Proceedings of ASME Turbo Expo 2011
GT2011
June 6-10, 2011, Vancouver, British Columbia, Canada

GT2011-46382

A NEW HYBRID METHOD FOR WIND POWER FORECASTING BASED ON WAVELET DECOMPOSITION AND ARTIFICIAL NEURAL NETWORKS

Maria Grazia De Giorgi*
University of Salento
Dep. Engineering for Innovation
Lecce, Italy

Marco Tarantino
University of Salento
Dep. Engineering for Innovation
Lecce, Italy

Antonio Ficarella
University of Salento
Dep. Engineering for Innovation
Lecce, Italy

ABSTRACT

Depending on their input, wind power forecasting models are classified as physical or statistical approaches or a combination of both.

Physical models use physical considerations, as meteorological information (Numerical Weather Prediction) and technical characteristics of the wind turbines (hub height, power curve, thrust coefficient). Statistical models use explanatory variables and online measurements, usually employing recursive techniques, like recursive least squares or artificial neural networks (ANNs) which perform a non-linear mapping and provide a robust approach for wind prediction.

In this paper a new hybrid method (mixing physical and statistical approaches) is proposed, based on the wavelet decomposition technique and on artificial neural networks, in order to predict power production of a wind farm in different time horizons: 1, 3, 6, 12 and 24 hours.

In particular, two approaches are compared, both based on the time series of on-line measured wind power and on the Numerical Weather Predictions; in the first approach, the forecast is carried out only through the training of a neural network which, in the second approach is, instead, used in combination with the wavelet decomposition technique, improving the performance especially over the short time horizons.

The error of the different forecast systems is investigated for various forecasting horizons and statistical distributions of the error are calculated and presented.

1. INTRODUCTION

Among the new sources of renewable energy, wind energy is undoubtedly the one which in the last years has got the best growth, becoming in various countries the true alternative to fossil fuels. At the end of 2009, worldwide nameplate capacity of wind-powered generators was 159.2 gigawatts (GW) with an energy production of 340 TWh, about 2% of worldwide electricity usage (against the 0.1% of 1997).

The increasing interest of the worldwide literature in the wind energy field is attested by several works which deal with this very attractive theme.

Morales et al.[1] propose a procedure to produce a set of plausible scenarios characterizing the uncertainty associated with wind speed at different geographic sites. This characterization constitutes an essential part within the decision-making processes faced by both power system operators and producers with a generation portfolio including wind plants at several locations. Zhou et al. [2] review the current state of the simulation, optimization and control technologies for the stand-alone hybrid solar–wind energy systems with battery storage. In [3] Xydis et al. perform a wind resource assessment study in the area of Central Peloponnese (inland) using Geographic Information Systems (GIS) tools and an exergy analysis. Hongxing [4] recommends an optimal design model for designing hybrid solar–wind systems employing battery banks for calculating the system optimum configurations and ensuring that the annualized cost of the systems is minimized while satisfying the custom required loss of power supply probability (LPSP). In [5] the hourly measured wind speed data for years 2003–2005 at 10 m, 30 m and 60 m height for Kingdom of Bahrain have been statically analyzed to determine the potential of wind power generation. Luickx et al.[6] define various elements that come into play when considering backup for electricity generation from wind power. The effects of several parameters, to be situated on the short-term operation of backup of wind power, are defined and

* Author of correspondence

analysed. Most important parameters are the load profiles, the wind power output profiles and the total amount of installed wind power.

The most important problem for the diffusion of wind energy is that it is characterized by a high variability, both in space and time. The short-term wind energy forecasts are very important to minimize the scheduling errors which impact grid reliability and market service costs. In a background as the one of energy, which is more and more oriented towards the economical concepts of Stock Exchange, the energy producers should be able to predict the amount of energy produced in the subsequent hours or days, with a good precision.

The natural consequence in the evolution of electric market has been the search of technical solutions, often based on historical series, able to predict the deliverable power in the short and middle period. The introduction, also in the energetic field, of the artificial neural networks, has given new impulse to use of systemic techniques, among the real time prediction systems. A requirement for a good model in a real time prediction system is the ability to keep acceptable reliability of the forecasts when the prediction length increases.

In fact, it's essentially the availability of an enough lead-time of forecasts which conditions the opportunity of the best use of the energetic source.

Because of the variability of wind, the skill of forecasting the wind speed in subsequent time intervals, must be evaluated in the various cases.

Different forecasting models have been developed in literature, each model has its own characteristics, and it can perform well in different situations.

Numerical weather prediction (NWP) models are good at predicting large-scale area wind speed and can achieve better results in long-term forecasting [7,8]. They use hydrodynamic atmospheric models which take into account physical phenomena such as frictional, thermal and convective effects.

Generally forecasts can be of two kinds, like Burton asserts [9]: short time predictions of wind speed on a time horizon variable from few seconds to minutes, which could be useful for the operative control of wind turbines, and long time forecasts on a time horizon from few hours to days. These last forecasts are useful in order to plan the energetic supply of wind parks. In [9] it's underlined that for short time forecasts the statistics techniques give good results, while for long time forecasts it's necessary to rely on meteorological methods.

The Auto Regressive Moving Average (ARMA) models are sophisticate methods which make a forecast based on a linear combination of n previous values. They are based upon the assumption that the value of wind speed at time k is a linear function of the two previous values at times (k-1) and (k-2) and that the coefficients of the linear function change every time. The initial assumption of this method is that the statistical properties of wind (average, auto-variance) don't change in the period taken into consideration for the prediction. This assumption can be limiting in the use of ARMA models.

Bossanyi [10] investigated the use of ARMA models for the forecast of wind speed from few seconds to few minutes obtaining a decrement of error up to 20% if compared with the statistical methods based on "persistence".

Instead, Boland et al. [11] applied ARMA models in order to predict the wind power with a prediction length of half an hour, also proposing a method for the selection of the optimal order of the model.

Very longer horizons were taken into account by Kavasseri and Seetharaman [12] who examined the use of fractional-Autoregressive Integrated Moving Average (ARIMA) models to forecast wind speeds on the day-ahead (24 h) and two-day-ahead (48 h) periods. They assessed forecast accuracy by computing three indexes, that's to say the daily mean error, the variance and the square root of the forecast mean square error, and their results indicate that significant improvements in forecasting accuracy are obtained with the proposed models compared to the persistence method.

Riahy and Abedi [13] proposed a new method, based on linear prediction, for wind speed forecasting. The method utilizes the 'linear prediction' method in conjunction with "filtering" of the wind speed waveform. This approach, oriented to minimize the absolute percentage error, is however applied only for a prediction time of few seconds ahead.

Recent techniques as Artificial Neural Networks (ANNs), neuro-fuzzy networks and wavelet-based methods are more and more used.

