Revista Brasileira de Climatologia
ISSN: 2237-8642 (Eletrônica)

MODELAGEM DA VELOCIDADE DO VENTO USANDO METODOLOGIAS ARIMA, HOLT-WINTERS E RNA NA PREVISÃO DE GERAÇÃO EÓLICA NO NORDESTE BRASILEIRO

CAMELO, Henrique do Nascimento – henriquecamelo13@yahoo.com.br
Universidade Federal do Rio Grande do Norte/UFRN

LUCIO, Paulo Sergio – pslucio@ccet.ufrn.br
Universidade Federal do Rio Grande do Norte/UFRN

LEAL JÚNIOR, João Verçosa – jblealjr@uece.br
Universidade Estadual do Ceará/UECE

RESUMO: O presente artigo mostra metodologias capazes de realizar previsões de velocidade do vento com boa acurácia as quais poderão ser aplicadas no setor de geração eólica. Para isso, dados de velocidade dos ventos das regiões de Fortaleza, Parnaíba e São Luís localizadas no Nordeste Brasileiro (NEB) foram aplicados aos modelos de previsão de séries temporais - Auto Regressivo Integrado de Médias Móveis (ARIMA), Holt-Winters (HW) e Redes Neurais Artificiais (RNA), e comparados entre si com a finalidade de identificar o de melhor performance. O período compreende janeiro de 2010 até dezembro 2014. O modelo RNA apresentou os melhores valores das medidas de acurácia, principalmente o erro percentual entre as séries observadas e simuladas cujos valores foram de 8,74% (Fortaleza) e 10,50% (São Luís). A localidade de Parnaíba os valores dos erros foram menores para o modelo ARIMA, com erro percentual de 9,70%. As séries temporais previstas pelos modelos conseguem representar características importantes, particularmente, as menores intensidades da média mensal da velocidade do vento (período chuvoso), seguido das maiores intensidades da média mensal da velocidade do vento (período seco).

Palavras- chaves: Séries temporais, NEB, energias renováveis, inteligência artificial.

WIND SPEED MODELING USING ARIMA, HOLT-WINTERS AND RNA METHODOLOGIES IN THE FORECAST OF WIND GENERATION IN NORTHEAST BRAZIL

ABSTRACT: The present article shows methodologies capable of performing predictions of wind speed with good accuracy which can be applied in the wind power generation sector. For this, wind velocity data from the Fortaleza, Parnaíba and São Luís regions located in the Brazilian Northeast (NEB) were applied to the time series forecast models - Integrated Automatic Regressive of Moving Averages (ARIMA), Holt-Winters (HW) and Artificial Neural Networks (RNA), and compared them with the purpose of identifying the one with the best performance. The period from January 2010 to December 2014. The RNA model presented the best values of the accuracy measurements, mainly the percentage error between the observed and simulated series, whose values were 8.74% (Fortaleza) and 10.50% (São Luís). The locality of Parnaíba was lower for the ARIMA model, with a percentage error of 9.70%. The time series predicted by the models can represent important characteristics, particularly the lower intensities of the monthly average wind speed (rainy season), followed by the higher monthly average wind speed (dry period).

Keywords: Time series, NEB, renewable energy, artificial intelligence.

1. INTRODUÇÃO

As pesquisas em energias renováveis são necessárias no Desenvolvimento Sustentável (DE). Essas formas de energias, que são

consideradas não-poluidoras da atmosfera, podem atenuar as mudanças climáticas, as quais muitos cientistas entendem que estão diretamente relacionadas ao aumento das emissões antropogênicas dos gases de efeito estufa (GEE), tais como, CO2, NOx e SOx, que causam efeitos negativos no clima como, por exemplo, o aumento do efeito estufa, chuvas ácidas, e a degradação da camada de ozônio (O3) (AKELLA et al., 2009; FRONDEL et al., 2010; PANWAR et al., 2011). As energias renováveis podem ser consideradas fontes inesgotáveis, limpas e livres. Os benefícios da utilização destas fontes são principalmente de ordem econômica e ambiental. Sendo assim, fazer com que as futuras gerações possam viver em um mundo saudável é um papel fundamental da humanidade, e para isso é preciso investir cada vez mais em pesquisas sobre energias limpas e renováveis, como uma forma de se alcançar o DE (BILGEN et al., 2008; RESCH et al., 2008; KAYGUSUZ et al., 2012).

Uma forma de energia renovável que vem favorecendo o crescimento da inserção de fontes limpas em matrizes energéticas de diversas nações trata-se da geração eólica. Por exemplo, atualmente o Brasil é considerado o maior gerador de energia eólica da América Latina de acordo com ranking divulgado em 2016 pela Global Wind Energy Council (GWEC), organização internacional especializada em energia eólica, que posicionou também o país na 5ª posição no ranking mundial de capacidade acumulada em aproximadamente 11 GW (GWEC, 2016). O Brasil é possível considerado um dos países que mais apresenta capacidade de produzir eletricidade a partir de fontes renováveis de energia. Exemplo disso, em sua matriz energética a principal fonte é hidráulica com aproximadamente 70%, de acordo com o Balanço Energético Brasileiro de 2014 (EPE, 2014). No contexto de potencial eólico, se estima, segundo o Altas do Potencial Eólico Brasileiro, cerca de 143 mil MW de potência eólica (CEPEL, 2010), dos quais 51% encontra-se sobre a região nordeste do país. De todo este potencial, o Estado do Ceará detém aproximadamente 18 mil MW de potência eólica aproveitável e médias anuais da velocidade do vento, a 50 m de altura, da ordem de 8,5 m/s (SEINFRA, 2001), o que tem feito com que o Estado receba bastante atenção de empresas do setor eólico. Esta média anual da velocidade do vento é bastante significativa quando comparada com as médias encontradas no Atlas Europeu do Vento, em especial na Alemanha, onde a velocidade média dos ventos é da ordem de 5 a 6 m/s no mesmo nível de altura (TROEN e PETERSEN, 1989).

Nos últimos anos têm sido realizadas pesquisas para fornecer a previsão de séries temporais da velocidade do vento para fins de geração de energia. Por exemplo, em Kavasseri e Seetharaman (2009), utilizaram o modelo Auto Regressivo Integrado de Médias Móveis (ARIMA), o qual também é conhecido por modelagem Box-Jenkins (BOX e JENKINS, 1978), para a previsão da velocidade do vento diária em quatro localidades da região de Dakota do Norte, nos EUA. Outra metodologia que vem sendo estudada para previsão de velocidade do vento é a utilização de Redes Neurais Artificiais (RNA) como, por exemplo, em (CADENAS e RIVERA, 2010), os autores realizaram previsão horária na região de La Venta, Oaxaca, México. Outro modelo bastante utilizado para previsão de velocidade do vento é a Suavização Exponencial (SE), o qual trabalha com médias ponderadas e fornece pesos maiores às observações mais recentes como, por exemplo, em Cadenas et al. (2010) os autores mostraram que a SES fornece menores erros quando comparado aos modelos ARIMA e RNA aplicados em região litorânea do México. Em Camelo et al. (2015) é possível

ISSN: 2237-8642 (Eletrônica)

identificar utilização do modelo Auto – Regressivo de Médias Móveis (ARMA) um modelo da classe do ARIMA, para obter previsão de velocidade do vento de boa acurácia em região serrana no Nordeste Brasileiro (NEB). Camelo et al. (2016) realizaram previsão de velocidade do vento em região litorânea do NEB com o uso do modelo Holt-Winters, o qual representa uma forma dos modelos da categoria de SE que leva em consideração características de tendência e sazonalidade comuns em séries temporais. Camelo et al. (2017) utilizaram outra categoria do modelo ARIMA, o qual inclui variáveis que podem se relacionar com a velocidade do vento para obter melhor acurácia na previsão desta variável, trata-se do modelo Auto-Regressivo Integrado de Médias Móveis com entradas Exógenas (ARIMAX). Neste estudo os autores mostraram que a inclusão da precipitação no modelo ARIMAX favorece previsões da velocidade do vento com maior precisão baseado em estudo de caso de regiões do NEB.

