

COMBINAÇÃO GEOMÉTRICA DE MODELOS BOX & JENKINS COM TEORIA WAVELET E ANÁLISE ESPECTRAL SINGULAR

Keila Mara Cassiano

Universidade Federal Fluminense
Rua Mário Santos Braga S/N - Valonguinho, - Niterói – RJ, 24.020-140.
keilamath@hotmail.com

Luiz Albino Teixeira Júnior

Universidade Federal da Integração Latino Americana
Av. Tancredo Neves, 6731, Foz do Iguaçu – PR, 85.867-970.
luiz.a.t.junior@gmail.com

Moisés Lima de Menezes

Universidade Federal Fluminense
Rua Mário Santos Braga S/N - Valonguinho, - Niterói – RJ, 24.020-140.
moises_lima@msn.com

Reinaldo Castro Souza

Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225, Gávea, Rio de Janeiro – RJ, 22.451-900.
reinaldo@ele.puc-rio.br

José Francisco Moreira Pessanha

Universidade do Estado do Rio de Janeiro
Rua São Francisco Xavier, 524, Maracanã, Rio de Janeiro – RJ, 20.550-013.
professorjfm@hotmail.com

RESUMO

Neste artigo são comparados quatro métodos de previsão para séries temporais: Box & Jenkins (BJ); Box & Jenkins com Teoria Wavelet (BJ-TW); Box & Jenkins com Análise Espectral Singular (BJ-SSA); e Combinação Geométrica dos métodos BJ-TW e BJ-SSA. Os métodos foram aplicados à série temporal do consumo residencial mensal de eletricidade em uma parte do Estado do Rio de Janeiro, Brasil. Os resultados mostram que a Combinação Geométrica dos métodos BJ-TW e BJ-SSA melhorou significativamente o desempenho sobre todos os outros métodos.

Palavras-chave: Modelagem Box & Jenkins, Teoria Wavelet, Análise Espectral Singular, Combinação Geométrica, Previsões Pontuais.

ABSTRACT

In this paper, we compare four forecasting methods: Box and Jenkins (BJ); Box and Jenkins with Wavelet Theory (BJ-WT); Box and Jenkins with Singular Spectrum Analysis (BJ-SSA); and Geometric Combination of BJ-WT and BJ-SSA methods. The methods were applied to the time series of monthly residential consumption of electricity in a part of the State of Rio de Janeiro, Brazil. The results show that the Geometric Combination of BJ-WT and BJ-SSA methods significantly improved the performance over all other methods.

Keywords: Box & Jenkins Modeling, Wavelet Theory, Singular Spectrum Analysis, Geometric Combination, Point Forecasts.

INTRODUÇÃO

A previsão do consumo de eletricidade tem um papel importante na expansão e na operação do sistema elétrico brasileiro [20]. Usualmente, ao invés de se prever o consumo total de eletricidade de determinada região, faz-se uma segmentação de seu mercado, dividindo-o em distintas classes de consumo (como, por exemplo, residencial, industrial, comercial) e, após isso, são efetivamente geradas as previsões para cada uma. Na perspectiva de previsão de séries temporais, [6] propôs a combinação geométrica de modelos preditivos bayesianos, o que implicou a necessidade de se obter pesos adaptativos com vistas a otimizar a estatística MSE (Mean Square Error). Uma vez que a Teoria Wavelet e a Análise Espectral Singular (SSA -Singular Spectrum Analysis) podem ser utilizadas em abordagens de pré-processamento de séries de tempo, com a finalidade de se obter ganhos no processo preditivo [1, 2, 4, 5, 6, 7, 9, 11, 15, 16, 18], este trabalho propõe a combinação geométrica de previsões de dois modelos Box & Jenkins: um integrado à Teoria Wavelet (para decomposição da série temporal de interesse antes de sua modelagem) e o outro integrado a SSA (para remoção de ruído da série temporal de interesse antes de sua modelagem). Estas duas abordagens são denotadas, respectivamente, por BJ-TW e BJ-SSA. Para a implementação dos referidos métodos, dada a sua relevância, utilizou-se a série temporal de consumo residencial mensal de eletricidade de uma concessionária de distribuição de energia elétrica que atende grande parte do Estado do Rio de Janeiro, Brasil. O presente artigo está estruturado em nove seções. Na Seção 2, apresenta-se a Teoria Wavelet e, na Seção 3, a Análise Espectral Singular. Na Seção 4 é apresentada uma breve descrição dos modelos Box & Jenkins. Na Seção 5 é definida a combinação geométrica de previsões pontuais. O método proposto, por sua vez, é apresentado na Seção 6. Na Seção 7, apresenta-se o estudo de caso. Por fim, tem-se os resultados e as discussões, na Seção 8, e, na Seção 9, as conclusões.

1. TEORIA WAVELET

2.1. VETOR WAVELET

Seja H um espaço de Hilbert (complexo, de dimensão infinita e, por simplicidade, separável [13]. Por operador sobre H , entende-se como sendo uma transformação linear e limitada de H para H [14]. Suponha que os operadores D e T satisfaçam aos seguintes axiomas: (A_0) D e T sejam shifts bilaterais de multiplicidade infinita (ou, simplesmente, shifts) sobre H [13]; e (A_1) $DT \neq TD$ e $DT^2 = TD$.

