Fault Diagnosis of PEMFC Systems Based on Decision-making Tree Classifier

Jiawei Liu, Qi Li*, Weirong Chen, Yu Yan, Lu Jiang
School of Electrical Engineering
Southwest Jiaotong University
Chengdu, China
liqi0800@163.com

***Abstract*—For purpose of solving the problem of fault diagnosis of the evaporatively cooled fuel cell system, a novel fast fault diagnosis method for an evaporatively cooled fuel cell system based on decision-making tree classifier is developed in the paper. The normalization method is utilized to filter the original datum. The decision-making tree classifier is used to classify the preprocessed data. It can effectively improve the diagnostic accuracy of the model. Example analysis shows that the novel approach can rapidly recognize two health states of membrane drying fault and hydrogen leakage fault. The diagnostic accuracy of the algorithm is 98.50%. The method proposed in this paper is suitable for online failure identification of evaporatively cooled fuel cell systems with large data samples and multi data dimensions.**

Keywords—decision-making tree, PEMFC, fault diagnosis, artificial intelligence

I. Introduction

Proton exchange membrane fuel cell (PEMFC) is a hopeful new generation technology. It has been widely used in electric vehicles, electric buses, hybrid electric vehicles, and tramcars [1-4]. In order to commercialize PEMFC systems on a large scale, improving the durability of fuel cells is one of the most important challenges [5-8].

Fault diagnosis is designed to improve the life of the fuel cell system by detecting faults. In order to achieve this goal, the diagnosis method must detect and identify some faults to avoid more serious failures. The detection phase is to determine whether there is a fault in the PEMFC systems. The identification phase is to determine the size and location of the fault, and the isolation phase is to overhaul and troubleshoot.

At present, artificial intelligence technology is gradually applied in the area of PEMFC failure diagnostics. Jiawei Liu et al. [9, 10] have designed a novel fault diagnosis approach for tramcar fuel cell system based on K-means clustering and a discrete hidden Markov model (DHMM). It can effectively detect the low pressure of air, hydrogen leak, the voltage over the range of outlet temperature signal of deionized ethylene glycol, normal state, the high inlet temperature of deionized ethylene glycol and the low pressure of deionized humidification pump of six kinds of health status. The diagnostic accuracy is 94.17%. Zhongliang Li et al. [11] have presented a PEMFC system failure detection strategy considering spatial heterogeneity and system dynamics. The single cell voltage value collected in the sliding window is used as the diagnostic variables. Shapelet transform is used to abstract the discriminant characteristics. The PEMFC fault detection and isolation (FDI) is realized by the sphere shaped multi-class support vector machines (SSM-SVM). The experiment results indicate that the approach can diagnose the fault of low air stoichiometric ratio, the fault of high voltage, normal-state, the fault of the membrane drying and the fault of low voltage. The average correct detection rate of the test sample is 96.13%. Lei Mao et al. [12-15] have proposed an optimal sensor selection approach based on PEMFC failure operating mode and sensor sensitivity. The experimental results show that the optimal sensor can successfully identify different degrees of fuel cell flooding faults. Zhixue Zheng et al. [16-18] have developed a PEMFC failure detection scheme based on reserve computing (RC). The laboratory results demonstrate that the RC strategy can identify a low flow rate of air, natural degradation, poor heat dissipation and poisoning of carbon monoxide. The classification accuracy of the data set is 92.43%.

Decision-making tree classifiers are instance-based aposteriori learning approaches. The algorithm infers the classification regulation of the decision-making tree presentation from a series of unordered and irregular conditions. Decision-making trees can also be represented as multiple “If-Then” rules. Generally, the recursion method of "top-down, divide and conquer" is adopted in the decision-making tree. The search space is divided into several mutually disjoint subsets. Attribute features are compared at the internal-nodes (non-leaf nodes) of the decision-making tree classifier. The lower branches from the node are determined according to different attribute values. The conclusion is drawn at the leaf node of the tree.

A novel fast fault diagnosis method for evaporatively cooled (EC) fuel cell system based on decision-making tree classifier is proposed for the first time. It can significantly improve the accuracy of diagnosis. The novel method uses a normalization strategy to process raw data and classifies feature vectors by decision-making tree classifier. The effectiveness of the novel method is verified through 200 sets of original fault samples.

This work was supported by National Natural Science Foundation of China (61473238) and NEEC Open-end Fund of China (NEEC-2017-B01).

