HAL open science

Diagnosis of fuel cells using instantaneous frequencies and envelopes extracted from stack voltage signal

Djedjiga Benouioua, Denis Candusso, Fabien Harel, Pierre Picard

To cite this version:

Djedjiga Benouioua, Denis Candusso, Fabien Harel, Pierre Picard. Diagnosis of fuel cells using instantaneous frequencies and envelopes extracted from stack voltage signal. International Journal of Hydrogen Energy, 2022, 47 (16), pp 9706-9718. 10.1016/j.ijhydene.2022.01.046 . hal-03595482

HAL Id: hal-03595482

https://hal.science/hal-03595482v1

Submitted on 3 Mar 2022

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.

Diagnosis of fuel cells using instantaneous frequencies and envelopes extracted from stack voltage signal

Djedjiga BENOUIOUA[1,2,3], Denis CANDUSSO[1,2], Fabien HAREL[4], Pierre PICARD[2,5]

[1] SATIE, Univ Gustave Eiffel, COSYS, FCLAB, Rue Ernest Thierry Mieg, 90010 Belfort Cedex, France.

[2] ITE EFFICACITY, 14-20 Boulevard Newton, Champs-sur-Marne, 77447 Marne la Vallée, France.

[3] Psychological & Healthcare Center SA, Rue de la Rôtisserie 2, 1204 Geneva, Switzerland.

[4] LICIT-ECO7 Lab, Univ Gustave Eiffel, ENTPE, Univ Lyon, FCLAB, Rue Ernest Thierry Mieg, 90010 Belfort Cedex, France.

[5] ENGIE, CRIGEN, 4 rue Jopéphine Baker, 93240 Stains, France.

Abstract:

The work carried out aims to diagnose fuel cells with reduced instrumentation and computation times. The article describes a non-intrusive, application-oriented diagnostic tool, based on the sole measurement of the stack voltage and requiring no specific external excitation of the electrochemical generator. The adopted data-driven method relies on well-suited signal analysis techniques (fast calculations of relevant fault signatures based on envelopes and instantaneous frequencies) and information processing (pattern recognition). A wide range of operating regimes can be identified (variations in flow rates, pressures, temperatures; combinations of simultaneous faults), even when they correspond to small deviations from nominal conditions. The portability of the method has been studied on two PEMFC stacks, designed for different applications: transport and stationary. Correct classification rates close to 98% are obtained in both cases.

Keywords:

Fuel cell; diagnosis method; pattern recognition; instantaneous envelopes; instantaneous frequencies.

Highlights:

- A data-driven and application-oriented method is developed to diagnose fuel cells.
- The input of the tool is the fuel cell stack voltage recorded at 11 Hz.
- Operating regime patterns are based on envelopes and instantaneous frequencies.
- Fault identifications are performed using machine learning approaches.
- High classification rates (98%) are obtained for two different PEMFC technologies.

1. Background and objectives

Durability and reliability are still major issues in the field of Fuel Cell (FC) generators, both for transport applications (objective: at least 6,000 hours of operation under dynamic loads for the automobile) and stationary applications (objective: 40,000 to 60,000 hours of operation). Besides, works on the detection and identification of sources of faults in FC gensets have shown that most of the constraints that modify and disrupt the operation are due to the own auxiliaries (i.e. electronics, valves, pumps, sensors, etc.) [1, 2]. In this context, the development of fault diagnosis tools at the FC system level appears to be a lever for optimizing the conditions of use of FC generators, for the longest possible service life. The value of a diagnostic module in power generation systems lies in its ability to identify and isolate the cause of a malfunction. Once the fault has been identified, the supervisory task triggers alarms, either for manual maintenance operations or for automatic corrective actions that react to the fault. A reliable diagnostic module allows: a gain in safety by avoiding accidents, a gain in production by reducing the downtime of the system in question and an increase in its lifespan. In the literature, a growing number of works carried out around FC systems propose diagnostic tools [3, 4], among which:

- Electrochemical characterization techniques such as polarization curve [5, 6], current interruption method [7] and Electrochemical Impedance Spectroscopy (EIS) [8] which have been widely used for the diagnosis of batteries and power converters [9]. These experimental methods have shown their effectiveness in obtaining electrochemical signatures of defects and in detecting different failure modes, but they have limitations for real time applications. Indeed, they disturb or interrupt the normal operation of the PEMFC, require precise controls and solicitations on the current or voltage, as well as additional equipment that is often expensive and bulky (for example, an electrochemical impedance spectrometer).

- Methods based on system modelling, sometimes called "internal methods" [10]. We can mention for example the models based on electrical equivalent circuits [11] which exploit electrochemical characterization measurements or current-voltage data, as well as analytical models and observers [12] that require an in-depth knowledge of the multi-physical mechanisms and numerous parameters governing the operating process of the energy converter. Most of these model-based methods rely on residuals (differences between observed and predicted values). But the non-linear nature of the FC physical phenomena, the reversible or non-reversible character of the degradations, the strong interactions between the various components of the converter make it indeed difficult to model the failures, to allow simulating the real performances of the considered system with precision and to generate accurate residuals, which is a prerequisite for the diagnostic stage of these approaches.

- Data-based methods avoid the complex modelling process and overcome the limitations of model-based methods. They seem to be more practical in most cases and they are therefore drawing the attention of many researchers involved in the field of FC diagnosis. These methods exploit signal processing tools to extract appropriate FC health indicators. However, the signals used might come from a possible large number of sensors and thus give rise to a considerable amount of information. Indeed, signals of different physical natures can be collected from a FC in a non-invasive way. Some measurements (current, voltage, temperature, pressure) can be used more easily than others (relative humidity, external magnetic field) in a control system for FCs. Methods such as the wavelet transform [13, 14] and the singularity analysis [15, 16] have thus shown their potential for diagnosing FCs. They are usually associated with pattern recognition approaches to perform the tasks of detection and identification of the defect(s). The works of Z. Zheng et al. [17, 18] provide useful descriptions of the many existing methods in the field. These methods sometimes require important and costly computing resources, which makes their implementation difficult in a real time diagnosis perspective for application purposes.

The new diagnostic tool we propose here is also based on a data-driven approach. It does not directly contribute to a better understanding and description of the physical phenomena present in FCs. It does, however, have a certain practical utility. It builds on a signal processing method consisting in determining local quantities extracted from the only voltage signal measured at the terminals of a FC stack. These quantities are envelopes and instantaneous frequencies, used in a relevant way to generate signatures related to operating regimes or faults. These descriptors of the FC voltage morphology are expected to be better suited than the Fourier Transform for dynamic time signals. As a matter of fact, envelopes and instantaneous frequencies will enable a representation of the voltage signal in both the time and frequency domains. They will also be less complex to generate than spectra calculated from wavelets or singularity analysis [13, 16]. Wavelet-based methods are relatively difficult to implement due to their mathematical complexity and the choice of a number of setting parameters that depend on the morphology of the analysed signal (parameters related to the mother / analysing wavelet, the selected decomposition scales, the signal size and sampling frequency). The tool also implements information processing techniques (more precisely, pattern recognition) to detect abnormal or at least undesired modes of operation, and to be able to initiate corrective measures later (control, maintenance). The objective of this work is to propose a diagnostic tool for FCs that can both be efficient for the identification of operating modes and that is based on mathematical foundations that are simple enough to allow practical implementations and limited computation times. The work presented here has led to a patent application [19] and, in particular, it addresses a current issue in the process of the FC state-of-health determination, namely the extraction of relevant information and the proposal of well-suited fault signatures from the morphology of common and simple measures available in the electrochemical generator.

