



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-01846859>

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

^aFCLAB (Fuel Cell Lab) Research Federation, FR CNRS 3539, rue Thierry Mieg, 90010 Belfort Cedex, France

^bFEMTO-ST (UMR CNRS 6174), ENERGY Department, UFC/UTBM/ENSMM, France

^cLaboratoire des Sciences de l'Information et des Systemes (LSIS), University of Aix-Marseille, France

^dCEA/LIST, 91191 Gif-sur-Yvette Cedex, France

^eCEA/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-compact 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

1 The environment and resource issues have been drawing increasing attention of the world.
2 Global warming and non-renewable resource exhaustion are two problems need to be addressed
3 urgently. One of the main causes of these issues is the high dependence of the fossil fuels in
4 the current energy structure. Since hydrogen can be produced from diverse sources, such as re-
5 newable resources and nature gas, large-scale use of hydrogen-based fuel cells is considered as
6 one of the most significant solutions dedicated to slacking the dependence on fossil fuels [1, 2].
7 Amongst various fuel cells, Polymer Electrolyte Membrane Fuel Cell (PEMFC) is potentially
8 beneficial for a wide range of applications, thanks to its attractive advantages, such as high effi-
9 ciency, high power density, in-situ zero-emission, low operating temperature, and quick response
10 to load changes [3, 4]. Especially, fuel cell electric vehicle (FCEV), which is equipped with
11 PEMFC stack as the main energy generator, is a powerful competitor in the future automobile
12

*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)

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

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

27 Several fault diagnosis strategies have been studied during the last decade [16, 17, 18, 19,
28 20, 21, 22, 23]. The general model based fault diagnosis theoretical base seems to be well estab-
29 lished and some positive results have been obtained by using these methods for some PEMFC
30 systems [16, 17]. Nevertheless, building a model with first principle is not a trivial task. The in-
31 ternal parameters, which are essential for modeling, are not evident to be found or estimated. In
32 addition, model structures and parameters may differ among different designs of fuel cell stacks
33 and other system components. Apart from the model based diagnosis, the application of data
34 based methodologies for the diagnosis of PEMFC systems has been drawing the attention of
35 researchers [18, 19, 20, 21, 22, 23]. Avoiding the sophisticated modeling process, the data based
36 diagnosis seems to be more practical in most cases. Actually, the data based diagnosis has been
37 utilized in a number of industrial processes [24, 25, 26].

38 Within the scope of data based fault diagnosis, a number of pattern classification tech-
39 niques have been widely used since FDI can be considered as a classification problem [25, 26].
40 Some classification based diagnosis strategies have been proposed for PEMFC systems (see
41 [18, 20, 21, 23] for instance). Different variables, feature extraction and classification meth-
42 ods have been studied using the historical data. In [18], the classification was supposed to be
43 carried out in the feature space which is generated using multifractal analysis on stack volt-
44 age. In [21], the fuzzy classification method was utilized to analyze Electrochemical Impedance
45 Spectroscopy. In our previous study [20], selecting individual cell voltages as the variables for
46 diagnosis, several classification and feature extraction techniques were compared from the per-
47 spectives of classification accuracy and computational complexity. The classification method
48 Support Vector Machine (SVM) was selected as the most suitable classification tool in the case.
49 The strategy was further developed by extending the capabilities of novel fault detection and
50 online adaptation [23].

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

60 industrial applications.

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

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

75 2. Development process of classification based online fault diagnosis

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

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

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

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

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

97 3. Diagnosis strategy development platform

98 As Fig. 2 shows, the development platform dedicated to online implementation of the diag-
99 nosis strategy consists of the following parts:

- 100 • PEMFC system

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

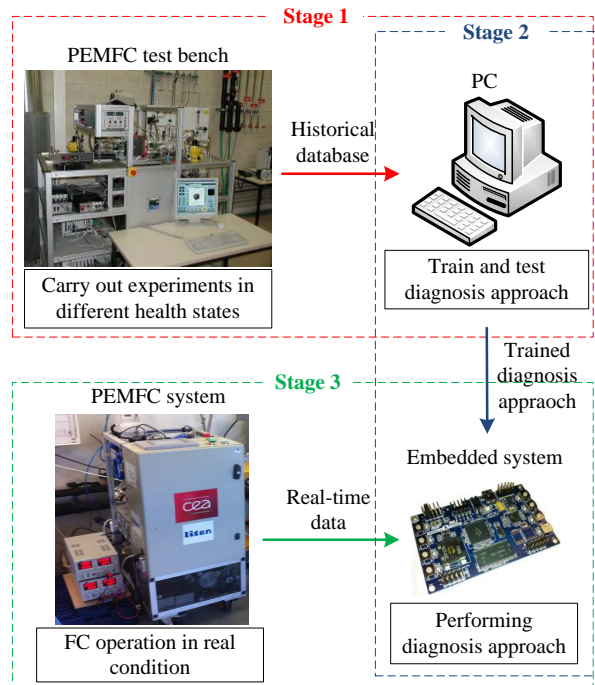


Figure 1. Developing process of classification based online fault diagnosis strategy

