



## PREVISÃO DE VELOCIDADE DE VENTO: UMA ABORDAGEM UTILIZANDO *BAGGING HOLT WINTERS* com DECOMPOSIÇÃO STL

**Tiago Mendes Dantas**

Departamento de Engenharia Industrial- PUC-RIO  
Rua Marquês de São Vicente, 225 – Gávea, Rio de Janeiro – RJ  
t.mendesdantas@gmail.com

**Fernando Luiz Cyrino Oliveira**

Departamento de Engenharia Industrial- PUC-RIO  
Rua Marquês de São Vicente, 225 – Gávea, Rio de Janeiro – RJ  
cyrino@puc-rio.br

### ESUMO

As fontes renováveis de energia constituem um importante tópico de interesse mundial. Dentre essas fontes, a energia eólica emerge como uma das mais promissoras e o Brasil apresenta um grande potencial de geração, graças às características do seu território, sobretudo na região Nordeste, onde os ventos sopram de maneira abundante. Com o objetivo de apresentar previsões mais precisas de velocidade para uma estação anemométrica da região, este trabalho apresenta uma abordagem de previsão de Séries Temporais que combina técnicas de Bagging, decomposição STL e o método clássico de Amortecimento Exponencial de Holt Winters. Os resultados indicam que o método proposto foi capaz de prever com significativa habilidade, superando o desempenho da abordagem de Holt Winters tradicional e ao método Neuro-Fuzzy.

**PALAVRAS CHAVE. Energia Eólica, Bagging, Holt Winters.**

**Área principal (EN-PO na Área de energia, EST-Estatística)**

### ESU CU

Nowadays, renewable sources of energy are an important issue in the world demands. Among these sources wind energy emerges as one of the most promising ones and, due to the characteristics of the Brazilian territory, the Northeast region has a huge potential, since there is an abundant wind flow in the area. In order to develop more accurate forecasts for an anemometric station, this work aims to implement a Times Series approach that combines Bagging Techniques, STL decomposition and the classic Exponential Smoothing Holt Winters Method. The results indicate that the proposed method was able to produce good forecasts, overcoming the classic Holt Winters and Neuro Fuzzy performances.

**KEYWORDS. Wind Energy. Bagging. Holt Winters.**

**Main area (EN-OR in Energy, EST-Statistics)**

## 1. Introdução

As fontes renováveis de energia representam uma das alternativas para atender o crescimento de demanda energética. Entre tais fontes, a energia eólica emerge como uma das mais promissoras, tendo experimentado um crescimento acelerado nos últimos anos. Suas vantagens decorrem principalmente do fato de ser de natureza barata, uma vez que o parque eólico é instalado, inesgotável e, sobretudo, limpa.

Atualmente, o vento é transformado em energia elétrica através de turbinas eólicas (aerogeradores). Estes, de maneira resumida, são geradores elétricos capazes de captar o vento, transformando a energia cinética gerada pela movimentação de suas pás (cata-vento) em energia elétrica. Em geral, os aerogeradores são dispostos em grupos para captar a maior quantidade de vento útil, constituindo assim um parque eólico.

Algumas críticas a esse tipo de gerador residem no fato de interferirem nos hábitos de certas aves, além de gerarem poluição sonora, visual e, sobretudo, da dificuldade de encontrar regiões que permitam a sua instalação já que devem ser alocados em locais com vento abundante, porém não muito fortes, de maneira a não comprometer a integridade do aerogerador.

Segundo o *Global Wind Energy Council (GWEC)*, a capacidade acumulada de geração de energia através do vento até dezembro de 2013 era de 318.105 MW. Ainda segundo o GWEC, a China é o país com a maior capacidade energética acumulada através de energia eólica com 91.412 MW, representando 28,7% do total de energia mundial gerada através do vento. Os EUA e Alemanha, tradicionais mercados desse tipo de energia, tinham capacidade de geração, em dezembro de 2013, de 61.091 MW e 34.250 MW, respectivamente, o que representava 19,2% e 10,8% do total mundial de energia eólica produzida, ver figura 1.