An application of ANNs for wind speed forecast with a time horizon of 1 hour is the one of Flores [14] on the base of the 3 previous values, obtaining a Mean Squared Error of 0.057, while a new approach to wind speed forecast in the next hour has been made by A. Sfetsos [15], whose method is based on a multi-step prediction of average values in 10 minutes intervals. Among the middle-long time forecasts, it can be underlined the work of S. Jayaraj [16], who applied an ANN in order to predict the wind speed in the subsequent hour, in the 24 subsequent hours and in the 48 subsequent hours: good performance have been obtained in forecasting the wind speed in the subsequent hour (with a maximum root mean squared error of 13% on the generated output), while the use of ANN in longer time intervals shows an increment of root mean squared error up to 23% in the case of the 24 subsequent hours.

Cadenas[17] applied the ARIMA models and the ANNs to a time series conformed by 7 years of wind speed measurements, with good results through a seasonal parameter introduced in the ARIMA model, using a prediction length of 1 hour and evaluating the performances by the mean squared error, the mean absolute error and the mean absolute percentage error. Potter et al. [18] and Johnson et. al [19] obtained an improvement of forecast quality for the brief prediction lengths by applying the Adaptive Neuro-Fuzzy Inference Systems (ANFIS), but a very short prediction length is tested: 2.5 minutes in [18] and 5 minutes in [19].

Barbounis and Theocharis [20] made a long-term wind speed and power forecasting in a wind farm using locally recurrent multilayer networks as forecast models with a prediction length from 15 minutes to 3 hours; they evaluated the performance by the normalized mean squared error, revealing an improvement in respect of the persistent and the ARMA models.

Bilgili et al. [21] applied resilient propagation (RP) artificial neural networks to predict the mean monthly wind speed of any target station using the mean monthly wind speeds of neighbouring stations which are indicated as reference stations. The maximum and the minimum mean absolute percentage errors were found to be, respectively, 14.13% and 4.49%. Another attempt to forecast the mean monthly wind speed by ANNs is the one of Kalogirou et al. [22], who used a multilayered artificial neural network for predicting the mean monthly wind speed in regions of Cyprus. Data for the period 1986-1996 have been used to train the neural network, whereas data for the year 1997 were used for validation; in this case the maximum percentage difference was of only 1.8%.

In [23] the profile of wind speed in Nigeria is modelled using artificial neural network. The ANN model consists of 3-layered, feed-forward, back-propagation network with different configurations, designed using the Neural Toolbox for MATLAB. The geographical parameters (latitude, longitude and altitude) and the month of the year were used as input data, while the monthly mean wind speed was used as the output of the network.

Beccali et al.[24] present a novel approach to wind speed spatial estimation on the isle of Sicily (Italy): an incremental self-organizing neural network (Generalized Mapping Regressor – GMR) is coupled with exploratory data analysis techniques in order to obtain a map of the spatial distribution of the average wind speed over the entire region.

Li et al.[25] illustrate a comprehensive comparison study on the application of different artificial neural networks in 1-h-ahead wind speed forecasting. Three types of typical neural networks, namely, adaptive linear element, back propagation, and radial basis function, are investigated.

In literature there are some interesting attempts to compare the new non linear forecast techniques with the linear ARMA models ([26],[27],[28]). According to Kariniotakis [26] the fuzzy-logic leads to improvements in the predictions of the wind power if compared with the simplest statistical techniques, but the forecast range that the author considered is however included between ten minutes and two hours. Palomares-Salas et al [27] based their comparison between ARIMA model and a back-propagation neural network on three parameters: the Pearson correlation coefficient associated with the original time-series and the forecasted series, the Index Of Agreement (IOA) of Willmot and the Root Mean Square Error. Their results show that ARIMA model is better than ANNs for short time-intervals to forecast (10 minutes, 1 hour, 2 hours and 4 hours).

Sfetsos [28] compares ARMA, ANNs and ANFIS techniques by evaluating the RMS error in the prediction length of 1 hour for the hourly wind speed.

The Multilayer Perceptron network (MLP) is the principal technique used in [29] behind other forecast methods like ARMA models and various kinds of ANNs. Two main forecast systems are presented, based not only on historical real data but also on numerical weather predictions, obtaining an average RMS error of about 14% in the horizon 12-24 hours.

Lei et al. [30] give a bibliographical survey on the general background of research and developments in the fields of wind speed and wind power forecasting. Based on the assessment of wind power forecasting models, further direction for additional research and application is proposed, such as deepen further study on artificial intelligence methods and improve their training algorithm aiming at more accurate results, combine different physical and statistical models to achieve good results both in long and short-term prediction, deepen further research on the practical application of the models, not only in theoretical, put forward new mathematical methods.

Another useful technique, proposed in the recent years for the short-term load forecasting of wind power systems, is wavelet decomposition. A certain regularity of the data is an important precondition for the successful application of ANNs [31]. When using classical statistical techniques, a stationary process is assumed for the data. For load time-series, an assumption of stationary series has to be discarded most of the time. Besides, one has to bear in mind that different kinds of non stationary series may exist [31]. In order to tackle the problem of non stationary series, wavelets have been utilized because they can produce a good local representation of the signal in both time and frequency domains [32]. Using this new representation of the original load-signal, one alternative is creating a model for the short-term load forecasting whose inputs are based on information from the original load sequence and from wavelet domain subseries, too [31]. Another alternative predicts the load's future behaviour by independently forecasting each subseries in the wavelet domain and the final forecast is obtained by returning to the original domain (inverse transform) [33,34]. Some other researchers proposed merging wavelets with ANNs (called wavenets) for short-term load forecasting [35,36].

Amjady and Keynia [37] propose a new hybrid method for the short-term load forecasting of power systems, composed of wavelet transform (WT), neural network (NN) and evolutionary algorithm (EA). WT can efficiently decompose the hourly load time series into its components. Each component is predicted by a combination of NN and EA and then by inverse WT the hourly load forecast is obtained.

In this paper a new hybrid method (mixing physical and statistical approaches) is proposed, based on the wavelet decomposition technique and on artificial neural networks, in order to predict the power production of a wind farm in different time horizons: 1, 3, 6, 12 and 24 hours.

In particular, two approaches are proposed, both based on the time series of on-line measured wind power and on the Numerical Weather Predictions (NWPs); in the first approach, the forecast is carried out only through the training of a neural network which, in the second approach is, instead, used in combination with the wavelet decomposition technique, improving the performance especially over the short time horizons. For each approach two forecast systems have been developed: the first includes only the wind speed coming from the NWPs, while the second contains also pressure and temperature.

The performance in the several testing sets has been evaluated by analyzing both the normalized absolute average percentage error and the frequency distribution of the normalized relative percentage error.

2. WIND FARM CHARACTERISTICS

This work is aimed to wind power forecasting and has been based on data measured on the three wind turbines of a park in a country of the Southern Italy.

