Analysis of PEM fuel cell experimental data using Principal Component Analysis and Multi linear regression

Latevi Placca, Raed Kouta, Denis Candusso, Jean-François Blachot, Willy Charon

• L. Placca, R. Kouta, D. Candusso, J-F. Blachot, W. Charon (mai 2010). *Analysis of PEM fuel cell experimental data using Principal Component Analysis and multi-linear regression*. International Journal of Hydrogen Energy. Vol. 35, n°10, pp. 4582-4591. Ed. Elsevier.

Analysis of PEM fuel cell experimental data using Principal Component Analysis and Multi linear regression

Latevi Placca[a, b, c, *], Raed Kouta[a, b], Denis Candusso[a, d], Jean-François Blachot[a,c], Willy Charon[a,b]

[a]FC LAB., Fuel Cell System Laboratory, Rue Thierry Mieg, 90000 Belfort, France.

[b]M3M research laboratory, University of Technology of Belfort-Montbéliard, 90010 Belfort, France.

[c]CEA, LITEN, 17, Rue des Martyrs - 38000 Grenoble, France.

[d]INRETS, The French National Institute for Transport and Safety Research, Laboratory of New Technologies (LTN), 25 Allée des Marronniers, 78000 Versailles - Satory, France.

[*]Corresponding author: Phone :(+33)384583654; fax :(+33)384583636

E-Mail addresses: latevi-atatoe.placca@utbm.fr (L. Placca), raed.kouta@utbm.fr (R. Kouta), denis.candusso@inrets.fr (D. Candusso), jean-francois.blachot@cea.fr (J.-F. Blachot), willy.charon@utbm.fr (W. Charon)

Abstract:

Polarisation curves performed at the Fuel Cell System Laboratory (FC LAB) at Belfort on a PEM fuel cell stack using a homemade fully instrumented test bench led to more than 100 variables depending on time. Visualising and analysing all the different test variables are complex. In this work, we show how the Principal Component Analysis (PCA) method helps to explore correlations between variables and similarities between measurements at a specific sampling time (individuals). To complete this method, an empirical model of the PEM fuel cell is proposed by linking the different input parameters to the cell voltage using Multiple Linear Regressions.

Keywords: Proton exchange membrane (PEM) fuel cell; Principal Component Analysis (PCA); Multiple Linear Regression; statistical analysis.

Contents

1 Introduction

Proton exchange membrane (PEM) fuel cells are considered to be the most promising fuel cell technology for transportation applications due to their low operating temperature and pressure resulting in a possible quick start-up [1]. However, to implement them in transportation systems, their durability and their reliability should be improved.

For this purpose, many studies were done on the durability of PEM fuel cells (PEMFC). W. Schmittinger and A. Vahidi [2] made a review of the main parameters influencing long term performance and durability of PEMFC. J. Wu et al. [3] summarised the different degradation mechanisms and mitigation strategies on a PEMFC. S. Zhang et al. [4] published the accelerated stress tests of membrane electrode assembly (MEA) durability in PEMFC. P. Pei et al. [5] used electrochemical surface area, internal resistance, particle size and hydrophobic nature of catalyst measurements to analyse an automotive PEMFC stack after 500 h accelerated lifetime test. Then, N. Yousfi-Steiner et al. [6] studied the PEMFC degradation and starvation issues and Z.-B. Wang et al. [7] made durability studies on performance degradation of Pt/C catalysts of PEMFC. Recently, Z.-M. Zhou et al. [8] evaluated and compared the durability of Pt-Pd/C and Pt/C catalysts using physical and electrochemical techniques. The Pt-Pd/C catalysts showed a better durability than the Pt/C ones. G. Chen et al. [9] developed an effective ex-situ method for characterising electrochemical durability of a gas diffusion layer under simulated PEMFC conditions.

Concerning fuel cell reliability, M. Tanrioven and M.S. Alam [10] developed a state-space generation model for a stand-alone PEM fuel cell. V. Mangoni et al. [11] proposed a reliability model of fuel cell by using a probabilistic approach considering the lack of data and the uncertainties of design. C. Wieland et al. [12] modelled PEMFC stacks reliability data through Petri nets and M. Gerbec et al. [13] made operational and process-safety analysis on a commercial PEMFC system.

In addition to durability and reliability studies, statistical methods were applied to PEMFC to optimise parametric performance. W.-L. Yu et al. [14] and B. Wahdame et al. [15] used “Design of Experiment” (DoE) techniques. In the first work, DoE was used to obtain optimal combination of the operating parameters (fuel cell operating temperatures, operating pressures, anode and cathode humidification temperatures, anode and cathode stoichiometric flow ratios) to improve the performance of a PEMFC. In the second work, DoE is applied to predict PEMFC stack voltage, fuel consumption, maximal electrical power and stack lifetime. I. Mohamed et al. [16] used genetic algorithms to maximise output power delivered by a PEM fuel cell stack, searching for the best configuration in terms of number of cells and cell surface area. Then, B. Wahdame et al. [17] used the response surface methodology to analyse and compare the results of two durability experiments. Their objective was to optimise the fuel cell operating conditions versus ageing time. S. Kaytakoglu and L. Akyalçin [18] used the Taguchi method to determine optimum working conditions for maximum power density of a fuel cell. S.-J. Wu et al. [19] used both Taguchi and neural network methods for PEMFC optimisation: the Taguchi method acquires the primary optimums of the operating parameters and the neural network constructs relationships between the control factors and the responses. A. Mawardi and R. Pitchumani [20] developed a

sampling-based stochastic model to analyse the effects of parameter uncertainty on the performance variability of PEMFC.

The present work is a contribution to statistical analysis of PEMFC data by analysing the effects of groups of test parameters on the performance of a fuel cell and by correlating the different parameters using "Principal Component Analysis" (PCA) on experimental data. With the results of this first analysis, an empirical model is deduced using multi-linear regression.

This paper is organised as follows: section 2 introduces the experimental set-up, the conditions of the tests, and the experimental data available for the analysis. Then principal component analysis is described and applied in section 3. Finally, multi linear regression results and discussion are presented in section 4.

2 Polarisation curve measurement

2.1 Experimental set-up

The PEMFC stack used for the tests is a 2.5 kW stack. With an active cell area of about 150 cm², the round fifty-cell stack has been assembled with commercially available membranes, gas diffusion layers (GDLs) and machined graphite flow distribution plates. The fuel cell can operate from atmospheric pressure to about 3 bars abs.

The test bench consists of a hydrogen line, an air line, temperature and gas humidification subsystems, a programmable electronic load, a control process and an interface (control software developed with Labview^{TM}). The temperature of the stack is controlled using a cooling water circuit which includes a pump, a proportional three–way valve, a cold exchanger and a heater. Another water circuit (iced water) is used to condense the water in the reactive gas streams at stack outlets. The whole test bench is operated using software setting the operating conditions: control of the gas flows and/or pressures, gas humidity levels, stack and gas temperatures, load current profiles, and normal or emergency shutdown. A detailed description of a similar fuel cell test bench can be found in references [21] and [22]. Below, Table 1 presents the different abbreviations used for the test data from the sensors, and their separation into inputs, outputs and others.