De acordo com [14], se D e T são shifts sobre H que satisfazem aos axiomas (A_0) e (A_1) , então qualquer vetor $\omega \neq 0 \in H$ que gera uma base ortonormal $\{D^m T^n \omega\}_{(m,n) \in \mathbb{Z} \times \mathbb{Z}}$ para H é um vetor ortonormal wavelet (ou, simplesmente, wavelet) e os vetores $\omega_{(m,n)} := D^m T^n \omega$, para todo $(m, n) \in \mathbb{Z} \times \mathbb{Z}$ são vetores ortonormais wavelet gerados a partir de ω . De acordo com [16], o subespaço fechado $W_m(\omega) := (\text{span}\{D^m T^n \omega\}_{n \in \mathbb{Z}})^\perp$ de H , onde $m \in \mathbb{Z}$, é chamado de subespaço de detalhes. Por sua vez, se D e T são shifts sobre H que satisfazem aos axiomas (A_0) e (A_1) , então qualquer vetor $\phi \neq 0 \in H$ que gera um subconjunto $\{D^m T^n \phi\}_{(m,n) \in \mathbb{Z} \times \mathbb{Z}} \subset H$ tal que $\langle D^{m'} T^{n'} \phi, D^j T^k \phi \rangle = 0$, sempre que $m' = j$ e $n' \neq k$, onde $m', n', j, k \in \mathbb{Z}$, e $\langle D^{m'} T^{n'} \phi, D^j T^k \phi \rangle \neq 0$, se caso contrário, é chamado de vetor escala e os vetores $\phi_{m,n} := D^m T^n \phi$, para todo $(m, n) \in \mathbb{Z} \times \mathbb{Z}$, são vetores escala gerados a partir de ϕ . De acordo com [16], o subespaço fechado $V_m(\phi) := (\text{span}\{D^m T^n \phi\}_{n \in \mathbb{Z}})^\perp$ de H , onde $m \in \mathbb{Z}$, é chamado de subespaço de aproximação.

2.2. TRANSFORMADA WAVELET

Por *transformada wavelet*, entende-se como sendo um produto interno [13], entre um vetor $x \in H$ e algum vetor *wavelet* $\omega_{m,n} \in W_m(\omega) \subset H$ ou algum vetor escala $\phi_{m,n}(\cdot) \in V_m(\phi)$, $(m,n) \in \mathbb{Z} \times \mathbb{Z}$. De acordo com [4], as transformadas *wavelet* podem ser agrupadas em duas categorias: *coeficientes de detalhes*, denotado por $\{d_{m,n}\}_{(m,n) \in \mathbb{Z} \times \mathbb{Z}}$, e *coeficientes de aproximação*, denotado por $\{a_{m,n}\}_{(m,n) \in \mathbb{Z} \times \mathbb{Z}}$. O coeficiente de detalhes $d_{m,n}$ e o coeficiente de aproximação $a_{m,n}$ são definidos, respectivamente, como $d_{m,n} := \langle x; \omega_{m,n} \rangle$ e $a_{m,n} := \langle x; \phi_{m,n} \rangle$, para todo $(m,n) \in \mathbb{Z} \times \mathbb{Z}$.

2.3. EXPANSÃO WAVELET

De acordo com [15], uma cadeia de subespaços fechados $\{V_m(\phi)\}_{m \in \mathbb{Z}}$ é uma *Análise Multirresolução Wavelet (AMR wavelet)*, com vetor escala ϕ se os axiomas (a), (b), (c), (d) e (e) são satisfeitos.

$$V_m(\phi) \subset V_{m+1}(\phi), \forall m \in \mathbb{Z};$$

$$\bigcap_{m \in \mathbb{Z}} V_m(\phi) = \{0\};$$

$$\left(\bigcup_{m \in \mathbb{Z}} V_m(\phi) \right)^{\perp} = H;$$

$$v \in V_m(\phi) \Leftrightarrow Dv \in V_{m+1}(\phi), m \in \mathbb{Z};$$

$$\{D^m T^n \phi\}_{n \in \mathbb{Z}} \text{ é uma base ortonormal de subespaço de aproximação } V_m(\phi), m \in \mathbb{Z}$$

Em [14] mostra-se que o espaço de *Hilbert* H pode ser ortogonalmente expandido tal como $H = \left(\sum_{m \in \mathbb{Z}} W_m(\omega) \right)^{\perp}$ e, em [15], por meio dos axiomas de uma *AMR wavelet* $\{V_m(\phi)\}_{m \in \mathbb{Z}}$, mostra-se que $V_{m_0}(\phi) = \left(\sum_{m=-\infty}^{m_0-1} W_m(\omega) \right)^{\perp}$, onde $m_0 \in \mathbb{Z}$. Com efeito, em [8] mostra-se que H pode ser expandido conforme em (1).

$$H = V_{m_0}(\phi) + \left(\sum_{m=m_0+1}^{+\infty} W_m(\omega) \right)^{\perp} \quad (1)$$

De acordo com [14], a expansão em (2) é chamada de expansão wavelet. Dada a identidade em (1) e o Teorema da Série de Fourier [13], segue que algum vetor $0 \neq x \in H$ possui uma expansão única através da série e Fourier [13], em termos da base ortonormal $\mathcal{C} \cup \{\omega_{m,n}\}_{(m,n) \in [m]_{m_0}^{+\infty} \times \mathbb{Z}}$ de H , definida por:

$$x = \sum_{n \in \mathbb{Z}} a_{m_0,n} \phi_{m_0,n}(\cdot) + \sum_{m=m_0}^{+\infty} \sum_{n \in \mathbb{Z}} d_{m,n} \omega_{m,n}(\cdot) \quad (2)$$

Onde $a_{m,n} := \langle x; \phi_{m,n} \rangle$ e $d_{m,n} := \langle x; \omega_{m,n} \rangle$ e as famílias $\{a_{m_0,n}\}_{n \in \mathbb{Z}}$ e $\{d_{m,n}\}_{(m,n) \in [m]_{m_0}^{+\infty} \times \mathbb{Z}}$ são, respectivamente, os coeficientes de aproximação e os de detalhes de $0 \neq x \in H$. De acordo com [15], $x_{V_{m_0}}(\phi) := \sum_{n \in \mathbb{Z}} a_{m_0,n} \phi_{m_0,n}(\cdot)$ e

$x_{W_m}(w) = \sum_{m \in \mathbb{Z}} d_{m,m} \omega_{m,m}$ são chamados respectivamente de componentes de aproximação (na escala m_0) e de componentes de detalhes na escala m de $0 \neq x \in \mathbb{R}^T$.