978-1-5386-8549-5/18/$31.00 ©2018 IEEE

II. EVAPORATIVELY COOLED FUEL CELL SYSTEM AND FAULT TYPE

A. *Evaporatively Cooled Fuel Cell Systems*

The EC fuel cell system in this paper consists of two 100 kW PEMFC stacks, each of which consists of 300 single cells. The topologies of the EC fuel cell systems are given in Fig. 1.

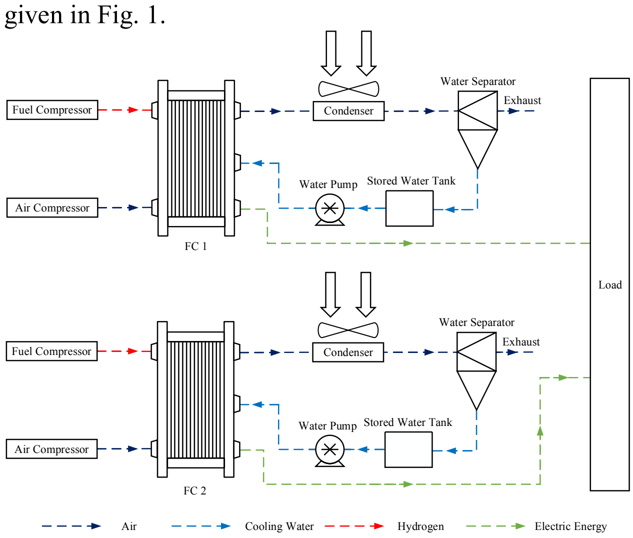


Fig. 1 Principle scheme of the EC fuel cell systems [13]

During the operation of the EC fuel cell systems, air and hydrogen are input into the cathode and anode of the PEMFC stack separately. Liquid water is input into the stack, too. Some of the liquid water evaporates to form a liquid/steam mixture, which can take heat away from the stack. In addition, on behalf of maintaining the system water balance, the heat exchanger is used to coagulate the liquid water at the outlet of the stack to restore and replace enough water to water chamber [13]. For purpose of collecting the complete information, numerous sensors are installed in the entrance and exit of the PEMFC stack. The performance of the stack is evaluated by collecting the flow rate, pressure, temperature and voltage of the PEMFC stack. The monitoring variables are summarized in Tab. 1.

Tab. 1 Measured variables

Variable	**Unit**	**Variable**	**Unit**
Anode inlet pressure #1	mbar	Cathode inlet temperature #1	℃
Anode inlet pressure #2	mbar	Cathode inlet temperature #2	℃
Anode outlet pressure #1	mbar	Cathode outlet temperature #1	℃
Anode outlet pressure #2	mbar	Cathode outlet temperature #2	℃
Primary water inlet pressure #1	mbar	Primary water inlet flow #1	SLPM
Primary water inlet pressure #2	mbar	Primary water inlet flow #2	SLPM
Cathode inlet pressure #1	mbar	Anode reactant flow	SLPM
Cathode outlet pressure #1	mbar	Cathode air inlet flow	SLPM
Cathode outlet pressure #2	mbar	Cathode stoichiometry	N/A
Load current	A	Coolant water inlet temperature	℃

B. *Fault Type*

1) Membrane Drying Failure

Relative humidity may affect the ohmic impedance of proton exchange membrane (PEM). PEM needs to contact with liquid water, and higher membrane relative humidity can ensure good ionic conductivity.

The changes in the relative humidity of the anode and the cathode, the overheating of the fuel cell working temperature, the bad humidification of the gas and the failure of the cooling system all may lead to the problems of the insufficient liquid water content in the membrane, the loss of the membrane electrode, the decrease of the conductivity and the increase of the PEM impedance, which lead to the membrane drying. The low water content will significantly affect proton conductivity, which may cause ohmic overpotential and output voltage drop of PEMFC. Continuous membrane drying will cause irreversible damage, and even lead to the excessive local temperature inside the PEMFC stack and shorten the remaining service life of PEMFC.

2) Hydrogen Leakage Failure

The hydrogen leak is caused by the failure of the PEMFC to obtain sufficient fuel supply at the load current, which will cause the operating voltage to drop.