Table 1: List of parameters and their classification

PARAMETERS	*NOMENCLATURE*	*TYPE OF PARAMETERS INPUT: in OUTPUT: out OTHER: oth*
t	Time [min]	in
U_i	Voltage of cell *i* [V]	out
TeH2	Temperature at stack inlet (anode side)[°C]	in
TsH2	Temperature at stack outlet (anode side) [°C]	out
TeAIR	Temperature at stack inlet (cathode side) [°C]	in
TsAIR	Temperature at stack outlet (cathode side) [°C]	out
TeEAU	Temperature at stack inlet (water loop) [°C]	in
TsEAU	Temperature at stack outlet (water loop) [°C]	out
PeAir	Pressure at stack inlet (cathode side) [bar rel.]	out
PsAir	Pressure at stack outlet (cathode side) [bar rel.]	in
PeH2	Pressure at stack inlet (anode side) [bar rel.]	out
PsH2	Pressure at stack outlet (anode side) [bar rel.]	in
DeAIR2000	Air flow in the air flow controller ranging from 0 to 2000 Nl/min	in
TinechH2	Temperature at the inlet of the condenser located at the anode stack outlet [°C]	in
TinechAIR	Temperature at the inlet of the condenser located at the cathode stack outlet [°C]	in
Courant	Load current [A]	in
ToutechAIR	Temperature at the outlet of the condenser located at the cathode stack outlet [°C]	out
DeltaPbanc	Absolute difference between PeAir and PeH2	oth
ToutechH2	Temperature at the outlet of the condenser located at the anode stack outlet [°C]	out
Ustack	Stack voltage [V]	out
PeEau	Pressure at stack inlet (water loop) [bar]	out
PsEau	Pressure at stack outlet (water loop) [bar]	out
DEau	Water flow in the cooling circuit (water loop) [l/min]	in
Patmo	Atmospheric pressure in the test room [mbar abs.]	in
DeH2_500	H2 flow in the flow controller ranging from 0 to 500 Nl/min	in
TR500	Iced water temperature at the inlet of air condenser [°C]	in
TR501	Iced water temperature at the outlet of air condenser [°C]	out
TR502	Iced water temperature at the inlet of H2 condenser [°C]	in
TR503	Iced water temperature at the outlet of H2 condenser [°C]	out
Tenceinte	Temperature in the test room [°C]	in
TCH01 … TCH14	Temperatures of the humidification systems [°C]	See below
TCH01 and TCH09	Temperatures of the air and H2 vaporisers [°C]	in
TCH02	Temperature at outlet of the H2 vaporiser [°C]	out
TCH03	Temperature in the air humidification column [°C]	in
TCH04	Temperature on the external surface of the humidification air column [°C]	in
TCH05	Temperature on the mirror hygrometer (external surface of the device box) [°C]	in
TCH06	Temperature of air at the outlet of the air humidification column [°C]	out
TCH07	Temperature of air pipe (surface) at the outlet of the air humidification column (near TCH06) [°C]	out
TCH08	Temperature in the H2 humidification column [°C]	in
TCH10	Temperature on the external surface of the H2 humidification column [°C]	in
TCH11	Temperature of air at the outlet of the H2 humidification column [°C]	out
TCH12	Temperature of H2 pipe (surface) at the outlet of the H2 humidification column (after TCH11) [°C]	out
TCH13	Temperature of air pipe (surface), stack inlet side [°C]	in
TCH14	Temperature of H2 pipe (surface), stack inlet side [°C]	in
HR100	Relative humidity at air inlet [%]	in
HR100FC	Relative humidity of air computed for the fuel cell temperature [%]	oth
HR101	Relative humidity at air outlet [%]	out
HR400	Relative humidity at H2 inlet [%]	in
HR400FC	Relative humidity of H2 computed for the fuel cell temperature [%]	oth
RL200	Water flow controller [g/h] for water injection into the air humidification column	in
RL201	Water flow controller [g/h] for water injection into the H2 humidification column	in
QL501	Iced water flow [l/min]	in
Icons	Current reference [A]	oth
Tpilecons	Stack temperature reference [°C]	oth
Igazcons	Flow reference for H2 and air ; only one value [expressed in A] computed from FSA and FSC	oth
DeH2_500cons	H2 flow reference in the flow controller ranging from 0 to 500 Nl/min	oth
DebAIR2000cons	Air flow reference in the flow controller ranging from 0 to 2000 Nl/min	oth

TtraceAIRcons	Temperature reference for heating cables on air line after humidification system [°C]	oth
QH2oAIRcalccons	Water flow for injection into air column (theoretical value) [g/s]	oth
RL200cons	Reference of water flow [g/h] for water injection into the air humidification column	oth
QH2oH2calccons	Water flow for injection into H2 column (theoretical value) [g/s]	oth
RL201cons	Reference of water flow [g/h] for water injection into the H2 humidification column	oth

2.2 Test methodologies

In order to avoid as far as possible any impact of recent events on the characterisation measurement, before any polarisation test, the fuel cell was operated during about 20-30 min at nominal steady-state conditions: 2.5kW electrical power.

After the conditioning procedure, two polarisation curves were recorded: by increasing and by decreasing the load current. The methodology used for the polarisation curve measurements was the following one: the polarisation curve record was performed by decreasing or incrementing the fuel cell current gradually by steps of 10A in the load current range: 0-100A (Fig. 1). Therefore, there are 11 load current stages for both loading cases. Before each current step, the gas flows, calculated from the stoichiometry rates corresponding to the new current value, were first prepared; they supplied the stack during a period of about 10 seconds. Therefore, the stack never starves. Then, the fuel cell had to supply the new current during the next 3 minutes (this value is approximate: the stationarity of the fuel cell performances may be reached more or less rapidly). 10 seconds after the end of this last stage, the new gas flows were prepared for the next current step. For currents below 20A, minimum flows were used. These minimum flows are the ones that correspond to 20A to ensure a better draining of the water out of the stack and thus to enable constant operation (without fluctuation) at low power output. All these duration parameters related to the polarisation curve procedure could be specified via the human-machine interface software developed in-lab. The choice of the duration parameters was made in order to obtain stationary conditions on each load step. When any cell voltage reached the minimal threshold cell voltage (200mV) or the stack voltage was lower than 15V, the polarisation curve test was stopped.