In order to define a prediction model for wind power, the most significant problem is to choose which parameter(s) will be used for the prediction.

In particular, an analysis of the historical series has been carried out, series represented by the following daily data (collected every 10 minutes): temperature (°C), wind speed (m/s), direction (degree), pressure (mmHg), wind power of each of the three turbines (kW), measured for five years.

First, it was necessary an accurate elaboration of the measured values in order to check, for each month, the days in which the parameters weren't available or were incorrect; of course an algorithm has been created in order to fill the missing parts of the historical series data and two vectors have been obtained: the hourly average wind speed for years I-V; the hourly average power for years I-V.

It's important to underline that the real values of wind power are reserved, so the normalized values are plotted in the following, adopting the range [0;1] to cover the whole range of measured powers.

Figures 1 and 2 depict the normalized hourly wind power versus, respectively, the hourly temperature and the hourly pressure in month of March, year V. It's evident the very low correlation between each of them and the wind power, so pressure and temperature haven't been taken into consideration in the developing of the forecast models.

Figures 3 and 4 show the normalized measured hourly wind power in the months of September and March, year IV. The determination coefficient of the best interpolation curve between the time series of wind power and wind speed is equal to 0.907, while the same coefficient is almost zero between power and pressure or power and temperature.

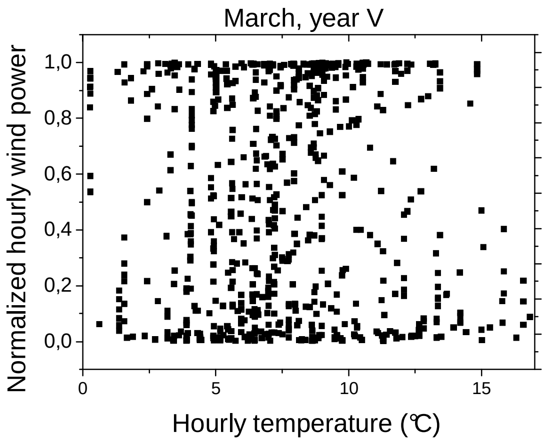

Figure 1. Normalized hourly wind power vs hourly temperature in the month of March, year V

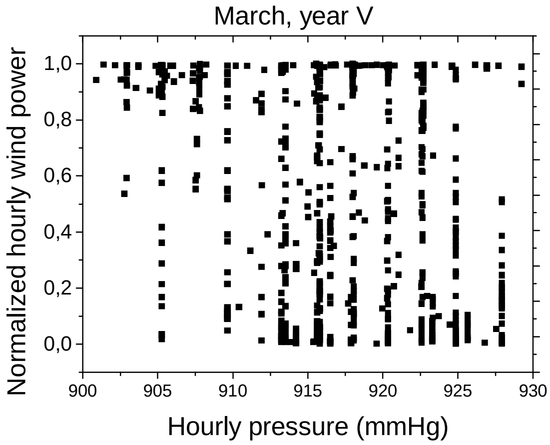

Figure 2. Normalized hourly wind power vs hourly pressure in the month of March, year V

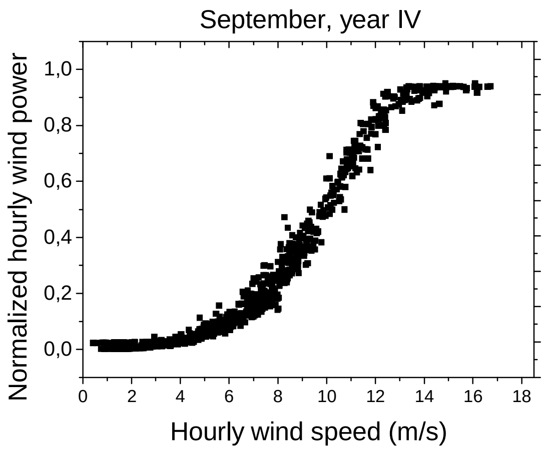

Figure 3. Normalized hourly wind power vs hourly wind speed in the month of September, year IV

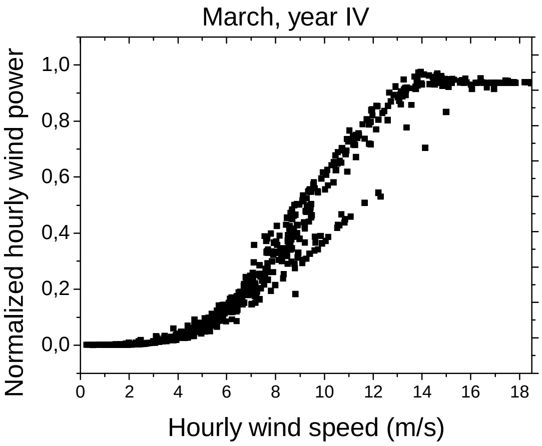

Figure 4. Normalized hourly wind power vs hourly wind speed in the month of March, year IV

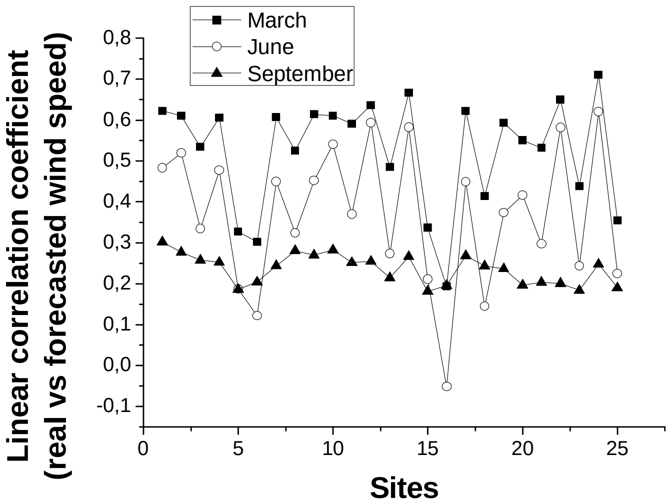

Figure 5. Linear correlation coefficient between the real hourly average wind speed and the forecasted one, year V

3. NUMERICAL WEATHER PREDICTION MODELS

Numerical Weather Prediction (NWP) models are usually developed and maintained by meteorological institutes and can be classified according to their space-temporal scale. Each NWP model tries to monitor the evolution of the atmosphere for its specific scale, even though high spatial resolution cannot be combined with high temporal resolution. In general, a NWP model with high spatial resolution (small spatial scale) will have a low temporal validity for its predictions (small temporal scale). A NWP model with a low spatial resolution (great spatial scale) will have a much greater temporal validity. The NWP models with great spatial and temporal scales are known as macro-scale models; they usually make predictions for the whole world (they are also known as global models) valid over one week. The NWP models with high spatial resolution, but with limited temporal resolution (validity) of a number of hours, are known as mesoscale models. Short-term wind power forecasting needs predictions from a NWP model with high spatial resolution.

