ISSN: 2230-9926

Available online at http://www.journalijdr.com

IJDR

International Journal of Development Research
Vol. 10, Issue, 09, pp. 40336-40343, September, 2020
https://doi.org/10.37118/ijdr.19845.09.2020

RESEARCH ARTICLE **OPEN ACCESS**

NONLINEAR AUTOREGRESSIVE NEURAL NETWORKS FOR FORECASTING WIND SPEED TIME SERIES

***Emanuel Abdalla Pinheiro, João Bosco Verçosa Leal Júnior, Lutero Carmo de Lima and Francisco Geraldo de Melo Pinheiro**

Course of Academic Masters in Applied Physical Sciences, State University of Ceará, Fortaleza, Ceará, Brazil

ARTICLE INFO

Article History:
Received 14th June 2020
Received in revised form
28th July 2020
Accepted 11th August 2020
Published online 30th September 2020

Key Words:
Wind electric generation, Forecasting methods, Wind speed, NARNET, Accuracy indexes.

****Corresponding author:***
Emanuel Abdalla Pinheiro

ABSTRACT

Increasingly, wind electric generation is becoming indispensable to complement the global energy matrix. In this way, it is necessary to develop forecasting methods for this type of renewable resource, in order to guarantee the appropriate use to be compatible with the needs of the electricity sector. In this paper, two hybrid models are proposed that combine Autoregressive Integrated Moving Average methodologies (ARIMA) with that of artificial neural networks (ANNs) to forecast time series of daily average wind speeds from data collected at 10 meters. The first model is a nonlinear autoregressive neural network, NARNET, of order (3,3,1) and the second a NARNET of order (4,3,1). The training of the models occurred with data collected from two cities in the northeast of Brazil, Acaraú and Guaramiranga, during a period of 100 days and the forecast occurred for a period of 126 consecutive days after the training. The forecast horizon is average, around days for 3 months. Moreover, the choice of the best forecasting model occurred through the analysis of statistical accuracy indices, among which are: the mean absolute error (MAE), the mean square error (MSE), the mean absolute percentage error (MAPE), the model suitability (FIT) and the correlation coefficient (R) and of determination (R^2). The forecast proved to be efficient in the localities of Acaraú with MSE errors of 0,2118 m/s, MAPE 10,75%, R 0,8685 and also for the locality of Guaramiranga MSE 0,2234 m/s, MAPE 12,16% and R 0,7233. In addition, the forecast curves for both locations showed good agreement with the observed values. Finally, the possibility of further improving hybrid modeling is discussed by changing the activation functions of the hidden layer.

Citation: ***Emanuel Abdalla Pinheiro, João Bosco Verçosa Leal Júnior, Lutero Carmo de Lima and Francisco Geraldo de Melo Pinheiro.*** "Nonlinear autoregressive neural networks for forecasting wind speed time series*", International Journal of Development Research*, 10, (09), 40336-40343.

INTRODUCTION

The forecast of wind speeds plays a fundamental role to face the challenges related to wind generation, thus it is extremely important to assist in planning and programming studies for the operation of the generation of the hydrothermal and wind system. According to Oliveira (2008), the efficient forecasting of wind and wind generation can contribute positively in the following ways: facilitating the commercialization in the electric energy market; subsidizing in the solution of the optimization problem of the dispatch of the generation of the hydrothermal and wind system; and providing data for wind farm generation control systems. A study that has become essential for wind generation are the methods of forecasting the wind speed time series. Good wind forecasting techniques ensure reliable measurements of wind generation, which contributes to maintaining the stability of the country's electricity grid.

In this scenario, the propositions of Box and Jenkins (BOX; JENKINS, 1976) for forecasting time series through ARIMA models were an important milestone in predicting linear time series, as they created one of the first algorithms for modeling a time series. More recently, new methodological proposals with artificial intelligence techniques based on artificial neural networks have been used to solve this problem (GOUVEIA, 2011). Recent researches performed by Catalão*et al.* (2011) proposed an approach to artificial neural networks to forecast wind energy in the short term in Portugal. The learning algorithm used was that of backpropagation having a three-layer structure. The MAPE obtained was 7,26%, while the average computational time was less than 5 seconds. The proposed approach performed well, according to the authors. Good achievements were obtained in Silva (2014), by applying ANNs with multilayer Perceptron architecture (MLP) in the forecast of wind speeds in Paracuru and Camocim localities, during a period of 630 days, modifying the activation functions of the hidden layers and obtaining values of correlation coefficients in the order of 0,95 and values of determination coefficient of 91% of the variability of the observed data and relative error values in the order of 0,07% for Paracuru locality. Sampaio (2020), hybrid models of ANNs with ARIMA models show efficiency rates of 84% for the monthly series and 82% for the hourly forecast series, for data collected at a height of 10 m in the city of Fortaleza. Camelo *et al*. (2017) uses hybrid ARIMA and ANN modeling, achieving significant reductions in error rates in the monthly wind forecast in the cities of Fortaleza, São Luís and Natal. Subsequently, Leal Junior *et al*. (2018) proposed a hybrid ANN and Holt-Winters methodology with excellent outcomes with a percentage error of approximately 5,0%, and also with the Nash-Sutcliffe efficiency coefficient of approximately 0,96, in the same locations. The main objective of this article is to develop a methodology of wind predictability, based on hybrid modeling combining ARIMA model with ANN, specifically, with multilayer Perceptron architecture to be applied in daily wind speed series for the medium term horizon. This model can help in the wind generation sector, for instance, acquiring important information on how the local wind potential can be exploited for energy generation, since it will be possible to make projections of the intensity of the wind speed, and in this way, plan how much electricity can be generated and able to meet the demand of the electrical system.

