

## Data Driven Model for a Fuel Cell stack development in a complex Multi-source Hybrid Renewable Energy System

G. Napoli, M. Ferraro, G. Brunaccini, G. Dispenza, V. Antonucci

CNR-ITAE, Consiglio Nazionale delle Ricerche  
Istituto di Tecnologie Avanzate per l'Energia "Nicola Giordano"  
Via Salita S. Lucia sopra Contesse 5-98126 Messina, Italy  
Phone:+39 090 624236, e-mail: [giuseppe.napoli@itaecnr.it](mailto:giuseppe.napoli@itaecnr.it)

**Abstract.** Fuel cells based on polymer electrolyte membrane are considered as the most hopeful clean power technology. The operating principles of polymer electrolyte membrane fuel cells (PEMFC) system involve electrochemistry, thermodynamics and hydrodynamics theory for which it is difficult to establish a mathematical model. In this paper a nonlinear data driven model of a PEMFC stack is developed using Neural Networks (NNs). The model presented is a black-box model, based on a set of measurable exogenous inputs and is able to predict the output voltage and cathode temperature of a high power module working at the CNR- ITAE. A 5 kW PEM fuel cell stack is employed to experimentally investigate the dynamic behaviour and to reveal the most influential factors.

### Key words

Polymeric Electrolyte Membrane Fuel Cell (PEMFC), Data Driven Model, Neural Networks (NNs), Nonlinear models, Backpropagation (BP).

### 1. Introduction

Many studies were devoted to the development of techniques able to model the fuel cell behaviour through the polarization curves by analyzing the steady-state performance.

However, the dynamic characteristics are also important factors for the fuel cell development, especially for systems where a number of variables (e.g. fuel feeding, load, stack temperature etc..) can frequently change.

The knowledge of the dynamic behaviour is therefore an important condition to evaluate the whole system performance. Moreover this knowledge gives also guiding principles for control strategies.

Modelling PEMFC is not an easy task considering an analytic approach [1] – [5]. Indeed, analytic models are based on the deep knowledge of physicochemical phenomena and usually depend on a lot of parameters.

In process industries it is very common to identify the required model from historical input-output data, usually

stored in the plant database. Models used in such a framework are generally known as Data driven model or Soft Sensors (SSs).

The wide availability of control systems allows to designers and operators an easy implementation of Data driven model with a minimum cost. Moreover a wide number of applications of SSs in the field of process industry, have been developed in recent years for plant output estimation [6] – [8], fault diagnosis [9], [10] and "what-if" analysis [11].

To identify the dynamical relationship between the input and output variables, Data driven model are often designed by using NNs, and Multi Layer Perceptrons (MLP) have been widely and successfully used to this end [12], [13].

It has been shown, in fact, that feed-forward NNs with one hidden layer can approximate any continuous function with good accuracy [14].

The objective of this paper is to achieve behavioural model that simulates the performance of the PEMFC stack fed with different flow rates, operating at different temperatures and under different load demands. Few works of this kind have been reported in literature [15] - [21].

The model obtained will be exploited as a component of complex Multi-source Hybrid Renewable Energy System simulator able to manage the energy flows between fuel cell stack and batteries under different scenarios of loads and whether data.

The NN based model, developed on the basis of the experimental data, can be hence used to investigate the influence of process variables for different operating conditions. By this approach the input/output empirical correlations of the PEMFC stack were expressed in terms of process operating variations.

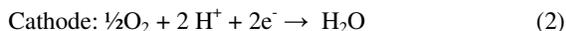
In the first part of the paper, a general description of the PEMFC stack is presented. Next, a brief introduction to Data driven model will be given. In the second part the test procedure is described and data driven modelling based on NNs applied to the fuel cell stack is proposed. Finally, simulation results obtained using the proposed model are compared to experimental data.

## 2. PEMFC Stack

PEM Fuel Cell is regarded as the most competitive candidate to become the main alternative energy source to the traditional forms of power conversion because of its well-known features.

It consists in a solid polymeric membrane which acts as electrolyte and two platinum porous catalyst electrodes coupled on both sides of the membrane.

