

MODELAGEM MATEMÁTICA E PREVISÃO DA VELOCIDADE DO VENTO EM PALMEIRA DOS ÍNDIOS/AL COM A UTILIZAÇÃO DOS MODELOS BOX & JENKINS

Lidiane Torres Pereira¹

torres.lidiane@outlook.com

Milton Perceus²

miltonperceus@uneal.edu.br

Eanes Torres Pereira³

eanes@computacao.ufcg.edu.br

Resumo

Neste artigo, são apresentadas uma análise descritiva, modelagem matemática e, em seguida, realizada uma previsão do tipo *One-step-ahead* dos dados de uma série temporal de velocidade do vento da cidade de Palmeira dos Índios (AL), tal série foi obtida no site do Instituto Nacional de Meteorologia (INMET), com o objetivo de estimar as condições do vento dessa localidade. A base de dados é composta por 82 observações e compreende o período de Janeiro de 2010 a Outubro de 2016, a torre da qual foram obtidos os dados tem 100m de altura, com anemômetros em 03 níveis (30m, 70m e 100m) e sensores de direção do vento (30m e 70m). Estatísticas descritivas foram utilizadas para analisar o comportamento da série, e foi obtido como média geral da velocidade o valor de 2,78m/s, com um desvio padrão de 0,93 e o coeficiente de variação 33,6%. A metodologia de Box-Jenkins foi adotada para analisar a série univariada de tempo, isto é, essa metodologia é utilizada na análise dos valores de uma única série temporal. São adotados para verificação do melhor modelo o Critério de Informação *Akaike* (AIC) e o Critério de Informação *Bayesiano* (BIC) juntamente com a análise da Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP). O modelo Auto-Regressivo Integrado de Média Móvel (ARIMA) foi empregado para realizar as previsões, após a análise de quatro modelos sugeridos, o modelo ajustado foi o ARIMA (1, 1, 0). O Erro Quadrático Médio (EQM) foi usado para quantificar o erro e verificar a escolha do melhor modelo, e teve resultado igual a 0,0563. Por fim, foram geradas as previsões de 12 amostras e comparadas com as 12 observações, destacando que as 12 amostras geradas na previsão apresentaram-se dentro do limite de confiabilidade.

Palavras-chave: Estatística descritiva. Modelos ARIMA. Previsão de séries temporais. Velocidade do vento.

Abstract

In this article, a descriptive analysis and a mathematical modeling of a wind speed time series are presented. Afterwards, an one-step ahead forecast is performed. The time series data were obtained from the INMET (Instituto de Meteorologia) website. The objective of this research was to estimate the wind conditions in Palmeira dos Índios (state of Alagoas) city. The dataset is composed of 82 samples and comprises the period from January of 2010 up to October of 2016. The data were obtained from a tower with 100m high, and anemometers located at 3 levels (30m, 70m, and 100m). Wind direction sensors were located at 2 levels (30m and 70m). Descriptive statistics to analyze the behavior of the series, and it was obtained a general average of 2.78m/s,

¹ Graduanda licenciatura em Matemática pela Universidade Estadual de Alagoas – UNEAL.

² Professor assistente UNEAL – Campus III; Doutorando Programa de Pós Graduação em Biometria e Estatística Aplicada – UFRPE.

³ Unidade Acadêmica de Sistemas e Computação – UASC; Centro de Engenharia Elétrica e Informática – CEEI; Universidade Federal de Campina Grande – UFCG.

with a standard deviation of 0.93 and a coefficient of variation of 33.6%. Box-Jenkins methodology was employed to analyse the univariate time series, i.e., that methodology was used to analyse the values of a single time series. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were employed to verify the best model. Besides Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) were also employed. Forecasting was performed by Autoregressive Integrated Moving Average (ARIMA), and after analysis of four suggested models, this model was adjusted to ARIMA(1, 1, 0). The Mean Squared Error (MSE) was used to quantify the model error and to verify the choice of the best model. The obtained MSE was 0.0563. Finally, forecasting for 12 samples were generated and compared with the 12 samples, it should be emphasized that the 12 forecasted samples were under the confidence limit.

Keywords: DescriptiveStatistic. ARIMA Models. Time series forecast. Wind Speed.

Introdução

Devido ao fato da água estar cada vez mais escassa, procura-se uma alternativa para a geração de energia elétrica que não seja por meio de usinas hidráulicas. Sendo essa nova fonte de energia, ao mesmo tempo, renovável, econômica e menos agressiva ao meio ambiente. Com essa perspectiva, está se expandindo a implantação de parques eólicos, com destaque no Brasil para a região Nordeste. Entre os parques eólicos instalados no Brasil temos o Complexo Eólico do Alto Sertão I (BA), o Parque Eólico Enacel (CE) e o Complexo Eólico Ventos de São Clemente (PE).

Em 2001 foi publicado o Atlas do Potencial Eólico Brasileiro, onde foi demonstrado que o Brasil possui valores consideráveis para a geração de energia eólica, sendo a região Nordeste uma das mais promissoras. A partir disso, vários estados interessaram-se e realizaram pesquisas que resultaram na elaboração de atlas desses estados. No estado de Alagoas o atlas foi publicado em 2008, resultado de um convênio entre a Eletrobrás (Centrais Elétricas Brasileiras) e o Estado de Alagoas, com a participação do Instituto de Tecnologia para o Desenvolvimento (LACTEC) e da Universidade Federal de Alagoas (UFAL).

