

USANDO A ANÁLISE ESPECTRAL SINGULAR NA PREVISÃO DA PRODUÇÃO MENSAL DE UM PARQUE EÓLICO

José Francisco Moreira Pessanha

Universidade do Estado do Rio de Janeiro - UERJ
Departamento de Estatística – Instituto de Matemática e Estatística Sala 6028-B
Rua São Francisco Xavier, 524 – Maracanã 20550-900, Rio de Janeiro – RJ
professorjtmp@hotmail.com

Wander Luiz de Santana Guedes

Universidade do Estado do Rio de Janeiro – UERJ
wander_lz@hotmail.com

Leandro Siller Loureiro

Universidade do Estado do Rio de Janeiro – UERJ
leandrosiller@gmail.com

Moisés Lima de Menezes

Universidade Federal Fluminense - UFF
Departamento de Estatística
Rua Mário Santos Braga s/n - Valonguinho, - Niterói – RJ, 24.020-140
moises_lima@msn.com

Resumo

A entrada em operação de novos aproveitamentos eólicos no Sistema Interligado Nacional (SIN) aponta para a necessidade de desenvolvimento de modelos capazes de prever a produção eólica. Neste trabalho é investigado o desempenho de uma técnica não paramétrica denominada Análise Espectral Singular (SSA) na previsão da produção mensal de um aproveitamento eólico. A aplicação da SSA é ilustrada por meio de uma série temporal de registros de velocidade do vento, disponibilizada pelo Sistema de Organização Nacional de Dados Ambientais (Projeto Sonda). Adicionalmente, compara-se o desempenho alcançado pela SSA com os resultados obtidos pela aplicação da abordagem paramétrica de Box & Jenkins tradicionalmente empregada na modelagem de séries temporais mensais.

Palavras-Chaves: Análise Espectral Singular; Modelagem Box & Jenkins; Energia Eólica.

Abstract

The entry into operation of new wind farms in the National Interconnected System (NIS) points out to the need for developing of wind power forecasting models. This work investigates the performance of a nonparametric technique called Singular Spectrum Analysis (SSA) to predict the monthly wind power output. The implementation of SSA is illustrated by means of a wind speed time series provided by the National Organization Environmental Data System (Sonda Project). Additionally, the performance achieved by the SSA was compared with the results obtained by the parametric approach of Box & Jenkins traditionally used in modeling monthly time series.

Keywords: Singular Spectrum Analysis; Box & Jenkins Modelling; Wind Power.

1. INTRODUÇÃO

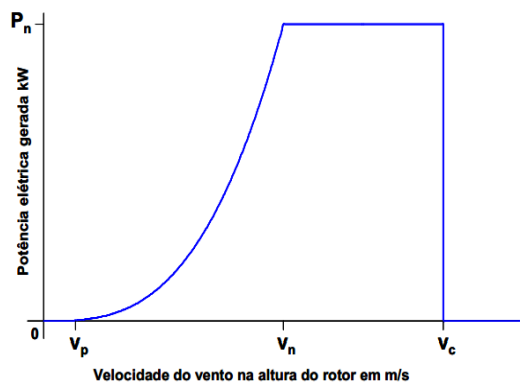
A entrada em operação de novos aproveitamentos eólicos no Sistema Interligado Nacional (BRASIL, 2013) aponta para a necessidade de desenvolvimento de modelos estatísticos capazes de prever a produção eólica. Neste trabalho é investigado o desempenho de uma técnica não paramétrica denominada Análise Espectral Singular ou *Singular Spectrum Analysis* - SSA (GOLYANDINA & ZHIGLJAVSKY, 2013) na previsão da produção mensal de um aproveitamento eólico.

Para ilustrar a aplicação da SSA considerou-se a série temporal formada por registros de velocidade do vento na cidade de Petrolina-PE, local de uma estação anemométrica monitorada pelo Sistema de Organização Nacional de Dados Ambientais (Projeto Sonda) do Instituto Nacional de Pesquisas Espaciais (Inpe). Adicionalmente, o desempenho alcançado pela SSA foi comparado com os resultados obtidos pela aplicação da abordagem paramétrica de Box & Jenkins (MORETTIN & TOLOI, 2006) tradicionalmente empregada na modelagem de séries temporais mensais.

A série temporal da produção eólica mensal analisada neste trabalho corresponde a um parque eólico fictício, mas foi obtida por meio da curva de potência de uma turbina eólica real, aplicada aos valores das medições da velocidade do vento disponibilizados na página do Projeto Sonda (<http://sonda.ccst.inpe.br/>). Destaca-se que todos os resultados foram obtidos com o auxílio dos programas R (R CORE TEAM, 2013) e Gretl (BAIOCCHI & DISTASO, 2003). O artigo está organizado em cinco seções. A seguir, o cálculo da série mensal de produção eólica é descrito na seção 2. Na sequência, na seção 3, tem-se uma breve apresentação da análise espectral singular. Os resultados alcançados são apresentados na seção 4. Por fim, na seção 5 encontram-se as principais conclusões do trabalho.