Este trabalho teve como objetivo comparar os modelos matemáticos de previsão, ARIMA, Holt-Winters e RNA aplicados as médias mensais da velocidade do vento em algumas localidades do NEB, e assim, identificar qual fornece melhores medidas de acurácia para obter maior precisão na previsão de velocidade do vento.

2. MATERIAL E MÉTODOS

2.1. REGIÃO DE ESTUDO E DADOS COLETADOS

Os dados observados da velocidade do vento nesse trabalho representam médias mensais entre o período de janeiro de 2010 até dezembro de 2014 ao nível de 10 m de altura em relação ao solo. As regiões de estudo são localizadas no Nordeste do Brasil (NEB), Fortaleza (lat -3,77 ° e lon -38,52°), Parnaíba (lat – 2,92 ° e lon – 41,75 °) e São Luís (lat -2,6 ° e lon -44,22 °). Os dados citados foram doados pelo Projeto Airmetar que surgiu da necessidade de ferramentas meteorológicas para consultar e analisar informações dos aeroportos do Brasil a partir da consulta e decodificação dos dados METAR (METeorological Aerodrome Report) fornecidos pela aeronáutica do Brasil. Vale ressaltar que esse projeto não possui fins lucrativos e os dados estão disponíveis gratuitamente no endereço eletrônico (http://www.airmetar.com.br/).

2.2. MODELOS UTILIZADOS

Os modelos matemáticos de previsão de séries temporais utilizados no trabalho foram:

(I) o modelo ARIMA possui características flexíveis, permitindo ajustes e adaptações em seus parâmetros, contribuindo para que diferentes tipos de séries temporais sejam representados, assumindo assim, uma estrutura de correlação linear entre os valores das séries temporais (MAKRIDAKIS et al., 1998).

A modelagem ARIMA é comumente aplicada em processos lineares não estacionários, supõem que as séries temporais não são estacionárias (que há

movimentos de tendência seja de crescimento ou decaimento de qualquer natureza). Para torná-las estacionárias, deve-se aplicar um número determinado de diferenças entre seus dados de ordem (*d*), geralmente uma ou duas. Para esses processos o modelo é comumente representado por ARIMA(*p, d, q*). O número de vezes que a série original tem que ser diferenciada antes de resultar uma série estacionária é denominada ordem de homogeneidade, e estes processos são descritos pelos modelos autorregressivos – integrados (número de diferenciações) – médias móveis de ordens *p*, *d*, e *q*, ARIMA(*p, d, q*) que podem ainda, ter incluído um operador sazonal, representado então por SARIMA(*p, d, q)(P, D, Q*). Para a estimação dos parâmetros do modelo ARIMA, determinado os valores de p, e q, pode-se estimar os p parâmetros φ (ou seja, coeficiente relativo ao filtro de médias móveis), os q parâmetros θ (ou seja, coeficiente relativo ao filtro autorregressivo estacionário), e os erros do modelo εt, também chamado de resíduos, conforme a (Equação1):

$$W_t = \varphi_1 W_{t-1} + \varphi_2 W_{t-2} + \ldots + \varphi_p W_{t-p} + \ldots \\ \ldots + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \quad (1)$$

Maiores detalhes sobre o modelo ARIMA em (MAKRIDAKIS et al., 1998). A seleção do modelo ARIMA mais adequado ao ajuste dos dados observados para cada localidade foi realizado através da utilização do critério Akaike (cuja abreviatura é AIC, do inglês Akaike Information Criterion) (AKAIKE, 1973). Esse critério é baseado na identificação do menor valor para AIC, o qual teoricamente resulta no melhor modelo para ajusta-se aos dados observados.

Para avaliar se os modelos ARIMA possuem viabilidade para a previsão de média mensal de velocidade do vento foi realizada análise dos resíduos. Para que o modelo ARIMA representado na (Equação 1) seja viável no tocante ao ajuste dos dados observados, o termo de erro εt, deste modelo deve se comportar como ruído branco, ou seja, possuir média zero, variância constante (homoscedasticidade) e ser não correlacionado, ou seja, os erros são independentes (GUJARATI, 2009). E ainda o termo εt deve seguir uma distribuição normal. Para identificação desses pressupostos existem alguns testes possíveis de serem aplicados sobre os resíduos. Dentre eles destaca-se o teste de verificação de normalidade dos resíduos por meio do Teste de Shapiro-Wilks, maiores detalhes em (GOIS et al., 2016), teste de independência dos resíduos através do Teste de Durbin-Watson, maiores detalhes em (GUJARATI, 2009), e o teste que verifica a igualdade da variância dos resíduos, ou seja, hipótese de homoscedasticidade através do Teste de Breusch-Pagan, maiores detalhes em (GUJARATI, 2009). Conforme a Tabela 1, tem-se a representação resumida dos testes utilizados, sendo que eles foram aplicados ao nível de significância de α = 0,05.

ISSN: 2237-8642 (Eletrônica)

Tabela 1 – Apresentação dos testes que identificam ruído branco para utilização do modelo ARIMA.

Teste de Shapiro-Wilks
H_0: amostra provém de uma população Normal. H_1: amostra não provém de uma população Normal. Tomar decisão: se o p-valor for maior do que α, ou seja, p > 0,05, (não rejeita H_0).
Teste Durbin-Watson
H_0: os resíduos são independentes. H_1: os resíduos não são independentes. Tomar decisão: se o p-valor for maior do que α, ou seja, p > 0,05, (não rejeita H_0).
Teste Breusch-Pagan
H_0: os resíduos possuem homoscedasticidade. H_1: os resíduos possuem heterocedasticidade. Tomar decisão: se o p-valor for maior do que α, ou seja, p > 0,05, (não rejeita H_0).

