

SINGULAR SPECTRUM ANALYSIS ASSOCIADA À MODELAGEM ARIMA E HOLT-WINTERS NA PREVISÃO DE CONSUMO DE ENERGIA

Moisés Lima de Menezes

Dep. Estatística - Universidade Federal Fluminense
Rua Mário Santos Braga, S/N, Campus Valonguinho, Centro, Niterói - RJ, 24.220-900.
moises_lima@msn.com

Keila Mara Cassiano

Dep. Estatística - Universidade Federal Fluminense
Rua Mário Santos Braga, S/N, Campus Valonguinho, Centro, Niterói - RJ, 24.220-900.
keilamath@hotmail.com

Rafael Morais de Souza

Dep. Ciências Contábeis - Universidade Federal de Minas Gerais
Av. Antônio Carlos, 6.627 - Campus Pampulha, Belo Horizonte – MG, 31.270-901.
rafael.morais.souza@gmail.com

Reinaldo Castro Souza

Dep. Eng. Elétrica - 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

Luiz Albino Teixeira Júnior

Inst. Latino-Americano de Tecnologia - Universidade Federal da Integração Latino-Americana
Av. Tancredo Neves, 6.731 - Bloco 4, Foz do Iguaçu – PR, 85.867-970.
luiz.a.t.junior@gmail.com

Resumo

Singular Spectrum Analysis (SSA) é uma técnica não-paramétrica que permite decompor uma série temporal em sinal e ruído. Neste artigo, os modelos Box & Jenkins e Holt-Winters são testados com e sem a abordagem SSA para a modelagem de uma série temporal de consumo residencial mensal de energia elétrica de uma concessionária do Rio de Janeiro. Três diferentes metodologias são utilizadas na abordagem SSA: Análise de Componentes principais (ACP), ACP associado com Análise de Cluster e Análise Gráfica dos Vetores Singulares. MAPE, MAE, RMSE e R^2 são estatísticas usadas para testar o poder preditivo dos modelos. Os resultados mostram um maior poder preditivo do modelo quando aplicado a séries filtradas em conjunto com a técnica SSA.

Palavras Chave: SSA, ARIMA, Holt-Winters, consumo de energia.

Abstract

Singular Spectrum Analysis (SSA) is a non-parametric technique to decompose a time series into signal and noise. In this article, the Box-Jenkins and Holt-Winters models are tested with and without SSA approach for modeling a time series of monthly residential electricity consumption from a dealership in Rio de Janeiro. Three different methodologies are used in the SSA approach: Analysis of Main Components (ACP), ACP associated with Cluster Analysis and Graphical Analysis of Singular Vectors. MAPE, MAE, RMSE and R^2 statistics are used to test the predictive power of the models. The results show a greater predictive power of the model when applied in conjunction with the filtered technique SSA series.

Keywords: SSA, ARIMA, Holt-Winters, electricity consumption.

1. INTRODUÇÃO

O crescimento da população, o progresso tecnológico e as economias emergentes têm promovido o aumento da demanda por energia elétrica em todo o mundo. Para lidar com essa tendência crescente, as previsões de carga são periodicamente revisadas, a fim de atualizar o planejamento da expansão de médio e longo prazo. Tradicionalmente, no planejamento da expansão de médio prazo, os modelos de Holt-Winters, Box & Jenkins e de regressão linear múltipla são considerados para a previsão de demanda de energia elétrica. Uma forma de melhorar a qualidade de ajuste dos modelos de previsão consiste na utilização de métodos de filtragem de dados no pré-processamento da série temporal. Entre os métodos disponíveis, a Análise Espectral Singular (*Singular Spectrum Analysis* – SSA) é um recente e poderoso método em séries temporais que incorpora elementos de análise clássica de séries temporais, estatística multivariada, geometria multivariada, sistemas dinâmicos e processamentos de sinais (ELSNER & TSONIS, 1996). SSA tem sido aplicada com sucesso em diversas áreas: da matemática e física a economia e matemática financeira, da meteorologia e oceanografia a ciências sociais (GOLYANDINA *et al.*, 2001).

SSA decompõe uma série temporal em uma soma de um pequeno número de componentes independentes e interpretáveis como: uma tendência de variação lenta, componentes oscilatórias e uma estrutura de ruído. SSA é uma ferramenta que pode ser usada para resolver os seguintes problemas: 1) encontrar tendências de diferentes resoluções; 2) suavizar séries temporais; 3) extrair componentes sazonais; 4) extrair ciclos com pequenos e grandes períodos; 5) extrair periodicidades com amplitudes variáveis; 6) extrair tendências complexas e periodicidades e 7) encontrar estrutura em séries temporais curtas (HASSANI *et al.*, 2012). Uma das vantagens do método proposto em SSA é sua abordagem não paramétrica; ou seja, não é necessário conhecer ou especificar o modelo paramétrico para a série temporal considerada.

Uma descrição minuciosa dos fundamentos teóricos e práticos da técnica SSA (com vários exemplos) pode ser encontrada em (BROOMHEAD & KING, 1986), (DANILOV & ZHGLJAVSKY, 1997) e (GOLYANDINA *et al.*, 2001). Uma introdução elementar ao SSA pode ser encontrada em (ELSNER & TSONIS, 1996).