2. ENERGIA EÓLICA

A relação entre a potência elétrica gerada por uma turbina eólica e a velocidade do vento é apresentada na forma da curva de potência ilustrada na Figura 1, na qual se observa que a conversão de energia eólica em energia elétrica se dá a partir da velocidade de partida v_p , usualmente entre 2,5 a 3 m/s, pois abaixo desses valores o conteúdo energético do vento não justifica o seu aproveitamento (AMARANTE et al, 2001). Ainda na Figura 1, vale notar que a potência da turbina é limitada ao valor nominal P_n , à velocidade nominal v_n , tipicamente velocidades entre 12 e 15 m/s. Para velocidades entre v_p e v_n a potência elétrica gerada por uma turbina eólica cresce com o cubo da velocidade do vento v de acordo com a equação (1), onde η_m é o rendimento mecânico da caixa de transmissão, η_g é o rendimento do gerador, ρ é a densidade do ar e C_p é o coeficiente de potência da turbina. Para velocidades superiores à velocidade de corte v_c , em geral 25 m/s, a turbina é retirada de operação com a finalidade de preservá-la de esforços mecânicos excessivos (CUSTÓDIO, 2009).



$$P(v) = \begin{cases} 0 & v < v_p \\ 0,5 C_p \rho A v^3 \eta_m \eta_g & v_p \leq v < v_n \\ P_n & v_n \leq v < v_c \\ 0 & v \geq v_c \end{cases} \quad (1)$$

Figura 1 - Curva de potência de um aerogerador

As velocidades características v_p , v_n e v_c em (1) podem ser encontradas nos catálogos de fabricantes de turbinas eólicas, juntamente com outros dados técnicos. Na implantação de um aproveitamento eólico é fundamental identificar a turbina que proporcione a extração mais eficiente da potência disponível no vento. A escolha da turbina adequada baseia-se no fator de capacidade, definido pela razão entre a produção média da turbina e a sua produção potencial, se operada constantemente a plena capacidade:

$$FC = P_{média} / P_{máxima} \quad (2)$$

O valor esperado da potência gerada em uma turbina eólica é dado por:

$$P_{média} = \int_0^{\infty} P(v) f(v) dv \quad (3)$$

onde $P(v)$ é a curva de potência em (1) e $f(v)$ é a função densidade de probabilidade de Weibull, determinada pelos parâmetros de escala ($c > 1$) e forma ($k > 0$) (SANSIGOLO, 2005):

$$f(v) = \frac{k}{c} \left[\frac{v}{c} \right]^{k-1} \exp \left[- \left[\frac{v}{c} \right]^k \right] \quad (4)$$

O fator de escala c tem relação direta com a velocidade média, enquanto o fator de forma k é um indicador da constância do vento, dado que quanto maior o valor de k , menor a variância. Tipicamente, k assume valores entre 2 e 3, mas em regiões de ventos alísios, como no Nordeste brasileiro, k pode atingir valores superiores a 6 (AMARANTE et al, 2001). Inserindo (1) em (3) obtém-se a seguinte expressão para o cálculo da potência média.

$$P_{média} = \int_{v_p}^{v_n} 0,5 [C_p \rho_m \rho_g] A v^3 f(v) dv + \int_{v_n}^{v_c} 0,5 [C_p \rho_m \rho_g] A v_n^3 f(v) dv \quad (5)$$

A potência gerada é máxima quando a velocidade alcança o valor nominal v_n :

$$P_{máxima} = 0,5 [C_p \rho_m \rho_g] A v_n^3 \quad (6)$$

Substituindo os resultados (5) e (6) em (2), obtém-se a seguinte fórmula para o cálculo do fator de capacidade (JANGAMSHETTI & RAU, 2001):

$$FC = \left(1/v_n^3 \right) \int_{v_p}^{v_n} v^3 f(v) dv + \int_{v_n}^{v_c} f(v) dv \quad (7)$$

O fator de capacidade está associado com a produção de energia, por exemplo, a produção mensal de energia elétrica pode ser estimada por meio da seguinte fórmula:

$$Energia = FC [P_{máxima} \text{ dias do mês}] \quad (8)$$

A expressão (8) permite estimar uma série temporal de produção eólica mensal, a partir de medições de velocidade do vento e dos parâmetros técnicos de uma turbina eólica. Por exemplo, na Fig. 2 apresenta-se a série temporal com estimativas da produção mensal de um parque eólico fictício com 75 aerogeradores a 125 m de altura em Petrolina-PE. No cálculo dessa série temporal foram considerados os registros anemométricos disponibilizados pelo Projeto Sonda para Petrolina e os dados técnicos do aerogerador Vestas V90 de 1.8 MW, cujas velocidades v_p , v_n e v_c são 4m/s, 12 m/s e 25 m/s respectivamente.

As medições de velocidade do vento disponibilizadas pelo Projeto Sonda são referentes às alturas de 25 m e 50 m. A velocidade na altura de 125 m (v_{125}) foi estimada por meio da lei da potência (CUSTÓDIO, 2009), a partir das medições em 25 m (v_{25}) e 50 m (v_{50}), conforme indicado a seguir, onde $\Delta = \ln(v_{50}/v_{25}) / \ln(50/25)$ depende das velocidades médias $\overline{v_{25}}$ e $\overline{v_{50}}$:

