APLICAÇÃO DA ANÁLISE ESPECTRAL SINGULAR PARA PREVISÃO DE PREÇOS DE PRODUTOS AGRÍCOLAS

Autoria: Valter de Senna, Carlos Alberto Orge Pinheiro

RESUMO

A pesquisa objetiva separar os componentes sinal e ruído de um conjunto M de séries temporais pelos métodos Análise Espectral Singular – AES e Análise Espectral Singular Multivariada – AESM e realizar previsões. As séries utilizadas são preços de produtos agrícolas no período de 22 de janeiro de 2010 a 04 de abril de 2014. As séries temporais apresentaram os menores valores para o Erro Quadrático Médio – EQMs para o método AESM. Isto ocorre uma vez que, ao contrário do método AES, este captura as estruturas que representam o comportamento mais abrangente ao levar em consideração os efeitos entre um conjunto M de séries.

1. Introdução

Existem algumas razões pelas quais os modelos clássicos não apresentam bom desempenho para modelagem e previsão das séries temporais econômicas. Primeiro, um modelo econômico que foi criado para explicar uma relação com um conjunto de pressupostos é inútil se os pressupostos não forem válidos. Com isso, para Plaut e Vautard (1994, p.210), os pressupostos dos modelos clássicos incluem não só aqueles que podem ser expressos como parâmetros do modelo, mas outros com a forma assintótica.

Além disso, muitos modelos utilizados na previsão de séries temporais econômicas são baseados em suposições restritivas de normalidade e linearidade dos dados observados. Acontece que os modelos clássicos de previsão, tais como modelos do tipo ARIMA, são baseados na suposição de estacionariedade da série e normalidade dos resíduos (Box e Jenkins, 1971; Brockwell e Davis, 2002). Assim, os modelos que não dependem destes pressupostos podem ser úteis para a modelagem e previsão de séries econômicas.

Climent, De Miguel e Olmeda (2000) em sua pesquisa consideraram as séries temporais econômicas como determinísticas e lineares. Neste caso, os modelos para séries temporais baseados em suposições de linearidade podem ser utilizados para modelagem e previsão. No entanto, muitas séries temporais econômicas apresentam comportamento não linear (Cao e Soofi, 1999; Hsieh, 1991; Scheinkman e LeBaron, 1989) e, portanto, os modelos lineares não são apropriados.

Com tudo isso, o método Análise Espectral Singular – AES, que é livre das suposições de linearidade e estacionariedade, é indicado para séries temporais lineares e não lineares, estacionárias e não estacionárias. Mais do que isso, a AES é um método não paramétrico de análise de séries temporais incorporando os elementos de análise de séries temporais clássicas, estatística multivariada, geometria multivariada, sistemas dinâmicos e processamento de sinais, conforme explicam Golyandina, Nekrutkin e Zhigljavsky (2001, p.1).

Em sua apresentação básica a AES consiste em dois estágios complementares: decomposição e reconstrução em que ambos incluem dois passos separados. No primeiro estágio a série temporal univariada é decomposta e no segundo a série original é reconstruída. Posteriormente, a série reconstruída poderá ser utilizada para a previsão. O conceito principal da AES é a separabilidade que caracteriza o quão bem, conforme a teoria clássica, os componentes sinal (tendência e sazonalidade) e ruído podem ser separados.

Seu uso é amplo, existindo em pesquisas de finanças, de acordo com Hassani, Dionisio e Ghodsi (2010), que consideraram a AES como um método de filtragem. Em pesquisa de diagnóstico biomédico o ruído foi extraído conforme Ghodsi, Hassani, Sanei e Hick (2009). Também tem sido usado como método de filtragem para e redução de ruído e previsão de consumo de energia elétrica em Kumar e Jain (2010). Outro aspecto importante para AES é que, ao contrário de outros métodos, é adequada para amostras de pequenas dimensões, conforme Golyandina, Nekrutkin e Zhigljavsky (2001, p.28).

Em outras situações, quando o interesse da pesquisa recaiu sobre a captura de estruturas que representassem o comportamento mais abrangente e que levassem em consideração os efeitos entre um conjunto M de séries temporais multivariadas, o método utilizado foi a Análise Espectral Singular Multivariada – AESM. Mantendo os mesmos estágios e passos da AES, a AESM foi inicialmente utilizada em dados atmosféricos. Para isso, grande parte das séries temporais foi extraída de variáveis associadas ao clima e representadas por localidades ou regiões num mapa, conforme pesquisas realizadas por (Keppenne e Ghil, 1993; Plaut e Vautard, 1994).