The weather predictions used in this paper come from a mesoscale NWP model, characterized by a grid resolution of 7 km; it's initialized at 00:00 (Italy time, GMT+1:00) and supplies the NWPs for the next 72 hours at hour intervals for the following variables: mean wind speed, direction, pressure, temperature and relative humidity at a quote of about 75 m. The NWPs have been available only for the year V, at 25 sites forming a square at the center of which the three turbines are approximately placed.

Thus, it's important to individuate among all the sites, the ones for which the forecasted NWPs data are in good agreement with the real data coming from the three turbines, in order to use these data in the training algorithms illustrated in the following sections.

To evaluate the sites with the best correlation with the real data recorded at the turbines, the linear correlation coefficient between the real hourly average wind speed and the forecasted one of all the sites has been calculated and thus, the five best correlated sites have been individuated and only the NWP data coming from these five sites have been used in the forecast systems described in the following. Figure 5 depicts the linear correlation coefficient evaluated for each of the 25 sites in the months of March, June and September (year V); the best sites individuated through this analysis are: 1, 12, 14, 22, 24.

4. ARTIFICIAL NEURAL NETWORKS

Neural networks are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the network function is determined largely by the connections between elements. It is possible to train a neural network to perform a particular function by adjusting the values of the connections (weights) between elements (Fig. 6 and 7). The basic component of such a system is a neuron. When they are in action, electrochemical signals are received through synapses to the neuron cell. Each synapse has its own weight that determines the contribution and extent to which the respective input affect the output of the neuron. The weighted sum of the input electrochemical signals is fed to the nucleus which in response sends electrical impulses that are transmitted to other neurons or to other biological units as actuation signals. Neurons are interconnected through synapses. The synaptic weights keep modifying during learning. Groups of neurons are organized to subsystems and they integrate to form the brain.

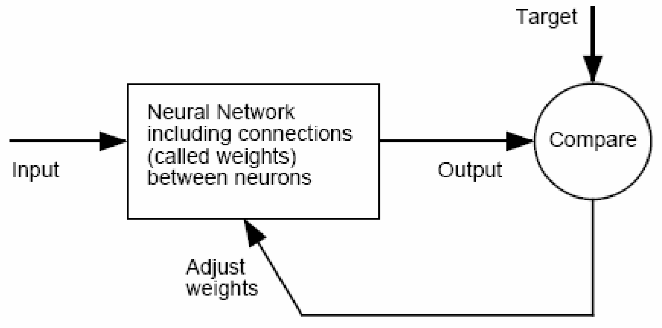

Figure 6. Scheme of operating in a neural network

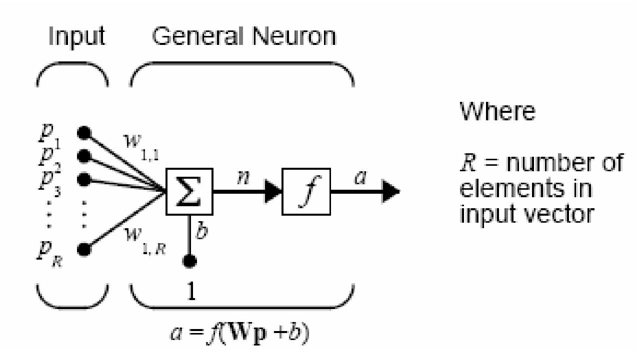

Figure 7. Model of a general neuron

In the ANN technique, a simulation of a small part of the central nervous system is done, which is rather a crude mathematical model of the biological nervous system. Inputs are fed into the corresponding neurons, and the electrochemical signals are altered by weights. The weighted sum is operated upon by an activation function, and outputs are fed to other neurons in the network. All these neurons are highly interconnected, and the activation values constitute the final output or may be fed to the next model. These connection weights are continuously modified during training to obtain the desired accuracy and generalization capabilities.

4.1 The Elman network

After a wide sensibility analysis described in [40] and oriented to evaluate the most suitable neural network to the forecast purposes of this paper (following the object to minimize the Mean Square Error), the Elman network has been chosen. It is characterized by a feedback from the first-layer output to the first layer input. This recurrent connection allows the Elman network to both detect and generate time-varying patterns (Fig. 8).

The aforementioned analysis [40] shows that the Elman network, compared to other neural networks and other kinds of forecast systems, is more suitable, especially in the final prediction length (24 h). This may well be due to the ability of the Elman network to both detect and generate time-varying patterns, (by the recurrent connection seen in Fig. 8) which appears practically negligible in the short-medium time horizons, yet becomes of great importance as the prediction length increases.

Several algorithms for training use the gradient of the performance function to determine how to adjust the weights to minimize performance. The gradient is determined using a technique called back-propagation, which involves performing computations backward through the network.

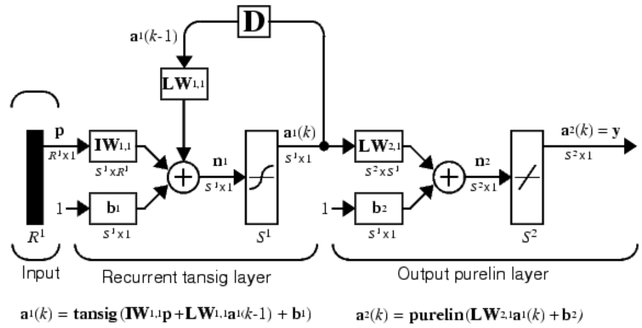

Figure 8. Typical architecture of an Elman Back Propagation network

The goal of the algorithm is to minimize the global error *E* defined as

$$E = \frac{1}{2}\sum_{k=1}^{n}\left(t(k) - o(k)\right)^2 \qquad (1)$$

where *o(k)* and *t(k)* are the output and target network for any *k* output node.

The Elman network has been used in all the methods applied.

5. THE WAVELET TRANSFORM

The Wavelet Transform (WT) can be mainly divided into two categories: Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). The CWT *W(a,b)* of a signal *f(x)* with respect to a wavelet $\Phi(x)$ is given by [31,32]:

$$W(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} f(x)\Phi(x-b)/a)dx \qquad (2)$$

where scale parameter *a* controls the spread of the wavelet and translation factor *b* determines its central position. $\Phi(x)$ is also called mother wavelet. A *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 [31,37]. This scheme is more efficient and just as accurate as the CWT [31]. It is known as the DWT:

$$W(m,n) = 2^{-m/2} \sum_{t=0}^{T-1} f(t)\Phi((t - n \cdot 2^m)/2^m)) \tag{3}$$

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=$n \cdot 2^m$); t is the discrete time index. In the proposed forecast method, a fast DWT algorithm based on the filters, e.g. decomposition low-pass, decomposition high-pass, reconstruction low-pass and reconstruction high-pass filters has been used. This algorithm was developed by Mallat [32]. Multiresolution via Mallat's algorithm is a procedure to obtain "approximations" and "details" from a given signal. An approximation is a low-frequency representation of the original signal, whereas a detail is the difference between two successive approximations. An approximation holds the general trend of the original signal, whereas a detail depicts high-frequency components of it [31]. Mathematical details of the Mallat's algorithm can be found in [32]. By successive decomposition of the approximations, a multilevel decomposition process can be achieved where the original signal is broken down into lower resolution components.