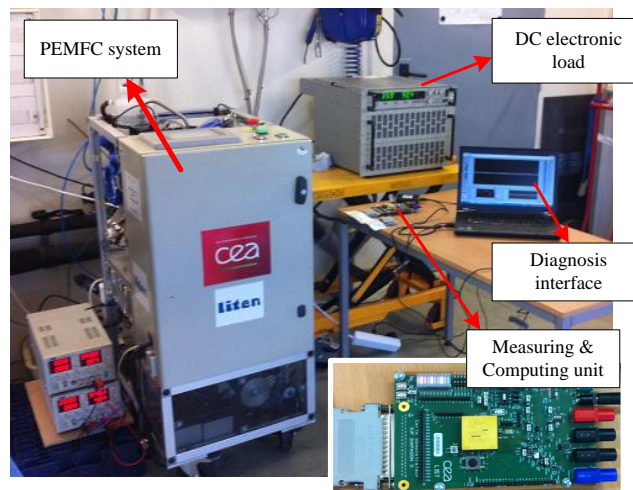


Figure 2. Overview of the development platform

102 a 64-cell stack which was fabricated by the French research organization CEA¹ specially
 103 for automotive application. The nominal operating conditions of the stack are summarized
 104 in Table 1. In normal or fault free state, all the operating parameters should be maintained
 105 at the nominal values with small variation caused by system noises. It is considered in our
 106 case that a fault in the PEMFC system occurs when the operating parameters are out of the
 107 normal range. Notice that the faults defined in our case are not limited to the fuel cell stack
 108 but cover the health states of the whole system. Certain faults, such as air pressure exceeds
 109 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 <i>Air</i>	2
Pressure at H_2 inlet	150 kPa
Pressure at <i>Air</i> 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

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

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

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

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

- 122 • DC load
 123 The load current can be flexibly varied through an electronic load.
- 124 • Measuring and computing unit
 125 The measuring and computing unit is devoted to measuring the variables for diagnosis and
 126 performing online the diagnosis approach. The core component is the specially-designed
 127 SiP chip (yellow square component shown in Fig. 2).

¹CEA: Alternative Energies and Atomic Energy Commission

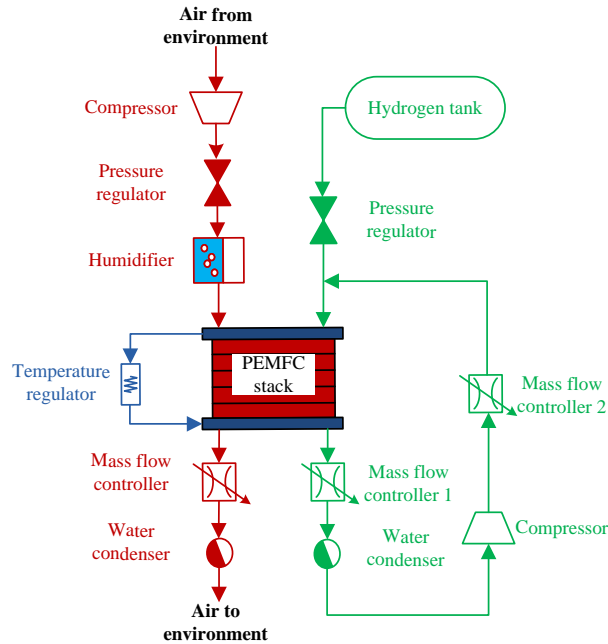


Figure 3. Schematic of the PEMFC system

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

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

139 • Diagnosis interface

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

145 **4. Diagnosis approach**

146 *4.1. Selection of the variables for diagnosis*

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

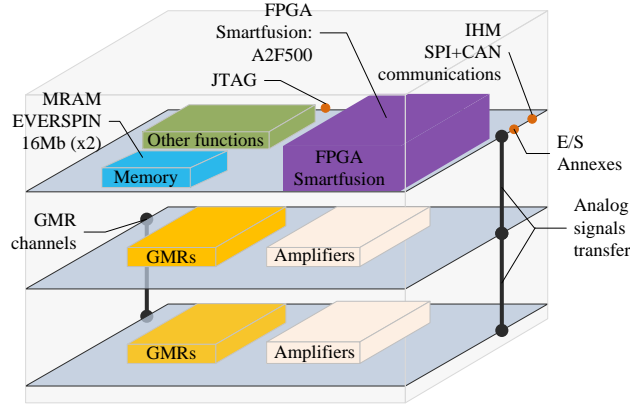


Figure 4. Embedded system designed for PEMFC system diagnosis [31]

