Short-term Wind Power Forecasting in Portugal by Neural Networks and Wavelet Transform

J.P.S. Catalão [a,b,]*, H.M.I. Pousinho [a], V.M.F. Mendes [c]

[a] *Department of Electromechanical Engineering, University of Beira Interior, R. Fonte do Lameiro, 6201-001 Covilha, Portugal*
[b] *Center for Innovation in Electrical and Energy Engineering, Instituto Superior Técnico, Technical University of Lisbon, Av. Rovisco Pais, 1049-001 Lisbon, Portugal*
[c] *Department of Electrical Engineering and Automation, Instituto Superior de Engenharia de Lisboa, R. Conselheiro Emídio Navarro, 1950-062 Lisbon, Portugal*

Received 27 February 2010; received in revised form 24 June 2010

Abstract

This paper proposes artificial neural networks in combination with wavelet transform 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. Results from a real-world case study are presented. A comparison is carried out, taking into account the results obtained with other approaches. Finally, conclusions are duly drawn.

Keywords: Wind power; Forecasting; Artificial neural networks; Wavelet transform

---

1. Introduction

Wind energy is gaining increasing importance throughout the world [1], and wind-driven power resources have become increasingly important in the planning and operation of electric power systems. 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 in 2010. The wind power generating capacity reached 3350 MW on November 2009 and continues growing [2]. Particularly, on 15 November at 7h30m, the contribution of wind power was the highest ever, reaching 71% of the total load [3].

Wind as the energy source has an intermittent nature. Integration of wind power into an electrical grid requires an estimate of the expected power from the wind farms at least one to two days in advance [4].

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) [5].

* Corresponding author. Tel.: +351 275 329914; fax: +351 275 329972.
E-mail address: catalao@ubi.pt (J.P.S. Catalão).

In the technical literature, several methods to predict wind power have been reported, namely physical [6] and statistical methods [7]. Physical method has advantages in long-term prediction while statistical method does well in short-term prediction [8].

Conventional statistical models are identical to the direct random time-series model, including auto regressive (AR), and auto regressive integrated moving average (ARIMA) [9] models. The persistence models are considered as the simplest time-series models, but 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 [8]. The persistence approach has proven to be a useful first approximation for short-term wind power forecasting. Hence, persistence and ARIMA approaches provide an important benchmark against which to compare alternative techniques.

In the recent years, some new methods are catching researcher's attention, namely data mining [10], artificial neural networks (NN) [11–14], fuzzy logic [15,16], evolutionary algorithms [17], and some hybrid methods [18,19]. 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, it has been reported that artificial-based models outperformed others in short-term prediction [8].

NN 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, NN 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 NN are data-driven. Three-layered feedforward NN are specially suited for forecasting, implementing nonlinearities using sigmoid functions for the hidden layer and linear functions for the output layer [20]. Hence, a three-layered feedforward NN trained by the Levenberg-Marquardt algorithm is considered in this paper.

This paper presents a successful application of using NN in combination with wavelet transform (WT) to forecast short-term wind power in Portugal. The proposed NNWT approach is compared with persistence, ARIMA and NN approaches, to demonstrate its effectiveness regarding forecasting accuracy and computation time.

This paper is organized as follows. Section II presents the proposed NNWT approach to forecast wind power. Section III provides the different criterions used to evaluate the forecasting accuracy. Section IV presents the numerical results from a real-world case study. Finally, Section V outlines the conclusions.

2. Proposed approach

The proposed NNWT approach to forecast short-term wind power is based on a combination of NN with WT. The WT is used to decompose the wind power series into a set of better-behaved constitutive series. Then, the future values of these constitutive series are forecasted using NN. In turn, the NN forecasts allow, through the inverse WT, reconstructing the future behavior of the wind power series and therefore to forecast wind power.

2.1. Wavelet transform

The WT convert a wind power series into a set of constitutive series. These constitutive series present a better behavior than the original wind power series, and therefore, they can be predicted more accurately. The reason for the better behavior of the constitutive series is the filtering effect of the WT.

A brief summary of WT is presented hereafter. For the sake of simplicity, one-dimensional wavelets are considered to illustrate the related concepts. Wavelet analysis is capable of revealing aspects of data that other signal analysis techniques miss, such as trends, breakdown points, discontinuities in higher derivatives and self-similarity. Furthermore, wavelet analysis can often compress or de-noise a signal without appreciable degradation [21]. These capabilities of WT can be useful in short-term wind power forecasting.

WTs can be divided in two categories: continuous wavelet transform (CWT) and discrete wavelet transform (DWT). The CWT $W(a,b)$ of signal $f(x)$ with respect to a mother wavelet $\phi(x)$ is given by [21]:

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x)\, \phi\left(\frac{x-b}{a}\right) dx \tag{1}$$

where the scale parameter a controls the spread of the wavelet and translation parameter b determines its central position. The $W(a,b)$ coefficient represents how well the original signal $f(x)$ and the scaled/translated mother wavelet match. Thus, the set of all wavelet coefficients $W(a,b)$, associated to a particular signal, is the wavelet representation of the signal with respect to the mother wavelet.

