ARIMA methodology applied to the very short-term forecast of wind power generation at Palmas Wind Farm

Metodologia ARIMA para previsão da geração de energia eólica de curtíssimo prazo aplicada ao Parque Eólico de Palmas

Paulo H. Soares[1]
Alexandre K. Guetter[1]

[1]**UFPR - Universidade Federal do Paraná – Departamento de Hidráulica e Saneamento**
soareshpaulo@gmail.com
guetter.dhs@ufpr.br

Abstract: Wind power is already a consolidated power source for electricity generation, with more than 300 GW installed worldwide. It shows impressive growth numbers, such as an increase of 715% in the world installed capacity in the period of 2003 to 2013. In Brazil, the share of wind power on the overall energy production has increased from less than 1% to almost 5% in last five years. However, the electricity generated from the wind is a source of uncertainties for power system operators and can generate undesired variations in the energy quality for the energy system as a whole, decreasing its efficiency. To decrease the uncertainties and to increase the efficiency of wind power generation are goals of wind power forecasting (WPF) models. This work shows the development of a probabilistic WPF model applied to Palmas Wind Farm, located in the state of Paraná, Brazil. The well-known autoregressive integrated moving average (ARIMA) methodology was employed to develop a very short-term WPF model. The performance of the model was evaluated using the mean absolute error (MAE) test, root mean squared error (RMSE) test and the Nash-Sutcliff (NS) index. It was found that the model is able to satisfactorily forecast up to three hours ahead, with a maximum error of 8.47%. The quality of the forecasts, evaluated by the MAE, RMSE and Nash-Sutcliffe index also showed satisfactory results. It is expected that the achievements of this work may be a reference for future works in the field.

Keywords: Wind power, electricity generation, wind power forecast, probabilistic models, renewable energy.

Resumo:

A energia eólica vem apresentando uma tendência de crescimento em sua capacidade instalada, tanto no Brasil quanto no mundo. O período entre 2003 e 2013 apresentou um crescimento de cerca de 715% na capacidade instalada mundial. No Brasil, a fatia de energia gerada pelo vento representa quase 5% do cenário energético atual, contra menos de 1% há cinco anos. A expansão da geração eólica traz consigo desafios para o sistema elétrico e seus operadores, na forma de incertezas e variações sobre a qualidade da energia gerada. Diminuir essas incertezas e contribuir para tornar os sistemas eólicos mais eficientes são objetivos da previsão da geração de energia eólica (PGEE), foco deste trabalho. As previsões apresentadas resultam de um modelo probabilístico desenvolvido com base nas séries históricas de dados de velocidade de vento da Usina Eólio-Elétrica de Palmas (COPEL), localizada no estado do Paraná, Brasil. O modelo preditivo faz uso da metodologia de modelagem ARIMA (modelos autorregressivos integrados e de médias móveis) para gerar os resultados apresentados neste trabalho. O desempenho da PGEE desenvolvida mostrou-se satisfatória para um horizonte de até três horas, com um erro máximo de 8,47%. Além de ser capaz de prever satisfatoriamente a geração da energia eólica, este modelo tem a pretensão de servir de referência para futuros trabalhos sobre o tema.

Palavras-Chave: Energia eólica, geração de energia elétrica, previsão de geração eólica, modelos probabilísticos, energias renováveis.

1 Introduction

When this work began in 2013, Brazil had 2.4 GW of installed wind power capacity. By March of 2015, this amount had increased to 5.7 GW, and it was expected that the installed wind power capacity would reach 7.9 GW by the end of 2015 [1]. Data from Brazilian Mines and Energy Ministry (MME) shows that, in July of 2015, wind power was responsible for 4.4% of the total energy produced in the country [2]. Five years earlier, the share of the energy production for wind power was only 0.72% [3].The Brazilian wind power production growth is aligned with the global tendency of increase in the wind power generation. According to the Renewable Energy Policy Network for the 21st Century (REN 21), the global wind power installed capacity reached 318 GW in 2013. In 2003 the world wind power installed capacity was only 39 GW. Thus, in ten years, the world wind power installed capacity grew 715%.

The increasing amount of wind power in the Brazilian and global energy markets was one of the main motivations for this research. The generation of electricity by wind has challenges that differ from those encountered through other sources. Unlike water, which can be stored in a reservoir, it is more difficult to control the amount of energy that is possible to be produced by wind [5]. On the other hand, wind power has also some advantages over other power generation sources. It does not emit gases to the atmosphere, because the source is clean and inextinguishable. Finally, its installation is relatively easy when compared to other power sources (except perhaps solar power) and it causes little environmental impact [5] [6].

Wind power is indeed a challenging matter for the energy system and its operators. Due to the peculiarities of wind power generation, the insertion of the produced electricity in the energy grid, if not done properly, can affect the quality of the whole power system. Harmonic distortions, voltage fluctuations and odd frequencies are examples of the possible problems [7]. The wind power forecast (WPF) arose as a solution to diminish these undesirable possibilities. There is more than one way to perform a WPF forecast, depending on the horizon and the type of the forecast. Very short-term and short-term forecasts usually employ statistical methods and depend only on the wind speed and wind direction time series. Mid-term and long-term forecasts employ numerical weather forecast methods to develop the WPF model [8]. Regardless of the type of definition, all WPF models aim at providing better certainty of how much wind power will be generated, allowing for an easier insertion of the electricity into the power grid [8] [9]. WPF can also be employed as a tool for planning the maintenance schedule for the wind farms. For all these reasons, WPF can be considered as a fundamental tool for wind power generation [8].

