HAL
open science

Online implementation of SVM based fault diagnosis strategy for PEMFC systems

Zhongliang Li, Rachid Outbib, Stefan Giurgea, Daniel Hissel, Samir Jemei, Alain Giraud, Sébastien Rosini

To cite this version:

Zhongliang Li, Rachid Outbib, Stefan Giurgea, Daniel Hissel, Samir Jemei, et al.. Online implementation of SVM based fault diagnosis strategy for PEMFC systems. Applied Energy, 2016, 164, pp.284-293. 10.1016/j.apenergy.2015.11.060 . cea-01846859

HAL Id: cea-01846859

https://cea.hal.science/cea-01846859v1

Submitted on 7 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Online implementation of SVM based fault diagnosis strategy for PEMFC systems

Zhongliang Li[a,b,*], Rachid Outbib[c], Stefan Giurgea[a,b], Daniel Hissel[a,b], Samir Jemei[a,b], Alain Giraud[d], Sebastien Rosini[e]

[a] *FCLAB (Fuel Cell Lab) Research Federation, FR CNRS 3539, rue Thierry Mieg, 90010 Belfort Cedex, France*
[b] *FEMTO-ST (UMR CNRS 6174), ENERGY Department, UFC/UTBM/ENSMM, France*
[c] *Laboratoire des Sciences de l'Information et des Systemes (LSIS), University of Aix-Marseille, France*
[d] *CEA/LIST, 91191 Gif-sur-Yvette Cedex, France*
[e] *CEA/LITEN, 38054 Grenoble, France*

Abstract

In this paper, the topic of online diagnosis for Polymer Electrolyte Membrane Fuel Cell (PEMFC) systems is addressed. In the diagnosis approach, individual cell voltages are used as the variables for diagnosis. The pattern classification tool Support Vector Machine (SVM) combined with designed diagnosis rule is used to achieve fault detection and isolation (FDI). A highly-compacted embedded system of the System in Package (SiP) type is designed and fabricated to monitor individual cell voltages and to perform the diagnosis algorithms. For validation, the diagnosis approach is implemented online on PEMFC experimental platform. Four concerned faults can be detected and isolated in real-time.

Keywords: PEMFC system, Fault diagnosis, SVM classification, System in Package, Online implementation

1. Introduction

The environment and resource issues have been drawing increasing attention of the world. Global warming and non-renewable resource exhaustion are two problems need to be addressed urgently. One of the main causes of these issues is the high dependence of the fossil fuels in the current energy structure. Since hydrogen can be produced from diverse sources, such as renewable resources and nature gas, large-scale use of hydrogen-based fuel cells is considered as one of the most significant solutions dedicated to slacking the dependence on fossil fuels [1, 2]. Amongst various fuel cells, Polymer Electrolyte Membrane Fuel Cell (PEMFC) is potentially beneficial for a wide range of applications, thanks to its attractive advantages, such as high efficiency, high power density, in-situ zero-emission, low operating temperature, and quick response to load changes [3, 4]. Especially, fuel cell electric vehicle (FCEV), which is equipped with PEMFC stack as the main energy generator, is a powerful competitor in the future automobile

*Corresponding author. Tel.: +33 (0)3 84 58 36 28, Fax: +33 (0)3 84 58 36 36. E-mail address: zhongliang.li@lsis.org (Z. LI), rachid.outbib@lsis.org (R. OUTBIB), stefan.giurgea@utbm.fr (S. GIURGEA), daniel.hissel@univ-fcomte.fr (D. HISSEL), alain.giraud@cea.fr (A.GIRAUD), Sebastien.Rosini@cea.fr (S.ROSINI)

market. Compared to the battery based electric vehicles, FCEVs have the advantages of quick recharging (refueling) and long running range [5, 6]. However, several bottlenecks of PEMFC technologies, such as reliability and durability, still exist and impede the widely commercial exploitation of the PEMFC products [7, 8].

Fault diagnosis, i.e., fault detection and isolation (FDI), is playing an increasingly important role in several kinds of modern industrial systems [9, 10, 11, 12]. It has been found that various faults involving different components of PEMFC systems can occur and cause performance degradations. For instance, the faults related to reactants supply subsystems and the ones related to water management. During the last decades, fault diagnosis devoted to improving the reliability and durability performance of PEMFC systems has drawn the attention of both academic and industrial communities [13, 14]. Through an efficient diagnosis strategy, more serious faults can be avoided thanks to an early fault alarm. With the help of diagnosis results, the downtime (repair time) can be reduced. Moreover, the precise diagnosis information can help to speed up the development of new technologies [15].

Several fault diagnosis strategies have been studied during the last decade [16, 17, 18, 19, 20, 21, 22, 23]. The general model based fault diagnosis theoretical base seems to be well established and some positive results have been obtained by using these methods for some PEMFC systems [16, 17]. Nevertheless, building a model with first principle is not a trivial task. The internal parameters, which are essential for modeling, are not evident to be found or estimated. In addition, model structures and parameters may differ among different designs of fuel cell stacks and other system components. Apart from the model based diagnosis, the application of data based methodologies for the diagnosis of PEMFC systems has been drawing the attention of researchers [18, 19, 20, 21, 22, 23]. Avoiding the sophistical modeling process, the data based diagnosis seems to be more practical in most cases. Actually, the data based diagnosis has been utilized in a number of industrial processes [24, 25, 26].

Within the scope of data based fault diagnosis, a number of pattern classification techniques have been widely used since FDI can be considered as a classification problem [25, 26]. Some classification based diagnosis strategies have been proposed for PEMFC systems (see [18, 20, 21, 23] for instance). Different variables, feature extraction and classification methods have been studied using the historical data. In [18], the classification was supposed to be carried out in the feature space which is generated using multifractal analysis on stack voltage. In [21], the fuzzy classification method was utilized to analyze Electrochemical Impedance Spectroscopy. In our previous study [20], selecting individual cell voltages as the variables for diagnosis, several classification and feature extraction techniques were compared from the perspectives of classification accuracy and computational complexity. The classification method Support Vector Machine (SVM) was selected as the most suitable classification tool in the case. The strategy was further developed by extending the capabilities of novel fault detection and online adaptation [23].

After elaborating the diagnosis strategy, the objective of this work is the online implementation. To do so, three aspects must be considered specially aiming at online implementation. First, for practical applications, such as FCEVs, reducing the volume and cost of hardwares is always required [27, 28]. The embedded system which fulfills the measurements and computation should be designed with compact layout and limited components. Second, the diagnosis approach programed in the embedded system is performed in real-time, which requires the diagnosis algorithms being saved in limited memory space and being handled in a sufficiently short diagnosis cycle [20]. Third, the diagnosis results should be sufficiently reliable and robust. The importance of online implementation cannot be over-emphasized to make the work closer to

industrial applications.

This study is dedicated to realizing the online implementation of a classification based fault diagnosis strategy for PEMFC systems. The approach employs individual cell voltages as the variables for diagnosis and SVM as the classification tool. Apart from the classification algorithm, a diagnosis rule is designed to obtain the diagnosis results based on the raw classification results. An embedded system of the System in Package (SiP) type is designed to precisely monitor the individual cell voltages and perform the diagnosis approach. The diagnosis approach is then integrated into the SiP and verified online in a PEMFC system. Four different faults which are generated deliberately are detected and isolated in real-time.