(II) Holt-Winters, no ano de 1957 Holt expandiu o modelo de suavização exponencial simples para lidar com dados que apresentavam tendência linear, e assim realizar previsões que fossem mais precisas que as realizadas com suavização exponencial simples, por ajustá-las a esta condição. Em 1960, Winters estendeu o modelo de Holt, incluindo uma nova equação que possibilitasse ser acrescentada nas previsões o comportamento da componente sazonal dos dados que se estivessem trabalhando, gerando assim o Método de Holt-Winters (MAKRIDAKIS et al., 1998). A equação referente ao método de suavização exponencial com sazonalidade e tendência linear, ou seja, com a componente sazonal sendo tratada de forma aditiva, é representada da seguinte forma:

$$a_t = \alpha\left(Y_t - s_{t-p}\right) + (1-\alpha)(a_{t-1} + b_{t-1}) \quad (2)$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta) + b_{t-1} \quad (3)$$

$$s_t = \gamma(Y_t - a_t) + (1-\gamma) + s_{t-p} \quad (4)$$

$$Y_{t+n} = a_t + b_t.h + s_{t-p+n} \quad (5)$$

em que: a_t (nível da série, cuja unidade nesse trabalho é m/s está relacionada de que forma a série temporal prevista evolui com o tempo, sendo que é identificado se a mesma varia lentamente no tempo ou, excepcionalmente, sofrer variações bruscas. Por exemplo, a série prevista pode ter movimento de crescimento lento de forma linear); b_t (tendência, cuja unidade nesse trabalho é m/s^2 está relacionada com o fato da série temporal prevista possuir movimentos de crescimento ou decrescimento que pode ocorrer em intervalos de tempos

distintos. Por exemplo, uma série temporal prevista hipotética possui uma tendência de crescimento linear no seu primeiro ano, já no segundo ano a mesma foi identificada com um decrescimento linear); s_t (componente sazonal, cuja unidade nesse trabalho é o m/s que está relacionada com o fato da série temporal prevista existir padrões cíclicos de variação que se repetem em intervalos relativamente constantes de tempo. Por exemplo, uma série temporal prevista hipotética durante um ano apresenta comportamento ondulatório a cada três meses); Y_{t+n} (previsão para o período n adiante, e possui unidade de m/s); *p* (período sazonal); n = 1, 2, 3, …, *h* (horizonte da previsão); E α, β, γ parâmetros exponenciais alisadores, do nível, da tendência e da sazonalidade, respectivamente. Vale ressaltar que os coeficientes (α, β, e γ) são encontrados de tal forma que ofereçam uma série temporal da previsão que minimize o erro quando comparado a série temporal observada (ou também chamada de real), e variam entre 0 e 1.

(III) As RNA foram apresentadas pela 1ª vez em 1943 pelo neurofisiologista McCulloch, e pelo matemático Walter Pitts (MCCULLOCH e PITTS, 1943). O modelo de Pitts e McCulloch era constituído de apenas uma saída e uma função de entrada, calculada por meio da soma ponderada de diversos valores. Esse modelo assemelha-se com o neurônio humano, onde sinais são recebidos pelo mesmo com intensidades diferentes, provocando uma determinada reação por parte da célula nervosa. Embora biologicamente inspiradas no neurônio humano, as RNAs encontraram aplicações em diferentes áreas científicas.

Conforme Zhang et al. (1998), em anos recentes uma das técnicas de inteligência computacional, comumente usada na tentativa de prever séries temporais é a RNA. Estas são baseadas na arquitetura e aprendizagem do cérebro humano. As RNAs funcionam conceitualmente de forma similar ao cérebro humano, ou seja, tentando reconhecer regularidades e padrões de dados sendo capazes de aprender com a experiência e fazer generalizações baseadas no seu conhecimento previamente acumulado. As RNAs em sua estrutura podem possuir ambos os modelos não-linear e linear e, assim obter melhores resultados quando comparado com outros modelo de previsão (ZHANG et al., 1998; HAYKIN, 2001; ZHANG et al., 2003).

Uma RNA pode ser pensada como uma rede de "neurônios" organizados em camadas. Os preditores (ou entradas **yt-i**) formam a camada inferior, e as previsões (ou saídas **yt+h**) formam a camada superior. Pode haver camadas intermediárias que contenham "neurônios ocultos". A (Figura 1) mostra um exemplo para a estrutura de uma RNA com 4 entradas e com 1 camada oculta. Os coeficientes ligados aos preditores são chamados de "pesos" e comummente representado por **wi**, Os pesos são selecionados através de um "algoritmo de aprendizagem", como é o caso do backpropagation (que surge do fato que o algoritmo se baseia na retropropagação dos erros para realizar os ajustes de pesos das camadas intermediárias), que minimiza o erro entre as séries temporais previstas e observadas (ZHANG et al., 1998).

Previsão de séries temporais é o processo de prever valores futuros de uma série temporal a partir do conhecimento de seus valores passados. Dentre as técnicas utilizadas para previsão de séries temporais estão aquelas baseadas

em diferentes arquiteturas de RNA, do tipo *feedforward* como a Multi-Layer Perceptron (MLP). As redes MLP foram originalmente concebidas para executar tarefas de natureza estática, não foram, portanto, idealizadas para tratar problemas temporais. O método de janela de tempo foi a primeira adaptação da rede MLP treinada com o algoritmo *backpropagation* para processamento dinâmico (ZHANG et al., 1998).

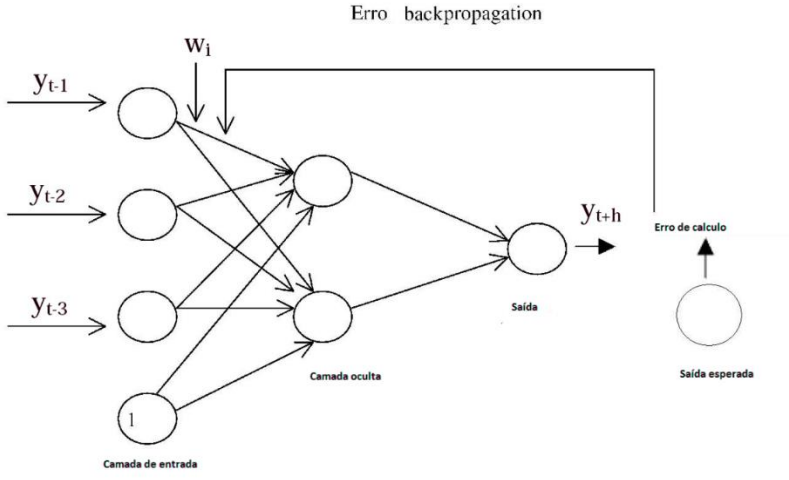


Figura 1 – Estrutura de uma RNA do tipo MLP com o algoritmo backpropagation com 3 entradas, 1 camada oculta e 1 saída.

Na estrutura da RNA como, por exemplo, mostrada (Figura 1), as saídas de nós em uma camada são entradas para a próxima camada. O resultado é, em seguida, modificada por uma função não linear antes da saída. Por exemplo, as entradas em neurónios j em camadas ocultas podem ser linearmente combinadas a partir da expressão:

$$z_j = b_j + \sum_{i=1}^{3} w_{i,j}\, y_i \qquad (6)$$

Na camada oculta, a (Equação 6) é modificada utilizando uma função não linear, tal como um sigmoide dada pela seguinte expressão:

$$s(z) = \frac{1}{1+e^{-z}} \qquad (7)$$

A (Equação 7), por sua vez, serve de entrada para a camada seguinte, sendo que essa estratégia tende a reduzir o efeito de valores extremos de entrada, proporcionando um melhor funcionamento da RNA. Os parâmetros b_1, b_2, b_3, e $w_{1,1}$,..., $w_{3,2}$ são descobertos na etapa de "aprendizagem" a partir dos dados observados da série temporal em questão. Vale ressaltar que os pesos

ISSN: 2237-8642 (Eletrônica)

inicialmente podem assumir valores aleatórios, que por sua vez, serão atualizados usando os dados observados. Ainda maiores detalhes sobre a utilização de RNA para previsão de séries temporais podem ser obtidos através de Zhang et al. (1998). Na literatura é possível destacar trabalhos sobre a utilização de RNA para previsão de série temporal de velocidade do vento como, por exemplo, em Flores et al. (2005), os autores realizaram a previsão da velocidade dos ventos em uma fazenda eólica usando redes neurais artificiais. A RNA baseou-se no algoritmo de aprendizagem *backpropagation*, cuja avaliação foi realizada com medição de dados reais de duas diferentes localizações.

