Hydrogen PEMFC stack performance analysis: a data-driven approach

Vitor V. Lopes[1], Augusto Q. Novais[1] and Carmen M. Rangel[2]

[1]INETI, Departamento de Modelação e Simulação, DMS, Paço do Lumiar, 22, 1649-038 Lisboa, Portugal
email: vitor.lopes@ineti.pt, augusto.novais@ineti.pt
[2]INETI, Unidade de Electroquímica de Materiais, DMTP, Paço do Lumiar, 22, 1649-038 Lisboa, Portugal
email:carmen.rangel@ineti.pt

Abstract

For low power fuel cells, it is paramount that management of reactants, water and heat, be realized in a passive fashion in order to minimize parasitic losses. Effective fuel, oxygen supply and water management for reliable performance are also greatly affected by cell geometry and materials.

Fuel cells are complex systems to optimize on a mere experimental basis. As an aid to this goal, data-driven analysis techniques, requiring no mathematical model to be fixed a priori, are gaining a reputation in other fields of work, where a phenomenological modeling approach might be intractable.

This work presents a characterization study of a 12W PEMFC series stack by means of a new data-driven technique, M-NMF. The stack was developed for low temperature operation, uses own designed flow field plates, integrated in a series configuration, and is operated for 12 combinations of hydrogen/air flowrate ratios, generating as many polarization curves.

M-NMF is applied, in combination with an alternating least squares algorithm, to the analysis of the overvoltage data matrix derived from the original experimental polarization data. From this analysis, it is possible to group and differentiate data according to similar overvoltage patterns and gain insight into their relative contribution to fuel cell performance

Keywords: *Fuel cells, polarization curves, data-driven techniques.*

1 Introduction

The characterization of fuel cell performance is often accomplished by the use of polarization curves, which measure the fuel cell overall losses against the cell current density, under given operation conditions [1].

The interpretation of these polarization curves is, usually, performed with the help of computational models to extract meaningful information from the observed overall overpotential. The simplest empirical models do not explicitly take into account the cell spatial dimensions and are based on the direct fit of the experimental data to a non-linear polynomial expression [2] that reflects the three major overpotential losses components in the fuel-cell: a) activation losses, due to the reaction energy barriers at the electrode-electrolyte interfaces; b) Ohmic losses due to cell internal resistance; and c) concentration losses, related to mass transfer limitations [3]. Thus, with this simple approach it is possible to characterise the observed fuel-cell losses and study how the different conditions affects each transfer mechanisms.

This paper proposes a new approach for the polarization curve data analysis based on data factorization techniques. The application of mathematical factorization techniques in electro-chemistry is relatively recent [4] and is being used in other related scientific research fields [5]. Data factorization methods are mathematical methods of reorganizing the information in a data-set of samples. One of the most commonly used is the principal component analysis (PCA), which computes the optimal set of basis vectors that can be used to approximate high-dimensional data in a least-squares sense [6]. By encoding the original data as linear combinations of the basis vectors, it transforms the data into a lower-dimensional space.

For some application, the PCs can be difficult to interpret in terms of physical meaning since the: a) basis vectors can have both positive and negative components, and b) data are represented as linear combinations of them, with positive or negative coefficients. To address this issue, several researchers suggested that the search for a representative basis should be more confined. Thus, new data factorization methods, such as non-negative matrix factorization (NMF), are being

developed to facilitate a rapid qualitative and quantitative data analysis and interpretation [7].

Standard non-negative matrix factorization techniques aim at decompose a given m-by-n data matrix (X) into a bilinear product of two matrices: a m-by-k matrix (W) composed by a reduced set of k basis vectors, and a k-by-n matrix (H) with the loadings needed to reconstruct the original matrix (X) in this new basis [8]. Mathematically, this factorization technique can be formulated as:

$$\begin{aligned} min \quad & \|X - WH\|_F^2 \\ \text{s.t.} \quad & W \geq 0 \\ & H \geq 0 \end{aligned} \quad (1)$$

In the standard NMF, the error different between the original data (X) and the low rank approximation (WH) is measured by the Frobenius norm, but other norms can be used. The low rank factors non-negativity constraints allow the representation of the original data as a sum of positive components, which constitutes an advantage to their interpretability, unlike PCA.