This work presents the development of a statistical WPF model based on the wind speed time series of Palmas Wind Farm. The model employed the ARIMA methodology, in order to forecast the future wind speed data. The ARIMA methodology chosen has been widely employed for forecasting purposes. Besides, this methodology allows future improvements of the model, through adding other features, such as support vector regression [10] or artificial neural networks.

The development of the model was performed through a computational approach. It was written according to two computational routines; the first one, employed for time series analysis and other pre-modeling procedures, was written using the Python programming language. The second, the forecast model itself, was written as a routine for the statistical software tool called R.

The model actually forecasts the wind speed and, after that, it converts the forecasted wind speed into wind power, through the power curve of the wind turbines. It can forecast up to three hours ahead with accuracy. It showed, for the most inaccurate case (three hours ahead forecast), a MAE (*Mean Absolute Error*) of 0.99 m/s during summertime and 1.19 m/s during wintertime. The Nash-Sutcliffe index showed positive results for the three hours ahead forecast, about 0.57, for both seasons.

The wind power forecast was equally able to provide precise forecasts, as shown in Figure 7. Also, the final difference between the actually measured energy and the forecasted energy for the selected period was less than 10%.

To develop a model able to predict wind power with satisfactory precision was the main goal of this work, which was conceived to be a baseline study, a reference for future studies developed by the team of the Department of Hydraulics and Sanitation (DHS) of Universidade Federal do Paraná. The secondary goal of this work is perhaps to promote this field of study within the research department.

This paper is organized as follows: section 2 exhibits the details of Palmas Wind Farm; section 3 presents the methodology behind the forecast model; section 4 is dedicated to results and interpretation of those results; finally, section 5 shows conclusions from this work.

2 Palmas Wind Farm

The models developed in this work were calibrated and validated based on the wind speed data of Palmas Wind Farm, which is located on the 26th km of the highway PR-280, in the State of Paraná, Brazil and whose georeferenced coordinates are 26° 34’46,8 S e 51° 41’51,0 W. The place is distant 320 km from Curitiba, the capital of the state. The nearest city, which is also named Palmas, is distant 30 km from the wind farm. The region where the wind farm is installed presents a plain relief with few and smooth hills and is located at 1350 m above sea level.

The main economic activity of the region is cattle raising, and the fields of Palmas are used for pasture. The wind farm installation exerted a negligible impact on the economic activity of the region. [11].

The wind farm consisted of five Enercon model E-40 wind turbines, whose technical information is presented in Table 1.

Table 1: Enercon E-40 Wind Turbine Specifications.

Hub height (m)	44.0
Rotor diameter (m)	40.3
Cut-in wind speed (m/s)	2.5
Rated wind speed (m/s)	12.0
Cut-out wind speed (m/s)	25.0

The wind farm was conceived through the joint effort of COPEL and Wobben Windpower, which is affiliated with Enercon GmbH in Brazil [12]. Figure 1 shows a photograph of one of the wind turbines of the wind farm.

Palmas Wind Farm was the first of its kind installed in the southern part of Brazil and its construction took only one week. The wind farm started to generate electricity in February, 1999 [13].

Figure 1: Enercon E-40 wind turbine in operation at the Palmas Wind Farm in 2014. Photo taken by the author.

3 Methodology

The essential data for this work was the wind speed time series, which was made available by COPEL. The database consisted by a ten minute wind speed data series and by a ten minute wind power data series, both ranging from January 2008 to December 2011. The wind speed data series was produced by the records of an anemometer installed in the wind farm, while the wind power data series were acquired from the SCADA system. The first three years of both time series (2008-2010) were used to develop the model, and the last year (2011) was used to evaluate the its performance.

3.1 Data series analysis

In order to correctly develop the forecast model, some initial procedures had to be performed. The first was to analyse the given time series, removing any possible inconsistent data and missing values. This step was performed through a computational routine, built in Python 2.7 and the corresponding packages for data analysis, such as Python Pandas.

After this first step, it was necessary to set the wind speed data to the correct height of the wind turbines. The anemometer of the site is placed in a 75 metres height pole; and the hubs of the wind turbines are placed 44 metres above the ground. Assuming that the wind behaviour follows a logarithmic profile, the interpolation was made by employment of the relation expressed in equation 1 [14].

$$\frac{U(z)}{U(h)} = \frac{ln\left(\frac{z}{z_0}\right)}{ln\left(\frac{h}{z_0}\right)} \quad (1)$$

where $U(z)$ represents the wind speed measured by the anemometer (m/s), $U(h)$ represents the wind speed in the desired height (m/s), z is the height of the anemometer (in metres), h is the height of the wind turbine (m) and z_0 represents the surface roughness (in metres), estimated as 0.03 metres [15].

The corrected ten minute time series was then transformed into an hourly time series, by taking its average, as shown in Figure 2.

The graphical view of the time series can be considered the first step in the development of the forecasting model. It offers the initial comprehension of the behaviour of the data series. Furthermore, a quantitative analysis was performed to understand the behaviour of the time series. Statistical parameters such as the mean, standard deviation, variance, skewness and kurtosis were assessed and will be discussed later.