Um exemplo do benefício promovido pela filtragem SSA pode ser encontrado em (HASSANI *et al.*, 2009), onde são apresentadas previsões para a produção industrial na Europa obtidas por meio de modelos ARIMA e Holt-Winters com e sem a filtragem SSA. Ainda em HASSANI *et al.*, (2009), destaca-se a análise gráfica dos vetores singulares. Neste artigo é investigado o uso da SSA na previsão do consumo mensal de energia elétrica, uma informação importante na operação e no planejamento de sistemas elétricos.

Este trabalho está dividido em cinco seções. Após esta breve introdução, na seção 2 tem-se uma descrição da metodologia SSA. Os métodos preditivos a serem utilizados são apresentados na seção 3. A série utilizada no experimento computacional é apresentada na seção 4, bem como os resultados e as discussões da aplicação da SSA. Por fim, na seção 5 são apresentadas as conclusões.

2. 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 tempo (GOLYANDINA *et al.*, 2001), (HASSANI *et al.*, 2012). A versão básica do método SSA pode ser dividida em duas etapas: *decomposição e reconstrução*.

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

Seja $Y_T = [y_1, \dots, y_T]_{1 \times T} \in R^T$, onde R^T é o conjunto dos vetores de dimensão T com elementos pertencentes ao conjunto dos números reais, uma série temporal (HAMILTON,

1994) e considere L tal que $2 \leq L \leq T$ de modo que L é um parâmetro a ser estimado e é chamado de *Comprimento de Janela* (GOLYANDINA *et al.*, 2001). Seja $F: R^T \rightarrow R^{L \times K}$ um mapa invertível (KUBRUSLY, 2001). Entende-se por *Incorporação* o procedimento no qual uma série temporal $Y_T \in R^T$ é levada pelo mapa F a uma matriz $\mathbf{X} = [X_1, \dots, X_K]_{L \times K} \in R^{L \times K}$, onde $X_k = [y_k, y_{k+1}, \dots, y_{k+L-1}]' \in R^L$, para todo $k \in [1, \dots, K]$, onde $K = T - L + 1$. A matriz \mathbf{X} , conhecida como matriz trajetória (HASSANI *et al.*, 2012), é uma matriz de *Hankel*, ou seja, os elementos $x_{i,j}$ tal que $i + j = \text{constante}$, são iguais.

Considere $\mathbf{S} = \mathbf{X}\mathbf{X}'$. Os autovalores de S dispostos em ordem de significância $\lambda_1 \geq \dots \geq \lambda_L \geq 0$ são obtidos e os respectivos autovetores U_1, \dots, U_L encontrados. Considere $V_l = \mathbf{X}'U_l/\sqrt{\lambda_l}$, como \mathbf{S} é um operador normal compacto semi-definido positivo, então a matriz Trajetória \mathbf{X} pode ser expandida através da SVD (1):

$$\mathbf{X} = \sum_{l=1}^L \mathbf{E}_l, \quad (1)$$

onde $\mathbf{E}_l = (\lambda_l)^{\frac{1}{2}}U_lV_l'$. Os conjuntos $\{\sqrt{\lambda_l}\}_{l=1}^L$ e $\{U_l\}_{l=1}^L$ são respectivamente o *espectro singular* e os *vetores singulares* de \mathbf{X} . A coleção $(\sqrt{\lambda_l}, U_l, V_l)$ é conhecida como *autotripla* da expansão SVD de \mathbf{X} . Os elementos da autotripla $(\sqrt{\lambda_l}, U_l, V_l)$ são definidos respectivamente por: valor singular, vetor singular à esquerda e vetor singular à direita de \mathbf{X} .

A contribuição de cada componente em (1) pode ser mensurada pela razão de autovalores $\lambda_l / \sum_{l=1}^L \lambda_l$. Considere d o posto (isto é, o número de autovalores não nulos) da matriz trajetória \mathbf{X} . Deste modo, segue que a identidade (1) pode ser reescrita como $\mathbf{X} = \sum_{l=1}^d \mathbf{E}_l$, onde $d \leq L$.

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

Considere a sequência $\{\mathbf{E}_l\}_{l=1}^d$ de matrizes elementares da expansão SVD. Agrupe as mesmas em m grupos disjuntos utilizando algum método [por exemplo, por meio de análise de componentes principais, análise gráfica de vetores singulares ou agrupamento hierárquico (*Clusterização Hierárquica*)] e assuma que o conjunto de índices gerado é dado por $\{I_1, \dots, I_m\}$, onde para todo i , $\{I_{i_1}, \dots, I_{i_{p_i}}\}$ e p_i é a cardinalidade do grupo I_i . Note que $\{\mathbf{E}_l\}_{l=1}^d = \cup_{i=1}^m \left\{ \mathbf{X}_{I_{i_j}} \right\}_{j=1}^{p_i}$, onde $m \leq d$. Com efeito, a matriz elementar \mathbf{X}_{I_i} arbitrária gerada a partir do grupo $\left\{ \mathbf{X}_{I_{i_j}} \right\}_{j=1}^{p_i}$ é tal que $\mathbf{X}_{I_i} = \sum_{j=1}^{p_i} \mathbf{X}_{I_{i_j}}$, de modo que a identidade (1) pode ser reescrita como em (2).

$$\mathbf{X} = \sum_{l=1}^L \mathbf{E}_l = \sum_{i=1}^m \mathbf{X}_{I_i}. \quad (2)$$

O objetivo do agrupamento é diminuir o número de componentes na expansão da matriz trajetória X . A contribuição de cada componente é mensurada pela razão de autovalores dada por $\sum_{j=1}^{p_i} \lambda_{I_{i_j}} / \sum_{l=1}^d \lambda_l$.