2. SINGULAR SPECTRUM ANALYSIS

Singular Spectrum Analysis (SSA) é um método de processamento de sinais que pode ser utilizado, dentre outras aplicações, na *remoção de ruído de séries de [7,9]*. A versão básica do método SSA pode ser dividida em duas etapas: *decomposição* e *reconstrução*.

3.1. DECOMPOSIÇÃO

A etapa da decomposição pode ser subdividida em duas subetapas: incorporação e decomposição em valores singulares (SVD - Singular Value Decomposition).

Seja $Y_T = [y_1, \dots, y_T] \in \mathbb{R}^T$ uma *série temporal* com cardinalidade igual a T e $F: \mathbb{R}^T \rightarrow \mathbb{R}^{L \times K}$ um *mapa invertível* [13]. Por *incorporação*, entende-se como sendo um procedimento no qual uma série temporal $Y_T \in \mathbb{R}^T$ é transformada pelo mapa F em uma matriz $X = [X_1, \dots, X_T]_{L \times K} \in \mathbb{R}^{L \times K}$, onde $X_k = [y_k, \dots, y_{k+L-1}]^T \in \mathbb{R}^L$, para todo k . Isto é, $Y_T \in \mathbb{R}^T \xrightarrow{F} X \in \mathbb{R}^{L \times K}$, onde $K = T - L + 1$. A matriz X é conhecida como *matriz trajetória* [9] e o parâmetro L , que assume algum valor inteiro no intervalo $2 \leq L \leq T$ é o *tamanho da janela* da matriz trajetória [7].

Considere o operador normal e compacto $S := XX^T$ [13]. Seja $\sigma(S)$ o espectro de S e $\{U_l U_l^T\}_{l=1}^L$ uma resolução de identidade sobre o espaço de Hilbert $(\mathbb{R}^L, (\cdot, \cdot))$ associada ao operador S , onde U_l é o autovetor associado ao autovalor $\lambda_l \in \sigma(S)$. Pode-se mostrar que S é um operador semi-definido positivo, de modo que $\lambda_l \geq 0$, para todo l . Seja $V_l := X^T U_l / \sqrt{\lambda_l}$ e considere a ordenação parcial: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$.

De acordo com [13], se $S := XX^T$ e S é um operador normal, compacto e semi-definido positivo, então a matriz trajetória X pode ser expandida através da *decomposição em valores singulares* (SVD). Como esta *condição suficiente* é verdadeira, segue que a matriz trajetória X pode ser expandida via *decomposição em valores singulares*, em (3).

$$X = \sum_{\lambda_l \in \sigma(S)} (\lambda_l)^{\frac{1}{2}} U_l V_l^T = \sum_{l=1}^L X_l \quad (3)$$

Onde: $X_l := \lambda_l^{1/2} U_l V_l^T$ e os conjuntos $\{\lambda_l^{1/2}\}_{l=1}^L$ e $\{U_l\}_{l=1}^L$ são, respectivamente, denominados por espectro singular e de vetores singulares da matriz trajetória X . A coleção (λ_l, U_l, V_l) é conhecida como autotripla na SVD da matriz trajetória X . A contribuição de cada componente em (1) pode ser mensurada pela razão de valores singulares, dada por $(\lambda_l)^{1/2} / \sum_{i=1}^L (\lambda_i)^{1/2}$. Considere que d seja o posto (isto é, o número de autovalores não nulos) da matriz trajetória X [13]. Segue que a identidade descrita em (3) pode ser reescrita tal como $X = \sum_{l=1}^d X_l$, onde $d \leq L$.

3.2. RECONSTRUÇÃO

A *etapa de reconstrução* pode ser subdividida em duas subetapas: *agrupamento* e *média diagonal*. A etapa de *agrupamento* consiste no procedimento de agrupar algumas sequências de matrizes elementares resultantes da decomposição SVD em grupos *disjuntos* e, após isso, somá-las, gerando novas matrizes elementares.

Tome a sequência $\{X_l\}_{l=1}^d$ de matrizes elementares na SVD, em [4]. Agrupe-as em $m \leq d$ grupos *disjuntos* e assuma que, após o agrupamento, o conjunto de índices gerado é dado por $\{I_1, \dots, I_m\}$, onde, para todo i , $I_i = \{I_{i1}, \dots, I_{ip_i}\}$ e p_i é a cardinalidade do grupo I_i .

