA equipment diagnosis results, LSTM (inputs do not include circulating pump status) diagnosis results and original LSTM network diagnosis results.

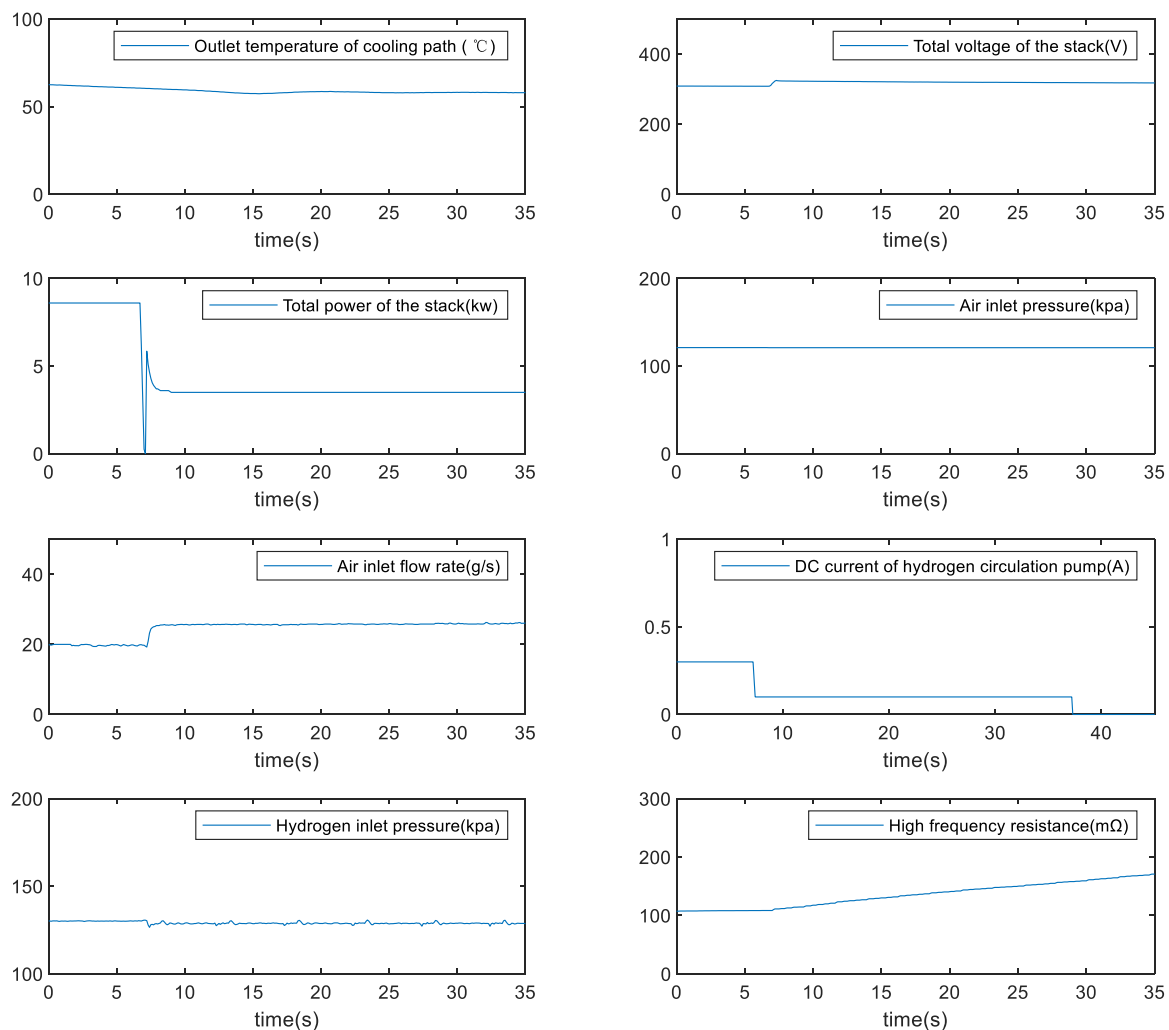


Fig. 16. Operating data of membrane drying.