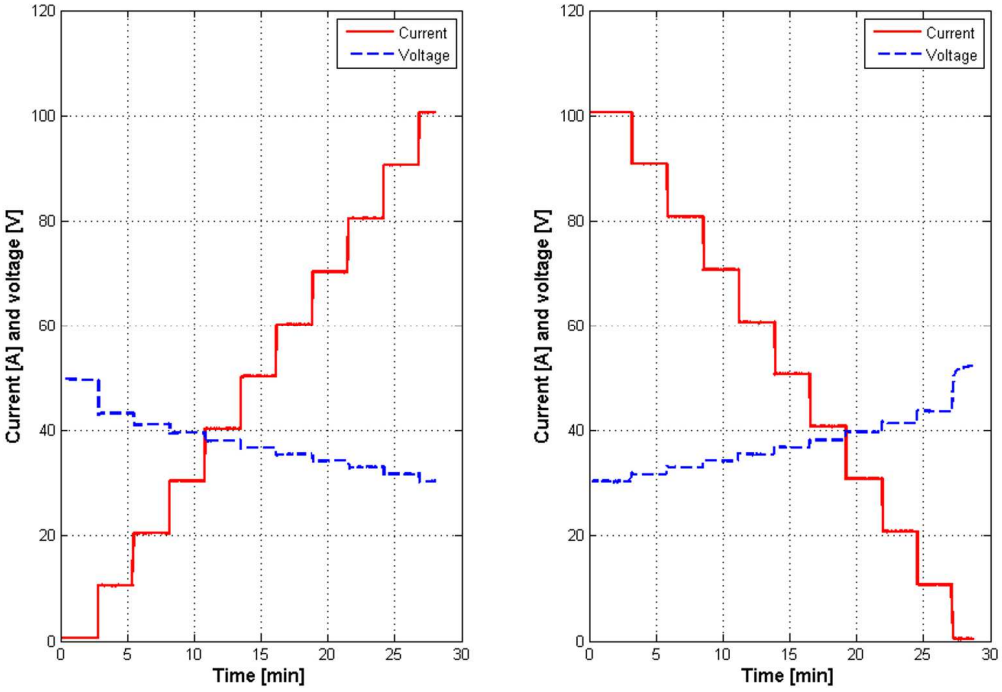

Fig. 1: Polarisation curve in increasing current (left) and in decreasing current (right) at 70% of Relative Humidity and a pressure of 1 bar abs.

Polarisation curves have been recorded with Relative Humidity (RH) rates of 70% and of 90% at both anode and cathode electrodes, and with a pressure of 1 bar and 2 bars by increasing and then decreasing the load. The huge amount of recorded data requires efficient methods of analysis. The Principal Component Analysis (PCA) was detected as having a lot of capabilities for providing high interesting information about the relationships between system variables. In fact, PCA is a method for reducing the dimensionality of a data set consisting of a large number of correlated variables while retaining as much as possible of the variation present in the data set. The aim of this work is to illustrate the PCA, and in the next section its presentation and implementation has been limited to the case of RH of 70% at both anode and cathode electrodes and a pressure of 1 bar in increasing and decreasing load modes.

3 Principal Component Analysis (PCA), results and discussion

The first part of this section is a short introduction to the PCA concepts. The objective is to show how the procedure can be used to perform the analysis of the fuel cell experiments. The second and third parts focus, respectively, on the implementation of the method on the total data (global PCA) and on one specific load current stage (local PCA), both for increasing and decreasing loads.

3.1 Short introduction to the PCA concepts

In the field of hydrogen energy, L. Trevani et al. [23] used PCA on UV-spectra, obtained over a very wide range of solution compositions and temperature to determine cumulative formation constants of Cu^{2+}(aq) complexes with Cl^{-}(aq). They found that PCA suggested 5 factors to be the minimum number required to achieve a description that is comparable with their experimental uncertainty. Then, A.A. Abreu et al. [24] used PCA to visualise the main differences between 4 biomass tests corresponding to different anaerobic sludges with different pHs ranging from 4.5 to 8.0. Their data set consisted of 13 variables and 32 samples.

The PCA determines the correlations between parameters and similarities between individuals. Table 1 distinguishes different parameters types (inputs, outputs, and other). Before performing a PCA, the variables need to be centred (relative coordinates) and scaled to 1 because they have different units. "Individual" is a very important concept in the PCA. In the real context, an individual consists of the set of measurements at a specific sampling time. The individuals are numbered from 1 to N corresponding to the first, the second,…, the N^{th} sampling time.

PCA defines new variables called principal components (PCs) that are linear combinations of the original variables. Fig. 2 shows the variances associated to the PCs. The variance analysis will be described later. The new variables are of unit length and orthogonal to each other. The correlation matrix between variables computed over the acquisition time is diagonalised. The values on the diagonal, i.e. the principal components, are the matrix eigenvalues and are ordered by decreasing values. The eigenvectors determine the linear combinations. The orthogonality of the eigenvectors is a consequence of the symmetry of the correlation matrix. Thus, principal components are not correlated over the acquisition time: this is the aim of the diagonalisation. The measurements can be represented in the hyperspace of the variables or in the new hyperspace of the principal components: it is simply a variable change. Further details about the methodology of the PCA can be found in [25].

The PCs show very interesting properties as demonstrated hereafter. Two 2D graphs are particularly important. Let us take Fig. 3 as an example of the representation. The typical results will be described later in this paper.

The first graph concerns the representation of the individuals in the hyperspace of the first and of the second principal axis. Individuals are distinguished on the graph with a number. The graph shows the evolution of the PCs with the acquisition time.

The second graph concerns the participation of the original variables within the value of the PCs. It can reveal the variables that are of importance in the physical behaviour.

It is interesting to display the two graphs on the same page. The first graph with the individual is set on the left side and the graph with the variables on the right side. On the left side graphs, the individuals show a group organisation. Three kinds of information are of importance: the behaviour of the group of individuals (homogeneous or not), the progression from the first individual to the last one and the transition between two groups of individuals (increasing, decreasing or steady). The right graph shows centred and reduced variables data projected onto the first two principal components. Due to the centring and scaling, the axes range from -1 to 1. The effective variables are separated into four groups numbered 1 to 4 that represent:

1. the most positive influencing variables according to the 1st PC,
2. the most negative influencing variables according to the 1st PC,
3. the most positive contributing variables according to the 2nd PC,
4. the most negative contributing variables according to the 2nd PC.

This representation helps to determine the variables that increase or decrease with the time studying the evolution of the PC that they affect mainly and their effect on that PC. It also identifies variables that have no effect on the individuals. To improve the visual analysis, the variables considered as "no-effect" are included in a green square.

The steps for performing the PCA are:

- Computation of the correlation matrix concerning the measurement variables.
- Solution of the eigenvalues problem which determines on the one side the linear combination coefficients defining the principal components and on the other side the variances of the principal components.
- Analysis of the variances of the principal components in order to check if they are statistically fitting in with a PCA. The total variance of the two main principal components has to be at least equal to 75 % to validate a PCA.
- Projection of the individuals on the two new "principal" axes.
- Elaboration of the two graph types.

