An Artificial Neural Network Approach for Short-Term Wind Power Forecasting in Portugal

J. P. S. Catalão, *Member, IEEE*, H. M. I. Pousinho, *Student Member, IEEE*, and V. M. F. Mendes

Abstract—**This paper presents an artificial neural network approach for short-term wind power forecasting in Portugal. The increased integration of wind power into the electric grid, as nowadays occurs in Portugal, poses new challenges due to its intermittency and volatility. Hence, good forecasting tools play a key role in tackling these challenges. The accuracy of the wind power forecasting attained with the proposed approach is evaluated, reporting the numerical results from a real-world case study.**

Index Terms—**Artificial neural networks, forecasting, wind power.**

I. Introduction

WIND-DRIVEN power resources have become increasingly important in the planning and operation of electric power systems [1]. In Portugal, the wind power goal foreseen for 2010 was established by the government as 3750 MW, representing about 25% of the total installed capacity by 2010 [2]. This value has been raised to 5100 MW, by the most recent governmental goals for the wind sector. Hence, Portugal has one of the most ambitious goals in terms of wind power, and in 2006 was the second country in Europe with the highest wind power growth.

The wind energy is free, so all wind-generated electric energy is accepted as it comes, i.e. as it is available. However, its availability is not known in advance. Because of the increasing penetration of wind resources in power systems, efforts have been made to predict the wind behavior and the corresponding electric energy production [1].

Short-term wind power forecasting is an extremely important field of research for the energy sector, as the system operators must handle an important amount of fluctuating power from the increasing installed wind power capacity. The time scales concerning short-term prediction are in the order of some days (for the forecast horizon) and from minutes to hours (for the time-step) [3].

In the technical literature, several methods to predict wind power have been reported, namely physical and statistical methods. The physical method requires a lot of physical considerations to reach the best prediction precision. For a physical model, the input variables will be the physical or meteorology information, such as description of orography, roughness, obstacles, pressure, and temperature. The statistical method aims at finding the relationship of the on-line measured power data. For a statistical model, the historical data of the wind farm may be used. Physical method has advantages in long-term prediction while statistical method does well in short-term prediction [4].

Conventional statistical models are identical to the direct random time-series model, including auto regressive (AR), and auto regressive integrated moving average (ARIMA) [5] models. The persistence models are considered as the simplest time-series models. They can surpass many other models in very short-term prediction. In spite of the unstable forecasting efficiency, they have been widely used in practice [4].

In the recent years, some new methods are catching researcher's attention, namely methods based on artificial intelligence like artificial neural network (ANN) [6], fuzzy logic and neuro-fuzzy [7], [8], evolutionary algorithms [9], and some hybrid methods [10], [11]. The accurate comparison of all the methods is quite difficult because these methods depend on different situations, and the data collection is a formidable task. However, there are some comparison and the approximate results, which proved that the artificial-based models outperformed others in short-term prediction [4].

ANNs are simple, but powerful and flexible tools for forecasting, provided that there are enough data for training, an adequate selection of the input-output samples, an appropriate number of hidden units and enough computational resources available. Also, ANNs have the well-known advantages of being able to approximate nonlinear functions and being able to solve problems where the input-output relationship is neither well defined nor easily computable, because ANNs are data-driven. Three-layered feedforward ANNs are specially suited for forecasting, implementing nonlinearities using sigmoid functions for the hidden layer and linear functions for the output layer [12].

This paper presents a successful application of using an ANN approach to forecast short-term wind power in Portugal. ANNs techniques are relatively easy to implement and show good performance being less time consuming than traditional time series techniques.

This paper is structured as follows. Section 2 presents the ANN approach. Section 3 provides the different criterions used to assess the behavior of the proposed approach. Section 4 presents the numerical results. Finally, Section 5 outlines the conclusions.

J. P. S. Catalão and H. M. I. Pousinho are with the University of Beira Interior, Covilha, Portugal (e-mail: catalao@ubi.pt; hmi-21@hotmail.com).

V. M. F. Mendes is with the Instituto Superior de Engenharia de Lisboa, Lisbon, Portugal (e-mail: vfmendes@isel.pt).

II. Artificial Neural Network Approach

ANNs are highly interconnected simple processing units designed in a way to model how the human brain performs a particular task [13]. Each of those units, also called neurons, forms a weighted sum of its inputs, to which a constant term called bias is added. This sum is then passed through a transfer function: linear, sigmoid or hyperbolic tangent.