Considere o conjunto $\cup_{i=1}^m \left\{ \mathbf{X}_{I_{i_j}} \right\}_{j=1}^{p_i}$ de matrizes elementares da expansão SVD. Dado que F é um mapa invertível, segue que F^{-1} aplicado sobre a matriz trajetória X é tal que:

$$\left\{ \begin{aligned} F^{-1}(\mathbf{X}) &= F^{-1}\left(\sum_{i=1}^m \mathbf{X}_{I_i}\right) = F^{-1}\left(\sum_{i=1}^m \sum_{j=1}^{p_i} \mathbf{X}_{I_{ij}}\right) \\ &= \sum_{i=1}^m \sum_{j=1}^{p_i} F^{-1}(\mathbf{X}_{I_{ij}}) = \sum_{i=1}^m \sum_{j=1}^{p_i} \left[y_t^{(I_{ij})} \right]_{1 \times T} \\ &= \sum_{i=1}^m \left[y_t^{(I_i)} \right]_{1 \times T} = \sum_{i=1}^m \left[y_t^{(i)} \right]_{1 \times T} = [y_t]_{1 \times T} \end{aligned} \right. \quad (3)$$

A componente $\left[y_t^{(i)} \right]_{1 \times T}$ em (3) é conhecida como *componente SSA* da série temporal $[y_t]_{1 \times T}$ gerada a partir da matriz elementar \mathbf{X}_{I_i} .

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

$$y_t^{(i)} = \begin{cases} \frac{1}{t} \sum_{l=1}^t x_{l,t-l+1}^{(i)}, & \text{se } 1 \leq t \leq L^* \\ \frac{1}{L^*} \sum_{l=1}^{L^*} x_{l,t-l+1}^{(i)}, & \text{se } L^* \leq t \leq K^* \\ \frac{1}{T - K^* + 1} \sum_{l=t-K^*+1}^{T-K^*+1} x_{l,t-l+1}^{(i)}, & \text{se } K^* \leq t \leq T \end{cases} \quad (4)$$

Cada componente $\left[y_t^{(i)} \right]_{1 \times T}$ concentra parte da energia da série temporal original $[y_t]_{1 \times T}$ que pode ser mensurada pela razão de autovalores $\sum_{j=1}^{p_i} \lambda_{I_{ij}} / \sum_{l=1}^d \lambda_l$. De acordo com HASSANI *et al.*, (2012), as componentes SSA podem ser classificadas em três categorias: *tendência*, *harmônica* (ciclo e sazonalidade) e *ruído* (GOLYANDINA *et al.*, 2001).

Um dos principais conceitos estudados em SSA é a propriedade de separabilidade (HASSANI *et al.*, 2012). Tal propriedade caracteriza quão bem separados estão as diferentes componentes umas das outras. Uma boa medida de separabilidade é a *Correlação Ponderada*. Por correlação ponderada (*weighted correlation*) ou *w*-correlação, podemos entender como uma função que quantifica a dependência linear entre duas componentes SSA $Y_T^{(1)}$ e $Y_T^{(2)}$, definida em (5):

$$\rho_{12}^{(w)} = \frac{\left(Y_T^{(1)}, Y_T^{(2)} \right)_w}{\left\| Y_T^{(1)} \right\|_w \left\| Y_T^{(2)} \right\|_w}, \quad (5)$$

onde $\left\| Y_T^{(i)} \right\|_w = \sqrt{\left(Y_T^{(i)}, Y_T^{(i)} \right)_w}$, $\left(Y_T^{(i)}, Y_T^{(j)} \right)_w = \sum_{k=1}^T w_k y_k^{(i)} y_k^{(j)}$, $(i, j) = (1, 2)$, $w_k = \min\{k, L, T - k\}$ (aqui é assumido que $L \leq T/2$).

Através da separabilidade, pode-se verificar estatisticamente se duas componentes SSA estão bem separadas, em termos de dependência linear. Se o valor absoluto da *w*-

correlação é pequeno (HASSANI *et al.*, 2012), então as componentes SSA correspondentes são classificadas como w -Ortogonal (ou quase w -Ortogonal); caso contrário, são ditas mal separadas. Salienta-se que comumente utiliza-se a correlação ponderada na fase de agrupamento SSA (GOLYANDINA *et al.*, 2001).

3. MÉTODOS PREDITIVOS

Nesta seção são apresentados os métodos preditivos a serem testados na previsão de séries filtradas e não filtradas via SSA. Os métodos testados neste artigo são os modelos de Holt-Winters e os modelos ARIMA.

3.1. MODELOS DE HOLT-WINTERS

De acordo com MORETTIN & TOLOI (2001), os métodos de amortecimento exponencial baseiam-se na premissa ponderação de que os dados são diferentemente ponderados. Comumente, as observações recentes contêm informações mais relevantes que as antigas, de forma que a ponderação dos dados (série temporal) decresce exponencialmente à medida que a observação torna-se mais antiga. Um caso particular dos métodos de amortecimento exponencial é o método de Holt-Winters multiplicativo, o qual realiza a modelagem de forma dinâmica (isto é, com parâmetros variantes no tempo) das componentes: nível ($a_1; t$), tendência ($a_2; t$) e sazonalidade ($\rho_{m(t)}$), conforme a equação (6).