When the hydrogen leakage fault is very serious, the PEMFC stack voltage gradually drops to zero and becomes negative, which will lead to the "reverse pole". If PEMFC continues to operate in the presence of "reverse pole", oxygen may be precipitated from the single battery at the anode, which will greatly reduce the voltage of PEMFC by entering the adjacent single cell through the cell gas pipeline. Hydrogen-oxygen mixing may cause hydrogen and oxygen to directly react and burn on the cathode catalyst, causing an explosion inside the fuel cell, lowering the open circuit voltage of the PEMFC, accelerating the degradation of the membrane electrode and shortening the remaining life of the fuel cell, and serious hydrogen leakage failure will damage the PEMFC.

III. BASIC THEORY OF DECISION-MAKING TREE CLASSIFIER

A. *Basic Idea and Representation of Decision-Making Tree Classifier*

The decision-making tree classifier categorizes patterns by formatting them from root node to the leaf node. A respective nonleaf node in the tree stands for a group of the value of a property. Its branches stand for every test result. Every leaf node in the tree stands for a category. The root node is as the top node of the tree.

In short, the decision-making tree of the tree-formed structure is similar to a flowchart. The top-down recursion is adopted. Start with the root node of the tree and test the comparison of property values on its inner nodes. The matching branch is then determined on the basis of the property value of the given case. The consequence is obtained at the leaf node of the decision-making tree. The procedure is repeated on the subtree with the updated node as its parent.

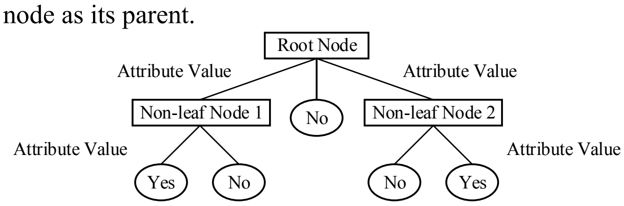


Fig. 2 Structure diagram of a decision-making tree classifier

Fig. 2 is the structure diagram of the decision-making tree classifier. Every non-leaf node stands for the input property of the learning set data. “Attribute Value” represents the value of the property. The leaf node represents the value of the object category attribute. The "No" and "Yes" in Fig. 2 separately stand for negative and positive examples in the example space.

B. Iterative Dichotomic Version 3 Algorithm

Heretofore, there are many decision-making tree generation algorithms. Among them, the most influential example learning algorithm in the world is the iterative dichotomic version 3 (ID 3) algorithm of J.R Quinlan.

The pioneering work of Quinlan is mainly to introduce the notion of common information in information theory in the training algorithms of the decision-making tree. He calls it information gain (IG), which is called the standard of attribute selection.

To exactly regulate the IG, a metric standard widely adopted in the information theory called entropy is defined, which represents the fineness of all sample dataset.

If the object properties have c unlike values, the entropy of the data set S involved in the c states can be defined as:

$$Entropy(S) = \sum_{i=1}^{c} -p_i \log_2 p_i \quad (1)$$

Where, p_i is the ratio of the sample number of the i th property value in the subset.

It can be obtained from the above equation: If entire instances in dataset S remain with the uniform category, then $Entropy(S) = 0$; If the number of samples in 2 classes isn’t adequate, then $Entropy(\mathrm{S}) \in (0,1)$.

Particularly, if the set S is a Boolean set, that is total instances in set S remain with 2 diverse categories. If the number of samples in the 2 classes is equivalent, there are $Entropy(\mathrm{S}) = 1$.

Entropy has been adopted as a standard to scale the set purity of learning samples. The definition of IG Gain(S, A) is:

$$Gain(\mathrm{S}, \mathrm{A}) = \mathrm{Entropy}(\mathrm{S}) = \sum_{v \in V(A)} \frac{|\mathrm{S}_v|}{|\mathrm{S}|} Entropy(S_v) \quad (2)$$

Where $V(A)$ is the range of particularity A ; S_v is a subcollection of the set S that is equivalent to v in attribute A .

After the concept of IG is introduced, the basic procedure of the ID3 algorithm will be described in detail. “Examples” can be set as training sample sets and “Attribute list” as candidate Attribute sets.