The diagnostic strategy is therefore developed taking into account the following requirements:

- A limited instrumentation: non-intrusive instrumentation provides a "generic" aspect to the diagnostic system, making it more easily adaptable to different geometries and sizes of FCs. The choice of the stack voltage as a useful signal for the diagnosis, that can reflect the health of the electrochemical generator, also goes in this direction.

- A simple method, more easily implementable from a perspective of real-time diagnosis with a view to the application. Fault signatures with reduced computation times, combined with a data-driven diagnostic approach, are purposely used. The mathematical tools implemented must result in a high-performance diagnostic algorithm, both in terms of classification rate (identification) of the operating modes considered and "calculation costs / time".

- The identification of a wide range of system faults. Complex situations involving two or even three simultaneous faults must be taken into consideration.

- The prevention of failures or irreversible degradations well in advance. The developed method must be sensitive to slight deviations from the nominal operating regime of the FC (caused in particular by the ancillaries of the generator).

This article is organized as follows. The two experimentally studied FCs and the test databases are presented in Section 2. The diagnostic strategy is described in Section 3. It includes a description of the mathematical tools used in the different steps of the method: generation of signatures related to the operating regimes of the FCs, selection of descriptors, and classification. In Section 4, the diagnostic method is applied to the two FC stacks investigated on a variety of operating points. The main conclusions of the paper are given in Section 5.

2. Tested fuel cells and establishment of databases for diagnosis

Two short PEMFC (Proton Exchange Membrane or Polymer Membrane) stacks, designed by two different manufacturers and intended for two distinct applications, are studied to evaluate the portability of the proposed diagnostic method. The stacks are tested under a variety of operating conditions imposed on test benches developed within the Hydrogen - Energy platform in Belfort, France (Fig 1). Details on the design of such test stands can be found in [20]. The first stack studied is a FC designed by CEA LITEN in Grenoble, France for an automotive application. It is equipped with metal gas distribution plates and is supplied with pure hydrogen. The second stack is manufactured by the German companies Riesaer Brennstoffzellentechnik - RBZ GmbH and Inhouse Engineering GmbH. It is dedicated to a stationary application (microcogeneration type, µCHP) and its bipolar plates are made of graphite. It is fed with a hydrogen-rich mixture, simulating a reformate. The main characteristics and nominal operating conditions relevant to the study are given for the two PEMFCs in Table 1.

The experimental protocol developed consists of deliberately introducing different regimes or faults during the operation of the FCs by deviating from their reference operating conditions through actions: on the reactive gas flows with the Factors of Stoichiometry at Anode and Cathode (FSA and FSC), on the reactant pressures (P), on the cooling circuit temperature (T), on the reactant Relative Humidity (RH) [15, 16]. The values of the parameter variations are specified in Table 2. The fault scenarios duplicate various failures that are representative of the application environments. The faults considered correspond mainly to potentially inefficient operations of the FC system ancillaries / actuators or sensors used in the Balance of-Plant (e.g. failure of the air supply subsystem in the FC generator related with the FSC parameter). Trade-offs concerning the parameter ranges explored have been found by considering the performance capabilities of the FCs (a total FC collapse or severe FC stack degradations had to be avoided), by taking into account the technological limitations of the FC test beds, as well as the possible durations of the experimental campaigns.

As mentioned in the introduction, the signal useful for our diagnostic tool is the stack voltage only. Two databases are thus constituted by recording (at 11 Hz) the voltage and the various physical signals reflecting the evolution of certain parameters of the auxiliaries that interact with the stack during its operation.

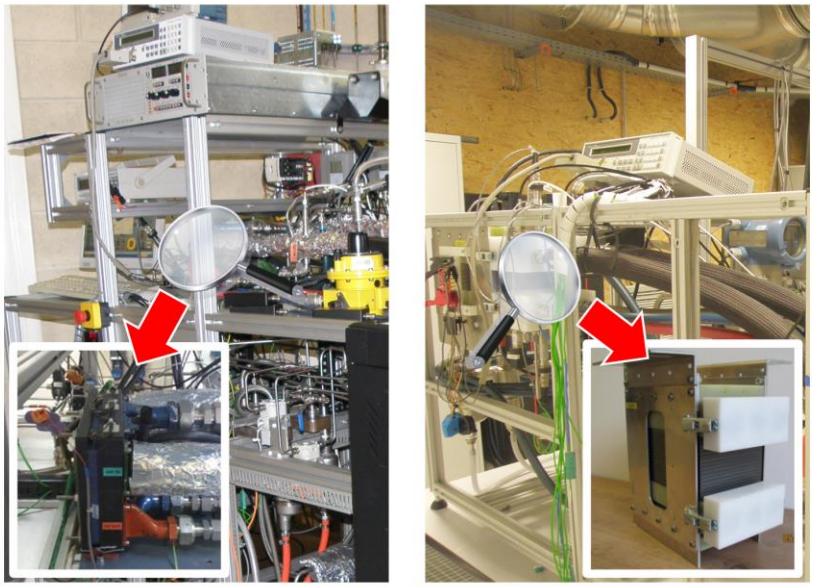

Fig. 1. Pictures of the two investigated PEMFCs and testbenches. Left: CEA LITEN stack. Right: RBZ - Inhouse Engineering GmbH stack.

Table 1. Characteristics and reference conditions of the two PEMFCs tested: stack from CEA LITEN designed for automotive purpose (referred to as "Auto") and stack from RBZ - Inhouse Engineering GmbH dedicated to stationary application (referred to as "µCHP").

Parameters	**Values of the "Auto" stack**	**Values of the "µCHP" stack**
Number of cells	8	12
Stack size	220 mm × 160 mm × 186 mm	270 mm × 170 mm x 270 mm
Active surface of the electrode	220 cm²	196 cm²
Cell current	110 A	80 A
Factor of Stoichiometry at Cathode, FSC (air)	2	2

Factor of Stoichiometry at Anode, FSA	1.5 (pure H_2)	1.3 (75% H_2 + 25% CO_2)
Absolute inlet pressure at anode and cathode, P	150 kPa	110 kPa
Temperature at the outlet of the FC temperature control loop, T	80°C	70°C
Relative Humidity at anode and cathode inlets, RH	50 %	50 %

Table 2. Experimental reference conditions (in blue and green) and alternative operating regimes (in bold, red, underlined; Nd = Not done) with the variations (Δ) of test conditions applied to the two FCs: CEA ("Auto") and RBZ - Inhouse Engineering GmbH ("µCHP").