Since the CWT is achieved by continuously scaling and translating the mother wavelet, substantial redundant information is generated. Therefore, instead of doing that, the mother wavelet can be scaled and translated using certain scales and positions usually based on powers of two. This scheme is more efficient and just as accurate as the CWT [22]. It is known as the DWT and defined as:

$$W(m,n) = 2^{-(m/2)} \sum_{t=0}^{T-1} f(t)\, \phi\left(\frac{t - n.2^m}{2^m}\right) \tag{2}$$

where T is the length of the signal $f(t)$. The scaling and translation parameters are functions of the integer variables m and n ($a = 2^m, b = n.2^m$); t is the discrete time index.

A fast DWT algorithm based on the four filters (decomposition low-pass, decomposition high-pass, reconstruction low-pass, and reconstruction high-pass filters), developed by Mallat [23], is considered in this paper.

Multiresolution via Mallat's algorithm is a procedure to obtain ''approximations'' and ''details'' from a given signal. An approximation holds the general trend of the original signal, whereas a detail depicts high-frequency components of it [22]. By successive decomposition of the approximations (Fig. 1), a multilevel decomposition process can be achieved where the original signal is broken down into lower resolution components.

"See Fig. 1 at the end of the manuscript".

A wavelet function of type Daubechies of order 4 (abbreviated as Db4) is used as the mother wavelet $\phi(t)$. The peculiarity of this wavelet system is that there is no explicit function, so we cannot draw it directly. The order of the Daubechies functions denotes the number of zero moments of the wavelet function. The wavelets are built based in a small function $\tau(t)$, given by:

$$\tau(t) = \sqrt{2} \sum_{n}^{N} l_n \tau(2t - n) \tag{3}$$

where $\tau(t)$ is the scale function or scaling wavelet and l_k denotes the low-pass filter coefficients which determine the characteristics of the resulting wavelet transform. The mother wavelet $\phi(t)$ is given by:

$$\phi(t) = \sqrt{2} \sum_{n}^{N} h_n \phi(2t - n) \tag{4}$$

where h_k denotes the high-pass filter coefficients closely related to the low-pass filter (l_k) mentioned above.

The scale function is used to capture the smooth, low frequency nature of the data, whereas the mother wavelets are used to capture the detailed and high frequency nature of the data. The coefficients of low-pass and high-pass filters are related by [24]:

$$h_n = (-1)^n l_{1-n} \tag{5}$$

The Daubechies wavelet Db4 offers an appropriate trade-off between wave-length and smoothness, resulting in an appropriate behavior for short-term wind power forecasting. Similar wavelets have been considered by previous researchers for load forecasting [21,22] and price forecasting [25,26]. Also, three decomposition levels are considered, as in [26], since it describes the wind power series in a meaningful way. The approach developed in this paper uses A3, D3 and D1 as inputs for the NN.

2.2. Artificial neural networks

NN are highly interconnected simple processing units designed in a way to model how the human brain performs a particular task [27]. Each of those units, also called artificial 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 NN. The artificial neurons are organized in a way that defines the network architecture. In feedforward networks, artificial neurons are often arranged in layers: an input layer, one or more hidden layers and an output layer [27].

Fig. 2 shows the architecture of a generic three-layered feedforward NN model.

"See Fig. 2 at the end of the manuscript".

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 artificial neurons 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 artificial neuron output layer with a pure linear transfer function.

The configuration chosen consists of a one hidden layer that uses a hyperbolic tangent sigmoid transfer function, given by:

$$f(s) = \frac{e^s - e^{-s}}{e^s + e^{-s}} \tag{6}$$

and a one unit output layer with a pure linear transfer function, given by:

$$f'(s') = s' \tag{7}$$

where s is the weighted input of the hidden layer, $f(s)$ is the hyperbolic tangent sigmoid transfer function of the hidden layer, s' is the weighted input of the output layer, and $f'(s')$ is the pure linear transfer function of the output layer.

This configuration has been proven to be a universal mapper, provided that the hidden layer has enough artificial neurons [28].

On one hand, if there are too few artificial neurons, the network will not be flexible enough to model the data well and, on the other hand, if there are too many artificial neurons, the network may over-fit the data. The number of artificial neurons in the hidden layer was chosen by trial and error. The best results were produced with six hidden artificial neurons. The number of model input parameters is four.

Forecasting with NN 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 and containing both inputs and the corresponding desired outputs, which is presented to the network. The adequate selection of inputs for NN training is highly influential to the success of training. In the learning process a NN 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. However, the standard backpropagation learning algorithm tends to converge slowly [29].

An algorithm that trains a NN 10 to 100 times faster than the standard 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 [14,20]. Hence, a three-layered feedforward NN trained by the Levenberg-Marquardt algorithm is considered in this paper.