MATERIALS AND METHODS

Collected data and study region: In the development of this research, the study regions are located in the state of Ceará in the Northeast of Brazil, in the cities of Guaramiranga and Acaraú. The data were collected from meteorological towers, Figure 1, installed at the Guaramiranga-A314 meteorological station, which is located at the point with the geographic coordinates: 3.121067° S and 40.087288° W, located at an altitude of 866 meters in relation to the level of the sea, and at the Acaraú-A360 weather station, which is located at the point with the geographical coordinates: 4.261351° S and 38.931068° W, located at an altitude of 67 meters above sea level. For each of the locations, the initial time series consists of 5424 raw data collected relating to wind speed over the period of 226 days, starting on June 30, 2019 until February 10, 2020. Daily, 24 measurements were taken every hour, always starting at 12 am and ending at 11 pm on the same day. Measurements were taken at a height of 10m in relation to the ground. These data have units of meter per second (m/s). Regarding the treatment of these data, the following methodology was used: the arithmetic mean was performed with the data collected from wind speeds during 24 hours on 226 days, giving rise to a new daily time series of mean wind speed which will be analyzed in this job.

The three steps below were followed to apply those data on the models proposed in this analysis, in order to make predictions of daily averages of wind speed:

(I) adjustment phase - in terms of daily averages, it consists of introducing the observed wind speed data into the forecasting models so that it is possible to provide the respective adjustments using the desired period. This phase can also be named the training phase or the identification and estimation phase. The days selected for adjustment were from 06/30/2019 to 10/07/2019, therefore 100 days;

(II) phase of the quality of the adjustments - this investigation is carried out using the methods of accuracy of the models, which will be discussed in the next section. Also known as the validation phase;

(III) forecasting phase – it shows the models ability to make projections of the wind speed for the period following the observed data, through the best models adjusted in terms of daily averages for the months following those of the initial adjustments. Phase (II) and (III) can be performed simultaneously, generating a single forecasting phase. The forecast days were from 10/08/2019 to 02/10/2020, therefore 126 days.

Figure 1. Standard weather station

Linear model - ARIMA: The linear model named autoregressive integrated moving average (ARIMA) was developed by Box and Jenkins in 1970 to predict time series. Valenzuela *et al*. (2008) affirm that the application of this requires the assumption that the variables have a linear auto-dependence relationship. The forecast of a time series is simply the establishment of the future values of the series. A forecast is a quantitative estimate, or set of estimates, about the likelihood of future events based on current and past information (SOUZA, 2004). Stationary linear processes are those that show constant mean over time or even varying around the mean (GUJARATI, 2010). For these processes it is possible to use basically three types of ARIMA models: autoregressive of order (p), denoted by AR(p), represented by equation 1, Moving Averages of order (q), denoted by MA(q), defined by equation 2, and Mixed ARMA models (p,q), equation 3.

$$Y_t = -a_1 Y_{t-1} - a_2 Y_{t-2} - \ldots\ldots - a_p Y_{t-p} + E_t \quad (1)$$

$$E_t = -c_1 E_{t-1} - \ldots - c_q E_{t-q} + Y_t \quad (2)$$

$$Y_t + a_1 Y_{t-1} + \ldots + a_p Y_{t-p} = E_t + c_1 E_{t-1} + \ldots + c_q E_{t-q} \quad (3)$$

where $a_i (i = 1, 2, 3, \ldots, p)$ and $c_j (j = 1, 2, 3, \ldots, q)$ represent the parameters of the model and E_t the random noise assumed independently and identically distributed with mean zero and constant variance σ^2. The observation at time *t* is represented by Y_t.

The application of these models is based on an iterative cycle, in which the choice of the order of the parameters (p,q) is performed based on the data itself. This whole process occurs from the following steps (WERNER, 2003):

1. Identification: consists of discovering which of the various versions of the Box-Jenkins models describes the behavior of the series;
2. Estimation: consists of estimating the parameters a_i of the autoregressive component, the parameters c_j of the moving averages component and the variance of E_t;
3. Verification: consists of assessing whether the estimated model is adequate to describe the behavior of the data. If the model is not suitable, the cycle is repeated, returning to the identification phase;
4. Forecasting: When a satisfactory model is obtained, it proceeds to the last stage of the Box-Jenkins methodology, which is the main objective of the methodology: making predictions.

Nonlinear model - Artificial Neural Networks: Artificial neural networks (ANNs) are noteworthy as general nonlinear models with the ability to learn complex patterns present in data sets of certain phenomena such as time series and, for this reason, make a great contribution to the study of time series prediction (HAYKIN, 2001). These networks are composed of processing units, known as neurons. These units are arranged in parallel constituting layers and are interconnected with neighboring layer units by connections associated with weights. By analogy to neurons in the human brain, the connections between units of a neural network are synapses, the weights being named synaptic weights (SILVA, 2014). The model of a neuron is shown in Figure 2, it identifies three basic parts: a set of synapses (or links of connections); a signal adder; and an activation function. This representation of the neuron is known as a single layer network, named Perceptron (MCCULLOCH; PITTS, 1940).