Some properties have to be mentioned:

- A global PCA does not consider local behaviours. It needs to be completed by a local PCA on each step of the current load. In this paper, the analysis is limited to the first stage that can give information about open circuit voltage (OCV). This helps to observe the behavioural changes in variables during time for a constant current.
- Relationship between changes in variables and physical processes will be attempted to confirm or deny the conclusions made by the PCA.

3.2 Global PCA

The results shown in this paper concern a global PCA performed on the data corresponding to the case of RH equal to 70% at both anode and cathode electrodes and a pressure of 1 bar in increasing (case 1) and decreasing (case 2) loads. Both of these cases lead to similar results which are represented on Fig. 2. It comes out that the first two principal components explain more than 80 % of the total variance, so the paper focuses on the first two PCs. By analysing linear combinations associated to these PCs, it appears that many basic variables are linearly correlated and there are also correlations between parameters.

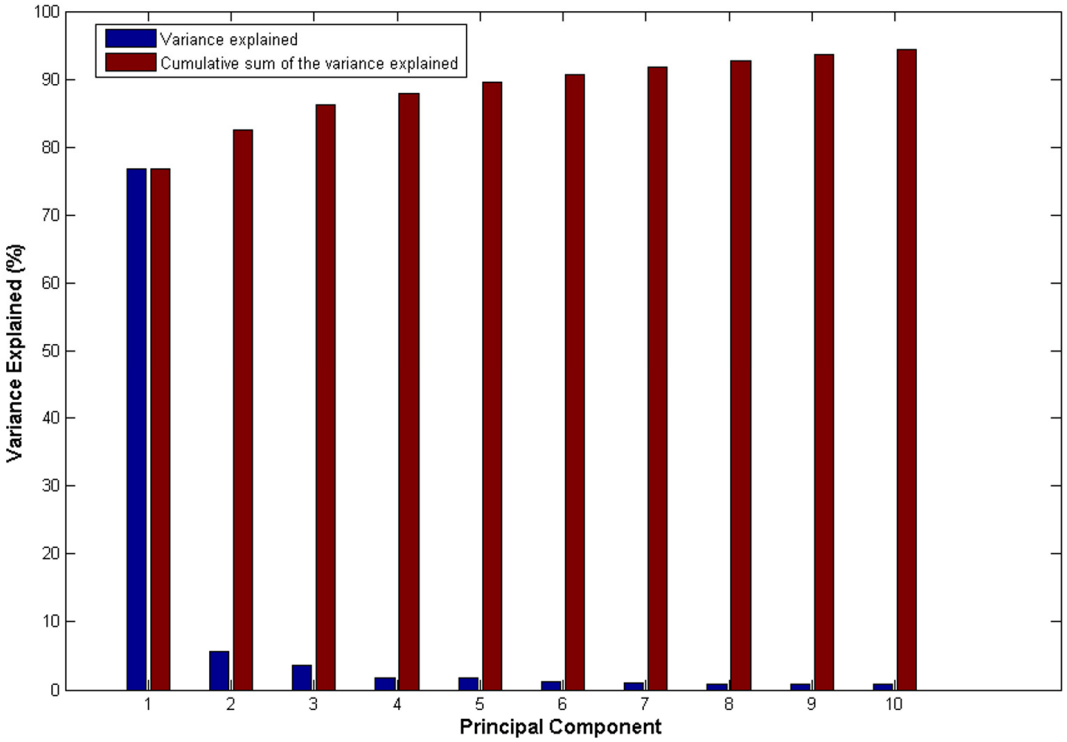

Fig. 2: Variance of principal components for the global PCA

Fig. 3 considers the increasing current case. The left side represents the individuals which are projected onto the first two principal axes. The right side shows the projection of the basic variables on the first two principal components axes.

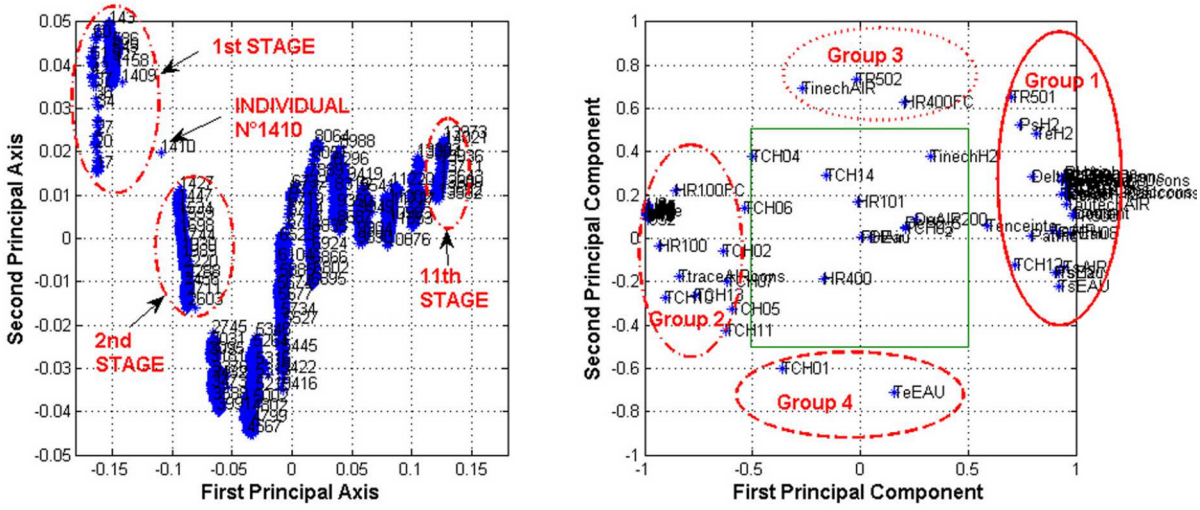

Fig. 3: PCA representation of individuals and variables in increasing load (case 1)

It can be observed on the left side of Fig. 3 that the individuals are separated into 11 groups corresponding to the 11 loading stages of current (cf Fig. 1). The first stage corresponds to OCV conditions. The figure shows a short transient state in the

beginning. The other groups are less spread and show more stable experimental conditions.

On the right side of Fig. 3, the 4 groups of variables can be identified. Considering the first principal component axis, the most contributing variables are

group 1: in the positive sense (+1): {DeltaPbanc, TCH12, TsAir, TsEau, TeAir, Patmo, ToutechAir, TCH09, RL200, RL200 cons, Deair 2000, DebAIR 2000 cons, Igazcons, DeH2_500, DeH2_500cons, PeH2, PsH2, TeH2, TR500,TR501, TR503, TR504, ToutechH2, QH20H2 calccons, QH20Aircalccons, RL 201, RL201cons, PeAIR, PeEAU, PsAIR, TCH08, Icons, courant}. See table 1 for variables meaning.

group 2: in the negative sense (-1): {HR100FC, HR100, TCH02, TtraceAIRcons, TCH10, TCH13, TCH11, TCH05, TCH07, Ustack, Ui}.