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

$$\Delta(\boldsymbol{x}) = -\left[\nabla^2 V(\boldsymbol{x})\right]^{-1} \nabla V(\boldsymbol{x}) \tag{8}$$

where $\nabla^2 V(\boldsymbol{x})$ is the Hessian matrix and $\nabla V(\boldsymbol{x})$ is the gradient vector.

Neglecting the second-order derivatives of the error vector, the Hessian matrix is given by:

$$\nabla^2 V(\boldsymbol{x}) = 2\, \boldsymbol{J}^T(\boldsymbol{x})\, \boldsymbol{J}(\boldsymbol{x}) \tag{9}$$

where $\boldsymbol{J}(\boldsymbol{x})$ is the Jacobian matrix.

The Gauss-Newton update is given by:

$$\Delta(\boldsymbol{x}) = -\left[\boldsymbol{J}^T(\boldsymbol{x})\, \boldsymbol{J}(\boldsymbol{x})\right]^{-1} \boldsymbol{J}^T(\boldsymbol{x})\, \boldsymbol{e}(\boldsymbol{x}) \tag{10}$$

where $\boldsymbol{e}(\boldsymbol{x})$ is the error vector.

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 be not invertible. This is overcome with the Levenberg-Marquardt algorithm, which consists in finding the update given by:

$$\Delta(\boldsymbol{x}) = -\left[\boldsymbol{J}^T(\boldsymbol{x})\, \boldsymbol{J}(\boldsymbol{x}) + \mu\, \boldsymbol{I}\right]^{-1} \boldsymbol{J}^T(\boldsymbol{x})\, \boldsymbol{e}(\boldsymbol{x}) \tag{11}$$

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 (11) 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 [29].

3. Forecasting accuracy evaluation

To evaluate the accuracy of the proposed NNWT approach in forecasting wind power, different criterions are used. This accuracy is computed in function of the actual wind power that occurred. These criterion allow comparing alternative techniques.

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 defined as follows [30]:

$$MAPE = \frac{100}{N} \sum_{h=1}^{N} \frac{\left|\hat{p}_h - p_h\right|}{\overline{p}} \tag{12}$$

$$\overline{p} = \frac{1}{N} \sum_{h=1}^{N} p_h \tag{13}$$

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:

$$SSE = \sum_{h=1}^{N} (\hat{p}_h - p_h)^2 \tag{14}$$

The SDE criterion is given by:

$$SDE = \sqrt{\frac{1}{N} \sum_{h=1}^{N} (e_h - \bar{e})^2} \tag{15}$$

$$e_h = \hat{p}_h - p_h \tag{16}$$

$$\bar{e} = \frac{1}{N} \sum_{h=1}^{N} e_h \tag{17}$$

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

A measure of the uncertainty of a model is the variability of what is still unexplained after fitting the model, which can be measured through the estimation of the variance of the error. The smaller this variance, the more precise is the prediction [25].

Consistent with definition (12), daily error variance can be estimated as [25]:

$$\sigma_{e,day}^2 = \frac{1}{24} \sum_{h=1}^{24} \left(\frac{|\hat{p}_h - p_h|}{\bar{p}} - (e_{day}) \right)^2 \tag{18}$$

$$e_{day} = \frac{1}{24} \sum_{h=1}^{24} \frac{|\hat{p}_h - p_h|}{\bar{p}} \tag{19}$$

4. Numerical results

The proposed NNWT approach has been applied for wind power forecasting in Portugal. The numerical results presented take into account all the wind farms in Portugal that have telemetry with the National Electric Grid (REN). Historical wind power data, available at the REN website (thus, public-domain information), are the main inputs to train the NN. For the sake of clear comparison, no exogenous variables (such as temperature or pressure) are considered.

Our forecaster (ANN architecture) predicts the value of the wind power subseries for 3 hours ahead, taking into account the wind power data of the previous 12 hours with a time-step of 15 minutes (therefore existing 48 measured power values, divided into four column vectors with 12 elements each). This procedure is repeated until the next 24 hours values are predicted.

It should be noted that the input layer is comprised by four artificial neurons, which correspond to four column vectors with 12 elements each, while the output layer is comprised by only one artificial neuron, which corresponds to a single column vector with 12 elements (3 hours, with a time-step of 15 minutes).

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 NNWT approach are shown in Figs. 3 to 6 respectively for the winter, spring, summer and fall days. Each figure shows the actual wind power, solid line, together with the forecasted wind power, dash-dot line.

"See Fig. 3 at the end of the manuscript".

"See Fig. 4 at the end of the manuscript".

"See Fig. 5 at the end of the manuscript".

"See Fig. 6 at the end of the manuscript".

Table 1 presents the values for the criterions to evaluate the accuracy of the proposed NNWT 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.

"See Table 1 at the end of the manuscript".

A good accuracy of the proposed NNWT approach was ascertained. The MAPE has an average value of 6.97%.