The paper is organized as follows: In Section 2, the general development of a classification based diagnosis strategy is summarized. Section 3 is dedicated to introducing the experimental platform. Then, the diagnosis algorithms, including the classification method SVM and the diagnosis rule, is presented in Section 4. The online implementation of diagnosis approach is described in Section 5. The diagnosis results are also provided and analyzed in the same section. Finally, the conclusion is made in Section 6.

2. Development process of classification based online fault diagnosis

The development process of the classification based online fault diagnosis strategy for PEMFC systems consists of three stages. Fig. 1 shows the different components and their tasks concerned in different stages, and the data flows in-between these components.

Algorithm training: In this stage, the historical data sampled in the experiments of different health states, i.e. normal operating state and different concerned faulty states, are analyzed using the computer and the software such as Matlab. The objective is to train and verify the diagnostic algorithms offline using the historical data.

Algorithm integration: In this stage, the programs for performing the diagnosis algorithms are coded and burnt into an embedded system, which is designed in consideration of specificities of both the objective PEMFC system and the characteristics of diagnosis algorithms. The embedded system equipped with diagnosis algorithms is then tested using the historical data. Thus, the obtained results can be compared with the results from the computer. The objective of this step is to ensure the algorithm can be loaded and run correctly in the embedded system.

Online realization: After the first two stages, the embedded system integrated with diagnosis approach is installed into the real PEMFC systems. The online tests are carried out using the real-time data. This step is operated just as the real situation. The objective is to make sure the different subsystems can cooperate as expected.

Generally speaking, in the literature (see for instance [18, 20, 21, 23]), the strategy is tested offline, i.e., the first stage. Here, we consider the feasibility of the second stage dedicated to the integration of the algorithm of classification, and the third stage, which is the online realization on the real process. Hence, the complete process of implementation is realized.

3. Diagnosis strategy development platform

As Fig. 2 shows, the development platform dedicated to online implementation of the diagnosis strategy consists of the following parts:

- PEMFC system
  The schematic of the whole PEMFC system is shown in Fig. 3. The core of the system is

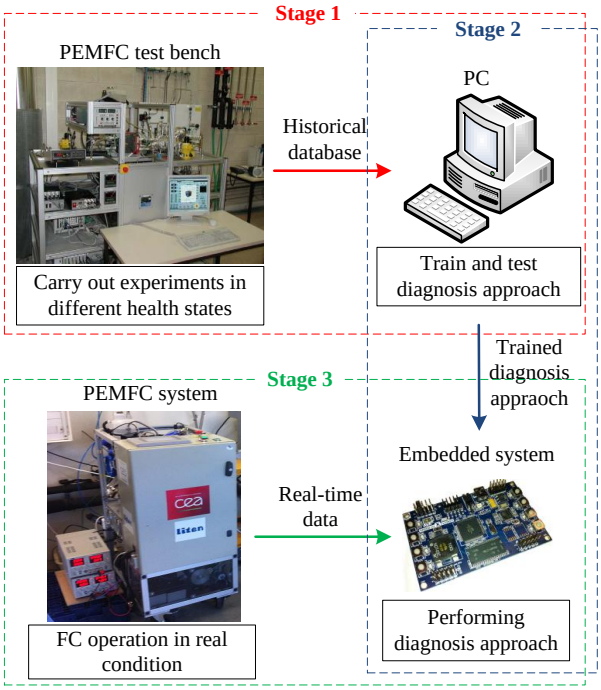


Figure 1. Developing process of classification based online fault diagnosis strategy

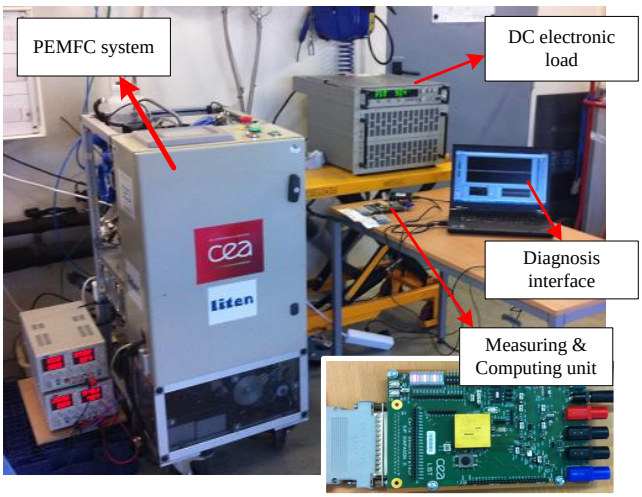


Figure 2. Overview of the development platform

a 64-cell stack which was fabricated by the French research organization CEA[1] specially for automotive application. The nominal operating conditions of the stack are summarized in Table 1. In normal or fault free state, all the operating parameters should be maintained at the nominal values with small variation caused by system noises. It is considered in our case that a fault in the PEMFC system occurs when the operating parameters are out of the normal range. Notice that the faults defined in our case are not limited to the fuel cell stack but cover the health states of the whole system. Certain faults, such as air pressure exceeds the normal range, may not cause obvious degradation with respect to fuel cell stack.

Table 1. Nominal conditions of the stacks

Parameter	Value
Stoichiometry H_2	1.5
Stoichiometry Air	2
Pressure at H_2 inlet	150 kPa
Pressure at Air inlet	150 kPa
Maximum differential of anode pressure and cathode pressure	30 kPa
Temperature (exit of cooling circuit)	65-70 °C
Anode relative humidity	50%
Cathode relative humidity	50%
Current	90 A
Voltage per cell	0.7 V
Electrical power	4032 W

The air is supplied from the environment. Thanks to the compressor, valve and mass flow regulator, the air flow rate and the pressure at inlet can be regulated. By using a humidifier, the hygrometry level of the fed air can be regulated to the required value.

Hydrogen is supplied from a high pressure tank. The pressure at the hydrogen inlet can be controlled thanks to the pressure regulator; namely a valve. The hydrogen flow rate can also be set through the regulator located at downstream of the stack.

The system is equipped with a temperature regulation subsystem in which the thermal exchanging medium is deionized water. The temperature measured at the water outlet is considered as the temperature of the stack.

The system is operated through the Labview interface. The parameters such as pressures, flow rates, relative humidities, temperatures of reactants, cell voltages can also be monitored and saved through the same interface.

- DC load
  The load current can be flexibly varied through an electronic load.

- Measuring and computing unit
  The measuring and computing unit is devoted to measuring the variables for diagnosis and performing online the diagnosis approach. The core component is the specially-designed SiP chip (yellow square component shown in Fig. 2).

