energies

MDPI

Article

Different Models for Forecasting Wind Power Generation: Case Study

David Barbosa de Alencar [1,*], Carolina de Mattos Affonso [1], Roberto Célio Limão de Oliveira [1], Jorge Laureano Moya Rodríguez [2], Jandecy Cabral Leite [3] and José Carlos Reston Filho [4]

[1] Department of Electrical Engineering, Federal University of Para—UFPA, Belém 66075-110, Brazil; carolina@ufpa.br (C.d.M.A.); limao@ufpa.br (R.C.L.d.O.)
[2] Department of Industrial Engineering, Universidade Federal da Bahia, Salvador 40170-115, Brazil; jorgemoyar@gmail.com
[3] Department of Research, Institute of Technology and Education Galileo of Amazon—ITEGAM, Manaus 69020-030, Brazil; jandecy.cabral@itegam.org.br
[4] Department of Postgraduate Curses, IDAAM., Manaus 69055-038, Brazil; jcreston@gmail.com
* Correspondence: david002870@hotmail.com; Tel.: +55-92-98125-1765

Received: 23 October 2017; Accepted: 23 November 2017; Published: 29 November 2017

Abstract: Generation of electric energy through wind turbines is one of the practically inexhaustible alternatives of generation. It is considered a source of clean energy, but still needs a lot of research for the development of science and technologies that ensures uniformity in generation, providing a greater participation of this source in the energy matrix, since the wind presents abrupt variations in speed, density and other important variables. In wind-based electrical systems, it is essential to predict at least one day in advance the future values of wind behavior, in order to evaluate the availability of energy for the next period, which is relevant information in the dispatch of the generating units and in the control of the electrical system. This paper develops ultra-short, short, medium and long-term prediction models of wind speed, based on computational intelligence techniques, using artificial neural network models, Autoregressive Integrated Moving Average (ARIMA) and hybrid models including forecasting using wavelets. For the application of the methodology, the meteorological variables of the database of the national organization system of environmental data (SONDA), Petrolina station, from 1 January 2004 to 31 March 2017, were used. A comparison among results by different used approaches is also done and it is also predicted the possibility of power and energy generation using a certain kind of wind generator.

Keywords: wind power; wind speed; time series; ARIMA; forecasting; wavelets

1. Introduction

The evaluation of wind potential in a region requires systematic data collection and analysis on wind speed and regime. Generally, a rigorous assessment requires specific surveys of the region where the wind farm will be placed [1–3]. There are three major markets for the field of global wind power generation: Europe, USA and China [4]. Wind energy penetration levels continue to rise, led by Denmark with a 40% use of this energy, followed by Uruguay, Portugal and Ireland with over 20%; Spain and Cyprus with about 20%; Germany with 16%; and the major markets of China, the US and Canada with 4%, 5.5% and 6% wind energy, respectively. The forecast of five years ahead is almost 60 GW of new wind power installations in 2017, rising to an annual market of 75 GW by 2021, and an accumulated installed capacity of more than 800 GW by the end of 2021 [5]. Wind energy is a clean and renewable alternative for the production of electric energy, presenting great social acceptance [6]. In the social feature, wind power plants do not cause major environmental impacts such as in hydroelectric plants and allow the compatibility between the production of electricity from the wind and the use of land for livestock and agriculture.

Wind generation occurs through the contact of the wind with the blades of the wind device. When rotating, these blades convert wind speed into mechanical energy that drives the rotor of the wind generator, which produces electricity. According to [7], tropical regions receive solar rays almost perpendicularly and are therefore warmer than the polar regions. Consequently, the warm air that is found in the low altitudes of the tropical regions tends to rise, being replaced by a mass of cooler air that comes from the Polar Regions.

Wind is the result of the displacement of air masses, caused by the effects of atmospheric pressure differences between two distinct regions and influenced by natural effects such as continentally, sea level, latitude, altitude, and soil roughness, among others [8]. According to [9], wind power temporal series always have non-linear and non-stationary characteristics and therefore it is very difficult to accurately forecast the power generated. In [10], it is established that accurate wind forecasting is decisive to have a reliable power system. However, the intermittent and unstable nature of the wind speed makes it very difficult to predict accurately. The objective of this paper is to develop a hybrid system composed by ARIMA model and two Neural Networks to forecast wind power. The proposed model was applied to a case study in Brazil, and results are according to reality.

2. Literature Review

2.1. Wind Power Generation Potential

The potential of electric energy produced from wind generation is obtained through the kinetic energy of the wind, which is converted into mechanical energy by a process that turns the wind force into a torque that acts on the rotor blades. The amount of energy generated by winds is a function of their speed (v) and mass (m) and is given by the kinetic energy equation [11,12]; thus, it is very important to make a good wind speed prediction. The power available in the wind, however, cannot be fully utilized by the wind turbine for the generation of electricity. According [7], to take into account this physical characteristic, an index called power coefficient C_p is introduced, which can be defined as the fraction of the available wind power that can be extracted by the rotor blades (see Figure 1).

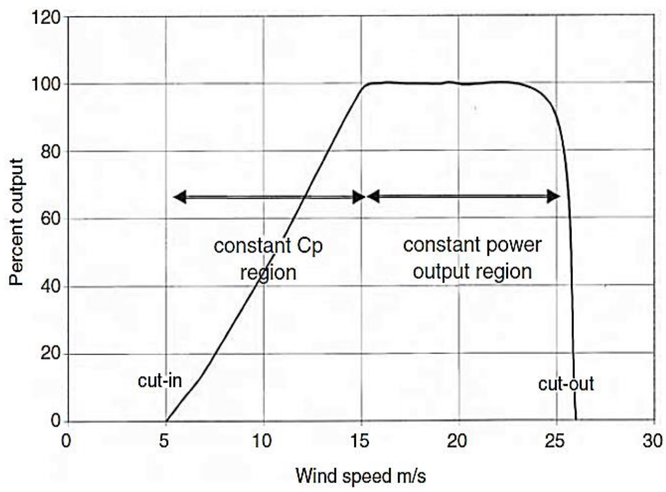


Figure 1. Regions of turbine speed control. Source: [11].

According to Betz's Law, no wind turbine can convert more than 59.3% of the kinetic energy of the wind into mechanical energy transformed at the rotor ($C_p \leq 59.3\%$), that is, only 59.3% of the energy contained in the air flow can theoretically be extracted by a wind turbine [13,14].

The potential of electric energy produced from wind generation is obtained through the kinetic energy of the winds, which is converted into mechanical energy by a process that turns the wind into torque acting on the rotor blades.

According to [11], the power curve regions can be described as follows:

- Optimum constant C_p region, where increasing power with increasing wind speed;
- Limited power region, generating a constant power, even in higher winds, by decreasing the C_p rotor efficiency; and
- Region of power shutdown, where power generation is decelerated to zero, and wind speed approaches the cut-out limit.

In [15], it is emphasized that wind speed prediction plays a vital role in the management, planning and integration of the energy system. In previous studies, most forecasting models have focused on improving the accuracy or stability of wind speed prediction. However, for an effective forecast model, considering only one criterion (precision or stability) is insufficient.

In [16], a new design methodology to predict wind farm energy production by means of a spiking neural network-based system is developed. Authors established that the calculation of the flow around wind turbines is a very complicated issue. They affirmed that turbine wakes are responsible for important power losses in wind farms and that randomness of wind speed and unexpected variations of wind speed may increase operating costs of the electricity grid as well as set potential threats to the reliability of electricity supply. A synergetic neural network (SNN)-based model for the prediction of wind farm energy production is developed. This model performs a prediction of the energy produced by one wind turbine of the wind farm, by using the wind speed and direction data coming from the three anemometric towers, during the whole day. This is a very useful model for analyzing turbines inside of a farm. Authors demonstrate that this model could accurately predict the energy produced by each wind turbine of the wind farm by means of testes conducted with experimental data that show low values of median absolute deviation (MAD), mean absolute percentage error (MAPE) and root mean square error (RMSE).

In this work, the focus is limited to predictive models of time series with the use of ARIMA filters, as well as the use of neural networks, however, it will also be analyzed the auto regressive (AR) models, moving average model (MA) and auto regressive moving-average (ARMA) model to enable a better understanding of the ARIMA Model.

2.2. *Types of Wind Energy Forecast*

There are different methods for predicting the wind power to generate. These methods are classified according to time scales and according to different methodologies that are available in the literature.

Time scales and methods for predicting wind power or energy, combining the literature can be divided into four categories [17–19]:

- Ultra-short-term forecast: From a few minutes to 1 h ahead.
- Short-term forecast: From one hour to several hours ahead.
- Medium term forecast: From several hours to one week ahead.
- Long-term forecast: From one week to one year or more ahead.

2.3. *Wind Speed Prediction Models*

Wind forecasting models can be broadly classified into the following three categories: (i) physical model; (ii) statistical and computational model; and (iii) hybrid model [6,20,21].

The artificial intelligence approach belongs to the statistical approach. The essence of the artificial intelligence approach is to establish the relationship between input and output by artificial intelligence methods, rather than using the analytical method. The model described in this form is usually a non-linear model. Many methods of artificial intelligence are better than conventional methods and have a good perspective of development [22].

2.4. Statistical Models and Artificial Neural Networks

Statistical models are easy to use and cheaper to develop compared to other models. Basically, statistical methods use the previous history of wind data to perform a forecast over the next few hours, they are good for short periods of time. The disadvantage of this method is that prediction error increases as time forecasting increases, i.e., statistical time series and methods of neural networks are primarily intended for short-term predictions [23,24]. According to [25], the sub classification of this approach can be defined as: models based on time series and methods based on neural networks. These forecasting methods are generally used for small forecast horizons (mesoscale), because, in these horizons, the correlation between the velocities of the winds, and consequently the generation, are greater. The statistical models most disseminated by researchers include: auto regressive (AR), auto regressive moving average (ARMA), and auto regressive integrated moving average (ARIMA).

Statistical methods, in many predictions, use the difference between predicted and actual wind speeds to adjust model parameters. The advantage of Artificial Neural Networks (ANNs) is to know the relation between inputs and outputs by a non-statistical approach. According to [14], neural networks can easily learn from the input and output mapping during the training phase. This allows the neural models to perform well, even without the researchers' knowledge of the problem.

The advantage related to the other methods is to provide relatively inexpensive models of statistical projection that do not require any data other than historical wind power generation data. However, the prediction accuracy for these models falls significantly when the time horizon is extended [23].

In [26], it is presented a new statistical method based on the AR model and analysis of independent components. Based on the results obtained, the proposed method obviously gives a greater precision compared to direct predictions.

2.5. Mixed Model or Hybrid Model

Usually, the combination of different approaches, such as physical and statistical approaches, models combining short and medium term, etc., is commonly referred to as a mixed or hybrid model approach [25]. According to [24], the object of hybrid models is to benefit themselves from the advantages of each model and obtain optimum overall forecast performance. Since the information contained in the individual forecasting method is limited, the hybrid method can maximize the available information, integrate information from individual models, and make the best use of the advantages of various forecasting methods, thereby improving prediction accuracy.

Many types of hybrid models were used to predict wind power. The types of combinations can be [23,24]:

- Combination of physical and statistical approaches;
- Combination of models for short term and medium term;
- Combination of alternative statistical models; and
- Combination of alternative models of artificial intelligence.

According to [14], mixed models are almost always used in horizons of short predictions (mesoscale) to adjust the results found through Numeric Weather Prediction (NWP) models

2.6. Methods for the Prediction and Classification of Time Series

Time series modeling is a dynamic research area that has attracted the attention of scientific communities over the past decades. According to [27], the main purpose of time series modeling is to collect carefully and rigorously study the past observations of a time series to develop an appropriate model, which describes the inherent structure of the series. This model is then used to generate future values for the series, i.e., to make predictions. Forecasting time series, therefore, can be termed as the act of predicting the future through the understanding of the past [28].