	Reference		**Δ FSC**		**Δ FSA**		**Δ P**		**Δ T**		**Δ RH**	
Stack	**Auto**	**µCHP**	**Auto**	**µCHP**	**Auto**	**µCHP**	**Auto**	**µCHP**	**Auto**	**µCHP**	**Auto**	**µCHP**
FSC	2	2	**1.3**	**2.6** **1.6**	2	2	2	2	2	2	2	2
FSA	1.5	1.3	1.5	1.3	**1.3**	**1.5** **1.2**	1.5	1.3	1.5	1.3	1.5	1.3
P (bar abs)	1.5	1.1	1.5	1.1	1.5	1.1	**1.3**	**Nd**	1.5	1.1	1.5	1.1
T (°C)	80	70	80	70	80	70	80	70	**75**	**72** **65**	80	70
RH (%)	50	50	50	50	50	50	50	50	50	50	**Nd**	**54** **46**

3. Principles of the diagnosis strategy

The diagnostic tool can first be presented in a global way, through the different main steps of the implemented diagnostic strategy: from the constitution of a reference database, to the

classification of the new observations allowing the recognition of operating regimes (Fig. 2). Essential information on the mathematical tools used in the different steps is provided in the following section.

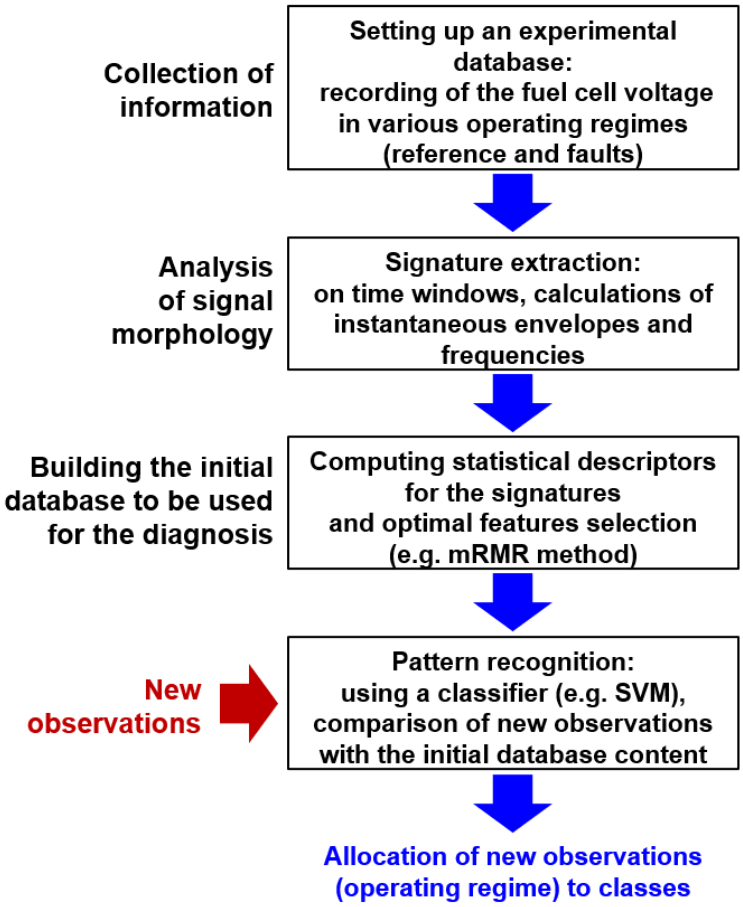

Fig. 2. The main steps of the diagnosis strategy.

3.1. Analysis of the FC voltage signal: key-notions about envelopes and instantaneous frequencies

Most of the information contained in a rough signal lies in its irregular structures. The fine analysis of these structures, thanks to the calculation of local quantities and characteristics, allows access to information latent in the signal. The contents of the voltage signals produced

by a FC are thus loaded with information, linked to various physical phenomena that take place in the electrochemical cell and that also depend on the operation of the auxiliaries in the complete generator. The analysis of such signals as a function of the operating conditions reveals oscillating phenomena whose characteristics change over time. It therefore seems interesting to describe them by showing their instantaneous characteristics. The time-frequency approach, which brings out the envelope and the frequency localised in time (instantaneous), seems to be very well adapted to characterise this type of signal.

In the rest of Subsection 3.1, details are given on the mathematical tools used for the extraction of patterns from the FC voltage signal. Three examples of signatures are given: the first two cases are based on known and simple "Chirp" type signals, and the third one is proposed from voltage records of a FC under different operating regimes. The remaining part of the section is devoted to the calculation and selection of descriptors from the patterns (envelopes and instantaneous frequencies), as well as to the discrimination of different operating regimes using a classifier.

3.1.1. Mathematical definitions of envelopes and instantaneous frequencies

Any real signal $x(t)$ can be associated with the corresponding analytical (complex) signal [21-25]:

$$z_x(t) = x(t) + j\, H\{x(t)\} \tag{1}$$

H denotes the Hilbert Transform (HT). This transform is used to extend a real signal into the complex domain. The HT associates to a signal $x(t)$ the signal $\check{x}(t)$ defined as follows, with vp the principal value of the integral in the sense of Cauchy:

$$\breve{x}(t) = \frac{1}{\pi}\, vp \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau}\, \mathrm{d}\tau \;=\; \int_{-\infty}^{+\infty} g(t-\tau)x(\tau)d\tau = g(t) * x(t) \qquad (2)$$

It can be related to a filtering operation of the signal $x(t)$ by a linear system of impulse response $g(t) = \frac{1}{\pi t}$, also called "quadrature filter", which is an all-pass filter (infinite bandwidth) introducing a phase shift of $\boldsymbol{\frac{\pi}{2}}$:

$$\breve{X}(f) = F\left\{\frac{1}{\pi t}\right\} \cdot \mathrm{F}\{x(t)\} = -j\, sgn(f)X(f) \qquad (3)$$

The Fourier Transform (FT) of the analytical signal is written as follows with $U(f)$, Heaviside step:

$$Z_x(f) = X(f) + j\,\breve{X}(f) = \big(1 + j\, sgn(f)\big)X(f) = 2U(f)X(f) \qquad (4)$$

The envelope, or instantaneous amplitude (modulus of the analytical signal), describes the amplitude modulation law:

$$Instantaneous\ Amplitude\ (t) = |z_x(t)| \qquad (5)$$

The instantaneous frequency (derived from the phase of the analytical signal) describes the phase modulation law:

$$Instantaneous\ Frequency\ (t) = \frac{1}{2\pi}\frac{d(arg(z_x))}{dt}(t) \qquad (6)$$

3.1.2. Example of calculation with a simple sinusoidal signal

As a basic example, let us consider a real $x(t)$ in the form of a sinusoidal with constant modulus A and frequency f_0 :

$$x(t) = A \sin(2\pi\, f_0\, t) \tag{7}$$

Then, the corresponding analytical (complex) signal $z_x(t)$ can be expressed as follows:

$$z_x(t) = A \sin(2\pi\, f_0\, t) + j\, A \sin\left(2\pi\, f_0\, t - \frac{\pi}{2}\right) \tag{8}$$

$$\Rightarrow\ z_x(t) = A \sin(2\pi\, f_0\, t) - j\, A \cos(2\pi\, f_0\, t) \tag{9}$$

$$\Rightarrow\ z_x(t) = A \exp(j\, 2\pi\, f_0\, t) \exp(-j\, \frac{\pi}{2}) \tag{10}$$

where the phase shift of $\frac{\pi}{2}$ appears clearly.