Table 2 shows a comparison between the proposed NNWT approach and three other approaches (persistence, ARIMA, and NN), regarding the MAPE criterion. The persistence approach assumes that the predicted value of the next step in the future is the last measured value. The ARIMA approach is developed using SPSS software. Parameter estimation is performed with the aid of this software. The configuration considered corresponds to an ARIMA (1,2,1).

"See Table 2 at the end of the manuscript".

The proposed NNWT approach provides the lowest average MAPE. The relative (negative or positive) values of forecast errors, considering ARIMA, NN and NNWT approaches, are shown in Figs. 7 to 10 respectively for the winter, spring, summer and fall days.

"See Fig. 7 at the end of the manuscript".

"See Fig. 8 at the end of the manuscript".

"See Fig. 9 at the end of the manuscript".

"See Fig. 10 at the end of the manuscript".

The ARIMA approach provides larger errors compared with NN and NNWT approaches. Comparing NN with the proposed NNWT approach, it is possible to conclude that the introduction of WT enables a reduction in the average MAPE.

In addition to the MAPE, stability of results is another important factor for the comparison of forecast approaches. Table 3 shows a comparison between the proposed NNWT approach and the three other approaches (persistence, ARIMA and NN), regarding daily error variances.

"See Table 3 at the end of the manuscript".

The average error variance is smaller for the proposed NNWT approach, indicating less uncertainty in the predictions. Improvement in the average error variance of the proposed approach with respect to the three other approaches is 79.8%, 41.3% and 7.8%, respectively.

Furthermore, the four plots of Fig. 11 provide average errors considering ARIMA, NN and NNWT approaches, for the four days analyzed.

"See Fig. 11 at the end of the manuscript".

Overall, the performance of the proposed NNWT approach is generally better than the performance of ARIMA and NN approaches, even if some other four representative days were used.

Moreover, the average computation time required by the proposed NNWT approach is less than 10 seconds using MATLAB, which is similar to the average computation time required by the NN approach but with the added advantage of the filtering effect. Instead, the ARIMA approach requires about 1 minute of computation time.

Hence, the proposed NNWT approach provides a powerful tool of easy implementation for short-term wind power forecasting.

5. Conclusions

A NNWT approach, based on the combination of artificial neural networks with wavelet transform, is proposed for short-term wind power forecasting in Portugal. The application of the proposed NNWT approach to wind power forecasting in Portugal is both novel and effective. The MAPE has an average value of 6.97%, outperforming persistence, ARIMA and NN approaches, while the average computation time is less than 10 seconds. Hence, the results presented confirm the considerable value of the proposed NNWT approach in forecasting wind power.

Acknowledgements

H.M.I. Pousinho thanks the Fundação para a Ciência e a Tecnologia (FCT) for a Ph.D. grant (SFRH/BD/62965/2009).

References

[1] Ackermann T. (ed.). Wind Power in Power Systems. John Wiley and Sons Ltd.: West Sussex, 2005.

[2] Melício R, Mendes VMF, Catalão JPS. Power converter topologies for wind energy conversion systems: Integrated modeling, control strategy and performance simulation. Renew Energy 2010;35(10):2165–74.

[3] REN – National Electric Grid. Monthly Report (in Portuguese), November 2009.

[4] Lazic L, Pejanovic G, Zivkovic M. Wind forecasts for wind power generation using the Eta model. Renew Energy 2010;35(6):1236–43.

[5] Costa A, Crespo A, Navarro J, Lizcano G, Madsen H, Feitosa E. A review on the young history of the wind power short-term prediction. Renew Sust Energy Rev 2008;12(6):1725–44.

[6] Lange M, Focken U, Physical Approach to Short-Term Wind Power Prediction. Springer-Verlag: Berlin, Heidelberg, 2006.

[7] Liu H, Tian H-Q, Chen C, Li Y-F. A hybrid statistical method to predict wind speed and wind power. Renew Energy 2010;35(8):1857–61.

[8] Ma L, Luan SY, Jiang CW, Liu HL, Zhang Y. A review on the forecasting of wind speed and generated power. Renew Sust Energy Rev 2009;13(4):915–20.

[9] Kavasseri RG, Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. Renew Energy 2009;34(5):1388–93.

[10] Kusiak A, Li W. Short-term prediction of wind power with a clustering approach. Renew Energy 2010;35(10): 2362–9.

[11] Mabel MC, Fernandez E. Analysis of wind power generation and prediction using ANN: A case study. Renew Energy 2008;33(5):986–92.

[12] Cadenas E, Rivera W. Short term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks. Renew Energy 2009;34(1):274–8.

[13] Ramirez-Rosado IJ, Fernandez-Jimenez LA, Monteiro C, Sousa J, Bessa R. Comparison of two new short-term wind-power forecasting systems. Renew Energy 2009;34(7):1848–54.

[14] Catalão JPS, Pousinho HMI, Mendes VMF. An artificial neural network approach for short-term wind power forecasting in Portugal. Eng Intell Syst Elect Eng Commun 2009;17(1):5–11.

[15] Sideratos G, Hatziargyriou ND. An advanced statistical method for wind power forecasting. IEEE Trans Power Syst 2007; 22: 258–65.