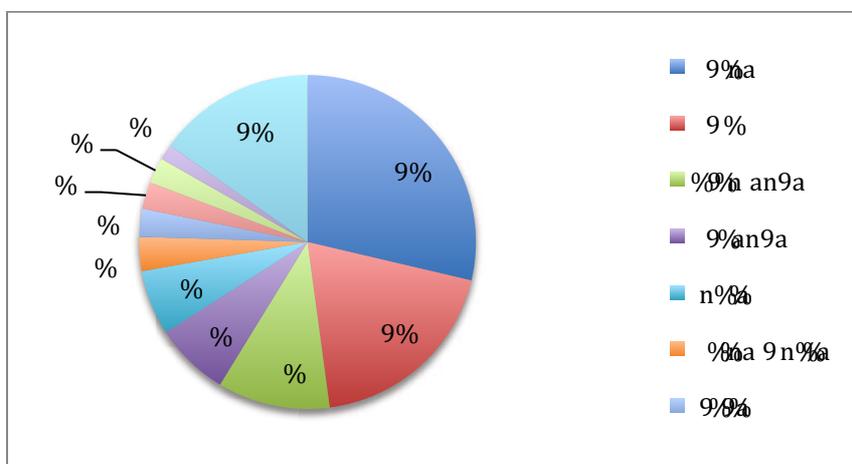


Figura 1 - Distribuição energia eólica mundial 2013

Dados do GWEC, considerando o ano de 2013, mostram que o Brasil figurou entre os 10 países que mais aumentaram a capacidade de geração elétrica via energia eólica, ao aumentar a capacidade de geração em 953 MW. Até o final de 2013 o país possuía capacidade de gerar um total de 3.461 MW.

É importante ressaltar que o Brasil está em uma posição favorável à utilização desse tipo de energia já que está em uma localização geográfica excelente, sobretudo no nordeste do país, onde os ventos sopram de maneira abundante.

Entretanto, apesar de todos os benefícios da utilização da energia eólica, dado sua natureza variável, é difícil a prever a velocidade do vento. Conseqüentemente, uma vez que energia eólica é conectada a um sistema elétrico de porte nacional, essa dificuldade gera implicações negativas nas operações de planejamento, manutenção e despacho. Portanto, existe

uma grande necessidade de se obter estimativas precisas para a velocidade do vento. Este trabalho insere-se neste esforço através da proposta de uma metodologia alternativa para previsão da velocidade do vento de uma estação anemométrica do Nordeste do Brasil.

## 2. Caracterização e contexto

A previsão da velocidade de vento tem papel fundamental para a operação do sistema elétrico. Tal fato advém da seguinte relação que o vento mantém com a capacidade de geração de energia:

$$P = 0,5 \times C_p \times \rho \times A \times V^3 \quad (1)$$

Onde:

$P$  é a potência eólica

$C_p$  é o coeficiente de potência que indica o quanto (%) da potência disponível no vento pode ser aproveitada na turbina

$\rho$  é a densidade do ar

$A$  é a área varrida pelo rotor

$V$  é a velocidade do vento

Fica claro na equação 1 a importância de prever de maneira correta a velocidade do vento, já que o mesmo tem uma relação cúbica com a potência eólica e, portanto, pequenas variações na velocidade do vento podem resultar em grandes variações na potência gerada.

O presente trabalho tem por objetivo apresentar um modelo que seja capaz de prever com maior acurácia a velocidade do vento para geração de energia.

Os dados foram obtidos através do projeto SONDA (Sistema de Organização Nacional de Dados Ambientais). O sistema fornece dados para diversas estações solares, anemométricas ou solares e anemométricas. Cada estação anemométrica fornece os seguintes dados com periodicidade de 10 minutos:

- Velocidade do vento aos 25 metros de altura em m/s;
- Velocidade do vento aos 50 metros de altura em m/s;
- Direção do vento aos 25 metros de altura em graus;
- Direção do vento aos 50 metros de altura em graus;
- Temperatura aos 25 metros de altura em graus Celsius;
- Temperatura aos 50 metros de altura em graus Celsius.