In the present work the authors use Daubechies wavelet of order 6 (abbreviated by Db6) as the mother wavelet $\Phi(t)$. This wavelet offers an appropriate tradeoff between wave-length (for evaluation of local behavior of signal) and smoothness, resulting in an appropriate behavior for short-term wind power forecasting. Mathematical details of Daubechies wavelet can be found in [38,39].

In [31] a discussion about the number of decomposition levels in the WT is presented where it is concluded that three levels of decomposition is the most promising choice for the short term wind power forecasting, because it has described the load series in a more thorough and meaningful way than the others. So, in the following three decomposition levels are considered.

6. THE FORECAST MODELS

6.1 First forecast approach: the Elman network on historical measured data and NWPs

The first forecast approach consists in the generation of the wind power forecast through the training of an Elman neural network. In particular, two forecast systems have been implemented.

In the system I, for each time instant t, the input vector is given by the average measured hourly power at that time and by the hourly wind speeds coming from the NWPs; in particular, for the five sites with the most correlated predicted variables (sites called A, B, C, D, E) the predicted wind speeds along the forecast horizon h are considered.

In the system II, for each time instant t, the input vector is given by the average measured hourly power at that time and by the hourly wind speeds, pressures and temperatures coming from the NWPs; in particular, for the above named sites (A, B, C, D, E) the predicted variables along the forecast horizon h are considered.

The target of both the systems is always made up of the sum of the average hourly powers along the forecast horizon h.

Because the NWPs were available only for the year V, the forecasting models have been applied with a training period of 8 months and on a testing period of 4 months, in 5 forecast horizons (1, 3, 6, 12, 24 hours).

The input/target schemes of the two systems are shown in Tab. 1 and 2, while Tab. 3 shows the final network parameters used in the training for each prediction length, obtained by an optimization process oriented to minimize the Mean Square Error.

Table 1. Input/target scheme of the forecast system I

Horizon (hours)	*Input*	*Unity of measurement*	*Target (kW)*
h	$v_{A,t+1} \cdots v_{A,t+h}$	m/s	$P_{t+1} + \ldots + P_{t+h}$
	$v_{B,t+1} \cdots v_{B,t+h}$		
	$v_{C,t+1} \cdots v_{C,t+h}$		
	$v_{D,t+1} \cdots v_{D,t+h}$		
	$v_{E,t+1} \cdots v_{E,t+h}$		
	P_t	kW	

Table 2. Input/target scheme of the forecast system II

Horizon (hours)	*Input*	*Unity of measurement*	*Target (kW)*
h	$v_{A,t+1} \cdots v_{A,t+h}$	m/s	$P_{t+1} + \ldots + P_{t+h}$
	$v_{B,t+1} \cdots v_{B,t+h}$		
	$v_{C,t+1} \cdots v_{C,t+h}$		
	$v_{D,t+1} \cdots v_{D,t+h}$		
	$v_{E,t+1} \cdots v_{E,t+h}$		
	$p_{A,t+1} \cdots p_{A,t+h}$	mmHg	
	$p_{B,t+1} \cdots p_{B,t+h}$		
	$p_{C,t+1} \cdots p_{C,t+h}$		
	$p_{D,t+1} \cdots p_{D,t+h}$		
	$p_{E,t+1} \cdots p_{E,t+h}$		
	$t_{A,t+1} \cdots t_{A,t+h}$	°C	
	$t_{B,t+1} \cdots t_{B,t+h}$		
	$t_{C,t+1} \cdots t_{C,t+h}$		
	$t_{D,t+1} \cdots t_{D,t+h}$		
	$t_{E,t+1} \cdots t_{E,t+h}$		
	P_t	kW	

Table 3. Elman network parameters used in the training process

		System I	System II
Training function		TRAINGDX	TRAINGDX
Adapt learning function		LEARNGD	LEARNGD
Performance function		MSE	MSE
Number layers		3	3
Neurons (layer 1)	*h=1*	6	21
	h=3	16	61
	h=6	31	121
	h=12	61	241
	h=24	121	481
Neurons (layer 2)	*h=1*	5	11
	h=3	8	31
	h=6	16	61
	h=12	31	121
	h=24	61	241
Neurons (layer 3) – output		1	1
Activation function hidden layer		TANSIG	TANSIG
Activation function output layer		PURELIN	PURELIN
Epochs		500	500

Notes

TRAINGDX = Gradient descent with momentum and adaptive learning rate backpropagation
LEARNGD = Gradient descent weight and bias learning function
MSE = Mean Squared Error
TANSIG = Hyperbolic tangent sigmoid transfer function
PURELIN = Linear transfer function

6.2 Second forecast approach: the wavelet decomposition and the Elman network on historical measured data and NWPs

The two systems (I and II) above described, based on the only application of an Elman neural network on the historical measured wind powers and on the NWPs, have been used in order to experiment a new forecasting approach based on:

- the six Daubechies wavelet employed to do the 3rd level discrete wavelet decomposition of the original hourly wind power time series (Fig. 9);
- the training of four Elman networks, like in the above described systems I and II, one for each of the four components obtained by the wavelet decomposition;
- the aggregation of the four partial forecast results of the previous step in order to generate the final forecast (Fig. 10).

This forecast approach has been applied on both the two systems I and II, keeping the same input/target schemes and the same network parameters illustrated in Tab. 1, 2 and 3. The wavelet decomposition of the time series of wind powers is, instead, depicted in Fig. 11. As discussed in the section 5, the so-called approximation (A3) is a low-frequency representation of the original signal and holds the general trend of the signal itself, while a detail is the difference between two successive approximations and depicts high frequency components of it.