[16] Potter CW, Negnevitsky W. Very short-term wind forecasting for Tasmanian power generation. IEEE Trans Power Syst 2006; 21(2): 965–72.

[17] Jursa R, Rohrig K. Short-term wind power forecasting using evolutionary algorithms for the automated specification of artificial intelligence models. Int J Forecast 2008;24(4):694–709.

[18] El-Fouly THM, El-Saadany EF, Salama MMA. Improved grey predictor rolling models for wind power prediction. IET Gener Transm Distrib 2007; 1(6): 928–37.

[19] Bessa RJ, Miranda V, Gama J. Entropy and correntropy against minimum square error in offline and online three-day ahead wind power forecasting. IEEE Trans Power Syst 2009; 24(4): 1657–66.

[20] Catalão JPS, Mariano SJPS, Mendes VMF, Ferreira LAFM. Short-term electricity prices forecasting in a competitive market: a neural network approach. Electr Power Syst Res 2007;77(10):1297–304.

[21] Amjady N, Keynia F. Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm. Energy 2009; 34(1): 46–57.

[22] Reis AJR, da Silva APA. Feature extraction via multiresolution analysis for short-term load forecasting. IEEE Trans Power Syst 2005; 20(1): 189–98.

[23] Mallat S. A theory for multiresolution signal decomposition—the wavelet representation. IEEE Trans Pattern Anal Mach Intell 1989; 11(7): 674–93.

[24] Uyar M, Yildirim S, Gencoglu MT. An effective wavelet-based feature extraction method for classification of power quality disturbance signals. Electr Power Syst Res 2008; 78: 1747–55.

[25] Conejo AJ, Plazas MA, Espínola R, Molina AB. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. IEEE Trans Power Syst 2005;20(2):1035–42.

[26] Amjady N, Keynia F. Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method. Int J Electr Power Energy Syst 2008; 30(9): 533–46.

[27] Haykin S. (ed.). Neural Networks: A Comprehensive Foundation. Prentice-Hall: New Jersey, 1999.

[28] Principe JC, Euliano NR, Lefebvre WC. (eds.). Neural and Adaptive Systems: Fundamentals Through Simulations. Wiley: New York, 2000.

[29] Hagan MT, Menhaj MB. Training feedforward networks with the Marquardt algorithm. IEEE Trans Neural Netw 1994; 5: 989–93.

[30] Shahidehpour M, Yamin H, Li Z. (eds.). Market Operations in Electric Power Systems: Forecasting, Scheduling and Risk Management. Wiley: New York, 2002.

Figure captions

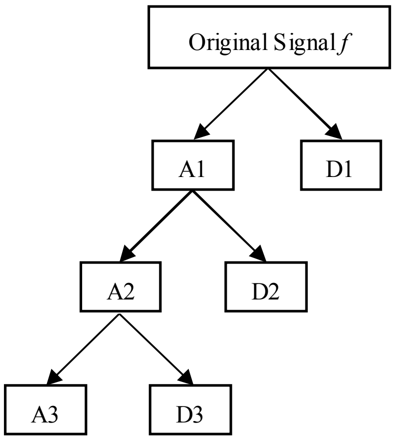

Fig. 1. Multilevel decomposition process.

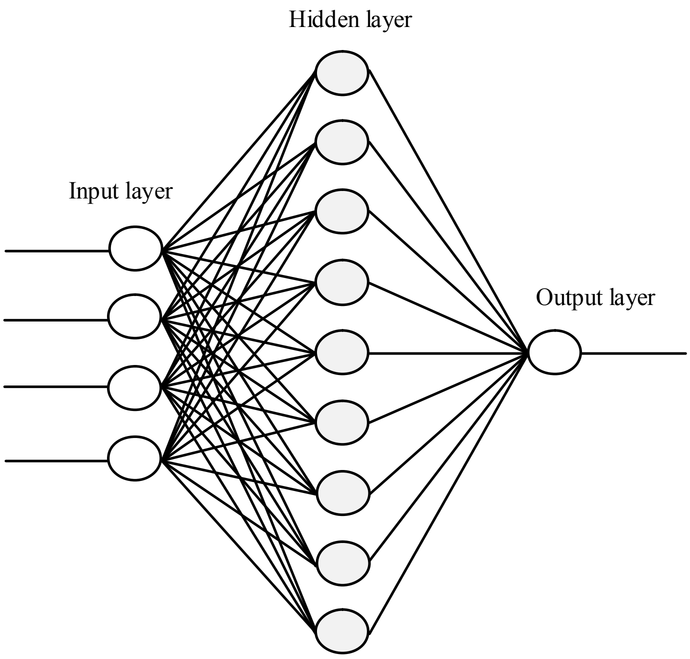

Fig. 2. Three-layered feedforward NN model.