From Groups 1 and 2, it comes out that the current and the cell voltage have opposite actions. Indeed, an increase in the load current involves a decrease on the cell voltage.

In addition, it can be noticed that Deair 2000 and DeH2_500 which are the air flow and the H_2 flow measured upstream the fuel cell respectively, are members of Group 1 as the current. It confirms that an increase of the load current can only be achieved with higher reactant flows. Similarly, PeH2 and PeAIR which are the pressures at stack inlet at anode side and at cathode side respectively, belong to this group. That is due to the fact that higher flows lead to higher pressures at stack inlets during the polarisation curve test.

Considering the second principal component axis, the most contributing variables are

group 3, variables {TinechAIR, TR502, HR400FC}, tends to put the individuals up and the opposite,

group 4 enclosing {TCH01, TeEAU}, tends to put the individuals down.

The influence of this latter grows until the fourth stage. After this stage, group 3 becomes stronger and helps the individuals to step up until the sixth stage.

The moving of the individuals to the right is due to the group 1 members. Before the fourth stage, group 2 tries to slow the right moving of the individuals. It is really efficient between the first stage of individuals and the second one but less efficient after the fourth one. The fifth stage of individuals has the highest vertical variability along the second PC. It seems that this stage is balanced. The decreasing loading case leads to the same conclusions.

3.3 Local PCA

The results of the global PCA cannot show what is occurring at a specific stage. They need to be completed by a local PCA to learn more about parameters variations at a stage. In fact, the variations of the parameters of the cell can help to deduce physical phenomena occurring in the cell. For the local PCA, the correlation matrix, the eigenvectors and eigenvalues are recalculated. Thus, the PCs used for the local PCA change from the global case.

In addition to the first two principal components, a third principal component will be considered in order to improve the visual analysis. These three components are sufficient for this analysis as they represent almost 79% of the total variance (Fig. 4).

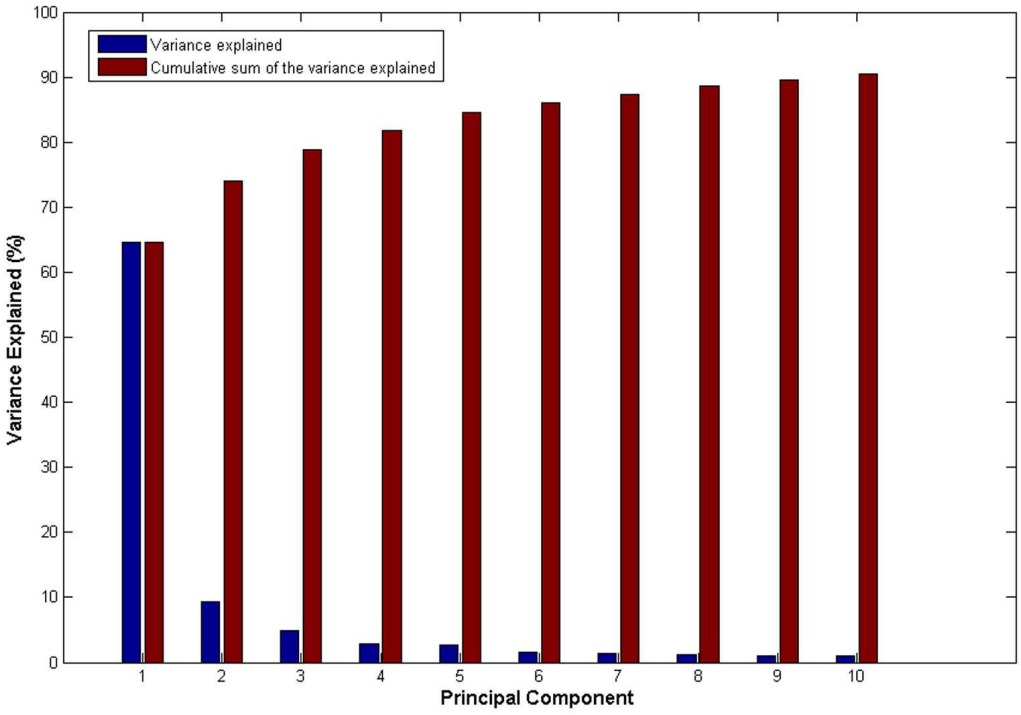

Fig. 4: Variance of principal components for the local PCA

To illustrate, local PCA is now presented on the first stage in increasing load. This period corresponds to OCV conditions whose study is interesting for example to analyse the MEA degradation [6].

Figures 5, 6 and 7 represent the PCA on the first and second, the first and third, and the second and third principal axes respectively. 3D graphs have been avoided as two dimensional plots are easier to describe and analyse.

Referring to the left side of Fig. 5, until the 80th individual (which is equivalent to a current time of about 7 seconds), the individuals are spread. In fact, this corresponds to a transient state of the cell. After this time, the graph is more centred and a steadier state occurs.

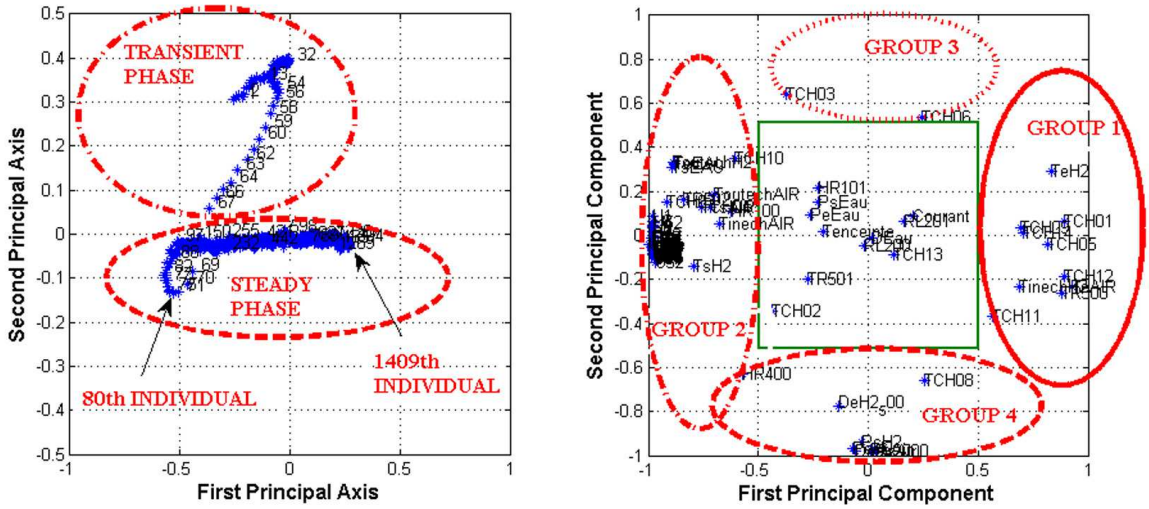

Fig 5: PCA on the first stage - First and second principal axes

With respect to the first PC, the most positive influencing variables for this stage are: {TinechH2, TCH09, TCH14, TeH2, TCH01, TCH05, TCH11, TCH12, TeAIR, TR500, TR503} which constitute group1. The most negative contributing ones are {U_i, Ustack, TsH2, TeEAU, ToutechH2, TsEau, TCH07, TR502, TCH04, TCH10, HR100, HR400, TsAIR, ToutechAIR, TinechAIR} (group 2). The load being equal to zero for this stage, the important parameters related to the voltage correspond to temperatures.