O sistema SONDA não apresentava dados completos em diversas estações. Em algumas estações existiam observações incompletas por meses. Tal problema fez com que a utilização de diversas estações anemométricas fosse inviável já que a falta de dados não seguia um padrão que fosse possível a estimação com certo nível de segurança, como por exemplo: falta de dados em datas diferentes em cada estação.

A estação de São João do Cariri na Paraíba, diferentemente das demais, apresentava dados praticamente completos (exceto em um mês de 2007) e com um nível de confiabilidade alto. Além disso, outros trabalhos já haviam sido feitos na mesma localidade. Ver Chou et al. (2006) e Pessanha, Silva e Souza (2010). Portanto, pelos motivos expostos, foi decidido utilizar os dados dessa estação, com horizonte temporal de 01 de janeiro de 2006 a 31 de dezembro do mesmo ano. A primeira semana de 2007 foi utilizada para testar a capacidade de previsão utilizando a abordagem proposta. O período também foi escolhido de maneira que fosse possível comparar os resultados com trabalhos anteriores na mesma estação anemométrica.

Na próxima seção é apresentado o procedimento do método de previsão proposto neste trabalho, seguido de breve descrição das técnicas utilizadas.

### 3. O procedimento *Bagging Holt Winters* com Decomposição STL

#### a. Decomposição STL (*Seasonal-Trend Decomposition Using Loess*)

O método STL desenvolvido por Cleveland et al.(1990) decompõe a série temporal em três componentes aditivos: Sazonalidade, Tendência e a diferença entre a soma dos dois primeiros e a série temporal, chamado de “*Remainder*”.

Segundo Hyndman e Athanasopoulos (2013), o método apresenta diversas vantagens sobre o método de decomposição clássica X12-ARIMA, dentre os quais podemos citar:

- Possibilidade de lidar com qualquer tipo de sazonalidade (X12-ARIMA permite apenas sazonalidade mensal e trimestral)
- Possibilidade de variação no componente sazonal ao longo do tempo e controle da taxa de variação por parte do usuário.
- Possibilidade de controle da suavidade ciclo-tendência por parte do usuário
- Robustez a *outliers* nos componentes sazonais e de ciclo-tendência.

A ideia do método é aplicar amortecimentos baseados em Regressão Local (*Loess-Locally Weighted Regression*) uma quantidade considerável de vezes gerando estimativas para os componentes de tendência e sazonalidade.

Após a aplicação do método, a decomposição da série pode ser descrita da seguinte maneira:

$$Y_t = T_t + S_t + R_t \quad (2)$$

Onde  $Y_t$  é a série temporal no instante  $t$ ,  $T_t$  é o componente de tendência no instante  $t$ ,  $S_t$  é o componente sazonal em  $t$ , e  $R_t$  é o *Remainder* no mesmo instante.

#### b. *Bootstrap*

O método de reamostragem de *Bootstrap*, criado por Efron (1979), tem sido amplamente utilizado em uma grande variedade de situações onde existe incerteza em relação a um estimador ou a qualidade de um método. Sua grande vantagem reside no fato de ser facilmente implementado em uma série de situações onde o cálculo de algum parâmetro de interesse não pode ser encontrado de maneira trivial.

Apesar de extremamente poderoso, o método foi desenvolvido inicialmente para casos onde os dados são independentes. Singh (1981) apresenta uma prova teórica, atestando a qualidade do método para casos onde os dados são IID. Entretanto, o autor também prova que ao utilizar dados dependentes, o método da forma elaborada por Efron (1979) é inadequado.

Um caso clássico de dados dependentes é o das Séries Temporais, em que a aplicação do método de *Bootstrap* requer cuidados, pois a estrutura de dependência deve ser preservada.

Uma alternativa para contornar o problema é a utilização de blocos. Dessa forma, os blocos são reamostrados mantendo a estrutura de dependência dentro dos blocos. Utilizando essa abordagem temos uma série de métodos desenvolvidos. Dentre os mais utilizados, podemos citar: *Nonoverlapping Block Bootstrap* proposto por Carlstein (1982), *Moving Block Bootstrap* proposto por Künsch (1989), *Circular Block Bootstrap* e o *Stationary Block Bootstrap*, propostos por Politis e Romano (1992) e (1994).