Note que $\{X_i\}_{i=1}^d = \mathbf{U}_{i=1}^m \left\{ X_{i,j} \right\}_{j=1}^{p_i}$, onde $m \leq d$. A matriz elementar X_{i_j} gerada a partir do grupo $\left\{ X_{i,j} \right\}_{j=1}^{p_i}$ é dada por $X_{i_j} = \sum_{j=1}^{p_i} X_{i,j}$, de modo que a identidade em (3) pode ser reescrita como em (4).

$$X = \sum_{i=1}^m X_{i_j} \quad (4)$$

É importante salientar que o procedimento de agrupamento pode ser realizado também sob a sequência $\mathbf{U}_{i=1}^m \left\{ X_{i,j} \right\}_{j=1}^{p_i}$ de igual forma ao realizado na sequência $\{X_i\}_{i=1}^d$ utilizando, inclusive, um método ou critério de agrupamento diferente do aplicado sob a sequência $\{X_i\}_{i=1}^d$ em (4). O objetivo do agrupamento é diminuir o número de componentes (ou matrizes elementares) na SVD da matriz trajetória X . A contribuição da componente X_{i_j} pode ser mensurada pela razão de valores singulares dada por $\sum_{j=1}^{p_i} (\lambda_{i,j})^{1/2} / \sum_{i=1}^d (\lambda_i)^{1/2}$.

Tome o conjunto $\mathbf{U}_{i=1}^m \left\{ X_{i,j} \right\}_{j=1}^{p_i}$ de matrizes elementares na SVD. Dado que F é um mapa invertível, segue que F^{-1} sobre a matriz trajetória X é tal que:

$$\begin{aligned} F^{-1}(X) &= F^{-1}\left(\sum_{i=1}^m X_{i_j}\right) = F^{-1}\left(\sum_{i=1}^m \left(\sum_{j=1}^{p_i} X_{i,j}\right)\right) = \sum_{i=1}^m F^{-1}\left(\sum_{j=1}^{p_i} X_{i,j}\right) \Rightarrow \\ &\Rightarrow \sum_{i=1}^m \left(\sum_{j=1}^{p_i} F^{-1}\left(X_{i,j}\right)\right) = \sum_{i=1}^m \left(\sum_{j=1}^{p_i} \left[y_t^{(i,j)}\right]_{1 \times T}\right) = \sum_{i=1}^m \left[y_t^{(i)}\right]_{1 \times T} \quad (5) \\ &= \sum_{i=1}^m \left[y_t^{(i)}\right]_{1 \times T} = \left[y_t\right]_{1 \times T}. \end{aligned}$$

A componente $\left[y_t^{(i)}\right]_{1 \times T}$ é conhecida como *componente SSA* da série temporal $\left[y_t\right]_{1 \times T}$ gerada, por meio do mapa F^{-1} , a partir da matriz elementar X_{i_j} .

Tome a matriz trajetória X e assuma que $L^* = \min(L, K)$ e que $K^* = \max(L, K)$. Considere que $x_{l,k}^{(i)}$ seja um elemento na linha l e coluna k na matriz X_{i_j} . O elemento $y_t^{(i)}$ da componente SSA $\left[y_t^{(i)}\right]_{1 \times T}$ é calculado por meio da *média diagonal*, que é definida em (4), a partir da matriz elementar X_{i_j} .

$$y_t^{(i)} = \begin{cases} \frac{\sum_{l=1}^{L^*} x_{l,t-l+1}^{(i)}}{L^*}, & \text{se } 1 \leq t < L^* \\ \frac{\sum_{l=t-K^*+1}^{L^*} x_{l,t-l+1}^{(i)}}{L^*}, & \text{se } L^* \leq t < K^* \\ \frac{\sum_{l=t-K^*+1}^{T-K^*+1} x_{l,t-l+1}^{(i)}}{T-K^*+1}, & \text{se } K^* \leq t \leq T \end{cases} \quad (6)$$

Cada componente $\left[y_t^{(i)}\right]_{1 \times T}$ concentra parte da energia da série temporal original $\left[y_t\right]_{1 \times T}$ que pode ser mensurada pela razão de valores singulares $\sum_{j=1}^{p_i} (\lambda_{i,j})^{1/2} / \sum_{i=1}^d (\lambda_i)^{1/2}$.

De acordo com [9], as componentes SSA $\left[y_t^{(i)}\right]_{1 \times T}$ podem ser classificadas em três categorias: *tendência*, *componentes harmônicas* (ciclo e sazonalidade) e *ruído*.

3. MODELAGEM BOX & JENKINS

De acordo com [17], um *processo estocástico estacionário de segunda ordem* é definido como uma família $\{Y_t\}_{t=1}^T$ de variáveis aleatórias tal que a média, a variância e a covariância são invariantes no tempo. Isto é, $Y_t \sim Dist(\mu, \sigma^2)$, para todo t . Seja $\{y_t\}_{t=1}^T$ uma realização do processo estocástico $\{Y_t\}_{t=1}^T$, isto é, *série temporal* estacionária de segunda ordem. Segue que o modelo de *Box & Jenkins* para $\{y_t\}_{t=1}^T$ é dado em (7).

$$y_t = \varphi y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (7)$$

O modelo, em (7), pode ser representado em termos de dois polinômios, os quais obtidos com o *shift defasagem* $\nabla = (1 - B)$, onde B é definido por $B^d y_t = y_{t-d}$, conforme em (8).

$$(1 - \varphi_1 B - \dots - \varphi_p B^p) y_t = (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t \quad (8)$$

Onde $\varphi_k \in \mathbb{R}$, $\theta_j \in \mathbb{R}$ denotam os parâmetros do modelo de Box & Jenkins, em (4), e ε_t , o ruído em t [19]. Em relação à verificação da estacionariedade de segunda ordem em uma série temporal, mostra-se, em [19], que tal verificação pode ser realizada através da análise de perfil do gráfico da função de autocorrelação simples ρ_k , a qual é definida em (9).