This paper proposes a new non-negative data factorization technique for the analysis and interpretation of a fuel cell polarization curve data set. Mathematically, this technique is based on the standard NMF method with additional constraints. When compared with the existing empirical polarization curve fitting methods, the proposed approach requires fewer assumptions since it is not based on the pre-specification of a semi-empirical model.

The paper is structured in four sections, addressing: the experimental methods used for the polarization curves data collection under different air-hydrogen ratios; the mathematical formulation of the proposed technique together with its main assumptions; the methods used for solving the mathematical optimization problem; and, the results obtained from the technique application to the experimental data-set.

2 Material and Methods

The PEMFC fuel cell under study is an open-air cathode type stack provided with the conventional configuration that comprises 2 end plates, 9 own design graphite bipolar plates, 8 membrane electrolyte assemblies (MEAs) and gaskets. Current collector plates are made of gold plated brass. The 8-cells stack was assembled using MEAs from 3M Corporation, using Nafion®111 as proton exchange membrane (PEM). The active electrode area is 3.8 cm². The stack uses a low power consumption air fan in the edge of the cathode manifolds, for combined high stoichiometric oxidant supply and

stack cooling purposes, avoiding the use of costly, large, heavy and power consuming air pumps and external cooling systems [9]. In this work, the operation conditions were varied in order to simulate situations that could arise when the stack is used in low-power portable PEMFC power applications.

Evaluation of the stack performance was undertaken in a purpose-built test station at UEQM/ INETI [10]. Polarization curves were run in a step-by-step mode, starting at the open circuit potential and taking the fuel cell to its maximum power. Hydrogen pressure was fixed at 500 mbar, the gas being humidified at 40ºC.

Hydrogen flowrates were varied between 0.2 and 0.4 Lmin⁻¹, for four different levels of oxygen coming from air vented by a fan, supplying flowrates from 8.8 to 40 Lmin⁻¹(3.5 to 8.0 O₂ Lmin⁻¹ equivalent).

Table 1. Table containing the experimental design flow-rates and the corresponding O_2 to H_2 feed-ratio.

Fan voltage (O_2 Lmin⁻¹ equivalent)	4.0 V	5.0 V	6.0 V	8.0 V
	3.5	4.6	5.8	8.0
H_2 Lmin⁻¹				
0.2	17.6	23.2	28.8	40.0
0.3	11.7	15.4	19.2	26.7
0.4	8.7	11.6	14.4	20.0

3 Non-negative data factorization method for polarization curve analysis

The direct application of the standard NMF technique to the polarization curves data might result in the extraction of a W basis, which might be difficult to interpret. Thus, to improve the extracted basis interpretability, it is necessary to add more constraints to the standard NMF mathematical formulation. These additional constraints are, by nature, empirical and reflect observations based on the characterization used in the empirical curve fitting methodology.

A polarization curve (E_{obs}) is modeled empirically to characterize the observed overpotential losses into three component terms: a) activation (E_{act}); b) ohmic (E_{ohm}); and c) concentration (E_{conc}) [3]:

$$\begin{aligned} E_{obs}(i) &= E_{oc} - E_{act}(i) - E_{ohm}(i) - E_{conc}(i) \\ E_{obs}(i) &= E_{oc} - E(i) \end{aligned} \quad (2)$$

where the E_{oc} represents the open-circuit cell potential, $E_{obs}(i=0)$.

From existent literature [2, 3], the following mathematical properties were observed:

O.1) overpotential losses are divided only in three components and are all positive;

O.2) the activation overpotential is modeled as a monotonically increasing concave function of the current;

O.3) the concentration overpotential is modeled as a monotonically increasing convex function of the current;

O.4) the ohmic overpotential has a positive linear dependence of the current.