Neste trabalho, apenas a o método *Moving Block Bootstrap* (MBB) foi utilizado. Foram selecionados blocos de tamanho 48 do componente *Remainder*, de maneira que possíveis resquícios de sazonalidade fossem captados.

A ideia do MBB é construir e selecionar aleatoriamente blocos de observações consecutivas. Podemos descrever o método da seguinte forma:

Seja

$$B_i = (Y_i, \dots, Y_{i+l-1})$$

o bloco de tamanho  $l$ , com início na observação  $Y_i$ , de um processo estacionário,  $1 \leq i \leq b$ , onde  $b = n - l + 1$ . Selecionando por amostra aleatória simples com reposição os blocos do conjunto  $\{B_1, \dots, B_b\}$ , temos uma nova amostra formada por  $B_1^* \dots B_k^*$ , onde  $k$  é o número de blocos selecionados. Os elementos contidos em  $B_i^*$  podem ser denotados por  $(Y_{(i-1)l+1}^*, \dots, Y_{il}^*)$ ,  $i = 1, \dots, k$ . Dessa forma, a amostra MBB, de tamanho  $m = k \times l$ , é a seguinte  $Y_1^*, \dots, Y_m^*$ .

### c. Método de Holt Winters

O método de alisamento exponencial inicialmente proposto por Holt (1957) e ampliado por Winters (1960), tem por objetivo captar o comportamento da série, separando seus componentes em tendência, sazonalidade e termo de erro. Podemos definir as equações do método aditivo da seguinte forma:

$$\begin{aligned} \ell_t &= \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \\ b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \\ s_t &= \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m} \\ \hat{y}_{t+h|t} &= \ell_t + hb_t + s_{t-m+h_m^+} \end{aligned} \quad (3)$$

Onde  $\ell_t$ ,  $b_t$ ,  $s_t$  são respectivamente o nível, a inclinação e o componente sazonal da série no tempo  $t$ ,  $m$  é o período da série,  $h$  é o horizonte de previsão e  $h_m^+ = [(h - 1) \bmod m] + 1$ . As constantes  $\alpha$ ,  $\beta^*$  e  $\gamma$  são parâmetros de suavização pertencentes ao intervalo  $[0,1]$ .

### d. Descrição do procedimento

O método de *Bagging* (*Bootstrap Aggregating*), proposto por Breinman (1996) tem sido bastante utilizado no campo de *Machine Learning*. Segundo o autor, o método tem por objetivo gerar múltiplas versões de preditores através de *Bootstrap*. A essas versões podem ser aplicadas as técnicas de interesse, de maneira que os resultados sejam agregados em um único resultado final. O autor ainda argumenta que se perturbações no conjunto de aprendizado geram grande diferenças na configuração dos preditores, então o método de *Bagging* pode melhorar a precisão do que se deseja prever.

Apesar de o uso técnica propiciar uma melhora substancial na capacidade preditiva de métodos como regressão linear e árvores de decisão, poucos trabalhos utilizando a *Bagging* foram desenvolvidos na área de séries temporais.

O procedimento utilizado nesse trabalho segue a recente abordagem de *Bagging* utilizando métodos de amortecimento exponencial proposta por Bergmeir, Hyndman e Benítez (2014).

O procedimento foi feito da seguinte maneira:

- A série temporal de velocidade de vento de São João do Cariri-PB é decomposta utilizando o método STL, resultando em três partes: Componente Sazonal, Tendência e “Remainder”.

- Ao último é aplicado o método *Moving Block Bootstrap*, com 29 replicações. O número de replicações segue a escolha feita por Bergmeir, Hyndman e Benítez (2014). Além disso, nos experimentos conduzidos por Breiman (1996) o número de replicações consideradas necessárias variou de 25 a 50. O autor comenta ainda que um número excessivo de replicações não aumenta a qualidade do resultado.
- A cada uma das replicações são adicionados os componentes sazonais e de tendência, previamente separados, resultando em 29 novas séries.
- A cada uma das séries é aplicado o método de Holt-Winters para prever o número desejado de passos à frente.
- O resultado final é obtido através da média simples de 30 previsões (série original mais 29 séries geradas via *Bootstrap*) e será chamado de Bagged Holt-Winters.