Esta pesquisa é inspirada na ideia, através dos métodos AES e AESM, de separar os componentes sinal e ruído de um conjunto M de séries temporais, representadas pelos preços dos produtos agrícolas AÇUC (açúcar), ALGO (algodão), ARRO (arroz), CAFE (café) e

SOJA (soja) e, em seguida, realizar previsões para os mesmos. O uso da AESM é dado pela ideia de que como a dinâmica das séries temporais de preços desses produtos passa por mudanças estruturais durante o período de tempo o método multivariado seja adequado por não ser sensível às mudanças dinâmicas.
A estrutura da pesquisa é como se segue. Na seção 2 é apresentada uma introdução aos métodos AES e AESM. Na seção 3 é descrita a metodologia empregada e a amostra. Os resultados dos testes de normalidade e normalidade multivariada das séries temporais são apresentados na seção 4. O desempenho para AES e AESM e os resultados para subséries ruído são considerado na seção 5. Finalmente, na seção 6 são apresentadas as considerações finais e sugestões.

2. Referencial teórico
Pode-se dizer que o principal objetivo da AES é decompor a série temporal univariada em um somatório de subséries, de modo que cada componente desta soma possa ser identificado tanto como tendência e periodicidade (sinal) além de ruído. Em seguida dá-se a reconstrução da série temporal original. Abaixo são apresentados os estágios e passos do método.

2.1 Estágio da decomposição para AES
Neste estágio o passo incorporação pode ser considerado como um mapeamento que transfere uma série temporal unidimensional $Y_T = (y_1,...,y_T)$ para a série multidimensional $X_1,...,X_K$ com vetores $X_i = (y_i,...,y_{i+L-1})^T \in R^L$, onde $K = T - L + 1$ e os vetores X_i são definidos como vetores defasados. Com isso, o único parâmetro da incorporação é o comprimento da janela L, um número inteiro que deve ser $2 \le L \le T-1$ conforme Golyandina, Nekrutkin e Zhigljavsky (2001, p.18). No entanto, conforme Golyandina (2010, p.263), resultados teóricos indicam que L deve ser suficientemente grande, mas, não superior a $T/2$. O resultado deste passo é a definição da matriz trajetória $X = [X_1,..., X_K] = (x_{ij})_{i,j=1}^{L,K}$, de forma que a mesma é uma matriz Hankel, uma vez que suas entradas são constantes ao longo das diagonais paralelas à diagonal secundária.
No passo decomposição em valores singulares – DVS da matriz trajetória é obtida uma soma de matrizes elementares. Assim, denota-se por $\lambda_1,...,\lambda_L$ os autovalores de XX^T em ordem decrescente de magnitude $(\lambda_1 \ge ... \ge \lambda_L \ge 0)$ e por $U_1,...,U_L$ os autovetores ortogonais. Ao estabelecer que $V_i = X^T U_i \big/ \sqrt{\lambda_i}$, a DVS da matriz trajetória pode ser escrita como:

$$X = X_1 + ... + X_d \qquad (1)$$

onde $X_i = \sqrt{\lambda_i} U_i V_i^T$ representa uma matriz de posto unitário ou comumente matriz elementar.

2.2 Estágio da reconstrução para AES
Neste estágio o passo agrupamento corresponde em dividir as matrizes elementares em grupos somando-as dentro de cada grupo (sinal e ruído). Ao deixar que $I = \{i_1,...,i_p\}$ seja um grupo de índices $i_1,...,i_p$, então, a matriz X_I que corresponde ao grupo I é definida por $X_I = X_{i_1},...,X_{i_p}$. Assim, o desdobramento do conjunto de índices $J = \{1,...,d\}$ em subconjuntos disjuntos $I_1,...,I_m$ corresponde a representação:

$$X = X_{I_1} + ... + X_{I_m} \qquad (2)$$

onde $X_{I_1},...,X_{I_m}$ são definidas como matrizes resultantes. Na expressão (2) tem-se uma nova decomposição de matrizes, esta é denominada como decomposição agrupada.