It is obvious that a successful time series forecast depends on an adjustment of an appropriate model. Great efforts have been made by researchers over many years to develop efficient models to improve forecast accuracy. As a result, the various models of important forecast time series have evolved in the literature.

According to [29], some characteristics are particular to this type of data, for example:

- Correlated observations are more difficult to analysis and require specific techniques.
- It is necessary to take into account the temporal order of observations.
- Complicating factors such as presence of trends and seasonal or cyclical variation may be difficult to estimate or remove.
- Model selection can be quite complicated, and the tools can be difficult to interpret.
- It is more difficult to deal with missed observations and discrepant data due to the sequential nature.

According to [27], one of the most popular and frequently used models of stochastic time series is the integrated auto regressive integrated moving average (ARIMA).

2.7. Stationary Time Series

One of the most frequent assumptions about a time series is that they are stationary, that is, they develop themselves randomly during time around a constant mean, reflecting some form of stable equilibrium [30]. However, in practice, most time series have in their nature some form of non-stationarity.

According to [31], in general terms, a stochastic process will be called a stationary if its mean and variance are constant over time and the covariance value between the two time periods depends only on the distance, the interval or the lag between the two periods and not of the actual time at which covariance is computed.

Stationary models are those who assume that the process is in equilibrium. According to [32], a process is considered to be weakly stationary if its means and variances remain constant over time, and the auto covariance function only depends on the time interval. A process is strongly stationary if all joint moments are invariant to time course.

2.8. Non-Stationary Time Series

If a time series is not stationary in the sense just defined, it is called a non-stationary time series. In other words, a non-stationary time series will have a mean that varies with time, or a variance that varies over time, or both [31].

According to [33], the non-stationary linear models studied in the literature are those that have a "non-explosive" behavior, thus presenting a homogeneous behavior. They are such series that, by taking a finite number of differences become stationary. These models are called ARIMA (Autoregressive Integrate Moving Average).

The non-stationary series have a tendency, being deterministic or stochastic in nature. In economics, price indices and product levels are examples of non-stationary series.

2.9. Box–Jenkins Models

The main Box–Jenkins models for estimation and prediction of time series and their mathematical formulation are presented in [29,34,35]. These models belong to the family of autoregressive moving average models (ARMA), subdivided into two other models: autoregressive (AR) and MA. However, these models are suitable for stationary series, i.e., those where the mean is constant all the time, but in general, the series are non-stationary; for example, economic series. Therefore, the model that will be presented for series whose behavior is non-stationary, is the ARIMA.

The purpose of the Box–Jenkins method is to identify and estimate a statistical model that can be interpreted as having been generated by the sample data. If this estimated model is used for forecasting, it must be admitted that its characteristics are constant over the period, and particularly over future periods. The simple reason for requiring stationary data is that any model that is inferred based on these data can be interpreted as stationary or stable and therefore provide a valid basis for the prediction [36,37].

2.10. Autoregressive Models (AR)

This form of an AR is intuitively attractive and has been widely applied to data sets in several areas [38–40]. They can be considered as:

- First-order autoregressive model: AR (1); and
- Auto-regressive model of order p: AR (p).

2.11. Moving Average Models (MA)

The MA model describes how an observation directly depends on one or more previous measurements [34,41,42]. They can be considered as:

- First-Order Moving Average Model: MA (1); and
- Moving Average Model of order q: MA (q).

2.12. Autoregressive Moving Average Models

A combination of the AR (p) and MA (q) models results in an autoregressive and moving average model, i.e., an autoregressive integrated moving average (ARMA). If a process consists of both parameters, MA and AR, it is called an ARMA process [34,43].

2.13. Autoregressive Integrated Moving Average Models (ARIMA)

When the process is non-stationary, the combination of autoregressive and moving average models results in an ARIMA (p, d, q) model, where d is the number of differences necessary to make the series Stationary.

According to [31], if it is necessary to differentiate a time series d times to make it stationary and apply the ARMA model (p, q), it can be said that the original time series is ARIMA (p, d, q); that is, it is an autoregressive integrated MA time series, where p denotes the numbers of autoregressive terms, d is the number of times the series must be differentiated before becoming stationary, and q is the number of moving average terms. According to [31], an ARIMA process (p, 0, 0) means a purely stationary AR (p) process; and an ARIMA (0, 0, q) means a purely stationary MA (q) process. Given the values of p, d and q, it is possible to tell which process is being modeled.

The question, of course, is: looking at a time series, how is it possible to know whether it follows a purely AR procedure (if so, what is the value of p) or a purely MA procedure (if so, what is the value of Q) or an ARMA process (and if so, what are the values of p and q) or an ARIMA process, in which case we need to know the values of p, d, and q. The Box–Jenkins methodology is very useful in answering the previous question [31].

According to [44], the basic idea of model identification is that if a time series is generated from an ARIMA process, it must have some theoretical properties of autocorrelation. By combining empirical autocorrelation patterns with theoretical ones, it is often possible to identify one or several possible models for a given time series

The basis of the Box–Jenkins methodological approach to temporal modeling series consists of three interactive phases: model identification, parameter estimation, and diagnostic and application testing [35]. The concepts of each step are presented in the literature according to [31,35,44,45].

In [46] there are listed several principles and criteria, which characterize a good model. They are described as follows: (1) Thrift, means that the model chosen should be as easy as possible, among others. (2) Identifiability, which allows satisfactory interpretation of the model. (3) Consistency with the data refers to the good fit of the model in relation to the data. (4) Consistent with the theory means harmony of claims with related economic theory or common sense. (5) Admissibility of data means that the model should not predict values that do not meet some final constraints. (6) Forecast Success refers to the precision of prediction obtained by the model.

2.14. Neural Networks

Recently, ANNs have been extensively studied and used in the prediction of time series. According to [47], a neural network is a massively parallel distributed processor made up of simple processing units (neurons), which have the natural propensity to store experimental knowledge and make it available for use. It resembles the brain in two respects: (1) knowledge is acquired by the network from its environment through a learning process; and (2) connecting forces between neurons, known as synaptic weights, are used to store the acquired knowledge.

According to [48], the arrangement of layered neurons and the pattern of binding between layers are called neural network architecture. The network architecture determines the number of connection weights and how the input signals are processed in the network. A neuron is an information processing unit that is fundamental to the functioning of a neural network. To achieve good performance, neural networks employ a massive interconnection of simple computational cells called "neurons" or "processing units".

The procedure used to carry out the learning process is called a learning algorithm; its function is to modify the synaptic weights of the network in an orderly way to achieve a desired design goal [47].

Another important parameter in the design of ANNs is the definition of the architecture of an ANN, since it restricts the type of problem that can be handled by the network. The architecture of an ANN consists of the way the neurons are structured and their connections, that is, the number of layers of the network, the number of neurons in each layer, the type of connection between the neurons [49].

2.15. Use of Wavelets

Nowadays, there are been used some hybrid models that combine the models related before with the use of wavelets. In [50] a comparison between two approaches, wavelet-ARIMA and wavelet-ANN models for temperature time series is made. The main conclusion of this paper is that the wavelet-ARIMA model is more effective than the wavelet-ANN model.

In [51], it is proposed a wavelet recurrent neural network with semi-parametric input data pre-processing for micro-wind power forecasting in integrated generation systems. In this paper, it is presented an improved micro wind output generation forecasting method by using a Wavelet Recurrent Neural Network predictor. Inputs speed data are pre-processed by a semi parametric based approach. The effectiveness of the proposed method is demonstrated through a wind power forecasting along an entire year.

There are still many discussions about the use of wavelets for analyzing data or for use them in forecasting combined whit other models such as ARIMA or Neural Networks. In [52], an analysis of the use of Wavelets for Time Series Forecasting is made. Authors arrived to important conclusions: they affirm that, for time series with a strong random component, wavelets generate only little improvements; and that, if the long-term structure is more important than the short-term oscillation, a denoising step plus ARIMA forecasting is the method of choice.

3. Materials and Methods

3.1. Database

The historical series of the meteorological variables used was obtained from national organization system of environmental data (SONDA) of the National Institute of Space Research (INPE). The series begins on 1 January 2004 and ends on 31 May 2017, this database displays the data per minute of the following variables: Air temperature, Air humidity, Atmospheric pressure, Average wind speed, Wind direction. The database provided by SONDA, can be accessed in the URL http://sonda.ccst.inpe.br/basedados/.

The complementary data were made available directly with station technicians, since the update up to the date studied is not available online. The site decision was made based on the Brazilian wind atlas [53], and by the analysis of the average speeds presented in the database. Figure 2 gives an example of the behavior of the time series wind speed of the database.

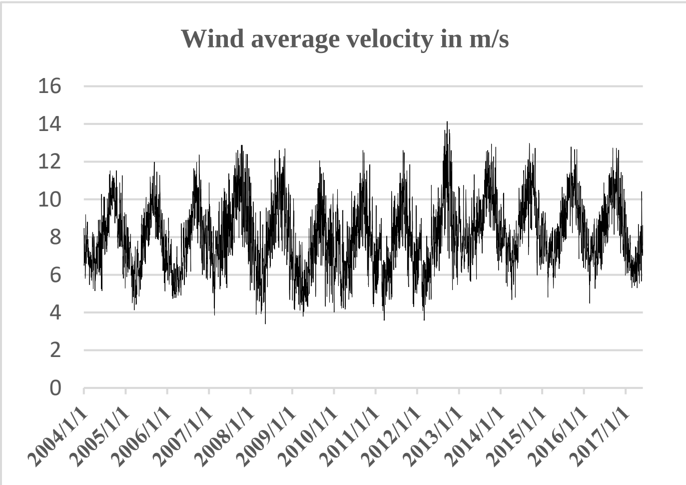


Figure 2. Time series components of the wind speed of the database in days. Source: http://sonda.ccst.inpe.br/basedados/.

In this paper, the results obtained for wind speed prediction will be presented using five models: ARIMA, ARIMA + WAVELET, ARIMA + NEURAL NETWORKS 1, ARIMA + NEURAL NETWORKS 1 + NEURAL NETWORKS 2 and NEURAL NETWORKS.

To allow the comparison of the different forecasting horizons, the application of the models follow an equal configuration for all horizons, so it is possible to verify which horizon has the best answer for the proposed model. The configurations will only be applied in the first step of the Model ARIMA. The applications of the neural networks in both the Hybrid model and in the comparative using only Neural Networks will follow standardization as demonstrated after the ARIMA model. The forecasts follow the pattern of Table 1.

Table 1. Pattern of forecast horizons.

Acronym	Forecasting	Magnitude	Data Amount	Base
UCP	Ultra-Short Term	Minutes	7200	5 days
CP	Short Term	Hours	8760	1 year
MP	Medium Term	Days	8736	13 years
		Weeks	1248	13 years
LP	Long Term	Months	312	13 years
		Years	13 *	13 years

* When forecasting the year horizon, due to the small amount of data available, it was necessary to use the horizon base in months, in which the technique of interpolation of the base results in months was applied, transforming them for years. In addition, on the horizon in years, the data moving average extrapolation technique was used five years ahead to allow the comparison of the expected results with real ones.

3.2. ARIMA Model

The first step of the proposed algorithm illustrated in Figure 3 is the ARIMA model (Auto Integrated Regressive of Moving Average), which results from the combination of three filters: the AR component, the Integration filter (I), and the MA component. The representation of this model is done through ARIMA notation of order (p, d, q). An ARIMA (1, 2, 0) representation indicates an order 1 for the AR (Self-Regressive) component, order 2 for component I (Integration or differentiation) and the last 0 for the MA, where: p is the number of seasonal auto-regressive terms; d is the number of seasonal differences; and q is the number of seasonal media moving.