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

158 4.2. Principle of the diagnosis approach

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

164 4.3. SVM classification

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

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

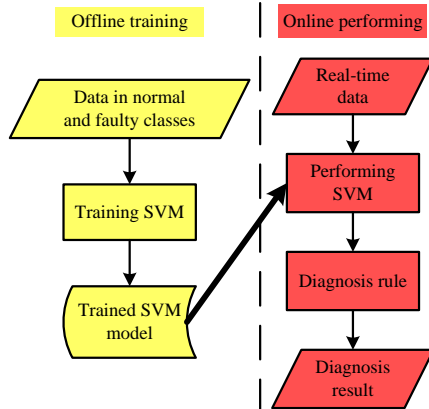


Figure 5. Flowchart of the diagnosis approach

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

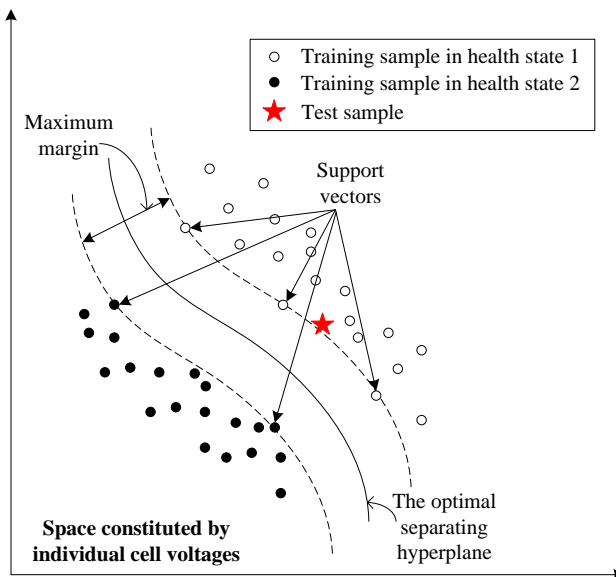


Figure 6. Schematic diagram of SVM dedicated to fault diagnosis

183

184

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

185 follows: Given a training dataset including N H -dimensional samples $\{\mathbf{x}_n | n = 1, \dots, N\}$. The
 186 training data are distributed in one normal class denoted by Ω_0 and C faulty classes denoted by
 187 $\Omega_1, \Omega_2, \dots, \Omega_C$. The number of the elements in set Ω_i is denoted as N_i . SVM model is trained
 188 based on the training dataset, and can be represented by a function \mathcal{F} of an arbitrary sample \mathbf{x} to
 189 its class index:

$$\mathcal{F}(\mathbf{x}) = i, \quad i \in \{0, 1, 2, \dots, C\} \quad (1)$$

190 For performing, a real-time sample can be classified into one of the known classes thanks to (1).
 191 The binary SVM corresponding to the first two classes is summarized as Algorithm 1. For
 192 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 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ distributed in 1st and 2nd classes.
- 3: Solve the quadratic problem:

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

with $g_n(\mathbf{x}) = 1$ if $\mathbf{x} \in \Omega_1$ and $g_n(\mathbf{x}) = -1$ if $\mathbf{x} \in \Omega_2$.

where $\mathbf{a} = [a_1, a_2, \dots, a_N]^T$ are the Lagrange multipliers, and the Gaussian kernel function is defined as

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

- 4: Save support vectors $\mathbf{x}_1^s, \mathbf{x}_2^s, \dots, \mathbf{x}_S^s$ 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 \mathbf{x} ,

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

where

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

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

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

199 4.4. Diagnosis rule

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

$$Fd_i(k) = \frac{\sum_{n=k-N_{lag}^i+1}^k (\mathcal{F}(\mathbf{x}_n) == i)}{N_{lag}^i}, \quad i \in \{1, \dots, C\} \quad (5)$$

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

Algorithm 2 Diagnosis rule

```

1: for  $i = 1$  to  $C$  do
2:   Collect  $\mathcal{F}(\mathbf{x}_{k-N_{lag}^i+1}), \dots, \mathcal{F}(\mathbf{x}_k)$ 
3:   Calculate  $Fd_i$  according to (5)
4:   if  $Fd_i \geq Th_i$  then
5:     Fault  $i$  detected
6:   else
7:     No fault  $i$ 
8:   end if
9: end for

```

213 5. Online implementation of the diagnosis strategy

214 5.1. Offline training and algorithm integration

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

222 variables for diagnosis. These cells are numbered by 1, 3, 5, 6, 7, 9, 10, 11, 12, 18, 19, 20, 21,
 223 22, counting from the negative pole of the stack.

224 The evolutions of individual cell voltages of the training data are shown in Fig. 7(a). The cell
 225 voltage details in a period of 100 s are shown in Fig. 7(b) (sample time 1 s). It can be observed
 226 that the magnitudes and behaviors of the voltages vary between different cells. Besides, a cycle
 227 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 (NI)
880	1675	Pressure of 1.3 bar at each side	Low pressure fault (F1)
1676	2618	Back to nominal condition	Normal state (NI)
2619	3499	Pressure of 1.7 bar at each side	High pressure fault (F2)
3500	4892	Back to nominal condition	Normal state (NI)
4893	6288	Lower relative humidity	Drying fault (F3)
6289	7518	Back to nominal condition	Normal state (NI)
7519	8287	St. Air 1.5	Low air stoichiometry fault (F4)
8288	8955	Back to nominal condition	Normal state (NI)