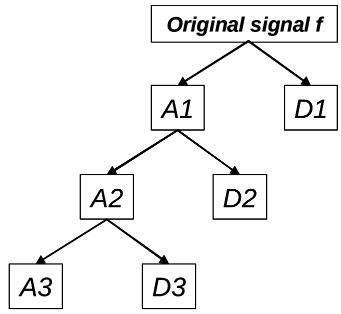

Figure 9. Multilevel decomposition process of the original hourly wind power time series (*f*): A and D stand for approximation and detail, respectively (f =A3+D1+D2+D3)

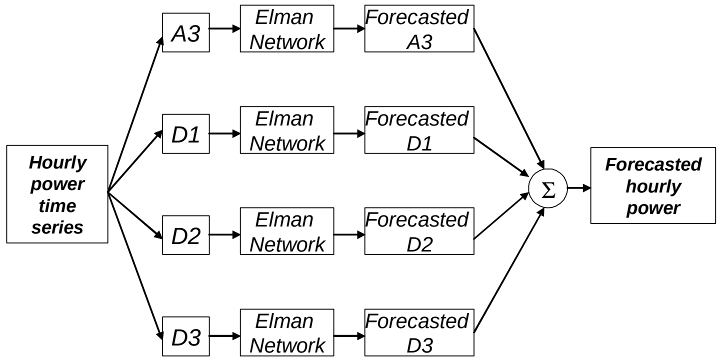

Figure 10. Architecture of the forecast approach based on the wavelet decomposition technique

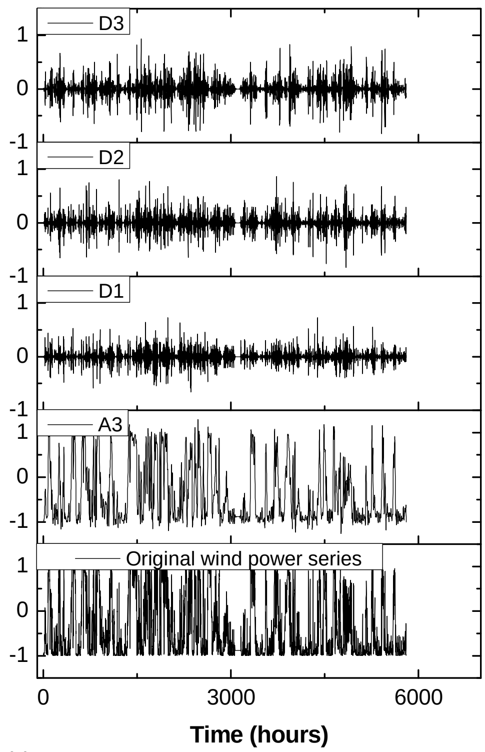

Figure 11. Six Daubechies wavelet decomposition (level 3) of the original wind power series (normalized values)

7. RESULTS

The performance of each forecast system has been evaluated through the normalized mean absolute percentage error

$$NMAPE = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{|P_i - T_i|}{Max_{i=1}^{n}(T_i)} \cdot 100 \quad (4)$$

where:
i = generic time instant;
n = number of observations;
P_i = predicted power at instant i;
T_i = real power at instant i.

Figure 12 illustrates the normalized absolute average error for the four forecast systems.

It's quite evident that the performances obtained by applying the wavelet decomposition technique in combination with the Elman network are highly better for the shorter forecast horizons, while there are no benefits when the prediction length exceeds 12 hours.

In particular, by comparing the two systems (I and II) implemented with the first forecast approach, it's noticeable that their performances are almost equivalent, with slightly better results for the systems II in the horizons of 12 and 24 hours, probably due to the further NWPs used in this system (pressure and temperature, besides the wind speed used in the system I too).

On the other hand, the second forecast approach leads to a very noticeable improvement in the shorter prediction lengths, and especially by applying the wavelet decomposition to the system I, in which the only NWP parameter is the wind speed.

However, considering only the error average values is not enough to evaluate differences in the performances of the forecast methods. Thus, a further analysis of the statistical distribution of normalized error has been performed in order to evaluate the curves of error distribution with a narrower shape.

Figures 13 and 14 depict, respectively, the probability that the error itself takes values in the ranges: [-10%; +10%] and [-20%; +20%].

Similarly to what described about the performance in terms of normalized average absolute error, the probability that the normalized error takes values in the ranges [-10% ; +10%] and [-20% ; +20%] is appreciably higher with the wavelet decomposition approach in the shorter time horizons.

By comparing the error distribution of the four forecasting models discussed in the paper (see Fig. 15, 16 and 17, concerning the time horizon of 1, 6 and 24 hours), it's evident that the two systems based on the wavelet decomposition technique are characterized by an error distribution with a narrower shape in the shorter time horizons, while the first approach, only based on the Elman artificial neural network, is preferable in the longest prediction length.

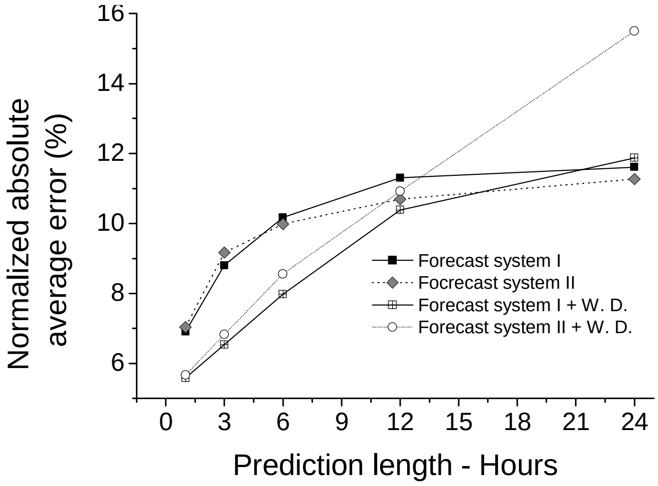

Figure 12. Normalized absolute average error vs prediction length

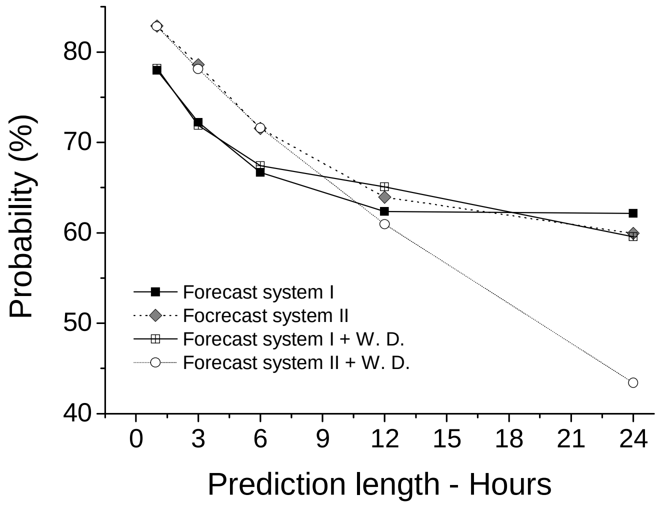

Figure 13. Probability that the normalized error takes values in the range [-10% ; +10%]