The basic chemical reactions are:



Resulting in :



Fuel Cells, like batteries, convert the energy contained into chemical species into DC power through electrochemical reactions. In Fig.1 a representative fuel cell system scheme is shown. A typical characteristic curve representing battery and Fuel Cell features is the polarization curve (Fig.2a), that is a representation of the device performance at different power generation levels. Fig. 2b shows the comparison of the voltage characteristics at a constant “discharge” current.

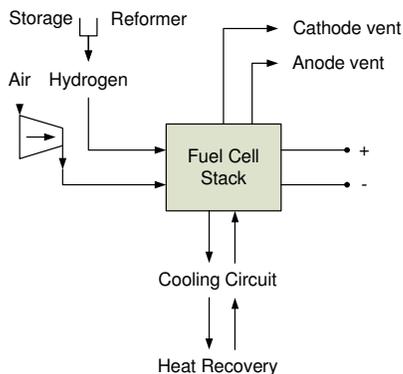


Fig. 1. Fuel Cells System scheme

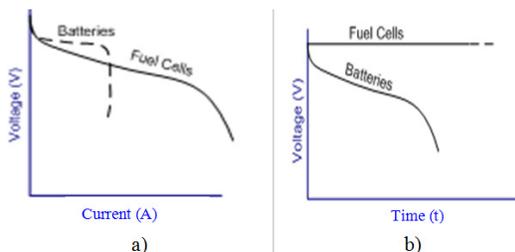


Fig. 2. Comparison between Batteries and Fuel Cell voltage.  
a) Polarization curve; b) Voltage at constant current

Differently from batteries, Fuel Cells are fed with an external source, like hydrogen stored as compressed gas

or into hydrides, or by hydrogen rich gas produced via hydrocarbons reforming.

In a Multi-source Hybrid Renewable Energy System the energy storage is a key issue for the management of the system itself and for the characteristics of the grid (e.g. micro-grids, smart-grids) which the system is connected to. The possibility to supply the fuel cell with hydrogen produced by electrolysis and natural gas as well, assures a more stable operation, providing a high quality power supplying. However, in order to allow a constant voltage operation, a converter (DC/DC or DC/AC) is needed.

The device in this study was a 5 kW NUVERA PEM fuel cell stack consisting of 50 cells with 500 cm<sup>2</sup> of geometric area each (Fig. 3).

The stack works with a Cathode Water Injection system that dissipate heat and humidifies the cathode side by mixing air stream and water.

In this work the stack was operated in a test station that includes an electronic load, gas mass flow controllers and pressure sensors and actuators, and a control and monitor software (Fig. 4).

The station allowed to modify a number of process variables and to obtain an experimental database about the PEMFC stack under different operating conditions.

The system was able to communicate with the user through a graphic LabVIEW interface developed by ITAE personnel built for the control and monitoring of the stack (Fig.5).

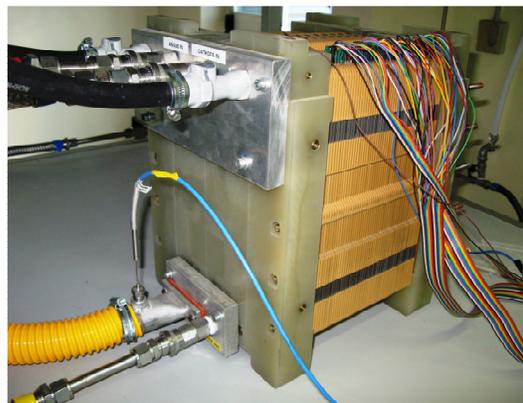


Fig. 3 Nuvera 5Kw PEM Stack



Fig. 4 Test Station

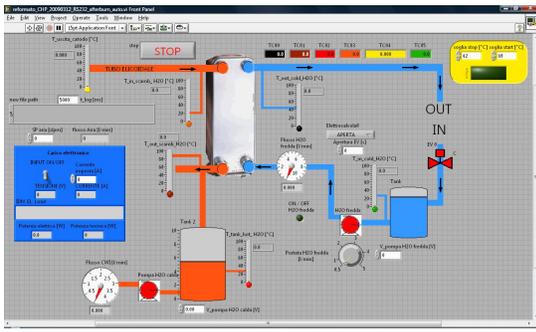


Fig. 5. Software interface of control and monitoring.