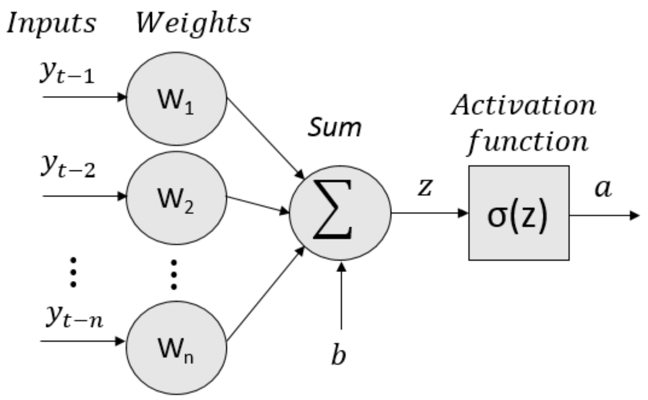

Figure 2. Perceptron model neural network representing *n* regressors at the input

Mathematically:

$$z = \sum_{i=1}^{n} w_i y_{t-i} + b \qquad (4)$$

$$a = \sigma(z) \qquad (5)$$

where $y_{t-i}(i = 1, 2, \ldots, n)$ are the signals received by the perceptron, w are the synaptic weights, $w_i(i = 1, 2, 3, \ldots, n)$ are the artificial equivalent to memory or knowledge of the model, b is a bias value that indicates the neuron's ease of "firing". The output signal z is a linear combiner responsible for unifying the input signals into a single value. Finally, $\sigma(z)$ represents the activation function applied to combination z, effecting the output signal a.

A neural network can be characterized by three main aspects: (1) the pattern of determination between the units (architecture), (2) the method of determining the definition weights (training or learning algorithm) and (3) its function of activation.

Architecture: the architecture of an ANN is how neurons are arranged in the network. The network architecture of the feedforward type is the most common. It is formed by layers of artificial neurons in such a way that the information enters through an input layer and is propagated by the network through intermediate layers always "forward", towards the output layer, having no cycles. Single layer feedforward networks have an input layer and a single neuron layer, which is the output layer itself. The multi-layer feedforward networks are made up of one or more hidden neuron layers, as shown in Figure 3. The prediction of time series is based on using data with time delays for its execution. Thereafter, the input layer is composed of the current data and the delayed values of the time series (Ferraz *et al.*, 2017).

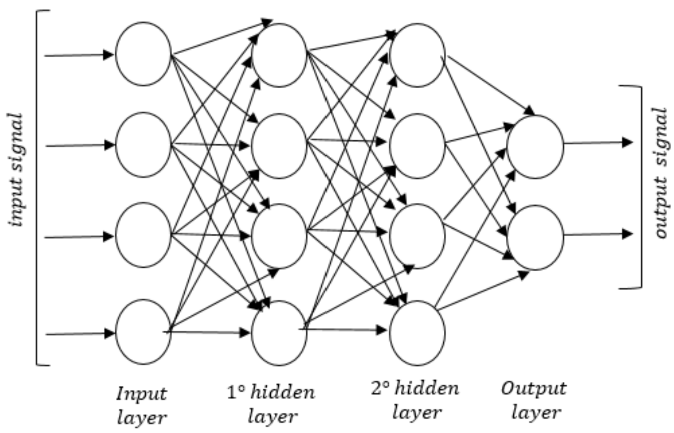

Figure 3. Artificial neural network with MLP architecture with four inputs, two hidden layers with four neurons each and two outputs

Training: The learning, or training, of a neural network is the process by which its free parameters, synaptic weights and polarization values are adjusted through a continuous form of stimulus by the external environment, with the specific type of learning defined due to the particular way in which the adjustments of free parameters occur (BRAGA, 2007). This is a process that takes a certain amount of time, depending on the processing capacity of the machine used. The training can end in three hypotheses: by the number of times, which are the number of iterations of the training, by a minimum acceptable error value or when there is an increase in the cross entropy error during the validation. During the training of the network, it is essential that the data is collected and presented correctly, because in this stage the model will be elaborated. Among the existing algorithms, the one used will be the backpropagation algorithm, which consists of two training phases. The first is the forward phase, in which a data pattern is presented to the network and the network processes the data, producing the output (response) signals. The second of these is the backward phase, which uses the error obtained between the forward phase response and the desired (known) result, to determine the adjustments to be made to the weights of the synaptic connections of the network neurons. This algorithm can be implemented by the Levenberg-Marquardt algorithm, which is an iterative optimization technique used to minimize functions expressed as quadratic sums of nonlinear functions (GOUVEIA, 2011).

Activation function: it aims to simulate the nonlinear functioning of a biological neuron. They are intended to process the signals that reach the neuron coming from synapses, deciding whether the neuron will activate. In case the neuron is not activated, that is, the minimum value is not reached, it should not have an effect on a posterior neuron. (RUFINO, 2014). The following describes some types of activation functions used in the literature for MLP architectures. The activation functions employed in this cases are: linear, equation 6, which are applied to the output layer, sigmoid type logistics, equation 7, and hyperbolic tangent, equation 8. These last two are applied to the hidden layers (SILVA, 2019).

$$\sigma_1(z) = z \tag{6}$$

$$\sigma_2(z) = \frac{1}{1+e^{-z}} \tag{7}$$

$$\sigma_3(z) = \frac{2}{1+e^{-2z}} - 1 \tag{8}$$

Hybrid models: The hybrid model is formed by the combination of two methodologies for solving the problem, thereby, it combines the ARIMA technique with the technique based on ANN. The ARIMA methodology adjusts the linear part of the model, while the artificial neural network is suitable for the nonlinear part of the model. This methodology will be denominated NARNET, a nonlinear autoregressive neural network, equation 9.