[1]CEA: Alternative Energies and Atomic Energy Commission

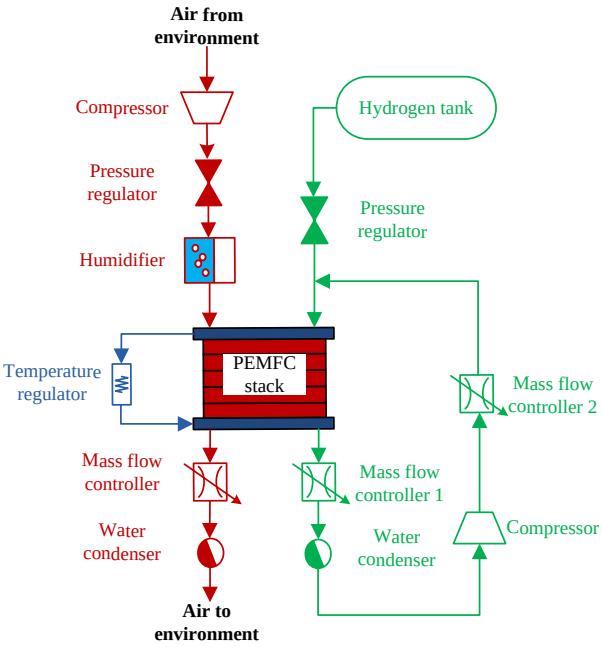


Figure 3. Schematic of the PEMFC system

The structure of the SiP is shown in Fig. 4. Here, the upper layer (see Fig. 4) can be seen as the "main board" which is equipped with a Smartfusion on-chip system developed by Microsemi. The device integrates an FPGA fabric, ARM Cortex-M3 Processor, and programmable analog circuitry [29]. Another two chips of 16 M memory are also added to the system. With the abundant connecting ports, several kinds of communications can be realized with other devices. This on-chip system is ideal hardware for embedded designers as it provides more flexibility than traditional fixed-function microcontrollers [29]. The other two layers, which are equipped with Giant MagnetoResistive (GMR) sensors, are used for measuring individual cell voltages precisely [30].

The block of SiP is settled on an electric board which is equipped only with some LED lights, test points, and connectors.

- Diagnosis interface
  The measurements and the calculation results obtained from the measuring and computing unit are exported to an output interface. In the presented application, the output interface has been materialized by an computer equipped with Labview software. With the help of Labview, the real-time cell voltage signals and the diagnosis results can be visualized on the screen. The real-time data can also be saved for advanced analysis.

4. Diagnosis approach

4.1. Selection of the variables for diagnosis

Individual cell voltages are selected as the variables for diagnosis. The selection is supported by the following factors: First, cell voltage signals are dependent synthetically on the conditions

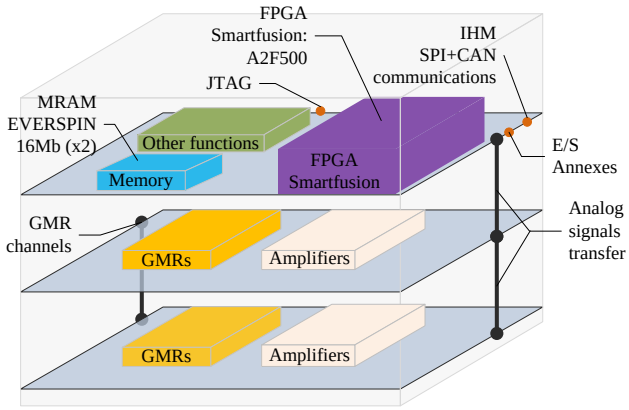


Figure 4. Embedded system designed for PEMFC system diagnosis [31]

of the fuel cells, such as the current density distribution, electrochemical characteristics, temperature, fluidic conditions, and aging effect. Voltage signals are therefore crucial for inspecting the health states of the fuel cells. Second, it is observed that the voltages of the cells located at different positions are usually different [32]. It is somehow necessary to monitor every single (or several) cell voltage(s) other than only the global stack voltage to get the knowledge of the local health states. Third, from our previous studies, it is observed that different faults lead to different magnitudes and distributions of cell voltages [33]. This characteristic is in accordance with the objective of the FDI. Moreover, the designed SiP is equipped with GMR sensors which facilitate the precise measurement of individual cell voltages.

4.2. Principle of the diagnosis approach

The principle of the proposed diagnosis approach can be summarized in Fig. 5. In the offline training process, the SVM classifier is trained based on the training dataset in which the data are collected in both normal and faulty states. In the online performing phase, the real-time data are handled using the trained SVM model. The diagnosis inference can be obtained based on the classification results and according to the diagnosis rule.

4.3. SVM classification

SVM is a classification method developed in the late 20th century [34]. It has been successfully used in a wide applications range during the last two decades [35]. Thanks to its excellent characteristics, SVM is a suitable diagnosis oriented classification tool. For instance, SVM has better generalization capability than conventional classification methods, such as artificial neural networks [35]. Concerning fault diagnosis, it is usually impossible to obtain the sufficient samples in faulty conditions. In such cases, SVM can provide more reliable results with a small number of learning samples. Regarding the training procedure of SVM, the global optimal solution can be guaranteed. In addition, the solution of SVM is usually represented using a small proportion of training samples, which makes the performing calculation light enough for real-time use [20].

The basic SVM theory comes from the binary classification problem. As Fig. 6 shows, the training samples, i.e. individual cell voltages, distributed in two classes are marked by rings

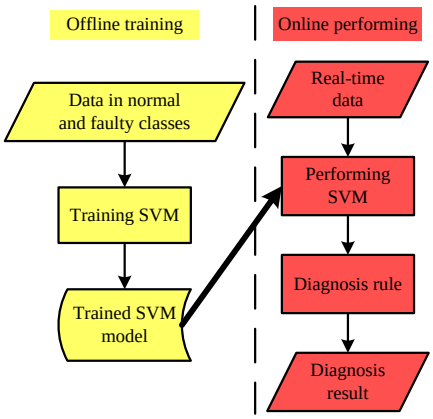


Figure 5. Flowchart of the diagnosis approach

and filled circles. The two classes can represent the normal state and a specific faulty state or two different faulty states. Notice that the space formulated by cell voltages is shown using a two-dimensional space to facilitate the visualization. Suppose we have some hyperplane which separates the two classes. SVM looks for the optimal hyperplane with the maximum distance from the nearest training samples. The samples that lie on the margin are called support vectors. In the performing phase, which class a test sample belongs to can be told according to its location.

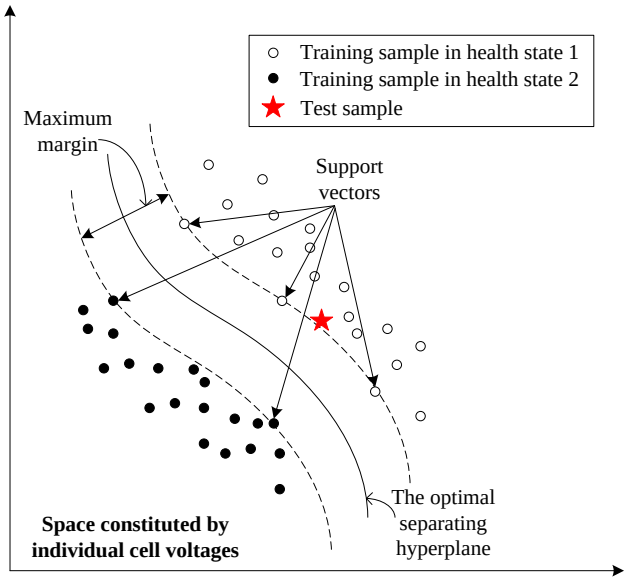


Figure 6. Schematic diagram of SVM dedicated to fault diagnosis

The training of a diagnosis oriented SVM procedure can be summarized mathematically as

follows: Given a training dataset including N H-dimensional samples $\{\boldsymbol{x}_n | n = 1, \ldots, N\}$. The training data are distributed in one normal class denoted by Ω_0 and C faulty classes denoted by $\Omega_1, \Omega_2, \ldots, \Omega_C$. The number of the elements in set Ω_i is denoted as N_i. SVM model is trained based on the training dataset, and can be represented by a function $\mathcal{F}$ of an arbitrary sample $\boldsymbol{x}$ to its class index:

$$\mathcal{F}(\boldsymbol{x}) = i, \ i \in \{0, 1, 2, \ldots C\} \tag{1}$$

For performing, a real-time sample can be classified into one of the known classes thanks to (1).