The control system allowed to change the input variables (e.g. gas flows combination, load and operating range of stack temperatures) and to monitor thermal and electrical characteristics (e.g. single cell voltage, anode and cathode temperatures, water temperature) by more than 20 sensors.

The software allowed also a complete data acquisition that was used for the experimental validation of the data driven model.

### 3. Data driven models

The objective in using data driven model for a PEMFC stack is to obtain the nonlinear input/output relationship that is represented in a set of historic data. To this aim models based on NNs have proved to be a suitable nonlinear modelling method. NNs and in particular Multi-Layer Perceptrons (MLPs), are tools generally accepted for modelling complex systems, due to their ability to approximate nonlinear I/O relations.

NNs are parallel computational devices consisting of groups of highly interconnected processing elements called neurons. Data driven model based on NNs are characterized by topology, computational characteristics of their elements, and training rules. The connections between the units are also known as synapses or weights. The pattern of the connections between the units determines the architecture of the network that depends largely on the desired task that the network is to perform. Although there are many learning algorithms that can be used to train multilayer NNs, backpropagation, with its variations, is currently the mainstay of artificial neural network learning.

A commonly used artificial neuron, shown in Fig. 6 is a simple structure, having just one nonlinear function of a weighted sum of several data inputs  $x_1, \dots, x_n$ ; this version, often called a perceptron, computes a ridge function:

$$y = \sigma(w_0 + \sum_{i=1}^n w_i x_i), \quad (4)$$

and assume that the function  $\sigma$  is a smooth, increasing, bounded function. Examples of sigmoids in common use are:

$$\begin{aligned} \sigma_1(u) &= \tanh(u) & \text{or} & \\ \sigma_2(u) &= 1/(1 + \exp(-u)) \end{aligned} \quad (5)$$

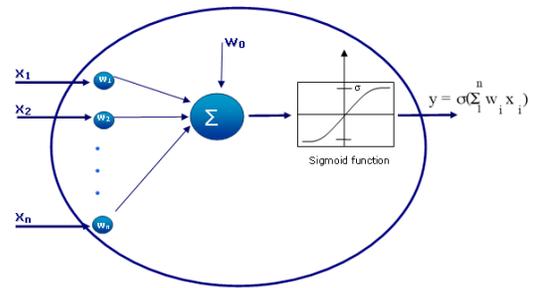


Fig. 6 Feedforward neuron

The weight-adjustment algorithm will use the derivatives of the sigmoid functions reported in (5). The weights  $w_i$  have to be selected or adjusted to make this ridge function approximate some known relation which may or may not be known in advance.

In order to learn functions more complex than ridge functions, one must use networks of perceptrons. Thus the general idea of feedforward networks is that they allow to realize functions of many variables by adjusting the network weights.

There are several neural network architectures useful for process modelling. The three-layer artificial neural network represents a feed-forward structure that is useful for many applications. In the three-layer NNs, the units are arranged into three layers: an input layer, which receives input data from outside the network; a hidden layer, which receives signals from the input layer; and an output layer, which receives signals from the hidden layer and represents the final output of the network (Fig. 7).

In the multi-layer NNs, the calculation of the network activation proceeds layer by layer from the input layer through the hidden layers to the output layer.

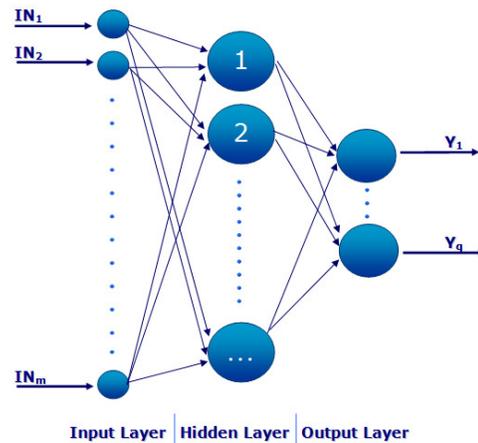


Fig. 7 Backpropagation network model

The general formulation for a NMA multi-input, multi-output model is:

$$[Y(k)] = f[IN_1(k), \dots, IN_1(k-n_1), IN_2(k), \dots, IN_2(k-n_2), \dots, IN_m(k), \dots, IN_m(k-n_m)] \quad (6)$$

where  $Y(k)$  is the output vector at time  $k$ ,  $IN_i$  ( $i = 1, 2, \dots, m$ ) are the input variables with their respective time delays  $n_i$ , and  $f$  is an unknown nonlinear function.