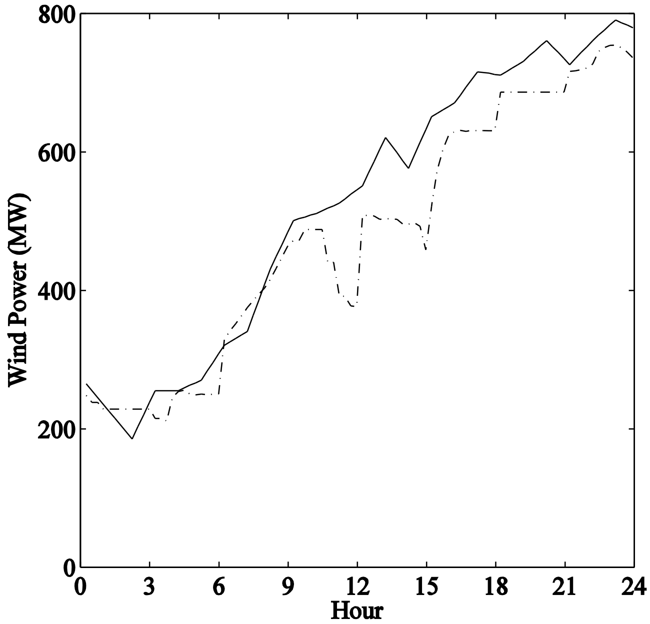

Fig. 3. Winter day: actual wind power (solid line) together with the forecasted wind power (dash-dot line).

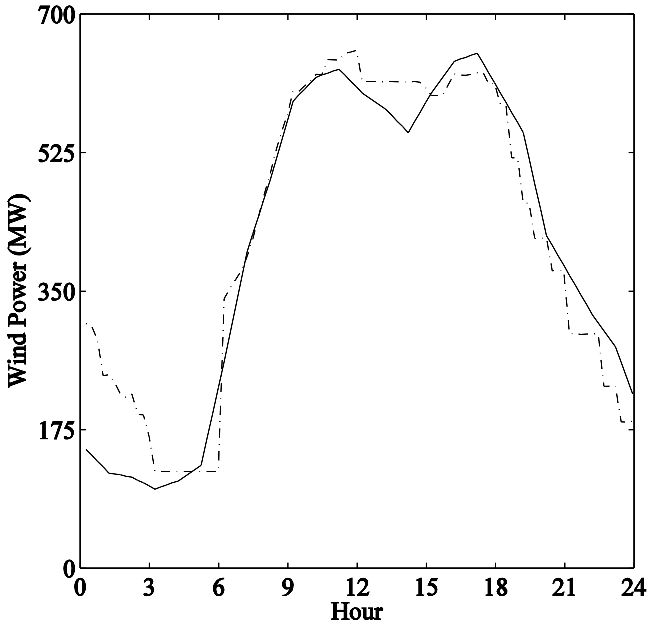

Fig. 4. Spring day: actual wind power (solid line) together with the forecasted wind power (dash-dot line).

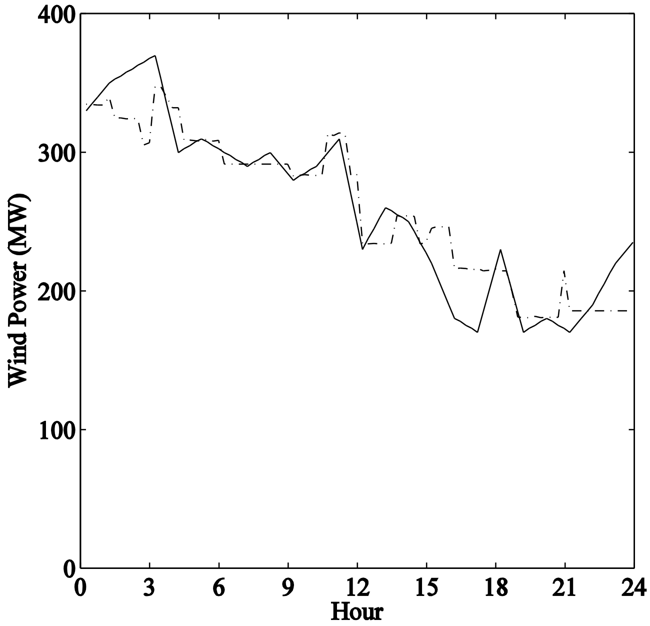

Fig. 5. Summer day: actual wind power (solid line) together with the forecasted wind power (dash-dot line).