Durante a elaboração do atlas foram analisadas as medições geradas pelas torres anemométricas instaladas nas localidades de Feliz Deserto, Roteiro, Maragogi, Palmeira dos Índios, Água Branca e Girau do Ponciano, pois situam-se em locais de diferentes relevo e vegetação. A seguir, podemos observar a figura da torre anemométrica instalada em Palmeira dos Índios:



Figura 1 - Estação Meteorológica de Palmeira dos Índios. Fonte: Portal INMET.

No município de Palmeira dos Índios tem duas estações meteorológicas, uma convencional e outra automática, ambas estão a uma latitude aproximada de $-9,420334^\circ$ e longitude $-36,620371^\circ$, e altitudes 274,9m e 278m, respectivamente.

Diz-se que para ser viável, a energia eólica deve ter seu potencial maior ou igual a 500 W/m^2 , a uma altura de 50m, o que requer uma velocidade mínima do vento de 7,0 m/s a 8,0m/s. E, frequentemente, a geração de energia elétrica se inicia a partir de 3,0m/s, abaixo disso fica inviável o potencial energético produzido (COSTA; LYRA, 2012).

O objetivo deste trabalho é verificar a viabilidade das condições do vento para a localidade de Palmeira dos Índios (AL). Diante disso, apresenta-se por meio de um estudo teórico, uma análise descritiva da série temporal e realização de modelagem matemática através de métodos estatísticos e dos modelos Box & Jenkins, conhecidos por modelos ARIMA (Auto - Regressivo Integrado de Média Móvel).

Utilizou-se o *software* MATLAB® R2016a (*Matrix Laboratory*) para realizar os testes e verificação dos modelos Box-Jenkins, geração dos gráficos e geração das previsões.

Metodologia

A série temporal utilizada foi adquirida do site do INMET (2016), no Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP). Destacando que são dados mensais de velocidades média do vento da Estação Meteorológica de Palmeira dos Índios (AL) e compreende o período de janeiro/2010 até outubro/2016, sendo que nesse período foram registradas 82 medições mensais. Tal torre está localizada próxima à divisa com o estado de Pernambuco, e tem 100m de altura, com anemômetros em 03 níveis (30m, 70m e 100m) e sensores de direção do vento (30m e 70m) (ELETROBRAS, 2008).

De acordo com Morettin e Tolo (1981), uma série temporal é qualquer conjunto de observações ordenadas no tempo. E fazem parte das séries temporais os seguintes elementos: a tendência, o ciclo, a sazonalidade e o ruído aleatório ou erro.

Segundo Pellegrini (2000), os dados apresentados como séries temporais podem ser representados por meio de modelagem matemática, podendo dessa forma, serem utilizados para gerar previsão. Tais modelos são considerados determinísticos ou estocásticos, serão determinísticos aqueles cuja previsão resultou em valores exatos, e estocásticos aqueles que utilizam-se de modelos probabilísticos para gerar previsões futuras. Os modelos estocásticos podem ser estacionários ou não estacionários. Segundo Pellegrini (2000, *apud* BOX *et al.*, 1994), diz-se estacionário, o processo onde há equilíbrio, que é aquele onde as

variáveis mantêm-se em um nível constante médio, isto é, a média é zero e a variância é constante. Conforme Cavalcanti (2007), é interessante que a série analisada seja estacionária, pois assim será maior a possibilidade de fixar medidas para a previsão com base no passado. Na Figura 2, apresenta-se a série temporal dos valores das médias mensais da velocidade do vento no período de janeiro/2010 a outubro/2016 na cidade de Palmeira dos Índios (AL).

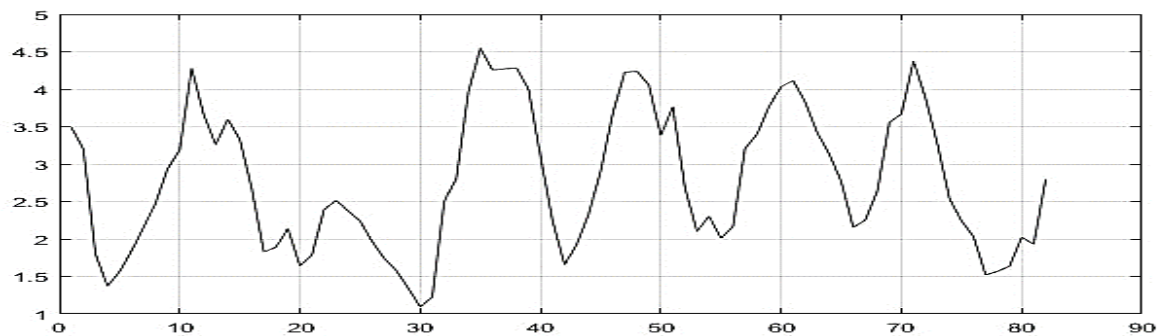


Figura 2: Gráfico dos dados médios mensais da série temporal coletados no período de jan/2010 a out/2016. Fonte: Autores, 2017.

Observando a figura 2, percebe-se que a série temporal em estudo segue um padrão não estacionário. Sendo assim, é necessário torná-la estacionária. E para isso, Gujarati e Porter (2011) citam, que podemos verificar a não estacionariedade da série pelo modelo auto-regressivo seguinte:

$$Y_t = \rho Y_{t-1} + e_t \quad -1 \leq \rho \leq 1 \quad (1)$$

No qual, e_t é o termo de erro estocástico ou ruído branco. Sendo que, a raiz unitária existe quando $\rho = 1$. Já, para $|\rho| < 1$ a série temporal Y_t é estacionária.