3.2 ARMA/ARIMA models theory

The methodology of the employment of the ARMA/ARIMA models was developed by George Box and Gwilym Jenkins in the 70's. One of the objectives of this methodology is to forecast future events of a certain phenomenon, by using the past time series of this event [16], [17].

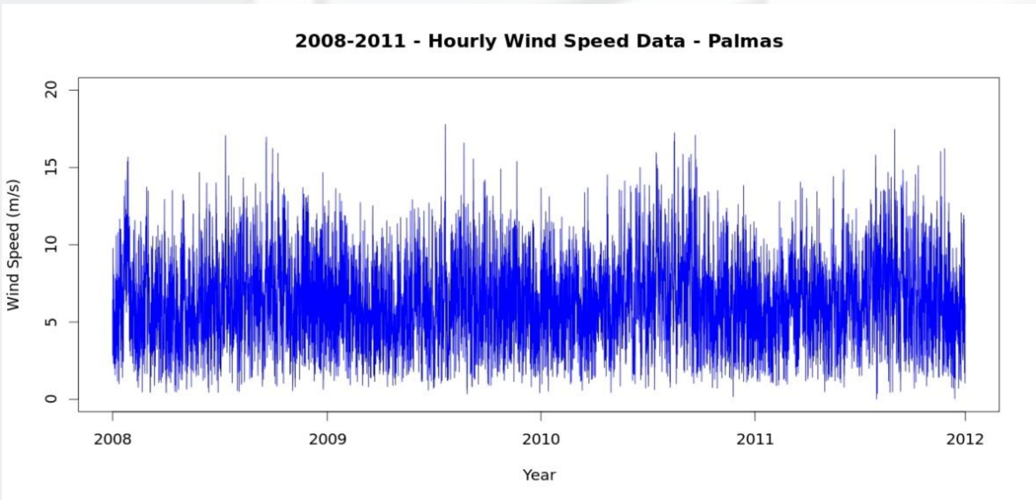

Figure 2: Original wind speed data series.

ARMA models join autoregressive models (AR) and moving average models (MA) together. They can be expressed through equation 2.

$$\tilde{z}_t = \emptyset_1 \tilde{z}_{t-1} + \emptyset_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2)$$

where $\tilde{z}$ represents an autoregressive and moving average process; $\emptyset_p$ represents the autoregressive parameters; $\tilde{z}_{t-1}$ represents the past values of the time series; θ_q represents the moving average parameters and a_{t-q} represents the random error (white noise process) [17].

This equation is the union of the models mentioned above. The autoregressive model represents a process through the linear combination of the p past values of the time series ($\tilde{z}_t$) plus a random white noise term (a_t) [17]. The p past values of the time series dictate the order of the AR model.

Moving average models are similar to the autoregressive models in the way they are dependent on q past values of the time series used to generate the model, although the model combines (and weights) the q values with the white noise process (a_t). Similar to the AR model, the order of an MA model is dictated by the value of q [17].

In both cases it is assumed that the models represent processes with mean zero and variance σ^2 [18].

The ARMA approach has an important characteristic: it was developed to model stationary time series. This often can be a problem since many real time series phenomena are not stationary.

But if the time series shows a non stationary behaviour, it is still possible to use an ARMA model for the time series. This can be done by differencing the time series in respect to a certain time interval, which aims at finding an homogenous behaviour of the time series between the range of time. Once this homogeneity is found, the series have to be integrated (or summed) again so an ARMA model can be adjusted; this integration step is part of the name of this approach: ARIMA models (*Autoregressive Integrated Moving Average models)* [17], whose definition is given by equation 3.

$$w_t = \emptyset_1 w_{t-1} + \dots + \emptyset_p w_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (3)$$

where $\emptyset_i$ represents a stationary autoregressive operator and w_t can be understood as the differenced time series. The definition of w_t is expressed in equation 4.

$$w_t = \nabla^{\mathrm{d}} \mathrm{z}_{\mathrm{t}} \quad (4)$$

in this case, d represents the number of differences needed to achieve the stationary state [17].

3.3 Stationarity of the time series

The stationarity of a time series can be evaluated graphically by the autocorrelation function (ACF) and also by testing the existence of unitary roots in the characteristic polynomial equation of the time series [17]. In this work, the stationarity of the wind speed time series were tested using both approaches. The ACF plot was compared to the Augmented Dickey Fuller (ADF) unity roots test. The aim of the test for a general AR model, given by $\varphi(B)z_t = a_t$, is to test the existence of a unit root in a general AR model, whose definition is shown in equation 5 [17], [19]:

$$\Delta z_t = \propto . z_{t-1} . \sum_{i=1}^{p} \varphi_i . X_{t-i} + a_t \quad (5)$$

The ADF test is a hypothesis test, where the null hypothesis, H_0 corresponds to the existence of a unit root in equation 5 and is equivalent to the statement of equation 6.

$$\propto = \sum_{i=1}^{p+1} \varphi_i = 1 \quad (6)$$

The existence of a unit root in the AR operator $\varphi(B)$ can be interpreted as a strong indicator of the non-stationary nature of the time series [17],[19].