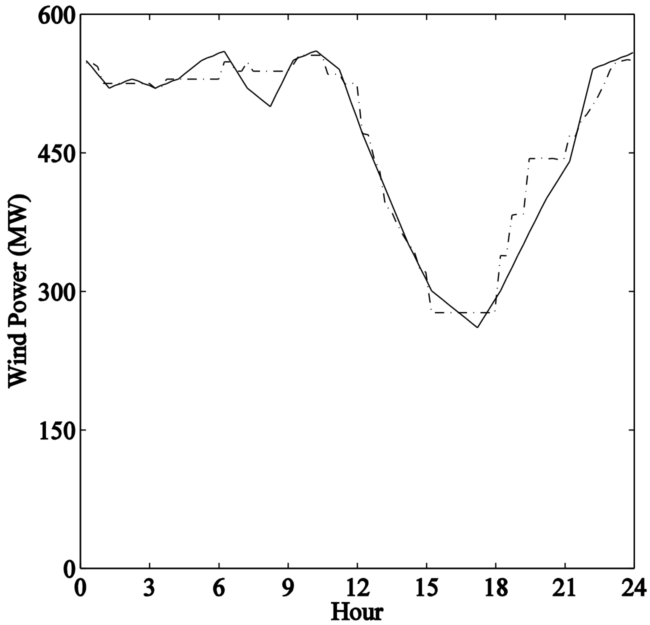

Fig. 6. Fall day: actual wind power (solid line) together with the forecasted wind power (dash-dot line).

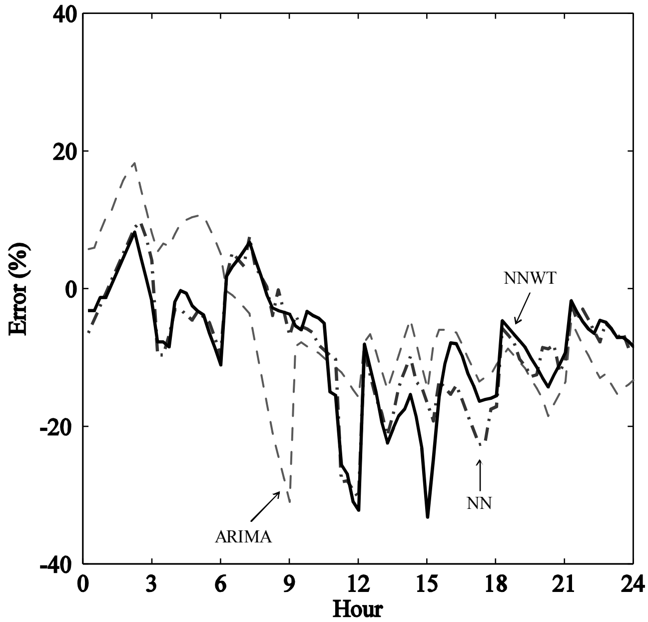

Fig. 7. Winter day: relative value of forecast errors considering ARIMA (dashed line), NN (dash-dot line) and NNWT (solid line) approaches.

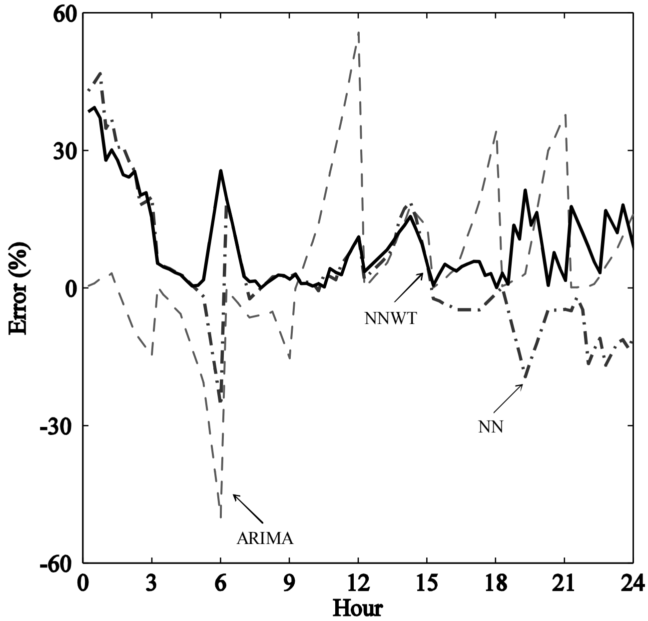

Fig. 8. Spring day: relative value of forecast errors considering ARIMA (dashed line), NN (dash-dot line) and NNWT (solid line) approaches.

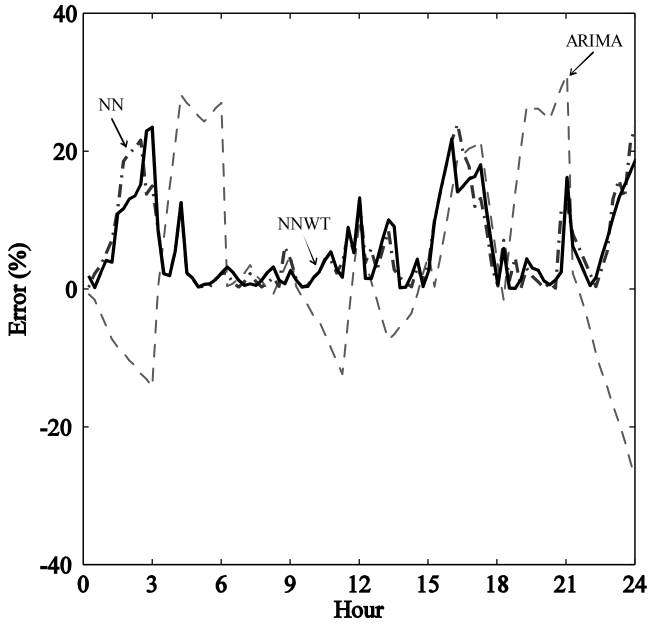

Fig. 9. Summer day: relative value of forecast errors considering ARIMA (dashed line), NN (dash-dot line) and NNWT (solid line) approaches.