O conceito de separabilidade apresenta destaque neste estágio. Desta forma, considerando que a série temporal original Y_T pode ser representada pela soma de duas subséries $Y_T = \left(Y_T^{(1)} + Y_T^{(2)}\right)$ representando sinal e ruído, respectivamente, a separabilidade das subséries $Y_T^{(1)}$ e $Y_T^{(2)}$ implica que os componentes obtidos pela DVS da matriz trajetória X podem ser agrupados em dois diferentes grupos de forma que a soma das matrizes em cada grupo origina as matrizes trajetórias $X^{(1)}$ e $X^{(2)}$ das subséries $Y_T^{(1)}$ e $Y_T^{(2)}$.

A separabilidade significa que cada linha da matriz trajetória $X^{(1)}$ é ortogonal a cada linha da matriz trajetória $X^{(2)}$, valendo também para as colunas. Para Golyandina, Nekrutkin e Zhigljavsky (2001, p.47) não ocorre separabilidade exata, mas, tão somente separabilidade aproximada. A qualidade da separabilidade aproximada é avaliada pela medida denominada correlação ponderada ou w-correlação.

Então, ao considerar as duas subséries $Y_T^{(1)}$ e $Y_T^{(2)}$, pode-se avaliar a qualidade da separação entre elas através da seguinte expressão:

$$\rho_{12}^{w} = \frac{\left\langle Y_T^{(1)}, Y_T^{(2)} \right\rangle_w}{\left\| Y_T^{(1)} \right\|_w \left\| Y_T^{(2)} \right\|_w} \qquad (3)$$

com $\left\| Y_T^{(i)} \right\|_w = \sqrt{\left\langle Y_T^{(i)}, Y_T^{(i)} \right\rangle_w}$ representando a norma da $i-ésima$ subsérie e $\left\langle Y_T^{(i)}, Y_T^{(j)} \right\rangle_w$ o produto interno entre um par de subséries, considerando que o ponderador w_k é definido por $w_k = \min\{k, L, T-k\}$ e que $L \le T/2$.

A expressão (3) pode apresentar valores entre 0 (zero) e 1 (um) . Desta maneira, ao considerar $\rho_{12}^{w} = 0$ tem-se que os componentes das subséries $Y_T^{(1)}$ e $Y_T^{(2)}$ são separáveis e de outra forma, quando $\rho_{12}^{w} = 1$ isto significa que os componentes das subséries $Y_T^{(1)}$ e $Y_T^{(2)}$ não são tão bem separáveis, ou seja, devem ser reunidos num mesmo grupo.

No segundo passo a média diagonal transforma a matriz obtida na decomposição agrupada em (2) para a forma de uma matriz Hankel, que pode ser posteriormente convertida a uma série temporal. Este procedimento é definido como média diagonal ou Hankelização da matriz. O resultado da Hankelização de uma matriz Z é a matriz Hankel $\mathrm{H}Z$.

Ao aplicar a Hankelização a todos os componentes na expressão (2) obtém-se a expansão $X = \widetilde{X}_{I_1} + ... + \widetilde{X}_{I_m}$ onde $\widetilde{X}_{I_1} = \mathrm{H}X$. Isto é equivalente à decomposição da série inicial $Y_T = \left(y_1,...,y_T\right)$ em um somatório de m séries; $y_t = \sum_{p=1}^{m} \widetilde{y}_t^{(p)}$, onde $\widetilde{Y}_T^{(p)} = \left(\widetilde{y}_1^{(p)},...,\widetilde{y}_T^{(p)}\right)$ corresponde à matriz X_{I_p}.

A série que resulta da operação acima, conforme $\widetilde{Y}_T^{(p)} = \widetilde{y}_1^{(p)},...,\widetilde{y}_T^{(p)}$, é obtida pela aplicação do procedimento de Hankelização em cada matriz (2). Com isso, se a Hankelização é aplicada a todos os componentes obtém-se a formulação:

$$X = \widetilde{X}_{I_1} + ... + \widetilde{X}_{I_m} \qquad (4)$$

O agrupamento adequado é responsável por uma decomposição em que as matrizes (4) são quase Hankel, levando a uma separabilidade aproximada. Sob a condição de que cada matriz (4) é uma matriz de Hankel, cada uma dessas matrizes determina unicamente a série $\widetilde{Y}_T^{(p)}$ e, portanto, a série inicial Y_T é decomposta na soma de m séries sendo responsável pela formulação:

$$\widetilde{Y}_T = \widetilde{y}_T^{(1)} + ... + \widetilde{y}_T^{(m)} \quad (5)$$

considerando $T = 1,...,T$ e para cada p a série $\widetilde{Y}_T^{(p)}$ é o resultado do processo de Hankelização da matriz X_{I_p}. Assim, na decomposição em (5) tem-se a soma de m componentes separáveis, já em forma de séries temporais.