Based on these observations, a modified NMF method (M-NMF) for the analysis of polarization curve data is described by the following optimization problem formulation:

$$min \qquad \|X - W H\|_F^2 + \lambda \|G\|_1^2 \qquad (3.1)$$
$$s.t. \qquad W \geq 0 \qquad (3.2)$$
$$H \geq 0 \qquad (3.3)$$
$$G \geq 0 \qquad (3.4)$$
$$\forall_{j,l=0\ldots n-1, v=1} \, w_{l,v} \leq w_{j,v} + g_{j,v}(i_l - i_j) \quad (3.5)$$
$$\forall_{j,l=0\ldots n-1, v=2,3} \, w_{l,v} \geq w_{j,v} + g_{j,v}(i_l - i_j) \quad (3.6)$$

where W is the n-by-3 basis matrix composed by the $w_{j,v}$ elements, G is composed by the $g_{j,v}$ elements and represents the sub-gradients of the matrix of W, and H is the 3-by-m coefficients matrix. The n-by-m data-set matrix X is constructed by the column-wise concatenation of the different polarization curves $E(i)$:

$$X = \begin{bmatrix} E_0(i_0) & \cdots & E_{m-1}(i_0) \\ \vdots & \ddots & \vdots \\ E_0(i_{n-1}) & \cdots & E_{m-1}(i_{n-1}) \end{bmatrix} \quad (4)$$

where the $i_0 = 0$ to i_{n-1} are the n values for the current intensity common to all polarization curves and sorted by increasing order.

The main differences between the M-NMF and NMF is the: a) penalization term in the objective function; b) constraints define by equations 3.4 to 3.6; and, c) the fixed number of extracted basis (3) equals the number of overpotential contributions, as identified by the empirical curve fitting methods. The imposed constraints are quite general, in the sense that there are no equations imposing a specific behaviour to the basis vectors and, thus, the number of possible alternative basis in the feasible solution space is still high. As a result, in the M-NMF the extracted basis vectors are expected to provide more physical insight about the underlying phenomena.

The penalization term, based on the norm-1 of the sub-gradient matrix entries, is introduced for regularization purposes and to prevent the existence of degenerate numerical solutions. As a side effect, this penalization induces an amount of smoothness on the basis vectors, since abrupt column-wise changes in the entries of the W matrix penalize the objective function. The penalization weight (λ) is set at a small constant (10^{-5}).

Equation 3.4 states that all sub-gradient matrix G entries must be non-negative, and corresponds to the observation that all overpotential contributions are increasing functions of the current. Equation 3.5 and 3.6 are the definitions of concave and convex properties for the W basis, respectively. Thus and in simple terms, the M-NMF aims at finding the best least squares basis characterized by three increasing functions, with one being concave and two convex.

4   Numerical method for solving the non-linear optimization problem

The M-NMF formulation defined by equations 3.1 to 3.6, is a non-linear optimization problem due to the WH bilinear term in equation 3.1. The numerical method used to solve this problem is based upon the alternating least squares method, i.e. it alternately fixes one matrix and improves the other:

$$\text{Find } W_{k+1} : f(W_{k+1}, H_k) \leq f(W_k, H_k)$$
$$\text{Find } H_{k+1} : f(W_{k+1}, H_{k+1}) \leq f(W_{k+1}, H_k) \quad (5)$$

where f function represents the objective function defined by the equation 3.1. In bound constrained optimization [11], this procedure is called the "block coordinate descent" method, where sequentially one block of variables is minimized under the respective constraints and the remaining block are fixed. For the M-NMF, there are only two block variables W and H and, hence, when one variable block is fixed, the resultant constrained penalized least square sub-problem admits one single unique solution.

The overall algorithm for solving the non-linear optimization problem is structured in Algorithm 1. Each least squares sub-problem can be solved in a straightforward manner by using standard semidefinite programming methods.

The convergence of the proposed algorithm is not a trivial result. Early convergence analysis for "block coordinate descent" methods requires that sub-problems to have unique solutions [11]. In the algorithm, both sub-problems are not strictly convex and, thus, each sub-problem may have multiple optimal solutions. For instance, when W_k is equal to zero, any H matrix is optimal. However, for the case of two blocks this uniqueness condition is not required [12] and, thus, any limit point of the

sequence $\{W_k, H_k\}$ generated by the algorithm is a stationary point of the optimization problem 3.1-3.6.

Algorithm 1. Alternating least squares algorithm implementation for the non-linear optimization method.