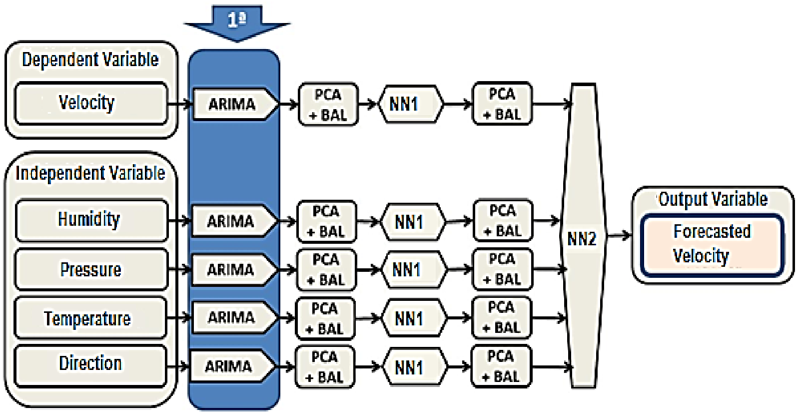


Figure 3. Block diagram of the first stage of the model—ARIMA.

Speed is classified as a dependent variable because its predicted value in ARIMA + NN1 + NN2 depends on the values of the variables humidity, pressure, temperature and direction.

ARIMA models (p, d, q) obtained for each of the time series are presented, respectively, in each application step according to Table 2.

Table 2. ARIMA model for each horizon.

Forecasting	Magnitude	ARIMA MODEL (p, d, q)
Ultra-Short Term	Minutes	(0, 1, 1) (0, 0, 0)
Short Term	Hours	(2, 0, 0) (2, 0, 1)
Medium Term	Days	(2, 0, 2) (1, 0, 1)
	Weeks	(0, 1, 1) (1, 0, 2)
Long Term	Months	(1, 0, 1) (1, 0, 0)
	Years	(0, 2, 1) (0, 1, 2)

Source: Generated by statistical package for the social sciences (SPSS) modeler.

In this step, the speed dependent variable and the independent variables described above, which are the Humidity, Pressure, Temperature and Direction, generate independent results one of each other. Independence is processed through the program input filter, the time interval of every database used was from 1 January 2004 to 31 May 2017; the estimation is based on the beginning of the data; and the forecast follows the multi pass proposal. First, the response analysis of the models for Step 1 (Minutes, Hours, Days, Weeks, Months, and Years) of each variable was performed. The reliability used was 95% and the delay numbers were 180 for minutes, 72 for hours, 21 for days, 12 for weeks, 38 months and 38 for years. Through extrapolation in the data, these delays represent three cycles of each representation, in order to test the result and then compare it with the variations of the different steps. The initial values found with the ARIMA model represented are the Minimum Error, Maximum Error, Mean Error, Standard Deviation and Linear Correlation. The mean speed of the wind (VMED), mean absolute error (MAE), root mean square error (RMSE), and mean absolute percent error (MAPE) were used to evaluate the prediction accuracy of the variable speed. In [54], these items are detailed. VMED is the calculation of the average wind speed:

$$\text{VMED} = \frac{1}{N}\sum_{i=1}^{N} V_i^{real} \tag{1}$$

where N is the number of samples, and V_i^{real} is the actual value of the Speed.

MAE expresses accuracy in the same data units, which helps to conceptualize the magnitude of the error. The equation is:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} \left| V_i^{real} - V_i^{prev} \right| \tag{2}$$

where, N is the number of samples, V_i^{real} is the actual value of Speed and V_i^{prev} is the speed predicted.

RMSE is a commonly used measurement of accuracy of time series values.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left(V_i^{real} - V_i^{prev} \right)^2} \tag{3}$$

MAPE expresses accuracy as a percentage of the error. Because this number is a percentage, it may be easier to understand than other statistics. For example, if the MAPE is 7, on average, the forecast is incorrect at 7%. The equation is:

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N} \frac{\left| V_i^{real} - V_i^{prev} \right|}{V_{med}} \times 100\% \tag{4}$$

Values by themselves cannot demonstrate positive or negative results these values are basis for model comparisons.

3.3. ARIMA + NN1 Model

The second step of the model illustrated with the block diagram in Figure 4 is performed in the first Neural Network—NN1. This step is done to predict explanatory variables and it uses ARIMA results as input variables, which are reduced through the principal component analysis (PCA), which finds linear combinations of the input fields reducing the components for using the main variables.

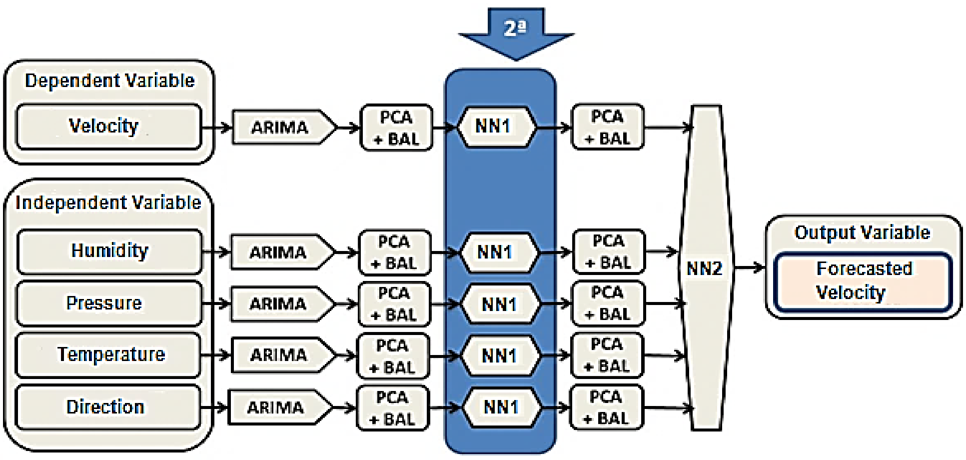


Figure 4. Block diagram of the second stage of the model—ARIMA + NN1.

The NN1 presents eight neurons in the input layers, two neurons in the hidden layer and one neuron in the output layer, configuration used for all variables. The back propagation error training algorithm was used, which adjusts the network weights in order to minimize the error between the actual values and the predicted outputs.

Data partitioning was 80% for training and 20% for testing. The stopping criterion used is the maximum training time per model. The network was trained with sigmoidal tangent activation function for all neurons.

The network follows a standardized programming with multilayer perceptron (MLP) with the topology 8-2-1, logistic activation function and back propagation algorithm. The network uses the data for each respective horizon, i.e., 180 for minutes, 72 for hours, 21 for days, 12 for weeks, 38 months, and 38 for years by applying the extrapolation in the data. These delays signify three cycles of each representation, and project one step forward, the recursive network re-inserts each projection at the input of the MLP and does that repetition automatically 20 times. In this step, the ARIMA + NN1 speed prediction is made and all values are analyzed based on errors, standard deviation and linear correlation.

3.4. ARIMA + NN1 + NN2 Model

The final step of the algorithm is represented by the block diagram in Figure 5. In this step, the final speed prediction is made.

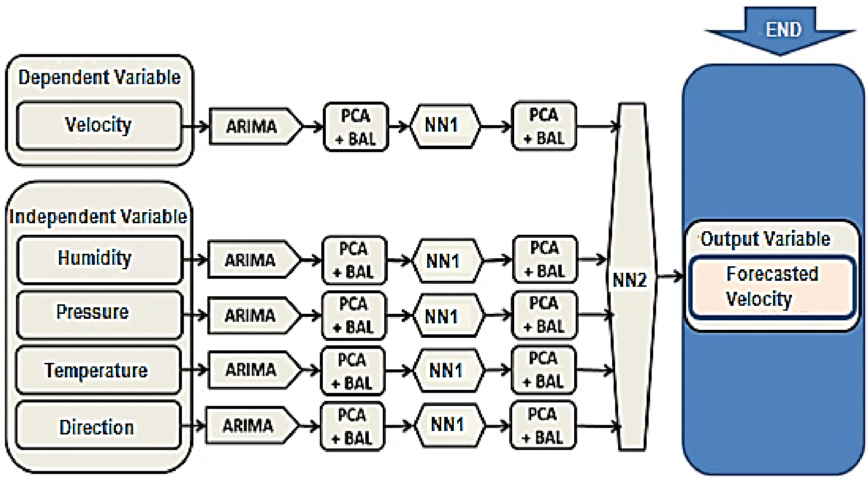


Figure 5. Block diagram of the final step of the model—ARIMA + NN1 + NN2.

The NN2 uses the outputs of the ARIMA + NN1 model as input to optimize the results, which adjusts the weights of the neural network in for minimizing the error between the actual values and the predicted outputs. Data partitioning was 80% for training and 20% for testing.

The network follows a standardized programming with MLP with the topology of 11 neurons in the input layer, eight neurons in the hidden layer and one neuron in the output layer, logistic activation function and back propagation algorithm. It uses values for each respective horizon; that is, 180 for minutes, 72 for hours, 21 for days, 12 for weeks, 38 months, and 38 for years by applying extrapolation to the data. These delays signify three cycles of each representation. They project one-step forward; the recursive network re-inserts each projection at the input of the MLP and do this repetition automatically 20 times.

The stopping criterion is the maximum training time per model. The network was trained with sigmoidal tangent activation function for all neurons.

3.5. Neural Networks Model

The model of neural networks was used to compare the results obtained by the proposed hybrid model; the configuration of the model has as input the environmental data of real values described as input variables.

The network follows a standardized programming with MLP with the topology formed with nine neurons in the input layer, seven neurons in the hidden layer and one neuron in the output layer, logistic activation function and backpropagation algorithm. It uses values for each respective horizon, that is, 180 for minutes, 72 for hours, 21 for days, 12 for weeks, 38 months, and 38 for years by applying extrapolation to the data. These delays represent three cycles of each representation, and project one step forward, the recursive network re-inserts each projection at the input of the MLP and does this repetition automatically 20 times.

3.6. Forecast of Wind Speed and Generated Power

The final objective of this work is to forecast the wind speed to predict the generated power. For a given wind speed, the power generated depends on the type of generator to use. The wind turbine chosen for the study was the WES100 model with 100 kW of power, capable of generating 100 kW at an average wind speed of (17 m/s), Figure 6 shows the power curve of the wind turbine.