$$v_{125} = v_{50} (125/50)^\Delta \quad (9)$$

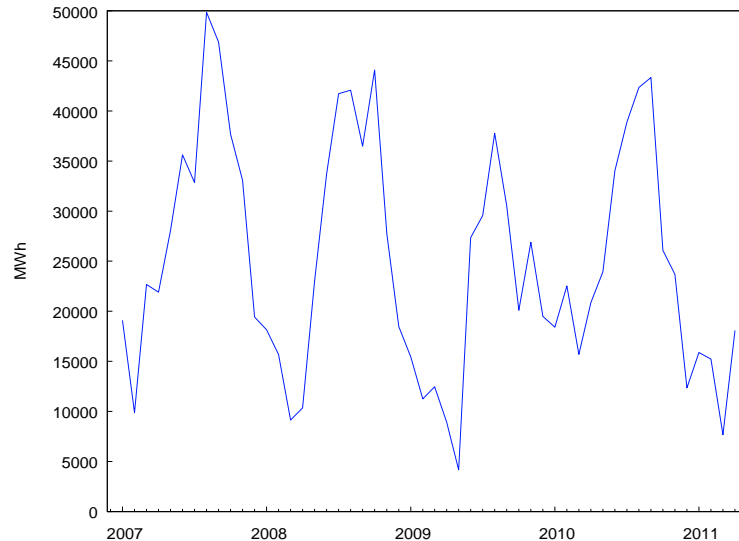


Figura 2 - Produção mensal de energia eólica entre janeiro/2007 e abril/2011

3. ANÁLISE ESPECTRAL SINGULAR

A Análise Espectral Singular (*Singular Spectrum Analysis - SSA*) é uma técnica não-paramétrica para análise e previsão de séries temporais (GOLYANDINA & ZHIGLJAVSKY, 2013). A SSA dispensa a premissa de estacionariedade da série temporal e sua aplicação envolve os estágios de decomposição e reconstrução, descritos resumidamente a seguir.

3.1. DECOMPOSIÇÃO

A decomposição de uma série temporal x_1, x_2, \dots, x_N é realizada em duas etapas: embutimento e decomposição em valores singulares (*singular value decomposition - SVD*).

Na etapa de embutimento uma janela móvel de tamanho L percorre a série temporal e a cada passo da janela obtém-se uma coluna da matriz trajetória em (10), cujo total de colunas é igual a $K=N-L+1$.

$$X = [X_1 \quad \dots \quad X_K] = (x_{ij})_{i,j=1}^{L,K} = \begin{matrix} \begin{matrix} x_1 & x_2 & \dots & x_K \end{matrix} \\ \begin{matrix} x_2 & x_3 & \dots & x_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \dots & x_N \end{matrix} \end{matrix} \quad (10)$$

O tamanho da janela L é um número inteiro no intervalo $2 \leq L \leq N-1$, sendo, em geral igual a $N/2$, se N é par, ou $(N+1)/2$, se N é ímpar (GOLYANDINA & ZHIGLJAVSKY, 2013). Por meio da SVD, a matriz trajetória é decomposta em uma soma de matrizes E_i $i=1, \dots, d$, em que d é o número de valores singulares não nulos:

$$X = \sum_{i=1}^d E_i = \sum_{i=1}^d \sqrt{\lambda_i} U_i V_i^T \quad (11)$$

onde U_i e λ_i $i=1, \dots, d$ correspondem, respectivamente, aos autovetores e autovalores da matriz $S = XX^T$, enquanto os vetores $V_i = X^T U_i / \sqrt{\lambda_i}$ $i=1, \dots, d$ correspondem às componentes principais da matriz S . Cada conjunto $\{\lambda_i, U_i, V_i\}$ $i=1, \dots, d$ é denominado uma *autotripla*.

3.2. RECONSTRUÇÃO

Inicialmente, as d componentes E_i são agrupadas em $m < d$ conjuntos disjuntos e mutuamente exclusivos, por exemplo, componentes de tendência, sazonalidade e irregular (ruído). Assim, para um grupo formado pela agregação de p componentes, $I = \{i_1, \dots, i_p\}$, obtém-se a matriz resultante Y_I :

$$Y_l = E_{i_1} + \dots + E_{i_p} \quad (12)$$

As matrizes resultantes são computadas para os grupos I_1, \dots, I_m , logo a expressão (11) pode ser reescrita como segue:

$$X = Y_{I_1} + \dots + Y_{I_m} \quad (13)$$

Na última fase, por meio do procedimento conhecido por média diagonal (14), cada matriz $Y_{I_i} \quad i=1, \dots, m$ é convertida em uma série temporal de tamanho N . Tal procedimento corresponde ao cálculo das médias envolvendo elementos y_{ij} cujas somas de índices $i+j$ são iguais.

$$\tilde{x}_q = \begin{cases} (1/q) \sum_{m=1}^q y_{m,q-m+1} & 1 \leq q \leq \min(L, K) \\ (1/\min(L, K)) \sum_{m=1}^{\min(L, K)} y_{m,q-m+1} & \min(L, K) \leq q \leq \max(L, K) \\ \frac{1}{N - \max(L, K) + 1} \sum_{m=q-\max(L, K)+1}^{N-\max(L, K)+1} y_{m,q-m+1} & \max(L, K) < q \leq N \end{cases} \quad (14)$$

As médias diagonais da matriz $Y_j \quad j=1, \dots, m$ formam a série $\tilde{X}^{(j)} = (\tilde{x}_1^{(j)}, \dots, \tilde{x}_N^{(j)})$. Portanto, a série inicial x_1, \dots, x_N é decomposta em uma soma de m séries:

$$x_n = \sum_{j=1}^m \tilde{x}_n^{(j)} \quad n = 1, \dots, N \quad (15)$$

Removendo as componentes irregulares (ruídos) obtém-se uma versão suavizada da série temporal, ou seja, uma estimativa do sinal \hat{x} contido na série temporal. Admitindo que o sinal seja formado pelas r primeiras componentes, o sinal \hat{x} é definido pela soma das r primeiras séries, enquanto o ruído $\underline{\hat{x}}$ corresponde ao total das demais componentes. Assim, uma série temporal x pode ser decomposta em duas parcelas aditivas $\hat{x} + \underline{\hat{x}}$:

$$\hat{x}_n = \sum_{j=1}^r \tilde{x}_n^{(j)} \quad n = 1, \dots, N \quad (16)$$

$$\underline{\hat{x}}_n = \sum_{j=r+1}^m \tilde{x}_n^{(j)} \quad n = 1, \dots, N \quad (17)$$