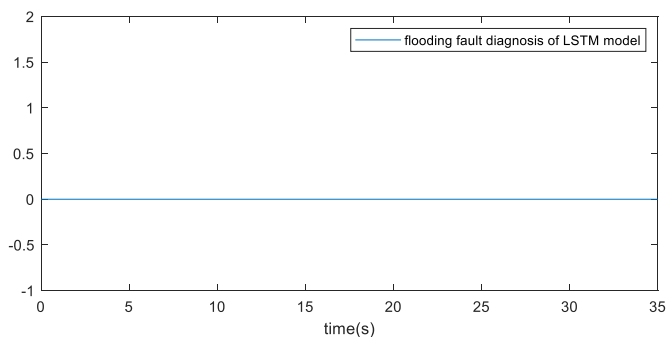


Fig. 17. LSTM model output of membrane drying.

## 6. Conclusion

The electrode flooding has been recognized as an important issue in the current fuel cell stack development. The conventional classification network has no ability to process time series (such as SVM network). It only relies on the information at a single point-in-time, which is considered insufficient in the prediction modeling work of PEM fuel cell. This may lead to unreasonable pre-diagnoses result. With regard to the above problems, LSTM network has been implemented in this work. A classifier is designed to diagnose the flooding faults of PEM fuel cell

stack. The modeling validation work is also carefully carried out. The modeling prediction shows reasonable agreement with the experimental tests. Therefore, this work has been used to develop controlling strategies and avoid flooding faults in the fuel cell system.

A close loop system relies heavily on the large number of sensors to capture and diagnose the operational state of fuel cell stack, which results in higher cost. In this study, the LSTM diagnosis network is proposed to integrate auxiliary system states and the sensor information in the inlet and outlet of fuel cell, which makes it possible to decrease the number of sensors used in the system, further leading to efficient fuel cell system development.

Furthermore, in order to promote adaptability of the modeling work under various vehicle conditions, more training data are needed to improve the diagnostic accuracy.

## Declaration of Competing Interests

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: All three authors are employed by Shanghai Hydrogen Propulsion Technology Co., Ltd.

## Acknowledgements

This work was supported by Shanghai Hydrogen Propulsion Technology Co., Ltd.