3.4 Choosing the most suitable model

Choosing a proper model requires a correct estimate of the number of AR and MA operators (p, q), as well the number of differences (d) needed for reaching the stationarity condition of the time series. This choice is not always obvious and can be improved by statistical tests such as the *Akaike Information Criteria* (AIC) and the *Baysean Information Criteria* (BIC), both employed in this work. The tests are based on the maximum likelihood of the values and on the number of parameters of the models [17],[20].

The AIC test can be defined by the expression in equation 7:

$$AIC = \frac{2ln(\theta) + 2r}{n} \quad (7)$$

where θ represents the value of the maximum likelihood; r is the number of parameters considered in the modelling and n is the size of the sample [17].

The BIC test can be defined by equation 8.

$$BIC = \ln(\hat{\sigma}_a^2) + r\frac{ln(n)}{n} \quad (8)$$

where $\hat{\sigma}_a^2$ represents an estimative of the value of the maximum likelihood; r and n are defined in the same way as in the AIC [17].

For both tests, the best model will always be the one that shows minor AIC or BIC values. In fact, the tests compare various possible ARMA/ARIMA models and indicate which of the selected models may be the best fit for the selected sample (time series) [20].

It is necessary to state that the tests are intended to support a decision, not to be the only way of making a decision on which is the best model.

3.5 Forecast Evaluation

Three statistical tests were employed in order to evaluate the results generated by the ARMA/ARIMA models:

a) Mean Absolute Error (MAE)

The MAE is a measure of the absolute difference between a forecast and the actual corresponding observation. The test is defined by equation 9.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(|for(i) - obs(i)|) \qquad (9)$$

where $for(i)$ means the forecasted values, $obs(i)$ means the observed values and n represents the size of the sample employed in the test. The use of the MAE test is supported by [21]

b) Root Mean Squared Error (RMSE)

The root mean squared error is a widely employed estimate of the performance of a model, mainly when the noise of the model tends to follow the normal distribution, but is still valid for the purposes of this work [22],[23].

The RMSE is defined by equation 10.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(|for(i) - obs(i)|)^2} \qquad (10)$$

The parameters $n, for(i)$ and $obs(i)$ are defined identically as in the MAE.

c) Nash-Sutcliffe Index (NS)

The Nash-Sutcliffe index is a measure of the efficiency of the model. The results of the index vary from $-\infty$ to 1; the latter being the goal for a perfect model. NS values close to 0 indicates that the performance of the model´s simulations is close to the mean of the data employed to build the model. Negative NS values indicate that the mean of the data is a better predictor than the model. The test is defined as follows [24]:

$$NS = 1 - \frac{\sum_{i=1}^{n}(|for(i) - obs(i)|)^2}{\sum(obs(i) - \overline{for(\iota)})^2} \qquad (11)$$

where $\overline{for(\iota)}$ means the average of the forecasted values. The other parameters have already been described.

3.6 Wind power forecast

The wind speed forecast is the fundamental result of the predictive model. From this result, the wind power forecast was performed for Palmas Wind Farm. It was done by adjusting the forecasted wind speed values to the characteristic power curve of the wind turbine.

The power curve shows the behaviour of the turbine according to different wind speeds. In this case, the wind turbine is an Enercon E40, whose power curve is shown in Figure 3.

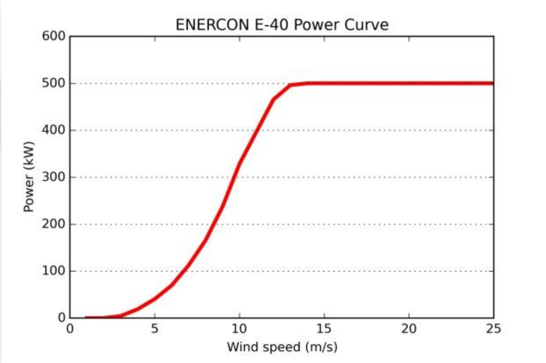

Figure 3: Enercon E40 wind turbine power curve.

Finally, the adjusted values are multiplied by the number of turbines of the wind farm (five in this case) leading to the final forecasted wind power output.

4 Results and discussion

In this section results of the model for two randomly chosen dates are presented, one in the winter of 2011 and another in the summer of the same year. The results consist of 100 forecasted values for the randomly selected dates.

4.1 The behaviour of the wind

The wind speed time series shown graphically in Figure 2 corresponds to the first look into the data series behaviour. This look is complemented by the frequency histogram shown in Figure 4, which classifies the occurrence frequency of observed wind speeds.

The histogram shows that most of the observed wind speeds occur within the range of 4 m/s and 10 m/s. and the average is about 6 m/s.

A Weibull distribution was fitted to the wind speed data distribution frequency, which revealed to be a good fit. The distribution has a shape factor of 2.67 and a scale factor of 7.17.

These factors are used in the Weibull distribution, which represents the probability of wind speed to be v for any time interval. This probability was calculated by equation 12.

$$w(v) = \left(\frac{a}{b}\right)\left(\frac{v}{b}\right)^{(a-1)} e^{-\left(\frac{v}{b}\right)^{a}} \qquad (12)$$

where a is the shape parameter and b is the scale factor for any value in the range $0 < v < \infty$ [25].