O conjunto das r componentes que formam o sinal deve apresentar elevada separabilidade em relação às demais componentes constituintes do ruído. A separabilidade entre as componentes (HASSANI, 2007) é avaliada pelo coeficiente de correlação ponderada (w -correlation) em (18). Portanto, idealmente, os coeficientes $\underline{\hat{x}}_{i,j}^{(w)}$ entre as r componentes que compõem o sinal e as componentes de ruído devem nulos.

$$\underline{\hat{x}}_{i,j}^{(w)} = \frac{\sum_{k=1}^N \min(k, L, N-k) \tilde{x}_k^{(i)} \tilde{x}_k^{(j)}}{\sqrt{\sum_{k=1}^N \min(k, L, N-k) \tilde{x}_k^{(i)2}} \sqrt{\sum_{k=1}^N \min(k, L, N-k) \tilde{x}_k^{(j)2}}} \quad (18)$$

3.3. PREVISÃO

Para realizar previsões por meio da SSA é necessário que a série satisfaça as relações de recorrência linear (*linear recurrence relations* - LRR), conforme indicado em Golyandina e Zhigljavsky (2013). Uma série $X_N = (x_1, \dots, x_N)$ satisfaz a LRR de ordem q se:

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_q x_{t-q} \quad t = q+1, \dots, N \quad (19)$$

O principal pressuposto é que o comprimento da janela L foi escolhido de forma a separar o sinal do ruído. As previsões são realizadas utilizando-se as r autotriplas escolhidas, conforme indicado a seguir:

$$\hat{x}_i = \begin{cases} \tilde{x}_i & i = 1, \dots, N \\ \sum_{j=1}^{L-1} a_j \hat{x}_{i-j} & i = N+1, \dots, N+M \end{cases} \quad (20)$$

em que o vetor de coeficientes $(a_1, a_2, \dots, a_{L-1})$ é dado por:

$$\frac{1}{\sum_{q=1}^r \sum_{j=1}^{L-1} \Delta_q^2 U_q^j} \quad (21)$$

onde U_q^j denota os primeiros $L-1$ elementos do vetor U_q , enquanto, Δ_q é o último elemento.

4. MODELAGEM DA PRODUÇÃO EÓLICA MENSAL

4.1. MODELAGEM ARIMA

A abordagem de Box & Jenkins é frequentemente empregada na previsão de séries temporais mensais e por esta razão serve de referência nas avaliações comparativas com outros métodos de previsão. Para verificar a qualidade das previsões, as quatro últimas observações (janeiro a abril/2011) foram omitidas no processo de identificação e estimação dos modelos. Em função da sazonalidade da produção eólica, considerou-se um modelo SARIMA de período igual a 12. Após a análise das funções de autocorrelação (Figuras 3 e 4) e a realização de testes de sobrefixação dos parâmetros (SOUZA & CAMARGO, 1996), o modelo SARIMA (1,0,0) x (1,1,0)₁₂ foi identificado como sendo o melhor modelo para descrever o comportamento da série analisada. As estimativas dos coeficientes do modelo obtidas pelo *software* Gretl encontram-se na Tabela 1.

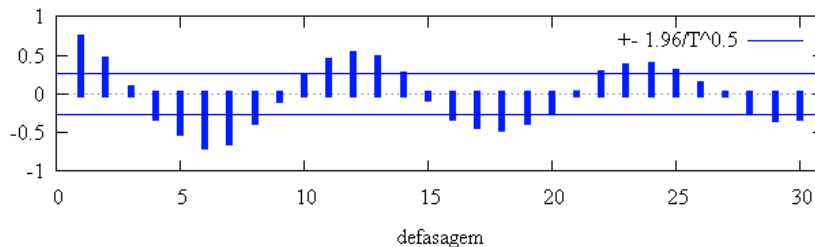


Figura 3 - FAC amostral da produção eólica mensal

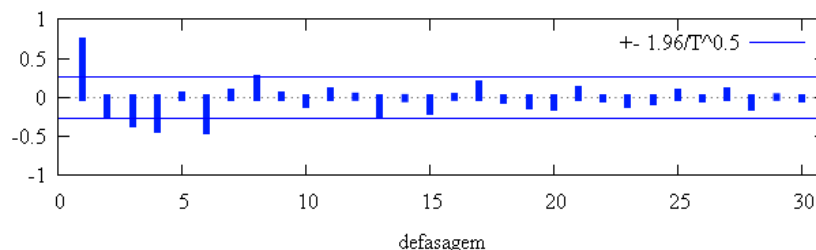


Figura 4 - FACP amostral da produção eólica mensal

Tabela 1 – Estimativas dos coeficientes do modelo SARIMA (1,0,0) x (1,1,0)₁₂

Parâmetro	Coefficiente	Erro Padrão	t	p-valor	Critério AIC
-----------	--------------	-------------	---	---------	--------------

$\hat{\alpha}_1$	0.5099	0.1442	3.537	0,0004	749,10
$\hat{\beta}_1$	-0.5511	0.1452	-3.796	0,0001	

Além da significância estatística do modelo ajustado, os resíduos não são autocorrelacionados (Figura 5) e seguem uma distribuição normal (Figura 6). Adicionalmente, na Figura 7 observa-se a boa aderência das estimativas e previsões aos valores realizados.