Na próxima seção são apresentados os resultados do método proposto.

#### 4. Resultados

A série em estudo apresenta velocidade média de vento de 5,31m/s e desvio padrão de 2,149m/s, ver Tabela 1. Através do teste de Dickey-Fuller (pvalor menor que 0,01, considerando hipótese alternativa de estacionariedade) podemos concluir que a série é estacionária. Além disso, segundo Custódio (2009), os ventos de uma maneira geral se distribuem segundo uma Distribuição de Weibull. O estudo conduzido por Chou et al. (2006) utiliza a Distribuição de Weibull nesta mesma série com fator de forma igual 2,5 e fator de escala igual a 5.

Medidas de Resumo (m/s)	
Mínimo	0,232
Máximo	11,508
Primeiro Quartil	3,637
Terceiro Quartil	6,857
Média	5,310
Mediana	5,220
Variância	4,617

**Tabela 1 - Medidas de resumo**

A série apresenta um padrão diário onde os ventos sofrem aumento de velocidade a partir de 9 horas chegando ao seu pico máximo por volta de 23 horas onde a partir desse momento começam a diminuir de intensidade, ver figura 2. Além da evidência visual, foram identificados via análise espectral picos nas frequências diárias. Com relação aos meses, é possível verificar que os ventos apresentam intensidades diferentes ao longo do ano. É possível perceber que os ventos sopram com maior intensidade de julho a janeiro, ver figura 3.



A avaliação da capacidade preditiva do modelo foi feita com dados da primeira semana de 2007 (*out of sample*). Para tal, foram utilizadas as métricas de comparação MAPE e RMSE. Os resultados mostram que a abordagem proposta (Bagged Holt-Winters) mostrou-se adequada para previsões, apresentando melhora na capacidade preditiva quando comparado com outros modelos: Modelo Ingênuo, Modelo Neuro-Fuzzy, proposto por Pessanha, Silva e Souza (2010) para a mesma os mesmos dados considerados neste trabalho) e Método de Holt-Winters Simples. A figura 5 apresenta o resultado 1 passo à frente (1 hora) durante toda a primeira semana.

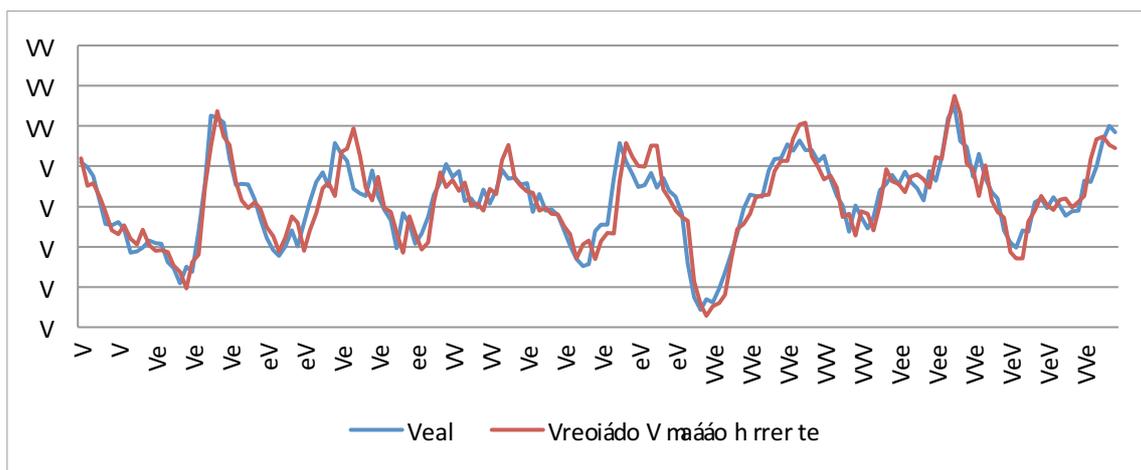


Figura 5 - Previsão 1 passo à frente para a semana de teste (*out of sample*)

A figura 6 apresenta os resultados da previsão quando é aumentada a janela de previsão, durante a primeira semana de 2007, para 6 passos à frente e 24 passos à frente. Os resultados apontam que o modelo ainda consegue prever regularmente bem quando o horizonte de previsão é aumentado, ver tabela 2.