1. Create the root node N of the decision-making tree;
2. If total samples remain with the same class C , N is replaced as a leaf node and noted as C category;
3. If the “Attribute list” is bare, N is returned as a leaf node and marked as the class with the largest number of samples in the sample;
4. The IG of each candidate particularity in “Attribute list” is computed. The matching attribute “ $Attribute^*$ ” of the maximum IG is chosen to mark the root node N ;
5. On the basis of every value V_i in the attribute “ $Attribute^*$ ” field, the corresponding branch is produced from the root node N and the S_i is a set of samples which satisfies the $Attribute^* = V_i$ condition in the examples set;
6. If S_i is empty, the matching leaf node is noted as the category with the maximal number of categories in the “Examples” sample datasets; else, the $Attribute^*$ is deleted from “Attribute list” and step 1 is returned to invent the subtree recursively.

IV. Example Analysis

A. Data Acquisition

The purpose of this paper is to diagnose the two health states of membrane drying fault and hydrogen leakage fault. 200 groups of data are collected in each health state and a total of 400 groups of samples are collected. The 200 groups of data are stochastically chosen as learning samples and the remaining 200 groups are adopted as test samples.

B. Feature Extraction

In Tab. 1, 20 variables of sensor monitoring are selected to form a 20-dimensional fault feature vector and the sample is normalized preprocessed.

C. Diagnostic Decision

The decision-making tree is created by the function of “*ctree* = ClassificationTree.fit(*P_train*, *T_train*)”, where the *ctree* is the node of affiliation, the *P_train* is the 20-dimensional normalized variables of the training sample, and the *T_train* is the label of the training set. The running environment is based on the MATLAB R2018b. The decision-making tree is created as shown in Fig. 3. If X3<0.521467，then node 2；if X3⩾0.521467, then node 1. Node 2 stands for class 1 and node 1 stands for class 2.

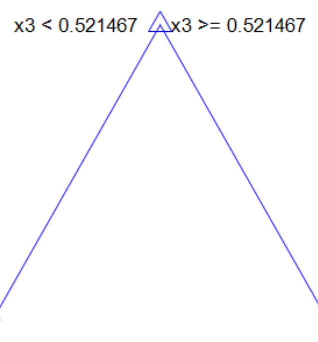
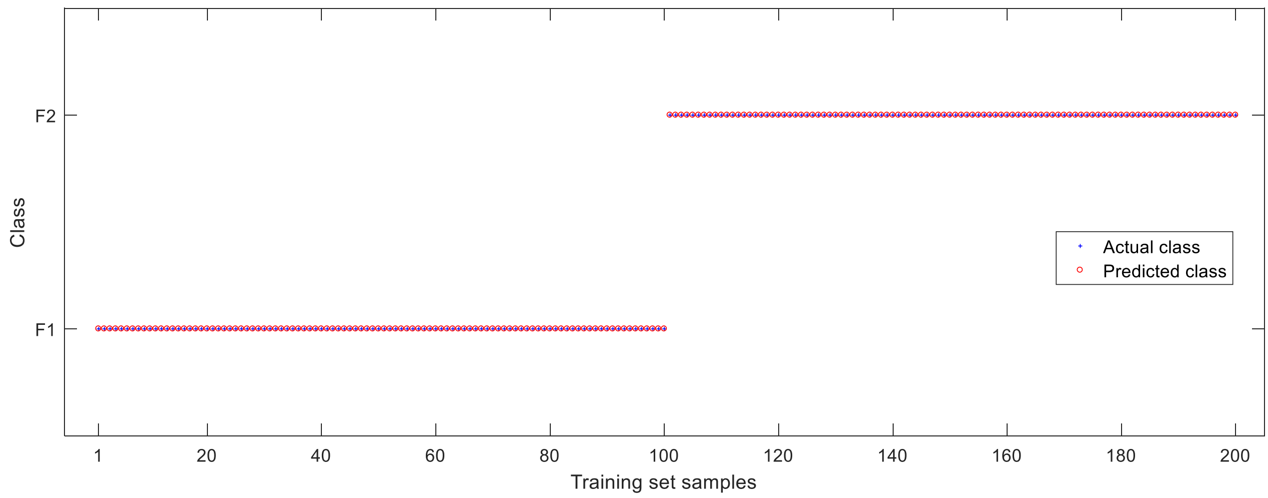
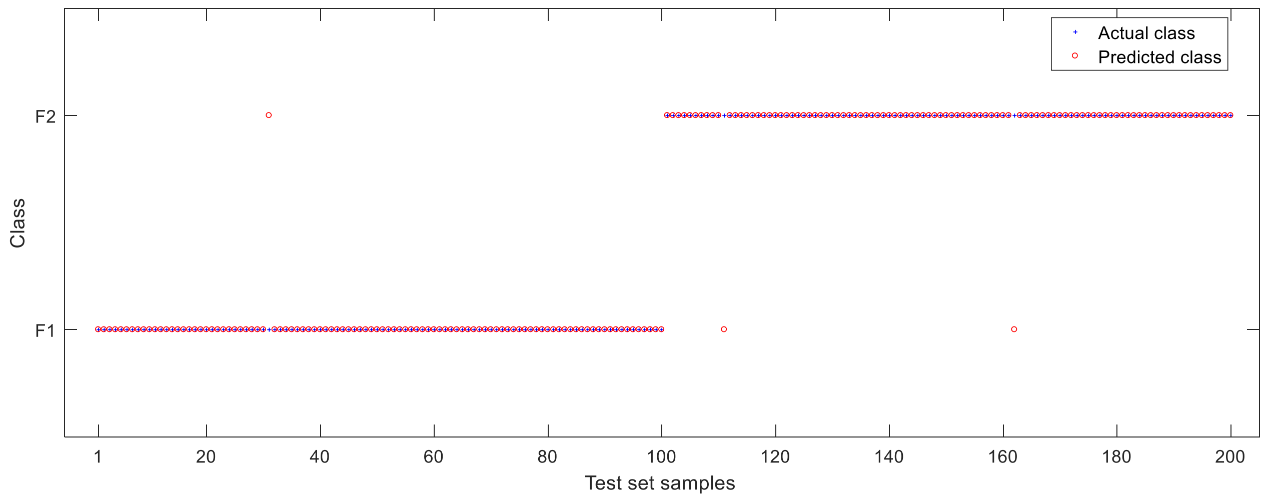