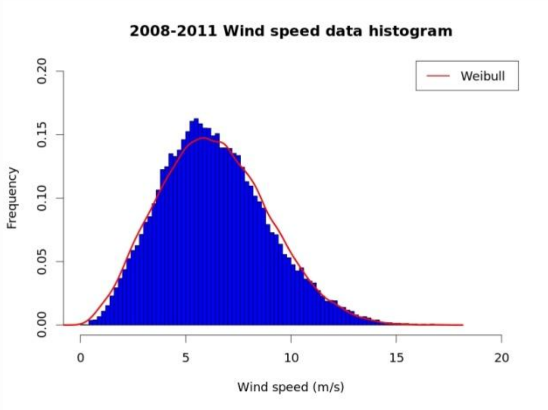

Figure 4: Wind speed data distribution histogram.

Quantitative data were also calculated in order to have the precise measure of the behaviour of the time series. This took into account statistics such as: maximum, minimum and average wind speed through the considered period of time; also the standard deviation, the skewness and the kurtosis.

Table 2 shows the main statistics for the wind speed time series. These results contribute to improve the comprehension of the histogram. The positive skewness value indicates that the majority of the data is grouped below the average wind speed of 6.37 m/s. The kurtosis value indicates a sharper peak in data distribution (relative to the normal distribution). Both cases match with the data shown in the histogram.

Table 2: 2008-2011 hourly wind speed data series statistics.

Max. wind speed (m/s)	17.78
Average wind speed (m/s)	6.37
Min. wind speed (m/s)	0.23
Standard deviation	2.56
Skewness	0.44
Kurtosis	0.07

4.2 Stationarity

Following the methodology presented in the previous section, it was then necessary to examine whether the time series would exhibit a stationary behaviour or not. Hence the autocorrelation function (ACF) was calculated and plotted, as shown in Figure 5. The confidence interval for the plot is 0.9 or 90%.

It can be observed from the Figure 5 that the ACF function does not present a constant decay with time, as expected for a white noise process [17]. Conversely, between lag 10 and 15 the function begins to increase. This was the first indicative that the time series have a non-stationary behaviour.

As discussed in the previous section, the stationarity of a time series also can be assessed by testing the presence of a unit root in the characteristic polynomial of the time series. In the present case this was evaluated performing the ADF test. Table 3 shows the outputs of the test.

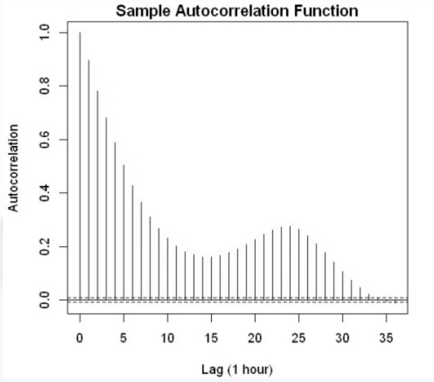

Figure 5: Results of ACF.

Table 3: ADF tests results.

	Test statistics	Critical value
Original	-5.53	-2.58
1st Difference	-41.62	-2.58

The ADF test results shows that the original wind speed time series may have no unit root lying in the unit circle. But the first difference of the time series exhibited more distant test statistics from the critical value. This fact and the conclusions taken from the ACF function suggested the development of an ARIMA model based on a differentiated time series instead on working with the original one.

4.3 Choosing the parameters of the model

The last step before the development of the WPF model was to decide the parameters of the model. As discussed in the previous section, to help in the decision process, the AIC and BIC tests were performed between a set of 20 possible ARIMA (p, d, q) models. The three best models are shown in Table 4:

Table 4: AIC & BIC tests results.

Model	σ^2	AIC	BIC
ARIMA (1,1,2)	1,269	80596,10	80628,8
ARIMA (2,1,1)	1,269	80606,59	80639,28
ARIMA (3,1,2)	1,268	80592,39	80641,44

The results from AIC and BIC agree with the ACF plot and with the ADF test, showing that a differentiation is necessary in order to obtain the better results. Thus, an ARIMA (1,1,2) was adjusted to forecast wind speed and its results are shown in the next section.

4.4 Forecasting model

As stated in the former sections, the WPF model is essentially a wind speed forecasting model built to perform an extra step, after the wind speed is forecasted, which is to calculate the wind power for the corresponding wind speed forecasted value. This way, it is necessary first to develop the ARIMA (1,1,2) to forecast the wind speed, then evaluate the outputs of the model and only after that, the wind power can be estimated.