1. Initialize H_1 with randomly selected values between 0 and 1
2. For $k = 1, 2, \ldots$

$$W_{k+1} : min \quad \|X - W_{k+1} H\|_F^2 + \lambda \|G\|_1^2$$
$$\text{s.t.} \quad W_{k+1} \geq 0$$
$$H = H_k$$
$$G \geq 0$$
$$\forall_{j,l=0\ldots n-1, v=1} \, w_{l,v,k+1} \leq w_{j,v,k+1} + g_{j,v}(i_l - i_j)$$
$$\forall_{j,l=0\ldots n-1, v=2,3} \, w_{l,v,k+1} \geq w_{j,v,k+1} + g_{j,v}(i_l - i_j)$$

$$H_{k+1} : min \quad \|X - W H_{k+1}\|_F^2$$
$$\text{s.t.} \quad W = W_{k+1}$$
$$H_{k+1} \geq 0$$

3. Iterate until convergence

5 Polarization curve data analysis and result discussion

The polarization data were obtained according to the experimental procedure presented in section 2, and consist of 12 polarization curves for different reactant hydrogen and air flowrates, as indicated. The different polarization curves labels were coded according to the hydrogen and air flowrates. For instance the polarization curve data for 0.4 Lmin^{-1} hydrogen flowrate and 8.0V fan air flowrate is coded as 0.4#8.0V.

Prior to the application of the M-NMF, the overpotential data matrix (X) is constructed by subtracting the open circuit potential from the observed potentials. Results presented in figures 1 to 4 show that M-NMF captures 99.97% of the original data variance of the original m-by-n data matrix (X).

Figure 1 presents the M-NMF basis vectors (W) and shows that: a) the 3-rank decomposition is enough to describe the original data variance; b) the M-NMF is able to find basis vectors with shapes similar to the overpotential contributions empirically identified; and, c) the ohmic component has a constant slope, although this is not explicitly imposed on the M-NMF formulation.
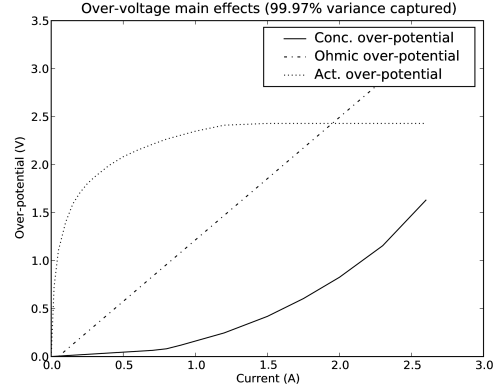

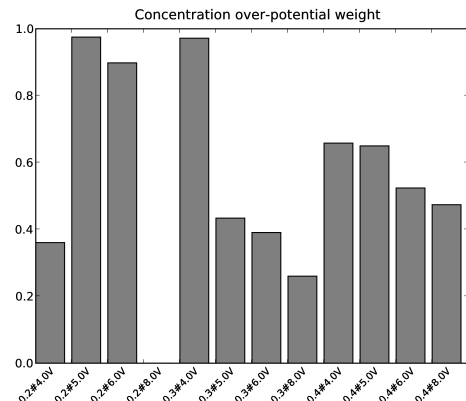Fig. 1. The M-NMF basis vectors (W) as a function of current.

Fig. 2. M-NMF concentration over-potential weights (from H matrix).

Figure 2 highlights the high variance distribution of the concentration overpotential for all experimental conditions.
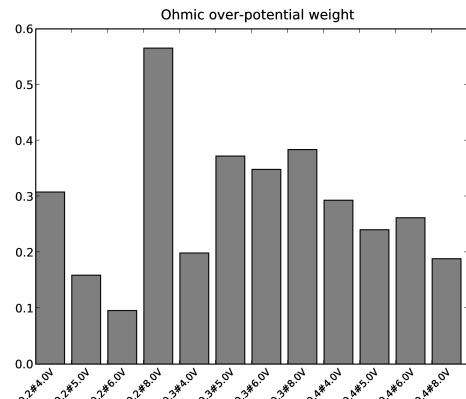

Fig. 3. M-NMF ohmic over-potential weights (from H matrix).