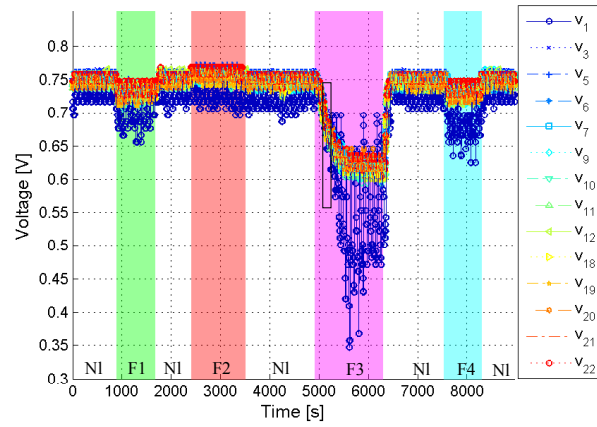
228 With the training dataset, the SVM classifier is trained. Parameters D and σ in Algorithm 1
 229 were initialized respectively as $D = 19000$ and $\sigma = 2000$ according to cross validation [37].

230 Classifying the training data with the trained SVM, the global classification accuracy rate is
 231 84.98%. More detailed results can be summarized as a confusion matrix, shown quantitatively in
 232 Table 3 and visually in Fig. 8. It could be observed that the false alarm rate (FAR), i.e., the rate
 233 of the samples in normal state wrongly diagnosed into the fault classes, is relatively low. The
 234 diagnosis accuracy for the data in F3 is also high. A considerable part of data in classes F1, F2,
 235 and F4 are wrongly classified into the normal class. From Fig. 8, it can be seen the classification
 236 results vibrate between the corresponding faults and normal classes. Actually, faults F1, F2, and
 237 F4 are relatively light ones compared to the faults such as F3. In these states, the data vary lightly
 238 from the normal state. The overlaps between the normal and faulty states exist. It can also be
 239 observed that some overlaps exist between class F1 and F4.
 240

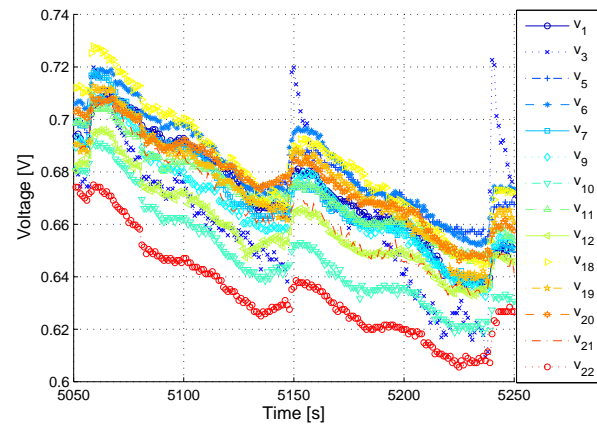
Table 3. Confusion matrix for classification results of training data

		Predicted classes				
		NI	F1	F2	F3	F4
Actual classes	NI	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

241 To improve the robustness and accuracy performance, the diagnosis rule is designed based
 242 on the classification results (refer Section 4.4). Parameters N_{lag}^i 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

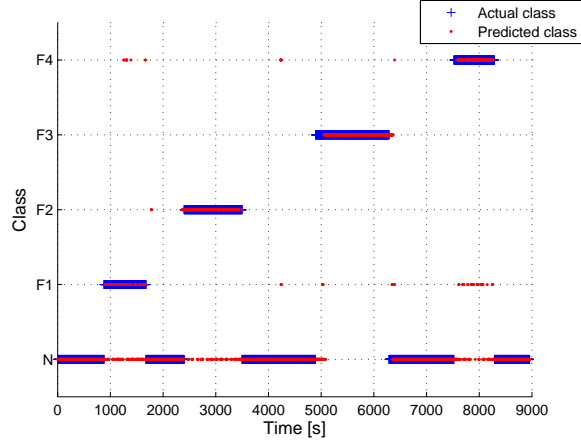


Figure 8. Classification results of the training data

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

245 The fault degrees and diagnosis results corresponding to the experiment of training data
 246 preparation are shown in Fig. 9(a) and Fig. 9(b). The global diagnosis accuracy rate reaches
 247 93.20%, which is significantly increased from the original classification result. The detailed di-
 248 agnosis results are also summarized quantitatively as a confusion matrix in Table 5. Comparing
 249 Table 3 and 5, it could be seen that the diagnosis results corresponding to F1, F2, and F4 are
 250 more accurate and consistent than the raw classification results. The phenomenon can also be
 251 observed visually by comparing Fig. 9(b) and 8. Concerning the data in normal state, the FAR is
 slightly increased.