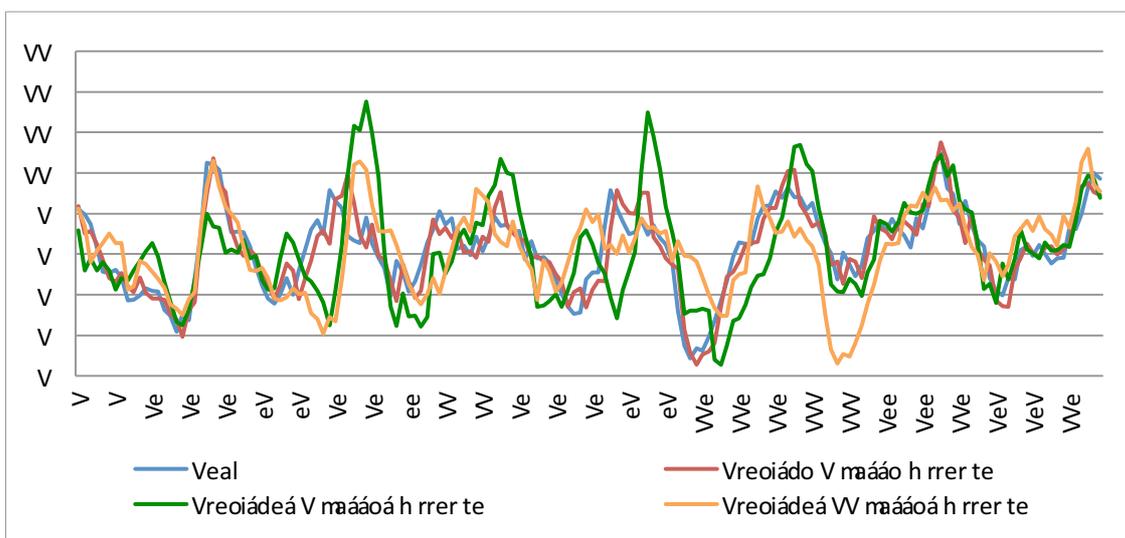


Figura 6 - Previsões Multi-Step para a semana de teste (*out of sample*)

Deve ser destacado que a abordagem utilizada foi capaz de obter melhor desempenho quando comparado com o modelo Ingênuo (previsão é igual ao instante imediatamente anterior), tipicamente usado para comparar a qualidade do modelo. Utilizando as métricas MAPE e RMSE, o modelo apresentado obteve ainda melhor desempenho em comparação com o modelo Neuro-Fuzzy, sendo capaz de prever com maior grau de acerto na comparação 1 passo à frente.

Finalmente, fica evidenciado que a abordagem que considera a agregação de resultados por amostras geradas via *Bootstrap* do componente *Remainder* (Bagged Holt-Winters), apresenta melhores resultados que a aplicação direta do método de Holt-Winters simples, como pode ser visto na Tabela 3.

MODELO	MAPE	RMSE
Bagged Holt-Winters 1 passo à frente	12,94%	0,90m/s
Bagged Holt-Winters 6 passos à frente	30,78%	1,78m/s
Bagged Holt-Winters 24 passos à frente	34,95%	2,08m/s

Tabela 2 – Resultados do modelo na semana de teste

MODELO	MAPE	RMSE
Bagged Holt-Winters	12,94%	0,90m/s
Holt-Winters	13,32%	0,93m/s
Modelo Ingênuo (Naive)	13,74%	0,93m/s
Modelo Neuro Fuzzy	16,63%	1,21m/s

Tabela 3 - Resultados de diferentes modelos na semana de teste 1 passo à frente

## 5. Considerações Finais

A abordagem proposta foi capaz de melhorar a qualidade das previsões de velocidade de vento na estação anemométrica de São João do Cariri-PB. Como era esperado, fica evidenciado que a medida que a janela de previsão aumenta, por exemplo, 24 horas, a qualidade da previsão decresce.

