Wind Forecasting and Wind Power Generation: Looking for the Best Model Based on Artificial Intelligence

Ronaldo R. B. de Aquino, Hugo T. V. Gouveia, Milde M. S. Lira, Aida A. Ferreira, Otoni Nobrega Neto, Manoel A. Carvalho Jr.
Department of Electrical Engineering
Federal University of Pernambuco (UFPE)
Recife, Brazil
{rrba@ufpe.br, hugotvg@gmail.com, milde@ufpe.br, aidaaf@gmail.com, otoni.nobrega@ufpe.br, macj@ufpe.br}

Abstract — Wind forecasting is extremely important to assist in planning and programming studies for the operation of wind power generation. Several studies have shown that the Brazilian wind potential can contribute significantly to the electricity supply, especially in the Northeast, where winds present an important feature of being complementary in relation to the flows of the San Francisco River. However, using wind power to generate electricity has some drawbacks, such as uncertainties in generation and some difficulty in planning and operation of the power system. This work proposes and develops models to forecast hourly average wind speeds and wind power generation based on Artificial Neural Networks, Fuzzy Logic and Wavelets. The models were adjusted for forecasting with variable steps up to twenty-four hours ahead. The gain of some of the developed models in relation to the reference models was of approximately 80% for forecasts in a period of one hour ahead. The results showed that a wavelet analysis combined with artificial intelligence tools provides more reliable forecasts than those obtained with the reference models, especially for forecasts in a period of 1 to 6 hours ahead.

Keywords – Wind Energy, Artificial Intelligence, Fuzzy Logic, Wind Forecasting, Neural Networks, Time Series Analysis, Wavelet Transforms.

I. INTRODUCTION

Among all alternative energy sources currently explored, wind energy is one of the most successful. One reason for this is the incentive policy by various countries, which ensures the purchase of electricity produced from wind power, even if it does not offer competitive prices. Germany and Denmark were the first countries to adopt policies to encourage the development of wind generation, followed by several countries, including Brazil, with the creation of the Incentive Program for Alternative Sources of Electric Energy – PROINFA.

The use of wind kinetic energy for electric power generation presents some inconveniences related to the uncertainties in the generation. Such inconveniences are basically caused by the variability of the wind speed. Therefore, the application of tools or techniques that are able to forecast the energy to be generated by those sources is essential to an appropriate integration of the wind source with the electric power system [1].

This paper suggests the application of Artificial Neural Networks – ANN, Adaptive Neuro-Fuzzy Inference Systems – ANFIS and multiresolution analysis by Wavelet Transform – WT toward the development of tools able to accomplish wind speed and wind power generation forecasting. The generated tools aim to subsidize planning of a hybrid generation system (hydrothermal and wind generations), mitigating some inconveniences of wind power generation [2].

The forecasting models evaluated in this study are developed from simple wind speed time series from the local wind site studied, due to a difficulty in obtaining more data from other sites. A similar kind of univariate model based on reservoir computing for wind forecasting was shown in [3].

The structure of this paper is as follows: in section II, the composition and analysis of the wind speed time series are shown. Then, in section III, the data preprocessing is presented. Section IV shows how the models are evaluated. The study and development of the models are accomplished and analyzed in section V. Comparisons between forecasts obtained with the proposed and reference models are done in section VI. Finally, in section VII, conclusions are drawn.

II. ARRANGEMENT FOR THE WIND SPEED TIME SERIES

In this paper, the created models are univariate, i.e., a single variable (wind speed) is used as an explanatory variable of the forecasting models.

The wind speed time series used to develop the models consist of hourly average values measured in three different cities in the state of Rio Grande do Norte. These cities are: MACAU, MOSSORÓ and NATAL. The data were measured at a 10m height between January and December 2008 and the size of each series is equal to 8784. Fig. 1 shows the wind speed series and their histograms.
In the univariate models, performance is directly related to the autocorrelation function, enabling us to identify the dependencies among the data. The autocorrelation function analysis also allows us to define the forecasting model inputs [1].

Fig. 2 depicts the autocorrelation function of the wind speed time series. We observed that the series autocorrelation has another maximum 24 hours after the initial peak. This behavior characterizes the winds’ cyclic behavior throughout the day.

### III. DATA PREPROCESSING

In this section, the basic procedures for processing the wind speed data are defined before creating the proposed forecasting models.