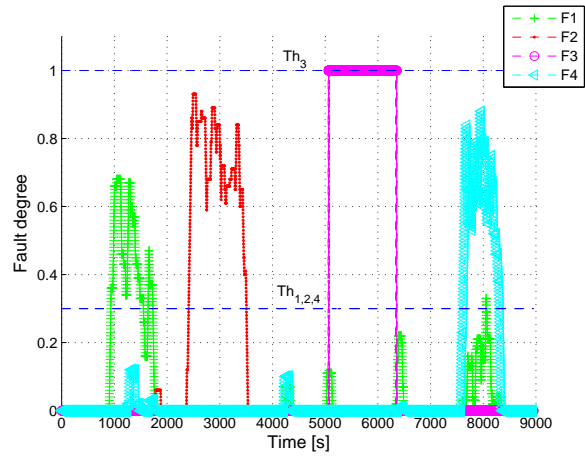
Table 4. Parameter initialization for the design of diagnosis rule

	F1	F2	F3	F4
N_{lag}^i	100	100	1	100
Th_i	0.3	0.3	1	0.3

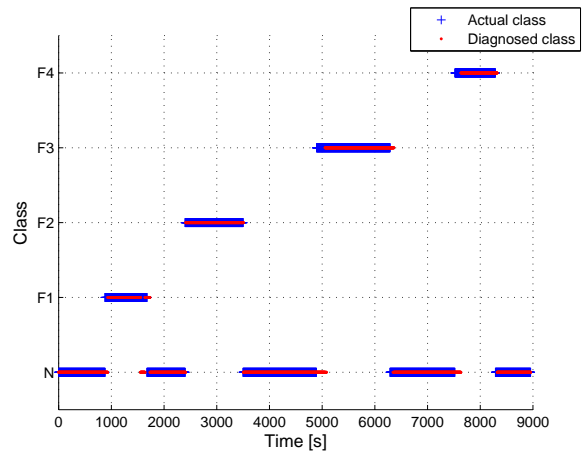
252

Table 5. Confusion matrix of diagnosis results of training data

		Predicted classes				
		NI	F1	F2	F3	F4
Actual classes	NI	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

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

257 5.2. Online validation

258 To realize online validation, the programed SiP was tested online with the real-time data
 259 during another experiment. The operations during this experiment are summarized in Table 6,
 260 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 (NI)
3661	4543	Pressure of 1.3 bar at each side	Low pressure fault (F1)
4544	5374	Back to nominal condition	Normal state (NI)
5375	6541	Pressure of 1.7 bar at each side	High pressure fault (F2)
6542	8128	Back to nominal condition	Normal state (NI)
8129	8909	St. Air 1.5	Low air stoichiometry fault (F4)
8910	9841	Back to nominal condition	Normal state (NI)
9842	11909	Lower relative humidity	Drying fault (F3)
11910	12459	Back to nominal condition	Normal state (NI)

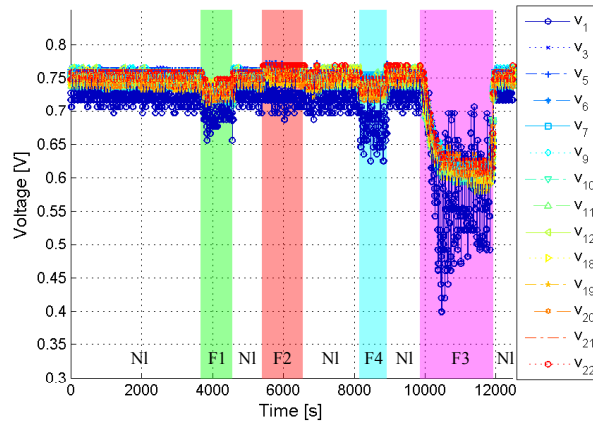
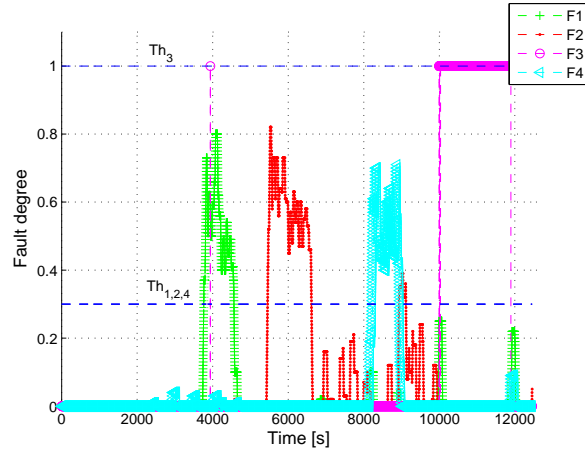
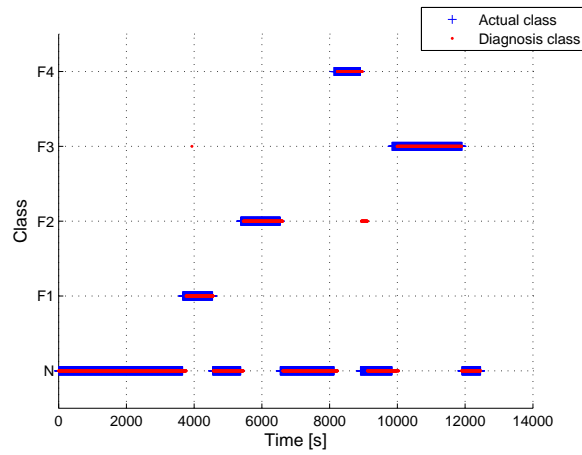