The envelope (or instantaneous amplitude) and the instantaneous frequency of the real signal $x(t) = A \sin(2\pi\, f_0\, t)$ can finally be obtained as follows:

$$Instantaneous\ Amplitude\ (t) = |z_x(t)| = A \tag{11}$$

$$Instantaneous\ Frequency\ (t) = \frac{1}{2\pi} \frac{d(arg(z_x))}{dt}(t) = f_0 \tag{12}$$

3.1.3. Examples of envelopes and instantaneous frequencies - case of frequency-modulated "Chirp" signals

First, as an illustration, the calculation of envelopes and instantaneous frequencies is applied to "Chirp" type signals. These are complex, pseudoperiodic signals, frequency-modulated around a carrier frequency and also amplitude modulated by an envelope a (with $a(t) \geq 0$) whose variations are slow compared to the oscillations of the phase $\varphi(t)$:

$$Chirp(t) = a(t)\ exp\{j\varphi(t)\} \tag{13}$$

Let us consider two examples of "Chirp" type signals. The first signal follows a linear modulation of the frequency f and the second follows a quadratic evolution of f.

- **Example 1:** "Chirp" with linear frequency modulation from $f_O = 1\ Hz$ to $f_{end} = 100\ Hz$ (Fig. 3).

The plot of Fig. 3 is done with Matlab™ using the available *chirp* and *hilbert* functions. It shows the signal under study and the estimated time trends for its envelope and instantaneous frequencies. The amplitude of the instantaneous envelope fluctuates around 1 (normalized "Chirp" signal), especially for instants corresponding to strong frequency values because of the relationship between the strong oscillation of the signal and the chosen sampling frequency (1 kHz). The estimation of the instantaneous frequency shows a linear temporal evolution of the local frequencies in the "Chirp" signal, from $f_O = 1\ Hz$ to $f_{end} = 100\ Hz$.

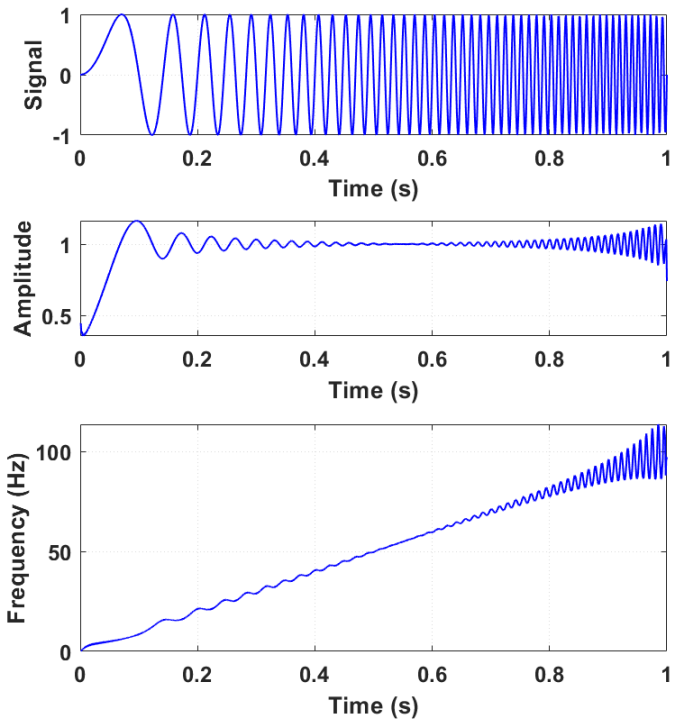

Fig. 3. First example of a "Chirp" signal with linear frequency modulation (top), with the time evolutions of its instantaneous envelope amplitude (middle) and frequencies (bottom).

Similar theoretical calculations to those presented in Section 3.1.2. can be carried out to explain the shapes of the curves observed in Fig. 3. The "Chirp" signal can thus be described by the following equation, related with the real signal $x(t)$ [25]:

$$x(t) = A\sin(2\pi\left[f_0 + \frac{f_{end}-f_0}{2T}t\right]t) \tag{14}$$

with $T = 1\ s$ in the example of Fig. 3.

The corresponding analytical signal $z_x(t)$ can be expressed as follows:

$$z_x(t) = A\exp\left(j\,2\pi\left[f_0 + \frac{f_{end}-f_0}{2T}t\right]t\right)\exp(-j\,\frac{\pi}{2}) \tag{15}$$

In this case, the theoretical expressions for the envelope (instantaneous amplitude), instantaneous phase and frequency of the real signal are as follows:

$$Instantaneous\ Amplitude\ (t) = |z_x(t)| = A \tag{16}$$

$$Instantaneous\ Phase\ (t) = arg(z_x) = 2\pi\left[f_0 + \frac{f_{end}-f_0}{2T}t\right]t - \frac{\pi}{2} \tag{17}$$

$$Instantaneous\ Frequency\ (t) = \frac{1}{2\pi}\frac{d(arg(z_x))}{dt}(t) = f_0 + \frac{f_{end}-f_0}{T}t \tag{18}$$

- **Example 2:** "Chirp" with concave quadratic frequency modulation from $f_O = 100\ Hz$ to $f_{end} = 25\ Hz$ (Fig. 4).

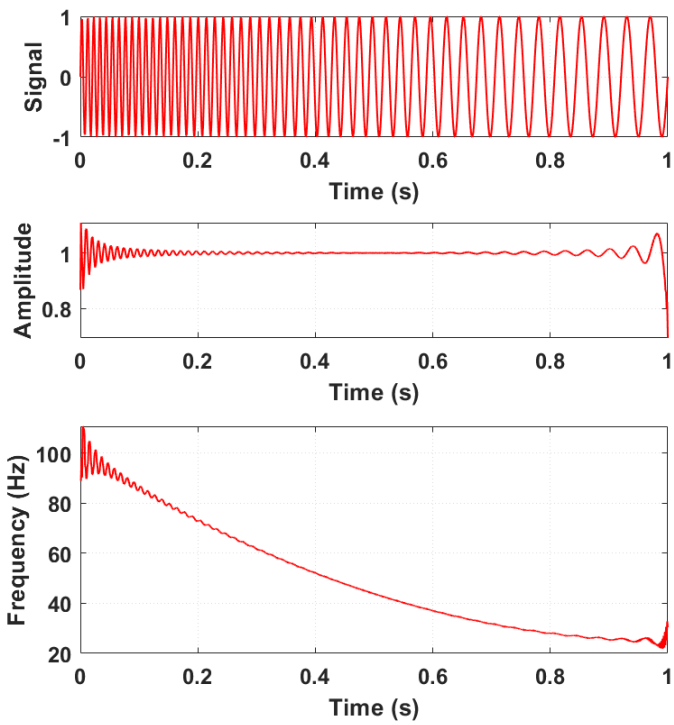

Fig. 4. Second example of a "Chirp" signal with quadratic frequency modulation (top), with the time evolutions of its instantaneous envelope amplitude (middle) and frequencies (bottom).

The two previous examples of "Chirp" signals show the interest of calculating the local quantities of an oscillatory signal, which are the instantaneous envelopes and frequencies. Their curves - by putting time and amplitude, time and frequency in bijective correspondence - make it possible to reveal any information possibly dissimulated or even hidden in the initial raw signal.