#### A. Normalization

The forecasting models based on ANN demand consistent treatment of the data in order to guarantee reasonably good performance and their effective application. In general, normalization is a straightforward and effective treatment for the data [1].

Normalization is carried out to assure that all variables used in the model inputs have equal importance during training; therefore the activation functions of artificial neurons are limited. In this paper, the data were normalized according to (1), also shown in [4], whose values ranged from 0 to 1.

$$\bar{X}(t) = \frac{X(t) - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}$$

where: $X$ and $\bar{X}$ are the wind speed time series non-normalized and normalized, respectively; $X_{\text{max}}$ and $X_{\text{min}}$ are the maximum and the minimum absolute value of the wind speed time series, respectively.

#### B. Multiresolution Analysis

The wind speed time series have a highly oscillatory behavior, so the development of models that provide satisfactory forecasts is a very complex task.

In order to mitigate the effects of wind speed variability and accomplish comparative studies among several forecasting models, a multiresolution analysis by WT is proposed in this paper. This analysis facilitates adjusting the parameters and training of the forecasting models.

The multiresolution analysis decomposes signals into different time scales resolutions. This analysis is accomplished through the pyramidal algorithm, which makes WT calculation more efficient. This algorithm makes it possible to obtain "approximation" and "details" of a given signal [1].

After preliminary tests of decomposition by wavelets using several wavelet families to identify which one could be the best to decompose the original wind speed signal, it was found that the Daubechies family wavelets were more robust [1].

Once defined the wavelet family to be used, an analysis was accomplished to decide on the best level to decompose the signal. In this analysis, it was found that the signals should be decomposed at level 3, because at higher decomposition levels, the detail coefficients did not present relevant information for the forecasting models [1].

#### C. Conversion of Wind Speed Forecasting to Wind Power Generation

The selected turbine was ENERCON E-70 E4, which is designed to operate at a 57m height. However, the wind speed time series were measured at a 10m height, consequently, the input and output patterns used to adjust and train the forecasting models are related to this height. Therefore, these models accomplish wind speed forecasting at a reference height of 10m.

Before transforming the wind speed forecasting into wind power forecasting using the wind turbine manufacturer's power curve, it is necessary to convert the wind speed forecasting from the height of 10m to the turbine height of 57m.

The conversion formula is given in (2), using the coefficient of roughness $\alpha = 0.1$. In this equation, $s$ represents the wind speed at a height of $H$ and $s_0$ represents the wind speed

$$s = s_0 \left(1 - \frac{H}{10m}\right)$$
forecasting at a reference height of \( H_0 \).

\[
s = s_0 \left( \frac{H}{H_0} \right)^\alpha
\]  

(2)

IV. PERFORMANCE EVALUATION

A. Definition of Forecast Errors

The four basic types of errors used in this paper to illustrate the performance of a wind speed forecasting model are: the mean absolute error – \( \text{MAE} \), mean absolute percentage error – \( \text{MAPE} \), the mean squared error – \( \text{MSE} \) and root mean squared error – \( \text{RMSE} \). The equations used to calculate them are defined as follows:

\[
\text{MAE}_s(k) = \frac{1}{N} \sum_{i=1}^{N} |e_i(t + k | t)|
\]   

(3)

\[
\text{MAPE}_s(k) = \frac{100}{N} \sum_{i=1}^{N} \frac{|e_i(t + k | t)|}{s(t+k)},
\]  

(4)

\[
\text{MSE}_s(k) = \frac{1}{N} \sum_{i=1}^{N} e_i(t + k | t)^2,
\]  

(5)

\[
\text{RMSE}_s(k) = \sqrt{\text{MSE}_s(k)},
\]  

(6)

where: \( N \) - number of data used for the model evaluation; \( k \) - forecast horizon (number of time-steps); \( e_i(t+k|t) \) - error corresponding to time \( t+k \) for the wind speed forecast made at origin time \( t \); \( s(t+k) \) - measured wind speed at time \( t+k \).