Wind speed (m/s)	Power kW
0	0,0
1	0,0
2	0,0
3	1,0
4	2,9
5	6,0
6	11,0
7	17,7
8	27,3
9	39,2
10	53,6
11	68,4
12	82,6
13	89,1
14	95,9
15	98,7
16	99,5
17	100,0
18	100,0
19	100,0
20	100,0
21	100,0
22	100,0
23	100,0
24	100,0

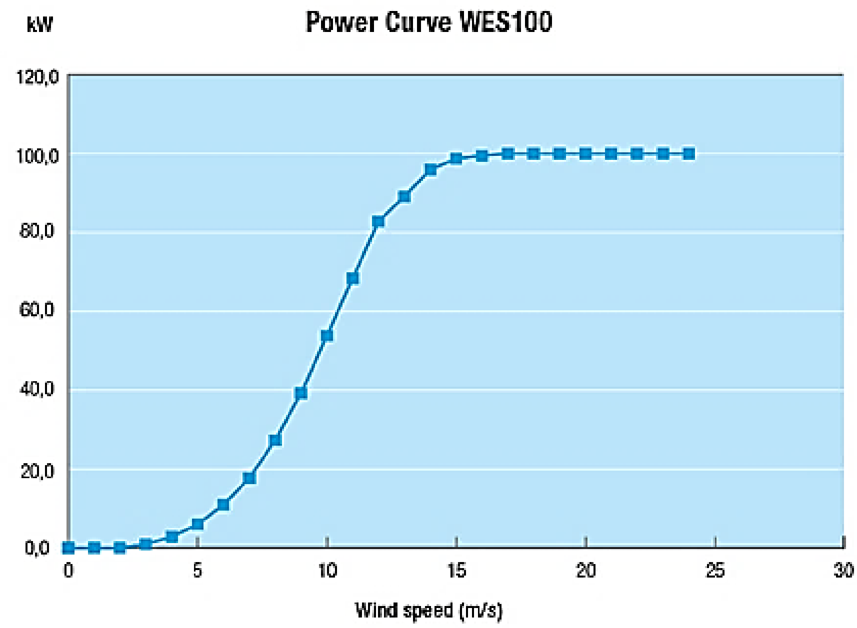


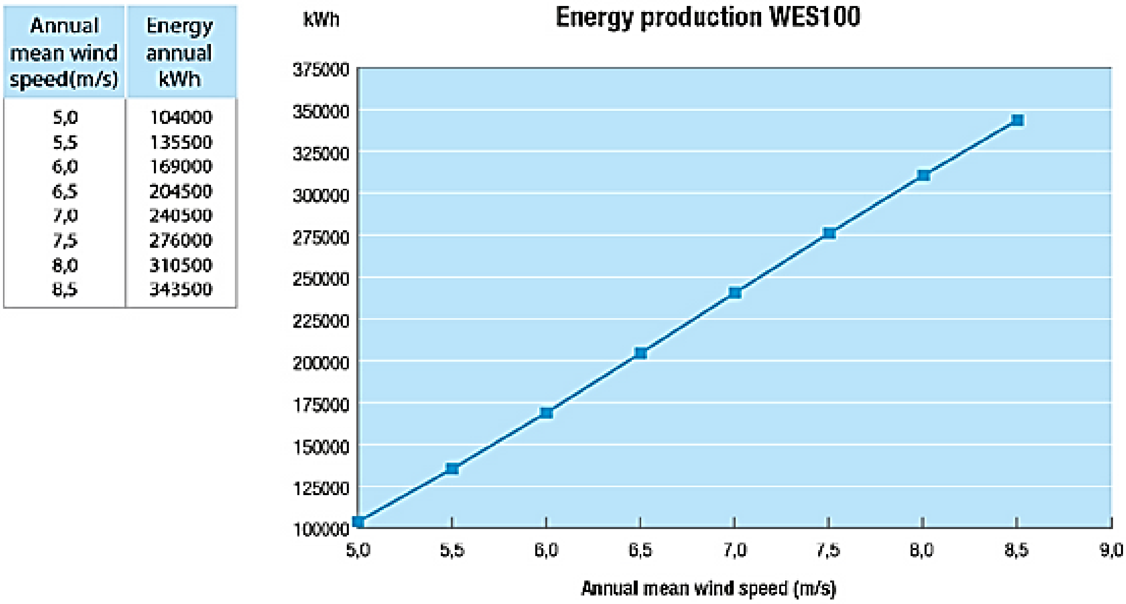
Figure 6. Power curve of the WES100 wind turbine. Source: http://www.energiapura.com/aerogerador-wes-100.

To obtain the generator power curve equation, curve expert software was used from the Power and Speed data. The equation that represents the generation curve is given by:

$$P = a + bx + cx^2 + dx^3 + ex^4 \tag{5}$$

where P = Generated Power; x = Wind Speed; a = 1.515151515153736 $\times$ 10^{-2}; b = $-$6.414141414141929 $\times$ 10^{-2}; c = $-$9.734848484848370 $\times$ 10^{-2}; d = 8.005050505050480 $\times$ 10^{-2}; and e = $-$1.893939393939385 $\times$ 10^{-3}.

Figure 7 shows the graph of annual generation of wind turbine generator wind energy solutions (WES) 100 in kWh.



Annual mean wind speed(m/s)	Energy annual kWh
5,0	104000
5,5	135500
6,0	169000
6,5	204500
7,0	240500
7,5	276000
8,0	310500
8,5	343500

Figure 7. Annual energy generation in Kwh. Source: http://www.energiapura.com/aerogerador-wes-100.

To obtain the energy curve equation provided by the generator, curve expert software was used from the Energy and Speed data. The annual energy equation generated according to the wind speed for this turbine is:

$$E = a + b \times v + c \times v^2 + d + e \tag{6}$$

where E = Generated Energy; v = wind speed; a = 5.738920454544906 $\times$ 10^5; b = $-$3.862506313128678 $\times$ 10^5; c = 9.348295454540566 $\times$ 10^4; d = $-$8.406565656562561 $\times$ 10^3; and e = 2.803030303030391 $\times$ 10^2.

4. Result Analysis

The results obtained will be shown for each forecast universe as described above. In the case of wind speed, there were used symmetric daubechies wavelets combined with ARIMA model for the forecast.

4.1. Ultra Short Term Forecast—CPU (Minutes)

The data used have a base containing 7200 rows, and a total of 36,000 data; such amount is equivalent to a universe of five days.

The multi-step ahead results show that some of the used models have a significant loss of precision, the higher the prediction step, the lower the precision. The ARIMA model already starts with very large errors, the absolute mean error for example in 5-min steps is 0.795 and the MAE percentage is 15.024%. For 20-min steps, the absolute mean error becomes 1.182, and the mean absolute percentage error reaches 20.591%. However, because the ARIMA model returns the prediction values with one-step delay, it will always be worse than the model with neural networks. The ARIMA + NN1 + NN2 model in the step forecast or 5 min, obtained an absolute mean error response of 0.199 and an absolute mean error response of 3.620%. For prediction of 20 min steps, the absolute mean error goes to 0.308 and the mean absolute percentage error goes to 5.305% (see Table 3); this result is superior to the other models, confirming its performance (see Table 3).

Table 3. Results of errors of predicted speed for multi-step forecast ultra-short term (minutes).

Model	Parameter to be Calculated	ARIMA	ARIMA + NN1	NEURAL NETWORKS	ARIMA + NN1 + NN2
Forecasting for 5 min	VMED (m/s)	5.290	5.290	5.555	5.489
	MAE (m/s)	0.795	0.397	0.265	0.199
	RMSE (m/s)	1.777	0.889	0.592	0.444
	MAPE (%)	15.024	7.512	4.769	3.620
Forecasting for 10 min	VMED (m/s)	6.428	6.428	6.317	6.345
	MAE (m/s)	1.132	0.566	0.377	0.283
	RMSE (m/s)	3.579	1.790	1.193	0.895
	MAPE (%)	17.607	8.804	5.972	4.459
Forecasting for 20 min	VMED (m/s)	5.739	5.689	5.843	5.804
	MAE (m/s)	1.182	0.616	0.411	0.308
	RMSE (m/s)	5.285	2.754	1.836	1.377
	MAPE (%)	20.591	10.825	7.027	5.305

The values presented in the table above show the results for each model. The results of ARIMA tend to predicted average based on the delay and the values with intervention of the model of neural networks are based on non-linear characteristics. The actual values were obtained until 31 May 2017, the forecast was made from the Minute 00 of 1 June 2017 extending to the 20th min of the same date; that is, a forecast of 20 step (minutes) forward.

4.2. Forecasted Speed at Ultra-Short Term (Minutes)

Figure 8 shows the results of the speed values in m/s predicted for ultra-short term in minutes.

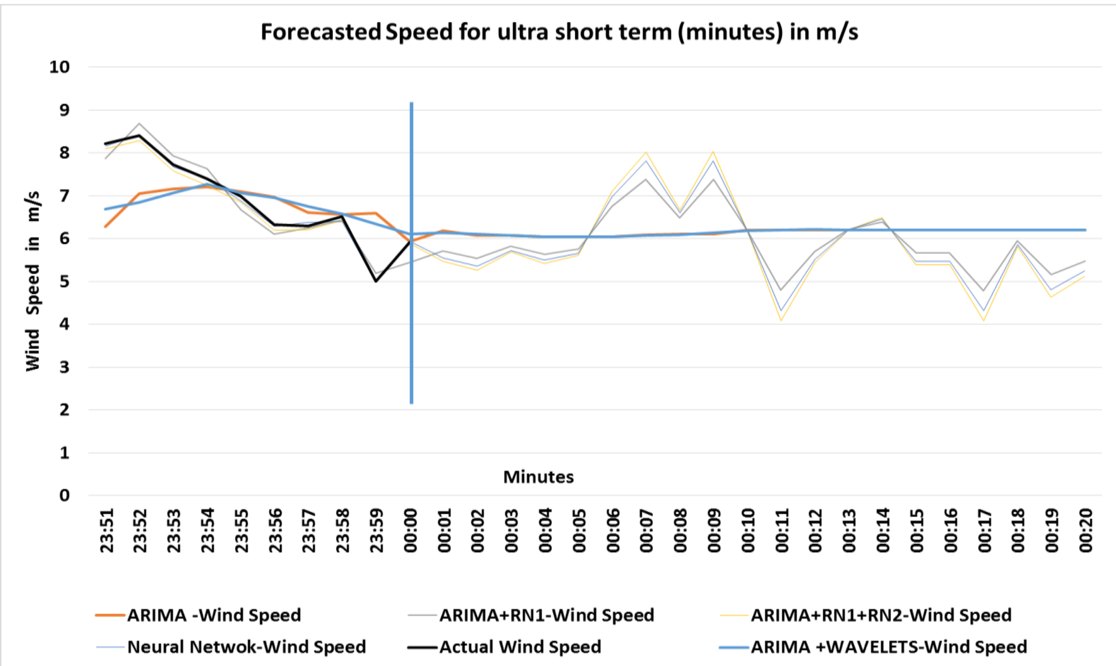


Figure 8. Estimated speed ultra-short term (minutes) in m/s.

In Figure 8 it can be appreciated that there is no great difference between ARIMA model and ARIMA + WAVELETS model, and the latter one will not be included in power and energy forecasting.

4.3. Estimated Power at Ultra-Short-Term (Minutes)

Figure 9 shows the generated power predicted for ultra-short-term (minutes); the results data are in KW.

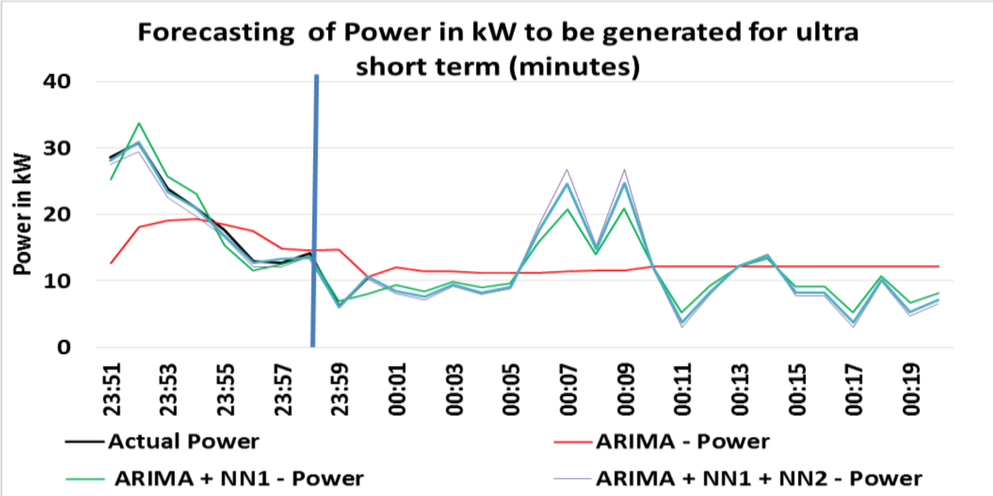


Figure 9. Generated power expected for ultra-short term (minutes) in KW.