When considering the second PC, the most positive contributing variables are {TCH03 and TCH06} (group 3) and the most negative contributing ones {HR 400, TCH08, DeH2_500, PsH2, PeH2, DeAIR2000, PeAir, and PsAir} (group 4).

In the beginning, group 3 variables are very influential and decrease with time until the 80th individual. Then the influence of "group 4" variables becomes more important up to the balance phase where the two groups compensate each other. Finally, groups 1 and 2 variables contribute to the uniform aspect of individuals at the steady phase. This is coherent with physical experimental results. For example, TsEAU which is the temperature at the stack outlet is part of group 2 and is influential at the steady phase during this first stage.

The consideration of the first and third principal axes on Fig. 6 completes these results. It confirms the separation of the individuals between a transient state period of time and a steady state around a current time of about 7 seconds. According to the first PCA axis, the positively and negatively contributing variables constitute the same group 1 and group 2 as previously. With regard to the 3rd principal component axis, there are no influential negative contributors and the most positive contributors are gathered in a new group 3: {TR501, TinechAIR, ToutechAIR}. This new group 3 increases with time after the start up transient period.

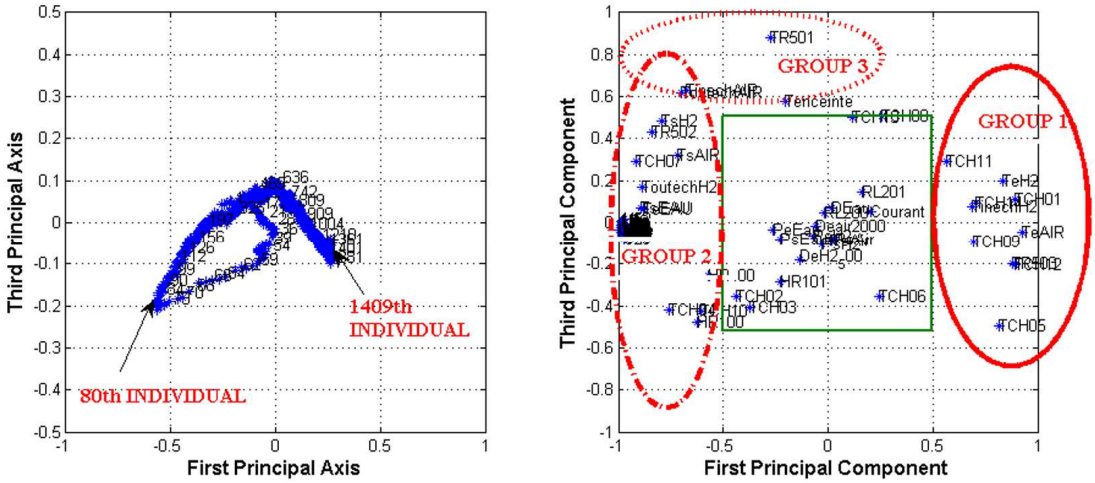

Fig 6: PCA on the first stage - First and third principal axes

The last projection to be taken into account is the 2nd - 3rd PC axes on Fig. 6.

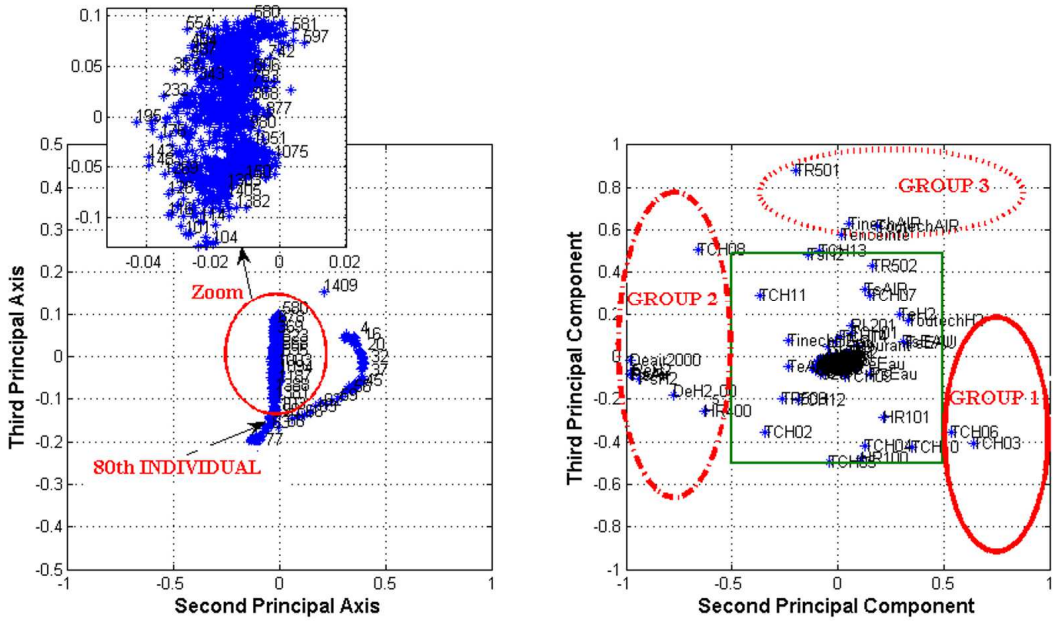

Fig 7 PCA on the first stage - Second and third principal axes

The occurrence of a short transient state followed by a more stable one is confirmed by Fig. 7. By considering the second principal component axis, the positively and the negatively contributing variables are the same group 3 and group 4 of Fig 5 respectively. They are named group 1 and group 2 on Fig. 7 to keep the same notation applied for other figures. With regard to the third principal component axis, the most positive contributing variables are the same group 3 of Fig. 6 and there are no negatively contributing variables as mentioned by Fig. 6. This is coherent with the previous results.

From Fig. 5, 6 and 7, the transient phase from the early individuals is due to the most positive contributing variables to the second PC and the second axis negative contributors move this transient state to a homogeneous state. This homogeneous state is characterised by the positive first axis contributors increment, the positive third axis contributors initial increment and later decrement, and the steadiness of the second PC.

The PCA shows the most contributing variables to the current and to the voltage. This can help to understand physical processes occurring in the cell and to improve the placement of the sensors. For example, TCH04 and TCH10 which are temperatures of external surface of the air and H_2 humidification air subsystems have opposite actions with the temperatures TCH01, TCH05, TCH09, TCH12 and TCH14. PCA can help to check the measures of the different sensors. Misbehaviours or badly located sensors can be detected and then replaced or repositioned.