Fig. 3 Decision-making tree classifier created



(a) Classification results of the training sample



(b) Classification results of the test sample

Fig. 4 Fault diagnosis for PEMFC system based on decision-making tree

The classification results of training datasets and test datasets are as shown in Fig. 4(a) and Fig. 4(b) respectively. Among them, the abscissa represents training samples and test samples respectively, and ordinates represent fault types. F1 represents the membrane drying fault and F2 represents the hydrogen leak fault.

As can be seen from the results, a total of 3 samples in the 200 samples of the test set are predicted to be incorrect (where one membrane drying fault sample is misclassified into hydrogen leak fault and two hydrogen leak fault samples are misclassified into membrane drying fault). The average diagnostic accuracy rate is 98.50%(197/200). It indicates that the created decision-making tree classifier can be applied to the failure recognition of the PEMFC systems.

V. Conclusion

Failure diagnosis of EC fuel cell systems is studied in this chapter. A novel fault diagnosis approach based on a decision-making tree classifier is also developed. The diagnosis of membrane drying fault and hydrogen leakage fault is realized. The conclusions are as follows:

1) An example analysis of 200 sets of raw fault sample data shows that the novel method can quickly diagnose the two health states of membrane drying fault and hydrogen leakage fault. The correct rate of classification is 98.50%.

2) Decision-making tree classifier tends to understand and accomplish. It is friendly to absorb the significance of the decision-making tree classifier after interpretation. The decision-making tree classifier is a white box model. Given an observation model, the corresponding logical expression can easily be derived from the resulting decision-making tree.

3) The novel method is especially suitable for engineering applications with large data samples and high data dimensions. It has good application value in the era of big data.

VI. Acknowledgment

I would like to thank Prof. Lei Mao with the

Department of Aeronautical and Automotive Engineering, Loughborough University, Loughborough, U.K. Prof. Mao Lei provides and discloses the access to Loughborough University Data Repository (https://lboro.figshare.com), so that I can develop the proposed approach based on the real industrial datum.

REFERENCES

[1] Y. Li, P. Pei, Z. Wu, P. Ren, X. Jia, D. Chen, and S. Huang, "Approaches to avoid flooding in association with pressure drop in proton exchange membrane fuel cells," *Applied Energy,* vol. 224, pp. 42-51, 2018.

[2] L. Mao and L. Jackson, "Effect of Sensor Set Size on Polymer Electrolyte Membrane Fuel Cell Fault Diagnosis," *Sensors,* vol. 18, p. 2777, 2018-08-23 2018.