$$\begin{cases} y_t = \mu_t + \epsilon_t \\ y_t = [a_{1,t} + a_{2,t} \times t] \times \rho_{m(t)} + \epsilon_t \end{cases} \quad (6)$$

onde ϵ_t é um *erro estocástico*, y_t é o valor observado em t e $\rho_{m(t)}$ é o fator sazonal em t relativo ao mês m . De acordo com HAMILTON (1994), a família $\{\rho_{m(\cdot)}\}_{m \in \Gamma}$ de fatores sazonais, onde Γ é o conjunto com todos os meses no ano, deve obedecer à restrição $\sum_{m \in \Gamma} \rho_{m(\cdot)} = L$, sendo L o tamanho do ciclo sazonal. No processo de estimação dos parâmetros da equação (6), são utilizados três hiperparâmetros (quantidades invariantes no tempo), denotados por α , β e λ os quais estão associados, respectivamente, às estimativas de nível, tendência e sazonalidade (quantidades variantes no tempo). De acordo com HASSANI *et al.* (2009), os valores ótimos dos hiperparâmetros α , β e λ se encontram na região convexa $[0,1]^3$ e tem como objetivo a minimização da estatística *MSE* (mean square error).

3.2. MODELOS DE BOX & JENKINS

De acordo com ELSNER & TSONIS (1996), um processo estocástico estacionário de segunda ordem é definido como uma família $\{Y_t\}_{t=1}^T$ de variáveis aleatórias cujos momentos (média, variância e covariância) são invariantes no tempo. Isto é: $Y_t \sim Dist(\mu; \sigma^2)$ para todo t . Considere que a sequência $\{y_t\}_{t=1}^T$ seja uma realização de $\{Y_t\}_{t=1}^T$ (isto é, uma série temporal estacionária de segunda ordem), onde y_t consiste na notação da observação em t . BOX & JENKINS (1970) propõem a equação (7) para a modelagem de sua dinâmica temporal $\{y_t\}_{t=1}^T$.

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}, \quad (7)$$

onde ϵ_t é um erro estocástico e y_t é o valor observado em t . O modelo (8) pode ser representado em termos de dois polinômios obtidos através do operador defasagem B (HAMILTON, 1994), definido por $B^d Y_t = Y_{t-d}$ e possui formulação geral definida em (8).

$$(1 - \phi_1 B - \dots - \phi_p B^p) y_t = (1 - \theta_1 B - \dots - \theta_q B^q) \epsilon_t, \quad (8)$$

onde $\phi_k, \forall k \geq 1$ e $\theta_j, \forall j \geq 1$ denotam os parâmetros do modelo *ARIMA* (autorregressivo e de médias móveis) ϵ_t e y_t são como definidos em (7). Em (9) as defasagens de y_t representam a

parte autoregressiva de ordem p ($AR(p)$), enquanto as defasagens dos erros representam a parte média móvel de ordem q ($MA(q)$). O exame de estacionariedade de segunda ordem de $\{y_t\}_{t=1}^T$ pode ser realizado por meio da análise da função de autocorrelação simples (ρ_k), definida em (9).

$$\rho_k = \frac{\sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2}, \quad (9)$$

onde \bar{y} é a média da série temporal e k é a defasagem (*lag*) da autocorrelação. A determinação das ordens p e q em (8) é realizada pela análise do perfil das funções de autocorrelação (ACF), ρ_k e autocorrelação parcial (PACF), ϕ_{kk} cujos padrões teóricos podem ser verificados em SOUZA & CAMARGO (2004). Uma vez identificada a ordem do modelo *ARIMA*, os coeficientes das partes $AR(p)$ e $MA(q)$ são estimados e as respectivas significâncias estatística são avaliadas (HAMILTON, 1994). Em especial, caso uma série temporal $\{y_t\}_{t=1}^T$ não estacionária na média, mas estacionária na variância e covariância, deve-se diferenciá-la d vezes, a fim de se gerar uma nova serie temporal que apresente estacionariedade na média (HAMILTON, 1994). Para tais classes de séries temporais, pode-se utilizar o modelo *ARIMA*($p; d; q$), cuja formulação geral é descrita em (10).

$$(1 - \phi_1 B^1 - \dots - \phi_p B^p)(1 - B)^d y_t = (1 - \theta_1 - \dots - \theta_q B^q) \epsilon_t, \quad (10)$$

Além dos modelos descritos em (8) e (10), ainda existem os modelos aplicados a séries com comportamento sazonais e modelos para séries que apresentam longa dependência: são os modelos *SARIMA*, e *ARFIMA* descritos em MORETTIN & TOLOI (2001).

4. EXPERIMENTO COMPUTACIONAL

No experimento computacional, considerou-se a série temporal de consumo residencial de energia elétrica, medida em kWh (com frequência mensal) de uma distribuidora de eletricidade que atende parte do estado do Rio de Janeiro apresentada na figura 1. A série considerada compreende o período de janeiro de 1995 a outubro de 2006. A amostra considerada na estimação dos modelos previsão (*in sample*) contém 136 observações e a amostra de teste (*out of sample*) é composta pelas últimas 6 observações. Foram realizadas previsões multi-step no horizonte de previsão igual a 6. Em função da diversidade de métodos utilizados, a implementação computacional contou com o apoio de diferentes *softwares*: MATLAB, para a abordagem SSA via análise de componentes principais sob a SVD a fim de se obter uma série temporal aproximada (menos ruidosa); *Caterpillar* SSA, para verificação detalhada da filtragem SSA via análise gráfica dos vetores singulares e seus gráficos de dispersão; *E-Views*, para as análises dos testes de BDS (independência) (BROCK *et al.*, 1996), (LIN, 1997) e modelagem *ARIMA*; Forecast Pro for Windows, para modelagem via método de Holt-Winters; R, para abordagem SSA usando clusterização hierárquica; e Microsoft Excel, para gerar os gráficos das séries obtidas e cálculo das estatísticas de aderência.