The forecast for this horizon takes into account an ultra-short time interval, which produces some abrupt changes in the speed variation. From the interval of 5 min to 7 min, there was an increase of approximately 15 kW according to the forecast ARIMA + NN1 + NN2, and a drop of approximately 20 kW from the range of 9 min to 11 min. After these peaks, forecasting tended to an average of approximately 10 kW and the ARIMA model tended to 12 kW. There were peaks in the prediction of up to 26 kW. The forecast of power generation for this horizon is important for electricity market compensation, real-time network operations and regulatory actions.

4.4. Short Term Forecast—(Hours)

It was also performed the short term forecast that covers the magnitude of time in hours. The data used for this horizon have a base containing 8760 lines, and a total of 43,800 data, these data are equivalent to a universe of 1 year. Table 4 shows the errors in the values of predicted speed for short term forecast.

Table 4. Results of errors for multi-season forecast short-term (hours).

Model		ARIMA	ARIMA+NN1	NEURAL NETWORK	ARIMA + NN1 + NN2
Forecasting for 5 h	VMED (m/s)	7.562	7.562	7.651	7.629
	MAE (m/s)	0.722	0.361	0.241	0.180
	RMSE (m/s)	1.614	0.807	0.538	0.403
	MAPE (%)	9.543	4.772	3.144	2.365
Forecasting for 10 h	VMED (m/s)	7.489	7.489	7.448	7.458
	MAE (m/s)	0.748	0.374	0.249	0.187
	RMSE (m/s)	2.366	1.183	0.789	0.591
	MAPE (%)	9.989	4.994	3.348	2.507
Forecasting for 20 h	VMED (m/s)	7.306	7.256	7.361	7.335
	MAE (m/s)	0.754	0.377	0.251	0.189
	RMSE (m/s)	3.374	1.687	1.125	0.843
	MAPE (%)	10.325	5.198	3.416	2.571

The short-term results show outcomes proportional to the ultra-short term results where the ARIMA model obtained a good result, and the neural network model was superior. The second stage of the ARIMA + NN1 model had a subtle improvement, and the Proposed Hybrid Model ARIMA + NN1 + NN2 is superior to the other models with results. Although the correlation in the Neural Networks model was subtly superior to the proposed Hybrid model, the result of the mean absolute error (MAE) has greater weight in the consideration of the best model. Nevertheless, the RMSE is also important because it shows the average magnitude of the estimated errors, has always positive value and the closer to zero, the higher the quality of the measured or estimated values. The smaller the discrepancy of the data, the better is the result. The ARIMA + NN1 + NN2 model returns a higher accuracy than other ones.

In Table 4, the results of errors in speed prediction for ultra-short term for 5 h, 10 h and 20 h for ARIMA, ARIMA + NN1, ARIMA + NN1 + NN2 and Neural Networks models are presented.

The average speed (VMED) results for the various prediction steps are very close due to the wind behavior which although varying throughout the day, the average takes into account all previous values. For example, the speed for five steps or 5 min in the ARIMA model is 7.562 m/s; this average is the sum of the values divided by the number of samples, in the 20-step or minute forecast the average speed is 7.306.

Although the speed is very close, the ARIMA model has a characteristic that tends to mean moving average of the data, even with this characteristic of the behavior of the average speed and the characteristic of the model, the average percentage error (MAPE) returns a very high value of forecast. In this forecast horizon for short-term (hours) multi-steps ahead, the ARIMA + NN1 + NN2 model in both the absolute mean error, the mean square error, and the mean error, obtained a better result than the other models.

4.5. Wind Speed Short Term Forecasting (Hours)

Figure 10 shows the results of the speed values in m/s predicted for short term forecasting in hours.

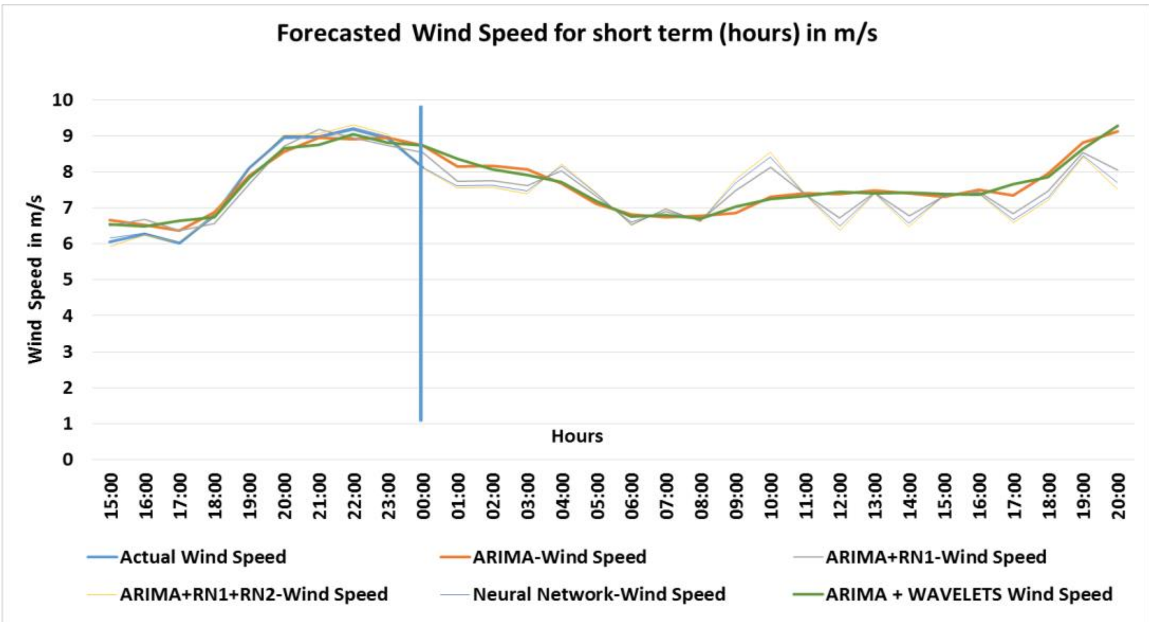


Figure 10. Short term (hours) forecasted Speed in m/s.

In Figure 10, it can be appreciated again that there is not a great difference between ARIMA model and ARIMA + WAVELETS model.

The actual values are until 31 May 2017, the forecast was made from Hour 01 of 1 June 2017 extending up to 20 h at the same date; that is, a forecast of 20 step (hours) forward.

4.6. Short Term (Hours) Forecasted Power in kW

Figure 11 shows the expected generated power for short term (hours); the results are given in kW.

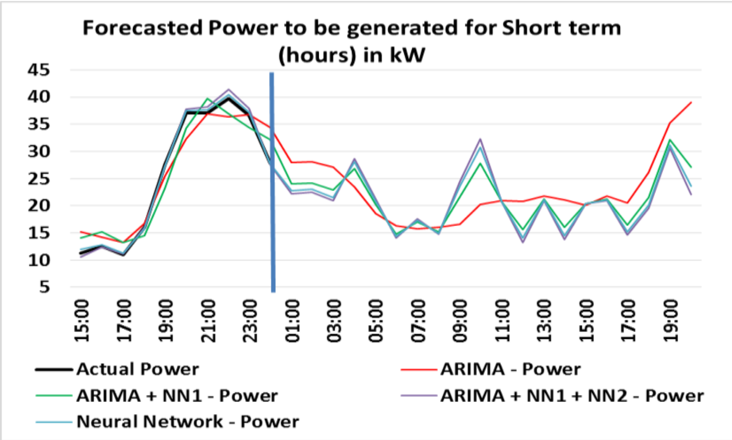


Figure 11. Generated power expected. Short-term (hours) in KW.

The generated power varies according to the expected wind speed, for the selected generator. For this generator can be generated a power of up to 40 kW. This power can be increased taking into account the speed and also adopting a wind generator of more capacity. The variation in the power generation prediction behavior was smoother than the ultra-short prediction, but there was still a variation of approximately 15 kW from the range of 8–10 h and a decrease in the same proportion in the 10 h for 12 h interval

This forecast horizon is important for planning the economic load dispatch, reasonable load decisions and operational safety in the electricity market.

4.7. Medium Term Forecast (Days)

The data used for this horizon, have a base containing 8736 lines, and a total of 43,680 data. These data are equivalent to a universe of 13 years.

All values of errors are considered important in the result analysis. For the medium term, the proposed Hybrid model ARIMA + NN1 + NN2 has a better result than other models. In Table 5, the results of the errors for Medium Term for 5 days, 10 days and 20 days are presented for ARIMA, ARIMA + NN1, ARIMA + NN1 + NN2 and Neural Networks models related to the speed predicted.

Table 5. Error results for multi-steps forecast medium term (days).

Model		ARIMA	ARIMA + NN1	NEURAL NETWORKS	ARIMA + NN1 + NN2
Forecasting for 5 days	VMED (m/s)	7.054	7.254	7.346	7.323
	MAE (m/s)	0.923	0.362	0.241	0.181
	RMSE (m/s)	2.065	0.809	0.539	0.404
	MAPE (%)	13.091	4.987	3.283	2.470
Forecasting for 10 days	VMED (m/s)	7.223	7.323	7.377	7.363
	MAE (m/s)	1.135	0.518	0.345	0.259
	RMSE (m/s)	3.590	1.637	1.091	0.819
	MAPE (%)	15.718	7.069	4.678	3.515
Forecasting for 20 days	VMED (m/s)	7.674	8.324	8.044	8.114
	MAE (m/s)	1.640	0.735	0.490	0.368
	RMSE (m/s)	7.333	3.289	2.193	1.645
	MAPE (%)	21.366	8.835	6.096	4.532

Results show that even the highest 20 h forecast for the ARIMA + NN1 + NN2 model is superior to the lower ARIMA and ARIMA + NN1 prediction, showing the efficiency of the proposed model.

4.8. Medium Term Forecasted Speed (Days)

Figure 12 shows the results of the speed values in m/s predicted for medium term in days.

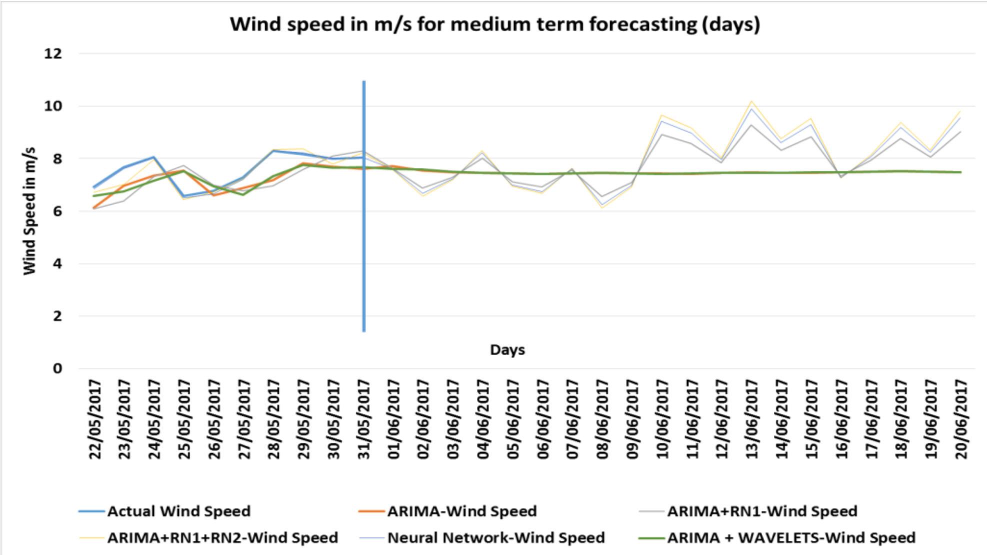


Figure 12. Speed values in m/s predicted for medium term in days.