Multilayer perceptrons are the best known and most widely used kind of ANN. The units are organized in a way that defines the network architecture. In feedforward networks, units are often arranged in layers: an input layer, one or more hidden layers and an output layer.

In order to find the optimal network architecture, several combinations should be evaluated. These combinations include networks with different number of hidden layers, different number of units in each layer and different types of transfer functions. The configuration chosen consists of a one hidden layer that uses a hyperbolic tangent sigmoid transfer function and a one unit output layer with a pure linear transfer function. This configuration has been proven to be a universal mapper, provided that the hidden layer has enough units [14]. On one hand, if there are too few units, the network will not be flexible enough to model the data well and, on the other hand, if there are too many units, the network may over-fit the data. Typically, the number of units in the hidden layer is chosen by trial and error, selecting a few alternatives and then running simulations to find out the one with the best results.

Forecasting with ANNs involves two steps: training and learning. Training of feedforward networks is normally performed in a supervised manner. One assumes that a training set is available, given by the historical data, containing both inputs and the corresponding desired outputs, which is presented to the network. The adequate selection of inputs for ANN training is highly influential to the success of training. In the learning process an ANN constructs an input-output mapping, adjusting the weights and biases at each iteration based on the minimization of some error measure between the output produced and the desired output. The error minimization process is repeated until an acceptable criterion for convergence is reached.

The most common learning algorithm is the backpropagation algorithm [15], in which the input is passed layer through layer until the final output is calculated, and it is compared to the real output to find the error. The error is then propagated back to the input adjusting the weights and biases in each layer. The standard backpropagation learning algorithm is a steepest descent algorithm that minimizes the sum of square errors. However, the standard backpropagation learning algorithm is not efficient numerically and tends to converge slowly.

An algorithm that trains an ANN 10 to 100 times faster than the usual backpropagation algorithm is the Levenberg-Marquardt algorithm. While backpropagation is a steepest descent algorithm, the Levenberg-Marquardt algorithm is a variation of Newton's method [16]. Hence, a three-layered feedforward ANN trained by the Levenberg-Marquardt algorithm is considered in this paper, as in [17].

Newton's update for minimizing a function $V(\boldsymbol{x})$ with respect to the vector $\boldsymbol{x}$ is given by:

$$\Delta(\boldsymbol{x}) = -\left[\nabla^2 V(\boldsymbol{x})\right]^{-1} \nabla V(\boldsymbol{x}) \tag{1}$$

where $\nabla^2 V(\boldsymbol{x})$ is the Hessian matrix and $\nabla V(\boldsymbol{x})$ is the gradient vector. Assuming that $V(\boldsymbol{x})$ is the sum of square errors, given by:

$$V(\boldsymbol{x}) = \sum_{h=1}^{N} e_h^2(\boldsymbol{x}) \tag{2}$$

then:

$$\nabla V(\boldsymbol{x}) = 2\, \boldsymbol{J}^T(\boldsymbol{x})\, \boldsymbol{e}(\boldsymbol{x}) \tag{3}$$

$$\nabla^2 V(\boldsymbol{x}) = 2\, \boldsymbol{J}^T(\boldsymbol{x})\, \boldsymbol{J}(\boldsymbol{x}) + 2\, \boldsymbol{S}(\boldsymbol{x}) \tag{4}$$

where $\boldsymbol{e}(\boldsymbol{x})$ is the error vector, $\boldsymbol{J}(\boldsymbol{x})$ is the Jacobian matrix given by:

$$\boldsymbol{J}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial e_1(\boldsymbol{x})}{\partial x_1} & \frac{\partial e_1(\boldsymbol{x})}{\partial x_2} & \cdots & \frac{\partial e_1(\boldsymbol{x})}{\partial x_n} \\ \frac{\partial e_2(\boldsymbol{x})}{\partial x_1} & \frac{\partial e_2(\boldsymbol{x})}{\partial x_2} & \cdots & \frac{\partial e_2(\boldsymbol{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial e_N(\boldsymbol{x})}{\partial x_1} & \frac{\partial e_N(\boldsymbol{x})}{\partial x_2} & \cdots & \frac{\partial e_N(\boldsymbol{x})}{\partial x_n} \end{bmatrix} \tag{5}$$

and where $\boldsymbol{S}(\boldsymbol{x})$ is given by:

$$\boldsymbol{S}(\boldsymbol{x}) = \sum_{h=1}^{N} e_h(\boldsymbol{x})\, \nabla^2 e_h(\boldsymbol{x}) \tag{6}$$