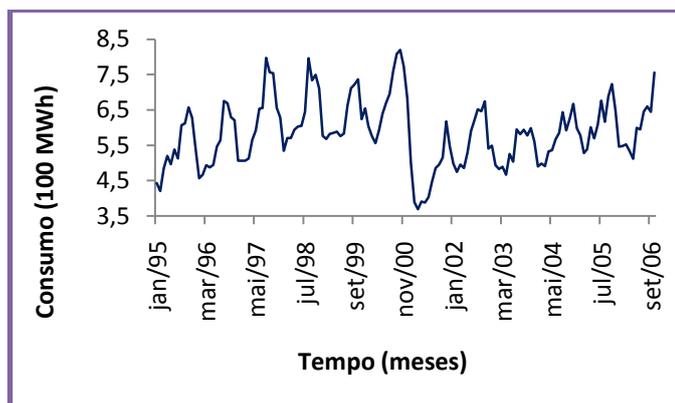


Figura 1. Série Mensal de Consumo de Energia Elétrica (kWh).

No estudo de caso, foram aplicadas três abordagens de filtragem via método SSA: análise de componentes principais (VASCONCELOS, 2001) sob SVD (ACP-SVD), análise de cluster integrada com ACP-SVD e análise gráfica de vetores singulares da SVD. A partir de cada abordagem de filtragem, foi gerada uma série temporal aproximada para a série temporal do consumo residencial mensal de energia elétrica que foi modelada por dois métodos preditivos: ARIMA e Holt-Winters Multiplicativo.

Os resultados obtidos ao longo dos experimentos computacionais realizados são comparados em termos das estatísticas de aderência: MAPE (erro médio percentual absoluto), MAE (erro médio absoluto), RMSE (raiz da média dos quadrados dos erros) e R^2 (coeficiente de determinação):

$$MAPE = \sum_{t=1}^T \left[\frac{|y_t - \hat{y}_t|}{y_t} \right] \times \frac{1}{T}, \quad MAE = \sum_{t=1}^T [|y_t - \hat{y}_t|] \times \frac{1}{T},$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^T (y_t - \hat{y}_t)^2}{T}}, \quad R^2 = 1 - \left(\frac{\sum_{t=1}^T (y_t - \hat{y}_t)^2}{\sum_{t=1}^T (y_t - \bar{y})^2} \right)$$

onde y_t é o valor observado e \hat{y}_t é a previsão, ambos para o instante t , onde $1 \leq t \leq T$.

4.1. ANÁLISE DE COMPONENTES PRINCIPAIS SOB SVD (ACP – SVD)

Na filtragem SSA via abordagem ACP-SVD, define-se um comprimento de janela ótimo igual a L para a matriz trajetória X , um truncamento ótimo na componente N na SVD, de modo que a soma das matrizes elementares remanescentes na SVD venha a gerar uma série temporal R_T classificada, via estatística BDS (BROCK, et al., 1996), como ruído. O objetivo é obter uma série temporal \tilde{Y}_T menos ruidosa que a série temporal original Y_T , removendo-se R_T . Trata-se, portanto, de um problema de otimização tal que os p -valores da estatística BDS são suas restrições. Nesta perspectiva, tem-se que os valores ótimos para L e N são, respectivamente, 71 e 50. Na Figura 2 são apresentados os logaritmos dos 71 autovalores dispostos em ordenação parcial decrescente e o ponto definido pelo valor ótimo de N .

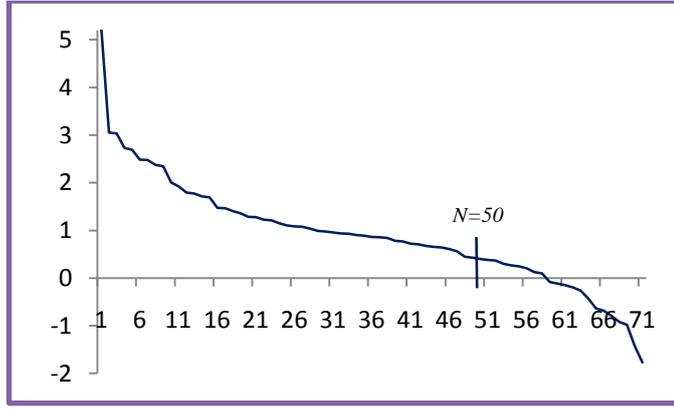


Figura 2. Logaritmo dos 71 Autovalores na SVD e o valor ótimo de N .

Assumindo as sequências $\{E_j\}_{j=1}^{50}$ e $\{E_j\}_{j=51}^{71}$ respectivamente as matrizes elementares na SVD que através de F^{-1} geram, respectivamente a série temporal aproximada S_T (menos ruidosa que a original Y_T) e uma série R_T constituída apenas de ruídos. Com efeito, tem-se que a série Y_T pode ser decomposta em termos de S_T e R_T conforme em (11).

$$Y_T = S_T + R_T \quad (11)$$

Removendo-se a série R_T , obtém-se a série filtrada \tilde{Y}_T via abordagem ACP-SVD tal que $\tilde{Y}_T = S_T$. A tabela I apresenta os resultados do teste BDS para a série temporal de ruídos R_T .

Tabela 1. Teste BDS para a Série de Ruídos R_T .