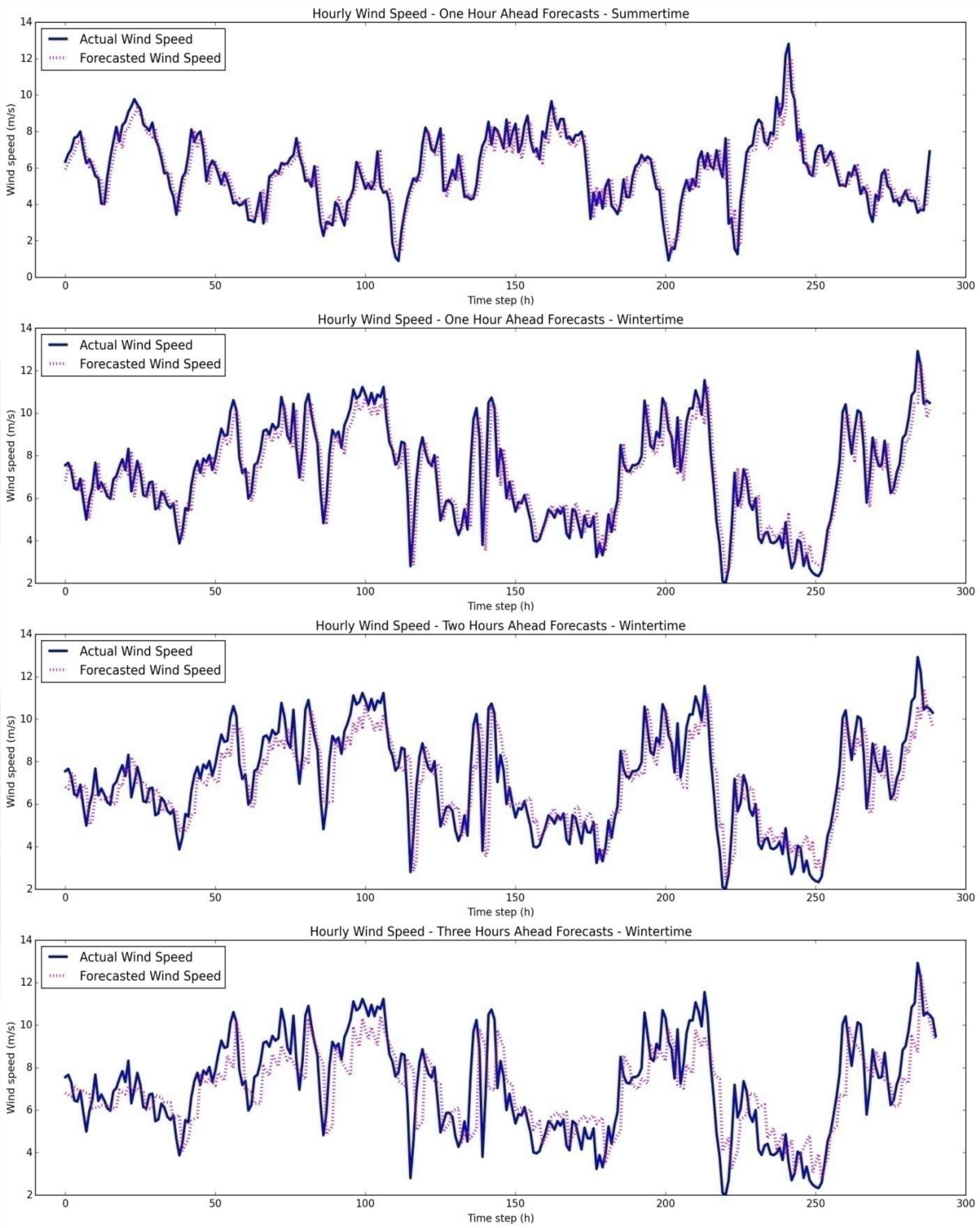

Figure 6: Forecasting results according to different time horizons and seasons of the year of 2011.

The ARIMA (1,1,2) model followed equation 13. As expected, this equation presents one autoregressive component and two moving average components. This characteristic is fixed for the model. The statistical software was able to recalculate and to update the coefficients (the numbers) of the equation within every new forecast.

$$w_t = 0.8609w_{t-1} - 0.8549a_{t-1} - 0.1427a_{t-2} \quad (13)$$

The model was designed in a way it could forecast the wind speed up to three hours in the future. This was done by adjusting the lags of the model. In fact the model could be extended to as many lags as desired, but after three lags (or three hours ahead) the outcomes of the model would be of little value.

In order to check its strength against the seasonal variability, the model had to forecast the wind speed of a random date in summer (summertime forecast) and of another random date in winter (wintertime forecast). Figure 6 shows the graphical results of the forecasts performed by the model.

As can be seen in Figure 6, the one hour ahead forecasts performed during summertime and during wintertime showed equally promising results. The results tend to degrade with the increase of the lag of time, as can be seen for the two and three hours ahead forecasts, where the distance between the forecasted values and the observed wind speed values (the blue line) increase. This result was expected for an autoregressive model.

Table 5 presents the results of evaluation tests for the hourly wind speed forecast. As expect from the observation of the forecast plot, the evaluators (MAE, RMSE and NS) show differences increasing with the increase of the horizon of the forecasts. However, these results can be considered satisfactory, at least for forecasts up to three hours ahead.

It can also be assessed from the results that the model is robust and is not greatly affected by seasonal variations.

Table 5: Wind speed forecasts evaluation results.

Summertime Forecasts Evaluation			
	MAE(m/s)	**RMSE (m/s)**	**NS**
Lag = 1	0.71	0.92	0.78
Lag = 2	0.85	1.09	0.69
Lag = 3	0.99	1.28	0.57
Wintertime Forecasts Evaluation			
	MAE(m/s)	**RMSE (m/s)**	**NS**
Lag = 1	0.81	1.07	0.79
Lag = 2	1.01	1.35	0.67
Lag = 3	1.19	1.54	0.56

4.5 Wind power forecast

The final step of the WPF model employs the wind turbine power curve in order to estimate the amount of wind power that could possibly be generated for the corresponding forecasted wind speed.