The binary SVM corresponding to the first two classes is summarized as Algorithm 1. For more details on SVM theory, the reader is referred, for instance, to [34] and references therein.

Algorithm 1 Binary SVM

Training:

1: Initialize D, σ.

2: Collect $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$ distributed in 1st and 2nd classes.

3: Solve the quadratic problem:

$$\begin{aligned} \min J(\boldsymbol{a}) &= \frac{1}{2} \sum_{\substack{n=1 \\ m=1}}^{N} a_n a_m g_n(\boldsymbol{x}) g_m(\boldsymbol{x}) k(\boldsymbol{x}_n, \boldsymbol{x}_m) - \sum_{n=1}^{N} a_n \\ \text{s.t.} \sum_{n=1}^{N} a_n g_n(\boldsymbol{x}) &= 0, \quad 0 \le a_n \le D \quad \text{for} \quad n = 1, \ldots, N \end{aligned} \tag{2}$$

with $g_n(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in \Omega_1$ and $g_n(\boldsymbol{x}) = -1$ if $\boldsymbol{x} \in \Omega_2$.
where $\boldsymbol{a} = [a_1, a_2, ..., a_N]^T$ are the Lagrange multipliers, and the Gaussian kernel function is defined as

$$k(\boldsymbol{x}_n, \boldsymbol{x}_m) = exp\left(-\frac{||\boldsymbol{x}_n - \boldsymbol{x}_m||^2}{\sigma}\right) \tag{3}$$

4: Save support vectors $\boldsymbol{x}_1^s, \boldsymbol{x}_2^s, \ldots, \boldsymbol{x}_S^s$ and corresponding g_n and a_n denoted by $\{g_n^s\}$ and $\{a_n^s\}$) for which $a_n > 0$, where S is the number of support vectors.

Performing:

For a new sample $\boldsymbol{x}$,

$$\mathcal{F}(\boldsymbol{x}) = \begin{cases} 1, \ \text{if } sign\left(\sum_{n=1}^{S} a_n^s g_n^s k(\boldsymbol{x}_n^s, \boldsymbol{x}) + b\right) = 1 \\ 2, \ \text{elsewhere} \end{cases} \tag{4}$$

where

$$b = \frac{1}{S} \sum_{j=1}^{S} \left(g_j^s - \sum_{n=1}^{S} a_n^s g_n^s k(\boldsymbol{x}_n^s, \boldsymbol{x}_j^s) \right)$$

To extend the binary classifier to multi-classification situations, there are several ways (see for instance [36] and references therein). The method "One-Against-One" has been adopted in this study. Actually, up to $C(C+1)/2$ binary SVMs can be constructed based on the training data

in $C+1$ classes. For classifying a new sample, first its classification results corresponding to all the binary SVMs are obtained. After that, the final classification result is obtained by voting all the binary classification results (see [36] for more details).

4.4. Diagnosis rule

In a general way, the classification results are used directly as the diagnosis results. However, in some practical cases, overlaps usually exist among different health states and the samples in classes can not be perfectly classified. In such cases, an additional diagnosis rule should be designed to provide more reliable and consistent diagnosis results. In this study, an additional degree of freedom is introduced to achieve this goal. The general idea is to use a sequence of classification results instead of a single one to determine the current health state. Specifically concerning the fault class i, $i \in \{1, \ldots, C\}$, at time k, the diagnosis results of the last N^i_{lag} samples (i.e. $\mathcal{F}(\boldsymbol{x}_{k-N^i_{lag}+1}), \ldots, \mathcal{F}(\boldsymbol{x}_k)$), named diagnosis window, are taken into account. Fault degree, denoted Fd, corresponding to a specific fault is defined as the rate of the fault is diagnosed:

$$Fd_i(k) = \frac{\sum_{n=k-N^i_{lag}+1}^{k} (\mathcal{F}(\boldsymbol{x}_n) == i)}{N^i_{lag}}, \ i \in \{1, \ldots, C\} \tag{5}$$

The fault degree of fault i, i.e., $Fd_i(k)$, is then calculated and compared to the pre-defined threshold denoted by Th_i. The fault occurrence at time k can be justified if the threshold is exceeded. The diagnosis rule can be expressed as Algorithm 2. Notice that diagnosis window size N^i_{lag} and threshold Th_i need to be initialized to realize this rule.

Algorithm 2 Diagnosis rule

```
for i = 1 to C do
    Collect F(x_{k-N^i_lag+1}), ..., F(x_k)
    Calculate Fd_i according to (5)
    if Fd_i ≥ Th_i then
        Fault i detected
    else
        No fault i
    end if
end for
```

5. Online implementation of the diagnosis strategy

5.1. Offline training and algorithm integration

Knowing that classification based diagnosis belongs to supervised learning methods, the data from various classes were needed for training. To prepare the training dataset, several faults were produced deliberately. Table 2 summarizes the operations in the experiment of training data preparation. Notice that all the concerned faults are recoverable ones caused by faulty operations. The faults consist of the ones of gas supply subsystems and those on water management. Constrained by the measuring capability of the initially designed SiP and configuration of connectors between the stack and SiP, the voltages of 14 cells can be measured and used as the

variables for diagnosis. These cells are numbered by 1, 3, 5, 6, 7, 9, 10, 11, 12, 18, 19, 20, 21, 22, counting from the negative pole of the stack.

The evolutions of individual cell voltages of the training data are shown in Fig. 7(a). The cell voltage details in a period of 100 s are shown in Fig. 7(b) (sample time 1 s). It can be observed that the magnitudes and behaviors of the voltages vary between different cells. Besides, a cycle property can be observed from the curves. Actually, it is related to the process of anode purge whose periodic time is 90 s.

Table 2. Experimental procedure for the preparation of the training dataset

Starting time	Ending time	Operation	Health state
0	879	Nominal condition	Normal state (Nl)
880	1675	Pressure of 1.3 bar at each side	Low pressure fault (F1)
1676	2618	Back to nominal condition	Normal state (Nl)
2619	3499	Pressure of 1.7 bar at each side	High pressure fault (F2)
3500	4892	Back to nominal condition	Normal state (Nl)
4893	6288	Lower relative humidity	Drying fault (F3)
6289	7518	Back to nominal condition	Normal state (Nl)
7519	8287	St. Air 1.5	Low air stoichiometry fault (F4)
8288	8955	Back to nominal condition	Normal state (Nl)

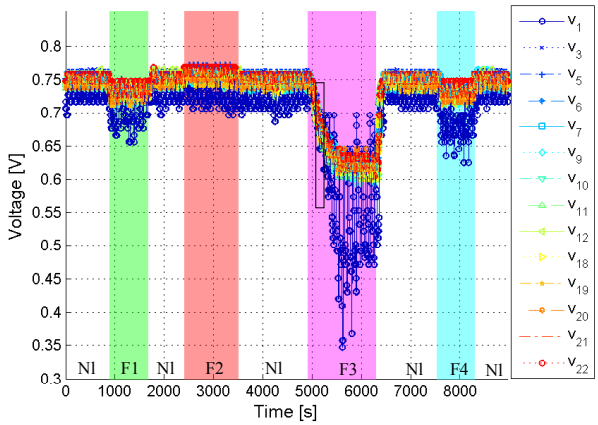
With the training dataset, the SVM classifier is trained. Parameters D and σ in Algorithm 1 were initialized respectively as $D = 19000$ and $\sigma = 2000$ according to cross validation [37].

