# Electrical Energy Consumption Forecast Using External Facility Data

Eugénia Vinagre, Luis Gomes, Zita Vale
GECAD - Knowledge Engineering and Decision Support Research Center
Polytechnic of Porto (IPP),
Porto, Portugal
empvm@isep.ipp.pt, lufog@isep.ipp.pt, zav@isep.ipp.pt

***Abstract*—Recent changes in the power systems gives place to the active consumers participation. The participation in demand response programs requires consumers to undertake strategic management of their consumption. Small and medium players should have the capability of performing day-ahead and hour-ahead load management which requires forecasting techniques applied to the consumption and generation. A good forecasting accuracy is very important for the quality of the management results but also very difficult to achieve. This paper proposes an artificial neural network based methodology to forecast the consumption in an office building. The considered building is equipped with a Supervisory Control and Data Acquisition (SCADA) system that stores data every 10 seconds. The stored data are used together with additional data, such as, the temperature and the solar radiation.**

## I. Introduction

The concept of Smart Grids (SG) has emerged during the last years and aims to address the most important problems in the power and energy field. Many relevant studies have been made in the area energy resources management but there are still many challenges to overcome. One of those challenges is the effectiveness in forecasting consumption and generation in order to ensure maximum efficiency of SG [1].

The balance between consumption and generation is a critical issue when dealing with massive Distributed Generation (DG) [2]. Energy consumption depends on several factors (e.g. weather conditions, annual seasons; consumer behavior, etc.), which makes its prediction a complex and difficult task [2]-[5]. Despite its complexity, the electrical consumption forecasting process should be evaluated in detail and not as a whole to ensure its full integration into demand response programs.

Demand response programs [6],[7], enable the active participation of small and medium players in the SG context. One of the main goals of the SG is to enable this participation autonomously without harming the comfort of consumers and the efficient operation of energy services [8].

In the SG context, resources forecasting is important at many levels, from the SG point of view to the small player perspective. This paper focuses on the small players' point of view. The forecasting of resources will be applied to an office building that uses a Supervisory Control and Data Acquisition (SCADA) system. The SCADA Office Intelligent Context Awareness Management (SOICAM) system manages the building energy consumption, including the control of HVAC devices [9]. SOICAM also stores the consumption data in a 10 seconds rate.

The main contribution of this paper is the development and implementation of a forecasting methodology that integrates the consumption data from the SCADA with additional data.

After this introductory section, Section II presents the proposed Artificial Neural Network (ANN). Section III describes the SOICAM system, where the ANN is applied. Section IV presents the case study scenario and Section V shows the obtained results. Section VI presents a discussion of the main conclusions of the paper.

## II. Artificial Neural Network for Consumption Forecasting

The configuration of a neural network is a complex task. There are many issues that may affect its performance, e.g. number of layers, number of nodes in each layer, connection between layers, activation functions, training algorithms, performance measures, etc. This section discusses some of the most important of these issues.

### A. Type of Neural Network

Neural networks are widely referenced in the literature as one of the best methodologies used in time series forecasting [10-13]. The present work uses Multilayer Perceptron (MLP) networks. The MPL has been used successfully in solving numerous different and highly complex problems. The topology of such networks is organized in layers (i.e. one input layer, one or several intermediate layers – hidden, and an output layer), and each layer includes 1 to *n* artificial neurons. The interconnections of nodes in layers is very important and determines the behaviour of the network. For most forecasting approaches, the networks are fully connected, i.e. all nodes in one layer are fully connected to all nodes in the following higher layer except for the output layer. Fig. 1 shows the architecture used in this paper. The architecture is used for the scenarios presented in Section IV. However, the Environment Context input change according to the scenario (i.e. can be used the temperature, solar radiation, and more).

The present work was done and funded in the scope of the following projects: EUREKA - ITEA2 Project SEAS with project number 12004; EUREKA - ITEA2 Project FUSE-IT with project number 13023; and UID/EEA/00760/2013 funded by FEDER Funds through COMPETE program and by National Funds through FCT.