Generally, to deduce the groups of correlated variables, experimenters used to represent all the variables versus current time. With PCA, in one or two graphs, correlated variables can be quickly detected.

The principal advantage of PCA is to show the correlations between the expected performance (Ustack, Ui) and all other parameters (inputs, outputs, other parameters) taking into account the time dependence in the experiments.

To complete the time dependent variables analysis, we suggest to use a mathematical model that does not consider the temporal evolution of parameters. This mathematical modelling is based on multi linear regression which is an efficient method that can be combined with PCA. This efficiency is due to the possibility to study effects of parameters on the performance from the modelling. The following section introduces multi linear regression analysis and its results.

4 Multiple Linear Regression, results and discussion

A multi linear regression attempts to model the relationship between two or more variables and a response by fitting a linear or a quadratic equation to observed data. Reference [26] details the implementation of this statistical method.

In this study, the variables are the input parameters of Table 1 (excluding time) and the response is at first the voltage of the stack and then the voltage of each cell in order to classify cells according to their delivered voltage.

A quadratic model (Eq. (1)) incorporating constant, linear, interaction and squared terms of variables was built on the first stage of the increasing loading cases.

$$y = \alpha_0 + \sum_{i=1}^{p} \alpha_i x_i + \sum_{\substack{j=1 \\ j \neq k}}^{p} \sum_{k=1}^{p} \beta_{jk} x_j x_k + \sum_{i=1}^{p} \gamma_i x_i^2 \qquad (1)$$

Where y is the voltage of a cell or the stack

$\alpha_i, \beta_{jk}, \gamma_i$ *are coefficients*

x_i *are the inputs considered*

p is the total number of inputs considered

The 29 input parameters involve 465 coefficients for each cell. For this reason, the whole table of coefficients is not given. Nevertheless, Table 2 presents the linear coefficients (α_i) of the dependence of the stack voltage on factors on the first stage in increasing load as for the local PCA. It shows as expected that coefficients of some inputs that increase with the voltage of the stack according to the local PCA (TeEAU, TCH04, TCH10, HR100) are positive. Thus, some inputs that decrease with the voltage of the stack (TeH2, TeAir, TinechH2, TCH01, TCH14) have negative coefficients.

Table 2: Linear coefficients of the dependence of the stack voltage on factors on the first stage in increasing load

INPUTS	*LINEAR COEFFICIENTS (α_i) OF THE REGRESSION*
TeH2	-7.69×10^{9}
TeAIR	-1.31×10^{9}
TeEAU	3.29×10^{9}
PsAir	2.98×10^{10}
PsH2	-1.53×10^{10}
DeAIR2000	1.08×10^{10}
TinechH2	-4.96×10^{10}
TinechAIR	-1.86×10^{9}
Courant	2.79×10^{11}
DEau	7.36×10^{13}
Patmo	-6.80×10^{8}
DeH2_500	5.50×10^{9}
TR500	2.78×10^{9}
TR502	-3.91×10^{9}
Tenceinte	1.39×10^{10}
TCH01	-4.81×10^{9}

TCH09	4.59×10^{8}
TCH03	9.06×10^{9}
TCH04	-6.01×10^{9}
TCH05	3.66×10^{9}
TCH08	4.68×10^{10}
TCH10	5.17×10^{9}
TCH13	-6.74×10^{9}
TCH14	-1.43×10^{9}
HR100	-7.08×10^{9}
HR400	2.54×10^{9}
RL200	-8.28×10^{7}
RL201	-5.58×10^{8}
QL501	1.26×10^{11}

Fig.8 shows that the expected voltage with the multiple regression fits well with the real one. The multiple regression statistics R^2 can be used to show the robustness of the regression law and model. The multiple regression correlation coefficient (R^2) is the percentage of variance explained by the linear regression in a sample of data. The global R^2 statistics is greater than 90% for every cell and for the whole stack. It shows the accuracy of the modelling at this stage for each cell voltage and also for the stack voltage.

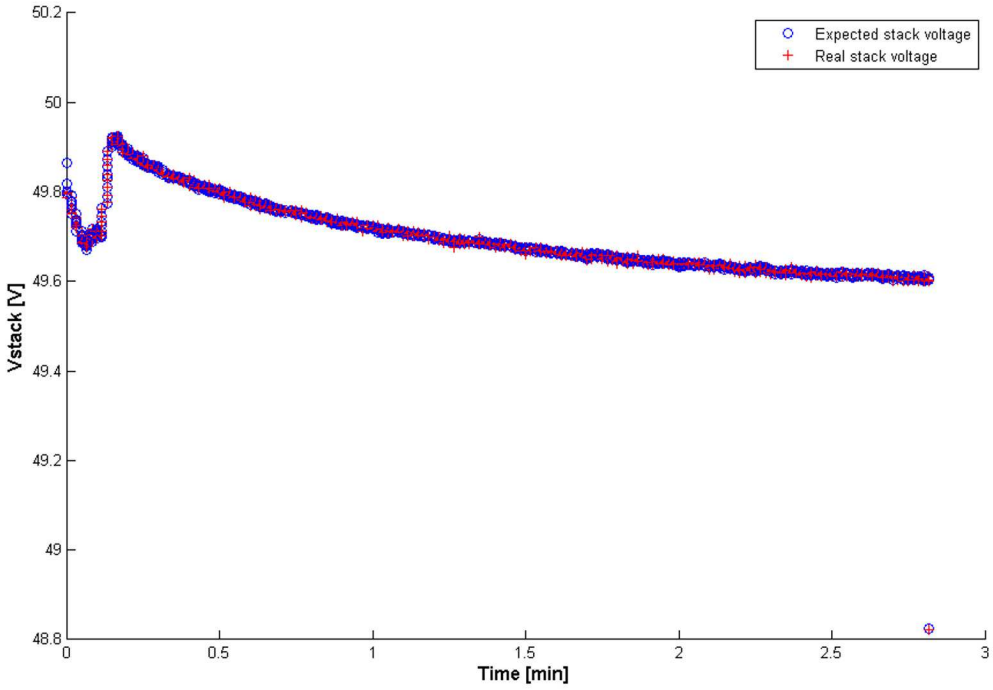

Fig 8 Multi-regression and real stack voltage graphs on the first stage in increasing load

5 Conclusion

The methods presented in this paper contribute on a very efficient way to the analysis and modelling of experimental data of a PEMFC. They are therefore very well adapted for fuel cell experimenters and designers. The Principal Component Analysis (PCA) helps to visualise the different parameters, their variations and their correlations with a minimum of graphs. It can be applied on the whole experimental data. The global PCA allows to classify groups of parameters which mostly contribute to the stack behaviour. The results show the 11 groups of individuals which correspond to the 11 stages of current load. This result helps to focus on specific stages of loading to run local PCA. We have limited this paper to the results of a PCA on the first load stage which confirms the transient physical phenomena involved during the fuel cell start-up period. The analysis on other stages can lead to fully new knowledge that will be presented in a next paper.