Figure 10. Cell voltages measured during online validation

261 The fault degrees and diagnosis results corresponding to the experiment for online validation
 262 are shown in Fig. 11(a) and Fig. 11(b). The global diagnosis accuracy of the online implemen-
 263 tation is 93.99%. Notice that the global diagnosis accuracy related to online validation is even a
 264 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

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

267 To show the effect of the proposed diagnosis rule, the raw classification results are also pro-
 268 vided here. The global classification accuracy of online validation data is 85.93%. The detailed
 269 classification results corresponding to online validation are summarized in Table 8. By com-
 270 paring Table 8 and 7, it can be seen that the diagnosis accuracy for the data in F1, F2 and F4
 271 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				
		NI	F1	F2	F3	F4
Actual classes	NI	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

272

273 5.3. Discussion

274 5.3.1. Delays caused by the diagnosis rule

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

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

288 5.3.2. Extendability of the approach

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

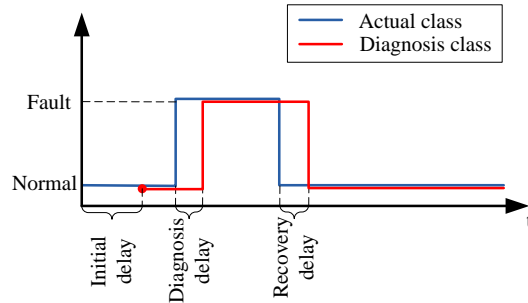


Figure 12. Delays introduced by diagnosis rule

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

304 5.3.3. Ageing effect

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

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

314 6. Conclusion

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

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

328 Acknowledgment

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

332 References

- 333 [1] J. Park, H. Oh, T. Ha, Y. I. Lee, K. Min, A review of the gas diffusion layer in proton exchange membrane fuel
334 cells: Durability and degradation, *Applied Energy* 155 (2015) 866 – 880. doi:http://dx.doi.org/10.1016/
335 j.apenergy.2015.06.068.
- 336 [2] H. Sun, C. Xie, H. Chen, S. Almheiri, A numerical study on the effects of temperature and mass transfer in
337 high temperature PEM fuel cells with ab-PBI membrane, *Applied Energy* 160 (2015) 937 – 944. doi:http:
338 //dx.doi.org/10.1016/j.apenergy.2015.02.053.
- 339 [3] J.-H. Jang, W.-M. Yan, H.-C. Chiu, J.-Y. Lui, Dynamic cell performance of kW-grade proton exchange membrane
340 fuel cell stack with dead-ended anode, *Applied Energy* 142 (2015) 108 – 114. doi:http://dx.doi.org/10.
341 1016/j.apenergy.2014.12.073.
- 342 [4] C. Raga, A. Barrado, A. Lazaro, C. Fernandez, V. Valdivia, I. Quesada, L. Gauchia, Black-box model, identifica-
343 tion technique and frequency analysis for pem fuel cell with overshooted transient response, *IEEE Trans. Power*
344 *Electron.* 29 (10) (2014) 5334–5346. doi:10.1109/TPEL.2013.2292599.
- 345 [5] A. Simons, C. Bauer, A life-cycle perspective on automotive fuel cells, *Applied Energy* 157 (2015) 884 – 896.
346 doi:http://dx.doi.org/10.1016/j.apenergy.2015.02.049.
- 347 [6] P. Xuewei, A. Rathore, Novel bidirectional snubberless naturally commutated soft-switching current-fed full-bridge