3.2. Examples of envelopes and instantaneous frequencies calculated from stack voltage signals

Voltage signals were acquired on the two FCs operated under different regimes. Some of these regimes can lead to improved performance while others can be considered as faults. Faults that occur in a FC can generally be attributed either to a natural ageing of the stack components or to a poor control of the FC system. Most faults in the stack can be directly related to unsuitable operating conditions, such as too high relative humidity or gas flow rates, or too high stack temperatures.

As an example, portions of signals extracted from stack voltage measurement are shown in Fig. 5. They were acquired on the FC designed for a stationary application (µCHP). Each one of the 9 extracts presented, related with the 9 operating conditions considered, include 1000 points corresponding to a duration close to 1.5 min (frequency of 11 Hz). The envelope curves and instantaneous frequencies calculated from the voltage signals are also shown in Fig. 5. They have been estimated on the pre-processed voltage signals (i.e. from which the linear or "trend" component, calculated on each interval of 1000 points corresponding to an operating regime, will have been removed). As can be seen, each regime imposes on the voltage signal, and on the envelope and instantaneous frequency signatures, relatively distinct imprints. This visual observation appears to be a first justification for the use of these local quantities as discriminating descriptors for the FC diagnosis. It should be noted that the time required to calculate the

envelopes and instantaneous frequencies is very short. With a desktop PC, it is less than 1 ms for an interval of 1000 points.

The voltages produced by the FC seem to have (in part) a random and non-stationary nature. As a result, the values of the instantaneous frequencies and amplitudes cannot be interpreted as "modulations" in the sense of deterministic signals (e.g. sinusoidal signals). These local quantities will simply be explored in the same way as signatures can be explored, giving rise to descriptors that feed a classification algorithm.

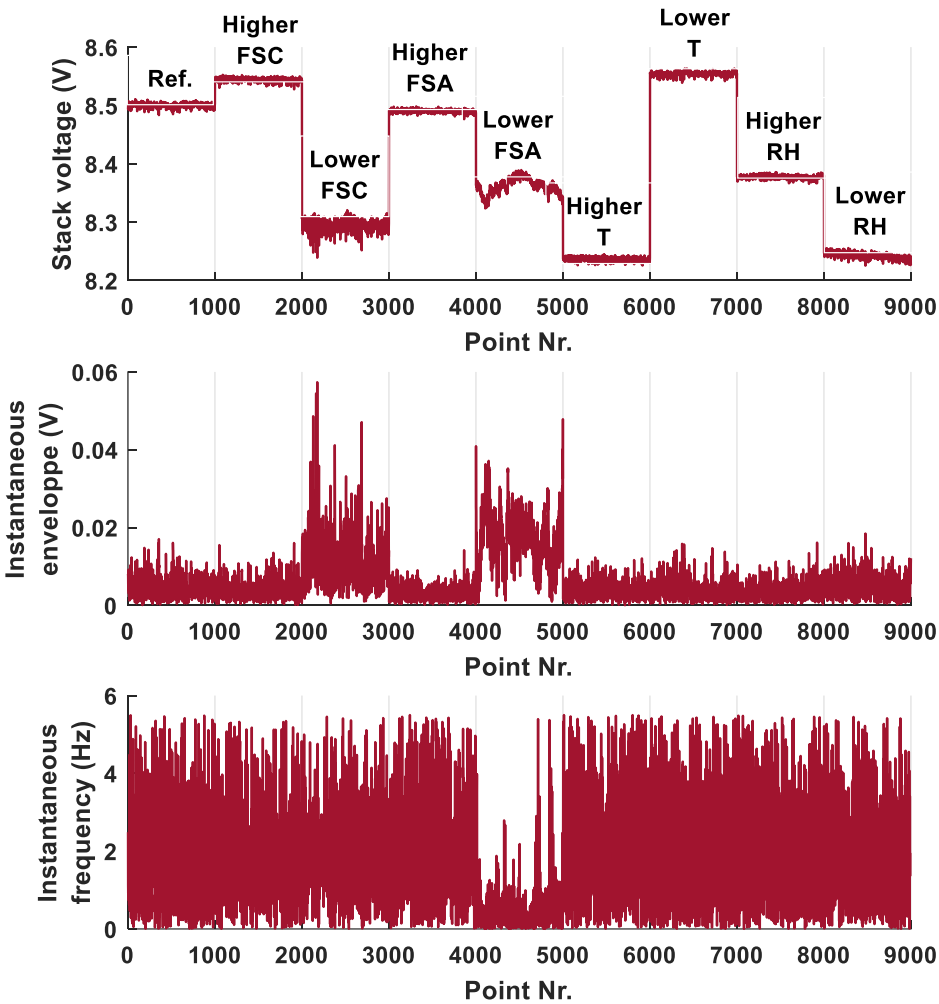

Fig. 5. Examples of voltage signals obtained for the FC dedicated to stationary application and operated in various regimes (top), amplitudes of instantaneous envelope per regime (middle), curve of instantaneous frequency per regime (bottom).

3.3. Method of calculation and selection of descriptors on operating regime signatures

In this study, 8 statistical parameters are estimated on each of the following three signals: raw FC voltage, instantaneous envelope amplitude, and instantaneous frequencies. These statistical parameters are: the maximum value, the minimum value, the mean, the standard deviation, the coefficient of variation, the median, the bias (or "skewness"), and the kurtosis. Initially, with 8 parameters / signal, 24 statistical descriptors result from the analysis of the three signals.

The stage of selecting the relevant parameters then intervenes to reduce the useful number of statistical descriptors and ultimately reduce the calculation time. The mRMR method ("minimal Redundancy, Maximum Relevance") is used in our study to reduce the dimensionality of the initial database [15, 16, 26-28].

3.3.1. Definition of the mRMR method

The mRMR method is based on the calculation of the mutual information $I(x,y)$ which measures the statistical dependence of a random variable x (taking its values in $\{x_1, \dots, x_i\}$) on another: y (taking its values in $\{y_1, \dots, y_j\}$). $I(x,y)$ is also called relative entropy, or Kullback divergence, and is calculated from the joint distribution $p(x,y)$ and the marginal distributions $p(x)$ and $p(y)$ [29]:

$$I(x,y) = \sum_{x_i} \sum_{y_j} p(x,y) \, log\left(\frac{p(x,y)}{p(x).p(y)}\right) \tag{19}$$

$I(x,y)$ is high if x and y are dependent, and zero if the variables are independent. The basic idea of the mRMR method is therefore to use this notion of mutual information to try to: - minimise

the redundancy (mR) between the characteristics and - maximise the relevance (MR) between the variable under study and the class under consideration. This can be achieved in the following way [26-28]:

- Minimising redundancy:

$$Redundancy(x) = \frac{1}{|S|^2} \sum_{x,y \in S} I(x,y) \quad (20)$$

with: $|S|$ the size of the set defined by the two variables x and y.

- Maximising relevance:

$$Relevance(x) = \frac{1}{|S|} \sum_{x \in S} I(C,x) \quad (21)$$

where: C is the class considered, with labels $\{c_1, \dots, c_n\}$.