Neglecting the second-order derivatives of the error vector, i.e., assuming that $\boldsymbol{S}(\boldsymbol{x}) \approx 0$, the Hessian matrix is given by:

$$\nabla^2 V(\boldsymbol{x}) = 2\, \boldsymbol{J}^T(\boldsymbol{x})\, \boldsymbol{J}(\boldsymbol{x}) \tag{7}$$

and substituting (7) and (3) into (1) we obtain the Gauss-Newton update, given by:

$$\Delta(\boldsymbol{x}) = -\left[\boldsymbol{J}^T(\boldsymbol{x})\, \boldsymbol{J}(\boldsymbol{x})\right]^{-1} \boldsymbol{J}^T(\boldsymbol{x})\, \boldsymbol{e}(\boldsymbol{x}) \tag{8}$$

The advantage of Gauss-Newton over the standard Newton's method is that it does not require calculation of second-order derivatives. Nevertheless, the matrix $\boldsymbol{J}^T(\boldsymbol{x})\, \boldsymbol{J}(\boldsymbol{x})$ may not be invertible. This is overcome with the Levenberg-Marquardt algorithm, which consists in finding the update given by:

$$\Delta(\boldsymbol{x}) = -\left[\boldsymbol{J}^T(\boldsymbol{x})\, \boldsymbol{J}(\boldsymbol{x}) + \mu\, \boldsymbol{I}\right]^{-1} \boldsymbol{J}^T(\boldsymbol{x})\, \boldsymbol{e}(\boldsymbol{x}) \tag{9}$$

where parameter μ is conveniently modified during the algorithm iterations.

When μ is very small or null the Levenberg-Marquardt algorithm becomes Gauss-Newton, which should provide faster convergence, while for higher μ values, when the first term within brackets of (9) is negligible with respect to the second term within brackets, the algorithm becomes steepest descent. Hence, the Levenberg-Marquardt algorithm provides a nice compromise between the speed of Gauss-Newton and the guaranteed convergence of steepest descent [16].

III. Forecasting Accuracy Evaluation

To evaluate the accuracy of the ANN approach in forecasting wind power, different criterions are used. This accuracy is computed in function of the actual wind power that occurred. The mean absolute percentage error (MAPE) criterion, the sum squared error (SSE) criterion, and the standard deviation of error (SDE) criterion, are defined as follows.

The MAPE criterion is given by:

$$MAPE = \frac{100}{N} \sum_{h=1}^{N} \frac{|\hat{p}_h - p_h|}{\bar{p}} \tag{10}$$

$$\bar{p} = \frac{1}{N} \sum_{h=1}^{N} p_h \tag{11}$$

where $\hat{p}_h$ and p_h are respectively the forecasted and actual wind power at hour h, $\bar{p}$ is the average wind power and N is the number of forecasted hours.

The SSE criterion is given by:

$$SSE = \sum_{h=1}^{N} (\hat{p}_h - p_h)^2 \tag{12}$$

The SDE criterion is given by:

$$SDE = \sqrt{\frac{1}{N} \sum_{h=1}^{N} (e_h - \bar{e})^2} \tag{13}$$

$$e_h = \hat{p}_h - p_h \tag{14}$$

$$\bar{e} = \frac{1}{N} \sum_{h=1}^{N} e_h \tag{15}$$

where e_h is the forecast error at hour h and $\bar{e}$ is the average error.

IV. Numerical Results

The proposed ANN approach has been applied for wind power forecasting in Portugal. Wind power forecasting is computed using only historical data. For the sake of clear comparison, no exogenous variables are considered.

The following days are randomly selected: July 3, 2007, October 31, 2007, January 14, 2008, and April 2, 2008, corresponding to the four seasons of the year. Hence, days with particularly good wind power behavior are deliberately not chosen. This results in an uneven accuracy distribution throughout the year that reflects reality.

Numerical results with the proposed ANN approach are shown in Figs. 1–4 respectively for the winter, spring, summer and fall days. Each figure shows the actual wind power, solid line, together with the forecasted wind power, dashed line.