348 isolated dc/dc converter for fuel cell vehicles, *IEEE Trans. Ind. Electron.* 61 (5) (2014) 2307–2315. doi:10.1109/
349 TIE.2013.2271599.
- 350 [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
351 durability: Degradation mechanisms and mitigation strategies, *Journal of Power Sources* 184 (1) (2008) 104 – 119.
352 doi:http://dx.doi.org/10.1016/j.jpowsour.2008.06.006.
- 353 [8] H. Chen, P. Pei, M. Song, Lifetime prediction and the economic lifetime of Proton Exchange Membrane fuel cells
354, *Applied Energy* 142 (2015) 154 – 163. doi:http://dx.doi.org/10.1016/j.apenergy.2014.12.062.
- 355 [9] B. Cai, Y. Liu, Q. Fan, Y. Zhang, Z. Liu, S. Yu, R. Ji, Multi-source information fusion based fault diagnosis of
356 ground-source heat pump using Bayesian network, *Applied Energy* 114 (2014) 1 – 9. doi:http://dx.doi.org/
357 10.1016/j.apenergy.2013.09.043.
- 358 [10] S. Yin, S. Ding, X. Xie, H. Luo, A Review on Basic Data-Driven Approaches for Industrial Process Monitoring,
359 *IEEE Trans. Ind. Electron.* 61 (11) (2014) 6418–6428. doi:10.1109/TIE.2014.2301773.
- 360 [11] Q. Liu, S. Qin, T. Chai, Multiblock Concurrent PLS for Decentralized Monitoring of Continuous Annealing Pro-
361 cesses, *IEEE Trans. Ind. Electron.* 61 (11) (2014) 6429–6437. doi:10.1109/TIE.2014.2303781.
- 362 [12] H. He, Z. Liu, Y. Hua, Adaptive Extended Kalman Filter Based Fault Detection and Isolation for a Lithium-Ion
363 Battery Pack, *Energy Procedia* 75 (2015) 1950 – 1955. clean, Efficient and Affordable Energy for a Sustainable
364 Future: The 7th International Conference on Applied Energy (ICAE2015). doi:http://dx.doi.org/10.1016/
365 j.egypro.2015.07.230.
- 366 [13] R. Petrone, Z. Zheng, D. Hissel, M. Péra, C. Pianese, M. Sorrentino, M. Becherif, N. Yousfi-Steiner, A review
367 on model-based diagnosis methodologies for PEMFCs, *International Journal of Hydrogen Energy* 38 (17) (2013)
368 7077–7091. doi:10.1016/j.ijhydene.2013.03.106.
- 369 [14] Z. Zheng, R. Petrone, M. Péra, D. Hissel, M. Becherif, C. Pianese, N. Yousfi Steiner, M. Sorrentino, A review on
370 non-model based diagnosis methodologies for PEM fuel cell stacks and systems, *International Journal of Hydrogen*
371 *Energy* 38 (21) (2013) 8914–8926. doi:10.1016/j.ijhydene.2013.04.007.
- 372 [15] Z. Li, Data-driven fault diagnosis for pemfc systems, Ph.D. thesis, Aix-Marseille University (2014).
- 373 [16] T. Escobet, D. Feroldi, S. de Lira, V. Puig, J. Quevedo, J. Riera, M. Serra, Model-based fault diagnosis in PEM fuel
374 cell systems, *Journal of Power Sources* 192 (1) (2009) 216–223.
- 375 [17] S. de Lira, V. Puig, J. Quevedo, A. Husar, LPV observer design for PEM fuel cell system: Application to fault
376 detection, *Journal of Power Sources* 196 (9) (2011) 4298 – 4305, CONAPPICE 2010. doi:http://dx.doi.
377 org/10.1016/j.jpowsour.2010.11.084.
- 378 [18] D. Benouioua, D. Candusso, F. Harel, L. Oukhellou, Fuel cell diagnosis method based on multifractal analysis of
379 stack voltage signal, *International Journal of Hydrogen Energy* 39 (5) (2014) 2236 – 2245. doi:http://dx.doi.
380 org/10.1016/j.ijhydene.2013.11.066.