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2}. \quad (9)$$

Onde \bar{y} é a média da série temporal $\{y_t\}_{t=1}^T$; e k , a defasagem da autocorrelação. A determinação dos parâmetros p e q pode ser realizada por meio de uma análise do perfil dos gráficos das funções de autocorrelação (ρ_k) - ACF (*autocorrelation function*) - e de *autocorrelação parcial* (ϕ_{kk}) - PACF (*partial autocorrelation function*) [8,19]. Em particular, supondo que a série temporal $\{y_t\}_{t=1}^T$ apresente *tendência* (isto é, não estacionariedade na média), deve-se diferenciá-la d vezes [6, 16, 22], a fim de encontrar uma série temporal transformada estacionária de segunda ordem (caso isso seja possível [19]). Supondo que $\{y_t\}_{t=1}^T$ apresente *não estacionariedade de segunda ordem*, segue que o modelo de *Box & Jenkins* para $\{y_t\}_{t=1}^T$ é dado em (10).

$$(1 - \varphi_1 B - \dots - \varphi_p B^p)(1 - B)^d y_t = (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t \quad (10)$$

Os modelos de *Box & Jenkins* que podem ser utilizados para a modelagem séries temporais que apresentam *sazonalidade*. Supondo que $\{y_t\}_{t=1}^T$ apresente *sazonalidade*, segue que o modelo *Box & Jenkins* para $\{y_t\}_{t=1}^T$ é dado em (11).

$$\varphi(B)(1 - \phi_1 B - \dots - \phi_p B^{pS}) \nabla^d (1 - B^S)^D y_t = \theta(B)(1 - \theta_1 B - \dots - \theta_q B^{qS}) \varepsilon_t \quad (11)$$

Onde: $\varphi(B) = (1 - \varphi_1 B - \dots - \varphi_p B^p)$, $\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$, d é a ordem das diferenças simples; D é a ordem das diferenças sazonais; S é o período sazonal; $\varphi_j \in \mathbb{R}$, e $\theta_j \in \mathbb{R}$ são, respectivamente, os coeficientes dos polinômios não sazonais; e $\phi_m \in \mathbb{R}$ e $\theta_n \in \mathbb{R}$ são, respectivamente, os coeficientes dos polinômios sazonais [3,8,19].

4. COMBINAÇÃO GEOMÉTRICA DE PREVISÕES PONTUAIS

O termo “combinação de modelos” concerne tanto à combinação de previsões quanto à combinação de densidades preditivas. Desse modo, seja $\{y_t\}_{t=1}^T$ uma série temporal. Suponha que exista uma coleção $\{M_j\}_{j=1}^m$, $m \geq 2$ de métodos preditivos plausíveis para a geração de previsões para $\{y_t\}_{t=1}^T$, dentro e fora da amostra. No entanto, existe o interesse do tomador de decisão em se determinar um método único para geração de previsões. De acordo com [6], existem duas possibilidades de se determiná-lo: 1. Escolher um método no conjunto $\{M_j\}_{j=1}^m$ baseado em critérios de seleção; e 2. Tomar k métodos em $\{M_j\}_{j=1}^m$, $k \leq m$, combiná-los.

Considere $\{y_t\}_{t=1}^T$ e k métodos de previsão em $\{M_j\}_{j=1}^m$, $k \leq m$. Assuma que ∇^k seja um conjunto com todas as previsões pontuais para $\{y_t\}_{t=1}^T$, na amostra e fora da amostra (com horizonte de previsão igual a h passos à frente), dos k métodos em $\{M_j\}_{j=1}^m$. Considere ainda que ∇_c seja o conjunto com todas as previsões pontuais para $\{y_t\}_{t=1}^T$, na amostra e fora da amostra (com horizonte de previsão igual a h passos à frente), de todas as combinações de previsões pontuais dos K métodos em $\{M_j\}_{j=1}^m$. Por *combinação de previsões pontuais* (genérica), entende-se como sendo uma função $\hat{y}_c: \nabla^k \rightarrow \nabla_c$ que faz o mapeamento de um vetor de previsões pontuais $\{y_{t,k}\}_{k=1}^k \in \nabla^k$ em uma previsão pontual combinada $\hat{y}_{t,c} \in \nabla_c$, onde t é um inteiro no intervalo $1 \leq t \leq T + h$. Em particular, uma *combinação geométrica de previsões pontuais* é uma função $\hat{y}_{CG}: \nabla^k \rightarrow \nabla_{CG}$ que faz o mapeamento de um vetor de previsões pontuais $\{y_{t,k}\}_{k=1}^k \in \nabla^k$ em uma previsão pontual geometricamente combinada $\hat{y}_{t,CG} \in \nabla_{CG}$ e é definida em (12).

$$\hat{y}_{CG,t} = \left[\prod_{j=1}^k (\hat{y}_{j,t})^{\rho_j} \right] \times \alpha + \beta \quad (12)$$

Onde $\{\rho_j\}_{j=1}^k$ é o conjunto de pesos geométricos adaptativos, e α e β são as constantes adaptativas multiplicativa e aditiva [2,6,10].

5. METODOLOGIA PROPOSTA

O método proposto é subdividido em três etapas: (1) modelagem via Box & Jenkins das componentes wavelet (ou subséries temporais wavelet); (2) modelagem via Box & Jenkins da série filtrada via abordagem SSA; (3) combinação geométrica das previsões pontuais das abordagens (1) e (2). Para realizar o ajuste da combinação geométrica das previsões pontuais, usamos o solver do Microsoft Excel para executar a otimização [10, 12]. Neste caso, os parâmetros ótimos da combinação geométrica são aqueles que minimizam a estatística MSE (Mean Square Error). Os três passos do método proposto são detalhados a seguir.