A detailed analysis reveals that: a) for a given H_2 flow, the concentration overpotential loss is at its lowest for the higher oxygen flowrates; b) the smallest concentration overpotential loss is found (case 0.2#8.0V) for the maximum O_2 to H_2 feedratio; c) for a given fixed hydrogen flowrate, the concentration overpotential loss decreases with

the increase of airflow, except for the case of the lowest hydrogen and lowest oxygen concentrations used. (case 0.2#4.0V).

For the ohmic overpotential, Figure 3 details a more complex variability pattern. For the 0.2 and 0.3 $Lmin^{-1}$ series the effect is reverse to that observed for the concentration overpotential, while the same trend is observed for the 0.4 $Lmin^{-1}$ series. However, the effect is less marked judging by the value of the respective weights. Results seem to indicate that a relative balance between minimum ohmic and concentration overpotentials is not easy to attain for the first two series, which cover most of the range of O_2 to H_2 feed-ratios used in this work.
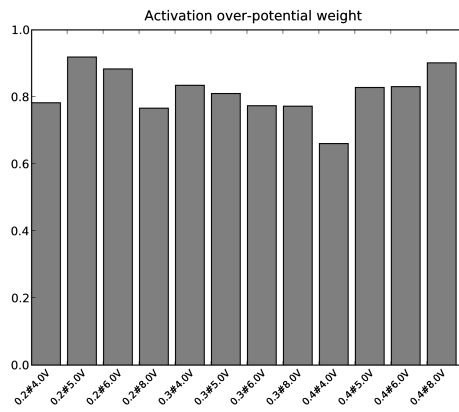

Fig. 4. M-NMF activation over-potential weights (matrix **H**).

Figure 4 shows that the activation overpotential component presents a smaller variability when compared with the other overpotential components, although exhibiting overall higher weight values. The minimum activation overpotential contribution is found for the lowest O_2 to H_2 feed-ratio used (0.4#4.0V).
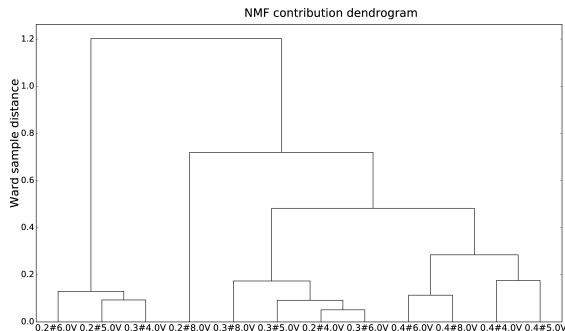

Fig. 5. Dendrogram based in the M-NMF weight contributions using Ward's clustering algorithm.

Figure 5, shows the dendrogram analysis of the M-NMF weight contributions calculated by Ward's method [13]. The analysis of this figure shows that the data can be divided into three different clusters, according to the contribution weights: a) 0.2#5.0V, 0.2#6.0V and the 0.3#4.0V; b) the data set for 0.4

$Lmin^{-1}$ hydrogen flowrate; c) 0.2#4.0V, 0.3#5.0V, 0.3#6.0V, 0.3#8.0V; and, d) the isolated 0.2#8.0V condition that presents a maximum ohmic and a minimum concentration overpotential weights.

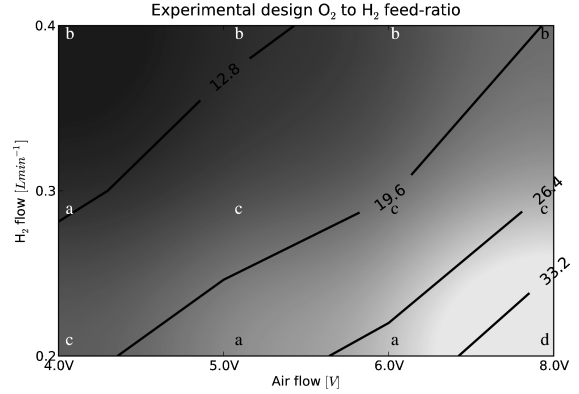Figure 6 shows a contour plot based on the feed ratios shown in Table 1, where these clusters are projected.

Fig. 6. O_2 to H_2 feed ratio contour plot (Table 1).