Classifying the training data with the trained SVM, the global classification accuracy rate is 84.98%. More detailed results can be summarized as a confusion matrix, shown quantitatively in Table 3 and visually in Fig. 8. It could be observed that the false alarm rate (FAR), i.e., the rate of the samples in normal state wrongly diagnosed into the fault classes, is relatively low. The diagnosis accuracy for the data in F3 is also high. A considerable part of data in classes F1, F2, and F4 are wrongly classified into the normal class. From Fig. 8, it can be seen the classification results vibrate between the corresponding faults and normal classes. Actually, faults F1, F2, and F4 are relatively light ones compared to the faults such as F3. In these states, the data vary lightly from the normal state. The overlaps between the normal and faulty states exist. It can also be observed that some overlaps exist between class F1 and F4.

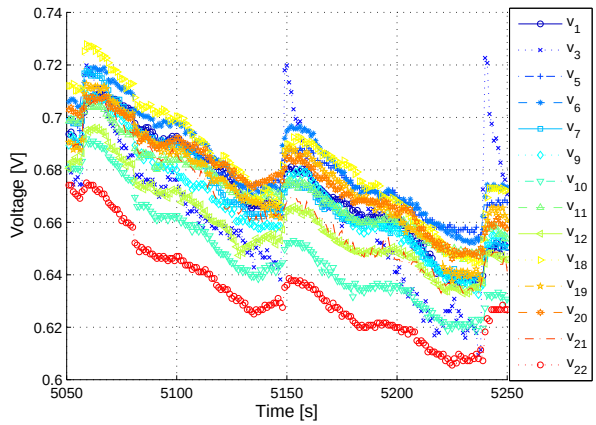
Table 3. Confusion matrix for classification results of training data

		Predicted classes				
		Nl	F1	F2	F3	F4
Actual classes	Nl	4762	29	25	68	11
	F1	406	373	0	0	17
	F2	300	0	800	0	0
	F3	157	11	0	1228	0
	F4	239	82	0	0	448

To improve the robustness and accuracy performance, the diagnosis rule is designed based on the classification results (refer Section 4.4). Parameters N^i_{lag} and Th_i were initialized based on



(a) Evolution in different health states



(b) Details in a short period

Figure 7. Cell voltages in training dataset

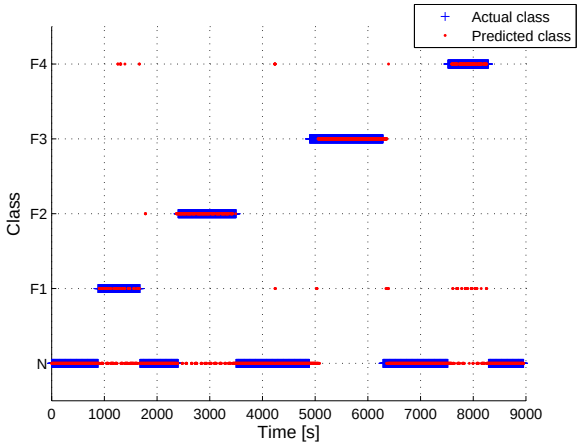


Figure 8. Classification results of the training data

the classification results and are shown in Table 4. Notice that the procedure for F3 is equivalent to use the original classification results.

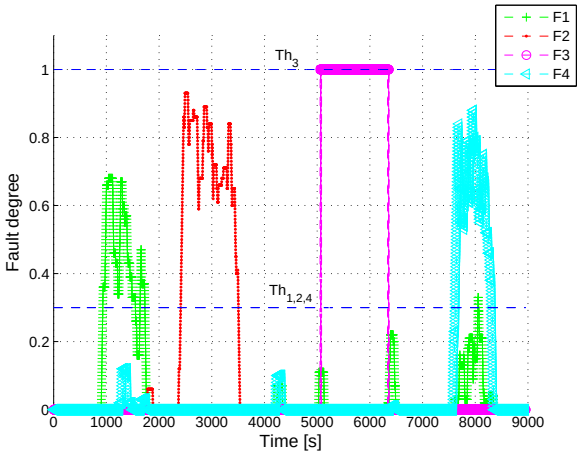
The fault degrees and diagnosis results corresponding to the experiment of training data preparation are shown in Fig. 9(a) and Fig. 9(b). The global diagnosis accuracy rate reaches 93.20%, which is significantly increased from the original classification result. The detailed diagnosis results are also summarized quantitatively as a confusion matrix in Table 5. Comparing Table 3 and 5, it could be seen that the diagnosis results corresponding to F1, F2, and F4 are more accurate and consistent than the raw classification results. The phenomenon can also be observed visually by comparing Fig. 9(b) and 8. Concerning the data in normal state, the FAR is slightly increased.

Table 4. Parameter initialization for the design of diagnosis rule

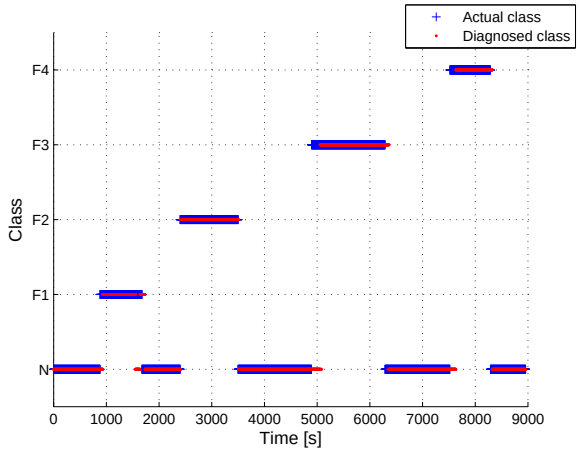
	F1	F2	F3	F4
N^i_{lag}	100	100	1	100
Th_i	0.3	0.3	1	0.3

Table 5. Confusion matrix of diagnosis results of training data

		Predicted classes				
		Nl	F1	F2	F3	F4
Actual classes	Nl	4729	57	7	68	34
	F1	138	658	0	0	0
	F2	10	0	1090	0	0
	F3	168	0	0	1228	0
	F4	114	13	0	0	642



(a) Fault degrees of the four concerned faults



(b) Diagnosis results

Figure 9. Fault degrees and diagnosis results corresponding to the training data

The diagnosis procedure tested by using a computer or by coding into the memory of SiP, respectively, provides 100% accordant results. Besides, the diagnosis algorithm could be calculated within a sampling cycle (i.e., 1s) using the SiP. That is to say, the diagnosis approach is successfully integrated into the SiP.