$$NARNET_{hybrid} = ARIMA_{linear} + ANN_{nonlinear} \tag{9}$$

The phases for the calculation of this modeling are divided into two stages, according to the block diagram of Figure 4. In the first phase, the current and past input regressors and previous output data are calculated. In the simplest case, of time series, the regressors are delayed outputs, such as *y(t-1), y(t-2)* to *y(t-n)*, known as standard regressors. By definition, all regressors are entered for both the linear and nonlinear function blocks of the nonlinearity estimator. In the second step, the nonlinearity estimator block maps the regressors to the output model using a combination of linear and nonlinear functions. Usually among the available nonlinearity estimators there are artificial neural networks, wavelet networks among others. For the hybrid model, both nonlinear estimation blocks are required. The computational implementation of the network was performed using MATLAB® software, which allows the construction and training of the hybrid. The parameters of the autoregressive nonlinear model are determined using the *nlarx*routine, like this, *nlarx(data_observed, [p 0 0], neuralnet)* in which the input data are, in the following order: observed data from the series, ARIMA model order and desired neural network. The observed data coincide with the data collected from the daily average wind speeds in the locations already described. The order of the ARIMA model (p,0,0) must be a autoregressive model, AR(p), as it is a simple univariate time series, in which there are no exogenous input variables.

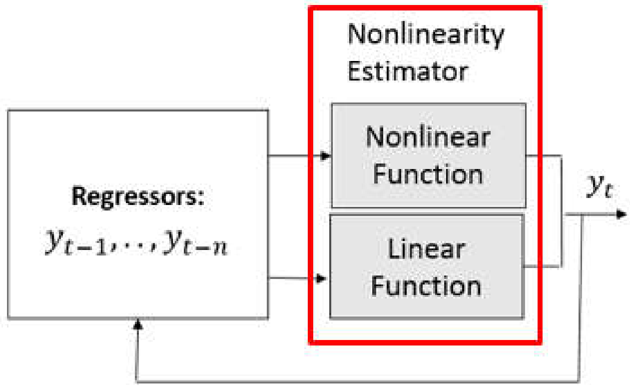

Figure 4. Block diagram represents the structure of the NARNET nonlinear model with a linear function representing the ARIMA model and the nonlinear function representing the ANN-based model

The designed neural network is a feedforward network with multi-layer Perceptron architecture with only one hidden layer and a single neuron in the output layer. The numerical value of p will indicate the number of regressors in the ANN input layer. The modeled NARNET hybrid network is trained, validated and allows new predictions to be made. The hybrid designation is NARNET (*p,o,s*) where *p* represents the number of inputs, *o* represents the number of neurons in the hidden layer and *s*, the number of neurons in the output layer.

Performance indices: Performance indices or measures of accuracy are important statistical analyzes that help in choosing the best model for forecasting the desired time series. Then, the proposed NARNET hybrid model will be subjected to these performance indices for both qualitative and quantitative assessment of the model. In order to be aware of such precision of the forecast made by the hybrid model, HYNDMAN and ATHANASOPOULOS (2018) suggest that the model's validation data are different from the data of its adjustments. Thereby, it is necessary to divide it into two phases before performing the forecast of the model, how it was discussed in previous section. There are several model validation techniques. Among these validation techniques, one can investigate the magnitude of the indices of certain performance indices (BROSILOW; JOSEPH, 2002). To calculate these measures, the following variables will be used: y representing the observed or collected wind speed, the average of y is the variable $\bar{y}$, while the variable evaluated or predicted by the model is $\hat{y}$. The estimated average speed value is represented by $\bar{\hat{y}}$.

Mean Absolute Error: it is the measure of the average magnitude of the forecast errors. It presents a variation between zero and infinite, and its optimal value is zero. It is measured in m/s and given by equation 10:

$$MAE = \frac{1}{n}\sum_{k=1}^{n}|y(k) - \hat{y}(k)| \tag{10}$$

Mean squared error: it represents the sum of the quadratic differences between the observed value and the estimated value divided by the amount of observed values, equation 11. The model that generates the lowest MSE is the closest to the real adjustment and, hence, the most appropriate. It is measured in m^2/s^2.

$$MSE = \frac{1}{n}\sum_{k=1}^{n}(y(k) - \hat{y}(k))^2 \quad (11)$$

Mean absolute percentage error: representation of the absolute mean error in percentage form, equation 12. This measure is important because of the ease of interpreting the percentages of the previous model in the face of a large amount of data (CAMELO *et al*., 2018) Values close to zero indicate good agreement with the observed data. It is defined mathematically by:

$$MAPE = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{y(k)-\hat{y}(k)}{y(k)}\right| \quad (12)$$

FIT:it is also known as quality of adjustment of the normalized mean square error (NRMSE). It represents the percentage of adjustment and varies between -∞, incorrect adjustment, 100%, perfect adjustment, equation 13. If the value is equal to zero, then the model will not be better to adjust the measured data than a straight line equal to average of collected data (MATHWORKS, 2020):

$$FIT = \left(1 - \frac{\|y-\hat{y}\|}{\|y-\bar{y}\|}\right) \times 100\% \quad (13)$$

The || || symbol indicates the norm of the difference between two vectors and corresponds to the Euclidean norm of the values, which can be calculated using equations 14 and 15:

$$\|y - \hat{y}\| = [\textstyle\sum_{k=1}^{n}[y(k) - \hat{y}(k)]^2]^{1/2} \quad (14)$$

$$\|y - \bar{y}\| = [\textstyle\sum_{k=1}^{n}[y(k) - \bar{y}]^2]^{1/2} \quad (15)$$