Ressalta-se que a utilização de RNA pelo software ambiente R versão 3.2.5 é fornecida por meio do pacote Forecast apresentados por (HYNDMAN e KHANDAKAR, 2008) por meio da função nnetar, que por sua vez, é simbolizada por NNAR(p, P, k)(12), em que, p (representa quantidade de termos auto-regressivo), P (é referente a dados sazonais), e k (representa o número de sub-camadas), e por fim, o número 12 é referente a quantidade de entradas **yi(t)** da primeira camada da RNA.

É importante ressaltar que a principal diferença entre RNA e as abordagens estatísticas, tais como os modelos ARIMA, é que as RNA não estabelecem hipóteses ou suposições sobre a distribuição ou propriedades dos dados e, desta forma, tendem a ser mais úteis em situações práticas. As RNAs constituem também uma abordagem inerentemente não-linear, fornecendo precisão quando modelam dados de padrões complexos. Existem vários tipos de RNAs, cada uma com diferentes objetivos, topologias e algoritmos de aprendizagem. Ainda para mais efeito de comparação com os modelos de séries temporais apresentado acima, os mesmo, são trabalhados com variáveis aleatórias que possuem uma determinada distribuição conjunta de probabilidade, já nas RNAs temos apenas sinais de entrada e saída da rede. A ausência de um modelo estatístico bem especificado impede, por exemplo, a construção de intervalos de confiança para as previsões geradas por RNA. A previsão gerada por RNA é sempre pontual, ao contrário do que ocorre com os modelos de séries temporais. Em modelos de previsão baseados em redes neurais os intervalos de confiança têm sido geralmente colocados tendo por base os modelos estatísticos lineares tradicionais. Entretanto, como as redes neurais são modelos não lineares, estas metodologias lineares nem sempre são adequadas de tal forma que novas propostas estão sendo testadas, tais como o método utilizando bayesian inference framework. Maiores detalhes sobre as características mencionadas neste parágrafo sobre o uso das RNAs obter em (ZHANG et al., 1998; HAYKIN, 2001; ZHANG et al., 2003).

Para a realização da previsão com os modelos ARIMA, Holt-Winters e RNA a partir dos dados citados em cada localidade foram utilizadas as seguintes etapas: (1) aplicação dos modelos aos dados da velocidade do vento no período de janeiro de 2010 até dezembro de 2014. Essa etapa consiste em fazer com que o modelos propostos possam captar características das séries temporais tais como, tendência, nível e sazonalidade; (2) previsão dos dados de velocidade do vento no período de janeiro de 2015 até dezembro de 2015, juntamente de seu intervalo de confiança.

ISSN: 2237-8642 (Eletrônica)

2.4. ESTATÍSTICA DE ERROS

Na avaliação da acurácia dos modelos de previsão propostos utilizou-se o Erro Médio Absoluto, o qual mede o valor médio de erro entre as séries observadas e ajustadas (CAMELO et al., 2017). Nessa análise essa variável será representada por MAE (do iglês *Mean Absolute Error*), cuja representação matemática é dada por:

$$MAE = \frac{1}{o}\sum_{i=1}^{o}\left|v_{ajustado} - v_{observado}\right|, \quad (8)$$

em que, vajustado representa valor individual da série ajustada e vobservado representa valor individual da série observada, e o é a ordem das séries. O MAE possui a unidade de m/s.

A Raiz do Erro Quadrático Médio (simbolizado por RMSE, sendo essa abreviatura do inglês Root Mean Squared Error) representa as diferenças individuais quadráticas entre as séries temporais observadas e ajustadas, que pode ser definido matematicamente por,

$$RMSE = \sqrt{\frac{1}{o}\sum_{i=1}^{o}\left(v_{ajustado} - v_{observado}\right)^2} \quad (9)$$

Da mesma forma o RMSE na (Equação 9) possui unidade de m/s. É importante ressaltar que o RMSE pode ainda ser interpretado da seguinte maneira: se houver grandes valores desta variável, os mesmos representam grandes erros nas variáveis ajustadas, e valores próximos de zero indicam ajustes quase perfeitos. Maiores detalhes sobre a Raiz do Erro Quadrático Médio podem ser obtidos em (CAMELO et al., 2017).

Outra forma de medida de erro é a Média do Erro Absoluto Percentual (simbolizado por MAPE, do inglês Mean Absolute Percentage Error). A grande vantagem de utilizar essa expressão está na sua representação em termos percentuais (%) que fornece um rápido entendimento. Já uma desvantagem que deve ser considerada está no sentido de que se, por acaso, o valor do observado for muito pequeno, qualquer discrepância faz o MAPE "explodir". A expressão utilizada com essa variável é representada por,

$$MAPE = \frac{1}{o}\sum_{i=1}^{o}\left|\frac{v_{ajustado} - v_{observado}}{v_{observado}}\right| \times 100 \quad (10)$$

Maiores detalhes sobre estatística de erros, em especial às (Equações 8, 9, e 10) que serão usadas nesse trabalho, podem ser obtidos em (CAMELO et al., 2017). Todos os cálculos realizados nesse estudo e além da parte gráfica, foram produzidos pelo software livre ambiente R versão 3.2.5 (R-PROJECT, 2016).

ISSN: 2237-8642 (Eletrônica)

3. RESULTADOS E DISCUSSÃO

A Tabela 2 apresenta a estatística de erros (MAE, RMSE e MAPE) na comparação entre observado e ajustado para os modelos ARIMA, HW e RNA no período de estudo. A análise da variável MAPE na localidade de Fortaleza apresentou o modelo RNA como o de menor valor em 8,74%, em Parnaíba o modelo ARIMA apresenta o menor valor de 9,70%. Enquanto na localidade de São Luís o menor valor encontrado de MAPE é atribuído ao modelo HW wm 8,73%. A análise da variável MAE mostra que em Fortaleza o modelo RNA apresenta o menor valor de 0,43 m/s, em Parnaíba o modelo ARIMA apresenta o menor valor de 0,49 m/s, e em São Luís o menor valor é apresentado pelo modelo HW em 0,31 m/s. A análise da variável RMSE mostra que em Fortaleza o modelo RNA apresenta o menor valor de 0,52 m/s, em Parnaíba o menor valor é dado pelo modelo ARIMA em 0,62 m/s, e em São Luís o menor valor é apresentado pelo modelo HW em 0,43 m/s. Portanto, de acordo com as medidas de acurácia, o modelo RNA forneceu melhores ajustes aos dados observados de Fortaleza, o modelo ARIMA melhores ajustes aos dados observados de Parnaíba, e por fim, o modelo HW apresentou melhores ajustes aos dados observados de São Luís.