#### 4. PEMFC Stack model

The typical steps of data driven models based on NNs can be summarized as it follows:

- data collection and filtering;
- variables and model structure selection;
- model identification;
- validation.

In this application data were been collected and filtered with the help of the system experts in order to represent the whole system dynamics. Moreover, a careful investigation of the available data was been performed to detect either missing data or outliers.

The initial step in any modelling approach based on NNs requires the training of a large number of candidate neural models with different structures for both the model (number and type of inputs) and the network (number of hidden neurons), and different I/O data set.

In this work the variables choice was been made taking into account the available know-how about the plant dynamics and is reported in Tab.1.

Tab. 1 Input and output variables

Variable	Ranges	Unit
Stack Current	[30÷110]	A
Hydrogen mass flow	[19÷104]	slpm
Nitrogen mass flow	[7÷39]	slpm
Air mass flow	[35÷184]	slpm
Cathode Temperature (OUT)	[54÷76]	°C
Stack Voltage (OUT)	[34÷41]	V

To obtain a significant stack voltage dynamic the load current was changed from 30 to 110 A, the air mass flow from 35 to 184 slpm, the hydrogen mass flow from 19 to 104 slpm and nitrogen mass flow from 7 to 39 slpm in a number of different combinations (Tab.2). The anode preheater was fixed to 60 °C.

Tab. 2 Gas flows and Stack Current combinations

Stoic. Ratio	Gas	Flow			
		Current			
		30A	60A	90A	110A
1.4	H <sub>2</sub>	19.6	39.2	58.8	72.8
	N <sub>2</sub>	7.26	14.52	21.77	26.96
	Air	35÷50	70÷100	105÷150	128.8÷184
1.6	H <sub>2</sub>	22.4	44.8	67.2	83.2
	N <sub>2</sub>	8.29	16.59	24.88	30.81
	Air	35÷50	70÷100	105÷150	128.8÷184
1.8	H <sub>2</sub>	25.2	50.4	75.6	93.6
	N <sub>2</sub>	9.33	18.66	28	34.66
	Air	35÷50	70÷100	105÷150	128.8÷184
2.0	H <sub>2</sub>	28	56	84	104
	N <sub>2</sub>	10.37	20.74	31.11	38.52
	Air	35÷50	70÷100	105÷150	128.8÷184

To train the different neural models, a set of patterns with sampling time T=10 sec., covering different process working points, was considered. A set of about 150 selected I/O samples was used for the NN training, while a different set of about 300 samples was used for testing and comparing the different models.

Data in software databases have different magnitudes, depending on the units adopted and on the nature of the process. This can cause larger magnitude variables to be dominant over smaller ones during the identification process. Data scaling was therefore needed. The z-score scaling method has been used for input variables (Fig. 8). The z-score normalization is given by:

$$x' = (x - mean_x) / \sigma_x \quad (7)$$

where:

- $mean_x$  is the estimation of the mean value of the unscaled variable;
- $\sigma_x$  is the estimated standard deviation of the unscaled variable.

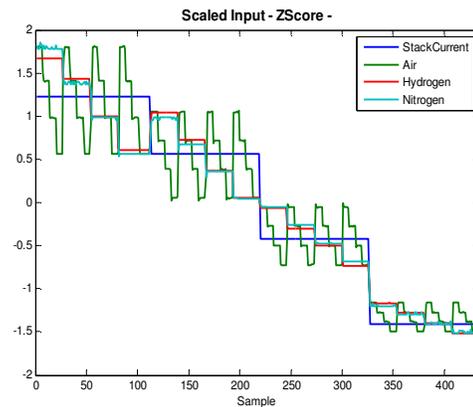


Fig. 8 Input variables

All the networks were trained using the Levenberg-Marquardt algorithm, with the early stopping approach to prevent overfitting.