In figure 7, a 3D plot displays the three overpotential components highlighting their relative contribution.
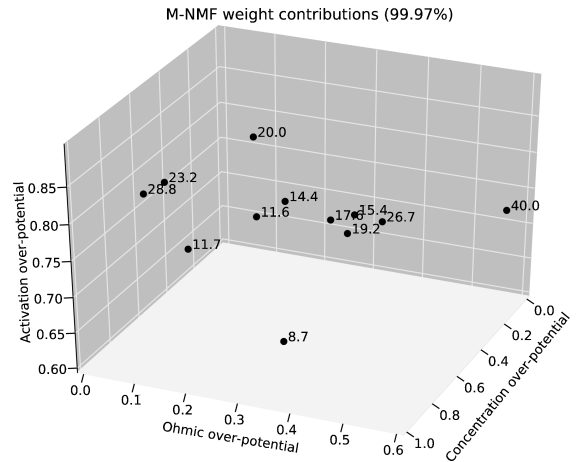

Fig. 7. M-NMF weight contributions 3d plot with the respective O_2 to H_2 feed ratio (labels) given by Table 1.

6   Discussion

Electrode kinetics affect fuel cell performance due to irreversibilities from activation, ohmic resistance and mass transfer, influenced by operating conditions, such as pressure, flowrates, humidification and temperature of reactant gases.

For fuel cells applications of low power, as in the present case, effective air supply and water management for reliable performance is greatly affected by cell geometry and materials.

The flow field configuration in current collectors and gas distribution plates have a great impact in critical issues related to water and thermal management. The graphite plate own-design flow field is expected to passively aid the water and thermal management of the cell, by reducing the system complexity and minimising the power consumption of auxiliary components. Gas diffusion electrodes supported by carbon made materials contain the catalyst formulation; apart from their influence on thermal and water management, other issues such as contamination (fuel/air) and catalyst de-activation, which are not dealt in here, also contribute to voltage loss.

Variations of membrane conductivity with temperature and water content are considered critical: drying during operation as a result of dragging of water by protons or over-saturated conditions may cause condensation at the electrodes, causing flooding with the consequent voltage degradation.

In the development of an open-cathode stack, as in the present case, one common problem is related to the flowrate requirement for air to be used as coolant as well as oxidant. The air flowrate and humidification should be such as to provide an adequate supply of oxygen to the MEA layer, while preventing the dehydration of the membrane (either excessive temperature or air flowrate) and flooding of the air mass transport paths. The use of an air fan is expected to enable MEA hydration level control, which will allow varying the air flowrate according to the stack operation conditions (power output, ambient air relative humidity and temperature).

In systems such the one under study, where air is simultaneously reactant and cooling gas, the situation is complex. Concentration voltage losses are expected if, on balancing these two effects, the resulting stoichiometry remains defective.

The obtained concentration, ohmic and activation overpotential results provide insight into the dynamics of the cell in terms of water transport, membrane hydration and temperature variation.

For 0.4 $Lmin^{-1}$ H_2, the activation overpotential is high throughout, which can be explained by sluggish cathode kinetics. As the air flow is increased, more oxygen is available but also temperature drops due to cooling effects, and hence an increase of the overpotential is observed. At low air flowrates, insufficient oxygen to meet both reaction and cooling demands, affects the water balance within the cell, which ensures the hydration of the membrane.

An increase in air flow, on the other hand, decreases the ohmic overpotential and thus contributes to a favourable water balance:

electrosmotic dragging of water from anode to cathode and back diffusion from cathode to anode. For the concentration overpotential a similar trend is observed.

For H_2 flowrates below 0.4 $Lmin^{-1}$, i.e. 0.3 and 0.2 $Lmin^{-1}$, the back diffusion of water from cathode to anode may predominate, thus contributing to a dryer cathode.

For 0.2 $Lmin^{-1}$, the overpotentials display a minimum value of the ohmic component, most likely related to the water balance: dry anode for low air flowrates and anode flooding for sufficiently high rates.

By comparison, the concentration overvoltage component exhibits a maximum value and overall reverse tendency. This same trend, although less marked, is equally observed for the activation overpotential.