The errors defined in equations (3) to (6) refer to the wind speed forecasts. The two types of errors used in this paper to illustrate the performance of the power generation forecasts are: the normalized mean absolute error – \( \text{NMAE} \) and normalized root mean squared error – \( \text{NRMSE} \). The equations used to calculate them are defined as follows:

\[
\text{NMAE}_P(k) = \frac{100}{N \cdot P_{inst}} \sum_{i=1}^{N} |e_P(t + k | t)|
\]   

(7)

\[
\text{NRMSE}_P(k) = \frac{100}{P_{inst}} \sqrt{\frac{1}{N} \sum_{i=1}^{N} e_P(t + k | t)^2},
\]  

(8)

where: \( e_P(t+k|t) \) - error corresponding to time \( t+k \) for the power generation forecast made at origin time \( t \); \( P_{inst} \) - installed wind power capacity.

B. Reference Models

It is worthwhile developing and implementing an advanced wind power forecasting tool if it is able to beat reference models, which are the result of simple considerations and not of modeling efforts. Probably the most common reference model used in the frame of wind power prediction or in the meteorological field is the persistence model [5]. This naive predictor tool states that the variable’s future value will be the same as the last one measured, i.e.

\[
y_{PERS}(t + k | t) = y(t),
\]  

(9)

where: \( y_{PERS}(t+k|t) \) - wind speed (or power generation) forecast for time \( t+k \) made at origin time; \( y(t) \) - measured wind speed (or power generation) at time \( t \).

In reference [6], the authors proposed a new reference – NR model whose forecasts are given according to the following equation:

\[
y_{NR}(t + k | t) = a_k y(t) + (1 - a_k) \mu(t),
\]  

(10)

where: \( \mu(t) \) - average of all the available observations of wind speed (or power generation) at time \( t \); \( a_k \) - is the correlation coefficient between \( y(t) \) and \( y(t+k) \).

C. Criteria for Comparison of Models

It might be of interest to highlight and to quantify the gain of preferring an advanced approach to the reference ones [6]. This gain, denoted as an improvement with respect to the considered reference model, is

\[
I_{ref, EC}(k) = 100 \left( \frac{EC_{ref}(k) - EC(k)}{EC_{ref}(k)} \right),
\]  

(11)

where: \( EC \) - considered Evaluation Criterion, which can be \( \text{MAE}, \text{MAPE}, \text{RMSE}, \text{NMAE} \) or \( \text{NRMSE} \).

Another way to evaluate the models’ performance is through the determination of the \( R^2 \) coefficient. This coefficient represents the ability of the model to explain the data’s variance. The value of \( R^2 \) must be between zero and one. The closer the unit is to the \( R^2 \) value, the better the forecast model. For each look-ahead time, the \( R^2 \) coefficient can be calculated by the following equation:

\[
R^2(k) = 1 - \frac{\sum_{i=1}^{N} e(t + k | t)^2}{\sum_{i=1}^{N} y(t + k)^2},
\]  

(12)

V. DEVELOPMENT OF THE FORECASTING MODELS

This paper proposes six different models for the forecasts of hourly average wind speeds. Four of them use ANN and the other two use ANFIS. The power generation forecasts for each
of the models is obtained using the forecasted speeds and the wind turbine power curve.

A. Models: ANNLM, ANNRP and ANFIS

These three models have the same entry and exit standards, that is, four inputs and one output. The input data are the last four hourly average speeds. The output is the forecast of the hourly average speed for a given horizon. Fig. 3 schematically illustrates the models.

![Figure 3. Illustration of ANNLM, ANNRP and ANFIS.](image)

The ANNLM and ANNRP models are formed by Multilayer Feedforward fully connected neural networks whose training algorithms are Levenberg-Marquardt and Resilient Propagation, respectively. These networks’ architecture consists of an input layer, an intermediate layer (hidden layer) and an output layer.

The number of neurons in the hidden layer of neural models is determined by varying the number of neurons in this layer. The amount that provided the best performance during training was thus selected. The hidden layer neurons use the Hyperbolic Tangent activation function. These models’ output layer consists of only one neuron, because there is only one exit. This neuron’s activation function is Sigmoid Logistics.

The ANFIS model also has four inputs and one output. The selection of the best inference system was performed using the subtractive clustering technique. The length of each cluster’s influence radius was determined by varying its value. The length that provided the best performance during training was thus selected.

B. Models: WTANNLM, WTANNRP and WTANFIS

These models differ from the previous ones in the kind of inputs. Here, the inputs are the wavelet coefficients from the multiresolution analysis, using the Daubechies wavelet family (db10) at level 3. This causes an increase in the number of inputs from 4 to 16, since the inputs are now composed by the approximation and detail coefficients, according to the illustration shown in Fig. 4.