The actual values are until 31 May 2017, and the forecast was made from 1 June 2017 extending through 20 June 2017 with 20 steps (Days) forward.

4.9. Estimated Power for Medium Term (Days)

Figure 13 shows the expected power generated for the medium term. The results are given in KW.

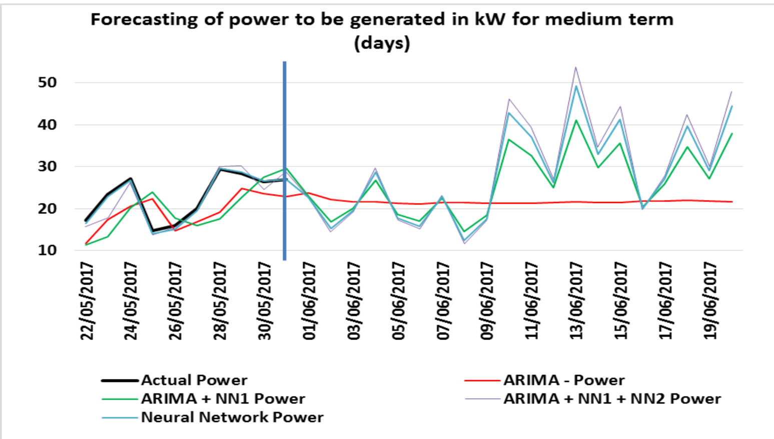


Figure 13. Power Generation Predicted for medium term (days) in kW.

The calculated actual values are until 31 May 2017, and the forecast was made from 1 June 2017 and extended until 20 June 2017 with 20 steps (days) forward. The variation of the generation capacity of the horizon days is greater in relation to the hour horizon. This is due to the behavior of the winds, which although not being part of the analysis of this paper, deserves an observation since the wind repeats its behavior during the cycle of one day, one month and one year, on proportional scales. It is possible to realize that there were generation peaks of more than 50 kW, close to 60% of the maximum capacity of the wind turbine. This generation value could increase if another wind generator of greater capacity would be used. The medium-term forecast horizon is important for unit commitment decisions, reserve commitment decisions and generator online/offline decisions.

4.10. Medium Term Forecast (Weeks)

The fourth forecast horizon was performed over the medium term, which covers the time quantity in weeks. The data used have a base containing 1248 rows, and a total of 6240 data, and such amount is equivalent to a universe of 13 years.

Results for the medium-term forecast show that the ARIMA + NN1 + NN2 hybrid model is superior both in relation to the compared models and in comparison, to the horizons already shown, the linear correlation keeps the result very strong considering the Pearson coefficient. The MAPE obtained shows a better result for this horizon due to the forecasted speed behavior. The mean absolute percentage error. The results of the errors in predicted speed for the medium term for 5 weeks, 10 weeks and 20 weeks for the ARIMA, ARIMA + NN1, ARIMA + NN1 + NN2 and Neural Networks models are presented in Table 6.

Table 6. Error of predicted speed for multi-step forecast—medium term (weeks).

Model		ARIMA	ARIMA + NN1	NEURAL NETWORKS	ARIMA + NN1 + NN2
Forecasting for 5 weeks	VMED (m/s)	7.112	7.112	7.502	7.404
	MAE (m/s)	1.169	0.585	0.390	0.292
	RMSE (m/s)	2.615	1.307	0.872	0.654
	MAPE (%)	16.443	8.221	5.196	3.948
Forecasting for 10 weeks	VMED (m/s)	6.686	6.686	7.218	7.085
	MAE (m/s)	1.595	0.798	0.532	0.399
	RMSE (m/s)	5.045	2.523	1.682	1.261
	MAPE (%)	23.862	11.931	7.368	5.630
Forecasting for 20 weeks	VMED (m/s)	6.598	6.748	7.259	7.131
	MAE (m/s)	1.683	0.827	0.551	0.413
	RMSE (m/s)	7.528	3.697	2.465	1.848
	MAPE (%)	25.514	12.250	7.592	5.796

Although, in the prediction of one step (week), the hybrid model proposed ARIMA + NN1 + NN2 had a loss of precision in relation to the medium term horizon forecast (days), these variations are relative to the statistical behavior of the data.

4.11. Predicted Wind Speed in Medium Term (Weeks)

Figure 14 shows the results of the expected speed values in m/s for medium term in weeks.

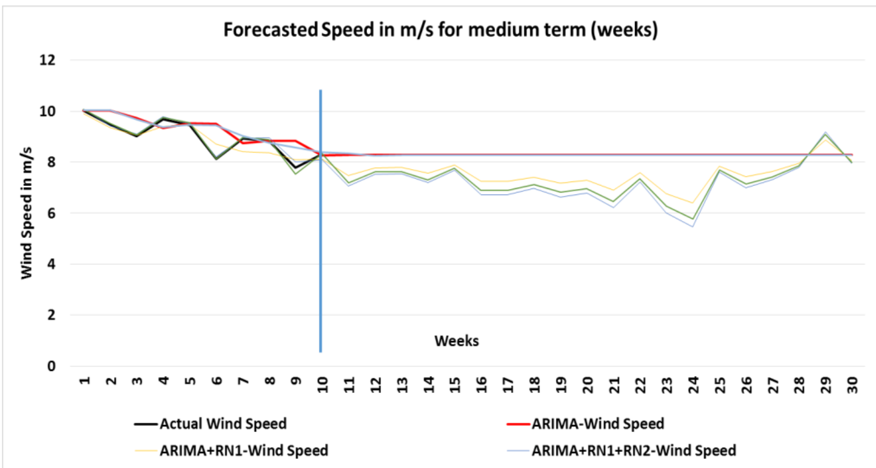


Figure 14. Forecasted Speed in m/s for medium term (weeks).

The actual values are up to 31 May 2017, and the forecast was made from 1 June 2017 extending for 20 weeks, which corresponds to the date of 12 October 2017, with a horizon of 20 step (weeks) forward.

4.12. Estimated Power Medium Term Weeks

Figure 15 shows the expected power generated for medium term (weeks); results are given in KW.

Actual calculated values are until 31 May 2017, and the forecast was made from 1 June 2017, extending it 20 weeks forward. Comparing the generation results with the other horizons, it is possible to notice that the behavior varies according to the winds, even though the variations are below 10 kW, the power generation forecast was maintained at an average near 20 kW and with peaks of 40 kW expected. These values may be higher with wind variation or adopting wind generator with higher generation capacity. The medium-term forecast horizon for weeks is also important for unit commitment decisions, reserve commitment decisions and generator online/offline decisions

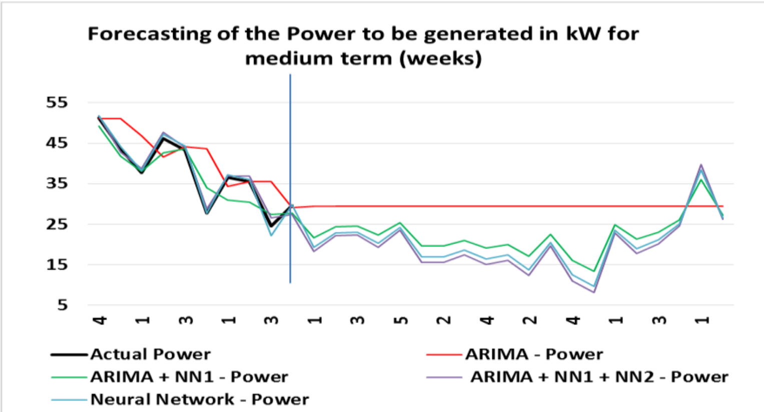


Figure 15. Forecasting of generated power in kW for medium term (weeks).

4.13. Long-Term Forecast (Months)

The fifth forecast horizon was performed in long term, the time quantity in months was used.

The data used have a base containing 312 rows, and a total of 1560 data, which is equivalent to a universe of 13 Years.

The long-term results show an improvement in the results in the proposed ARIMA + NN1 + NN2 hybrid model, which is superior both in relation to the compared models and in comparison, to the previous horizons, the linear correlation has a result considered perfect according to the Pearson's coefficient. These values have higher results than the multi-step forecast, since these represent responses to one-step only. Table 7 presents the results of the long-term errors in predicted speed for 5 months, 10 months and 20 months for the ARIMA, ARIMA + NN1, ARIMA + NN1 + NN2 and Neural Networks models.

Table 7. Errors for multi-step forecast of speed for long-term (months).

Model		ARIMA	ARIMA + NN1	NEURAL NETWORK	ARIMA + NN1 + NN2
Forecasting for 5 months	VMED (m/s)	9.202	9.202	8.923	8.993
	MAE (m/s)	0.838	0.419	0.279	0.209
	RMSE (m/s)	1.873	0.937	0.624	0.468
	MAPE (%)	9.104	4.552	3.130	2.329
Forecasting for 10 months	VMED (m/s)	9.271	9.271	8.870	8.971
	MAE (m/s)	1.202	0.601	0.401	0.300
	RMSE (m/s)	3.800	1.900	1.267	0.950
	MAPE (%)	12.963	6.481	4.516	3.349
Forecasting for 20 months	VMED (m/s)	9.387	9.387	9.003	9.099
	MAE (m/s)	1.455	0.657	0.438	0.328
	RMSE (m/s)	6.506	2.937	1.958	1.468
	MAPE (%)	15.498	6.996	4.863	3.608

The ARIMA + NN1 + NN2 hybrid model maintained the same evolution of the medium term horizon (weeks), which had the best performance in relation to the previous horizons, but had a loss of precision in multi steps.

4.14. Long-Term (Months) Forecast Speed

Figure 16 shows the results of the speed values in m/s predicted for long term in months.

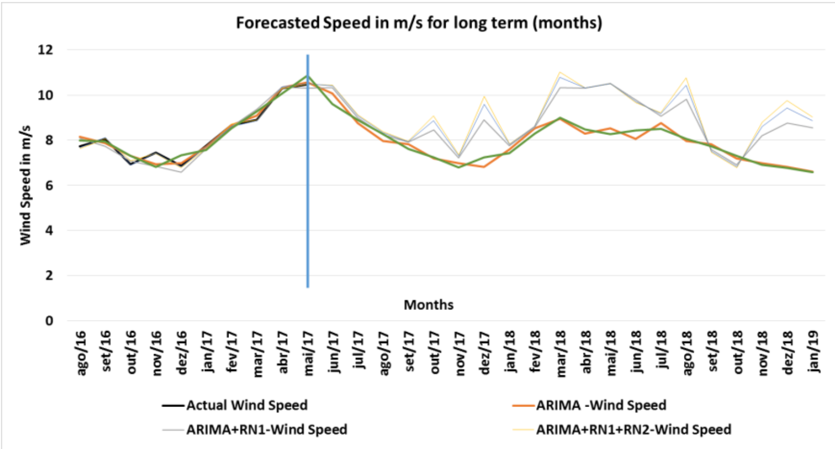


Figure 16. Estimated speed for long-term forecast (months) in m/s.

The actual amounts are up to 31 May 2017, the forecast was made from 1 June 2017 and it was extended for 20 months, covering until January 2019, a horizon of 20 step (months) forward.

4.15. Long-Term (Months) Forecasted Power

Figure 17 shows the generated power expected for long term (months); the results are given in KW.