978-1-4799-7560-0/15 $31.00 © 2015 IEEE
DOI 10.1109/SSCI.2015.101

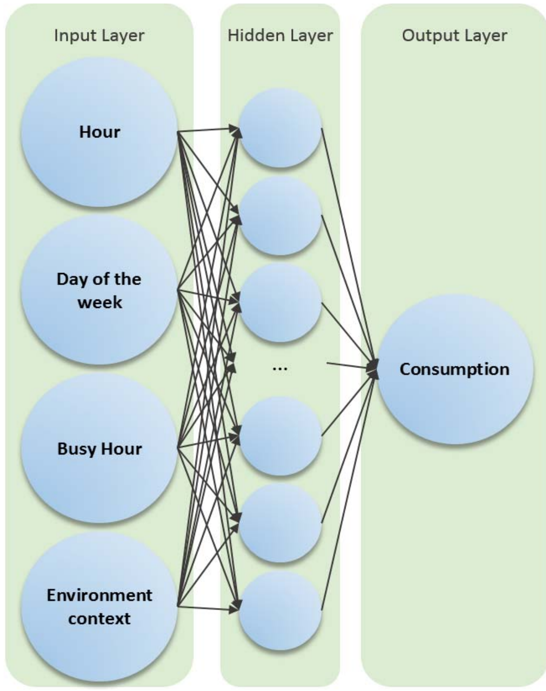

Fig. 1. Architecture of MLP - Multilayer Percepton

## B. Hidden layers and nodes

The hidden layers and the respective nodes play a very important role in many successful applications of neural networks. The hidden nodes in the hidden layers are what allows neural networks to detect the feature, to capture the pattern in the data and to perform complicated nonlinear mapping between input and output variables. Much theoretical work has been done in the area, showing that most authors use only one hidden layer for forecasting purposes [10].

The problem of determining the optimal number of hidden nodes is very important to obtain a good prediction. In general, networks with fewer hidden nodes are preferable as they usually have better generalization ability, this dealing better with the overfitting problem. But networks with too few hidden nodes may not have enough power to model and learn the data. There is no rule to select this parameter, although some systematic approaches are reported. For example, Jeff Heaton [10] proposes the following guidelines:

- The number of hidden neurons should be between the size of the input layer and the size of the output layer.
- The number of hidden neurons should be 2/3 the size of the input layer, plus the size of the output layer.
- The number of hidden neurons should be less than twice the size of the input layer.

Other authors have proposed that the number of hidden nodes should be calculated according to the input node number (*n*) in the following proportions: 2*n* [11], *n* [12], *n*/2 [13]. However, the most common way to determine the number of hidden nodes is through experience or trial and error.

## C. Activation function

The activation function, also called the transfer function, determines the relationship between inputs and outputs of a node in a network [14]. In general, the activation function introduces a degree of nonlinearity that is essential for most ANN applications. The main activation functions referenced in the literature are:

- Sigmoid (logistic) function

  f(x) = (1 + exp(-x))-1 (1)

- Hyperbolic tangent (tanh) function

  f(x)=(exp(x)-exp(-x))/(exp(x)+exp(-x)) (2)

- Sine or Cosine function

  f(x)=sin(x) or f(x)=cos(x) (3)

- Linear function

  f(x)=x (4)

Among them, the logistic transfer function is the most popular choice for forecasting problems.

## D. Training algorithm

The training of the MLP networks is usually done using the backpropagation algorithm. Backpropagation is a supervised training and deterministic algorithm that implements the decreasing gradient method in the sum of squared errors. The mechanism of the algorithm is to correct the error during training (i.e. the output layer errors are propagated to the inner layers to correct the function of the weights in the current point) [15]. Another relevant issue in training an ANN is the specification of an objective function or a cost function. Typically MSE (mean squared error) is used. MSE formulation is specified in (1).

$$MSE = \frac{1}{n}\sum_{t=1}^{n}\left(F_t - A_t\right)^2 \qquad (1)$$

where *n* is the number of samples and *F*t is our estimation of *A*t

## E. Forecast accuracy measures

Although there can be many performance measures for an ANN forecast, such as the modeling time and training time, the ultimate and the most important measure of performance is the prediction accuracy it can achieve beyond the training data. An accuracy measure is often defined in terms of the forecasting error, which is the difference between the actual (desired) and the predicted value. There are a number of measures of accuracy in the forecasting literature and each has advantages and limitations [16]. The MAPE (Mean Absolute Percentage Error) and the SMAPE (Symmetric Mean Absolute Percentage Error) are among the most frequently used measures, therefore they have been applied in the case study of this paper.