Segundo Gujarati e Porter (2011), devemos realizar a subtração de Y_{t-1} de ambos os lados da Eq. (1), para tornar a série estacionária, assim, teremos:

$$\begin{aligned} Y_t - Y_{t-1} &= \rho Y_{t-1} - Y_{t-1} + e_t \\ Y_t - Y_{t-1} &= (\rho - 1) Y_{t-1} + e_t \\ \Delta Y_t &= \delta Y_{t-1} + e_t \end{aligned} \quad (2)$$

Em que:

$$\delta = (\rho - 1)$$

Δ = operador de diferença

Assim, de acordo com os autores, estimamos a Eq. (2), testamos as hipóteses nula (H_0) e alternativa (H_1), em que:

$$\begin{aligned} H_0 : \rho = 1 &\leftrightarrow H_0 : \delta = 0 \\ H_1 : \rho < 1 &\leftrightarrow H_1 : \delta < 0 \end{aligned} \quad (3)$$

Como, para $\delta = 0$, $\rho = 1$. Então, há uma raiz unitária e a série é não estacionária. Nesse caso, como $\delta = 0$, a Eq. (2), ficará assim:

$$\Delta y_t = (y_t - y_{t-1}) = e_t \quad (4)$$

Conforme dito anteriormente, e_t é o termo de erro de ruído branco, sendo ele estacionário, significando que “as primeiras diferenças de uma série temporal de passeio aleatório são estacionárias” (GUJARATI; PORTER, 2011).

Vários são os testes de raiz unitária existentes na literatura específica, e neste trabalho optou-se pela utilização do Teste de Dickey-Fuller Aumentado (ADF), concomitantemente com o Teste de Phillips-Perron (PP). Sendo que, foram tabulados por Dickey e Fuller, valores críticos por simulações de Monte Carlo, e com isso desenvolvida uma estatística τ (tau) para realizar o teste de raiz unitária.

Em conformidade com Cazarotto (2006, apud Dickey–Fuller, 1979) tem-se que, a estatística gerada deverá ser comparada com tais valores críticos.

E segundo o Portal Action (2017), a estatística té calculada por:

$$T = \frac{\hat{\delta}}{se(\hat{\delta})} \quad (5)$$

Em que:

$\hat{\delta}$ = estimador para δ

$se(\hat{\delta})$ = estimador para o desvio padrão do erro de δ .

Assim, para que a hipótese nula $H_0 : \delta = 0$ seja rejeitada, o p-valor (valor da probabilidade) tem que ser menor que 5%, isto é, menor que 0,05 (GUJARATI; PORTER, 2011). Além disso, ainda concordando com Gujarati e Porter (2011), quando o valor da estatística τ aumenta, o p-valor diminui. Sendo, com isso, possível rejeitar a hipótese nula H_0 com maior confiança.

O Teste Dickey-Fuller Aumentado (ADF), é estimado conforme Gujarati e Porter (2011), através da seguinte regressão:

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^m \alpha_i \Delta Y_{t-i} + e_t \quad (6)$$

Em que:

e_t = termo de erro de ruído puro

$\Delta Y_{t-1} = (Y_{t-1} - Y_{t-2})$

$\Delta Y_{t-2} = (Y_{t-2} - Y_{t-3})$, assim sucessivamente...

Para realizar o teste, estima-se a equação (6) e em seguida a estatística τ (tau) e por fim, é consultada a tabela dos valores críticos de Dickey-Fuller. Com isso, analisa-se se o valor da estatística τ é menor do que o valor crítico tabelado, caso seja, a hipótese nula $H_0: \delta = 0$ é rejeitada, o que torna a série temporal estacionária, pois não possui raiz unitária. Porém, se o valor da estatística τ é maior do que o valor crítico tabelado, a hipótese nula $H_0: \delta = 0$ não será rejeitada e a série temporal será não estacionária.

No teste PP, emprega-se os métodos estatísticos não paramétricos para abordar a correlação serial dos termos de erro sem inserir os termos de diferença defasados. Diante disso, Alves (2010) cita que o teste PP poderá ser aplicado quando os resíduos estiverem correlacionados serialmente.

Assim, conforme Alves (2010, *apud* Cirino e Lima, 2006), o teste PP é expresso por:

$$Z\delta = t\beta \left(\frac{\gamma_0}{f_0}\right)^{\frac{1}{2}} - \frac{T(f_0 - \gamma_0)S(\hat{\beta})}{2f_0^{\frac{1}{2}}S} \quad (7)$$

Em que:

$\hat{\beta}$ = estimador do parâmetro

$S(\hat{\beta})$ = estimador do erro padrão do parâmetro

γ_0 = estimador da variância do resíduo

t = número de observações

f_0 = estimador do espectro do resíduo na frequência zero

S = erro padrão da regressão do teste

Evidencia-se que, de acordo com Alves (2010) o teste PP também testa a hipótese nula $H_0: \delta = 0$. Assim, tem-se que na estatística $Z\delta$ pode ser empregada os mesmos valores críticos do teste ADF. Com isso, caso a hipótese nula $H_0: \delta = 0$ seja aceita, a série será não estacionária, do contrário, ela será estacionária.