- 381 [19] J. Kim, I. Lee, Y. Tak, B. Cho, State-of-health diagnosis based on hamming neural network using output voltage
382 pattern recognition for a PEM fuel cell, *International Journal of Hydrogen Energy* 37 (5) (2012) 4280–4289. doi :
383 10.1016/j.ijhydene.2011.11.092.
- 384 [20] Z. Li, R. Outbib, S. Giurgea, D. Hissel, Y. Li, Fault detection and isolation for Polymer Electrolyte Membrane
385 Fuel Cell systems by analyzing cell voltage generated space, *Applied Energy* 148 (2015) 260 – 272. doi:http:
386 //dx.doi.org/10.1016/j.apenergy.2015.03.076.
- 387 [21] Z. Zheng, M.-C. Péra, D. Hissel, M. Becherif, K.-S. Agbli, Y. Li, A double-fuzzy diagnostic methodology dedicated
388 to online fault diagnosis of proton exchange membrane fuel cell stacks, *Journal of Power Sources* 271 (0) (2014)
389 570 – 581. doi:http://dx.doi.org/10.1016/j.jpowsour.2014.07.157.
- 390 [22] J. Hua, J. Li, M. Ouyang, L. Lu, L. Xu, Proton exchange membrane fuel cell system diagnosis based on the mul-
391 ti-variate statistical method, *International Journal of Hydrogen Energy* (2011) 1–10doi:10.1016/j.ijhydene.
392 2011.05.075.
- 393 [23] Z. Li, R. Outbib, S. Giurgea, D. Hissel, Diagnosis for pemfc systems: A data-driven approach with the capabilities
394 of online adaptation and novel fault detection, *IEEE Trans. Ind. Electron.* 62 (8) (2015) 5164–5174. doi:10.
395 1109/TIE.2015.2418324.
- 396 [24] T. Rauber, F. de Assis Boldt, F. Varejao, Heterogeneous Feature Models and Feature Selection Applied to Bearing
397 Fault Diagnosis, *IEEE Trans. Ind. Electron.* 62 (1) (2015) 637–646. doi:10.1109/TIE.2014.2327589.
- 398 [25] M. Amar, I. Gondal, C. Wilson, Vibration Spectrum Imaging: A Novel Bearing Fault Classification Approach,
399 *IEEE Trans. Ind. Electron.* 62 (1) (2015) 494–502. doi:10.1109/TIE.2014.2327555.
- 400 [26] D. You, X. Gao, S. Katayama, WPD-PCA-Based Laser Welding Process Monitoring and Defects Diagnosis by
401 Using FNN and SVM, *IEEE Trans. Ind. Electron.* 62 (1) (2015) 628–636. doi:10.1109/TIE.2014.2319216.
- 402 [27] C. Chan, J. Jiang, G. Chen, X. Wang, K. Chau, A novel polyphase multipole square-wave permanent magnet motor
403 drive for electric vehicles, *IEEE Trans. Ind. Appl.* 30 (5) (1994) 1258–1266. doi:10.1109/28.315237.
- 404 [28] S. Kang, K. Min, Dynamic simulation of a fuel cell hybrid vehicle during the federal test procedure-75 driving
405 cycle, *Applied Energy* 161 (2016) 181 – 196. doi:http://dx.doi.org/10.1016/j.apenergy.2015.09.
406 093.
- 407 [29] Website: Smartfusion introduction.
408 URL <http://www.microsemi.com/products/fpga-soc/soc-fpga/smartfusion>
- 409 [30] F. Rothan, C. Condemine, B. Delaet, O. Redon, a. Giraud, A low power 16-channel fully integrated GMR-
410 based current sensor, *European Solid-State Circuits Conference* (2012) 245–248doi:10.1109/ESSCIRC.2012.
411 6341304.
- 412 [31] A. GIRAUD, Embedded smart sensors for measurement and diagnosis of measurement and diagnosis of multicells
413 PEMFC, report - project DIAPASON (June 2012).
- 414 [32] P. Rodatz, F. Büchi, C. Onder, L. Guzzella, Operational aspects of a large PEFC stack under practical conditions,
415 *Journal of Power Sources* 128 (2) (2004) 208–217. doi:10.1016/j.jpowsour.2003.09.060.
- 416 [33] Z. Li, S. Giurgea, R. Outbib, D. Hissel, Fault diagnosis and novel fault type detection for pemfc system based
417 on spherical-shaped multiple-class support vector machine, in: *Advanced Intelligent Mechatronics (AIM), 2014*
418 *IEEE/ASME International Conference on*, 2014, pp. 1628–1633. doi:10.1109/AIM.2014.6878317.
- 419 [34] J. C. Platt, Sequential Minimal Optimization : A Fast Algorithm for Training Support Vector Machines, *Technical*
420 *Report MSR-TR-98-14*, Microsoft Research (1998) 1–21.
- 421 [35] A. Widodo, B.-s. Yang, Support vector machine in machine condition monitoring and fault diagnosis, *Mechanical*
422 *Systems and Signal Processing* 21 (2007) 2560–2574. doi:10.1016/j.ymsp.2006.12.007.
- 423 [36] C.-w. Hsu, C.-j. Lin, A Comparison of Methods for Multiclass Support Vector Machines, *IEEE Trans. Neural Netw.*
424 13 (2) (2002) 415–425.
- 425 [37] G. C. Cawley, N. L. Talbot, Fast exact leave-one-out cross-validation of sparse least-squares support vector ma-
426 chines, *Neural Networks* 17 (10) (2004) 1467 – 1475. doi:http://dx.doi.org/10.1016/j.neunet.2004.
427 07.002.
- 428 [38] N. Wagner, E. Glzow, Change of electrochemical impedance spectra (EIS) with time during CO-poisoning of the
429 Pt-anode in a membrane fuel cell , *Journal of Power Sources* 127 (12) (2004) 341 – 347, eighth Ulmer Electro-
430 chemische Tage. doi:http://dx.doi.org/10.1016/j.jpowsour.2003.09.031.
- 431 [39] C. Farrell, C. Gardner, M. Ternan, Experimental and modelling studies of CO poisoning in PEM fuel cells, *Journal*
432 *of Power Sources* 171 (2) (2007) 282 – 293. doi:http://dx.doi.org/10.1016/j.jpowsour.2007.07.006.
- 433 [40] X. Zhang, P. Pisu, An unscented kalman filter based approach for the health-monitoring and prognostics of a
434 polymer electrolyte membrane fuel cell, in: *Annual Conference of the Prognostics and Health Management Society*
435 2012, Vol. 3, 2012, p. 9.
- 436 [41] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, N. Zerhouni, Prognostics and Health Management of PEMFC
437 State of the art and remaining challenges, *International Journal of Hydrogen Energy* 38 (35) (2013) 15307 – 15317.
438 doi:http://dx.doi.org/10.1016/j.ijhydene.2013.09.051.
- 439 [42] R. Petrone, D. Hissel, M. Péra, D. Chamagne, R. Gouriveau, Accelerated stress test procedures for PEM fuel cells

440 under actual load constraints: State-of-art and proposals, International Journal of Hydrogen Energy 40 (36) (2015)
441 12489 – 12505. doi:<http://dx.doi.org/10.1016/j.ijhydene.2015.07.026>.