MAPE shows the error in percentage, which facilitates the comparison of results between different time series. This measure is calculated as shown in (2).

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{A_t - F_t}{A_t}\right| \qquad (2)$$

where $At$ represents the actual values and $Ft$ the predicted values by the neural network. SMAPE also shows the error percentage but it is less sensitive to large amplitudes of the series of values range to predict. The formula for its calculation is shown in (3).

$$SMAPE = \frac{1}{N}\sum_{t=1}^{N}\frac{|F_t - A_t|}{(|A_t| + |F_t|)/2} \quad (3)$$

## III. SCADA Office Intelligent Context Awareness Management System

Smart grids enable the small and medium players' active participation in the network operation. The use of Demand Response (DR) programs and events and the application of Demand Side Management (DSM) techniques are useful tools to achieve the desired participation from those players. Some on-going studies are trying to figure out the better and most efficient way to bring those players to the smart grid context.

The SCADA Office Intelligent Context Awareness Management (SOICAM) was developed for test bed experiences [9]. The system is implemented in real facilities in the GECAD research center, located in the Institute of Engineering – Polytechnic of Porto (ISEP/IPP), Portugal. The system is implementation in real facilities used by more than 30 researchers in a daily basis.

The SOICAM system is divided in two levels: the physical level (infrastructures) and the operational level. The operational level uses a Multi-Agent System (MAS) approach to represent, control and manage the 3 integrated facilities. The system also includes a microgrid dedicated agent to aggregate the facilities and their services and resources. The use of 3 real facilities and an aggregator agent (microgrid agent) is not enough for some tests and proof of concepts. For this reason the SOICAM uses simulation in order to integrate more agents that otherwise cannot be placed in the system. Fig. 2 shows the architecture of SOICAM. Buildings I, N and F are ISEP buildings. Houses 1, 2 and 3 are simulated facilities. The simulated facilities use real electrical energy consumption measurements. The use of real measurements enables the profile simulation of the simulated facility. As can be seen in Fig. 2, SOICAM uses 3 web services to provide different services to the software agents. The use of web services allows the efficiency of computation inside the agents.

The real resources, placed inside the facilities, were integrated in the agents using three load groups: Lighting Group; HVAC Group; and Electrical Sockets Group. The separation of loads facilitates the analysis by the system and improves the control and management aspects. For security reasons, SOICAM only allows the control of HVAC systems. The control of lights and electrical sockets is not allowed. Even so, the SOICAM can communicate with the facility users in order to ask for and achieve lower consumption in these two groups. This limitation was built to prevent the cut of important work and ensure the minimal conditions in the facilities. The control of the HVAC system is also limited to a number of offices, for instance, the control of HVAC systems inside the two server and data center rooms is not available. Otherwise, an inappropriate use or a system malfunction could represent heavy lost to the research center.