- Passo 1: Inicialmente, realizamos a decomposição wavelet de nível d [11] da série temporal $\{y_t\}_{t=1}^T$. Então obtemos uma subsérie temporal wavelet de aproximação $\{y_{V_{m_0}(\phi),t}\}_{t=1}^T$ com parâmetro igual a m_0 e uma subsérie temporal wavelet de detalhes $\{y_{W_{m_0}(\phi),t}\}_{t=1}^T$, para cada parâmetro inteiro no intervalo $m_0 \leq m \leq m_0 + (p - 1)$. Após isso, executamos sua modelagem, em que é utilizado um modelo particular. Box & Jenkins [1, 3, 8, 10] para cada subserie temporal wavelet, de tal

forma que para gerar os seguintes conjuntos de previsões pontuais $\{\hat{Y}_{W_{m_0}(\phi),t}\}_{t=1}^T$ e $\{\hat{Y}_{W_{m_0}(\omega),t}\}_{t=1}^T$, para cada parâmetro inteiro m_0 intervalo $m_0 \leq m \leq m_0 + (p - 1)$, onde h é o horizonte de previsão. Finalmente, nós somamos estas previsões pontuais, para cada instante t , como em (13).

$$\hat{Y}_{wavelet,t} = \left[\hat{Y}_{W_{m_0}(\phi),t} + \sum_{m=m_0}^{m_0+(p-1)} \hat{Y}_{W_{m_0}(\omega),t} \right] \quad (13)$$

Onde: $\hat{Y}_{wavelet,t}$ é a previsão pontual em t . Esta é a abordagem Modelagem Box & Jenkins com teoria Wavelet.

- Passo 2: Inicialmente, a partir da série temporal $\{Y_t\}_{t=1}^T$, realizamos o procedimento SSA (como a abordagem descrita na seção 3) a fim de gerar a série temporal filtrada (ou aproximada) denotada por $\{\tilde{Y}_t\}_{t=1}^T$. Após isso, realizamos a modelagem Box & Jenkins, gerando o conjunto $\{\hat{Y}_{SSA,t}\}_{t=1}^{T+h}$ de previsões pontuais, onde h é o horizonte de previsão. Essa é a abordagem, Modelagem Box & Jenkins com SSA.

- Passo 3: Realizamos a combinação geométrica das previsões pontuais dos conjuntos $\{\hat{Y}_{wavelet,t}\}_{t=1}^{T+h}$ e $\{\hat{Y}_{SSA,t}\}_{t=1}^{T+h}$, como descrito na seção 4. No método proposto, o ajuste da combinação geométrica ocorre por meio de um processo de otimização não linear do Problema de Programação Matemática [10,12] descrito em (14).

Função Objetivo: $\min_{\alpha, \beta, \rho_1, \rho_2} \{MSE\}$

$$\text{Sujeito a } \begin{cases} \hat{Y}_{CG,t} = [(\hat{Y}_1)^{\rho_1} \times (\hat{Y}_2)^{\rho_2}] \times \alpha + \beta \\ MSE = \sum_{t=1}^T \frac{(Y_t - \hat{Y}_{CG,t})^2}{T} \end{cases} \quad (14)$$

Onde:

$$\rho_1 = \rho_{wavelet} \quad \rho_2 = \rho_{SSA}$$

$$\hat{Y}_1 = \hat{Y}_{wavelet} \quad \hat{Y}_2 = \hat{Y}_{SSA}$$

7. ESTUDO DE CASO

No estudo de caso, considerou-se a série temporal mensal de consumo residencial de eletricidade de uma concessionária de distribuição que serve parte do Estado do Rio de Janeiro. A série temporal em consideração abrange o período de janeiro de 1995 a outubro de 2006 e sua representação gráfica está representada na figura 1. A cardinalidade desta série de tempo é igual a 142 onde 130 observações são utilizadas na amostra de treinamento e 12 observações são utilizadas na amostra de teste. Portanto, o horizonte de previsão adotado é de 12 meses a frente. Neste trabalho, usamos os seguintes programas: MATLAB, E-Views e Microsoft Excel.

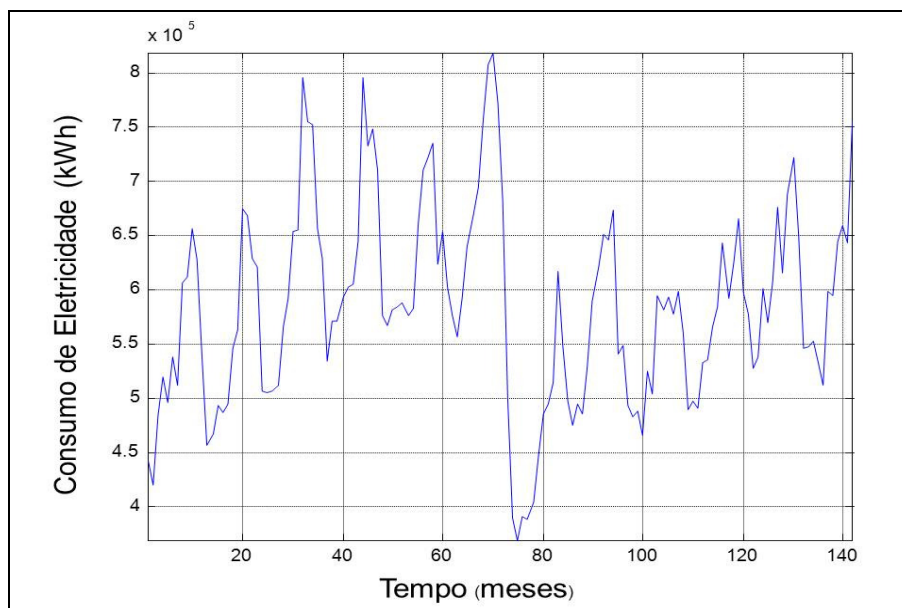


Figura 1: Série Mensal de Consumo Residencial de Eletricidade (KWh)