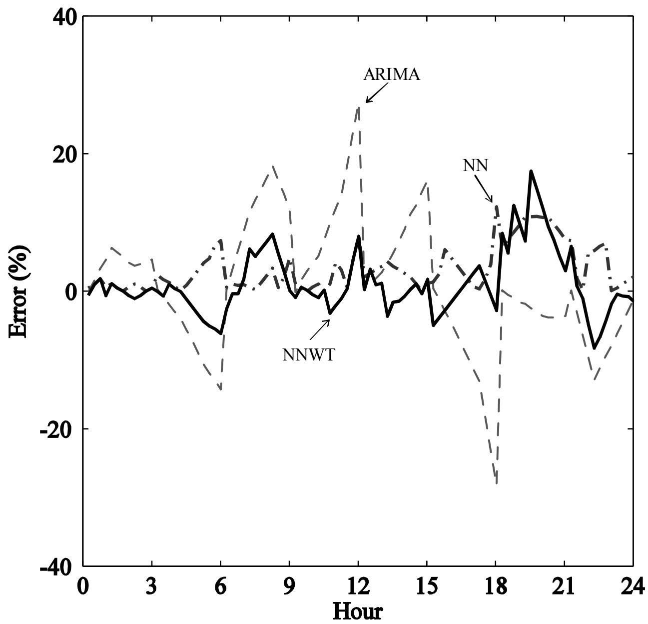

Fig. 10. Fall day: relative value of forecast errors considering ARIMA (dashed line), NN (dash-dot line) and NNWT (solid line) approaches.

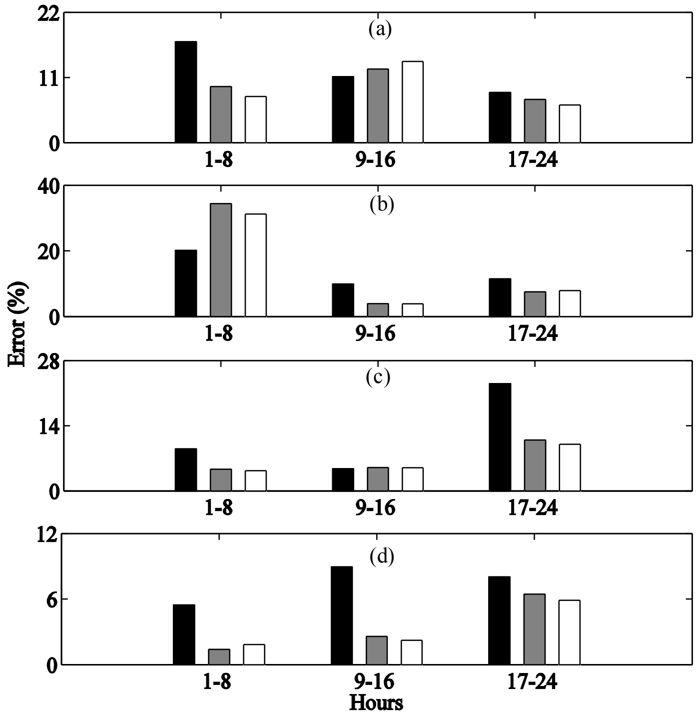

Fig.11. Average errors within three time intervals, considering ARIMA (black rectangle), NN (grey rectangle) and NNWT (white rectangle) approaches for the days analyzed: (a) Winter, (b) Spring, (c) Summer, and (d) Fall.

Tables

Table 1

Statistical analysis of the daily forecasting error

Day	MAPE (%)	$\sqrt{\text{SSE}}$ (MW)	SDE (MW)
Winter	9.23	606.15	38.80
Spring	9.55	533.62	37.61
Summer	5.97	218.11	15.99
Fall	3.14	211.38	15.97

Table 2

Comparative MAPE results

	Winter	Spring	Summer	Fall	Average
Persistence	13.89	32.40	13.43	16.49	19.05
ARIMA	10.93	12.05	11.04	7.35	10.34
NN	9.51	9.92	6.34	3.26	7.26
NNWT	9.23	9.55	5.97	3.14	6.97

Table 3

Daily forecasting error variance

	Winter	Spring	Summer	Fall	Average
Persistence	0.0074	0.0592	0.0085	0.0179	0.0233
ARIMA	0.0025	0.0164	0.0090	0.0039	0.0080
NN	0.0044	0.0106	0.0043	0.0010	0.0051
NNWT	0.0055	0.0083	0.0038	0.0012	0.0047