Abordagem	Dim.	Estatística BDS	Estatística Z	Probab.
ACP - SVD	2	-0,003026	-0,570768	0,5682
	3	0,001916	0,227310	0,8202
	4	0,001747	0,174020	0,8618
	5	0,004415	0,422002	0,6730
	6	0,008716	0,863864	0,3877

É possível verificar na tabela 1 que a hipótese nula do teste BDS (independência) não é rejeitada ao nível de 5% de significância. Portanto, não há evidência empírica de que a série temporal de ruído R_T possua qualquer estrutura de dependência temporal. Com base no teste BDS, verifica-se que, de fato, série temporal aproximada $\tilde{Y}_T = S_T$ é menos ruidosa que a série temporal original de consumo Y_T .

4.2. ACP – SVD INTEGRADA COM ANÁLISE DE CLUSTER

Na filtragem SSA via abordagem ACP-SVD integrada com análise de cluster, a filtragem ocorre em duas etapas. Na etapa 1, realiza-se o mesmo procedimento utilizado na abordagem ACP-SVD, visto na Seção 3.1. Na etapa 2, os vetores singulares remanescentes (isto é, aqueles que não foram removidos na etapa 1) são agrupados (de forma excludente) em k clusters (neste artigo, usou-se $k = 3$), via análise hierárquica de cluster (ALDENDERFER & BLASHFIELD, 1994), gerando 3 componentes SSA para a série temporal \tilde{Y}_T . O teste BDS é realizado em cada componente a fim de identificar a que possui propriedades estatísticas de ruído (geralmente a componente SSA 3). O objetivo final desta abordagem é obter uma série temporal aproximada \tilde{Y}_T menos ruidosa que a série temporal original Y_T , removendo-se a componente identificada como ruído.

Nesta perspectiva, na primeira etapa, é obtida uma matriz trajetória X de Y_T através do processo de incorporação SSA cujo comprimento de janela ótimo encontrado é igual a 68 (isto é, $L = 68$). Como consequência, foram gerados 68 vetores singulares na SVD da matriz

trajetória X , dos quais, 39 vetores singulares foram classificados como ruído, com base no teste BDS (a 5% de significância) apresentado na tabela 2. Tais vetores foram removidos. Após isso, os 29 vetores singulares remanescentes foram agrupados em 3 clusters, via análise de cluster, conforme a tabela 3.

Tabela 2. Teste BDS para a Componente SSA 3.

Abordagem	Dim.	Estatística BDS	Estatística Z	Probab.
	2	-0,000679	-0,151957	0,8792
ACP - SVD	3	-0,002588	-0,364554	0,7154
Integrada com	4	0,006310	0,747067	0,4550
Análise de Cluster	5	0,010440	1,187179	0,2352
	6	0,006928	1,053869	0,2919

Tabela 3. Componentes SSA obtidas através da Análise de Cluster

COMPONENTE	VETORES SINGULARES
SSA 1	1,2,3,4,5,6,7
SSA 2	8,9,10,11
SSA 3	12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29

4.3. ANÁLISE GRÁFICA DOS VETORES SINGULARES

A análise dos vetores singulares da série temporal na base definida pelos vetores singulares resultantes da SVD permite identificar as componentes de tendência, sazonalidade e ruído da série temporal Y_T . O problema geral aqui consiste em identificar e separar as componentes oscilatórias das componentes que fazem parte da tendência. De acordo com GOLYANDINA et al. (2001), a análise gráfica de tais coordenadas aos pares permite identificar por meio visual as componentes harmônicas da série (LIN, 1997), (SOUZA & CAMARGO, 2004).

De igual forma ao experimento computacional sob a abordagem ACP-SVD, considerou-se o comprimento de janela ótimo L igual a 71 e um ponto ótimo de truncamento d igual a 50, gerando 50 vetores singulares. O software utilizado para esta abordagem foi Caterpillar SSA (GISTATGROUP, 2010). Por meio da análise gráfica dos pares de vetores singulares é possível classificá-los conforme seu comportamento.

Considere um harmônico puro com frequência igual a ω , fase igual a δ , amplitude igual a ξ e o período $\rho = 1/\omega$ definido como um divisor do comprimento de janela L e K . Se o parâmetro ρ assume um valor inteiro, então ρ é classificado como período do harmônico (MORETTIN & TOLOI, 2001). As coordenadas da série temporal em duas componentes ortogonais descrevem uma sequência de fatores (GOLYANDINA et al., 2001). As funções *seno* e *coseno* com frequências, amplitudes e fases iguais resultam em um diagrama de dispersão que exhibe um padrão circular. Desta feita, o diagrama de dispersão exhibe um polígono regular com ρ vértices. Para uma frequência $\omega = \frac{m}{n} < 0,5$ com m e n inteiros e primos entre si, os pontos são vértices de um polígono regular de n vértices. Assim, a identificação das componentes que são geradas por um harmônico é reduzida à análise pictórica do padrão determinado nos diferentes pares de componentes (GOLYANDINA et al., 2001).

A Figura 3 apresenta os 16 primeiros vetores singulares. É possível verificar que as componentes 1, 4 e 5 são componentes de tendência. É possível identificar ainda que as componentes 2, 3, 14 e 15 constituem componentes harmônicas. Para as demais componentes, há necessidade de uma análise mais profunda, como a mostrada na figura 4.