Table I presents the values for the criterions to evaluate the accuracy of the proposed ANN approach in forecasting wind power. The first column indicates the day, the second column presents the MAPE, the third column presents the square root of the SSE, and the fourth column presents the SDE.

A good accuracy of the proposed ANN approach was ascertained. The MAPE has an average value of 7.26%. Moreover, the average computation time is less than 5 seconds. All the cases have been run on a PC with 1 GB of RAM and a 2.0-GHz-based processor.

For comparison purposes, the average MAPE value for a persistence approach would be 19.05%. The persistence approach assumes that the predicted value of the next step in the future is the last measured value. The persistence approach has proven to be a useful first approximation for short-term wind power forecasting and provides a benchmark against which to compare alternative techniques.

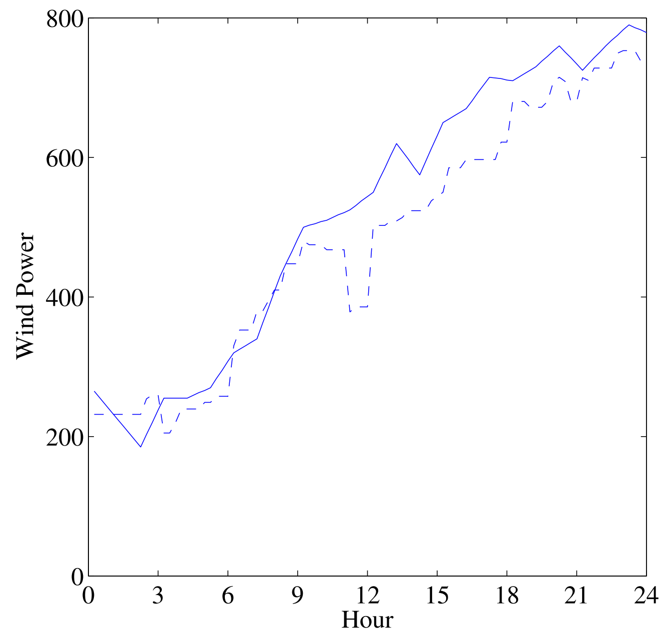

Fig. 1. Winter day: actual wind power, solid line, together with the forecasted wind power, dashed line, in megawatt.

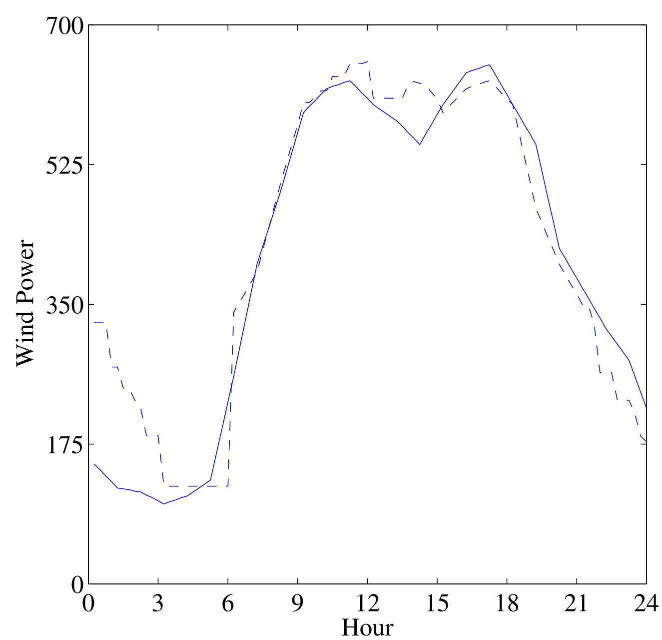

Fig. 2. Spring day: actual wind power, solid line, together with the forecasted wind power, dashed line, in megawatt.