Results of the WPF model are displayed graphically in Figure 7. Again the difficulties of the model in forecasting extreme values can be noted. The degrading of the forecasts, with the increase of the time lag, can also be seen.

Despite these difficulties it is worth noting that the general behaviour of the forecasts tends to follow the actual wind power time series, even for higher forecast horizons. Small differences between the actual energy produce by the wind farm and the forecasted energy has shown that the model is able to forecast wind power with a satisfactory level of precision, as shown in Table 6.

Table 6: Differences of the wind power generated and forecasted.

	Measured	Forecast (lag 1)	Forecast (lag 2)	Forecast (lag 3)
Energy (kWh)	115290.86	114474.22	112618.56	105524.43
Error (kWh)	–	816.64	2672.30	9766.43
Error (%)	–	0.7%	2.31%	8.47%

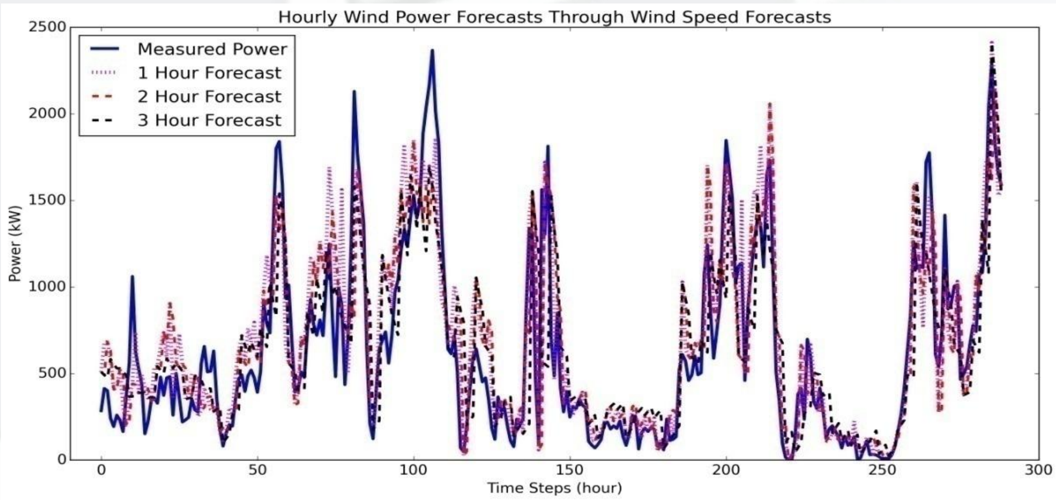

Figure 7: Wind power forecasts performed through wind speed forecasts.

It can be inferred from Table 6 that the error between the observed and forecasted energy is small (0.7%) for the one hour ahead (lag = 1) wind power forecast. The error increases with the lengthening of the horizon of the forecast. For the two hours ahead forecast, the error percentage is 2.31%, while for the three hours ahead this number reaches 8.47%, which can still be considered acceptable for forecasting purposes.

5 Conclusions

The goal of this work was to show the development of a wind power forecasting (WPF) model as well its results. As stated in the introductory section of this paper, the model was designed to be a baseline study for the newly established wind power research group of the Department of Hydraulics and Sanitation (DHS) at Universidade Federal do Paraná.

The first goal, the development of the WPF model, was successfully achieved. The designed model has shown its ability to forecast the wind speed and the wind power for Palmas Wind Farm, as long as forecasts do not exceed three hours ahead in time.

The model presented degrading results with the increase of the horizon of the forecast. But the results still exhibited acceptable MAE, RMSE and NS results. The degrading phenomenon is also expected for an autoregressive model.

The wind power forecasting model also showed reasonably enough forecasts; and the computed maximum difference between the forecasted energy and the measured energy was only 8.47% for the three hours ahead forecasts.

Summarizing, the main goal of the work was accomplished and now it is expected that this model will be improved in future works, by adding to it different approaches, such as the Artificial Neural Networks (ANN), which can possibly expand the forecast horizon of the model as well improve the reliability of the model's forecasts.

Acknowledgements

To the Graduate Program in Environmental Engineering and Water Resources (PPGERHA) of Universidade Federal do Paraná (UFPR) for the opportunity of developing this work and to CAPES foundation, for the provided assistantship.

References

[1] BRASIL. Portal Brasil. **Infraestrutura**. Brasília 2015. Retrieved from: <http://www.brasil.gov.br/infraestrutura/2015/04/capacidade-de-energia-eolica-deve-crescer-62-em-2015-no-pais> Accessed in 08 out 2015.

[2] BRASIL. Ministério de Minas e Energia. **Boletim Mensal de Monitoramento do Sistema Elétrico Brasileiro – Agosto de 2015.** Brasília, 2015. Retrieved from: <http://www.mme.gov.br/documents/10584/2027273/BoBolet+de+Monitoramento+do+Sistema+El%C3%A9trico+-+Agosto-2015.pdf/5efc24ca-ea15-445b-845ad724b3308cc9>. Accessed in 08 out 2015.

[3] BRASIL, Ministério de Minas e Energia, **Monitoramento do Sistema Elétrico Brasileiro – Boletim de Julho/2010.** Brasília, 2010. Retrieved from: <http://www.mme.gov.br/documents/10584/1832844/BoleBol_de_Monitoramento_do_Sistema_Elxtrico_julho_2010.pdf/cd199247-66a6-46a6-8fe3-a4a21a1cc5a6> Accessed in 08 out 2015.