5.2. Online validation

To realize online validation, the programed SiP was tested online with the real-time data during another experiment. The operations during this experiment are summarized in Table 6, and the measured cell voltages, i.e., the variables for diagnosis are plotted in Fig. 10.

Table 6. Experimental procedure for online validation

Starting time	Ending time	Operation	Health state
0	3660	Nominal condition	Normal state (Nl)
3661	4543	Pressure of 1.3 bar at each side	Low pressure fault (F1)
4544	5374	Back to nominal condition	Normal state (Nl)
5375	6541	Pressure of 1.7 bar at each side	High pressure fault (F2)
6542	8128	Back to nominal condition	Normal state (Nl)
8129	8909	St. Air 1.5	Low air stoichiometry fault (F4)
8910	9841	Back to nominal condition	Normal state (Nl)
9842	11909	Lower relative humidity	Drying fault (F3)
11910	12459	Back to nominal condition	Normal state (Nl)

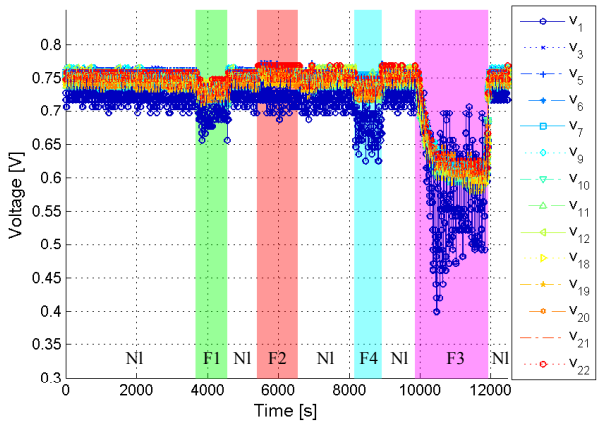
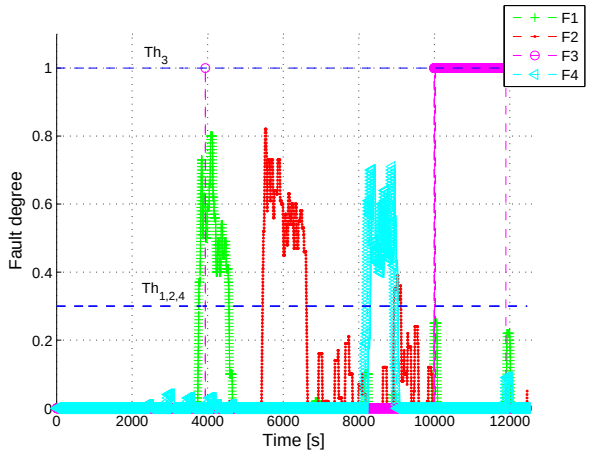
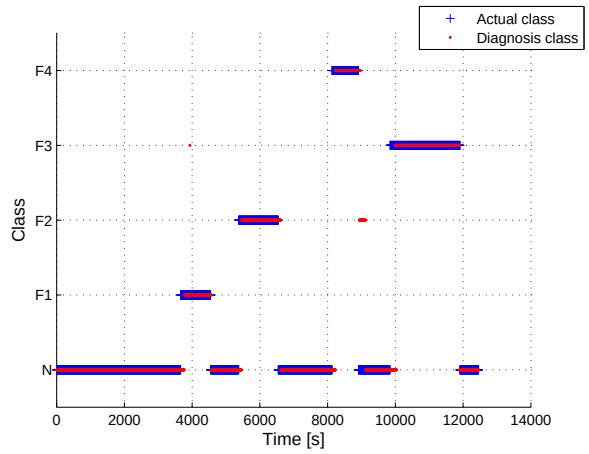


Figure 10. Cell voltages measured during online validation

The fault degrees and diagnosis results corresponding to the experiment for online validation are shown in Fig. 11(a) and Fig. 11(b). The global diagnosis accuracy of the online implementation is 93.99%. Notice that the global diagnosis accuracy related to online validation is even a little higher than that for training data. The main reason is that the proportion of normal data in



(a) Fault degrees of the four concerned faults



(b) Diagnosis results

Figure 11. Fault degrees and diagnosis results corresponding to the online validation

Table 7. Confusion matrix of diagnosis results corresponding to online validation

		Diagnosed classes				
		Nl	F1	F2	F3	F4
Actual classes	Nl	7266	19	250	0	28
	F1	101	782	0	0	0
	F2	77	0	1090	0	0
	F3	173	0	0	1890	0
	F4	101	0	0	0	680

the validation dataset is bigger than that of training data. Similarly, the online diagnosis results are also summarized quantitatively as the confusion matrix shown in Table 7.

To show the effect of the proposed diagnosis rule, the raw classification results are also provided here. The global classification accuracy of online validation data is 85.93%. The detailed classification results corresponding to online validation are summarized in Table 8. By comparing Table 8 and 7, it can be seen that the diagnosis accuracy for the data in F1, F2 and F4 classes can be increased from the raw classification results by using the proposed diagnosis rule. However, the FAR is also a little increased.

Table 8. Confusion matrix of classification results corresponding to online validation

		Diagnosed classes				
		Nl	F1	F2	F3	F4
Actual classes	Nl	7302	31	220	0	10
	F1	446	432	0	1	4
	F2	477	0	690	0	0
	F3	132	28	5	1890	8
	F4	363	14	14	0	390

5.3. Discussion

5.3.1. Delays caused by the diagnosis rule

Although the robustness performance and accuracy are improved by using the proposed diagnosis rule, it should be mentioned that some delays are also introduced. These delays are the main factors which cause the diagnosis error. As Fig. 12 shows, when the diagnosis rule is launched, to formulate the first diagnosis window for the detection of fault i ($i = 1, \ldots, C$), the *initial delay* which is of the length N^i_{lag} is introduced. When fault i occurs, the fault can be diagnosed after $N^i_{lag} Th_i$ sample periods in ideal case. The data during the *diagnosis delay* are wrongly diagnosed into the normal state. When fault i is eliminated, the diagnosis result of the corresponding fault will continue $N^i_{lag}(1-Th_i)$ more sample periods. The data related to *recovery delay* are wrongly diagnosed into the fault state and lead to the increase of FAR.

Notice that the three delays are all determined by diagnosis window size N^i_{lag}. To shorten the delays, N^i_{lag} must be initialized with a small value. On the contrary, a relatively big N^i_{lag} is usually required to improve the robustness and accuracy performance. Hence, a compromise needs to be made to parameterize N^i_{lag}.

5.3.2. Extendability of the approach

Here, four different faults are concerned in the study. It should be noticed that other types of faults than the given ones could also be encountered. For instance, the fault of catalyst poisoning which is usually caused by the CO mixed in the entered hydrogen [38, 39], the faults related to the temperature subsystem, and the faults occurs at the electric circuit [20]. The discriminative information contained in the individual cell voltage signals is the key factor that determines whether a fault can be efficiently detected and isolated from other faults. Actually, in our previous study [20], it has been demonstrated that the CO catalyst poisoning fault, the fault related to temperature management, and the fault in electric circuit fault can be accurately diagnosed thanks to our diagnostic approach.