2.3 Algoritmo de previsão para AES

O algoritmo de previsão é dado conforme Golyandina, Nekrutkin e Zhigljavsky (2001, p.95):
(a) Série temporal $Y_T = (y_1,...,y_T)$, $T > 2$.
(b) Comprimento da janela $L = T/2$.
(c) Espaço Linear $\ell_r \subset R^L$ de dimensão $r < L$. Supõe-se que $e_L \notin \ell_r$ onde $e_L = (0,0,...,1)^T \in R^L$.
(d) Número M de pontos para previsão.
Na sequencia os autores definem notações e comentários:
(a) $X = [X_1,...,X_K]$ é a matriz de trajetória da série temporal Y_T.
(b) $P_1,...,P_r$ é uma base ortonormal em ℓ_r.
(c) $\widehat{X} = [\widehat{X}_1 : ... : \widehat{X}_K] = \sum_{i=1}^{r} P_i P_i^T X$. O vetor $\widehat{X}_i$ é a projeção ortogonal de X_i dentro do espaço ℓ_r.
(d) $\widetilde{X} = H\widehat{X} = [\widetilde{X}_1 : ... : \widetilde{X}_K]$ é o resultado da Hankelização da matriz X.
(e) Para qualquer vetor $Y \in R^L$ denota-se por $Y_\Delta \in R^{L-1}$ o vetor composto dos últimos componentes $L-1$ do vetor Y, enquanto $Y^\nabla \in R^{L-1}$ o vetor dos primeiros componentes $L-1$ do vetor Y.
(f) Estabelece-se que $v^2 = \pi_1^2 + ... + \pi_r^2$, considerando que π_i é o último componente do vetor $P_i (i = 1,...,r)$.
(g) Supondo que $e_L \notin \ell_r$, os autores explicam que isto implica que ℓ_r não é um espaço vertical, então, $v^2 \leq 1$. Assim, pode ser provado que o último componente y_L de qualquer vetor $Y = (y_1,..., y_L)^T \in \ell_r$ é uma combinação linear dos primeiros componentes $(y_1,...,y_{L-1})$, conforme:

$$y_L = a_1 y_{L-1} + ... + a_{L-1} y_1$$

O vetor $A = (a_1,...,a_{L-1})$ pode ser expresso como:

$$A = \frac{1}{1-v^2} \sum_{i=1}^{r} \pi_i P_i^\nabla$$

e a dose não depende da escolha de uma base $P_1,...,P_r$ no espaço linear ℓ_r. Nas notações acima a série temporal $Y_{T+M} = (y_1,...,y_{T+M})$ é definida pela expressão:

$$y_i = \begin{cases} \widetilde{y}_i & para \quad i = 1,...,T \\ \sum_{j=1}^{L-1} a_j y_{i-j} & para \quad i = T+1,...,T+M \end{cases} \quad (6)$$

Os números $y_{T+1},...,y_{T+M}$ dos termos M da previsão recorrente AES. A definição do operador linear $P^{(r)} : \ell_r \to R^L$ é dada pela fórmula:

$$P^{(r)} Y = \begin{pmatrix} Y_\Delta \\ A^T Y_\Delta \end{pmatrix}, Y \in \ell_r$$

estabelecendo que:

$$Z_i = \begin{cases} \widetilde{X}_i & para \quad i = 1,...,K \\ P^{(r)} Z_{i-1} & para \quad i = K+1,...,K+M \end{cases} \tag{7}$$

com a matriz $Z = [Z_1,...Z_{K+M}]$ representando a matriz da trajetória das séries Y_{T+M}. Portanto, (7) pode-se considerar como a forma do vetor em (6).

2.4 Estágio da decomposição para AESM

Embora a técnica AESM siga a estrutura da AES contendo os mesmos estágios e passos, por utilizar um conjunto M de séries temporais isto acaba requerendo algumas particularidades na formação da matriz XX^T e na definição do comprimento da janela L. Tal particularidade ganha destaque uma vez que através da definição adequada do comprimento da janela é possível capturar a periodicidade da série temporal.