Correlation coefficient (R): shows the degree of adjustment of the model in relation to the real model, equation 16. It can be defined as a mathematical parameter that enables the degree of comparison of two variables: the observed data and the estimated data. It can assume values that vary between -1 and 1. When R=1, it means a perfect positive correlation between the two variables (maximum values occur at the same time). Though, when R=-1, it means a perfect negative correlation between the two variables, that is, if one increases, the other decreases. If R=0 means that the two variables are not related, further details can be obtained about statistical correlation in Fonseca (1985).

$$R = \frac{\sum_{k=1}^{n}(\hat{y}(k)-\bar{\hat{y}})(y(k)-\bar{y})}{\sqrt{\sum_{k=1}^{n}(\hat{y}(k)-\bar{\hat{y}})^2 \sum_{k=1}^{n}(y(k)-\bar{y})^2}} \quad (16)$$

Determination coefficient (R^2):it is the square of Pearson's correlation coefficient. It describes the proportion of the total variation in the observed data that can be explained by the model, equation 17. It ranges from 0 to 1, with higher values indicating better agreement with the real model (LEGATES, 1999).

$$R^2 = 1 - \frac{\sum_{k=1}^{n}(y(k)-\hat{y}(k))^2}{\sum_{k=1}^{n}(y(k)-\bar{y})^2} \quad (17)$$

Some authors, such as Coelho *et al*. (2015), consider that R^2 values between 0,8 and 1,0 can be considered sufficient for many identification applications. It is worth remembering that the relative evaluation of the model's performance is always important, however these measures should not be used exclusively, as Willmott (1981) argued. Thereafter, it is always appropriate to quantify the error in terms of units of the variable. It is in this context that the importance of the absolute error measures presented as the mean square error, or MSE, and the mean absolute error, or MAE, is inserted.

RESULTS AND DISCUSSIONS

The predicted data obtained took into consideration two hybrid models NARNET with the configurations of Table 1. Both hybrid models were informed to the time series winds speeds obtained in the cities of Acaraú and Guaramiranga, approve in the Material and Methods section. These data went through a testing and validation phase, generating a new time series by hybrid modeling. Finally, the new wind speed time series is used to forecast new values in a forecast range. Model 1, NARNET (3,3,1), is a nonlinear autoregressive neural network in which the order of the autogressive term is 3, AR(3). Therefore, the neural network has 3 regressors in the input layer, 3 neurons in the hidden layer and one neuron in the output layer. While model 2, NARNET (4,3,1), is a nonlinear autoregressive neural network in which the order of the autogressive term is 4, AR(4). The neural network has 4 regressors in the input layer, 3 neurons in the hidden layer and one neuron in the output layer.

In both models, the initialization of the synaptic weights is random and the neuron activation functions in the hidden layer are sigmoidal logistic, whereas in the output layer, the activation function is linear. The number of neurons in the hidden layer is usually half the number of inputs in the network. Nevertheless, better achievements were obtained with the number of inputs being the same number of neurons in the hidden layer. Table 2 shows the accuracy measures in the comparison between two time series, the one observed from the data collected and the one adjusted by the NARNET hybrid model between the periods from June 30, 2019 to February 10, 2020, in the two locations studied. The values of the accuracy indices refer to the complete period, that is, the training phase, together with the forecasting phase, as discussed in the Material and Methods section.

For the locality of Acaraú, the best performance model was NARNET (3,3,1), because the absolute accuracy indices, MAE and MSE, were lower and the relative accuracy indices, MAPE, R and R^2, are more close to the desired values, respectively, 0 and 1 for the correlations. It is worth noting that the differences are small, but significant. The NARNET model (4,3,1) presents a slightly lower MAPE of 10,75% against 10,93% of model 1. In all other criteria, model 1 is better adjusted to the predicted data. The MAE of model 1 is 0,3642 m/s against 0,3678 m/s of model 2. FIT is the best indicator for the adequacy of the estimated series in relation to the observed series. Hence, the FIT of 49,69% shows that model 1 is more adequate than model 2, 48,33%. The correlation coefficients are also very close to 0,8685 of the first against 0,8620 of the second.

Table1. Parameters of the two hybrid NARNET models adopted in wind speed predictions in the locations of Acaraú and Guaramiranga during the study period

Configuration parameters	Model 1 Narnet (3,3,1)	Model 2 Narnet (4,3,1)
Numberof Inputs	3	4
Number of neurons: hidden layer and output	3-1	3-1
Training algorithm	Backpropagation Levenberg-Marquardt	Backpropagation Levenberg-Marquardt
Learning Rate– Prediction Rate	44%-56%	44%-56%
Activation functions hidden layer/output	Logistic - Linear	Logistic – Linear
Synaptic weights initialization	Random	Random
Maximum number of training epocs	300	300
StopCriteria	Increase in crossentropy	Increase in crossentropy

Table 2. Accuracy measures of the predictions of the NARNET (3,3,1) and NARNET (4,3,1) models for the locations of Acaraú city and Guaramiranga city in relation to the collected observations

Performance Indices	Model 1: narnet (3,3,1)		Model 2: narnet (4,3,1)	
	Acaraú	Guaramiranga	Acaraú	Guaramiranga
MAE (m/s)	0,3642	0,3633	0,3678	0,3584
MSE (m^2/s^2)	0,2118	0,2252	0,2234	0,2250
MAPE (%)	10,93	12,65	10,75	12,16
FIT (%)	49,69	30,05	48,33	30,08
R	0,8685	0,7212	0,8620	0,7233
R^2	0,7543	0,5201	0,7430	0,5231