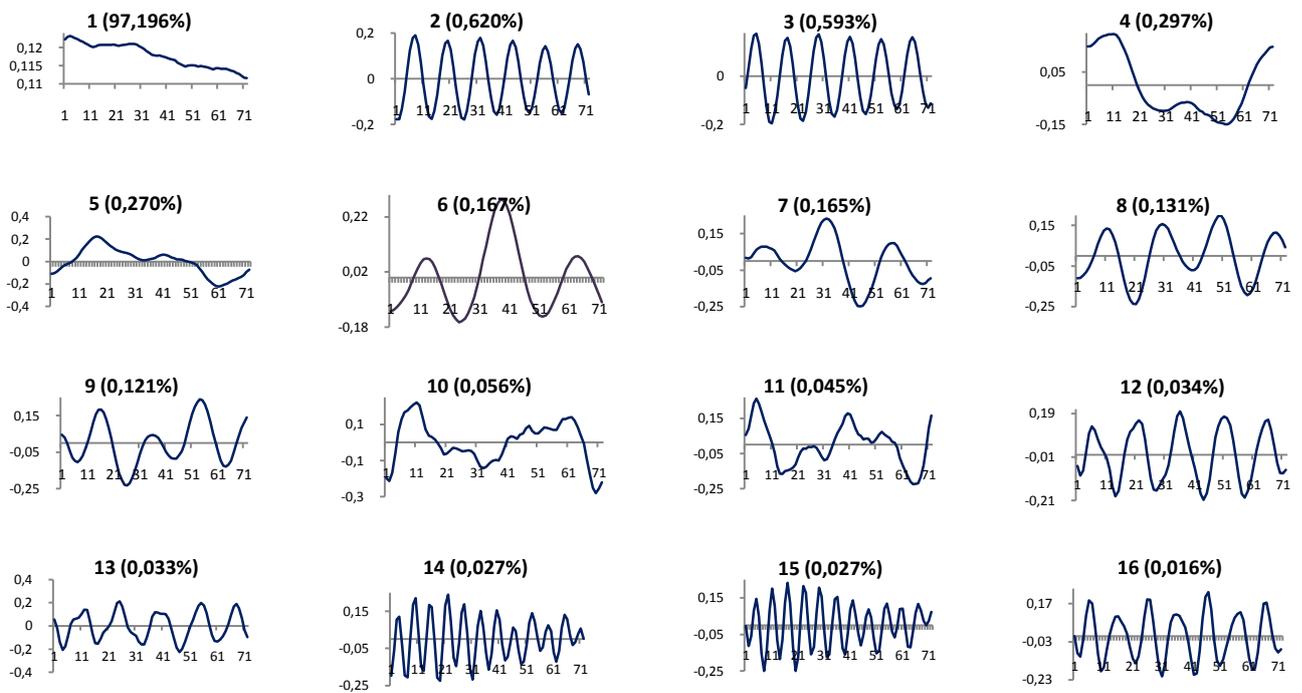


Figura 3. Os 16 primeiros Vetores Singulares.

A Figura 4 apresenta três pares de vetores singulares, verifica-se que os vetores singulares 2 e 3 são componentes harmônicas com período igual a 12 meses, enquanto os vetores singulares 14 e 15 são componentes harmônicas com período igual a 6 meses. Por sua vez, os vetores singulares 6 e 7 são componentes harmônicas e podem ser representados (de forma aproximada) por uma combinação de senos e cossenos.

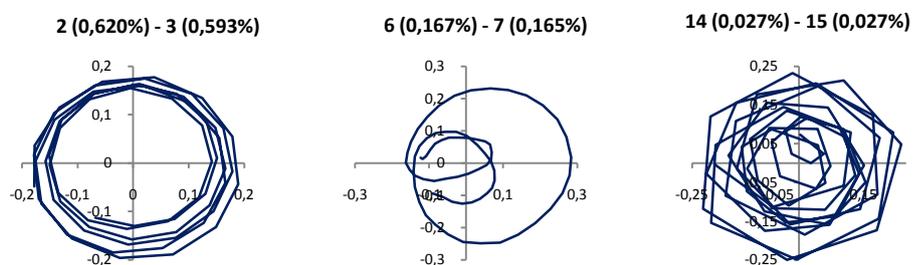


Figura 4. Gráficos de Dispersão de Pares de Vetores Singulares Harmônicos.

Por exclusão, os vetores singulares que não são classificados como componente de tendência ou harmônica, via análise gráfica, são classificados como ruído. Após a análise gráfica, dos 50 vetores singulares da SVD, a divisão mostrada na tabela 4 é obtida.

Tabela 4. Agrupamento dos Vetores Singulares via Análise Gráfica.

COMPONENTE	VETORES SINGULARES
TENDÊNCIA	1,4,5,10,11
HARMÔNICA	2,3,6-9,12-18
RUÍDO	19-50

A tabela 5 mostra a matriz de correlação ponderada das três componentes obtidas usando a análise gráfica dos Vetores Singulares. Segundo os valores apresentados nesta tabela, as três componentes sugerem uma estrutura de separabilidade que as enquadra como bem separáveis.

Tabela 5. Correlação Ponderada entre as três componentes SSA.