[3] Z. Li, R. Outbib, S. Giurgea, D. Hissel, S. Jemei, A. Giraud, and S. Rosini, "Online implementation of SVM based fault diagnosis strategy for PEMFC systems," *Applied Energy,* vol. 164, pp. 284-293, 2016.

[4] Q. Li, T. Wang, C. Dai, W. Chen, and L. Ma, "Power Management Strategy Based on Adaptive Droop Control for a Fuel Cell-Battery-Supercapacitor Hybrid Tramway," *IEEE Transactions on Vehicular Technology,* vol. 67, pp. 5658-5670, 2018.

[5] Z. Li, R. Outbib, S. Giurgea, D. Hissel, and Y. Li, "Fault detection and isolation for Polymer Electrolyte Membrane Fuel Cell systems by analyzing cell voltage generated space," *Applied Energy,* vol. 148, pp. 260-272, 2015.

[6] Z. Li, R. Outbib, S. Giurgea, and D. Hissel, "Diagnosis for PEMFC Systems: A Data-Driven Approach With the Capabilities of Online Adaptation and Novel Fault Detection," *IEEE Transactions on Industrial Electronics,* vol. 62, pp. 5164-5174, 2015.

[7] Z. Li, R. Outbib, D. Hissel, and S. Giurgea, "Data-driven diagnosis of PEM fuel cell: A comparative study," *Control Engineering Practice,* vol. 28, pp. 1-12, 2014.

[8] Z. Li, S. Giurgea, R. Outbib, and D. Hissel, "Fault diagnosis and novel fault type detection for PEMFC system based on spherical-shaped multiple-class support vector machine," in *2014 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*: IEEE, 2014, pp. 1628 - 1633.

[9] J. Liu, Q. Li, W. Chen, and T. Cao, "A discrete hidden Markov model fault diagnosis strategy based on K-means clustering dedicated to PEM fuel cell systems of tramways," *International Journal of Hydrogen Energy,* vol. 43, pp. 12428-12441, 2018.

[10] J. Liu, Y. Yan, Q. Li, and W. Chen, "mRVM based fault diagnosis strategy for PEMFC systems of hybrid tramway," in *2017 Chinese Automation Congress (CAC)* Jinan, China: IEEE, 2017, pp. 1096 - 1100.

[11] Z. Li, R. Outbib, S. Giurgea, and D. Hissel, "Fault diagnosis for PEMFC systems in consideration of dynamic behaviors and spatial inhomogeneity," *IEEE Transactions on Energy Conversion,* pp. 1-9, 2018.

[12] L. Mao, L. Jackson and B. Davies, "Effectiveness of a Novel Sensor Selection Algorithm in PEM Fuel Cell On-Line Diagnosis," *IEEE Transactions on Industrial Electronics,* vol. 65, pp. 7301-7310, 2018.

[13] L. Mao, L. Jackson and S. Dunnett, "Fault Diagnosis of Practical Polymer Electrolyte Membrane (PEM) Fuel Cell System with Data-driven Approaches," *Fuel Cells,* vol. 17, pp. 247-258, 2017.

[14] L. Mao, L. Jackson and B. Davies, "Investigation of PEMFC fault diagnosis with consideration of sensor reliability," *International Journal of Hydrogen Energy,* 2017.

[15] L. Mao, B. Davies and L. Jackson, "Application of the Sensor Selection Approach in Polymer Electrolyte Membrane Fuel Cell Prognostics and Health Management," *Energies,* vol. 10, p. 1511, 2017.

[16] Z. Zheng, S. Morando, M. Pera, D. Hissel, L. Larger, R. Martinenghi, and A. Baylon Fuentes, "Brain-inspired computational paradigm dedicated to fault diagnosis of PEM fuel cell stack," *International Journal of Hydrogen Energy,* vol. 42, pp. 5410-5425, 2017.

[17] Z. Zheng, R. Petrone, M. C. Pera, and D. Hissel, "Diagnosis of a commercial PEM fuel cell stack via incomplete spectra and fuzzy clustering," in *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society* Vienna, Austria, 2013, pp. 1595-1600.

[18] R. Petrone, Z. Zheng, D. Hissel, M. C. Péra, C. Pianese, M. Sorrentino, M. Becherif, and N. Yousfi-Steiner, "A review on model-based diagnosis methodologies for PEMFCs," *International Journal of Hydrogen Energy,* vol. 38, pp. 7077-7091, 2013.