These correlation coefficients are better as they approach the unit, thus the value of 0,8685 is a better indicator than 0,8620. Regarding the location of Guaramiranga, the measurement values are close, with a slightly better advantage for the NARNET model (4,3,1) compared to the NARNET model (3,3,1). The absolute errors 0,3633 m/s of model 1 against 0,3584 m/s of model 2. Quadratic errors practically equal and difference of 0,002 m^2/s^2 between them. The MAPE of the second model, 12,16%, is already better than the first model, 12,65%, however the FIT indices of 30,05% and 30,08% and the correlation coefficients 0,7212 and 0,7233 indicate a slight improvement in the second model compared to the first. Making an analogy with the research developed by Sampaio (2020), the hybrid models tested also consisted of ARIMA and ANN hybrids and present efficiency rates of 84% for monthly series and 82% for hourly forecast series, for data collected at 10 m in height in the city of Fortaleza. These values are very close to the values found in this paper for the case of the locality of Acaraú. The determination coefficients of Acaraú were around 75%. None theless the models used by Sampaio (2020) had many regressors in the input layer and many neurons in the hidden layer, this possibly requires more processing capacity of the software used and more time for the simulations. The hybrids (3,3,1) and (4,3,1) are much simpler and easier to be simulated. The estimates of the daily averages of the time series of Guaramiranga did not show such a promising outcome in relation to the adjustments of the time series of the locality of Acaraú. This fact can be perceived by the coefficient of determination, because while the series in the location of Acaraú obtained values of R^2 above 0,75, for the other location the values are much closer to 0,53. Therefore, the coefficients of determination are closer to 1 for the location of Acaraú. Nevertheless, the values for the second location are still good. Similar studies like Camelo *et al*. (2017) show MAPE values around 10,36% for prediction with models based on a hybrid ARIMA and ANN model with twelve regressors in the input layer, six in the hidden layer with sigmoid activation function and a neuron in the output layer. In this case, the study was based on monthly and annual averages of time series between January 2010 and December 2014 in the locations of Fortaleza, Parnaíba, São Luís and Natal located in the Northeast of Brazil and in this regard they differ from this paper. Comparing with others literatures based on short forecast horizons like Catalão *et al*. (2018) that used neural networks with architecture similar to this study, the MAPE values are close to 7,26%, considering a horizon of 5 seconds. In this sense, there is still a relative proximity of the results, although with works with average forecasting horizons, they consider much more data in the collections, thus they tend to present slightly worse findings in relation to short-term studies, which use a database much smaller data. In the figures below, Figures 5(a), 5(b), 5(c) and 5(d), it is possible to observe the comparisons between the training and forecasting phases of the observed and observed models. The training phase of the first 100 daily averages, corresponding to 06/30/2019 to 10/07/2019, while the forecast predicts 126 subsequent averages, corresponding to 10/08/2019 to 02/10/2020. In Figures 5(a) and 5(b), relative to the daily averages of the city of Acaraú, it is possible to identify that the adjusted time series, black and red lines, can perfectly follow the profile of the observed time series, blue line, in practically the entire period analyzed in particular with many similarities in values, minimum and maximum, indicating the ability of the NARNET hybrids (3,3,1) and (4,3,1) to represent seasonality. Although the graphical representations of these time series are very similar, even in the case of large quantities of measurements, and thereby, make it difficult to visualize differences in their values, it is still possible to identify differences, in particular, between the 100th and the 110th the average speed in the adjusted model is around 4,4 m/s, while the observed data have values close to 5,0 m/s. The adjustment also does not follow the peak of the observed time series that corresponds to the value of 5,4 m/s on day 62 and another very close on day 125. It is important to note that on day 175 there is a sharp reduction in the average speed of the observed series of wind speed and the adjusted series can follow this trend correctly. The interpretations for Guaramiranga, Figures 5(c) and 5(d) are similar to those found for Acaraú, that is, the series adjusted is able to represent particular characteristics of the series observed, for instance, in the item similarity of

Table 3. Summary of the most suitable models found in the research according to the greatest number of criteria analyzed

Cities	Best model	Best indices
Acaraú	Narnet(3,3,1)	MAE MSE FIT R R^2
Guaramiranga	Narnet(4,3,1)	MAE MSE MAPE FIT