Tabela 2 – Estatística de erros (MAE, RMSE e MAPE) na comparação entre observado e ajustado para os modelos ARIMA, HW e RNA no período de janeiro de 2010 até dezembro de 2014.

ARIMA	**MAE (m/s)**	**RMSE (m/s)**	**MAPE (%)**
Fortaleza	0,56	0,72	11,81
Parnaíba	0,49	0,62	9,70
São Luís	0,39	0,51	10,74
Holt-Winters	**MAE (m/s)**	**RMSE (m/s)**	**MAPE (%)**
Fortaleza	0,48	0,62	10,68
Parnaíba	0,51	0,72	10,34
São Luís	0,31	0,43	8,73
RNA	**MAE (m/s)**	**RMSE (m/s)**	**MAPE (%)**
Fortaleza	0,43	0,52	8,74
Parnaíba	0,51	0,64	9,85
São Luís	0,38	0,45	10,50

Comparativamente com a literatura é possível destacar (HU et al., 2013), onde os autores utilizaram o modelo ARIMA para previsão de média mensal da velocidade do vento no período de 2001 a 2006. Nesse estudo foi identificado que o menor valor de MAPE encontrado para regiões montanhosas da China é de aproximadamente 13,70%. A utilização de RNA para previsão da velocidade do vento em Portugal (CATALÃO et al., 2009) encontram valor de MAPE da

ISSN: 2237-8642 (Eletrônica)

ordem de 7,26%. O algoritmo utilizado foi o backpropagation com uma estrutura de três camadas. A previsão da velocidade do vento com o uso do modelo HW é possível destacar previsão de médias horárias realizada por Malgorzata e Piotr (2014). Os autores encontraram para as variáveis MAE e RMSE, os seguintes valores de 2,22m/s e 2,93m/s respectivamente, com horizonte de previsão de 6h em regiões da Polônia. A comparação entre os modelos ARIMA e RNA para previsão de velocidade de vento em termos de médias horárias de um parque eólico na província de Hubei da China é realizada por (LING e XU, 2011). Nesse estudo os autores demonstraram que tais modelos foram satisfatório, princialmente o RNA que forneceu melhores previsões, segundo o valor de MAPE no horizonte de 4 horas é de 7,37%.

A Tabela 3 mostra o modelo selecionado e a análise dos resíduos realizada a partir dos testes citados anteriormente. Os resultados mostram que para todas as localidades, o modelo ARIMA é viável para representação dos dados observados, em função de satisfazer a condição de ruído branco, ou seja, são homocedásticos (Teste Breusch-Pagan); independentes (Teste Durbin-Watson); e possuem distribuição normal (Teste Shapiro-Wilks), sendo essas informações obtidas em função de cada p-valor fornecido por cada teste, os quais são maiores do que o nível de significância de $\alpha = 0,05$. Com isso ressalta-se que esses resultados são indicativos de que existe viabilidade para a utilização do modelo nas localidades de estudo. Em todas as localidades os melhores ajustes na utilização da modelagem Box e Jenkins foram fornecidos com o uso do modelo SARIMA, ou seja, nessas localidades é necessário considerar a componente sazonal do modelo ARIMA. E em termos de RNA a partir do pacote *Forecast* do software *R*, a função nnetar, forneceu os seguintes termos (2, 1, 2), (1, 1, 2), e (1, 1, 2) respectivamente para Fortaleza, São Luís e Parnaíba. Ressalta-se a dificuldade de se calcular os intervalos de confiança para RNA conforme já comentado, ou seja, em função da ausência de um modelo estatístico bem especificado. Dessa forma, ainda não é possível o cálculo do intervalo de confiança para RNA no software *R* pelo pacote citado acima para a previsão da série temporal considerada.

ISSN: 2237-8642 (Eletrônica)

Tabela 3 – Apresentação dos modelos ARIMA para cada localidade e aplicação dos testes sobre os resíduos.

Localidade (Modelo)	Shapiro-Wilk	Durbin-Watson	Breusch-Pagan
Fortaleza SARIMA(0,1,0)(1,0,0)	p-valor: 0,128. Decisão: os resíduos possuem distribuição normal.	p-valor: 0,725. Decisão: os resíduos são independentes.	p-valor: 0,518. Decisão: as variâncias dos resíduos são iguais (homoscedasticidade).
Parnaíba SARIMA(0,1,0)(1,0,0)	p-valor: 0,167. Decisão: os resíduos possuem distribuição normal.	p-valor: 0,190. Decisão: os resíduos são independentes.	p-valor: 0,590. Decisão: as variâncias dos resíduos são iguais (homoscedasticidade).
São Luís SARIMA (1,0,1)(1,1,1)	p-valor: 0,078. Decisão: os resíduos possuem distribuição normal.	p-valor: 0,5328. Decisão: os resíduos são independentes.	p-valor: 0,980. Decisão: as variâncias dos resíduos são iguais (homoscedasticidade).

A (Figura 2) mostra os modelos de previsão para Fortaleza, sendo que na (Figura 2A) representa o modelo SARIMA, (Figura 2B) o modelo HW, e na (Figura 2C) o modelo RNA. Em todos os modelos as séries temporais ajustadas (representadas nas linhas contínuas cor vermelha) consegue acompanhar o perfil da série temporal observada (representada na linha contínua cor preta), com semelhanças entre máximos e mínimos da velocidade do vento como, por exemplo, no mês de abril de 2014 os valores ajustado e observado para o RNA são aproximadamente igual a 4,5 m/s. Da mesma forma as séries temporais previstas para 2015 (representadas nas linhas contínuas cor azul) obtidas pelos modelos HW e RNA, tendem acompanhar a série temporal observada, ou seja, também há coincidência entre máximos e mínimos da velocidade do vento como, por exemplo, o mínimo da previsão para o modelo HW foi registrado para o mês de abril de 2015 no valor de aproximadamente 4,0 m/s, este valor é próximo dos valores encontrados para o mesmo mês dos anos de 2013 e 2014. Já para o modelo RNA, o máximo da previsão foi de aproximadamente 7,0 m/s, sendo que este valor é semelhante aos encontrados para o mesmo mês dos anos de 2013 e 2014. Ressalta-se que na previsão para os meses do ano 2015 nos modelos ARIMA e HW ocorre aumento no intervalo de confiança (representado na linha seccionada cor preta em 95%), sendo que isso poderá ser um indicativo que os modelos propostos possuem dificuldades em prever valores distantes da série observada.