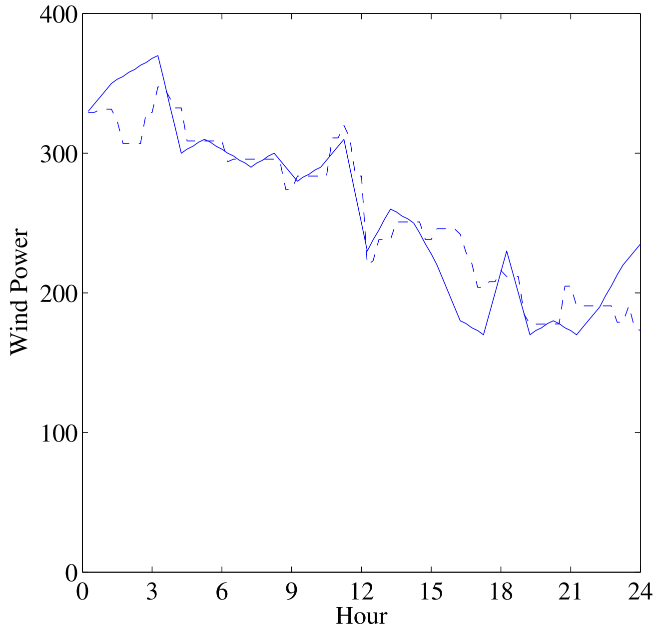

Fig. 3. Summer day: actual wind power, solid line, together with the forecasted wind power, dashed line, in megawatt.

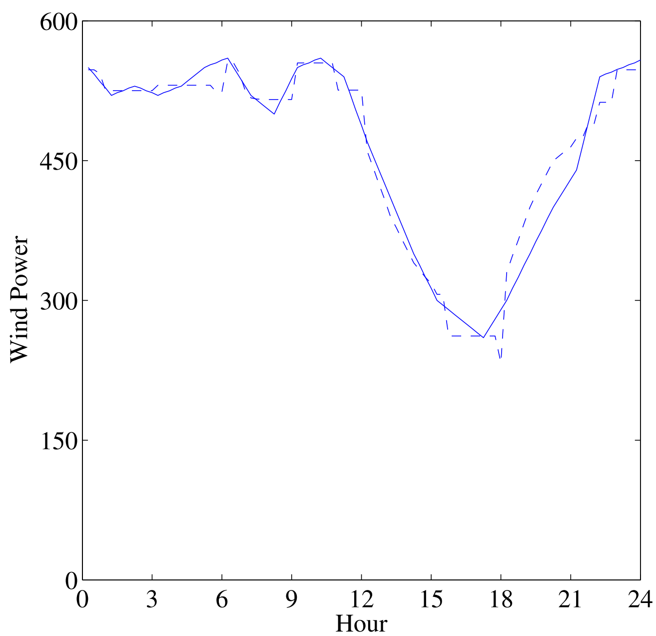

Fig. 4. Fall day: actual wind power, solid line, together with the forecasted wind power, dashed line, in megawatt.

TABLE I
STATISTICAL ANALYSIS OF THE DAILY FORECASTING ERROR

Day	MAPE	$\sqrt{\text{SSE}}$	SDE
Winter	9.51%	593.71	34.78
Spring	9.92%	578.13	42.49
Summer	6.34%	232.56	17.11
Fall	3.26%	207.10	14.85

Hence, the proposed ANN approach provides a powerful tool of easy implementation for forecasting wind power, enhancing forecasting accuracy over the persistence approach.

V. CONCLUSIONS

An ANN approach is proposed for short-term wind power forecasting in Portugal. The application of the proposed approach to wind power forecasting has been proven to be successful. The MAPE has an average value of 7.26%, while the average computation time is less than 5 seconds. Hence, the proposed approach presents a good trade-off between forecasting accuracy and computation time, outperforming the persistence approach.

VI. ACKNOWLEDGMENT

The authors gratefully acknowledge the contribution of Mr. N. M. S. Martins.

VII. REFERENCES

[1] L. A. F. M. Ferreira, "Evaluation of short-term wind predictability," *IEEE Trans. Energy Convers.*, vol. 7, pp. 409-417, Sep. 1992.

[2] A. Estanqueiro, R. Castro, P. Flores, J. Ricardo, M. Pinto, R. Rodrigues, and J. Peças Lopes, "How to prepare a power system for 15% wind energy penetration: the Portuguese case study," *Wind Energy*, vol. 11, pp. 75-84, Jan.-Feb. 2008.

[3] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsen, and E. Feitosa, "A review on the young history of the wind power short-term prediction," *Renew. Sust. Energy Rev.*, vol. 12, pp. 1725-1744, Aug. 2008.

[4] L. Ma, S. Y. Luan, C. W. Jiang, H. L. Liu, and Y. Zhang, "A review on the forecasting of wind speed and generated power," *Renew. Sust. Energy Rev.*, vol. 13, pp. 915-920, May 2009.