Análise Estatística Descritiva

Os dados da série foram modelados através da aplicação de conceitos da estatística descritiva para analisar como eles se comportam, dentre os conceitos aplicados encontram-se o de média aritmética, utilizado para calcular a média geral da velocidade do vento no período estudado (CRESPO, 2002); o de variância, usado para verificar o grau de variabilidade dos valores em estudo (IEZZI, 2004); o de desvio padrão para comparar diferenças entre conjuntos de valores (IEZZI, 2004).

Além do coeficiente de variação (CV) que é utilizado na comparação de duas ou mais séries de valores, e é definido pela razão entre o desvio padrão e a média aritmética (CRESPO, 2002).

Modelagem ARIMA

Em Werner e Ribeiro (2003, *apud* Fava, 2000) tem-se que os modelos ARIMA são formados pela combinação de três elementos que são denominados de “filtros”, são eles: elemento auto-regressivo (AR), filtro de integração (I) e o elemento de médias móveis (MA). Sendo que uma série pode ser composta pelos três filtros ou pelas subclasses deles. Os modelos Box-Jenkins também conhecidos por modelos ARIMA (Auto-regressive Integrated Moving Averager), isto é, Auto-regressivos Integrados de Médias Móveis, são utilizados em análises de séries univariadas de tempo, ou seja, o método é empregado para analisar valores de uma única série temporal, para fins de previsão. A partir disso, Albuquerque e Moraes (2006) citam que, o método ARIMA é subdividido nas etapas seguintes: especificação matemática do modelo (AR, MA, ARMA, ARIMA) – onde AR significa modelo Auto-regressivo de ordem “p”, MA significa modelo de Médias Móveis de ordem “q”, e ARMA significa modelos Auto-regressivos de Médias Móveis de ordem “p” e “q”; identificação do modelo – obtenção de (p, d, q) – e FAC e FACP, que são as Funções de Autocorrelação e Função de Autocorrelação Parcial; estimação (existem vários critérios para seleção dos modelos, neste trabalho utilizaremos os critérios AIC – critério de informação Akaike –, e BIC – critério de informação Bayesiano) e posteriormente será feita a seleção dos parâmetros; análise do erro quadrado médio

(EQM), que serve para analisar o ajustamento do modelo aos dados da série; checagem do modelo por meios estatísticos e previsões cabíveis.

Nos Modelos Auto-regressivos de ordem p (AR (p)) a série temporal Y_t é formada por valores passados regredidos e pelo ruído aleatório e_t . Onde “p” indica a ordem do modelo, ou seja, o número de defasagens aplicadas. E é dado por:

$$\tilde{Y}_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t \quad (8)$$

Onde, $\tilde{Y}_t = Y_t - e_t$. Sendo que, ϕ_i é o parâmetro que estabelece a relação de \tilde{Y}_t com o valor Z_{t-i} para $i = 1, 2, \dots, p$.

Já nos Modelos de Médias Móveis de ordem q (MA (q)) a série \tilde{Y}_t é resultado de combinações entre os ruídos brancos e os períodos atuais e passados. Diante disso, tem-se que:

$$\tilde{Y}_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} \quad (9)$$

No qual, $\tilde{Y}_t = Y_t - e_t$, “q” representa a ordem do modelo, θ_i é o parâmetro que descreve como \tilde{Z}_t se relaciona com o valor e_{t-i} , para $i = 1, 2, \dots, q$.

Os Modelos Auto-regressivos de Médias Móveis de ordem p e q (ARMA (p, q)) surgem da mistura de um componente AR (p) com outro MA (q), dessa forma o modelo ARMA (p, q) requer um número menor de termos, os quais podem ser referidos na seguinte equação:

$$\tilde{Y}_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \mu_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (10)$$

Para os Modelos Auto-regressivos Integrados de Médias Móveis (ARIMA (p, d, q)) temos que, caso as séries temporais sejam não estacionárias devemos transformá-las em estacionárias. Desse modo, Werner e Ribeiro (2003, *apud* Morentin & Tolo, 1987), mencionam que tal transformação compreende em tomar diferenças sucessivas da série original até conseguir uma série estacionária. Então, a primeira diferença de Y_t é dada por:

$$\Delta Y_t = Y_t - Y_{t-1} \quad (11)$$

E a segunda, por:

$$D^2 Y_t = D [D Y_t] = D [Y_t - Y_{t-1}] = Y_t - 2Y_{t-1} - Y_{t-2} \quad (12)$$

Ainda concordando com os autores citados, tomar uma ou duas diferenças, geralmente, será o suficiente para tornar a série estacionária. Desse modo, denomina-se ordem de integração, o número “d” de diferenças. Com isso, incluindo o termo de ordem de integração, podemos utilizar os modelos ARIMA (p, d, q) da seguinte forma:

$$W_t = \phi_1 W_{t-1} + \dots + \phi_p W_{t+p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (13)$$

Onde, $W_t = \Delta^d Y_t$.

Segundo Souza (2006), seja o coeficiente de autocorrelação de ordem k, definido por:

$$\rho_k = \frac{\text{cov}(y_t, y_{t-k})}{v(y_t)} = \frac{\gamma_k}{\gamma_0} \quad (14)$$

Assim, a função de autocorrelação (FAC) define-se pela sequência de pares (k, ρ_k), k= 1, 2, ... Vale ressaltar, que os valores negativos de k não são considerados expressamente. De acordo com Gujarati e Porter (2011), ρ_k é um número sem unidades ou puro, e situa-se entre -1 e +1.