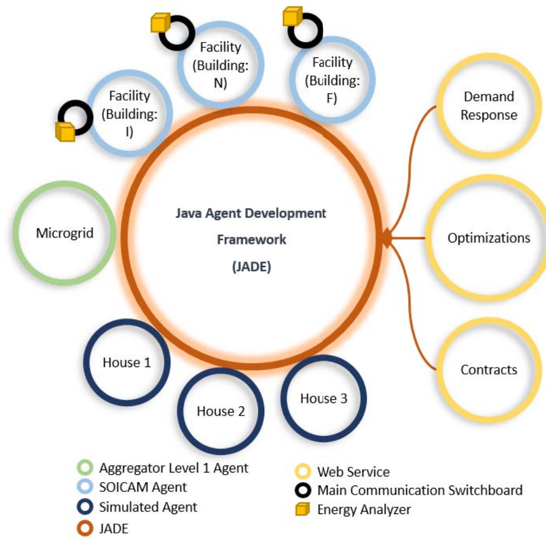


Fig. 2. SOICAM Operational Level architecture

The Physical Level, in Fig. 3, is composed by multiple electrical switchboards with three-phase energy analyzers. Each phase is used to measure one of the three load groups: Lighting Group; HVAC Group; and Electrical Sockets Group. The HVAC control is done using a digital auxiliary port, built in the energy analyzer, and a relay (24V/DC to 230V/AC) and a step relay in order to minimize the impact and deterioration of the relay. Each building contains a Main Communication Hardware that can be a Programmable Logic Controller or a custom built hardware. The Main Communication Hardware communicates with the Energy Analyzers, located in several electrical switchboards, through RS-485 communication using the Modbus/RTU protocol. The time step of the measures is different for each building depending on the communication hardware (from 10 seconds to 40 seconds).

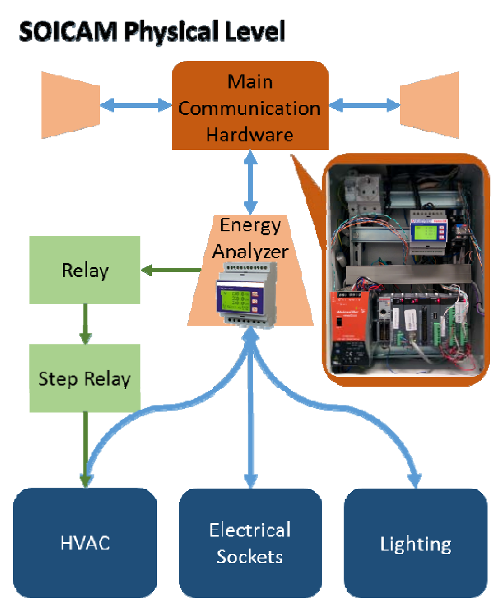


Fig. 3. SOICAM Physical Level architecture

The SOICAM data is stored in a SQL database. The data register has started in June, 2014. The register from buildings I and N only includes the consumption, reaching more than 30 researchers daily. F building has a 1kW wind turbine and a 1kW photovoltaic array. This building does not include any office, and it is only used for distributed generation.

The use of SOICAM for DSM and DR tests is vital to validate new methodologies and algorithms. The data produced is also particularly important for user profiling and data analysis. The system also stores outside data regarding environmental conditions. The data provided by ISEP weather station (ISEP Meteo website [17]) is stored using 5 minute periods. The mixture of outside data and SOICAM data can be very important to improve the quality of forecasting algorithms, as studied in this paper.

## IV. Case Study Scenario

This section explains the preprocessing and the processing of the data to be used in the three scenarios of the case study.

### A. Preprocessing

For the present case study, real data of environmental conditions and energy consumption of the building N has been used. The observations are recorded each 10 seconds, and are grouped in three groups (HVAC, lighting and loads connected to sockets) as described in Section III.

The analyzed time series regard the period from October 2014 to April 2015. Initially these data were cleaned and integrated with context data (i.e. day of the week and hour of work) and external environmental data (i.e. temperature, temperature felt, radiation, precipitation and humidity) for the same period. The observations were organized into three basic scenarios as shown in table I:

TABLE I. Scenarios created for the study of temporal series of power consumption

| *Scenario* | *Training data set* | *Test /validation data set* |
|---|---|---|
| S1 | October 2014 – March 2015 | April 2015 |
| S2 | January 2015 – March 2015 | April 2015 |
| S3 | March 2015 | April 2015 |

The data normalization was performed automatically by the software (i.e. tolbox *nnet matlab*) in the processing stage.

### B. Processing

To implement the neural network we use the *feedforwardnet* function of Matlab. This function creates a multilayer feedforward network. The following configuration have been used:

- Number hidden layers: 1;
- Number of neurons in the hidden layer: 10;
- Training function: Levenberg-Marquardt backpropagation [18];
- Transfer function for hidden layers: Hyperbolic tangent sigmoid transfer function [19];
- Transfer function for the output layer: Linear transfer function [13].

This configuration has been adopted for this case study after several experiments. The results achieved with the specified configuration were always better when compared with the results of the other alternative configurations. It is also noteworthy to refer that good results were achieved when the network was configured with 5 and 8 neurons in the hidden layer; however, these were still slightly worse than those obtained with 10 neurons in the hidden layer.