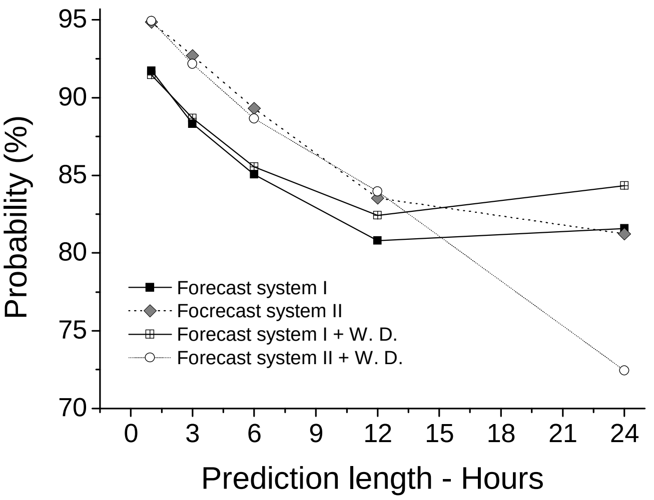

Figure 14. Probability that the normalized error takes values in the range [-20% ; +20%]

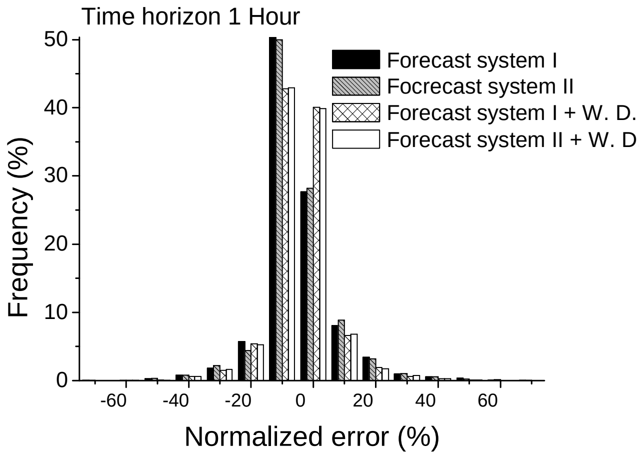

Figure 15. Error distribution (prediction length of 1 hour)

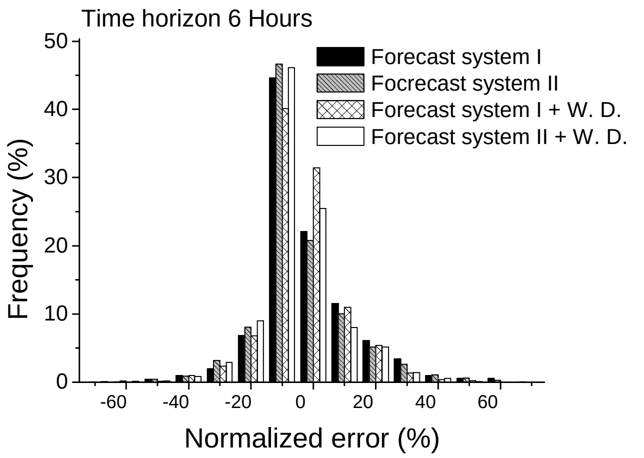

Figure 16. Error distribution (prediction length of 6 hours)

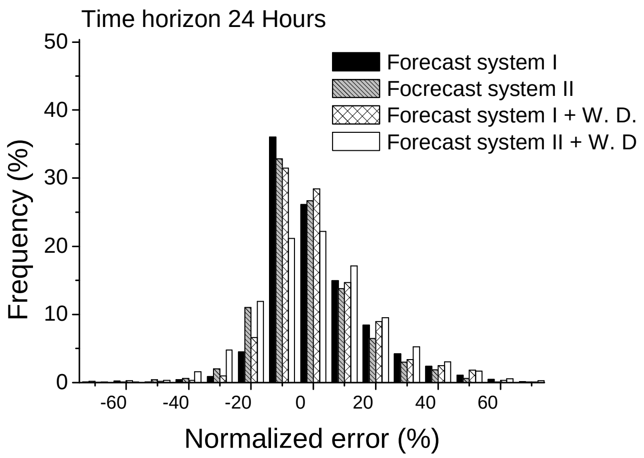

Figure 17. Error distribution (prediction length of 24 hours)

However, it's also important to underline that the wavelet decomposition approach, although characterized by better performances, it's more complex and requires four training operations to generate a forecast, while the first approach assures a somewhat short computational time. As example, on a standard desktop PC (Pentium 4, 3 GHz, RAM 1 Gb), the training operations of the neural networks described in the first forecast approach (that's to say without the wavelet decomposition) take about 10-15 minutes; of course, it's necessary to multiply by 4 this value to obtain the total computation time of the forecast systems based on the wavelet decomposition.

8. CONCLUSIONS

In this paper a new hybrid method (mixing physical and statistical approaches) is proposed, based on the wavelet decomposition technique and on artificial neural networks, in order to predict the power production of a wind farm in different time horizons: 1, 3, 6, 12 and 24 hours.

In particular, two approaches are proposed, both based on the time series of on-line measured wind power and on the Numerical Weather Predictions; in the first approach, the forecast is carried out only through the training of a neural network which, in the second approach is, instead, used in combination with the wavelet decomposition technique, improving the performance especially over the short time horizons. For each approach two forecast systems have been developed: the first includes only the wind speed coming from the NWPs, while the second contains also pressure and temperature.

The performance in the several testing sets has been evaluated by analyzing both the normalized absolute average percentage error and the frequency distribution of the normalized relative percentage error.

In particular, by comparing the two systems (I and II) implemented with the first forecast approach, it's noticeable that their performances are almost equivalent, with slightly better results for the systems II in the horizons of 12 and 24 hours, probably due to the further NWPs used in this system (pressure and temperature, besides the wind speed used in the system I too).

On the other hand, the second forecast approach leads to a very noticeable improvement in the shorter prediction lengths, and especially by applying the wavelet decomposition to the system I, in which the only NWP parameter is the wind speed.

By comparing the error distribution of the four forecasting models discussed in the paper, it's evident that the two systems based on the wavelet decomposition technique are characterized by an error distribution with a narrower shape in the shorter time horizons, while the first approach, only based on the Elman artificial neural network, is preferable in the longest prediction length.

However, it's also important to underline that the wavelet decomposition approach, although characterized by better performances, it's more complex and requires four training operations to generate a forecast, while the first approach assures a somewhat short computational time.

A further future validation of the wavelet decomposition technique in the short-term wind power forecasting could be experimented by testing it on other prediction systems like Auto Regressive Moving Average (ARMA) models or Adaptive Neuro-Fuzzy Inference Systems (ANFIS), investigating the error in prediction for various forecasting horizons and comparing the results obtained with and without the wavelet decomposition itself. This kind of tests could be very useful in order to evaluate if the benefit due to the wavelet decomposition depends upon the forecast technique used, and so to select the best method (among ANNs, ARMA and ANFIS) on which to carry out a short-term wind power forecast system.