[4] REN21, Renewables. Global status report. Renewable Energy Policy Network for the 21st Century, France, 2013.

[5] LU, M. S.; CHANG, C. L.; LEE, W. J.; WANG, L. Combining the wind power generation system with energy storage equipments. In: **Industry Applications Society Annual Meeting, 2008. IAS'08. IEEE**. IEEE, 2008. p. 1-6.

[6] VILLELA, I. A. de C.; SILVEIRA, J. L. Ecological efficiency in thermoelectric power plants. **Applied thermal engineering**, Elsevier, v. 27, n. 5, p. 840–847, 2007.

[7] ANAYA-LARA, O.; JENKINS, N.; EKANAYAKE, J.; CARTWRIGHT, P.; HUGHES, M. **Wind energy generation: modelling and control**. John Wiley & Sons, 2009.

[8] MONTEIRO, C.; KEKO, H.; BESSA, R.; MIRANDA, V.; BOTTERUD, A.; WANG, J.; CONZELMANN, G. **A quick guide to wind power forecasting: state-of-the-art 2009**. Argonne National Laboratory (ANL), 2009.

[9] WANG, J.; BOTTERUD, A.; MIRANDA, V.; MONTEIRO, C.; SHEBLE, G. Impact of wind power forecasting on unit commitment and dispatch. In: **Proc. 8th Int. Workshop Large-Scale Integration of Wind Power into Power Systems**. 2009. p. 1-8.

[10] RODRIGUES, S. B.; TEIXEIRA, L. L.; JÚNIOR, L. A. T.; SILVA, A. C. L. d.; SIQUEIRA, P. H.; FRANCO, E. M. C. Forecasting of the relative displacements in block I11 of the Itaipu hydroelectric. **Espaço Energia**, Issue 22, 2015.

[11] CAMARGO, A. S. G. d. **Análise da operação das Usinas Eólicas de Camelinho e Palmas e avaliação do potencial eólico de localidades no Paraná**. 2005. Dissertação de Mestrado, CEFET - Paraná.

[12] WOBBEN Windpower. Apresentação da Empresa. 2016. Retrieved from: <http://www.wobben.com.br/empresa/wobben/apresentacao-da-empresa/>

[13] COPEL: Companhia Paranaense de Energia. Usina Eólica de Palmas. 2013. Retrieved from: <http://www.copel.com/ci/_edicoes/299/usina_eolica_de_palmas.jsp.>

[14] BARTHELMIE, R. J.; PALUTIKOF, J. P.; DAVIES, T. D. Estimation of sector roughness lengths and the effect on prediction of the vertical wind speed profile. **Boundary-Layer Meteorology**, v. 66, n. 1-2, p. 19-47, 1993.

[15] LIMA, J. M.; GUETTER, A. K.; FREITAS, S. R.; PANETTA, J.; LAMOSA, D.; OLIVEIRA, F. Previsão da geração de energia eólica no horizonte de curtíssimo prazo utilizando um modelo regional de previsão do tempo. In: XXII

SEMINÁRIO NACIONAL DE PRODUÇÃO E TRANSMISSÃO DE ENERGIA ELÉTRICA, Brasília, 2013.

[16] DA SILVA, L. M. O. **Uma aplicação de árvores de decisão, redes neurais e KNN para a identificação de modelos ARMA não-sazonais e sazonais**. 2005. Tese de Doutorado. PUC-Rio.

[17] BOX, G. E.; JENKINS, G. M.; REINSEL, G. C. **Time series analysis: forecasting and control**. John Wiley & Sons, 2008.

[18] COMMANDEUR, J. J.; KOOPMAN, S. J. **An introduction to state space time series analysis**. OUP Oxford, 2007.

[19] YANG, X.; ZHANG, L. A note on self-normalized Dickey-Fuller test for unit root in autoregressive time series with GARCH errors. **Applied Mathematics-A Journal of Chinese Universities**, v. 23, n. 2, p. 197-201, 2008.

[20] EMILIANO, P. C. **Fundamentos e aplicações dos critérios de Informação:** akaike e bayesiano. 2009. Dissertação de Mestrado. Universidade Federal de Lavras.

[21] ERDEM, E.; SHI, J. ARMA based approaches for forecasting the tuple of wind speed and direction. **Applied Energy**, v. 88, n. 4, p. 1405-1414, 2011.

[22] CHAI, T.; DRAXLER, R. R. Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. **Geoscientific Model Development**, v. 7, n. 3, p. 1247-1250, 2014.

[23] GOMES, J.; MINE, M. R. M. Estimativa de parâmetros de modelos hidrológicos utilizando o filtro de Kalman. **Revista Brasileira de Recursos Hídricos**, p. 67–86, 1998.

[24] MCCUEN, R. H.; KNIGHT, Z.; CUTTER, A. G. Evaluation of the Nash–Sutcliffe efficiency index. **Journal of Hydrologic Engineering**, American Society of Civil Engineers, v. 11, n. 6, p. 597–602, 2006.

[25] PATEL, M. R. **Wind and solar power systems: design, analysis, and operation**. CRC press, 1999.