A trial and error approach was used to select the best model among possible candidates. Several sub-optimal neural models were been obtained, corresponding to a different number of hidden neurons and/or different set of learning patterns. The neural model with the higher correlation coefficients has 12 hidden neurons and indicators of performance using train and test data are reported in Tab.3.

Fig.9 shows the comparison between acquired data and corresponding estimations using the test data for the model based on the proposed procedure.

Fig. 10 and 11 and Fig. 12 and 13 show the same comparison for the Stack Voltage and Cathode Temperature using a test and a train data subset respectively.

In Fig. 14 and 15 the estimation capabilities of the model, on the same set of test data used in Fig. 9 were analyzed by the 4-plot analysis of the residual.

In more detail, the 4-plot analysis was based on visual inspection of the time plot, lag plot, histogram and normal probability plot.

These four graphics can be used to test specific properties of the model residual. In particular, the histogram and the normal plot will be used to investigate the normal distribution of the model residual. If the model identification works properly, the histogram will be approximately zero centred and bell-shaped, while the normal plot will be close to a straight line. The lag plot is used to search for any dependence of the model residual on its past values.

Any discrepancy from the 4-plot ideal aspect is a symptom of divergence from the stated hypothesis for the model residual.

Tab.3 Performance of the best NN model.

	<b>Variable</b>	<b>Correlation Coefficient</b>	<b>Mean Error</b>	<b>Standard Deviation</b>
<b>Test</b>	<b>VStack</b>	0.9962	0.0086	0.1980
	<b>TCath</b>	0.9843	0.1097	0.5855
<b>Train</b>	<b>VStack</b>	0.9978	0.0064	0.1398
	<b>TCath</b>	0.9958	0.0363	0.4772

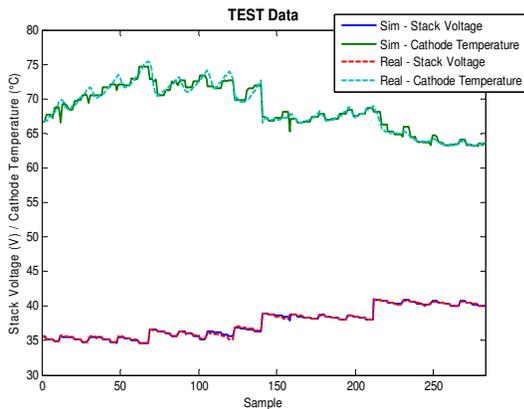


Fig.9 Comparison of Stack Voltage and Cathode Temperature acquired data and corresponding model estimation for the Test data.

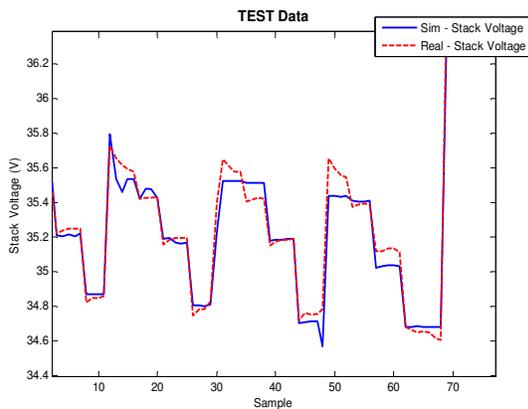


Fig.10 Comparison of Stack Voltage acquired data and corresponding model estimation for a Test data subset.

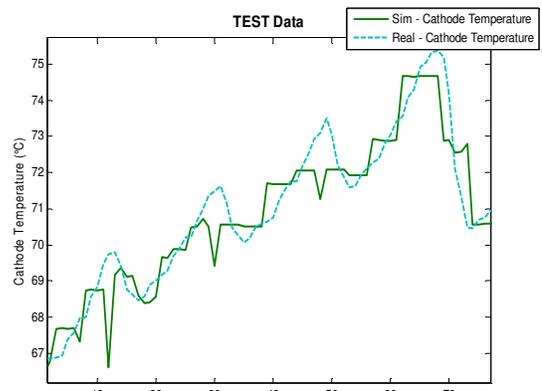


Fig.11 Comparison of Cathode Temperature acquired data and corresponding model estimation for a Test data subset.