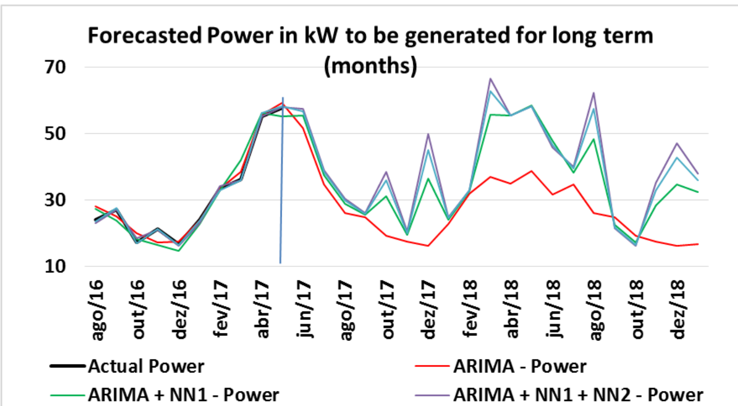


Figure 17. Predicted Power Generated for long Term forecast (months) in KW.

Actual calculated values are until 31 May 2017, the forecast was made from 1 June 2017, and it was extended 20 months forward. In this horizon, the forecast amplitude of the generated power is greater, due to the wind behavior, which varies the complete cycle, in April, May and June. There were obtained peaks of more than 60 kW of predicted power, almost 70% of the total capacity of the wind generator. If another higher capacity generator replaces the generator, the power generation will also be higher. The long-term forecast horizon is important for maintenance planning, operation management, optimum operating cost, and feasibility study of wind farm projects.

4.16. Long-Term Forecast (Years)

The sixth forecast horizon was performed in the long term, the time quantity was used in years.

For the long-term horizon of magnitude in years, because the database has only 13 rows, which is considered insufficient to make a forecast with the models and techniques proposed, the base of months was used, which contains 312 rows, with a total of 1560 data. This amount is equal to a universe of 13 years. After the application of the models was done, the interpolation of the data transforming the results from months to years, and for comparison of the results, the extrapolation was also used through the moving average of the data in years to enable the comparative tests to be carried out.

The results for long term in years have similar behavior to the other horizons keeping the proposed Hybrid Model ARIMA + NN1 + NN2 with superior performance compared to the other models analyzed. Table 8 presents the results of the five-year long-term errors for the ARIMA, ARIMA + NN1, ARIMA + NN1 + NN2 and Neural Networks models. For this horizon, it was only possible to forecast for five-steps (years) due to limited database of 13 years.

Table 8. Error results for multi-steps long term forecast (years).

Model		ARIMA	ARIMA + NN1	NEURAL NETWORK	ARIMA + NN1 + NN2
Forecast for 5 years	VMED (m/s)	8.802	7.874	8.289	8.215
	MAE (m/s)	1.226	0.421	0.294	0.221
	RMSE (m/s)	2.741	0.941	0.658	0.494
	MAPE (%)	13.927	6.378	3.552	2.688

The result for the long-term horizon in years had a great performance even having a very small database; the result was possible using statistical and mathematical techniques of extrapolation and interpolation and due to the consistency of the applied models.

4.17. Long-Term Forecasted Speed

Figure 18 shows the results of the speed values in m/s predicted for Long Term forecast in years.

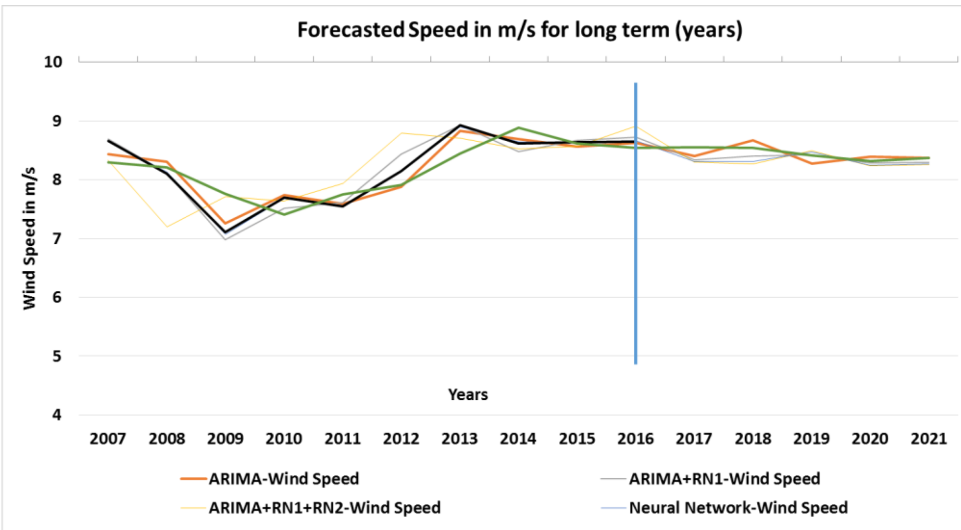


Figure 18. Estimated Speed in m/s for long-term forecast (years). Source: Authors.

The actual values are until 31 December 2016, the forecast was made from 1 January 2017, and it was extended for five years, covering until 2021, which is a horizon of 20 steps (years) forward. The annual speed has a smaller variation than the velocities of the other horizons, since the wind has a variation according to the seasons of the year, but this paper does not intend to treat this subject.

4.18. Long-Term (Years) Power Generation Forecast

Figure 19 shows the power generated predicted for long term (years) forecast; the results are given in kW.

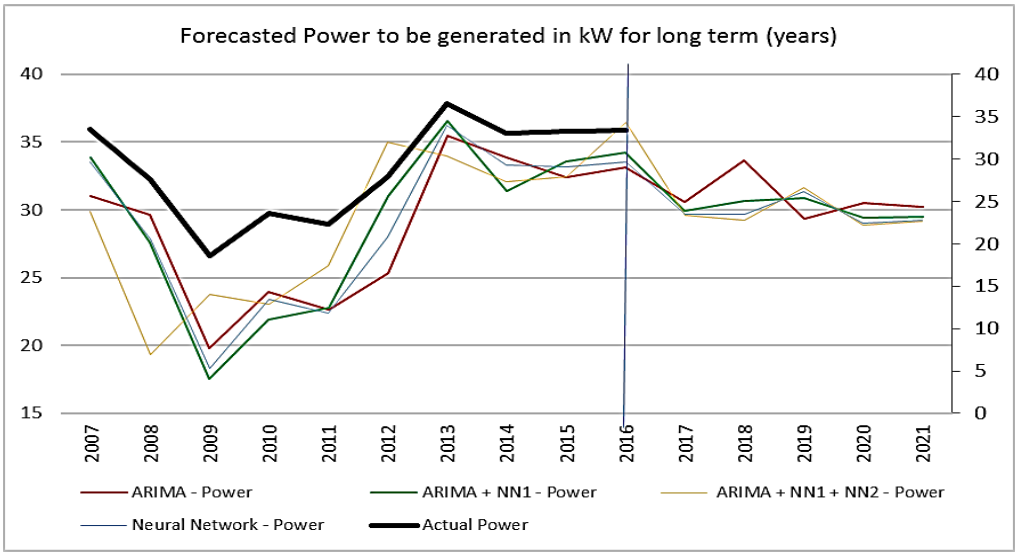


Figure 19. Predicted generated power in kW for long-term forecasting (years). Source: Authors.

The calculated actual values are until 31 December 2016, the forecast was made from 1 January 2017, and it was extended five years forward. This forecast horizon extends for five steps or five years; it is a lower forecast than the previous horizons due to the amount of historical data. However, it is possible to perceive that the amplitude of forecast of generation is greater, due to the behavior of the wind the in the months of April, May and June, with peaks of more than 60 kW of expected power, almost 70% of the total capacity of the wind generator, if the generator is replaced by a larger one, the power generation will also be bigger. The long-term forecast horizon for steps in years is important for maintenance planning, operation management, optimum operating cost and feasibility study for wind farm projects.

4.19. Average Yearly Energy Predicted for Long-Term Forecast (Years)

Figure 20 shows the predicted average annual energy forecast for long-term horizon (years); the results are given in kWh.

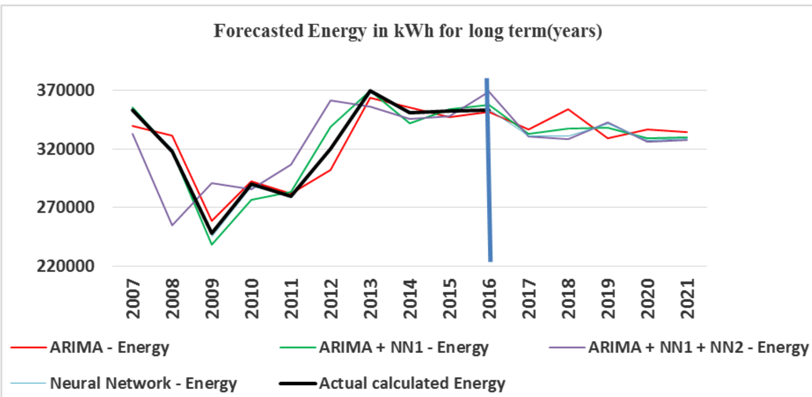


Figure 20. Average annual energy Predicted for Long-term forecast (years) in kWh. Source: Authors.

The energy forecast for the horizon shows a peak of approximately 365,000 kWh for the ARIMA model and 362,000 kWh for the ARIMA + NN1 + NN2 model, the amplitude of the forecast is not very large, and does not have very abrupt forecast variations.

5. Conclusions

- Accurate and reliable wind speed prediction is vital for wind farm planning and operational planning for electrical networks. To improve the accuracy of wind speed prediction, many forecasting approaches have been proposed; however, these models typically do not account for the importance of data pre-processing and are limited by the use of individual models.
- Achieving accurate forecasts of wind speed and power is still a critical problem. Since wind power is proportional to the wind speed cubed, the wind power potential assessment is summarized as wind speed prediction.
- There are many models and their variants for predicting wind speeds, both simple and hybrid, but none of them cover the full range of forecasting possibilities from ultra-short-term forecasts to several years ahead.
- To forecast the wind speed and the possible power to be generated, four prediction models were used:

  ARIMA

  ARIMA + NN1

  ARIMA + NN1 + NN2

  NEURAL NETWORKS

  The four types of forecast were made according to the revised literature:

  Ultra-short-term forecasting

  Short-term forecasting

  Medium-term forecasting

  Long-term forecasting

- Of the models used, the hybrid model of ARIMA + NN1 + NN2 was the one that presented the best results with the smallest errors in the prediction of wind speed in all forecast horizons, as can be seen in the table and graphs presented in this paper. For the prediction for a five-step forward, the best response to the MAE was obtained for the hour horizon with a result of 0.180, and the worst response obtained was for the weeks horizon with a response of 0.292. The best response for the RMSE was obtained for the hour horizon with 0.403, and the worst response was for the weeks horizon with an error of 0.654. For the mean absolute percentage error (MAPE) responses, the best response was obtained for the month's horizon with 2.329%, and the worst response for the weeks horizon with 3.948%. For prediction of 20 steps forward, the best response to absolute mean error (MAE) was obtained for the hour horizon with a result of 0.189, and the worst response was for the week's horizon with an error of 0.413. The best response for RMSE was obtained for the hour horizon with 0.843, and the worst response was for the week's horizon with a response of 1.848. For the mean absolute percentage error (MAPE) responses, the best response was obtained for the hour horizon with 2.571%, and the worst response for the week's horizon with 5.796%.

The use of hybrid models is shown to be more efficient by considering the linear and non-linear characteristics of the modeled signals. The higher the forecasting step, the lower the guarantee of results, but the market needs quick responses and the trend of this response speed feature is increasing in the world scenario. In addition to an efficient model for the need for demand, the database is crucial for results with optimum accuracy. With the prediction of wind speed, it is possible to predict the wind generation of the analyzed region, depending on the generators to be installed. These results of wind speed prediction and therefore of wind power generation potential are unprecedented in the literature, more so, with the combination of models and term predictions, which is undoubtedly a novelty.