Neste estudo fica claro que a abordagem utilizada não só é capaz de prever com significativa habilidade, como também mostra que utilizar uma combinação de previsões geradas através de aplicação de *Moving Block Bootstrap* no componente *Remainder* da série decomposta via STL e prevista via Holt-Winters aditivo produz melhores resultados preditivos do que a aplicação simples e direta do método de Holt-Winters na série de dados considerada.

Como extensões deste trabalho pretende-se utilizar outras técnicas de decomposição da Série Temporal, além de outros modelos de previsão após a aplicação de técnicas de *Bootstrap* na série decomposta. São ainda objetos de estudos futuros, a utilização de informações de estações anemométricas próximas, além da utilização de variáveis explicativas como temperatura e direção do vento.

## Referências

- Amarante, O. A. C., Brower, M., Zack, J. & de S.a, A. L.** (2001). Atlas do Potencial Eólico Brasileiro. Ministério de Minas e Energia/ Eletrobrás/CEPEL/Cresesb.
- Bergmeir, C., Hyndman, R., Benítez, J.** (2014) Bagging Exponential Smoothing Methods Using STL Decomposition and Box-Cox Transformation. Working Paper 11/14.
- Breiman, L.** (1996) Bagging predictors. *Machine Learning* 24(2):123–140
- Chou, S.C., Chagas, D.J., Paiva Neto, A.C., Ventura, G.M., Lyra, A. & Pereira, E.B.** (2006). Regime eólico de quatro estações no Nordeste. XIV Congresso Brasileiro de Meteorologia, Florianópolis
- Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I.** (1990) STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics* 6:3-73.
- Cordeiro, C., Neves, M.** (2006) The Bootstrap methodology in time series forecasting. In: Rizzi, A., Vichi, M. (eds.) *Proceedings of CompStat2006*, pp. 1067–1073.
- Custódio, R.S** (2009) Energia eólica para produção de energia elétrica. Eletrobras, Rio de Janeiro
- Efron, B.** (1979) Bootstrap methods: another look at the Jackknife. *The Annals of Statistics.*, 7, 1–26
- Global Wind Energy Council** (2014) GWEC Global Wind Report – Annual Market Update 2013.
- Hastie, T., Tibshirani, R., Friedman, J.** (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer
- Holt, C.C.** (1957). Forecasting trends and seasonals by exponentially weighted averages. Carnegie Institute of Technology, Pittsburgh ONR memorandum no. 52.
- Hyndman, R., Athanasopoulos, G.** (2013) *Forecasting: principles and practice*. URL <http://otexts.com/fpp/>
- Hyndman, R.J. & Khandakar, Y.** (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 26(3).
- Hyndman, R, Koehler, A, Snyder, R., Grose, S.,** (2002) A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting* 18(3):439–454
- Künsch, H.R.** (1989) The jackknife and the bootstrap for general stationary observations. *Annals of Statistics* 17(3):1217–1241
- Pessanha, J.F., Silva, I.V.G. & Souza, R.C.** (2010). Previsão de Velocidade de Vento com Sistema de Inferência Neuro-Fuzzy Adaptativo. SPOLM 2010.
- Politis, D.N. & Romano, J.P.** (1992). A circular Block-Resampling Procedure for Stationary Data. *Exploring the Limits of Bootstrap.*, Wiley New York, 263–270



**Politis, D.N. & Romano, J.P.** (1994). The Stationary Bootstrap. Journal of the American Statistical Association., 89, No. 428, 1303–1313

**Winters, P.R.** (1960). Forecasting sales by exponentially weighted moving Average. Manangement Science, 6, 324-342.