ISSN: 2237-8642 (Eletrônica)

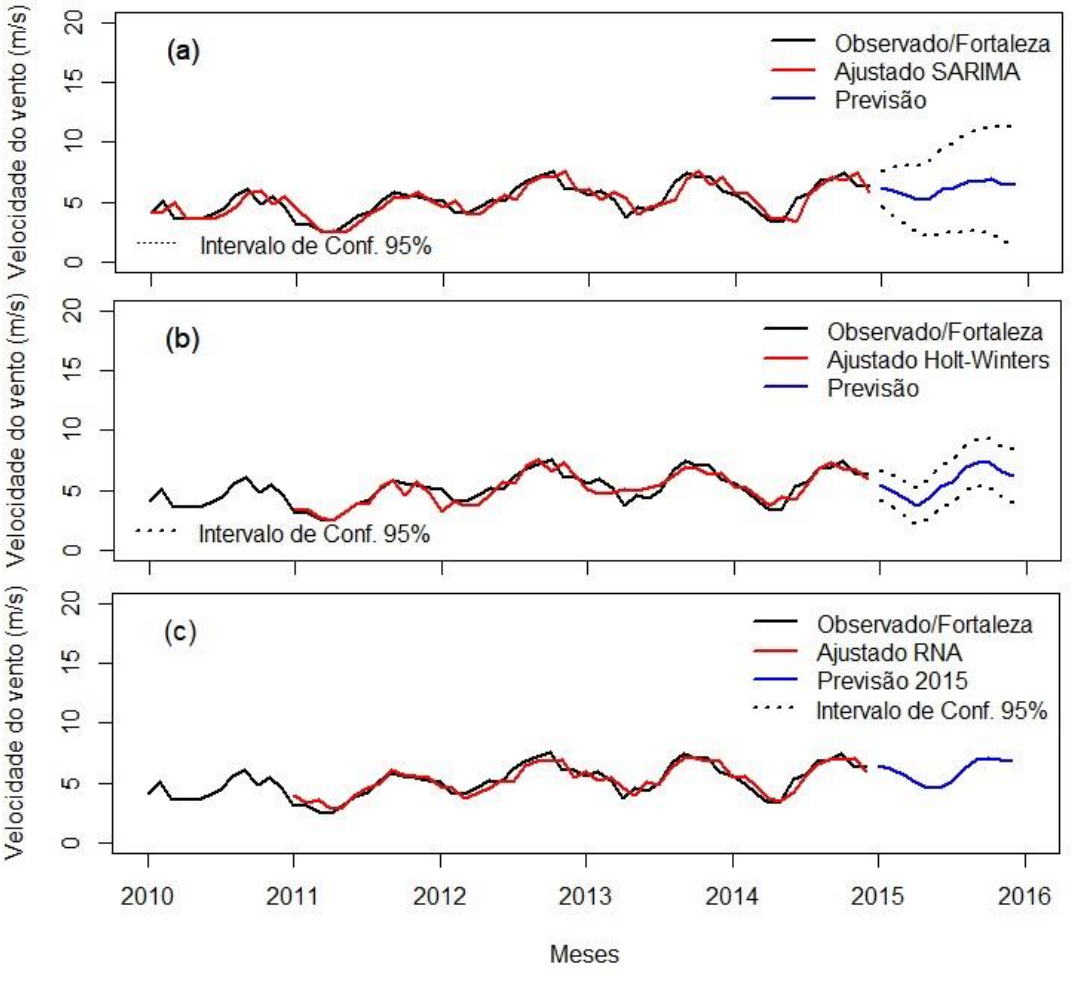


Figura 2 – Apresentação dos modelos de previsão para médias mensais de velocidade do vento em Fortaleza. (A) Modelo SARIMA, (B) Modelo Holt-Winters, e (C) Modelo RNA.

De acordo com estudo realizado a respeito da normal climatológica em todo território brasileiro pelo Instituo Nacional de Meteorologia (INMET), o qual pode ser consultado em (INMET, 2016), para a cidade de Fortaleza os meses de maior precipitação, ou também chamados de períodos chuvosos estão entre fevereiro e maio, com pico no mês de abril. Nesse período segundo Camelo et al. (2008) a média mensal da intensidade da velocidade do vento em Fortaleza é menor quando comparado ao período de menores totais de chuva, período seco entre os meses de outubro até dezembro. Ainda na (Figura 3), é possível verificar que os modelos conseguem captar menores intensidades da média mensal da velocidade do vento no período chuvoso em especial, na (Figura 3B) relativo ao mês de abril no valor de aproximadamente 4,0 m/s, e também maiores intensidades para o período seco como é o caso do mês de outubro com aproximadamente 7,0 m/s.

A (Figura 3) mostra os modelos de previsão da média mensal da velocidade do vento em Parnaíba. Em que na (Figura 3A) representa o modelo

ISSN: 2237-8642 (Eletrônica)

ARIMA, (Figura 3B) representa o modelo HW, e na (Figura 3C) o modelo RNA. No tocante aos ajustes (representados nas linhas contínuas cor vermelha) aos dados observados (representado na linha contínua cor preta), é possível identificar semelhanças entre máximos e mínimos para as duas séries em todos os modelos. No entanto, em termos da série temporal prevista para os meses de 2015, apenas o modelo HW apresentou comportamento similar a série temporal observada. Os intervalos de confiança de 95% (linha seccionada cor preta) aumentam em função dos meses previstos para 2015 relativo aos modelos ARIMA e HW, esse resultado poderá ser indicativo de que para esses modelos há dificuldades em prever valores distantes da série observada.