COMPONENTE	TENDÊNCIA	HARMÔNICA	RUÍDO
TENDÊNCIA	1	0,014	0
HARMÔNICA	0,01	1	0,028
RUÍDO	0	0,028	1

4.4. RESULTADO DOS MÉTODOS DE PREVISÃO

Para verificação do poder preditivo incorrido com o uso da abordagem SSA na série temporal de consumo residencial de energia elétrica, foram considerados os métodos preditivos *ARIMA* e Holt-Winters. Nas tabelas 6 e 7, encontram-se os resultados das seguintes estatísticas de aderência: R^2 , *MAPE*, *MAE* e *RMSE*. Nas referidas tabelas, HW refere-se à modelagem Holt-Winters e *ARIMA* refere-se à modelagem Box & Jenkins. Pode-se perceber nestas duas tabelas que o poder preditivo dos modelos quando usadas as séries filtradas via SSA supera o dos modelos quando aplicados às séries originais sem a filtragem. Esta percepção está caracterizada pelas medidas de erro (*MAPE*, *MAE* e *RMSE*) menores para os modelos a partir das séries filtradas via SSA e do coeficiente R^2 maior para estes modelos.

Tabela 6. Estatísticas de Aderência dos Métodos Testados (*In Sample*)

	MODELO	R^2	MAPE	MAE	RMSE
SÉRIE TEMPORAL ORIGINAL	HW	0,7784	0,058	0,328	0,433
	ARIMA	0,8073	0,054	0,310	0,403
ACP - SVD	HW	0,8943	0,034	0,303	0,397
	ARIMA	0,8973	0,032	0,303	0,364
SSA ACP ASSOCIADA A ANÁLISE DE CLUSTER	HW	0,9021	0,044	0,301	0,377
	ARIMA	0,8995	0,045	0,315	0,356
ANÁLISE GRÁFICA	HW	0,9601	0,024	0,205	0,296
	ARIMA	0,9626	0,015	0,160	0,243

Tabela 7. Estatísticas de Aderência dos Métodos Testados (*Out of Sample*)

	MODELO	R^2	MAPE	MAE	RMSE
SÉRIE TEMPORAL ORIGINAL	HW	0,6981	0,014	0,470	0,390
	ARIMA	0,7028	0,015	0,450	0,400
ACP - SVD	HW	0,8596	0,014	0,360	0,380
	ARIMA	0,8458	0,015	0,310	0,330
SSA ACP ASSOCIADA A ANÁLISE DE CLUSTER	HW	0,8995	0,011	0,225	0,353
	ARIMA	0,8896	0,012	0,281	0,380
ANÁLISE GRÁFICA	HW	0,9258	0,007	0,190	0,340
	ARIMA	0,9449	0,008	0,120	0,240

5. CONCLUSÕES

Neste artigo foi considerada a filtragem SSA de uma série de consumo de energia antes de sua modelagem. O uso da filtragem SSA integrada ao modelo *ARIMA* ou Holt-Winters de previsão promove ganhos preditivos verificados em todas as estatísticas consideradas. Em particular, o modelo *ARIMA* integrado com o método SSA com abordagem análise gráfica de vetores singulares (cujos resultados estão destacados em vermelho) obteve melhores resultados. Estes resultados fortalecem a hipótese de que a filtragem de séries temporais por SSA antes de sua modelagem apresentam um ganho significativo nas previsões.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Aldenderfer, M. S. and Blashfield, R. K. *Cluster Analysis*. Sage Publications, Califórnia, 1984.
- [2] Box, G. E. P. and Jenkins, G. M. *Time Series Analysis: Forecasting and Control*. Holden-Day, 1970.
- [3] Brock, W. A., Dechert, W., Scheinkman, J., and LeBaron, B. A test for Independence based on the correlation dimension. *Econometric reviews*, 15 (3), 1996, pp. 197-235.
- [4] Broomhead, D. S. and King, G. P. Extracting qualitative dynamics from exponential data. *Physica D*. 20, 1986. pp. 217-236.
- [5] Danilov, D. and Zhigljavsky, A. *Principal Components of time Series: The caterpillar Method*. University of St. Petersburg Press. (In Russian), 1997.
- [6] Elsner, J. B. and Tsonis, A. A. *Singular Spectrum Analysis. A New Tool in Time Series Analysis*. Plenum Press, 1996.
- [7] Gistatgroup. Caterpillar SSA. Petersburg University, Department of Mathematics, Russia, 2010. [Online]. Available at: <http://www.gistatgroup.com/cat/index.html>
- [8] Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. *Analysis of time series structure: SSA and related techniques*. Chapman & Hall/CRC. New York, USA, 2001.
- [9] Hamilton, J. *Time Series Analysis*. Princeton University Press, 1994.
- [10] Hassani, H., Heravi, S., and Zhigljavsky, A. Forecasting European Industrial Production with Singular Spectrum Analysis. *International Journal of Forecasting*. 25, 2009, pp. 103-118.
- [11] Hassani, H., Heravi, S., and Zhigljavsky, A. "Forecasting UK Industrial Production with Multivariate Singular Spectrum Analysis," presented at the 2012 International Conference on the Singular Spectrum Analysis and its Applications, Beijing, China, 2012.
- [12] Kubrusly, C. S. *Elements of Operator Theory*. Birkhäuser, Boston, 2001.
- [13] Lin, K. The ABC's of BDS. *Journal of Computational Intelligence in Finance*, 97 (jul/Aug.). 1997, pp. 23-26.
- [14] Morettin, P. A. e Tolo, C. M. C. *Análise de Séries Temporais*. Edgard Blücher, São Paulo, 2001.
- [15] Souza, R. C. e Camargo, M. E. *Análise de Séries Temporais: os Modelos ARIMA*, 2ª.ed., Gráfica e Editora Regional, Rio de Janeiro, 2004.
- [16] Vasconcelos, S. (2011). Análise de Componentes Principais [Online]. Available at: <http://www.ic.uff.br/aconci/PCA-ACP.pdf>