The score of a variable is the combination of these two factors. It can be calculated in two different ways:

$$Score(x) = \frac{Relevance(x)}{Redundancy(x)} \quad (22)$$

or else:

$$Score(x) = Relevance(x) - Redundancy(x) \quad (23)$$

3.3.2. Application of the mRMR method

As an example, let us consider the case of the study conducted with the FC dedicated to stationary application. From the 24 statistical parameters initially considered, the "relevant" statistical descriptors (i.e. which lead to better class discrimination) are now only 4:

- the coefficient of variation of instantaneous frequencies,
- the average of the instantaneous amplitudes,
- the maximum value of the raw voltage,
- the minimum value of the raw voltage.

3.4. Identification of operating regimes using a classifier

This is in this final stage of the diagnostic strategy that methods from the field of pattern recognition need to be used. Statistical descriptors were estimated from the three useful signals: the raw voltage signal, the instantaneous frequency curve, and the instantaneous amplitude curve. The descriptors, that were deemed "relevant" following the application of the mRMR method, form a learning base in which the operating regimes under study are characterised. These are now assigned to groups, commonly called "classes". The supervised classification technique, called multi-class Support Vectors Machines (SVM), is the classifier used in our study. SVMs have been shown to be effective in many fields such as medical diagnosis, bioinformatics, information retrieval, computer vision, finance, and also diagnosis and prognosis of FCs [30, 31]. SVMs are a class of algorithms, based on the search for the optimal margin hyperplane, initially defined for cases of binary discrimination between linearly separable data (i.e. with 2 classes that can be labelled {-1, 1}). A diagram illustrating this principle is given in Fig. 6. The mathematical foundations of the SVM approach are detailed in several works, notably in [32-34].

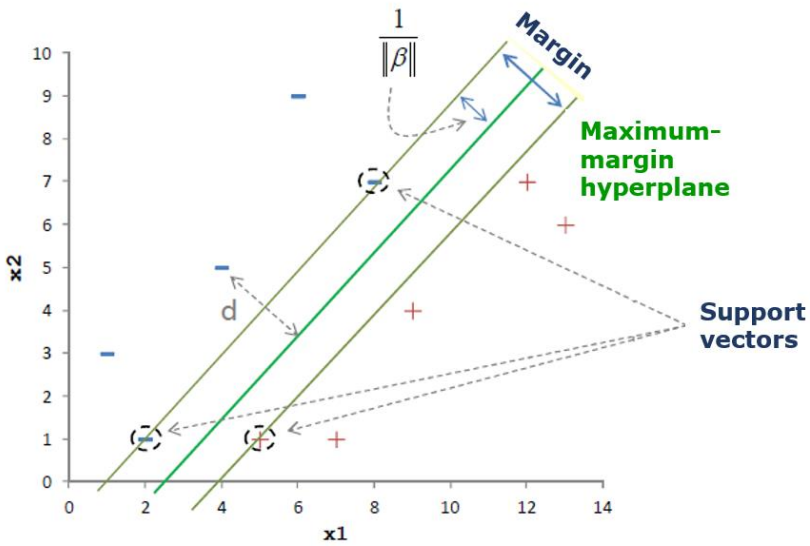

Fig. 6. The SVM classifier approach. Illustration of the notions of margin, support vectors, and separating hyperplane (case of a binary classification) [32].

To address the basic mathematical formulation of the SVM approach, let us consider the simple case of a binary classification where the task is to separate data $x_{i \in R}$ into two classes y labelled +1 or -1 ($y \in \{+1, -1\}$). The objective is to find a linear separation that distinguishes x_i belonging to the +1 class from x_i belonging to the -1 class. The classifier f (or separation function) takes the form of a linear combination of the variables x_i :

$$y = f\left(x_1, x_2, \dots, x_p; \beta_1, \beta_2, \dots, \beta_p\right) \tag{24}$$

$$f(x) = x_i^T \beta + \beta_0 = x_1\beta_1 + x_2\beta_2 + \cdots + \beta_0 \tag{25}$$

with: $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ and β_0 being the $(p+1)$ parameters to be estimated.

The principle of margin maximisation is formulated to choose an optimal hyperplane that should maximise the distance between the separation boundary and the points of each class that are closest to it (Fig. 6). Maximising the margin is therefore equivalent to minimising the norm of the parameter vector β, i.e. [32]:

$$\min_{\beta,\beta_0} \frac{1}{2}\|\beta\|^2 \text{ with } \|\beta\| = \sqrt{\beta_1^2 + \cdots + \beta_p^2} \tag{26}$$

under constraint, with $i = 1, \ldots, n$:

$$y_i \times (x_i^T \beta + \beta_0) \geq 1 \tag{27}$$

The constraints indicate that all points are on the right side; at worst, they are on the line defining the support vectors. Solving this problem (i.e. computing the β-coefficients of the hyperplane) can be done using Lagrange multipliers. Originally, SVMs are binary classifiers. However, there are methods that extend their applications to the case of multiple classes and non-linearly separable data by introducing kernel functions. In our case, we have adopted the "One versus all" approach to multi-class SVMs. In Fig. 7, we present an example of a classification result for non-linearly separable data into 8 classes. The computation was performed using the LIBSVM toolbox, available in the Matlab™ environment, and developed by Chih-Chung Chang and Chih-Jen Lin [35].

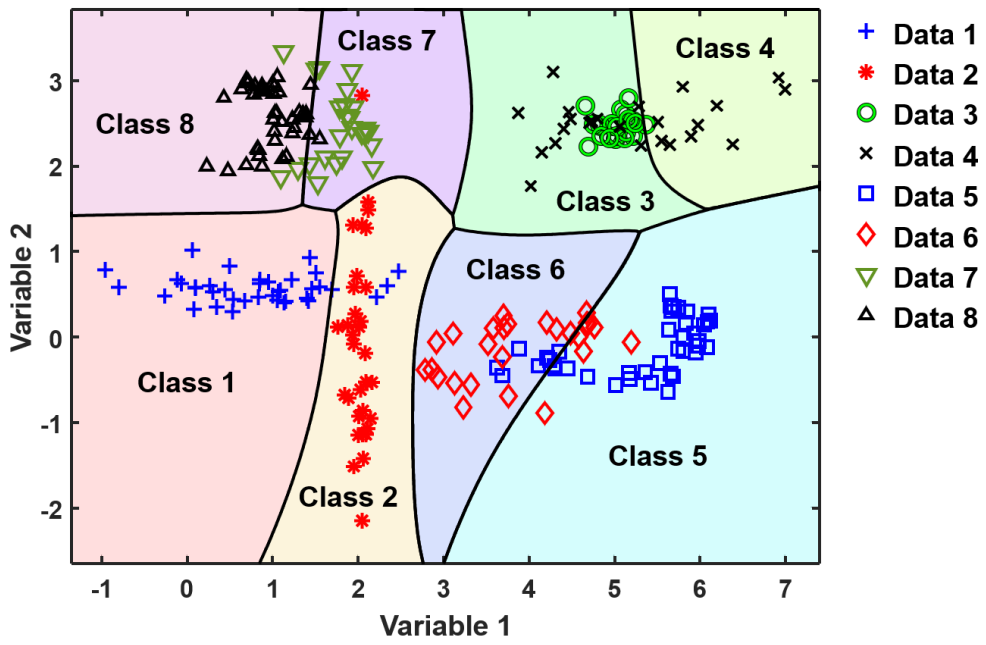

Fig. 7. Example of a multi-class classification of non-linearly separable data using the LIBSVM toolbox [35].