Seja o coeficiente da função de autocorrelação parcial (FACP) de ordem k, frequentemente simbolizado por ϕ_{kk} , este, mede a correlação entre Y_t e Y_{t-k} , após a influência de Y_{t-1} , Y_{t-2} , ..., Y_{t-k+1} , sobre Y_t .

Souza (2006) diz que, “o coeficiente ϕ_{jj} , j = 1, 2,... é dado pelo último coeficiente β_{jj} de cada uma das auto-regressões” a seguir:

$$\begin{aligned} Y_t &= \beta_{11} Y_{t-1} + \mu_t \rightarrow \phi_{11} = \beta_{11} \\ Y_t &= \beta_{11} Y_{t-1} + \beta_{22} Y_{t-2} + \mu_t \rightarrow \phi_{22} = \beta_{22} \\ &\vdots \\ Y_t &= \beta_{k1} Y_{t-1} + \beta_{k2} Y_{t-2} + \dots + \beta_{kk} Y_{t-k} + \mu_t \rightarrow \phi_{kk} = \beta_{kk} \end{aligned} \quad (15)$$

Conforme Costa (2009), a autocorrelação de uma dada variável é definida pela distância ou *lag* (atraso) com que se deseja medi-la. Pinto (2006) cita que observando a FAC é possível identificar somente o parâmetro q, que é a ordem do modelo MA(q). E para identificação dos outros modelos (AR (p) e ARMA (p, q)) analisa-se a FACP. Vale lembrar, que o modelo ARIMA (p, d, q) é definido inicialmente com a identificação do parâmetro d, que segundo Pinto (2006) pode ser feita analisando a estacionariedade da série temporal ou analisando as FAC e FACP após a quantidade de d diferenciações.

Para identificação dos critérios de seleção de modelos, utilizamos os denominados critérios de informação *Akaike* – AIC, e critério de informação *Bayesiano* – BIC. Tais critérios são baseados na variância estimada $\hat{\mu}_t$, no tamanho da amostra e nos valores de p e q. Tem-se ainda, que são estimados vários pares (p, q) e escolhido o par de menor valor para AIC e para BIC (SOUZA, 2006). Esses critérios são dados por:

$$AIC = \ln (\hat{\sigma}_3^2) + \frac{2(p + q)}{n} \quad (16)$$

$$BIC = \ln (\hat{\sigma}_3^2) + \frac{(p + q)\ln (n)}{n} \quad (17)$$

É importante frisar que os critérios de informação são usados para complementar a FAC e a FACP (SOUZA, 2006).

Para verificar a escolha do melhor modelo e quantificar o erro, utilizou-se o erro quadrático médio (EQM), e segundo Pasquotto (2010), dada a série temporal Y_t , no instante t, tem-se que o erro de previsão e_t para os próximos instante t+1, t+2, ..., t+h, será compreendido pela diferença entre o valor estimado e o valor real da série. Assim sendo, o erro total de previsão é dado por:

$$E = \sum_{i=1}^h e_{t+i} \quad (18)$$

Onde e_{t+i} é o erro de previsão um passo à frente.

Pasquotto (2010), salienta ainda, que essa medida é pouco utilizada devido a sinais opostos dos erros e_{t+i} , o que pode causar a anulidade do somatório. Dessa forma, eleva-se cada erro ao quadrado e obtém-se, o erro quadrado médio (EQM) da previsão, o qual é dado por:

$$EQM = \frac{\sum_{i=1}^h e_{t+i}^2}{n} \quad (19)$$

Resultados e Discussão

A série temporal em estudo é composta por 82 amostras de dados, que foram obtidas no site do INMET e representam a velocidade do vento mensal do período compreendido entre Janeiro de 2010 a Outubro de 2016, na cidade de Palmeira dos Índios (AL). Desses dados, 70 amostras foram utilizadas para realizar os testes, ficando as outras 12 amostras para comparação com os resultados que serão gerados na previsão.

Primeiramente, utilizou-se a análise descritiva para mostrar o comportamento dos dados da série em estudo, para tal, calculou-se as medidas descritivas de velocidade mínima e máxima, média, mediana, desvio padrão, variância e coeficiente de variação. Tais valores podem ser observados na Tabela 1.

Tabela 1 - Medidas descritivas dos dados da série temporal

<i>Estatísticas</i>	<i>Valores</i>
Vel. Máxima	4,55
Vel. Mínima	1,09
Média	2,78
Mediana	2,65
Desvio Padrão	0,93
Variância	0,87
Coef. de Variação	33,6%

Fonte: Autores (2017).

Observa-se na tabela 1, que a média geral foi igual a 2,78, variando entre 4,55m/s (vel. máx.) e 1,09m/s (vel. mín.), com um desvio padrão de 0,93, variância de 0,87 e coeficiente de variação de 33,6%. Segundo Correa (2003), o coeficiente de variação encontrado é de alta dispersão, pois é maior que 30%.