The definition of the input and output neurons is shown by Table II, including the number of nodes and their description.

The ANN_h, ANN_l and ANN_s networks were tested with each of the external variables. Each of the networks defined in Table II were tested according to the scenarios in Table I. The network was tested in the forecasting, hour by hour, of each day of the month of April 2015. The output of the ANN returns the forecasting of 24 hours, between midnight and 23:59, of the next day.

TABLE II. Neurons definition input / output

| *Id ANN* | *Input neurons* | | *Output neurons* | |
|---|---|---|---|---|
| | *No* | *Description* | *No* | *Description* |
| ANN_1 | 3 | - Context data* | 1 | - Total consumption/h |
| ANN_2 | 3 | - Context data | 3 | - HVAC consumption<br>- lights consumption<br>- sockets consumption |
| ANN_h | 4 | - Context data;<br>- External environmental data** | 1 | - HVAC consumption |
| ANN_l | 4 | - Context data;<br>- External environmental data | 1 | - light consumption |
| ANN_s | 4 | - Context data;<br>- External environmental data | 1 | - sockets consumption |

* *The context data refer to: hour; day of the week and hour of work (the hour of work is characterized as: time of high, medium and low occupancy of the building);*

** *The environment data refer to: temperature, temperature felt, solar radiation, precipitation and humidity.*

## V. Results

The first tests were made with the ANN_1 network in order to select which scenario was more favorable for the prediction of consumption. We conclude that the most appropriate series to train the network, leading to the best results, was the S1 serie (see Fig. 4 and Table III).

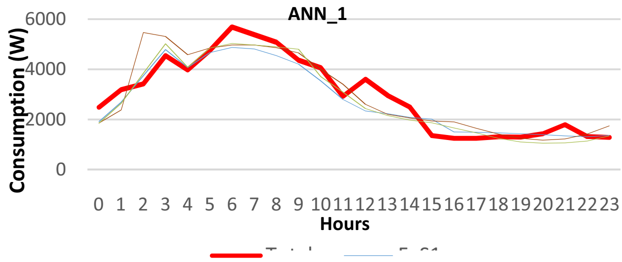


Fig. 4. ANN_1 forecasting results using the three alternative time series

The variation of people in the "research building" is quite irregular throughout the year. Because of this, power consumption also suffers large variations over time, making it difficult to obtain more accurate estimates (i.e. below 10%).

TABLE III. RESULTS OF ANN_1 USING THE THREE TIME SERIES

| *Forecasted time serie* | *MAPE %* | *SMAPE %* |
|---|---|---|
| *F_S1* | 14.0 | 14.3 |
| *F_S2* | 20.5 | 19.8 |
| *F_F3* | 16.2 | 17.2 |

The following tests are focused on ANN_2 network. These tests have also resulted in good results; however, this is only verified when the individual values obtained by the three outputs are summed to calculate the total consumption per hour. The results obtained for each of the sub consumption groups (i.e. HVAC, Lights and Sockets) were not satisfactory, as shown by Fig. 5 and Table IV.

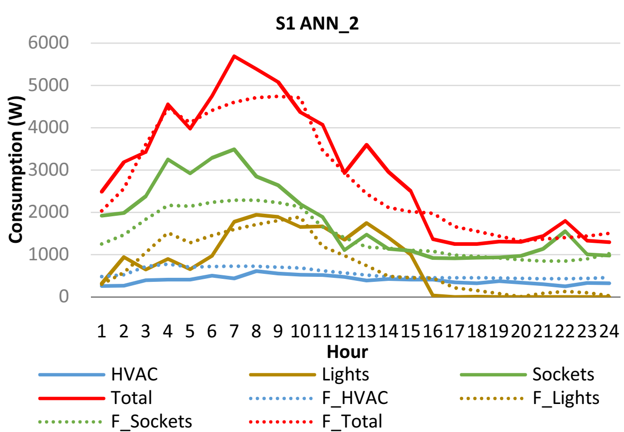


Fig. 5. ANN_2 forecasting results using the S1 time series

TABLE IV. RESULTS OF ANN_2 USING S1

| *Forecasting* | *MAPE %* | *SMAPE %* |
|---|---|---|
| *F_HAVAC* | 44.8 | 34.6 |
| *F_Lights* | 59.1 | 55.9 |
| *F_Sockets* | 18.1 | 20.5 |
| *F_Total* | 15.0 | 15.3 |