One of the complex issues related to the effect of the operating conditions on cell performance is the non-linear variation of the saturated water vapour pressure with temperature. Relative humidity (RH) versus temperature of exit air and hydrogen in PEM fuel cells shows that for most operating conditions it will result in either a dry or flooded condition for the fuel cell.

7   Conclusions

The deconvolution of fuel cell polarization curves into 3 major components reflecting as many main causes for voltage drop, was achieved with good results by the use of a modified NMF method (M-NMF) developed in this work, together with an algorithm to solve the associated non-linear optimization problem.

For the activation overpotential, as expected, results confirmed the limitation on cell performance of the cathodic kinetics (much lower than for the anode), exhibiting the average highest component weights throughout. This can be attributed to the low working temperature of the fuel cell which is also associated to a high cathodic overvoltage.

Air stoichiometry is revealed to be extremely high for systems in which air plays a dual role, that of reactant and cooling gas. For minimal voltage loss oxygen-hydrogen feed-ratios are paramount in order to find a balance between the electrosmotic drag of water to the cathode and water back diffusion from the cathode.

Air blown at 4V is revealed as insufficient for the fuel cell under study, in order to provide air to be used as coolant as well as oxidant.

Maximum oxygen flowrates used in this work are found to be useful only when working at high currents. At low currents, overvoltage limitations

may bring about the drying out of the anode, more so when lower hydrogen flow ratios are used.

This paper illustrates how a M-NMF designed for the analysis of polarization curve data and an algorithm to solve the associated nonlinear optimization problem, can be applied to experimental data and be instrumental in their interpretation.

Further work is planned that will take advantage of this capability in addressing larger sets of operating conditions and of data.

References

[1] J. Wu, X. Yuan, M. Blanco, J. Martin, J. Zhang, H. Wang, *Diagnostic tools in PEM fuel cell research: Part I Electrochemical techniques.* Int. J. Hydrogen, 33 (2008) 1735 – 1746

[2] L. Pisani, G. Murgia, M. Valentini, B. D'Aguanno, *A new semi-empirical approach to performance curves of polymer electrolyte fuel cells*, J. Power Sources 108 (2002) 192 – 203

[3] M. Hoffman D. Noren, *Clarifying the Butler–Volmer equation and related approximations for calculating activation losses in solid oxide fuel cell models*, J. Power Sources, 152 (2005) 175 – 181

[4] S. Brown and R. Bear, Jr. *Chemometric techniques in electrochemistry: A Critical Review*, CRC Crit. Rev. Analyt. Chem., 24, (1993) 99 – 131

[5] Y. Nia, S. Kokot *Does chemometrics enhance the performance of electroanalysis?* Analytica Chimica Acta, 626 (2), 130 – 146, 2008

[6] I. T. Jolliffe *Principal component analysis*, 2nd ed., Springer series in statistics. New York Springer-Verlag, 2002

[7] P. Hopke *The evolution of chemometrics*, Anal. Chim. Acta, 500(1-2), 365 – 377, 2003.

[8] M. Berrya, M. Brownea, A. Langvilleb, V. Paucac, R. Plemmonsc *Algorithms and applications for approximate non-negative matrix factorization.* Comp. Stat. & Data Anal., 53 (2007) 155 – 173.

[9] D. T. Santa Rosa, D. G. Pinto, V. S. Silva, R. A. Silva, C. M. Rangel, *High performance PEMFC stack with open-cathode at ambient pressure and temperature conditions*, Int. J. of Hydrogen Energy, 32, 4350-4357, 2007.

[10] C. M. Rangel, R. A. Silva, P. P. Luz; *Células de combustível para aplicações na indústria automóvel*, Ciência e Tecnologia de Materiais, 13, 12-26 (2001).

[11] D. Bertsekas. *Nonlinear programming*. Athena Scientific, Belmont, 2nd edition, 1999.

[12] L. Grippo and M. Sciandrone. *On the convergence of the block nonlinear Gauss-Seidel method under convex constraints*. Operations Research Letters, 26 127-136, 2000

[13] Ward, J.H., *Hierarchical grouping to optimize an objective function*. J. American Stat. Assoc., 58: 236-244, 1963.