Em seguida, analisando o gráfico da figura 2, observa-se que a série em questão não é estacionária, pois as médias da velocidade têm altas variações no decorrer do tempo. E para confirmar essa hipótese, foram aplicados os testes de raiz unitária ADF e PP, que confirmaram a não estacionariedade da série, pois em ambos os testes o p-valor foi igual a 0,3950. E como citado anteriormente, para aceitar a hipótese de estacionariedade, o *p-valor* teria que ser menor que 5%, ou seja, menor que 0,05.

Com isso, foi necessário aplicar diferenciações a série para amenizar o efeito de não estacionariedade, e após a primeira diferenciação foram realizados novamente os testes ADF e PP, e desta vez verificou-se a estacionariedade da série. Como o valor da ordem de integração da série equivale a quantidade de diferenciações, assim, encontramos o valor de d que é igual a um ($d = 1$).

Como pretende-se encontrar um modelo ARIMA adequado, analisamos a seguir, os gráficos da FAC e FACP da série diferenciada, figuras 2 e 3 respectivamente:

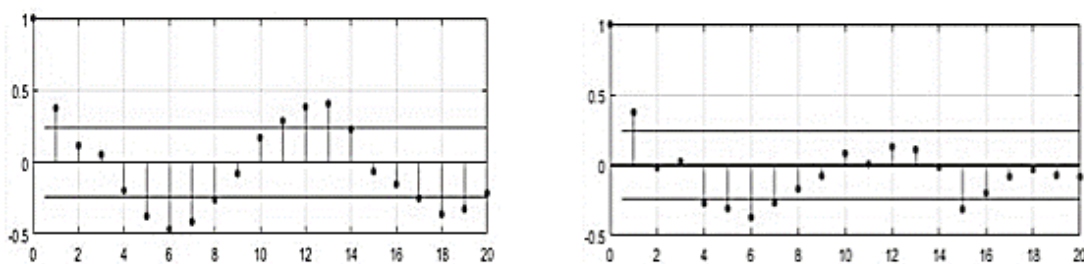


Figura 3 – FAC. Fonte: Autores (2017). Figura 4 – FACP. Fonte: Autores (2017).

Observando as figuras, temos que na figura 3, há um corte no primeiro lag da FAC, o que sugere um modelo MA (1). A figura 4, mostra-nos que há um corte no primeiro lag da FACP, só que neste caso, a suposição é de um modelo AR (1) ou um modelo ARMA (1, 0).

Sendo o modelo ARIMA composto pela combinação de modelos, há a necessidade de estimar os parâmetros de todos os modelos possíveis e para isso são utilizados os critérios AIC e BIC, como mostra a tabela 2:

Tabela 2 - Parâmetros estimados dos critérios AIC e BIC

	<i>AIC</i>	<i>BIC</i>
<i>ARIMA (0, 1, 0)</i>	110,1083	112,3424
<i>ARIMA (1, 1, 0)</i>	101,5416	106,0098
<i>ARIMA (0, 1, 1)</i>	101,8879	106,3562
<i>ARIMA (1, 1, 1)</i>	103,5266	110,2289

Fonte: Autores (2017).

O modelo escolhido foi o modelo ARIMA (1, 1, 0), pois é o que se refere aos menores valores para AIC e BIC. Ainda assim, deve ser comprovada a adequabilidade do modelo ARIMA (1, 1, 0) através da análise de resíduos para constatar se o modelo é bem

ajustado a série de dados real, pois, os critérios AIC e BIC são apenas indicativos de tal modelo.

Verifica-se na Figura 5, o gráfico dos resíduos e observa-se que os dados giram em torno de zero, indicando que o modelo é ajustado.

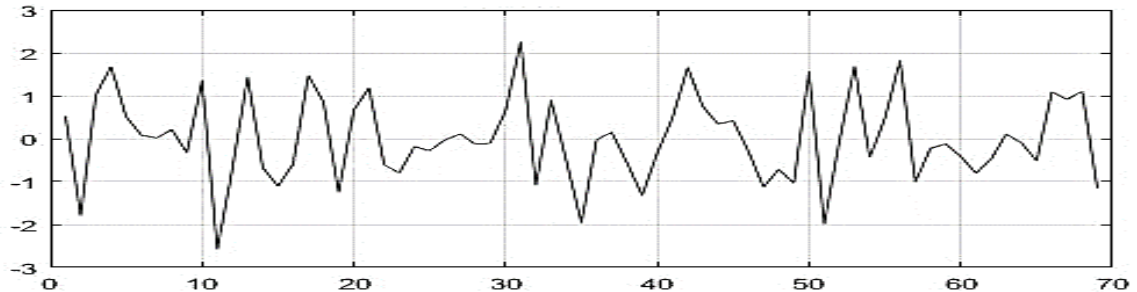


Figura 5: Resíduos padronizados. Fonte: Autores (2017).

Na figura 6, apresentamos o gráfico QQ PLOT da amostra de dados para verificação se os dados seguem uma distribuição normal.