8. RESULTADOS E DISCUSSÕES

8.1. MODELAGEM BOX & JENKINS

O modelo Box & Jenkins (abordagem tradicional) utilizado para a modelagem da série temporal de consumo residencial mensal de eletricidade, no formato do E-Views, é dado por:

$$\text{CONSUMO} = \text{AR}(1) \text{ MA}(3) \text{ AR}(4) \text{ AR}(8) \text{ MA}(10)$$

8.2. MODELAGEM BOX & JENKINS COM TEORIA WAVELET

Na abordagem BJ-TW, foi feita a decomposição wavelet de nível 2 [11] com o uso da base ortonormal wavelet de Daubechies com momento igual a 40 (db 40) [4], de forma a gerar uma subsérie temporal e aproximação ao de nível 2, denotada por A_2 ; uma subsérie temporal de detalhes de nível 2, denotada por D_2 ; e uma subsérie temporal de detalhes de nível 1, denotada por D_1 , da série temporal de consumo residencial mensal de eletricidade. A Figura 2 apresenta os gráficos, de cima para baixo, das subséries temporais wavelet A_2 , D_2 e D_1 , respectivamente.

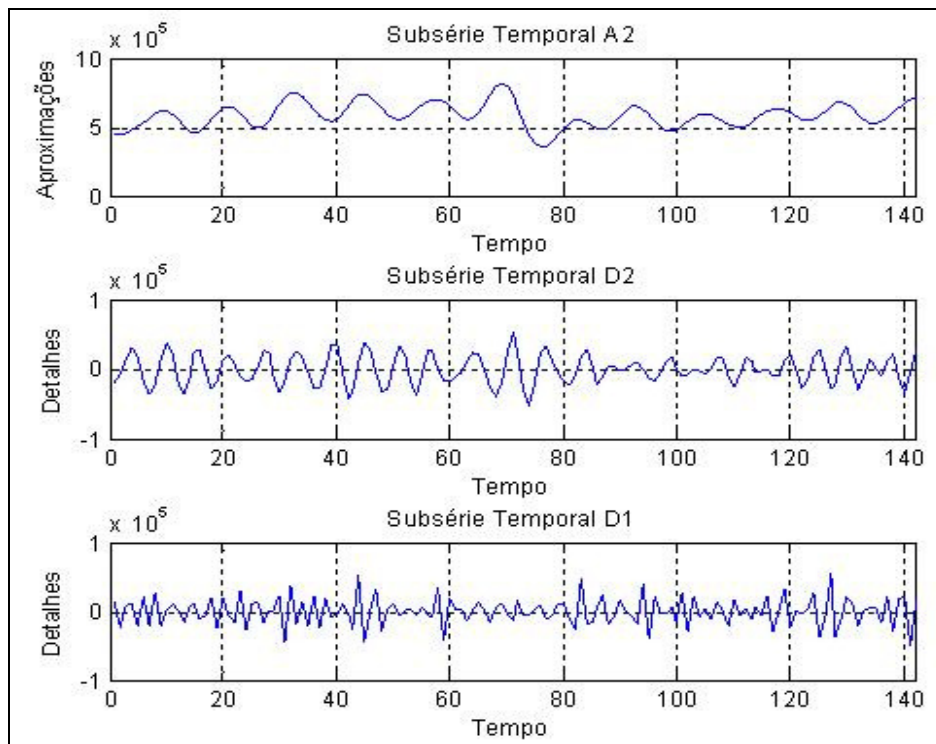


Figura 2: Subsérie Wavelet da Série Residencial de Consumo de Eletricidade.

Para a modelagem das subséries temporais waveletA2, D2 e D1, os modelos Box & Jenkins utilizados, no formato do E-Views, são dados, respectivamente, por:

$$\begin{aligned} \text{LOG (A2)} &= \text{C AR(1) AR(2) AR(4) AR(5) AR(6) AR(9) MA(1) MA(2) MA(4) MA(5)} \\ \text{d(d(D2))} &= \text{AR(1) AR(2) AR(3) AR(4) MA(2) MA(3) MA(4)} \\ \text{d(d(d(D1)))} &= \text{AR(1) AR(2) AR(4) AR(6) MA(1) MA(2) MA(3)} \end{aligned}$$

Onde o símbolo LOG significa transformação logarítmica e o d, diferença.

8.3. MODELAGEM BOX & JENKINS COM SSA

Na abordagem BJ-SSA, foi extraída uma série temporal de resíduos via método SSA, cujos parâmetros ótimos são $L = 71$ e $m = 50$, da série temporal de consumo residencial mensal de eletricidade. Aplicou-se a mesma o teste BDS, cujos resultados empírico estão na Tabela 1.

Tabela 1: Teste BDS para a série temporal de resíduos extraída via SSA.

Dimensão	Estatística BDS	Escore Z	p-valor
2	-0.00302	-0.57076	0.5682
3	0.00191	0.22731	0.8202
4	0.00174	0.17402	0.8618
5	0.00441	0.42200	0.6730
6	0.00871	0.86386	0.3877

Note, na Tabela 1, que a hipótese nula (independência) não é rejeitada a um nível de 5% de significância. O melhor modelo Box & Jenkins com SSA para a modelagem da série temporal de consumo residencial mensal de eletricidade extraída, no formato do E-views, é dado por:

$$\text{CONSUMO SSA} = \text{AR(1) AR(2) MA(1) MA(2) AR(10) MA(10) MA(11)}$$

8.4. COMPARAÇÃO DOS RESULTADOS

Na Tabela 2, têm-se as estatísticas de aderência MAPE e MAE dos quatro métodos de previsão implementados.