![Figure 4. Illustration of WTANNLM, WTANNRP and WTANFIS.](image)

C. Databases Partition

Before performing the models’ training, a matrix of input and output patterns was set up for each series of speeds, and for each forecasting horizon (1 to 24 hours ahead). For each row of these matrices standards, the first columns are the inputs and the last column is the output. For ANNLM, ANNRP and ANFIS, these arrays have five columns, while for WTANNLM, WTANNRP and WTANFIS, these arrays have seventeen columns.

As stated earlier, the training standards of the models using neural networks must be unvarying, so the matrices were normalized before performing the ANNLM, ANNRP, WTANNLM and WTANNRP training.

The database was divided into 60% training sets, 30% validation and 10% testing sets.

D. Training for the Chosen Architecture

In projects involving ANN, there is a correlation among the number of patterns used in the training set, the complexity of the problem to be handled, the number of free parameters (weights and bias) and the learning process, as shown in [7]. However, deterministic rule exists relating these variables; in spite of some researchers prefer certain practical rules. In this work, these variables are restricted to the number of free parameters, i.e. the ANN architecture, because the number of training patterns is limited to the size of the database and the complexity of the problem is intrinsic to the kind of application.

In this regard, to select the ANN architecture, the rule in [4] was used. The rule states that architectures must be defined by changing the number of neurons in the hidden layer from 3 to 15, and selecting the one with the best performance (lowest MAPE) in the validation set during the training process.

E. Choosing the Best ANN for Each Neural Model

The validation method is also known as k-fold cross-validation, where k represents the number of partitions randomly generated from the examples to train, test and validate the ANN. In this method, the samples are divided into mutually exclusive k partitions. For each of k iterations, n different partitions are used to test the ANN and all the others (k – n) are used to train and validate [1].

It is expected that by applying this technique, the error’s mean value in each experiment on the test set represents the expected output of the ANN created by the k experiments, as shown in [7, 8]. Here, the k value was set equal to 10.

F. Choosing the Best Inference System for Each Neuro-Fuzzy Model

The methodology used to choose the best Neuro-Fuzzy model was proposed in [4]. From the training set, the subtractive clustering technique was used in order to generate the initial Fuzzy Inference Systems – FIS with the cluster radius ranging from 0.3 to 0.7, and an increment equal to 0.1. Then, ANFIS was used to adapt the membership functions of the generated FIS. Training and validation sets were used to adapt these functions. Finally, the selected models were those that had the lowest MAPE for the test set.
G. Results of the Chosen Models

The best MAPE values for the models chosen for each of the analyzed sites are presented below.

Fig. 5 shows the results for MACAU. We can observe that for forecasts up to 7 hours ahead, models that use WT provide better forecasts than those that do not use it. From 13 hours ahead, the Neuro-Fuzzy models provide better forecasts than the others, with the WTANFIS model providing the best forecasts for all considered time steps.

Fig. 6 shows the results for MOSSORÓ. We can see that for forecasts up to 9 hours ahead, models that use WT provide better forecasts than those that do not use it. From 19 hours ahead, the Neuro-Fuzzy models provide better forecasts than the others, with the WTANFIS model providing the best forecasts for all considered time steps. For MOSSORÓ, MAPE values are higher due to the large number of speeds less than 1 m/s, which increases the MAPE.

Fig. 7 shows the results for NATAL. It is clear, that for forecasts up to 4 hours ahead, models that use WT provide better forecasts than those that do not use it. From 13 hours ahead, the Neuro-Fuzzy models provide better forecasts than the others, with the WTANFIS model providing the best forecasts for all considered time steps.

VI. FORECASTS AND COMPARISONS BETWEEN THE MODELS

This section presents comparisons between WTANNLM and WTANFIS performances with those obtained using reference models (9) and (10) for MACAU. WTANNLM and WTANFIS were chosen because they presented, in most cases, a smaller MAPE for test sets.

A. Wind Speed Forecasts for MACAU

The period chosen for the MACAU forecasts evaluation is formed by 8016 hourly average speed from 01/01/2009 to 30/11/2009.