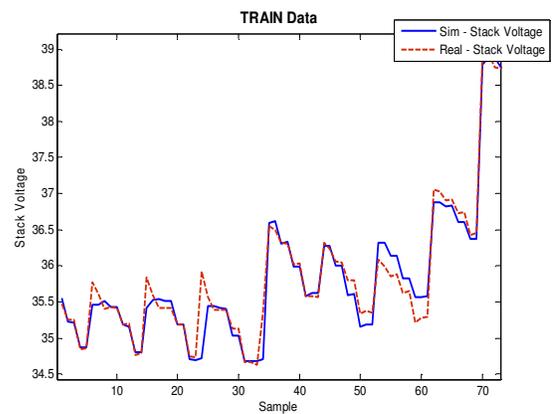


Fig.12 Comparison of Stack Voltage acquired data and corresponding model estimation for a Train data subset.

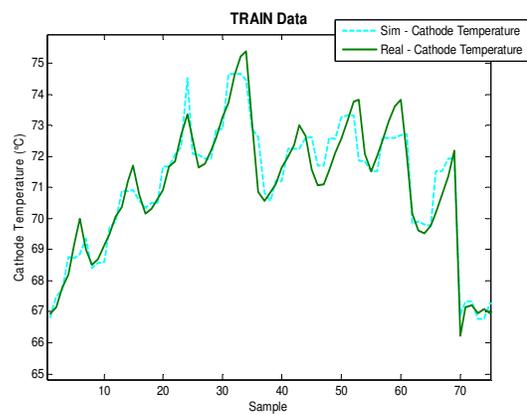


Fig.13 Comparison of Cathode Temperature acquired data and corresponding model estimation for a Train data subset.

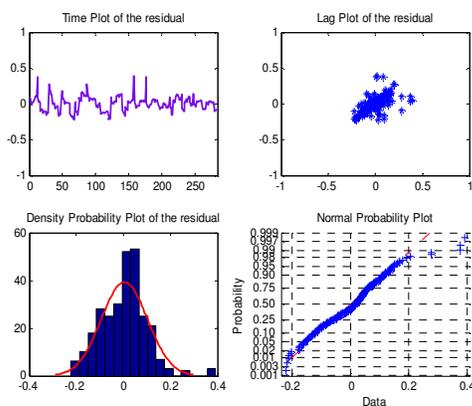


Fig.14 Stack Voltage. 4-Plot analysis for the model residual

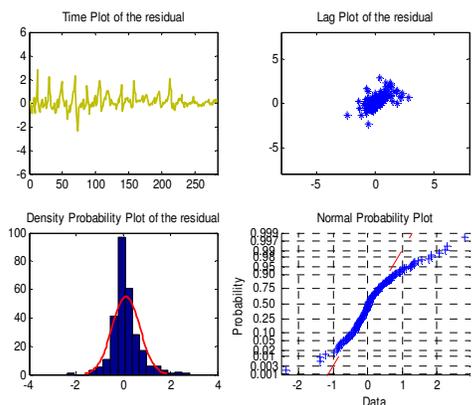


Fig.15 Cathode Temperature. 4-Plot analysis for the model residual

## 5. Conclusion

This work investigated the dynamic performance of a 5 kW NUVERA PEM fuel cell stack behaviour with the aim to provide an integration in a Multi-source hybrid alternative energy systems and to study the related issues such as control strategies, microgrid connection, stack versatility and its ability to follow the load, etc.. The developed model includes all important operating characteristics of the process using non-parametric approach.

The dynamic behaviour showed particular characteristics during different stages of air/hydrogen/nitrogen flow rate, temperature and current load variation.

The data driven obtained model performed quite satisfactory and stack voltage and cathode temperature dynamics were simulated with accuracy. The trained NN model is computationally fast and easy to use, especially in the case where physical models are not readily available.