Assim, para o primeiro estágio a incorporação pode ser considerada como um mapeamento que transfere um conjunto M de séries temporais unidimensionais $Y_{T_i}^{(i)} = \left(y_1^{(i)},...,y_{T_i}^{(i)}\right)$, com $i = 1,...,M$, para uma matriz multidimensional $\left[X_1^{(i)},...,X_{K_i}^{(i)}\right]$ com vetores $X_j^{(i)} = \left(y_j^{(i)},...,y_{j+L_i+1}^{(i)}\right)^T \in R^{L_i}$, onde $K_i = T_i - L_i + 1$. Os vetores $X_j^{(i)}$ são chamados de vetores defasados. Semelhante à técnica AES a matriz $X^{(i)}$ é uma matriz Hankel. Neste passo, considerando um conjunto M de séries temporais, com $T = 1,...,T$, são definidas as matrizes trajetórias $X^{(i)}$, para $i = 1,...,M$ em cada série temporal $Y_{T_i}^{(i)}$, todas com a mesma dimensão $(L \times (T - L + 1))$. O resultado deste passo é a formação de um bloco de matrizes trajetórias X_V, conforme:

$$X_V = \begin{bmatrix} X^{(1)} \\ \vdots \\ X^{(M)} \end{bmatrix} \tag{8}$$

o bloco de matrizes trajetórias X_V representa um formato vertical.

No segundo passo é realizada a DVS do bloco de matrizes trajetórias $X_V X_V^T$ obtendo uma soma de matrizes elementares. Assim, denota-se por $\lambda_{V_1},...,\lambda_{V_{M\times L}}$ os autovalores de $X_V X_V^T$ em ordem decrescente de magnitude $\left(\lambda_{V_1} \geq ... \geq \lambda_{V_{M\times L}} \geq 0\right)$ e por $U_{V_1},...,U_{V_{M\times L}}$ os autovetores ortogonais. A matriz $X_V X_V^T$, de dimensão $(ML \times ML)$, é dada conforme:

$$X_V X_V^T = \begin{bmatrix} X^{(1)}X^{(1)T} & X^{(1)}X^{(2)T} & \cdots & X^{(1)}X^{(M)T} \\ X^{(2)}X^{(1)T} & X^{(2)}X^{(2)T} & \cdots & X^{(2)}X^{(M)T} \\ \vdots & \vdots & \ddots & \vdots \\ X^{(M)}X^{(1)T} & X^{(M)}X^{(2)T} & \cdots & X^{(M)}X^{(M)T} \end{bmatrix} \tag{9}$$

A estrutura em (9) é similar a matriz de variância-covariância obtida na literatura clássica da análise estatística multivariada conforme Hassani e Mahmoudvand, (2013, p.59). A matriz $X^{(i)}X^{(i)T}$ é a mesma utilizada na AES para uma única série temporal $Y_{T_i}^{(i)}$. Semelhante ao obtido em AES, a DVS nesse passo é dada por:

$$X_V = X_{V_1} + ... + X_{V_{M\times L}} \tag{10}$$

onde $X_{V_i} = \sqrt{\lambda_{V_i}} U_{V_i} V_{V_i}^T$ representa a matriz elementar e $V_{V_i} = X_V^T U_{V_i} \big/ \sqrt{\lambda_{V_i}}$.

2.5 Estágio da reconstrução para AESM

Semelhante a AES o agrupamento corresponde em dividir as matrizes elementares $X_{V_1},...,X_{V_{d_V}}$ em grupos disjuntos somando-as dentro de cada grupo. Assim, o desdobramento do conjunto de índices $J=\{1,...,d_V\}$ em subconjuntos disjuntos $I_1,...,I_m$ corresponde a representação:

$$X_V = X_{I_1} + ... + X_{I_m} \qquad (11)$$

onde $X_{I_1},...,X_{I_m}$ são definidas como matrizes resultantes.

Assim, como um caso simples que apresenta os componentes sinal e ruído, são usados dois grupos de índices, conforme $I_1=\{1,...,r\}$ e $I_2=\{r+1,...,d_V\}$, o primeiro grupo associado ao componente sinal e o último ao ruído.