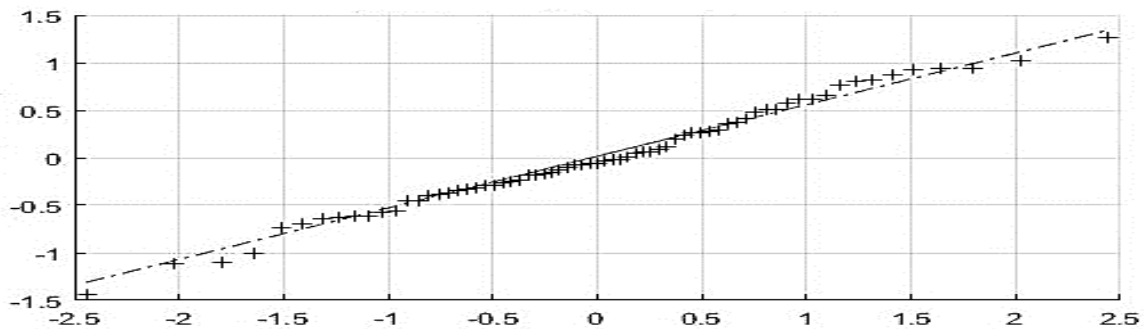


Figura 6 - Gráfico dos dados da amostra versus o padrão normal. Fonte: Autores (2017).

Considera-se que os ruídos, representados pelo sinal positivo (+), encontram-se quase sobrepondo-se a linha reta dos dados reais, neste caso, considera-se que os resíduos do modelo ARIMA (1, 1, 0) seguem uma distribuição normal.

Analisando as figuras 7 e 8, da FAC e FACP dos resíduos do modelo ARIMA (1, 1, 0), respectivamente, nota-se que na figura 7, assim como, na figura 8, os lags significativos encontram-se dentro do intervalo de confiança e não apresentam correlação serial.

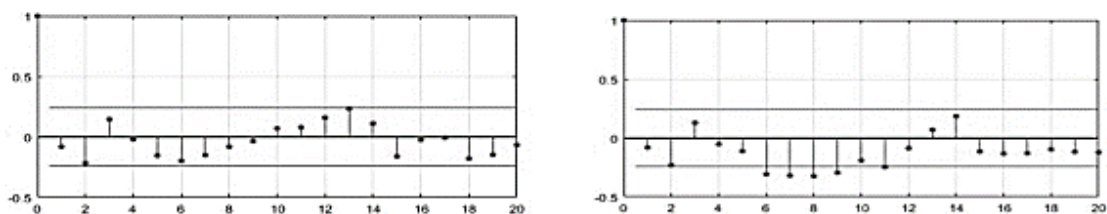


Figura 7(esquerda) – FAC dos resíduos do modelo ARIMA (1, 1, 0). Figura 8(direita) – FACP dos resíduos do modelo ARIMA (1, 1, 0). Fonte: Autores (2017).

Realizou-se, em seguida, a previsão da amostra para os meses de nov/2015 a out/2016 (12 observações) como observamos na figura 9:

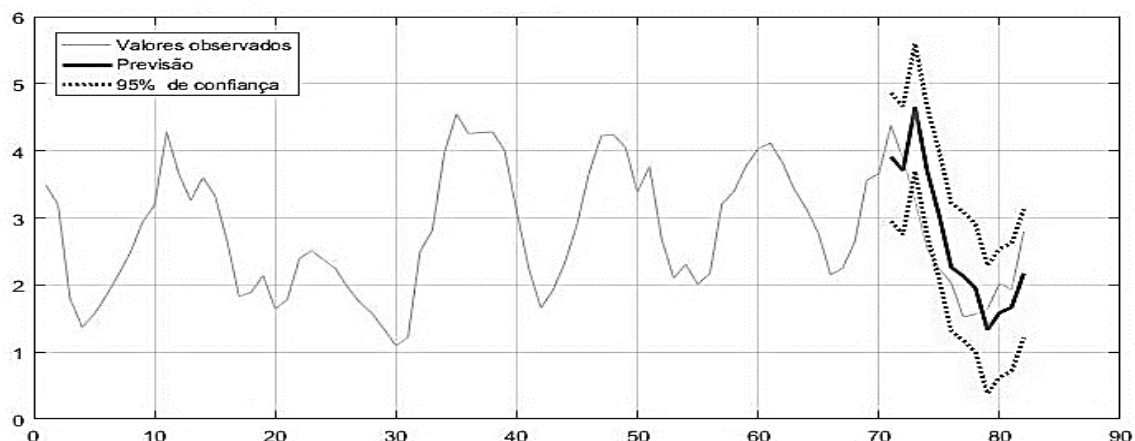


Figura 9 - Previsão e 12 observações com modelo ARIMA (1, 1, 0) dos dados médios mensais com intervalo de confiança. Fonte: Autores (2017).

Destaca-se que, os valores previstos apresentam-se semelhantes à série original e dentro dos limites de confiabilidade, com EQM do modelo estimado igual a 0,0563.

Conforme observa-se na Figura 9, conseguiu-se estimar um modelo de séries temporais entre os modelos ajustados. Destacou-se que o modelo auto-regressivo de ordem 1 com uma diferença, ARIMA (1, 1, 0), foi o que explicou melhor as condições de velocidades do vento para a cidade de Palmeira dos Índios, entre os modelos ARIMA selecionados.

Considerações finais

A energia eólica está despertando cada vez mais o interesse dos investidores, uma vez que, este tipo de energia é renovável, sustentável e inesgotável.

Neste trabalho, objetivou-se fazer um estudo sobre as condições de velocidade do vento para a cidade de Palmeira dos Índios (AL), utilizando o método Box-Jenkins.

Diante do exposto, dentre os modelos possíveis encontrados, utilizou-se o modelo ARIMA (1, 1, 0), pois de acordo com os critérios AIC e BIC, este foi o que apresentou menor valor, sendo desta forma, o que melhor se ajusta.