REFERENCES

[1] Morales JM, Minguez R, Conejio AJ. A methodology to generate statistically dependent wind speed scenarios. Applied Energy 2010;87: 843-855.

[2] Zhou W, Chengzhi L, Zhongshi L, Lu L, Yang H. Current status of research on optimum sizing of stand-alone hybrid solar–wind power generation systems. Applied Energy 2010;87: 380-9.

[3] Xydis G, Koroneos C, Loizidou M. Exergy analysis in a wind speed prognostic model as a wind farm sitting selection tool: A case study in Southern Greece. Applied Energy 2009;86: 2411-2420.

[4] Hongxing Y, Wei Z, Chengzhi L. Optimal design and techno-economic analysis of a hybrid solar–wind power generation system. Applied Energy 2009;86: 163-9.

[5] Jowder F. Wind power analysis and site matching of wind turbine generators in Kingdom of Bahrain. Applied Energy 2009;86: 538-545.

[6] Luickx PJ, Delarue ED, D'haeseleer WD. Considerations on the backup of wind power: operational backup. Applied Energy 2008;85: 787-799.

[7] Saylor DJ, Rosen JN, Hu T, Li X. A neural network approach to local downscaling of GCM output for assessing wind power implications of climate change. Renewable Energy 2000;19: 359-378.

[8] Sideratos G, Hatziargyriou ND. An Advanced Statistical Method for Wind Power Forecasting. IEEE Transactions on power systems 2007; 22.

[9] Burton NJ, Bossanyi E. Wind energy handbook, Wiley 2001.

[10] Bossanyi E. Short-term stochastic wind prediction and possible control application. Proceedings of the Delphi Workshop on "Wind Energy Application" Greece 1985: 137-142.

[11] Boland J, Ward K, Korolkowiecz M. Modelling the Volatility in Wind Farm Output. School of Mathematics and Statistics-University of South Australia.

[12] Kavasseri RG, Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. Renewable Energy 2009;34: 1388-1393.

[13] Riahy GH, Abedi M. Short term wind speed forecasting for wind turbine applications using linear prediction method. Renewable Energy 2008;33: 35-41.

[14] Flores AT, Tapia G. Application of a control algorithm for wind speed prediction and active power generation. Renewable Energy 2005;33: 523-536.

[15] Sfetsos A. A novel approach for the forecasting of mean hourly wind speed time series. Renewable Energy 2002;27: 163-174.

[16] Jayaraj KP, Padmakumari K, Sreevalsan E, Arun P. Wind speed and power prediction using artificial neural networks. European Wind Energy Conference 2004 (EWEC).

[17] Cadenas E, Rivera W. Wind speed forecasting in the South Coast of Oaxaca, Mexico. Renewable Energy 2006;32: 2116-2128.
[18] Potter C, Ringrose M, Negnevitsky M. Short-term wind forecasting techniques for power generation. Australasian Universities Power Engineering Conference 2004 (AUPEC).
[19] Johnson P., Negnevitsky M, Muttaqi KM.. Short-term wind forecasting using Adaptive Neural Fuzzy Inference. System Australasian Universities Power Engineering Conference 2008 (AUPEC).
[20] Barbounis TG, Theocharis JB. Locally recurrent neural networks for long-term wind speed and power prediction. Neurocomputing 2006;69: 466–496.
[21] Bilgili M, Sahin B, Yasar A. Application of artificial neural networks for the wind speed prediction of target station using reference stations data. Renewable Energy 2007;32: 2350-2360.
[22] Kalogirou S, Neocleous C, Pashiardis S, Schizas C. Wind Speed Prediction Using Artificial Neural Networks. European Symposium on Intelligent Techniquesy 1999.
[23] Fadare DA. The application of artificial neural networks to mapping of wind speed profile for energy application in Nigeria. Applied Energy 2010;87: 934-942.
[24] Beccali M, Cirrincione G, Marvuglia A, Serporta C. Estimation of wind velocity over a complex terrain using the Generalized Mapping Regressor. Applied Energy 2010;87: 884-893.
[25] Li G, Shi J. On comparing three artificial neural networks for wind speed forecasting. Applied Energy 2010;87: 2313-2320.
[26] Kariniotakis G, Nogaret E, Stavrakakis G. Advanced short-term forecasting of wind power production. Proceedings of the 1996 European Union Wind Energy Conference EUWEC'97, Dublin, Ireland, 1997: 751-4.
[27] Palomares-Salas JC, De la Rosa JJG, Ramiro JG, Melgar J, Aguera A, Moreno A. Comparison of Models for Wind Speed Forecasting. Research Unit PAIDI-TIC-168. University of Cádiz.
[28] Sftetos A. A comparison of various forecasting techniques applied to mean hourly wind speed time series. Renewable Energy 2000;21: 23-35.
[29] Ramirez-Rosado IJ, Ferandez-Jimenez LA, Monteiro C, Sousa J, Bessa R. Comparison of two new short-term wind power forecasting systems. Renewable Energy 2009;34: 1848-1854.
[30] Lei M, Shiyan L, Chuanwen J, Hongling L, Yan Z. A review on the forecasting of wind speed and generated power. Renewableand Sustainable Energy Reviews 2009;13: 915-920.
[31] Rocha ReisAJ, Alves da Silva AP. Feature extraction via multiresolution analysis for short-term load forecasting. IEEE Trans Power Syst 2005;20(1): 189–98.
[32] Mallat S. A theory for multiresolution signal decomposition – the wavelet representation. IEEE Trans Pattern Anal Mach Intell 1989;11(7):674–93.
[33] Zhang BL, Dong ZY. An adaptive neural-wavelet model for short term load forecasting. Electric Power Syst Res 2001;59(2):121–9.
[34] Kim CI, Yu IK, Song YH. Kohonen neural network and wavelet transform based approach to short-term load forecasting. Electr Power Syst Res 2002;63(3): 169–76.
[35] Bashir Z. Short term load forecasting by using wavelet neural networks.IEEE Proc Can Conf Electr Comput Eng 2000;1:163–6.
[36] Huang CM, Yang HT. Evolving wavelet-based networks for short-term load forecasting. Proc IEE Gener Trans Distrib 2001;148(3):222–8.
[37] Amjady N, Keynia F. 2009, Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm. Energy 2009;34(1): 46-57.
[38] The Math Works, MATLAB 2006. See: <http://www.mathworks.com>.
[39] Shensa MJ. The discrete wavelet transform: wedding the a trous and Mallat Algorithms. IEEE Trans Signal Process 1992;40(10):2464–82.
[40] De Giorgi MG, Ficarella A, Tarantino M. Error analysis of short term wind power prediction models. Applied Energy 2011; 88(4):1298-1311.