No passo seguinte a média diagonal transforma a matriz obtida na decomposição agrupada em (9) para a forma de uma matriz Hankel, que pode ser posteriormente convertida a uma série temporal. Considera-se $\widetilde{X}^{(i)}$ uma aproximação da matriz $X^{(i)}$ obtida a partir do passo média diagonal. Se $\widetilde{x}_{mn}^{(i)}$ é um elemento da matriz $\widetilde{X}^{(i)}$ o $j-ésimo$ termo da série reconstruída $\widetilde{Y}_{T_i}^{(i)} = \left(\widetilde{y}_1^{(i)},...,\widetilde{y}_j^{(i)},...,\widetilde{y}_{T_i}^{(i)}\right)$ é obtido pela média aritmética $\widetilde{x}_{mn}^{(i)}$ para todo (m,n) de modo que $m+n-1=j$.

2.6 Algoritmo de previsão para AESM

A previsão obtida a partir de um conjunto M de séries temporais é dada:

$$\left[\hat{y}_{j_1}^{(1)},...,\hat{y}_{j_M}^{(M)}\right]^T = \begin{cases} \left[\widetilde{y}_{j_1}^{(1)},...,\widetilde{y}_{j_M}^{(M)}\right], & j_i = 1,...,T_i \\ \left(I_{MxM} - WW^T\right)^{-1} W U^{\nabla M^T} Z_h & j_i = N_i+1,...,N_i+h \end{cases} \qquad (12)$$

com $U_j^{(i)\nabla}$ representando os primeiros $L_i - 1$ componentes do vetor $U_j^{(i)}$ e $\pi_j^{(i)}$ os últimos componentes do vetor $U_j^{(i)}$ com $(i=1,...M)$. Além dessas considerações a matriz $U^{\nabla M}$ é dada conforme:

$$U^{\nabla M} = \begin{bmatrix} U_j^{(i)\nabla} \\ \vdots \\ U_j^{(M)\nabla} \end{bmatrix} \qquad (13)$$

e a matriz W representada por:

$$W = \begin{bmatrix} \pi_1^{(1)} & \pi_2^{(1)} & \cdots & \pi_r^{(1)} \\ \pi_1^{(2)} & \pi_2^{(2)} & \cdots & \pi_r^{(2)} \\ \vdots & \vdots & \cdots & \vdots \\ \pi_1^{(M)} & \pi_2^{(M)} & \cdots & \pi_r^{(M)} \end{bmatrix} \qquad (14)$$

Além disso, $Z_h = \left[Z_h^{(1)},...,Z_h^{(M)}\right]^T$ e $Z_h^{(i)} = \left[\hat{y}_{Ni-Li+h+1}^{(i)},...,\right]\hat{y}_{Ni+h+1}^{(i)}$ com $(i=1,...M)$.

Desta forma, se o algoritmo de previsão em AES era definido com base nas fórmulas recorrentes lineares a previsão para um conjunto M tem por base a fórmula recorrente multilinear.

3. Metodologia

Nesta pesquisa são aplicados os métodos AES e AESM para decompor e reconstruir a partir de um conjunto M de séries temporais e realizar previsões através dos algoritmos vistos nas seções 2.3 e 2.6. Assim, foi realizada a separação entre sinal e ruído para posterior previsão de cada série temporal individualmente. Em seguida, o desempenho das previsões obtidas para diferentes passos à frente é avaliado.

3.1 Amostra
As séries temporais escolhidas, nesta pesquisa, são preços de produtos agrícolas, conforme: AÇUC (açúcar), ALGO (algodão), ARRO (arroz), CAFE (café) e SOJA (soja). Eles foram obtidos no banco de dados do Centro de Estudos Avançados em Economia Aplicada da Escola Superior de Economia Luiz de Queiroz – ESALQ e correspondem ao período de 22 de janeiro de 2010 a 04 de abril de 2014, cuja periocidade totaliza 204 semanas. Todas as séries na pesquisa são apresentadas na forma logarítmica.
A amostra acima foi dividida em dois grupos. O primeiro grupo se refere ao conjunto das séries temporais que são utilizadas pela AES e AESM, e o segundo grupo, composto das 12 últimas semanas da amostra, foi utilizado para avaliação de desempenho da previsão realizada.