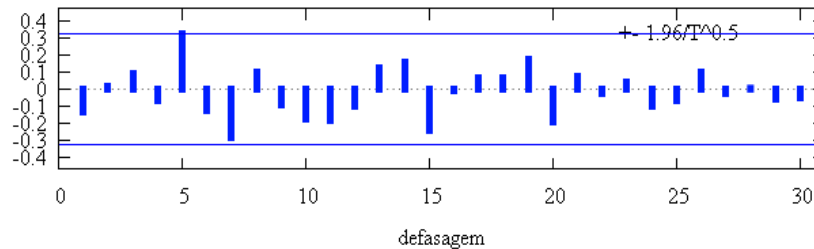


Figura 5 - FAC dos resíduos do modelo SARIMA(1,0,0)x(1,1,0)₁₂

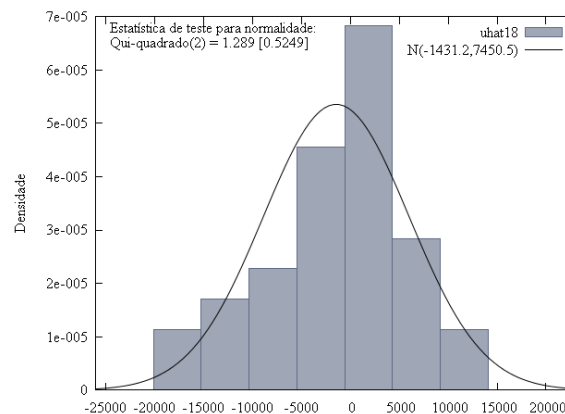


Figura 6 - Teste de normalidade dos resíduos

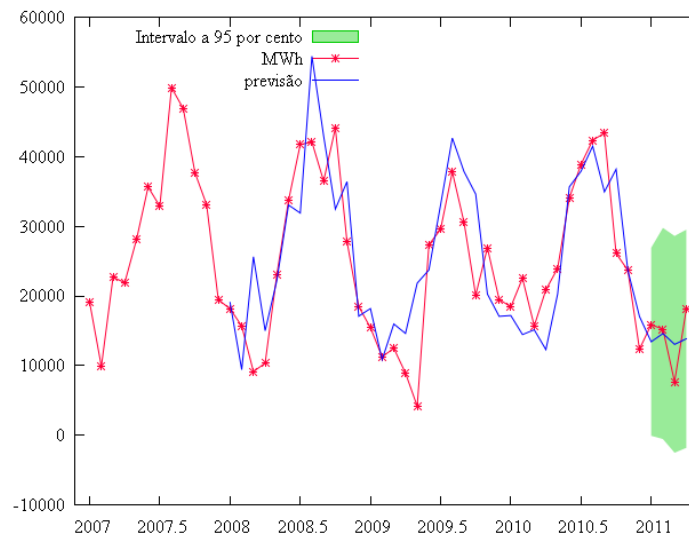


Figura 7 - Estimativas e previsões segundo o modelo SARIMA (1,0,0) x (1,1,0)₁₂

4.2. MODELAGEM SSA

A modelagem foi realizada em ambiente R por meio do pacote Rssa. Dado que a série tem 48 observações, considerou-se uma janela de tamanho $L=24$. Na Figura 8a são apresentados os logaritmos dos 24 autovalores dispostos em ordem decrescente. Segundo HASSANI (2007), o primeiro valor singular está associado à componente de tendência,

enquanto os autovalores com magnitudes semelhantes estão relacionados com períodos harmônicos específicos. A proximidade dos valores singulares 2 e 3 indica que eles estão associados com a componente sazonal da série. Além disso, após o quarto valor singular, visualiza-se um decaimento lento e contínuo que caracteriza o início do ruído.