[5] R. G. Kavasseri and K. Seetharaman, "Day-ahead wind speed forecasting using f-ARIMA models," *Renew. Energy*, vol. 34, pp. 1388-1393, May 2009.

[6] I. J. Ramirez-Rosado, L. A. Fernandez-Jimenez, C. Monteiro, J. Sousa, and R. Bessa, "Comparison of two new short-term wind-power forecasting systems," *Renew. Energy*, vol. 34, pp. 1848-1854, Jul. 2009.

[7] G. Sideratos and N. D. Hatziargyriou, "An advanced statistical method for wind power forecasting," *IEEE Trans. Power Syst.*, vol. 22, pp. 258-265, Feb. 2007.

[8] C. W. Potter and W. Negnevitsky, "Very short-term wind forecasting for Tasmanian power generation," *IEEE Trans. Power Syst.*, vol. 21, pp. 965-972, May 2006.

[9] R. Jursa, K. Rohrig, "Short-term wind power forecasting using evolutionary algorithms for the automated specification of artificial intelligence models," *Int. J. Forecast.*, vol. 24, pp. 694-709, Oct.-Dec. 2008.

[10] T. H. M. El-Fouly, E. F. El-Saadany, and M. M. A. Salama, "Improved Grey predictor rolling models for wind power prediction," *IET Gener. Transm. Distrib.*, vol. 1, pp. 928-937, Nov. 2007.

[11] R. Bessa, V. Miranda, and J. Gama, "Improvement in wind power forecasting based on information entropy-related concepts," presented at the 2008 IEEE Power & Energy Society General Meeting, Pittsburgh, Pennsylvania, Jul. 2008.

[12] J. P. S. Catalão, S. J. P. S. Mariano, V. M. F. Mendes, and L. A. F. M. Ferreira, "An artificial neural network approach for short-term electricity prices forecasting," in *Proc. 14th Int. Conf. on Intelligent System Applications to Power Systems*, pp. 411-416, Nov. 2007.

[13] S. Haykin, *Neural Networks: A Comprehensive Foundation*, New Jersey: Prentice-Hall, 1999.

[14] J. C. Principe, N. R. Euliano, and W. C. Lefebvre, *Neural and Adaptive Systems: Fundamentals Through Simulations*, New York: Wiley, 2000.

[15] B. R. Szkuta, L. A. Sanabria, and T. S. Dillon, "Electricity price short-term forecasting using artificial neural networks," *IEEE Trans. Power Syst.*, vol. 14, pp. 851-857, Feb. 1999.

[16] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Trans. Neural Netw.*, vol. 5, pp. 989-993, Nov. 1994.

[17] J. P. S. Catalão, S. J. P. S. Mariano, V. M. F. Mendes, and L. A. F. M. Ferreira, "Short-term electricity prices forecasting in a competitive market: A neural network approach," *Elect. Power Syst. Res.*, vol. 77, pp. 1297-1304, Aug. 2007.

VIII. Biographies

J. P. S. Catalão (M'04) received the M.Sc. degree from the Instituto Superior Técnico, Lisbon, Portugal, in 2003 and the Ph.D. degree from the University of Beira Interior, Covilha, Portugal, in 2007.

He is currently an Assistant Professor at the University of Beira Interior. He is the author or co-author of more than 70 scientific papers presented at international conferences or published in reviewed journals.

Dr. Catalão is an Associate Editor for the *International Journal of Power and Energy Systems*, and a Member of the Editorial Board of *Electric Power Components & Systems*. Also, he is a reviewer for IEEE Transactions on Power Systems, and other IEEE and *International Journals*.

H. M. I. Pousinho (S'09) received the electromechanical engineering degree (with honors) from the University of Beira Interior, Covilha, Portugal, in 2008.

He is currently finishing his M.Sc. degree at the University of Beira Interior, in collaboration with the Instituto Superior de Engenharia de Lisboa, Lisbon, Portugal.

V. M. F. Mendes received the M.Sc. and Ph.D. degrees from the Instituto Superior Técnico, Lisbon, Portugal, in 1987 and 1994, respectively.

He is currently a Coordinator Professor with Aggregation at the Instituto Superior de Engenharia de Lisboa, Lisbon, Portugal. He is the author or co-author of more than 110 scientific papers presented at international conferences or published in reviewed journals.