4. Final application of the diagnostic strategy

As a reminder, all the operating regimes and parameter variations (Δ) considered for the validation of the developed diagnostic tool are summarised in Table 2, for the two PEMFCs tested: CEA stack (automotive application - "Auto") and RBZ - Inhouse Engineering FC (micro-cogeneration application - "µCHP"). Statistical descriptors are therefore determined on the basis of the raw voltage signals, envelopes and instantaneous frequencies corresponding to the different operating regimes studied. Only the most "relevant" descriptors, identified by the mRMR method, are used in the pattern recognition classification step (implementation of the SVM algorithm). Additional information related to the diagnostic process applied to the two FCs is given in Table 3. The diagnostic tool allows the discrimination of a wide range of operating

regimes, for both stacks, with very good classification rates of about 98% (97.5% for the "Auto" PEMFC and 98.8% for the "µCHP" PEMFC). The classification rates per operating regime studied and per stack are indicated in Tables 4 and 5, showing the forms of the confusion matrices from the pattern recognitions. It can be seen that most of the operating regimes are identified at 100%. The classification results are of the same order of magnitude, and even slightly better, than those presented in [16] and obtained from the same types of FC operating regimes but on the basis of signatures (singularity spectra) that were more complex to generate.

Table 3. Additional information on the diagnostic process (signature calculations, pattern recognition) applied to the two FCs.

	"Auto" Stack	**"µCHP" Stack**
Number of classes considered (operating regimes)	8	9
Number of individuals considered in the learning phase (1 individual = 1000 points)	8×20	166
Number of individuals taken into account for the test phase	80 (=8×10)	85 (=8×10+5)
Number of descriptors (relevant statistical parameters)	5	4
Envelope and instantaneous frequency signal calculation time	0.3 ms / individual	0.4 ms / individual
Calculation time for statistical parameters	0.2 ms / individual	0.3 ms / individual
Computing time related to the learning phase (building the SVMs)	0.2 s	0.2 s
Computing time related to the test phase (classifying individuals)	0.8 ms for the 80 individuals	1 ms for the 85 individuals
Accuracy of diagnosis (rate of correct classification)	97.5 % (2 out of 80 individuals are misclassified)	98.8 % (1 out of 85 individuals is misclassified)

Table 4. Classification results obtained from the experimental data based on the CEA stack ("Auto" application). The delta symbol (Δ) refers to the change in the parameter.

	Reference class	Δ FSC class	Δ FSA class	Δ P class	Δ (FSA + P) class	Δ (FSC + P) class	Δ (FSC + FSA) class	Δ (FSC + FSA + P) class
Reference	9	0	0	0	1	0	0	0
Δ FSC	0	**10**	0	0	0	0	0	0
Δ FSA	0	0	**10**	0	0	0	0	0
Δ P	0	0	0	**10**	0	0	0	0
Δ (FSA + P)	0	0	0	0	**10**	0	0	0
Δ (FSC + P)	0	0	0	0	0	**10**	0	0
Δ (FSC + FSA)	0	0	0	0	1	0	9	0
Δ (FSC + FSA + P)	0	0	0	0	0	0	0	**10**

Table 5. Classification results obtained for the operating regimes of the RBZ - Inhouse Engineering stack ("µCHP" application).

	Reference class	Δ FSC (high level) class	Δ FSC (low level) class	Δ FSA (high level) class	Δ FSA (low level) class	Δ T (high level) class	Δ T (low level) class	Δ RH (high level) class	Δ RH (low level) class
Reference	**10**	0	0	0	0	0	0	0	0
Δ FSC (high level)	0	**10**	0	0	0	0	0	0	0
Δ FSC (low level)	0	0	**10**	0	0	0	0	0	0
Δ FSA (high level)	0	0	0	**10**	0	0	0	0	0
Δ FSA (low level)	0	0	0	0	**10**	0	0	0	0
Δ T (high level)	0	0	0	0	0	**10**	0	0	0
Δ T (low level)	0	0	0	0	0	0	**10**	0	0
Δ RH (high level)	0	0	0	0	0	0	0	**10**	0
Δ RH (low level)	0	0	1	0	0	0	0	0	4

5. Conclusions

In this paper, we have described an application-oriented diagnosis tool designed for the identification of the operating regimes of FC systems. The proposed method is based on signal analysis and information processing techniques. It involves four main steps: experimental tests for data collection (constitution of a reference experimental database, linked to different FC regimes), analysis of the morphology of the stack voltage signal and extraction of signatures (envelopes and instantaneous frequencies), calculation and selection of statistical descriptors for the constitution of a reduced database useful for the diagnosis (application of the mRMR method), and finally classification of the new observations to identify the regime (by pattern recognition, e.g. using SVMs). The tool has been applied to two FCs and the results obtained indicate that it allows the discrimination of a wide range of operating regimes. Correct classification rates close to 98% are obtained for both stacks.

To summarise, the tool has the following properties:

(i) It is non-intrusive as it is based on the sole measurement of the voltage measured at the terminals of the full stack.

(ii) It uses minimal instrumentation (only one voltage sensor is needed to monitor the FC state-of-health, unlike other existing diagnostic methods which rely on more sophisticated instrumentation). It does not require any specific external excitation of the system under study.

(iii) It relies on mathematical tools that can lead to limited computation times, especially with regard to the generation of the signatures reflecting the morphology of the FC voltage. Envelopes and instantaneous frequencies are well-suited to describe dynamic time signals in both the time and frequency domain. They are also less complex to implement than wavelets or singularity spectra.

(iv) It can be implemented in a "preventive" framework because it allows the detection of small deviations from the nominal operating conditions.

(v) It can identify a wide range of faults with a very good accuracy. We have thus been able to consider: some variations in gas flow rates and pressures, some changes in FC temperature and humidity levels, the presence of carbon monoxide at anode, and even the combinations of simultaneous faults. However, in this case, an extended learning base is required.

(vi) It is generic in nature and can potentially be used for a large number of systems regardless of type, geometry, size or application. The results of the classification of a voltage data set, based on the experiments carried out on two stacks from two different manufacturers and intended for different applications (transport and stationary), show the effectiveness and portability of the developed diagnostic strategy.

Acknowledgments:

The work performed was done within the French ANR project, titled “Decentralized energy production”, directed by EFFICACITY, the French R&D Institute for urban energy transition, and Université Gustave Eiffel (former IFSTTAR).

References

[1] Y. Zhang, J. Jiang. Bibliographical review on reconfigurable fault-tolerant control systems. Annual Reviews in Control 2008;32:229-252.

[2] C. Lebreton, M. Benne, C. Damour, N. Yousfi-Steiner, B. Grondin-Pérez, D. Hissel, J.-P. Chabriat. Fault tolerant control strategy applied to PEMFC water management. Int. J. of Hydrogen Energy 2015;40(33):10636-10646.

[3] D. Hissel, M.-C. Péra. Diagnostic & health management of fuel cell systems: Issues and solutions. Annual Reviews in Control 2016;42:201-211.