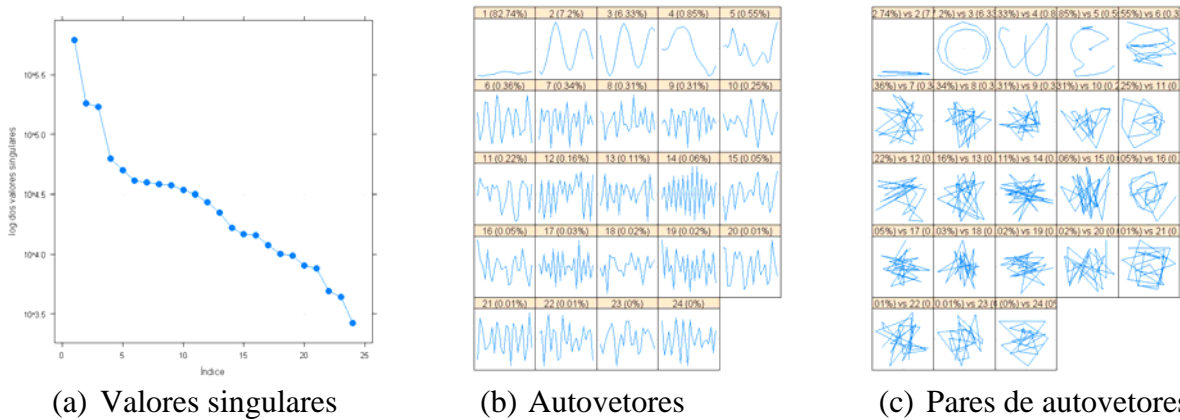


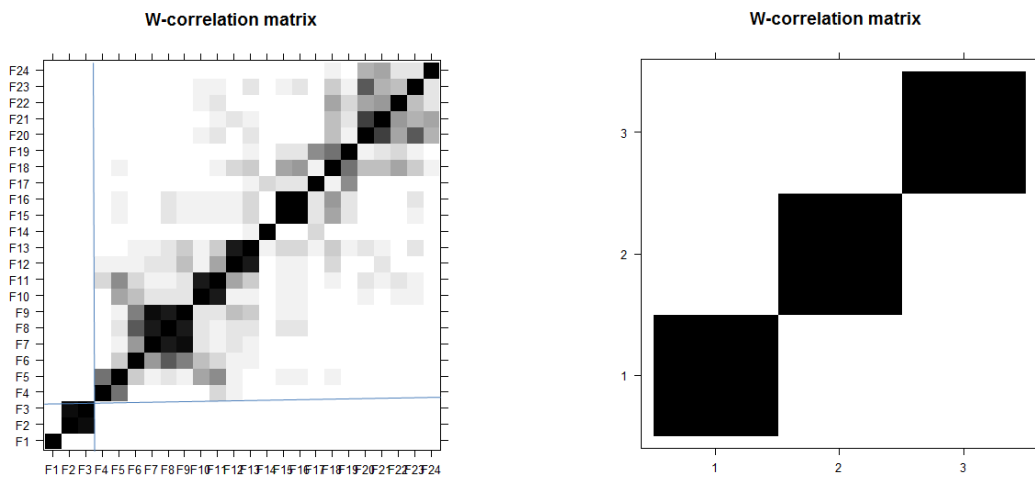
Figura 8 – Resultados da decomposição realizada pela SSA

Na Figura 8b observa-se que o primeiro autovetor captura a tendência da série, enquanto os autovetores 2 e 3 apresentam perfis que sugerem sazonalidade. Ainda na Figura 8b observa-se que até o quarto autovetor as curvas são suaves, enquanto que a partir do quinto autovetor as curvas tendem a apresentar vértices, um indício de que nesse ponto pode haver uma separação entre o sinal e o ruído. Verifica-se também que os três maiores autovetores concentram mais de 95% da variação total da série temporal analisada.

Os gráficos ilustrados na Figura 8c são utilizados na identificação dos períodos das componentes harmônicas da série temporal. Neste caso, é possível identificar formas geométricas nos pares de autovetores (2-3). Tal observação sugere que os autovetores 2 e 3 constituem componentes harmônicas. Por outro lado, a partir do par (3-4) na primeira linha e terceira coluna da Figura 8c, os gráficos resultantes dos cruzamentos dos autovetores apresentam formas sem nenhuma possibilidade de identificação, uma observação que ratifica a ideia do ruído iniciar-se a partir da quarta componente.

Na Figura 9a tem-se a representação gráfica da matriz de correlações ponderadas (w -correlation), na qual se pode verificar que a primeira componente não possui correlação com nenhuma outra componente. Observa-se também, que as autotriplas 2 e 3 são altamente correlacionadas, o que ratifica a interpretação anterior sobre a componente sazonal. Ainda na Figura 9a, a linha vertical separa as componentes do sinal e do ruído a partir da terceira componente. Já na Figura 9b visualiza-se a boa separação entre as componentes tendência (primeira autotripla), sazonal (autotriplas 2 e 3) e ruído (demais autotriplas).

As componentes identificadas são ilustradas na Figura 10. O sinal contido na série é ilustrado pela curva tracejada na Figura 11 e corresponde à soma das componentes de tendência e sazonalidade.



(a) Entre todas as componentes

(b) Entre tendência, sazonalidade e ruído

Figura 9 - Matriz w-correlação

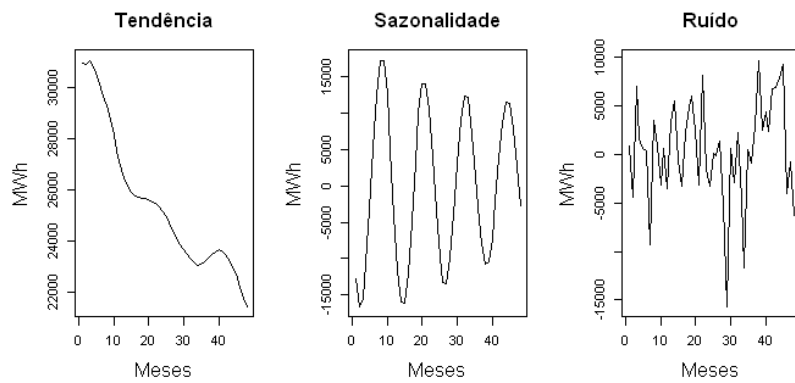


Figura 10 – Componentes da série analisada

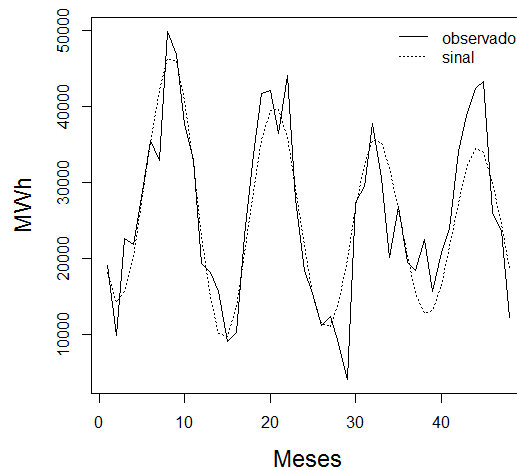


Figura 11 – Série analisada e sinal

Na sequência, na Figura 12 são apresentados os valores suavizados ao longo do período de janeiro de 2007 até dezembro de 2010 e as previsões para o período de janeiro até abril de 2011 obtidas pelo algoritmo de previsão recorrente.