Finally, ANN_h, ANN_l and ANN_s networks are tested, with the aim of analyzing and selecting the external variables which could help improving the results obtained by the network ANN_2. The conclusion is that not all the sub consumer groups were similarly influenced by external variables. Additionally, the scenario that has achieved the best results was not equal for all networks. The achieved results led to the following conclusions:

- HAVC Network (ANN_h) - is strongly influenced by temperature and the best results were obtained with the S3 time series;
- Light Network (ANN_l) - is strongly influenced by precipitation and the best results have been achieved by using the S1 time series.
- Sockets Network (ANN_s) – the precipitation is the variable that had the most positive influence on the results; and S3 time series was the one that enabled the achievement of the best forecasting results.

Fig. 6 and Table 5 show the results of the networks that have achieved the best results.

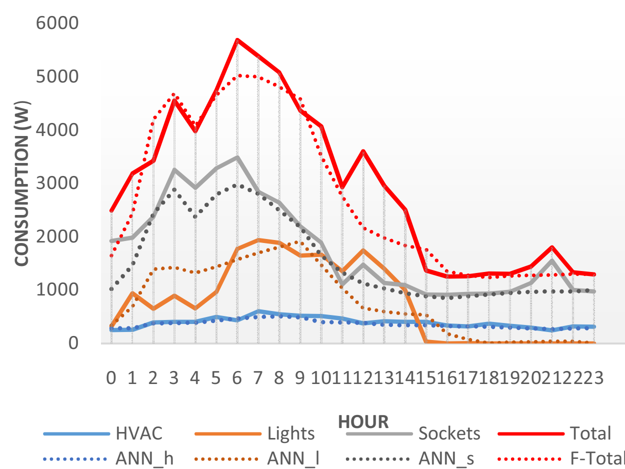


Fig. 6. Forecasting results of the best considered neural networks

TABLE V. RESULTS OF THE BEST NEURAL NETWORKS

| *Forecasting* | *MAPE %* | *SMAPE %* |
|---|---|---|
| *ANN_h* | 10.3 | 10.8 |
| *ANN_l* | 18.1 | 19.6 |
| *ANN_s* | 12.6 | 14.2 |
| *F_Total* | 13.6 | 14.9 |

The fact that the networks are not equally influenced by the same variables has made it impossible to extract conclusions that allow a direct reconfiguration of ANN_2 in order to obtain better results. The average processing time of these networks is of 2 seconds, which allows running the forecasting process by the three networks separately and sequentially. If, in the future, the performance of the process becomes critical to the SOICAM system, the process can be distributed across multiple servers to improve the runtime of the forecast.

## VI. Conclusions

The main conclusion from this case study is that it is possible to get a detailed forecast of energy consumption for every hour of the following day without loss of efficiency when compared to the results obtained by the forecast using only the total consumption. Subsequently, more accurate forecasts using external environmental data, and focused on specific consumption groups can be performed.

As future work, the study of the same time series according to the same methodology, to predict each 5 minutes of the next hour is proposed. When the historical data in this series completes a cycle exceeding one year, it is proposed to study network behavior with the inclusion of the variable "epoch of year" with respect to its efficiency and its performance.