The use of wavelets does not improve the wind speed forecasting; ARIMA + WAVELETS forecasting results are very closed to ARIMA forecasting ones.

Acknowledgments: The authors gratefully acknowledge the support to this research done by FAPEAM; UFPA; CNPq Productivity of Research Funds Process 301105/2016-2; and ITEGAM.

Author Contributions: All authors contributed to this research. David Barbosa de Alencar conceived designed and performed the experiments; Carolina de Mattos Affonso and Jandecy Cabral Leite suggested the study idea and shared in writing and revising the paper; José Carlos Reston Filho contributed to the development and application of software; Jorge Laureano Moya Rodríguez revised all aspects concerned the forecasting; and David Barbosa de Alencar, Carolina de Mattos Affonso, Roberto Célio Limão de Oliveira, Jorge Laureano Moya Rodríguez, Jandecy Cabral Leite and José Carlos Reston Filho analyzed the data and interpreted the results. All authors revised and approved the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nagababu, G.; Simha, R.R.; Naidu, N.K.; Kachhwaha, S.S.; Savsani, V. Application of OSCAT satellite data for offshore wind power potential assessment of India. *Energy Procedia* **2016**, *90*, 89–98. [CrossRef]
2. Gao, X.; Yang, H.; Lu, L. Study on offshore wind power potential and wind farm optimization in Hong Kong. *Appl. Energy* **2014**, *130*, 519–531. [CrossRef]
3. Adaramola, M.S.; Agelin-Chaab, M.; Paul, S.S. Assessment of wind power generation along the coast of Ghana. *Energy Convers. Manag.* **2014**, *77*, 61–69. [CrossRef]
4. Kaplan, Y.A. Overview of wind energy in the world and assessment of current wind energy policies in Turkey. *Renew. Sustain. Energy Rev.* **2015**, *43*, 562–568. [CrossRef]
5. GWEC. Strong Outlook for Wind Power. 2017. Available online: http://www.gwec.net/global-figures/graphs/ (accessed on 2 June 2017).
6. Manwell, J.F.; McGowan, J.G.; Rogers, A.L. *Wind Energy Explained: Theory, Design and Application*; John Wiley & Sons: Hoboken, NJ, USA, 2010.
7. Dutra, R. *Energia Eólica–Princípios e Tecnologia*; Centro de Referência para Energia Solar e Eólica Sérgio de Salvo Britto: Rio de Janeiro, Brazil, 2011.
8. Fundação Estadual do Meio Ambiente (FEAM). Utilização da Energia Eólica no Estado de Minas Gerais: Aspectos Técnicos E O Meio Ambiente. 2013. Available online: http://www.feam.br/images/stories/materia/comunicado_tecnico2_gemuc.pdf (accessed on 23 February 2017).
9. Wang, C.; Zhang, H.; Fan, W.; Ma, P. A new chaotic time series hybrid prediction method of wind power based on EEMD-SE and full-parameters continued fraction. *Energy* **2017**, *138*, 977–990. [CrossRef]
10. Wang, D.; Lou, H.; Grunder, O.; Lin, Y. Multi-step ahead wind speed forecasting using an improved wavelet neural network combining variational mode decomposition and phase space reconstruction. *Renew. Energy* **2017**, *113*, 1345–1358. [CrossRef]
11. Patel, M.R. *Wind and Solar Power Systems: Design, Analysis, and Operation*; CRC Press: Boca Raton, FL, USA, 2005.
12. Bernstein, J. *Wind Resource Analysis Program 2002*; Minnesota Department of Commerce: Saint Paul, MN, USA, 2002.
13. Thomas, T.; Cheriyan, E.P. Wind energy system for a laboratory scale micro-grid. In Proceedings of the 2012 IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, 1–2 March 2012.
14. Oliveira, J.B.D. Sistemas Inteligentes e Wavelets para Previsão de Vento e Geração eólica. Master's Thesis, Universidade Federal de Pernambuco, Recife, Brazil, 2008.
15. Wang, J.; Heng, J.; Xiao, L.; Wang, C. Research and application of a combined model based on multi-objective optimization for multi-step ahead wind speed forecasting. *Energy* **2017**, *125*, 591–613. [CrossRef]

16. Brusca, S.; Capizzi, G.; Sciuto, G.L.; Susi, G. A new design methodology to predict wind farm energy production by means of a spiking neural network–based system. *Int. J. Numer. Model. Electron. Netw. Devices Fields* **2017**. [CrossRef]
17. Wang, X.; Guo, P.; Huang, X. A review of wind power forecasting models. *Energy Procedia* **2011**, *12*, 770–778. [CrossRef]
18. Dongmei, Z.; Yuchen, Z.; Xu, Z. Research on wind power forecasting in wind farms. In Proceedings of the 2011 IEEE Power Engineering and Automation Conference (PEAM), Wuhan, China, 8–9 September 2011.
19. Zhao, X.; Wang, S.; Li, T. Review of evaluation criteria and main methods of wind power forecasting. *Energy Procedia* **2011**, *12*, 761–769. [CrossRef]
20. Gasch, R.; Twele, J. *Wind Power Plants: Fundamentals, Design, Construction and Operation*; Springer Science & Business Media: Berlin, Germany, 2011.
21. Foley, A.M.; Leahy, P.G.; Marvuglia, A.; McKeogh, S.J. Current methods and advances in forecasting of wind power generation. *Renew. Energy* **2012**, *37*, 1–8. [CrossRef]
22. Lei, M.; Shiyan, L.; Chuanwen, J.; Hongling, L.; Yan, Z. A review on the forecasting of wind speed and generated power. *Renew. Sustain. Energy Rev.* **2009**, *13*, 915–920. [CrossRef]
23. Wu, Y.-K.; Hong, J.-S. A literature review of wind forecasting technology in the world. In Proceedings of the 2007 IEEE Lausanne Power Tech, Lausanne, Switzerland, 1–5 July 2007.
24. Chang, W.-Y. A literature review of wind forecasting methods. *J. Power Energy Eng.* **2014**, *2*, 161–168. [CrossRef]
25. Soman, S.S.; Zareipour, H.; Malik, O.; Mandal, P. A review of wind power and wind speed forecasting methods with different time horizons. In Proceedings of the 2010 North American Power Symposium (NAPS), Arlington, TX, USA, 26–28 September 2010.
26. Firat, U. Wind speed forecasting based on second order blind identification and autoregressive model. In Proceedings of the 2010 Ninth International Conference on Machine Learning and Applications (ICMLA), Washington, DC, USA, 12–14 December 2010.
27. Adhikari, R.; Agrawal, R. *An Introductory Study on Time Series Modeling and Forecasting*; Lambert Academic Publishing: Saarbrücken, Germany, 2013.
28. Raicharoen, T.; Lursinsap, C.; Sanguanbhokai, P. Application of critical support vector machine to time series prediction. In Proceedings of the 2003 International Symposium on Circuits and Systems, Bangkok, Thailand, 25–28 May 2003.
29. Ehlers, R.S. *Análise de Séries Temporais*; Universidade Federal do Paraná: Curitiba, Brazil, 2007.
30. Morettin, P.A.; Toloi, C. *Análise de Séries Temporais*; Editora Edgar Blucher: São Paulo, Brazil, 2006.
31. Gujarati, D.N.; Porter, D.C. *Econometria Básica-5*; Mc Grow Hill: New York, NY, USA, 2011.
32. Souza, F.M. *Modelos Box & Jenkins Aplicados a Previsão de Demanda de Leitos Hospitalares*; Programa de Pós-Graduação em Estatística e Modelagem Quantitativa, Universidade Federal de Santa Maria (UFSM): Santa Maria, Brazil, 2006.
33. Figueredo, C.J. Previsão de Séries Temporais Utilizando a Metodologia Box & Jenkins e Redes Neurais para Inicialização de Planejamento e Controle de Produção. Master's Thesis, Universidade Federal do Paraná, Curitiba, Brazil, 2008.
34. Hipel, K.W.; McLeod, A.I. *Time Series Modelling of Water Resources and Environmental Systems*; Elsevier: Amsterdam, The Netherlands, 1994; Volume 45.
35. Makridakis, S.; Wheelwright, S.C.; Hyndman, R.J. *Forecasting Methods and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
36. Heckman, J.J.; Serletis, A. Introduction to Econometrics with Theory: A Special Issue Honoring William A. Barnett. *Econ. Rev.* **2015**, *34*, 1–5. [CrossRef]
37. Pokorny, M. *An Introduction to Econometrics*; Blackwell: Oxford, UK, 1987.
38. Akaike, H. Fitting autoregressive models for prediction. *Ann. Inst. Stat. Math.* **1969**, *21*, 243–247. [CrossRef]
39. Creal, D.; Koopman, S.J.; Lucas, A. Generalized autoregressive score models with applications. *J. Appl. Econ.* **2013**, *28*, 777–795. [CrossRef]
40. Marcano, J.L.L.; Bell, M.A.; Beex, A.L. Classification of ADHD and non-ADHD using AR models. In Proceedings of the 2016 IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016.

41. Box, G.E.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis: Forecasting and Control*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
42. Chatfield, C. *The Analysis of Time Series: An Introduction*; CRC Press: Boca Raton, FL, USA, 2016.
43. Marelli, D.; You, K.; Fu, M. Identification of ARMA models using intermittent and quantized output observations. *Automatica* **2013**, *49*, 360–369. [CrossRef]
44. Zhang, G.P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* **2003**, *50*, 159–175. [CrossRef]
45. Bigović, M. Demand forecasting within Montenegrin tourism using Box-Jenkins methodology for seasonal ARIMA models. *Tour. Hosp. Manag.* **2012**, *18*, 1–18.
46. Kovačić, Z.J. *Analiza Vremenskih Serija*; Ekonomski Fakultet Beograd: Beograd, Serbia, 1995.
47. Haykin, S. Network, Neural: A comprehensive foundation. *Neural Netw.* **2004**, *2*, 41.
48. Olaofe, Z.; Folly, K. Wind power estimation using recurrent neural network technique. In Proceedings of the 2012 IEEE Power Engineering Society Conference and Exposition in Africa (PowerAfrica), Johannesburg, South Africa, 9–13 July 2012.
49. De Pádua Braga, A.; de Leon Ferreira, A.C.P.; Ludermir, T.B. *Redes Neurais Artificiais: Teoria e Aplicações*; LTC Editora: Rio de Janeiro, Brazil, 2007.
50. Nury, A.H.; Hasan, K.; Alam, M.J.B. Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh. *J. King Saud Univ. Sci.* **2017**, *29*, 47–61. [CrossRef]
51. Bonanno, F.; Capizzi, G.; Sciuto, L.; Napoli, C. Wavelet recurrent neural network with semi-parametric input data preprocessing for micro-wind power forecasting in integrated generation Systems. In Proceedings of the 2015 International Conference on Clean Electrical Power (ICCEP), Taormina, Italy, 16–18 June 2015.
52. Schlüter, S.; Deuschle, C. *Using Wavelets for Time Series Forecasting: Does It Pay Off?* IWQW Discussion Paper Series; No. 04/2010; University of Erlangen-Nuremberg: Erlangen, Germany, 2010.
53. Atlas Do Potencial Eólico Brasileiro. Ministério de Minas e Energia. 2001. Available online: http://www.cresesb.cepel.br/index.php?link=/atlas_eolico_brasil/atlas.htm (accessed on 23 February 2017).
54. Wu, L.; Shahidehpour, M. A hybrid model for day-ahead price forecasting. *IEEE Trans. Power Syst.* **2010**, *25*, 1519–1530.