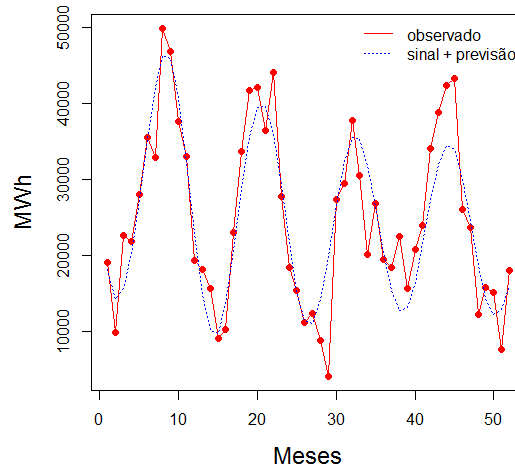


Figura 12 – Estimativas e previsões da produção eólica mensal

Os resultados apresentados acima tomaram por base uma janela com tamanho $L=24$ e $r=3$. Conforme sugerido por Pepelyshev & Zhigljavsky (2010) vale avaliar a estabilidade das previsões para diferentes valores dos parâmetros L e r . As previsões obtidas para diferentes valores de L e r , juntamente com as raízes dos respectivos erros quadráticos médios (RMSE – *Root Mean Squared Error*) no período de previsão, são apresentados na Tabela 2.

Tabela 2 – Resultados da SSA para diferentes valores de L e r

Data	$L=24$ $r=3$	$L=24$ $r=4$	$L=20$ $r=3$	$L=20$ $r=4$	$L=16$ $r=3$	$L=16$ $r=4$	$L=12$ $r=3$	$L=12$ $r=4$	Observado
Jan/2011	14.397	18.285	16.223	21.516	15.831	16.306	15.044	7.186	15.879
Fev/2011	12.282	16.042	14.322	20.246	14.330	14.453	13.083	4.358	15.220
Mar/2011	12.946	16.461	14.922	21.536	15.386	15.770	14.354	5.989	7.672
Abr/2011	16.133	19.282	17.781	25.039	18.687	19.831	18.488	13.029	18.048
RMSE	3.252	4.616	3.659	8.633	3.896	4.169	3.540	7.443	

Na Tabela 2, vale notar que o menor RMSE ocorre quando $L=24$ e $r=3$, embora neste caso não sejam observadas as melhores previsões em cada mês. Ainda na Tabela 2, observa-se que o RMSE com 4 componentes é sistematicamente superior ao obtido quando $r=3$, independentemente do tamanho da janela L . Os resultados também indicam que o RMSE é mais sensível ao número de componentes r e menos sensível ao tamanho da janela L .

Outra forma para avaliar a estabilidade das previsões e ainda obter intervalos de previsão com 95% de confiança consiste em utilizar técnicas de reamostragem *Bootstrap* (GOLYANDINA & ZHIGLJAVSKY, 2013). Admitindo que os resíduos ζ sejam estacionários, a distribuição empírica dos resíduos é usada para gerar Q simulações *bootstrap* da série de ruídos. Na sequência, as séries de resíduos simuladas são somadas ao sinal \hat{x} estimado pela SSA e obtidas Q séries temporais $x_q = \hat{x} + \zeta_q$ $q=1, \dots, Q$. De cada série x_q é gerada uma previsão pelo método LRR h passos à frente e, ao final, são produzidas Q previsões. Os quantis de 2,5% e 97,5% das distribuições empíricas das previsões definem os limites dos intervalos de previsão. A média das Q previsões fornece uma estimativa do valor esperado da série temporal h passos à frente. A seguir, na Figura 13 são apresentados os intervalos de previsão com 95% de confiança no período de janeiro até abril de 2011, gerados por meio de reamostragem *Bootstrap* com Q igual a 10.000 simulações.

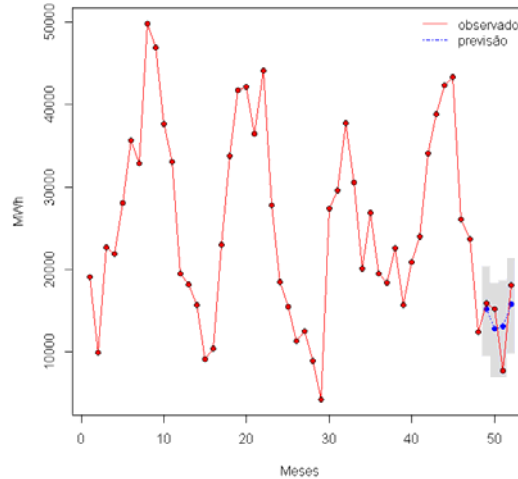


Figura 13 – Previsões e intervalos de previsão da produção eólica mensal

4.3. COMPARAÇÃO DE RESULTADOS

A seguir, na Tabela 3 são apresentadas as previsões e os valores do RMSE alcançados pelos distintos modelos avaliados para o período de janeiro a abril de 2011. A SSA (com ou sem reamostragem *Bootstrap*) alcançou um RMSE ligeiramente menor que o encontrado pelo modelo ARIMA. Adicionalmente, os intervalos de previsão na Tabela 4, mostram o desempenho superior da abordagem SSA com *Bootstrap*, pois esta obteve intervalos mais estreitos que os encontrados pelo modelo ARIMA e com limites positivos.