To complete the PCA, empirical models can be built from the experimental data using multi linear regression. As an example, each cell voltage of the stack has been modelled as an expression of the input parameters of the stack. The whole stack voltage is also linked to input parameters at the previous specific stages of PCA during increasing and decreasing loads.
For increasing load, the multi linear regression modelling has been proved to be accurate (robust) at the first load stage for each cell and for the stack voltage. It also confirms the weak influence of parameters interactions as shown by the PCA analysis:

the fact of having a large amount of the variance explained by the first principal components, which are a linear combination of parameters, means that the variance caused by parameter interactions (terms β_{jk} on the linear regression model) is low.

Here again a deep analysis can lead to interesting results about the correlations and interactions between variables, between inputs and outputs.

6 References

[1] J. Larminie, A. Dicks, Fuel Cell Systems Explained, Wiley, West Sussex England, 2000.

[2] W. Schmittinger, A. Vahidi, A review of the main parameters influencing long-term performance and durability of PEM fuel cells, J. Power Sources 180 (2008)1-14.

[3] J. Wu, X. Zi Yuan, J.J. Martin, H. Wang, J. Zhang, J. Shen, S. Wu, W. Merida, A review of PEM fuel cell durability: Degradation mechanisms and mitigation strategies, J. Power Sources 184 (2008) 104-119.

[4] S. Zhang, X. Yuan, H. Wang, W. Mérida, H. Zhu, J. Shen, S. Wu, J. Zhang, A review of accelerated stress tests of MEA durability in PEM fuel cells, Int. J. Hydrogen Energy 34 (2009) 388-404.

[5] P. Pei, X. Yuan, P. Chao, X. Wang, Analysis on the PEM fuel cells after accelerated life experiment, Int. J. Hydrogen Energy (2009), doi:10.1016/j.ijhydene.2009.09.103.

[6] N. Yousfi-Steiner, Ph. Moçotéguy, D. Candusso, D. Hissel, A review on PEM Fuel Cell catalyst degradation and starvation issues: causes, consequences and diagnostic for mitigation, J. Power Sources 194 (2009)130-145.

[7] Z.-B. Wang, P.-J. Zuo, Y.-Y. Chu, Y.-Y. Shao, G.-P. Yin, Durability studies on performance degradation of Pt/C catalysts of proton exchange membrane fuel cell, Int. J. Hydrogen Energy 34 (2009) 4387-4394.

[8] Z-M. Zhou, Z.-G. Shao, X.-P. Qin, X.-G. Chen, Z.-D. Wei, B.-L. Yi, Durability study of Pt-Pd/C as PEMFC cathode catalyst, Int. J. Hydrogen Energy, 35 (2010) 1719-1726.

[9] G. Chen, H. Zhang, H. Ma, H. Zhong, Electrochemical durability of gas diffusion layer under simulated proton exchange membrane fuel cell conditions, Int. J. Hydrogen Energy, 34 (2009) 8185-8192.

[10] M. Tanrioven, M.S. Alam, Impact of load management on reliability assessment of grid independent PEM fuel cell power plants, J. Power Sources 157 (2006) 401-410.

[11] V. Mangoni, M. Pagano, G. Velotto, Fuel cell reliability model based on uncertain data, International Conference on Clean Electrical Power, 2007. ICCEP '07. 21-23 May 2007 Page(s):730 – 735, IEEE 1-4244-0632-3/07, 2007.

[12] C. Wieland, O. Schmid, M. Meiler, A. Wachtel, D. Linsler, Reliability computing of polymer-electrolyte-membrane fuel cell stacks through Petri nets, J. Power Sources 190 (2009) 34-39.

[13] M. Gerbec, V. Jovan, J. Petrovcic, Operational and safety analyses of a commercial PEMFC system, Int. J. Hydrogen Energy 33 (2008) 4147-4160.

[14] W.-L. Yu, S.-J. Wu, S.-W. Shiah, Parametric analysis of the proton exchange membrane fuel cell performance using design of experiments, Int. J. Hydrogen Energy 33 (2008) 2311-2322.

[15] B. Wahdame, D. Candusso, X. François, F. Harel, J.M. Kauffmann, G. Coquery, Design of experiment techniques for fuel cell characterisation and development, Int. J. Hydrogen Energy 34 (2009) 967-980.

[16] I. Mohamed, N. Jenkins, Proton exchange membrane (PEM) fuel cell stack configuration using genetic algorithms, J. Power Sources, 131 (2004) 142-146.

[17] B. Wahdame, D. Candusso, X. François, F. Harel, M.-C. Pera, D. Hissel, J.-M. Kauffmann, Comparison between two PEM fuel cell durability tests performed at

constant current and under solicitations linked to transport mission profile, Int. J. Hydrogen Energy, 32 (2007) 4523-4536.
[18] Süleyman Kaytakoglu, Levent Akyalçin, Optimization of parametric performance of a PEMFC, J. Power Sources 32 (2007) 4418-4423.
[19] S.-J. Wu, S.-W. Shiah, W.-L. Yu, Parametric analysis of proton exchange membrane fuel cell performance by using the Taguchi method and a neural network, Renewable Energy 34 (2009) 135-144.
[20] A. Mawardi, R. Pitchumani, Effects of parameter uncertainty on the performance variability of proton exchange membrane (PEM) fuel cells, J. Power Sources 160 (2006) 232-245.
[21] D. Hissel, M.C. Pera, D. Candusso, F. Harel, S. Begot, in: X.-W. Zhang (Ed.), Advances in Fuel Cells: Research Signpost, North Carolina State University, 2005, ISBN 81-308-0026-8.
[22] F. Harel, X. François, S. Jemeï, S. Moratin, Conception et réalisation d'un banc d'essai pour piles à combustibles à membrane de faibles puissances, INRETS Report, LTE no. 0310, May 2003, Belfort, France.
[23] L. Trevani, J. Ehlerova, J. Sedlbauer, P.R. Tremaine, Complexation in the CU(II)-LiCl-H2O system at temperatures to 423 K by UV-Visible spectroscopy, Int. J. Hydrogen Energy, (2009), doi:10.1016/j.ijhydene.2009.10.046.
[24] A.A. Abreu, A.S. Danko, J.C. Costa, E.C. Ferreira, M.M. Alves, Inoculum type response to different pHs on biohydrogen production from L-arabinose, a component of hemicellulosic biopolymers, Int. J. Hydrogen Energy, 34 (2009) 1744-1751.
[25] I.T. Jolliffe, Principal Component Analysis, Second Edition, Springer, 2002, ISBN 0-387-95442-2.
[26] A. Gelman, J. Hill, Data analysis using regression and multilevel/hierarchical models, Cambridge University Press, 2007, ISBN 0-521-86706-1.