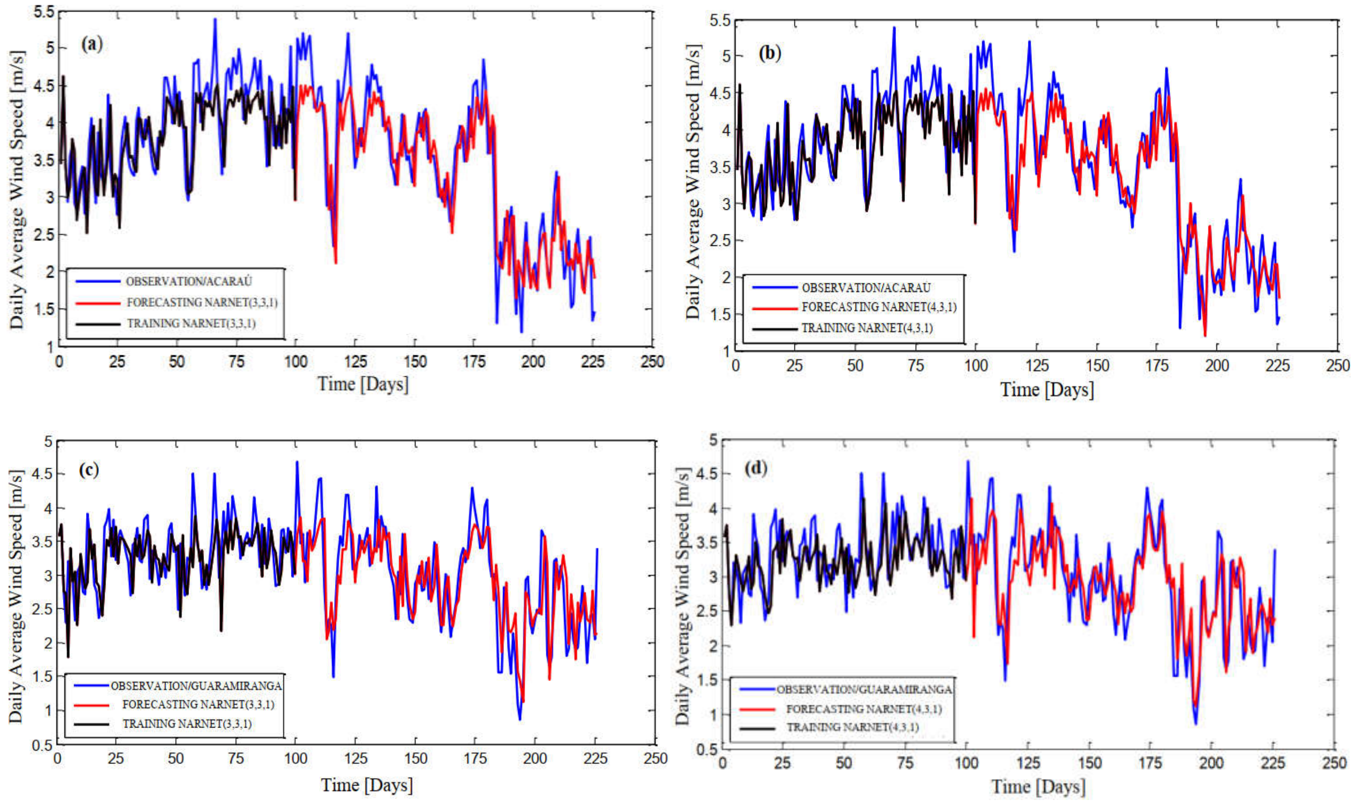

Figure 5. Comparison between the series of data estimated by the hybrid model NARNET and the observed data collected in the same period for a height of 10 meters in the locality of: a) Acaraú Model NARNET (3,3,1); b) Acaraú Model NARNET (4,3,1); c) Guaramiranga Model NARNET (3,3,1); d) Guaramiranga Model NARNET (4,3,1)

maximum and minimum values between these series, and thus, a positive point of the hybrid model NARNET (3,3,1) and (4,3,1) in the representation of seasonality of wind speeds in Guaramiranga. Although the two series, observed and adjusted, have very close values to each other, even so, it is possible to identify differences in the graphical display, like the peak values of the observed series are not predictedby either of the two NARNET models. This occurs on days 60, 70 and 110, on these days the observed values of the average wind speed reach 4,7 m/s, the best forecast of the NARNET model (4,3,1) is 4,3 m/s. Despite this, both models follow the observed series well without leaving great graphic distances from the observed model. This fact is very noticeable in the case of the graphs of Guaramiranga, as there are no holes between the predicted and collected graphs. The minimum values of the observed data are present in the predicted models. Finally, the two models proposed for this second location, by graphic analysis, are very similar, this fact is also possible to observe by the accuracy measures already mentioned. Summarizing the achievements of the work, Table 3 presents the best outcomes for each location according to the largest number of fitness criteria.

Conclusions

The hybrid NARNET models, on which this study is based, provide satisfactory performance, both in the training period and in the forecast period of the model, either due to the high level of agreement between the estimated data series and the observed data series and their correlation coefficients were selected, either by the low values of the quantified errors. The hybrid model NARNET (3,3,1) obtained the best findings for the locality of Acaraú with a correlation coefficient of approximately 0,87 and MSE of 0,21m/s. The second hybrid model NARNET (4,3,1) resulted in better forecasts for the locality of Guaramiranga with MAPE of 12,16% and MSE of 0,225 m/s. The FIT accuracy indexes of 49,69% for model 1 in the locality of Acaraú and 30,08% for model 2 in the locality of Guaramiranga show the difficulty of performing the time series modeling of wind speeds with a forecast horizon medium.

This research can help in several areas of interest in forecasting time series, as is the case of the wind sector, and it is possible to acquire important information about the local wind potential, which can be used to generate electricity, or that is, an attempt to provide guarantees to the decision makers of the sector in the exploration of the winds of a given location. For future studies, two specific guidelines can be followed. The first is to use this NARNET model to forecast time series of wind speeds in the short forecast horizon and verify the outcomes obtained. The second is to verify changes in the activation functions of neurons in the hidden layer in order to improve the efficiency of the neural network in identifying the time series with regard to the nonlinear part of it. Among the options for tests, radial basis functions and wavelet functions can be used, as new works suggest good findings in forecasting with these types of activation functions.

Acknowledgment: The authors are grateful to course of Academic Masters in Applied Physical Sciences from the State University of Ceará (UECE) by opportunity and idealization of this study.

REFERENCES

Box, G.E.P. Jenkins,G.M. 1976. Time Series Analysis: Forecasting and Control. 5th Edition. Wiley, 712 p.

Braga, A. P. Carvalho, A.P.L.F. Ludermir, T.B. 2007. Redes Neurais Artificiais: Teoria e aplicações. 2sted., Rio de Janeiro: LTC.