[4] J. Wang, B. Yang, C. Zeng, Y. Chen, Z. Guo, D. Li, H. Ye, R. Shao, H. Shu, T. Yu. Recent advances and summarization of fault diagnosis techniques for proton exchange membrane fuel cell systems: A critical overview. J. of Power Sources 2021;500:229932.

[5] M.G. Santarelli, M.F. Torchio, P. Cochis. Parameters estimation of a PEM fuel cell polarization curve and analysis of their behavior with temperature. J. of Power Sources 2006;159(2):824-835.

[6] C-Y. Lim, H. Haas. Diagnostic method for an electrochemical fuel cell and fuel cell components. Patent US20060051628(A1). https://patents.google.com/patent/US20060051628A1/en (last access: 9/9/2021).

[7] W.J. Wruck, R.M. Machado, T.W. Chapman. Current interruption-instrumentation and applications. J. of Electrochemical Society 1987;134(3):539-546.

[8] E. Ivers-Tiffe, A. Weber, A. Schichlein. Electrochemical impedance spectroscopy. In: W. Vielstich, H.A. Gasteiger, A. Lamm. Handbook of fuel cells: fundamentals, technology and applications. Vol. 2. New York: Wiley, pp. 220-235, 2003.

[9] X. Zhang, T. Zhang, H. Chen, Y. Cao. A review of online electrochemical diagnostic methods of on-board proton exchange membrane fuel cells. Applied Energy 2021;286:116481.

[10] R. Petrone, Z. Zheng, D. Hissel, M.-C. Péra, C. Pianese, M. Sorrentino, M. Becherif, N. Yousfi-Steiner. A review on model-based diagnosis methodologies for PEMFCs. Int. J. of Hydrogen Energy 2013;38(17):7077-7091.

[11] X. Yuan, H. Wang, J.C. Sun, J. Zhang. AC impedance technique in PEM fuel cell diagnosis - A review. Int. J. of Hydrogen Energy 2007;32:4365-4380.

[12] H. Yuan, H. Dai, X. Wei, P. Ming. Model-based observers for internal states estimation and control of proton exchange membrane fuel cell system: A review. Journal of Power Sources 2020;468:228376.

[13] N. Yousfi-Steiner, D. Hissel, P. Moçotéguy, D. Candusso. Non-intrusive diagnosis of polymer electrolyte fuel cells by wavelet packet transform. Int. J. of Hydrogen Energy 2001;36(1):740-746.

[14] N. Yousfi-Steiner, P. Moçotéguy, L. Gautier, D. Hissel, D. Candusso. Detection of defects in an electrochemical device. Patent WO/2010/149935 A1. https://patents.google.com/patent/WO2010149935A1/en (last access: 9/9/2021).

[15] D. Benouioua, D. Candusso, F. Harel, L. Oukhellou. PEMFC stack voltage singularity measurement and fault classification. Int. J. of Hydrogen Energy 2014;39(36):21631-21637.

[16] D. Benouioua, D. Candusso, F. Harel, P. Picard, X. François. On the issue of the PEMFC operating fault identification: Generic analysis tool based on voltage pointwise singularity strengths. Int. J. of Hydrogen Energy 2018;43(25):11606-11613.

[17] Z. Zheng, R. Petrone, M.-C. 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. Int. Journal of Hydrogen Energy 2013;38 :8914-8926.

[18] Z. Zheng, S. Morando, M.-C. Péra, D. Hissel, L. Larger, R. Martinenghi, A. B. Fuentes. Brain-inspired computational paradigm dedicated to fault diagnosis of PEM fuel cell stack. Int. J. of Hydrogen Energy 2017;42(8):5410-5425.

[19] D. Benouioua, D. Candusso, F. Harel, P. Picard. Method for determining the operating state of a system, method for configuring a classifier used to identify such a state, and device for determining such a state. Patent WO2019048803A1. https://patents.google.com/patent/WO2019048803A1/en. Last access: 9/9/2021.

[20] D. Hissel, M.-C. Péra, D. Candusso, F. Harel, S. Bégot. Characterization of polymer electrolyte fuel cell for embedded generators. Test bench design and methodology. In: Zhang Xiang-Whu, editor. Chapter of advances in fuel cells, research Signpost. North Carolina State Univ; 2005. p. 127-148.

[21] B. Boashash. Estimating and Interpreting the Instantaneous Frequency of a Signal - Part 1: Fundamentals. Proceedings of the IEEE, Vol. 80, Nr. 4, pp. 520-538, 1992.

[22] H. Bukac. Instantaneous Frequency: Another Tool of Source of Noise Identification. Int. Compressor Enginnering Conference, Purdue University, USA, 2004.

[23] J-C. Cexus. Analyse des signaux non-stationnaires par transformée de Huang, opérateur de Teager-Kaiser et transformation de Huang-Teager (THT). PhD Thesis, Univ. de Rennes, 2005.

[24] B. Picinbono, W. Martin. Représentation des signaux par amplitude et phase instantanées. Annales des Télécommunications 1983;38:179-190.

[25] P. Delachartre, D. Vray. Analyse temps fréquence des signaux. Lecture at INSA Lyon - CREATIS Lab. https://www.creatis.insa-lyon.fr/~vray/doc_cours/temps-frequence_expose_3_Slides.pdf. Last access: 9/9/2021.

[26] H. Peng, F. Long, C. Ding. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on Pattern Analysis and Machine Intelligence 2005;27(8):1226-1238.

[27] C. Ding, H. Peng. Minimum redundancy feature selection from microarray gene expression data. J. of Bioinformatics and Computational Biology 2005;3(2):185-205.

[28] Hanchuan Peng's web site. Information on mRMR (minimum redundancy maximum relevance feature selection). 2017. http://home.penglab.com/proj/mRMR/.

[29] J-M. Brossier. Théorie de l'information. Lecture at ENSIMAG, Grenoble, France 2014. http://www.gipsa-lab.grenoble-inp.fr/~jean-marc.brossier/TheorieInformation-Ensimag-2014.pdf. Last access: 9/9/2021.

[30] Z. Li, R. Outbib, S. Giurgia, D. Hissel, S. Jemeï, A. Giraud, S. Rosini. Online implementation of SVM based fault diagnosis strategy for PEMFC systems. Applied Energy 2016;164:284-293.

[31] K. Chen, S. Laghrouche, A. Djerdir. Aging prognosis model of proton exchange membrane fuel cell in different operating conditions. Int. J. of Hydrogen Energy 2020;45(20):11761-11772.

[32] R. Rakotomalala. SVM - Support Vector Machine. Lecture at Univ. Lumière Lyon 2, France, 2017. https://eric.univ-lyon2.fr/~ricco/cours/slides/svm.pdf. Last access: 9/9/2021.

[33] C-M. Bishop. Pattern Recognition and Machine Learning. Springer, 2007.

[34] C. Cortes, V. Vapnik. Support-Vector Network. Machine Learning 1995;20(3):273-297.

[35] C-C. Chang, C-J. Lin. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2011, 2:27:1--27:27. Software available at: https://www.csie.ntu.edu.tw/~cjlin/libsvm/. Last access: 9/9/2021.