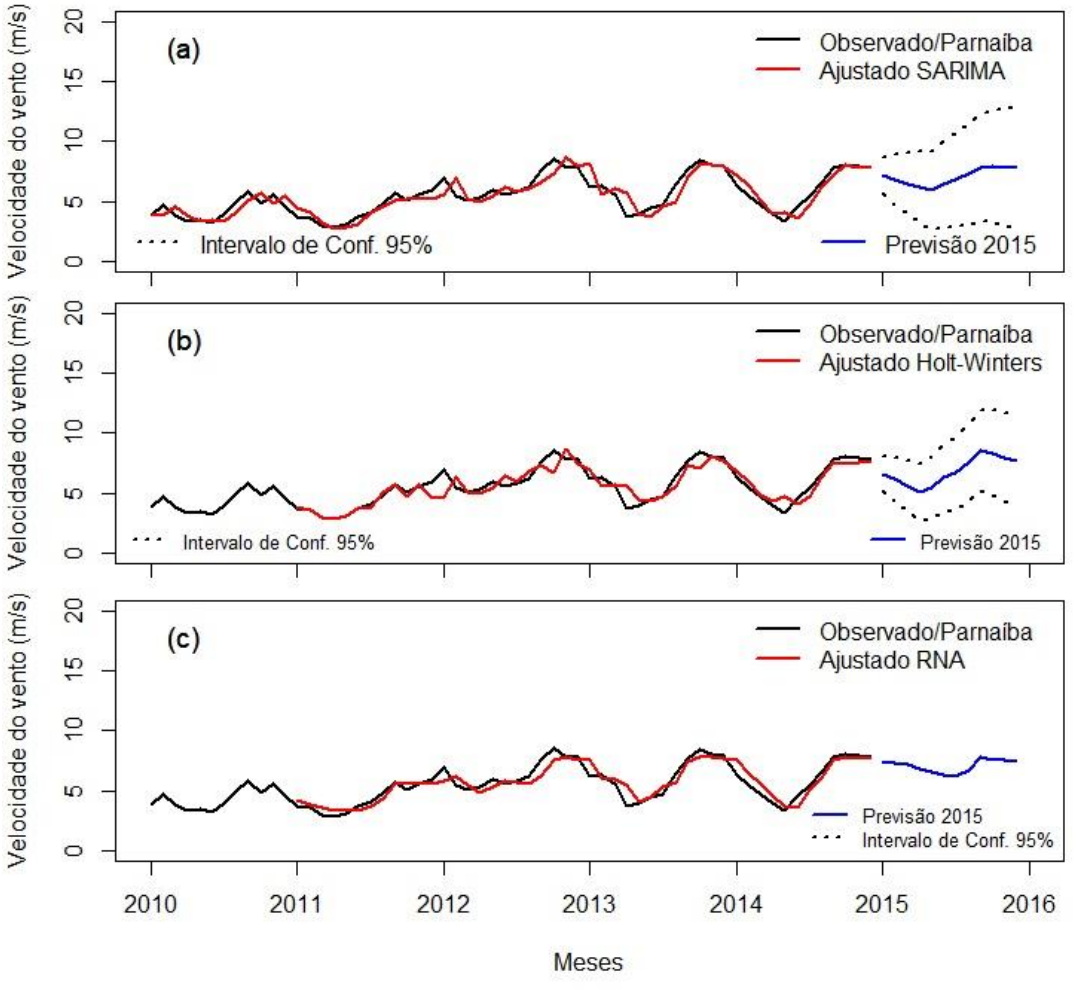


Figura 3 – Apresentação dos modelos de previsão para médias mensais de velocidade do vento em Parnaíba. (A) Modelo SARIMA, (B) Modelo Holt-Winters, e (C) Modelo RNA.

A (Figura 4) mostra os resultados dos modelos de previsão ARIMA na (Figura 4A), HW na (Figura 4B) e RNA na (Figura 4C) todos para a localidade de São Luís. Como é possível identificar as séries temporais ajustadas (linha contínua na cor vermelha) se assemelham por demais a série temporal observada (linha contínua cor preta) em termos de equivalências de máximos e mínimos. No tocante as séries temporais previstas para os meses de 2015

ISSN: 2237-8642 (Eletrônica)

(representadas na linha contínua cor azul) tendem acompanhar a série temporal observada, porém o intervalo de confiança de 95% aumenta para os modelos ARIMA e HW em função dos meses, isso pode ser um indicativo de que os modelos propostos para São Luís poderá ter dificuldades em apresentar previsões para períodos distantes da série temporal observada.

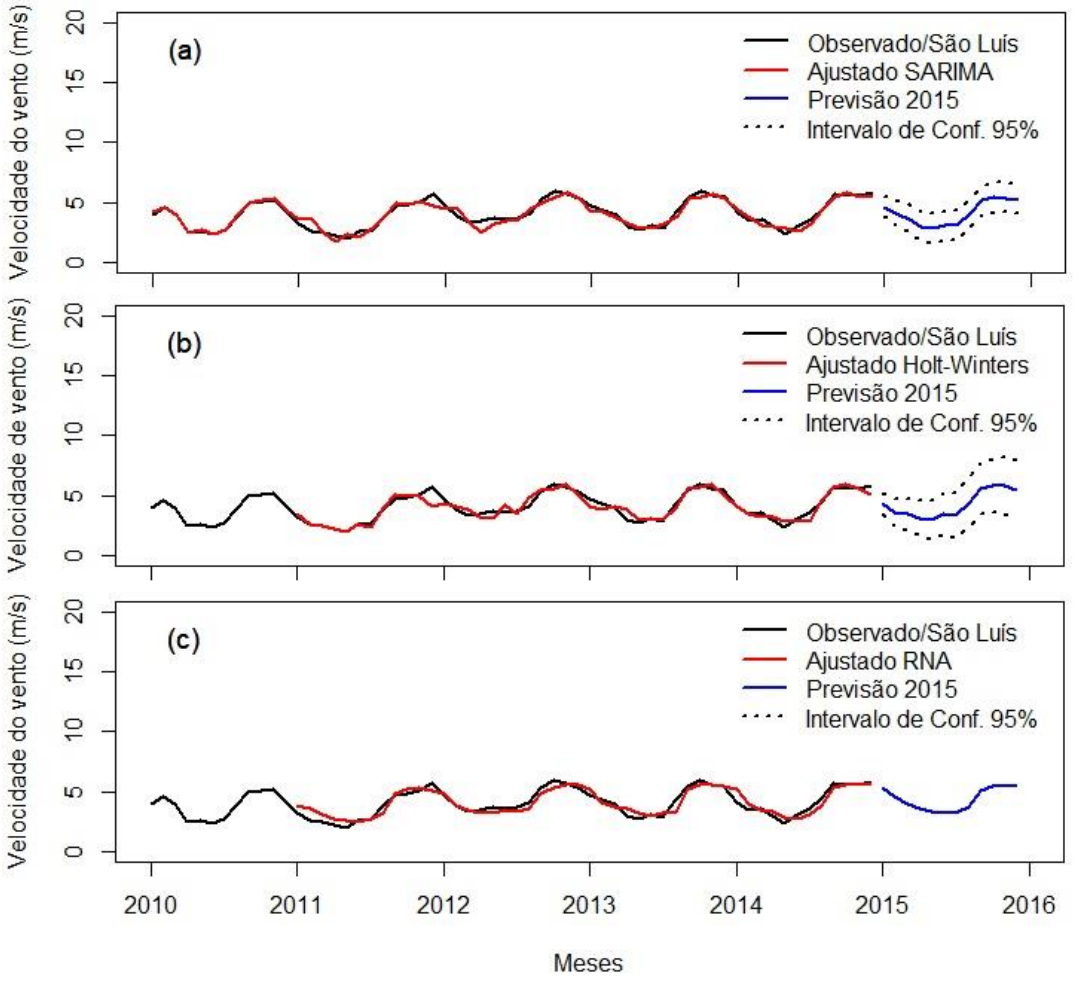


Figura 4 – Apresentação dos modelos de previsão para médias mensais de velocidade do vento em São Luís. (A) Modelo SARIMA, (B) Modelo Holt-Winters, e (C) Modelo RNA.

4. CONCLUSÕES

A comparação dos modelos de previsão de séries temporais aplicados as médias mensais da velocidade do vento nesse trabalho revelou que nas localidades de Fortaleza e São Luís, apresentaram menores valores das estatísticas de erros para o modelo RNA. Já para a localidade de Parnaíba o modelo ARIMA apresentou os melhores resultados para a estatística de erros. No tocante aos ajustes aos dados observados os três modelos apresentaram semelhanças entre máximos e mínimos da velocidade do vento. Para as séries temporais previstas os modelos conseguem captar características importantes das séries temporais observadas como é o caso de menores intensidades da média mensal da velocidade do vento no período chuvoso e, também maiores

intensidades da média mensal da velocidade do vento para o período seco. Para trabalhos futuros poderá ser testado, a partir da disponibilidade de dados da velocidade do vento, nessas regiões de estudo e em outras localidades do nordeste brasileiro, previsões com maiores quantidades de séries temporais e, portanto, identificar se com isso haverá uma redução na análise de erros e consequentemente melhores ajustes aos dados observados.

5. REFERÊNCIAS BIBLIOGRÁFICAS

AKAIKE, H. Information theory and an extension of the maximum likelihood principle. Proc. 2a Inter. Symposium on Information Theory, 267-281, Budapest, 1973.