Tabela 3 - Valores previstos para o período de janeiro a abril/2011 pelos três modelos

Mês/Ano	Série Real MWh	SARIMA		SSA		SSA + <i>Bootstrap</i>	
		Previsão	Diferença	Previsão	Diferença	Previsão	Diferença
Jan/2011	15.878,6	13.410,1	2.469,5	14.397,4	1.481,2	15.209,6	669,0
Fev/2011	15.220,4	14.596,9	623,5	12.282,4	2.938,0	12.819,3	2.401,1
Mar/2011	7.671,9	13.042,0	-5.370,1	12.946,3	-5.274,4	13.079,0	-5.407,2
Abr/2011	18.047,8	13.839,8	4.208,0	16.132,9	1.914,9	15.825,2	2.222,6
RMSE		3.641,0		3.252,0		3.177,7	

Tabela 4 – Intervalos de previsão

Mês/Ano	Série Real MWh	SARIMA		SSA + <i>Bootstrap</i>	
		Limite inferior	Limite superior	Limite inferior	Limite superior
Jan/2011	15.878,6	-88,5	26.908,6	9.422,7	20.417,8
Fev/2011	15.220,4	-555,5	29.749,3	6.756,9	18.367,0
Mar/2011	7.671,9	-2.511,7	28.595,7	6.836,4	18.635,4
Abr/2011	18.047,8	-1.816,6	29.496,2	9.675,0	21.326,9

5. CONCLUSÕES

A perspectiva de maior inserção de fontes eólicas no Sistema Interligado Nacional (SIN) aponta para a necessidade de desenvolvimento de modelos capazes de prever a produção de energia em parques eólicos. Neste trabalho foi investigado o uso da Análise Espectral Singular (SSA) na previsão da produção eólica mensal, uma informação relevante para o planejamento de médio prazo da operação do SIN e para a comercialização de energia. A SSA é uma alternativa não paramétrica aos tradicionais modelos ARIMA e suas variantes.

Em termos de previsões pontuais, os resultados apresentados neste trabalho mostram que as previsões calculadas pela SSA são compatíveis com as encontradas pela abordagem ARIMA. Contudo, os intervalos de previsão obtidos pela SSA com *Bootstrap* foram mais estreitos que os encontrados pela modelagem ARIMA.

Apesar dos resultados promissores obtidos pela SSA, no presente trabalho foi analisada a série temporal de um parque eólico fictício. Portanto, mais investigações devem ser realizadas com séries temporais da produção de parques eólicos em operação. Em estudos futuros recomenda-se comparar o desempenho da SSA com os modelos auto-regressivos periódicos PAR(p), os mesmos utilizados na modelagem de vazões e energia natural afluyente no contexto do planejamento da operação energética do SIN (MACEIRA et al, 2006, MENEZES et al, 2014). A possibilidade de combinar a SSA com a reamostragem *Bootstrap* pode ser uma alternativa para a geração de séries sintéticas da produção eólica.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] AMARANTE, O. A. C.; BROWER, M.; ZACK, J.; SÁ, A.L. *Atlas do Potencial Eólico Brasileiro*. Brasília: Ministério de Minas e Energia, 2001.
- [2] BAIOCCHI, G.; DISTASO, W. GRETL: Econometric software for the GNU generation, *Journal of Applied Econometrics*, 18, p. 105-110. 2003.
- [3] BRASIL, Ministério de Minas e Energia, Empresa de Pesquisa Energética (2013) *Plano Decenal de Expansão de Energia 2022*, Brasília: MME/EPE. Página consultada a 18.06.2014 em http://www.epe.gov.br/PDEE/20140124_1.pdf.
- [4] CUSTÓDIO, R. S. *Energia eólica para produção de energia elétrica*. Rio de Janeiro: Eletrobrás. 2009.
- [5] GOLYANDINA, N.; ZHIGLJAVSKY, A. *Singular Spectrum Analysis for Time Series*. New York: Springer. 2013.
- [6] HASSANI, H. Singular Spectrum Analysis: Methodology and Comparison. *Journal of Data Science*, 5, p. 239 - 257, January. 2007.
- [7] JANGAMSHETTI, S.H.J.; RAU, V.G. Optimum siting of wind turbine generators. *IEEE Transactions on Energy Conservation*, v. 16, n. 1, p. 8-13, March. 2001.
- [8] MACEIRA, M.E.P.; PENNA, D.D.J.; DAMÁZIO, J.M. Geração de cenários sintéticos de energia e vazão para o planejamento da operação energética, *Cadernos do IME, Série Estatística*, Rio de Janeiro, v. 21, p. 11-35, Dezembro. 2006.
- [9] MENEZES, M.L.; SOUZA, R.C.; PESSANHA, J.F.M. Combining singular spectrum analysis and PAR(p) structures to model wind speed time series, *Journal of Systems Science & Complexity*, 27, p. 29-46. 2014.
- [10] MORETTIN, P. A.; TOLOI, C. M. C. *Análise de Séries Temporais*. 2. ed. São Paulo: Ed. Blucher. 2006.
- [11] PEPELYSHEV, A.; ZHIGLJAVSKY, A. Assessing the stability of long-horizon SSA forecasting, *Statistics and Its Interface*, 3, p. 321-327. 2010.
- [12] R CORE TEAM (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>. 2013
- [13] SANSIGOLO, C. A. Distribuição de probabilidade de velocidade e potência do vento, *Revista Brasileira de Meteorologia*, v. 20, n. 2, p. 207-214. 2005.
- [14] SOUZA, R.C.; CAMARGO, M. E. *Análise e Previsão de Séries Temporais: Os Modelos ARIMA*. Ijuí, RS: SEDIGRAF. 1996.