## References


[1] M. Paun, G. Lorenz, P. Schlosser, ―Smart grids and networks of the future - eurelectric views , CIRED Workshop 2010: Sustainable Distribution Asset Management & Financing, Lyon, França, 7-8 de Junho de 2010

[2] Rueda, J.L.; Guaman, W.H.; Cepeda, J.C.; Erlich, I; Vargas, A, "Hybrid Approach for Power System Operational Planning With Smart Grid and Small-Signal Stability Enhancement Considerations," Smart Grid, IEEE Transactions on , vol.4, no.1, pp.530,539, March 2013;

[3] Nakawiro, W.; Erlich, I; Rueda, J.L., "A novel optimization algorithm for optimal reactive power dispatch: A comparative study," Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), 2011 4th International Conference on, vol., no., pp.1555,1561, 6-9 July 2011;doi: 10.1109/DRPT.2011.5994144

[4] Venayagamoorthy, G.; Rohrig, K.; Erlich, I, "One Step Ahead: Short-Term Wind Power Forecasting and Intelligent Predictive Control Based on Data Analytics,"Power and Energy Magazine, IEEE , vol.10, no.5, pp.70,78, Sept.-Oct. 2012; doi: 10.1109/MPE.2012.2205322

[5] Fortmann, J.; Engelhardt, S.; Kretschmann, J.; Feltes, C.; Erlich, I, "New Generic Model of DFG-Based Wind Turbines for RMS-Type Simulation," Energy Conversion, IEEE Transactions on , vol.29, no.1, pp.110,118, March 2014; doi: 10.1109/TEC.2013.2287251

[6] Vale, Z.; Morais, H.; Ramos, S.; Soares, J.; Faria, P., "Using data mining techniques to support DR programs definition in smart grids," Power and Energy Society General Meeting, 2011 IEEE , vol., no., pp.1,8, 24-29 July 2011; doi: 10.1109/PES.2011.6039081

[7] Ramos, S.; Morais, H.; Vale, Z.; Faria, P.; Soares, J., "Demand response programs definition supported by clustering and classification techniques," Intelligent System Application to Power Systems (ISAP), 16th International Conference on , vol., no., pp.1,6, 25-28 Sept. 2011

[8] L. Gomes, P. Faria, H. Morais, Z. Vale, C. Ramos, "Distributed Agent-Based Intelligent System for Demand Response Program Simulation in Smart Grids," IEEE Intelligent Systems, vol. 29,pp. 56-65, January/February 2014.

[9] L. Gomes, F. Fernandes, P. Faria, M. Silva, Z. Vale, C. Ramos, "Contextual and Environmental Awareness Laboratory for Energy Consumption Management," Clemson University Power Systems Conference 2015 (PSC 2015), Madren Conference Center, Clemson, South Carolina, USA, 10-13 Mar. 2015.

[10] Jeff Heaton, "Introduction to Neural Networks for Java", 2nd Edition, Heaton Research, 2008; ISBN: 1604390085 9781604390087

[11] Lippmann, R.P., "An introduction to computing with neural nets", IEEE ASSP Magazine, 4–22, April-1987

[12] Wong, F.S.,"Time series forecasting using backpropagation neural networks", Neurocomputing 2, 147–159, 1991

[13] Tang, Z., Fishwick, P.A., "Feedforward neural nets as models for time series forecasting", ORSA Journal on Computing 5 (4), 374 – 385, 1993

[14] Phat, V.N.; Trinh, H., "Exponential Stabilization of Neural Networks With Various Activation Functions and Mixed Time-Varying Delays", Neural Networks, IEEE Transactions on vol.21, no.7, pp.1180-1184, July 2010; doi: 10.1109/TNN.2010.2049118

[15] Rumelhart, D. E; Hinton G. E; Williams, R. J., "Learning representations of back- propagating error" London: Nature, v. 323, pp. 533-536, 1986

[16] Makridakis, S., Wheelwright, S.C., McGee, V.E., "Forecasting: Methods and Applications", 2nd ed. John Wiley, New York, 1983

[17] Meteo.isep.ipp.pt, 'Meteorologia no Instituto Superior de Engenharia do Porto', 2015. [Online]. Available: http://meteo.isep.ipp.pt/. [Accessed: 02- Oct- 2015].

[18] Hagan, M.T., and M. Menhaj, "Training feed-forward networks with the Marquardt algorithm," IEEE Transactions on Neural Networks, Vol. 5, No. 6, 1999, pp. 989–993, 1994.

[19] Vogl, T.P., J.K. Mangis, A.K. Rigler, W.T. Zink, and D.L. Alkon, "Accelerating the convergence of the backpropagation method," Biological Cybernetics, Vol. 59, 1988, pp. 257–263.