Em relação às estatísticas calculadas da série, destacamos que a média durante o período analisado foi de 2,78, oscilando entre 4,55m/s (vel. máx.) e 1,09m/s (vel. mín.). O desvio padrão foi de 0,93 e o coeficiente de variação 33,6%.

Ressalta-se sobretudo, que de acordo com a Figura 9, após a previsão das 12 amostras, os dados da previsão apresentaram-se, de forma geral, dentro dos limites de confiabilidade, e que os resíduos do modelo mostraram-se com características de ruído branco, proporcionando assim, ótimo ajuste.

Como o valor da média geral foi menor que 3m/s, concordando com Costa e Lyra (2012), é interessante a utilização de outras metodologias para verificar a viabilidade de produção de energia com os ventos dessa região já que a média da velocidade obtida no período da pesquisa foi de 2,78m/s, ressaltando que a pesquisa relatada neste trabalho não é conclusiva.

Referências

- ALVES, Greiciano José; COELHO, Alexandre Bragança; GONÇALVES, Lílian Valeriano. *Efetividade do hedge no mercado de café arábica para as praças de Caratinga – MG e São Sebastião do Paraíso – MG*. In: 48º Congresso da Sociedade Brasileira de Economia, Administração e Sociologia Rural – SOBER, Campo Grande, *Anais...* Campo Grande, MS: SOBER, 2010. 16 p.
- CAVALCANTI NETTO, M. A. *A previsão com a metodologia de Box-Jenkins*. 2007. (Desenvolvimento de material didático ou instrucional - Material didático).
- CAZAROTTO, Simone. *Teste de raiz unitária em modelo painel: uma aplicação a teoria da paridade real de juros na América Latina*. 2006. 71 f. Dissertação (Mestrado em Economia) – Universidade Federal de Santa Catarina, Programa de Pós-Graduação em Economia, Florianópolis, 2006.
- CORREA, Sonia M. B. B. *Probabilidade e Estatística*. 2 ed. Belo Horizonte: PUC Minas Virtual, 2003. 116 p.
- CRESPO, Antonio A. *Estatística Fácil*. 18 ed. São Paulo: Saraiva, 2002.
- COSTA, G. B.; LYRA, R. F. da F. Análise dos padrões de vento do Estado de Alagoas. *Revista Brasileira de Meteorologia*, v. 27, n.1, 31 – 38, 2012.
- COSTA, Kleber Carlos de Oliveira. *Análise de DFA e de agrupamento do perfil de densidade de poços de petróleo*. 2009. 89 f. Dissertação (Mestrado em Engenharia de Petróleo) – Universidade Federal do Rio Grande do Norte, Natal, 2009.
- GUJARATI, Damodar N.; PORTER, Dawn C. *Econometria básica*. 5 ed. São Paulo: AMGH Editora Ltda, 2011.
- IEZZI, Gerson et al. *Coleção Matemática: Ciência e Aplicações*. 3v. 2 ed. São Paulo: Atual editora, 2004.
- INMET. *Instituto Nacional de Meteorologia*. Disponível em: <<http://www.inmet.gov.br>> Acesso em: 12 ago. 2016.
- MORAES, Marcel Castro de; ALBUQUERQUE, Aldous Pereira. *Previsão para o preço futuro do cacau através de uma série univariada de tempo: uma abordagem utilizando o método ARIMA*. In: 44º Congresso da Sociedade Brasileira de Economia, Administração e Sociologia Rural - SOBER, Fortaleza, *Anais...* Fortaleza, Ceará: SOBER, 2006. 17 p.
- MORETTIN, Pedro Alberto; TOLOI, Clélia Maria de Castro. *Modelos para previsão de séries temporais*. 1v. Rio de Janeiro: Impa, 1981.
- PASQUOTTO, Jorge Luis Durgante. *Previsão de séries temporais no varejo brasileiro: uma investigação comparativa da aplicação de redes neurais recorrentes de elman*. 2010. 200 f. Dissertação (Mestrado em Administração) – Universidade de São Paulo, Faculdade de Economia, Administração e Contabilidade, São Paulo, 2010.

- PELLEGRINI, Fernando Rezende. *Metodologia para implementação de sistemas de previsão de demanda*. 2000. 146 f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal do Rio Grande do Sul, Escola de Engenharia, Porto Alegre, 2000.
- PINTO, Rodrigo Bastos. *Finanças comportamentais e metodologia Box & Jenkins: uma aplicação no mercado brasileiro*. 2006. 81 f. Dissertação (Mestrado em Administração) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Administração, Rio de Janeiro, 2006.
- PORTAL ACTION. *Teste de Phillips-Perron*. Disponível em: <<http://www.portalaction.com.br/series-temporais>>. Acesso em: 05 jan. 2017.
- SOUZA, Francisca Mendonça. *Modelos Box Jenkins aplicado a previsão de demanda de leitos hospitalares*. 2006. 83 f. Monografia (Especialização em Estatística e Modelagem Quantitativa) – Universidade Federal de Santa Maria, RS, Centro de Ciências Naturais e Exatas, Santa Maria, 2006.
- WERNER, Liane; RIBEIRO, José Luis Duarte. Previsão de demanda: uma aplicação dos modelos Box-Jenkins na área de assistência técnica de computadores pessoais. *Gestão e produção*, v. 10, n.1, 47 – 67, 2003.