The different MAE, RMSE and MAPE values are shown in Figs. 8, 9 and 10, respectively. We can observe that the models proposed in this paper provide better forecasts than reference models for all look-ahead times. Another interesting feature of the proposed models is that up to 12 hours ahead, the average forecasts errors are significantly smaller than those obtained with the reference models. From 12 to 24 hours ahead, the average errors of the proposed models hardly vary with the forecast horizon, while for the reference models these errors reduce.
The improvement in MAPE with respect to a persistence model for the wind speed forecasts is shown in Fig. 11. It can be observed that for forecasts up to 6 hours ahead, the proposed models’ improvements are greater than 60%.

Fig. 12 shows the cumulated squared errors for wind speed forecasts one hour ahead.

Scatter diagrams along with the determination coefficients for wind speed forecasts one hour ahead are shown in Fig. 13. In these diagrams, the closer the points are on the line, the better the forecasts.

B. MACAU’s Power Generation Forecasts

Fig. 14 shows the power curve used to perform power generation simulations in this paper. It is a wind turbine rated at 2300 kW and the hub height is equal to 57 meters.

Manufacturers provide only a few points of the curve (red points), so to estimate the power (in kW) provided at any speed between the cut-in and cut-out, the curve was adjusted in four different intervals through the least squares method.
The different $NMAE$ and $NRMSE$ values are shown in Figs. 15 and 16, respectively. We can observe that for the proposed models, the maximum $NMAE$ and $NRMSE$ are about 6.5% and 12% of installed wind power capacity, respectively. For the persistence model, the maximum $NMAE$ and $NRMSE$ are about 13% and 22% of installed wind power capacity, respectively.

According to [9], typical model results for single wind forecasting are: $NMAE$ around 6–9% and $NRMSE$ around 10–13% of the installed capacity for the first 6 hours, rising to 13–16% and 18–22% for 48 hours ahead, respectively.

In [10], the $NMAE$ value for six hours ahead is around 15%, while in [11], the $NMAE$ is around 9% at the same time horizon. In [12], $NMAE$ values for forecasts from 1 to 4 hours ahead are 5.98%, 9.47%, 12.00% and 15.21%, respectively.

The average $NMAE$ value for 3 hours ahead forecasts with the hybrid model proposed in [13] is equal to 1.65%. In this same study, the average $NMAE$ values for the persistence and new reference models are 5.24% and 5.17%, respectively.

The improvement in $NMAE$ with respect to the persistence model is shown in Fig. 17. We can see that for forecasts up to 15 hours ahead, improvements brought on by the proposed models are greater than 50%. For 1 hour ahead, there is an almost 80% improvement for $WTANFIS$.

Fig. 18 shows the cumulated squared errors for power generation forecasts one hour ahead. Periods in which the graphs have a greater slope refer to the months of the year in which the average wind speeds are higher due to seasonality. It can be observed that the slope is less pronounced for proposed models, especially for $WTANFIS$. 

Figure 14. Wind turbine power curve.

Figure 15. $NMAE$ of power generation forecasts.

Figure 16. $NRMSE$ of power generation forecasts.

Figure 17. $NMAE$ improvement with respect to persistence model of power generation forecasts.
Scatter diagrams along with the determination coefficients for power generation forecasts one hour ahead are shown in Fig. 19. A lower variance of power generation forecasts obtained with the proposed model can be observed. Most WTANNLM model power generation forecasts are lower than the observed values. The WTANFIS model presented the best determination coefficient.

VII. CONCLUSIONS

The models developed in this paper are based on time series analysis of wind speeds through computational intelligence. The forecast horizons presented fall within short-term forecasts for up to twenty-four hours ahead, which is an appropriate horizon to subsidize the operation planning of power systems.

All developed models provided good forecasts for all the look-ahead times considered, especially for the shorter ones (from 1 to 6 hours ahead). We found that the forecasts’ quality is strongly influenced by the series autocorrelation for both adopted reference models and those models that do not use WT.

Decomposition of wind speed series through WT allowed models to extract relevant information about the cyclical and seasonal behavior of wind speeds. The information presented in the wavelet transforms through its approximation and detail components was crucial to significantly improve forecasts made with models that use these signals as inputs.

Through the analysis, we found that the behavior of forecast errors for different horizons was very similar for all the three sites. The adopted methodology for the models’ development was quite adequate, which ensured statistically reliable forecasts, i.e., the models acquired the ability to generalize without becoming biased.

REFERENCES