AKELLA, A. K.; SAINI, R. P.; SHARMA, M. P. Social, economical and environmental impacts of renewable energy systems. Renewable Energy, v. 32, n. 2, p. 390-396, 2009.

BILGEN, S.; KELES, S.; KAYGUSUZ, A.; SARI, A.; KAYGUSUZ, K. Global warming and renewable energy sources for sustainable development: a case study in Turkey. Renewable and Sustainable Energy Reviews, v. 12, n. 2, 372-396, 2008.

BOX, G. P. E; JENKINS, G. M. Time Series Analysis: Forecasting and Control. Edição revista. São Francisco: Holden Day, 1978.

CADENAS, E.; JARAMILHO, O. A.; RIVERA, W. Analysis and forecasting of wind velocity in chetumal, quintana roo, using the single exponential smoothing method. Renewable Energy, v. 35, n. 5, p. 925-930, 2010.

CADENAS, E.; RIVERA, W. Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA – ANN model. Renewable Energy, v. 35, n. 12, p. 2732-2738, 2010.

CAMELO, H. N.; TEIXEIRA, W. F.; LUCIO, P. S.; LEAL JUNIOR, J. B. V. Previsão de Velocidade Média do Vento através da utilização de modelagem Auto-Regressiva de Médias Móveis (ARIMA) em Região Serrana no Estado do Ceará - Brasil. Revista Brasileira de Energia, v. 4, n. 3, p. 17-29, 2015.

CAMELO, H. N.; CARVALHO, P. C. M.; LEAL JUNIOR, J. B. V.; ACCIOLY FILHO, J. B. P. Análise estatística da velocidade de vento do estado do Ceará. Revista Tecnologia, v. 29, n. 2, p. 211-223, 2008.

CAMELO, H. N.; LUCIO, P. S.; LEAL JR., J. B. V. Modelagem de média mensal de velocidade do vento para região litorânea no nordeste Brasileiro através do método aditivo Holt-Winters com vias a previsão de geração eólica. Revista Brasileira de Energia Renováveis, v. 5, n. 4, p. 587-604, 2016.

CAMELO, H. N.; LUCIO, P. S.; CARVALHO, P. C. M.; LEAL JR., J. B. V. Previsão de velocidade do vento no nordeste brasileiro através do modelo ARIMAX. Revista Ciência e Tecnologia, v. 20, n. 36, p. 9-17, 2017.

CATALÃO, J. P. S.; POUSINHO, H. M. I.; MENDES, V. M. F. An artificial neural network approach for short - term wind power forecasting in Portugal. Engineering Intelligent Systems, v. 17, n. 1, p. 1-5, 2009.

CEPEL – Centro de Pesquisas de Energia Elétrica. 2010. Disponível em: <http://www.cepel.br>. (Consultado em 12/01/2015).

GOIS, G.; SOUZA, J. C. S.; DE OLIVEIRA-JÚNIOR, J. F.; DELGADO, R. C.; LYRA, G. B.; TEODORO, P. E. Models to estimate incident solar radiation on Seropédica, Rio de Janeiro. Bioscience Journal (on line),v. 32, n. 2, p. 505-513, 2016.

EPE – Empresa de Pesquisa Energética. 2014. Disponível em: <http://www.epe.gov.br>. (Consultado em: 12/05/2015).

FLORES, P.; TAPIA, A.; TAPIA, G. Application of a control algorithm for wind speed prediction and active power generation. Renewable Energy, v. 30, n. 4, p. 523-536, 2005.

FRONDEL, M.; RITTER, N.; SCHMIDT, C. M.; VANCE, C. (2010). Economic impacts from the promotion of renewable energy technologies: The German experience. Energy Policy, v. 38, n. 8, p. 4048-4056, 2010.

GWEC. Global Wind Statistics 2016, http://www.gwec.net/wp-content/uploads/vip/GWEC_PRstats2016_EN_WEB.pdf; 2017 [accessed 18 October 2017].

GUJARATI, D. N.; PORTER, D. C. Basic Econometrics. 4ªed. McGraw-Hill: New York, 2009.

HAYKIN, S. Redes neurais: princípios e prática. Bookman: New York, 2001.

HYNDMAN, R. J; KHANDAKAR, Y. Automatic Time Series Forecasting: The forecast Package for R. Journal of Statistical Software, v. 27, n. 3, p. 1-22, 2008.

HU, J.; WANG, J.; ZENG, G. A hybrid forecasting approach applied to wind speed time series. Renewable Energy, v. 60, p. 185-194, 2013.

INMET – Instituto Nacional de Meteorologia. (2016). Página eletrônica:<http://www.inmet.gov.br/portal/index.php?r=clima/normaisclimatologicas>. Consultado em 11/05/2016.

KAYGUSUZ, K. Energy for sustainable development: A case of developing countries. Renewable and Sustainable Energy Reviews, v. 16, n. 2, p. 1116-1126, 2012.

KAVASSERI, R.; SEETHARAMAN, K. Day-ahead wind speed forecasting using f-ARIMA models. Renewable Energy, v. 34, n. 5, p. 1388-1393, 2009.

LING, C.; XU, L. "Comparison between ARIMA and ANN Models Used in Short-Term Wind Speed Forecasting," in Power and Energy Engineering Conference (APPEEC), 2011 Asia-Pacific, 2011, pp. 1-4.

LIU, H.; HONG-QI, T.; YAN-FEI, L. Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. Applied Energy, v. 98, p. 415-424, 2012.

MAKRIDAKIS, S. G.; WHEELWRIGHT, S. C.; HYNDMAN, R. J. Forecasting: methods and applications. 3ª Ed., New York: Wiley, 656 p, 1998.

MCCULLOCH, W.; PITTS, W. A. A Logical Calculus of Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics, v. 5, n. 1, p. 115-133, 1943.

MALGORZATA, J.; PIOTR, O. Comparative analysis of wind speed short term forecasts for wind farms. Agricultural Engineering, v. 18, n. 4, p. 81-90, 2014.

PANWAR, N. L.; KAUSHIK, S. C.; KOTHARI, S. Role of renewable energy sources in environmental protection: A review. Renewable and Sustainable Energy Reviews, v. 15, n. 3, p. 1513-1524, 2011.

RESCH, G.; HELD, A.; FABER, T.; PANZER, C.; TORO, F.; HAAS, R. Potentials and prospects for renewable energies at global scale. Energy Policy, v. 36, n. 11, p. 4048-4056, 2008.

R-PROJECT, R Development Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

SEINFRA – Secretaria de Infraestrutura do Estado do Ceará. 2001. Atlas do Potencial Eólico do Estado do Ceará. Fortaleza – CE, Brasil. Disponível em: <http://www.seinfra.ce.gov.br/>. (Consultado em 12/01/2015).

TROEN, I.; PETERSEN, E. L. European wind atlas. Riso National Laboratory; 1989.

ZHANG, G.; PATUWO, B. E; HU, Y. M. Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting, v. 4, n. 1, p.35-62, 1998.

ZHANG, L.; LUH, P. B. & KASIVISWANATHAN, K. "Energy clearing price prediction and confidence interval estimation with cascaded neural netoworks" IEEE Transactions on Power Systems, v. 18, n. 1, p. 99-105, 2003.