Brosilow, C. Joseph, B. 2002. Techniques of Model-Based Control.Upper Saddle River, New Jersey: Prentice-Hall.

Camelo, H.N.; Carvalho, P.C.M.; Leal Junior, J.B.V. 2017. Métodos de previsão de séries temporais e modelagem híbrida ambos aplicados em médias mensais de velocidade do vento pra regiões do nordeste do Brasil. Revista Brasileira de meteorologia, v. 32, n. 4, p. 565-574.

Camelo, H.N.; Leal Junior, J.B.V; Carvalho, P.C.M. 2018. Modelagem Box-Jenkins aplicada a previsão de velocidade do vento em regiões do nordeste brasileiro para fins de geração eólica. Revista do Centro de Ciências Natureza e Exatas, v. 40, n. 16.

Carvalheiro, L. C. O. et al. Validação da Temperatura da Superfície do Mar a partir de medições de satélite. Available online at: < https://repositorio.aemet.es/bitstream/20.500.11765/5374/1/6B_Carvalheiro.pdf >.

Catalão, J. P. S.; Pousinho, H. M. I.; Mendes, V. M. F. 2011. Short-term wind power forecasting in Portugal by neural networks and wavelet transform. Renewable Energy, v. 36, n. 4, april, p. 1245-1251.Available online at: < https://doi.org/10.1016/j.renene.2010.09.016>.

Coelho, A. A. R.; Coelho, L. dos S..2015. Identificação de sistemas dinâmicos lineares. 221p. 2st ed. Florianópolis. Ed. da UFSC.

Compare. MATHWORKS, 2020. Available online at: <https://www.mathworks.com/help/ident/ref/compare.html#mw_d5add924-0181-4ed8-bfc6-2b7486bb8f3b>.

Ferraz, R.S.F., da C .Cruz, F., Ferraz, R.S.F., Correia, A.F.S., e de Simas, E.F. 2017. Previsão multi-passos da velocidade do vento através de redes neurais artificiais. The 12th Latin-American Congress on EletricityGeneration and Transmission, CLAGTEE, 2017.

Gouveia, H. T. V. 2011. *Previsão de ventos e geração eólica do sistema ne: analisando diversos sítios e buscando a melhor modelagem através da inteligência artificial*.116f. Dissertação (Mestrado) – Universidade Federal de Pernambuco, Recife.

Gujarati, D.N. 2010. Econometria Básica. 5 st ed. São Paulo: McGraw-Hill.

Haykin, S. S. 2001. Redes neurais: princípios e prática . 2st ed. Porto Alegre, RS: Bookman.

Hyndman, R. J.; Athanasopoulos, G. 2018. Forecasting: principles and practice. 2nd ed. Otexts, 2018. Available online at:< https://otexts.org/fpp2/>.

Leal Junior, J.B.V.; Camelo, H.N.; Carvalho, P.C.M. 2018. Modelo híbrido de previsão de séries temporais para possíveis aplicações no setor de geração eólica. Revista do Centro de Ciências Natureza e Exatas, v. 40, p. 01-06.

Legates, D. R., Davis, R. E. 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Geophys. Water resources research, v. 35, n. 1, january,p. 233-241.

Mcculloch, W. S.; Pitts, W. 1943.A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics, v. 5 p.115–133, 1943.

Oliveira, J. B. 2008. *Sistemas inteligentes e wavelets para previsão de vento e geração eólica*. 92 f. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Federal de Pernambuco, Recife.

Rojas, I.; Valenzuela, O.; Rojas, F.; Guillen, A.; Herrera, L.J.;Pomares, H.; Marquez, L.; Pasadas, M. 2008. Soft-computing techniques and ARMA model for time series prediction. Neurocomputing, v. 71, n. 4-6, p. 519–537.

Rufino, M. P. 2014.*Implementação de uma Rede Neural Artificial para Classificação de Sinais Mioelétricos*. 31f. Monografia (Curso de Graduação em Engenharia Elétrica) – Universidade Federal da Paraíba, João Pessoa.

Sampaio, A. F. 2020.*Modelo híbrido com ARIMA e RNA aplicado para a previsão de curto prazo de velocidade do vento direcionado ao estudo de potencial eólico em região do estado do Ceará*, 57 f. Dissertação (Mestrado) - Curso de Mestrado Acadêmico em Ciências Físicas Aplicadas, Centro de Ciência e Tecnologia, Universidade Estadual do Ceará, Fortaleza.

Silva, C. Q. 2019.*Redes neurais aplicadas no reconhecimento de símbolos matemáticos manuscritos online*. 54f. Monografia (Graduação) - Curso de Matemática Aplicada, Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo.

Silva, S.C. da. 2014.*Utilização de redes neuras na previsão de vento do litoral do Ceara* UECE, 105 f. Dissertação (Mestrado) - Curso de Mestrado Acadêmico em Ciências Físicas Aplicadas, Centro de Ciência e Tecnologia, Universidade Estadual do Ceará, Fortaleza.

Souza, R. C.; Camargo, M. E. 2004.Análise e previsão de séries temporais. Os modelos arima. 2st. ed. Rio de Janeiro: Gráfica e Editora Regional.

Werner, L.; Ribeiro, J. L. D. 2003. Previsão de demanda: uma aplicação Box-Jenkins na área de assistência técnica de computadores pessoais. Gestão & Produção, v. 10, n. 1, p. 47-67.

Willmott, C. J. 1981. On the validation of models. Journal PhysicalGeography, v. 2, n. 2,p. 184-194.

****