## References

[1] Rowe, A., Li, X., 2001. Mathematical modelling of proton exchange membrane fuel cells. *Journal of Power Sources*, 102: 82-88.

[2] Bıyıkoglu Atilla. Review of proton exchange membrane fuel cell models. *International Journal of Hydrogen Energy* 2005;30:1181–212.

[3] Bernardi DM, Verbrugge MW. A mathematical model of a gas diffusion electrode bonded to a polymer electrolyte. *AICHE* 1991;37:1151–62.

[4] T.E. Springer, T.A. Zawodinski, S. Gottesfeld, Polymer electrolyte fuel cell model, *J. Electrochem. Soc.* 138 (1991) 2334.

[5] S. Um, C.Y. Wang, K.S. Chen, Computational fluid modeling of proton exchange membrane fuel cells, *J. Electrochem. Soc.* 147 (2000) 12.

[6] L. Fortuna, S. Graziani, A. Rizzo and M.G. Xibilia, “Soft Sensors for Monitoring and Control of Industrial Processes,” series *Advances in Industrial Control*, Springer-Verlag, London Ltd, (2007).

[7] Liu J, On line soft sensor for polyethylene process with multiple production grades. 16th IFAC World Congress, Prague (2005).

[8] A. Di Bella; S. Graziani; G Napoli; M.G. Xibilia. A Comparative Analysis of the Influence of Methods for Outliers Detection on the Performance of Data Driven Models. In: *Instrumentation and Measurement Technology Conference Proceedings, 2007, IMTC 07, IEEE. Warsaw, 1-3 Maggio 07*, p. 1-5, (2007).

[9] Borairi M, Wang H, (1998) Actuator and sensor fault diagnosis of nonlinear dynamic systems via genetic neural networks and adaptive parameter estimation techniques. *Proc. of the IEEE International Conference on Control Applications* 1:278–282

[10] Frank PM, (1990) Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy – a survey and some new results. *Automatica* 26(3):459–474

[11] L. Fortuna, S. Graziani, G. Napoli, M.G. Xibilia - “Virtual Instruments for the what-if analysis of a process for pollution minimization in an industrial application” - *Proc. of the 14-th Mediterranean Conf. on Control and Automation, MED 06, Ancona, 28-30 June 06*

[12] S. Chen, S.A. Billings. P.M. Grant, Non Linear Systems identification using Neural Networks, *Int. Journal of Control*, Vol. 51, no. 6, pp. 1191-1214, Elsevier, 1990.

[13] K.S. Narendra, K. Parthasarathy, Identification and control of dynamical systems using neural networks, *IEEE Trans. Neural Networks* 1 (1990) 4-27.

[14] Haykin S, *Neural networks: a comprehensive foundation*, 2nd edn. Prentice Hall (1999)

[15] Cha´vez-Ramı´rez AU, et al., High power fuel cell simulator based on artificial neural network, *International Journal of Hydrogen Energy* (2009), doi:10.1016/j.ijhydene.2009.09.071.

[16] Jemie S, Hissel D, Pera MC, Kauffman JM. On-board fuel cell power supply modeling on the basis of NN methodology. *Journal of Power Sources* 2003;124: 479–86.

[17] El-Sharkh MY, Rahman A, Alam MS. Neural networks based control of active and reactive power of stand alone PEM fuel cell power plant. *Journal of Power Sources* 2004; 135:88–94.

[18] Bao Cheng, Ouyang Minggao, Yi Baolian. Modeling and control of air stream and hydrogen flow with recirculation in a PEM fuel cell systemdII. Linear and adaptive nonlinear control. *International Journal of Hydrogen Energy* 2006;31: 1897–913.

[19] S. Ou, L.E.K. Achenie, “A hybrid Neural Network model for PEMfuel cells”, *Journal of Power Sources*, 2004.

[20] Saengrung A, Abtahi A, Zilouchian A. Neural network model for a commercial PEM fuel cell system. *Journal of Power Sources* 2007;172:749–59.

[21] Ogaji SOT, Singh R, Pilidis P, diacakis M. Modeling fuel cell using artificial intelligence. *Journal of Power Sources* 2006; 154:192–7.