Tabela 2. Estatísticas de Aderência.

Método	In Sample		Out of Sample	
	MAPE	MAE	MAPE	MAE
BJ	7.27%	36129.3	11.92%	63976.2
BJ-TW	1.88%	11370.8	6.45%	41281.4
BJ-SSA	5.60%	32069.5	10,36%	59196.5
Combinação Geométrica de BJ-TW e BJ-SSA	1.84%	11124.3	5.67%	36642.6

9. CONCLUSÕES

Neste trabalho foram comparados os métodos de previsão: Box & Jenkins (BJ); Box & Jenkins com Teoria Wavelet (BJ-TW); Box & Jenkins com Análise Espectral Singular (BJ-SSA); e Combinação Geométrica dos métodos BJ-TW e BJ-SSA na previsão de uma série de consumo residencial mensal de eletricidade em uma parte do Estado do Rio de Janeiro. Pelos resultados exibidos na Tabela 2, verifica-se que o uso da teoria wavelet e SSA resultou em ganhos preditivos quando comparados com as previsões fornecidas pelo modelo Box & Jenkins tradicional. Contudo, a combinação geométrica dos métodos BJ-TW e BJ-SSA incorreu em ganhos preditivos ainda maiores, tanto dentro da amostra quanto fora da amostra, gerando as melhores previsões pontuais no experimento. Tais resultados permitem afirmar que a combinação geométrica de BJ-TW e BJSSA pode melhorar significativamente o desempenho preditivo sobre os outros três métodos, com a vantagem de ser um método implementado de forma operacionalmente simples no processo de previsões pontuais para séries temporais. Uma vez que o método pode ser aplicado a quaisquer tipos de séries temporais, tal inovação fornece uma boa contribuição para a área de previsão, pois superou a qualidade de previsão do método mais tradicionalmente usado (BJ) e superou até mesmo a qualidade dos outros métodos mais modernos da literatura (BJ-TW e BJ-SSA).

10. REFERÊNCIAS

- [1] Armstrong, S. J. (2004). Principles of Forecasting: A handbook for Researchers and Practitioners. Massachusetts: Eletronic Services available at <http://www.wkap.nl>
- [2] Bates, J. M., and Granger, C. W. J.. (1969). Combination Forecasts. *Operation Research Quarterly*.
- [3] Box, G. E. P., and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day.
- [4] Daubechies, I. (1988). Orthonormal Bases of Compactly Supported Wavelet. *Comm. Pureand Applied Math.*, 41, pp. 909-996.
- [5] Elsner, J. B., and Tsonis, A. A. (1996). *Singular Spectrum Analysis. A New Tool in Time Series Analysis*. Plenum Press, New York.
- [6] Faria, A. E., and Mubwandarika, E. (2008). Multimodality on the Geometric Combination of Bayesian Forecasting Models. *International Journal of Statistics and Management System*. 3, 1-25.
- [7] Golyandina, N., Netrutkin, V., and Zhigljavsky, A. (2001). *Analysis of Time Series Structure: SSA and Related Techniques*. Chapman & Hall / CRC. New York - London.
- [8] Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press.
- [9] Hassani, H. (2007). Singular Spectrum Analysis: Methodology and Comparison. *Journal of Data Science*. 5, 239-257.
- [10] Junior, L. A. T (2009). *Combinação Geométrica de Métodos Preditivos; Aplicação a Série de Consumo Residencial de Energia Elétrica*. Dissertação de Mestrado. PUC-Rio.
- [11] Junior, L. A. T, Pessanha, J. F. M., and Souza, R. C. (2011). Análise Wavelet e Redes Neurais Artificiais na Previsão de Velocidade de Vento. XVII Simpósio Brasileiro de Pesquisa Operacional, Ubatuba, Brazil.
Available at <http://www.xliiisbpo.iltc.br/pdf/87357.pdf>
- [12] Junior, L. A. T, Pessanha, J. F. M., and Souza, R. C. (2011). Combinação Linear de Previsões com Ajuste MINIMAX. XIV Simpósio de Pesquisa Operacional e Logística da Marinha, Rio de Janeiro, Brasil.
- [13] Kubrusly, C. S. (2001). *Elements of Operator Theory*. Birkhauser, Boston.
- [14] Kubrusly, C. S., and Levan, N. (2006). Abstract Wavelets Generated by Hilbert Space Shift Operators. *Adv. Math. Sci. Appl.* 14, 643-660.
- [15] Levan, N., and Kubrusly, C. S. (2003). A Wavelet Time-Shift-Detail Decomposition. *Mathematics and Computers in Simulation*. 63, 73-78.
- [16] Mallat, S. (1998). *A Wavelet Tour of Signal Processing*. Academic Press, San Diego.
- [17] Morettin, P.A., Toloí, C.M.C (2001) *Análise de Séries Temporais*. Edgard Blucher, São Paulo.
- [18] Newbold, P., and Granger, C. W. J. (1974). Experience with Forecasting Univariate Times Series and the Combination of Forecasts. *Journal of the Royal Statistical Society*. V.137, pp.131-165.
- [19] Souza, R.C., Camargo, M.E. (2004). *Análise de Séries Temporais: os modelos ARIMA*. 2a. edição. Gráfica e Editora Regional, Rio de Janeiro.
- [20] Wang, X., and McDonald, J. R. (1994). *Modern Power System Planning*. McGraw Hill.