## Reference

- [1] Zhan Z, Wang C, Fu W, et al. Visualization of water transport in a transparent PEMFC. *Int J Hydrog Energy* 2012;37(1):1094–105.
- [2] Hissel D, et al. Model-based diagnosis for proton exchange membrane fuel cells. *Math Comput Simul* 2010;81(2):158–70.
- [3] Iranzo A, Boillat P. Liquid water distribution patterns featuring back-diffusion transport in a PEM fuel cell with neutron imaging. *Int J Hydrog Energy* 2014;39(30):17240–5.
- [4] Seong J Y, Bae Y C, Sun Y K. Water activities of polymeric membrane/water systems in fuel cells. *J Power Sources* 2006;157(2):733–8.
- [5] Li H, et al. A review of water flooding issues in the proton exchange membrane fuel cell. *J Power Sources* 2008;178(1):103–17.
- [6] Petrone R, et al. A review on model-based diagnosis methodologies for PEMFCs. *Int J Hydrog Energy* 2013;38(17):7077–91.
- [7] Zheng Z, et al. A double-fuzzy diagnostic methodology dedicated to online fault diagnosis of proton exchange membrane fuel cell stacks. *J Power Sources* 2014;271:570–81.
- [8] Mohammadi M, et al. Fuzzy logic and passivity-based controller applied to electric vehicle using fuel cell and supercapacitors hybrid source. *Energy Procedia* 2014;50:619–26.
- [9] Silva RE, et al. Proton exchange membrane fuel cell degradation prediction based on adaptive neuro-fuzzy inference systems. *Int J Hydrog Energy* 2014;39(21):11128–44.
- [10] Cheah MJ, Kevrekidis IG, Benziger JB. Water slug formation and motion in gas flow channels: the effects of geometry, surface wettability, and gravity. *Langmuir* 2013;29(31):9918–34.
- [11] Liu X, et al. Water flooding and pressure drop characteristics in flow channels of proton exchange membrane fuel cells. *Electrochim Acta* 2007;52(11):3607–14.
- [12] Banerjee R, Kandlikar SG. Experimental investigation of two-phase flow pressure drop transients in polymer electrolyte membrane fuel cell reactant channels and their impact on the cell performance. *J Power Sources* 2014;268:194–203.
- [13] Banerjee R, et al. Experimental validation of two-phase pressure drop multiplier as a diagnostic tool for characterizing PEM fuel cell performance. *Int J Hydrog Energy* 2014;39(31):17791–801.
- [14] Anderson R, et al. A critical review of two-phase flow in gas flow channels of proton exchange membrane fuel cells. *J Power Sources* 2010;195(15):4531–53.
- [15] Pei P, et al. Hydrogen pressure drop characteristics in a fuel cell stack. *Int J Hydrog Energy* 2006;31(3):371–7.
- [16] Kim J, Tak Y. Implementation of discrete wavelet transform-based discrimination and state-of-health diagnosis for a polymer electrolyte membrane fuel cell. *Int J Hydrog Energy* 2014;39(20):10664–82.
- [17] Li Z, Outbib R, Hissel D, et al. Data-driven diagnosis of PEM fuel cell: a comparative study. *Control Eng Pract* 2014;28(jul.):1–12.
- [18] Li Z, Outbib R, Hissel D, et al. Diagnosis for PEMFC systems: a data-driven approach with the capabilities of online adaptation and novel fault detection. *IEEE Trans Ind Electron* 2015;62(8):5164–74.
- [19] Zhou Shangwei, Dhupia JS. Online adaptive water management fault diagnosis of PEMFC based on orthogonal linear discriminant analysis and relevance vector machine. *Int J Hydrog Energy* 2020;45(11).
- [20] Zhang S, Wang Y, Liu M, et al. Data-based line trip fault prediction in power systems using LSTM networks and SVM. *IEEE Access* 2017;1 -1.
- [21] Wang Q, Guo Y, Yu L, et al. Earthquake prediction based on spatio-temporal data mining: an LSTM network approach. *IEEE Trans Emerg Top Comput* 2017;1 -1.
- [22] Shao B, Hu X, Bian G, et al. A multichannel LSTM-CNN method for fault diagnosis of chemical process. *Math Probl Eng* 2019;2019(3):1–14.
- [23] Ijaodola OS, El-Hassan Z, Ogungbemi E, et al. Energy efficiency improvements by investigating the water flooding management on proton exchange membrane fuel cell (PEMFC). *Energy* 2019;179(JUL.15):246–67.
- [24] Wei D, Haijiang W, Xiao-Zi Y, Jonathan JM, Daijun Y, Jinli Q, Jianxin M. A review on water balance in the membrane electrode assembly of proton exchange membrane fuel cells. *Int J Hydrog Energy* 2009;34(23):9461e78.
- [25] Thomas AZ, John D, Valerio Judith, Shimshon G. The water content dependence of electro-osmotic drag in proton-conducting polymer electrolytes. *Electrochim Acta* 1995;40(3):297e302.
- [26] Yong HP, Jerald AC. An experimental investigation of electro-osmotic drag coefficients in a polymer electrolyte membrane fuel cell. *Int J Hydrog Energy* December 2008;33(24):7513e20.
- [27] Zhiping L, Zhangyong C, Yuxia Z, Zhen L, Jing L. Electro-osmotic drag coefficient and proton conductivity in Nafion® membrane for PEMFC. *Int J Hydrog Energy* 2010;35(7):3120–4.
- [28] Springer TE, Zawodzinski TA, Gottesfeld S. Polymer electrolyte fuel cell model. *J Electrochem Soc* 1991;138(8) J. Electrochem. Soc. August.
- [29] Qiao Z, Paul M, Jay B. Diffusion and interfacial transport of water in Nafion. *J. Phys Chem B* 2011;115(12):2717–27.
- [30] Majsztrik P, Bocarsly A, Benziger J. Water permeation through nafion Membranes: the role of water activity. *J Phys Chem B* 2008;112(51):16280–9.
- [31] Soowhan K, Mench MM. Investigation of temperature-driven water transport in polymer electrolyte fuel cell: thermo-osmosis in membranes. *J Membr Sci* 2009;328:113–20.
- [32] German C, Mariano A, Ana MCL. Effect of water content in the gas diffusion layer of H<sub>2</sub>/O<sub>2</sub> PEM fuel cell. *J Mater Sci Eng* 2016;6(7e8):213–21.
- [33] Hussaini IS, Wang CY. Visualization and quantification of cathode channel flooding in pem fuel cells. *J Power Sources* 2009;187(2):444–51.
- [34] Hsieh SS, Huang YJ, Her BS. Pressure drop on water accumulation distribution for a micro PEM fuel cell with different flow field plates. *Int. J. Heat Mass Transf.* 2009;52(23-24):5657–9.
- [35] Graves A. Supervised sequence labelling with recurrent neural networks. Vol 385. 2012th ed. Berlin, Heidelberg: Springer Berlin Heidelberg.