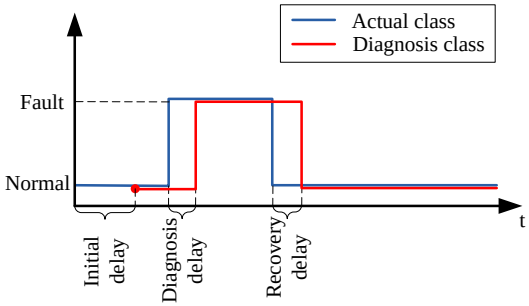


Figure 12. Delays introduced by diagnosis rule

Nevertheless, it should also be emphasized that the data corresponding to the concerned faults must be collected in prior to implement an efficient diagnosis. It is usually difficult or impossible to carry out the experiments in all the faulty cases. This could be considered as the main drawback of all the data based approaches. To alleviate this defect, we proposed previously a procedure which enables the recognition of a novel fault even the corresponding data have not appeared in the training phase [33].

5.3.3. Ageing effect

The performance degradations involving a PEMFC system also result from the ageing effect other than the regular faults. When the ageing effect is taken into account, the diagnosis results could be impacted. For instance, the data in normal state are not stationary. Using the originally developed diagnosis approach, the normal data might eventually be diagnosed as the ones in faulty states as time goes on. In order to maintain the performance, a self-adaptation method is proposed for the diagnosis in our previous work [23].

Concerning the ageing effect, several studies have been launched and focused on the strategy of prognosis which is dedicated to describing the degradation tendency related to ageing effect and predicting the residual useful life far ahead [40, 41, 42].

6. Conclusion

In this paper, a classification based diagnosis strategy is proposed for PEMFC systems and the results of online realization are demonstrated. According to the obtained results, the following conclusions can be made:

1. A satisfying online diagnosis results can be obtained with individual cell voltages serving as the variables for diagnosis.
2. In the diagnostic approach, SVM classification method and the designed diagnosis rule are performed successively in the diagnostic process. The efficiency of the approach is validated in the perspectives of diagnosis accuracy and online implementability. More especially, it is demonstrated that, by post-processing raw classification results using the proposed diagnosis rule, more robust and accurate diagnosis results can be obtained.
3. The specially designed SiP can fulfill the tasks of precisely measuring individual cell voltages and implementing the diagnostic approach online. Thanks to the compact design, it is promising to be used in practical applications like FCEVs.

Acknowledgment

This work is a contribution to the ANR DIAPASON2 project (fuel cell diagnosis methods for vehicle and stationary applications 2nd phase). The authors would like to thank their partners for their contribution.

References

[1] J. Park, H. Oh, T. Ha, Y. I. Lee, K. Min, A review of the gas diffusion layer in proton exchange membrane fuel cells: Durability and degradation , Applied Energy 155 (2015) 866 – 880. doi:http://dx.doi.org/10.1016/j.apenergy.2015.06.068.

[2] H. Sun, C. Xie, H. Chen, S. Almheiri, A numerical study on the effects of temperature and mass transfer in high temperature PEM fuel cells with ab-PBI membrane , Applied Energy 160 (2015) 937 – 944. doi:http://dx.doi.org/10.1016/j.apenergy.2015.02.053.

[3] J.-H. Jang, W.-M. Yan, H.-C. Chiu, J.-Y. Lui, Dynamic cell performance of kW-grade proton exchange membrane fuel cell stack with dead-ended anode , Applied Energy 142 (2015) 108 – 114. doi:http://dx.doi.org/10.1016/j.apenergy.2014.12.073.

[4] C. Raga, A. Barrado, A. Lazaro, C. Fernandez, V. Valdivia, I. Quesada, L. Gauchia, Black-box model, identification technique and frequency analysis for pem fuel cell with overshooted transient response, IEEE Trans. Power Electron. 29 (10) (2014) 5334–5346. doi:10.1109/TPEL.2013.2292599.

[5] A. Simons, C. Bauer, A life-cycle perspective on automotive fuel cells, Applied Energy 157 (2015) 884 – 896. doi:http://dx.doi.org/10.1016/j.apenergy.2015.02.049.

[6] P. Xuewei, A. Rathore, Novel bidirectional snubberless naturally commutated soft-switching current-fed full-bridge isolated dc/dc converter for fuel cell vehicles, IEEE Trans. Ind. Electron. 61 (5) (2014) 2307–2315. doi:10.1109/TIE.2013.2271599.

[7] J. Wu, X. Z. Yuan, J. J. Martin, H. Wang, J. Zhang, J. Shen, S. Wu, W. Merida, A review of PEM fuel cell durability: Degradation mechanisms and mitigation strategies, Journal of Power Sources 184 (1) (2008) 104 – 119. doi:http://dx.doi.org/10.1016/j.jpowsour.2008.06.006.

[8] H. Chen, P. Pei, M. Song, Lifetime prediction and the economic lifetime of Proton Exchange Membrane fuel cells , Applied Energy 142 (2015) 154 – 163. doi:http://dx.doi.org/10.1016/j.apenergy.2014.12.062.

[9] B. Cai, Y. Liu, Q. Fan, Y. Zhang, Z. Liu, S. Yu, R. Ji, Multi-source information fusion based fault diagnosis of ground-source heat pump using Bayesian network, Applied Energy 114 (2014) 1 – 9. doi:http://dx.doi.org/10.1016/j.apenergy.2013.09.043.

[10] S. Yin, S. Ding, X. Xie, H. Luo, A Review on Basic Data-Driven Approaches for Industrial Process Monitoring, IEEE Trans. Ind. Electron. 61 (11) (2014) 6418–6428. doi:10.1109/TIE.2014.2301773.

[11] Q. Liu, S. Qin, T. Chai, Multiblock Concurrent PLS for Decentralized Monitoring of Continuous Annealing Processes, IEEE Trans. Ind. Electron. 61 (11) (2014) 6429–6437. doi:10.1109/TIE.2014.2303781.

[12] H. He, Z. Liu, Y. Hua, Adaptive Extended Kalman Filter Based Fault Detection and Isolation for a Lithium-Ion Battery Pack, Energy Procedia 75 (2015) 1950 – 1955, clean, Efficient and Affordable Energy for a Sustainable Future: The 7th International Conference on Applied Energy (ICAE2015). doi:http://dx.doi.org/10.1016/j.egypro.2015.07.230.

[13] R. Petrone, Z. Zheng, D. Hissel, M. Péra, C. Pianese, M. Sorrentino, M. Becherif, N. Yousfi-Steiner, A review on model-based diagnosis methodologies for PEMFCs, International Journal of Hydrogen Energy 38 (17) (2013) 7077–7091. doi:10.1016/j.ijhydene.2013.03.106.

[14] Z. Zheng, R. Petrone, M. Péra, D. Hissel, M. Becherif, C. Pianese, N. Yousfi Steiner, M. Sorrentino, A review on non-model based diagnosis methodologies for PEM fuel cell stacks and systems, International Journal of Hydrogen Energy 38 (21) (2013) 8914–8926. doi:10.1016/j.ijhydene.2013.04.007.

[15] Z. Li, Data-driven fault diagnosis for pemfc systems, Ph.D. thesis, Aix-Marseille University (2014).

[16] T. Escobet, D. Feroldi, S. de Lira, V. Puig, J. Quevedo, J. Riera, M. Serra, Model-based fault diagnosis in PEM fuel cell systems, Journal of Power Sources 192 (1) (2009) 216–223.

[17] S. de Lira, V. Puig, J. Quevedo, A. Husar, LPV observer design for PEM fuel cell system: Application to fault detection, Journal of Power Sources 196 (9) (2011) 4298 – 4305, CONAPPICE 2010. doi:http://dx.doi.org/10.1016/j.jpowsour.2010.11.084.

[18] D. Benouioua, D. Candusso, F. Harel, L. Oukhellou, Fuel cell diagnosis method based on multifractal analysis of stack voltage signal, International Journal of Hydrogen Energy 39 (5) (2014) 2236 – 2245. doi:http://dx.doi.org/10.1016/j.ijhydene.2013.11.066.

[19] J. Kim, I. Lee, Y. Tak, B. Cho, State-of-health diagnosis based on hamming neural network using output voltage pattern recognition for a PEM fuel cell, International Journal of Hydrogen Energy 37 (5) (2012) 4280–4289. doi:10.1016/j.ijhydene.2011.11.092.

[20] Z. Li, R. Outbib, S. Giurgea, D. Hissel, Y. Li, Fault detection and isolation for Polymer Electrolyte Membrane Fuel Cell systems by analyzing cell voltage generated space, Applied Energy 148 (2015) 260 – 272. doi:http://dx.doi.org/10.1016/j.apenergy.2015.03.076.

[21] Z. Zheng, M.-C. Péra, D. Hissel, M. Becherif, K.-S. Agbli, Y. Li, A double-fuzzy diagnostic methodology dedicated to online fault diagnosis of proton exchange membrane fuel cell stacks, Journal of Power Sources 271 (0) (2014) 570 – 581. doi:http://dx.doi.org/10.1016/j.jpowsour.2014.07.157.

[22] J. Hua, J. Li, M. Ouyang, L. Lu, L. Xu, Proton exchange membrane fuel cell system diagnosis based on the multivariate statistical method, International Journal of Hydrogen Energy (2011) 1–10doi:10.1016/j.ijhydene.2011.05.075.

[23] Z. Li, R. Outbib, S. Giurgea, D. Hissel, Diagnosis for pemfc systems: A data-driven approach with the capabilities of online adaptation and novel fault detection, IEEE Trans. Ind. Electron. 62 (8) (2015) 5164–5174. doi:10.1109/TIE.2015.2418324.

[24] T. Rauber, F. de Assis Boldt, F. Varejao, Heterogeneous Feature Models and Feature Selection Applied to Bearing Fault Diagnosis, IEEE Trans. Ind. Electron. 62 (1) (2015) 637–646. doi:10.1109/TIE.2014.2327589.

[25] M. Amar, I. Gondal, C. Wilson, Vibration Spectrum Imaging: A Novel Bearing Fault Classification Approach, IEEE Trans. Ind. Electron. 62 (1) (2015) 494–502. doi:10.1109/TIE.2014.2327555.

[26] D. You, X. Gao, S. Katayama, WPD-PCA-Based Laser Welding Process Monitoring and Defects Diagnosis by Using FNN and SVM, IEEE Trans. Ind. Electron. 62 (1) (2015) 628–636. doi:10.1109/TIE.2014.2319216.

[27] C. Chan, J. Jiang, G. Chen, X. Wang, K. Chau, A novel polyphase multipole square-wave permanent magnet motor drive for electric vehicles, IEEE Trans. Ind. Appl. 30 (5) (1994) 1258–1266. doi:10.1109/28.315237.

[28] S. Kang, K. Min, Dynamic simulation of a fuel cell hybrid vehicle during the federal test procedure-75 driving cycle, Applied Energy 161 (2016) 181 – 196. doi:http://dx.doi.org/10.1016/j.apenergy.2015.09.093.

[29] Website: Smartfusion introduction.
URL http://www.microsemi.com/products/fpga-soc/soc-fpga/smartfusion

[30] F. Rothan, C. Condemine, B. Delaet, O. Redon, a. Giraud, A low power 16-channel fully integrated GMR-based current sensor, European Solid-State Circuits Conference (2012) 245–248doi:10.1109/ESSCIRC.2012.6341304.

[31] A. GIRAUD, Embedded smart sensors for measurement and diagnosis of measurement and diagnosis of multicells PEMFC, report - project DIAPASON (June 2012).

[32] P. Rodatz, F. Büchi, C. Onder, L. Guzzella, Operational aspects of a large PEFC stack under practical conditions, Journal of Power Sources 128 (2) (2004) 208–217. doi:10.1016/j.jpowsour.2003.09.060.

[33] Z. Li, S. Giurgea, R. Outbib, D. Hissel, Fault diagnosis and novel fault type detection for pemfc system based on spherical-shaped multiple-class support vector machine, in: Advanced Intelligent Mechatronics (AIM), 2014 IEEE/ASME International Conference on, 2014, pp. 1628–1633. doi:10.1109/AIM.2014.6878317.

[34] J. C. Platt, Sequential Minimal Optimization : A Fast Algorithm for Training Support Vector Machines, Technical Report MSR-TR-98-14, Microsoft Research (1998) 1–21.

[35] A. Widodo, B.-s. Yang, Support vector machine in machine condition monitoring and fault diagnosis, Mechanical Systems and Signal Processing 21 (2007) 2560–2574. doi:10.1016/j.ymssp.2006.12.007.

[36] C.-w. Hsu, C.-j. Lin, A Comparison of Methods for Multiclass Support Vector Machines, IEEE Trans. Neural Netw. 13 (2) (2002) 415–425.

[37] G. C. Cawley, N. L. Talbot, Fast exact leave-one-out cross-validation of sparse least-squares support vector machines, Neural Networks 17 (10) (2004) 1467 – 1475. doi:http://dx.doi.org/10.1016/j.neunet.2004.07.002.

[38] N. Wagner, E. Glzow, Change of electrochemical impedance spectra (EIS) with time during CO-poisoning of the Pt-anode in a membrane fuel cell , Journal of Power Sources 127 (12) (2004) 341 – 347, eighth Ulmer Electrochemische Tage. doi:http://dx.doi.org/10.1016/j.jpowsour.2003.09.031.

[39] C. Farrell, C. Gardner, M. Ternan, Experimental and modelling studies of CO poisoning in PEM fuel cells, Journal of Power Sources 171 (2) (2007) 282 – 293. doi:http://dx.doi.org/10.1016/j.jpowsour.2007.07.006.

[40] X. Zhang, P. Pisu, An unscented kalman filter based approach for the health-monitoring and prognostics of a polymer electrolyte membrane fuel cell, in: Annual Conference of the Prognostics and Health Management Society 2012, Vol. 3, 2012, p. 9.

[41] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, N. Zerhouni, Prognostics and Health Management of PEMFC State of the art and remaining challenges, International Journal of Hydrogen Energy 38 (35) (2013) 15307 – 15317. doi:http://dx.doi.org/10.1016/j.ijhydene.2013.09.051.

[42] R. Petrone, D. Hissel, M. Péra, D. Chamagne, R. Gouriveau, Accelerated stress test procedures for PEM fuel cells

under actual load constraints: State-of-art and proposals, International Journal of Hydrogen Energy 40 (36) (2015) 12489 – 12505. doi:http://dx.doi.org/10.1016/j.ijhydene.2015.07.026.