4. Resultados dos testes de normalidade
Os testes de Anderson-Darling (A-D) e de Shapiro Wilk (S-W) são usados para testar se uma amostra de dados tem origem de uma população com uma distribuição específica. Todos os dois testes tendem a funcionar bem na identificação de uma distribuição como não-normal quando a distribuição em questão está distorcida. No entanto, são menos exigentes quando a distribuição é uma distribuição t e a não-normalidade é devido à curtose. Em geral, entre os dois testes baseados na função de distribuição empírica, o teste A-D tende a ser mais eficaz na detecção de desvios na cauda da distribuição. Na pesquisa os dois testes são utilizados para uma visão abrangente dos resultados.
Os testes rejeitam a hipótese de normalidade quando o valor p for menor ou igual a 0,05. Assim, o teste de normalidade permite afirmar com confiança de 95% que os dados não se ajustam à distribuição normal. A Tabela 1 representa os resultados do teste de normalidade para um nível de 5% de significância. Como pode ser visto, a partir dos resultados, todas as séries não estão distribuídas normalmente.

Tabela 1 – Teste de normalidade

	AÇUC	ALGO	ARRO	CAFE	SOJA
Número de Observações	192	192	192	192	192
Shapiro-Wilk	0,93	0,81	0,96	0,95	0,96
p(valor)	0,00	0,00	0,01	0,00	0,00
Anderson-Darling	4,59	11,76	2,62	2,75	3,23
p(valor)	0,00	0,00	0,00	0,00	0,01

Fonte: Dados obtidos pelos autores

As séries temporais econômicas podem apresentar uma estrutura com tendências não-lineares e sazonalidade complexa. Vale ressaltar que os métodos AES e AESM não assumem linearidade ou normalidade dos dados observados. Para avaliar o aspecto da normalidade do conjunto de dados, fez-se o uso do teste Doornik-Hansen-Omnibus (DHO) que é um teste de normalidade multivariada. O teste foi aplicado entre pares formados pelas séries temporais. Os resultados representados na Tabela 2 indicam que há fortes evidências de não-normalidade multivariada para um nível de 5% de significância entre cada par de série.

Tabela 2 – Teste de normalidade multivariada DHO e p(valor)

	AÇUC	ALGO	ARRO	CAFE	SOJA
AÇUC		11,5	17,76	9,47	26,46
		0,00	0,00	0,00	0,00
ALGO			26,4	13,1	20,6
			0,00	0,00	0,00
ARRO				22,4	14,6
				0,00	0,00
CAFE					19,7
					0,00

Fonte: Dados obtidos pelos autores

5. Desempenho para AES e AESM

Para a modelagem da AES e AESM foi utilizado o software Lingo, versão 11. Este software é comercial e oferece grande quantidade de algoritmos e características para a construção das matrizes necessárias para o método. Assim, para realização dos dois estágios e, conforme descrito na seção 2.1 o valor de L deve ser igual a $T/2$ em AES. Já para AESM, conforme Hassani e Mahmoudvand, (2013, p.68), a definição do comprimento da janela L é dada por:

$$L = \frac{1}{M+1}(T+1) \qquad (15)$$

com M representando o número de séries temporais e T o número de observações.
Então, o passo agrupamento que corresponde dividir as matrizes elementares em grupos disjuntos somando-as dentro do grupo sinal, de um lado e ruído do outro, irá requerer a separação, com base no conceito de separabilidade ponderada.
Então, com o objetivo de minimizar a correlação ponderada, conforme (3), com base numa escolha binária dos autovalores para definição dos grupos sinal e ruído, a definição das subséries sinal para cada índice do mercado acionário foi realizada por um processo de otimização através do software Lingo. Para isso, o valor de L foi definido em 32 para a AES e em 96 para AESM. A Tabela 3, com base nos resultados obtidos, indica que a correlação ponderada entre sinal e ruído para cada série temporal foi reduzida. Tal separação é importante uma vez que a subsérie sinal é utilizada no algoritmo de previsão para definição dos passos à frente.

Tabela 3 – Correlação entre sinal e ruído e Erro Quadrático Médio

	Parâmetros		Correlação Sinal / Ruído		EQM	
Série	L	h	AESM	AES	AESM	AES
AÇUC	32 / 96					
		1			7,87E-04	9,15E-03
		3			6,45E-04	1,75E-02
		6			3,68E-03	6,90E-02
		12			4,29E-03	4,84E-01
			5,072E-04	3,621E-03		

ALGO	32 / 96					
		1			8,57E-05	1,76E-02
		3			2,65E-04	2,71E-01
		6			3,95E-04	5,26E-02
		12			1,26E-03	5,80E-03
			5,470E-04	5,033E-04		
ARRO	32 / 96					
		1			3,08E-03	3,47E-03
		3			6,77E-03	1,48E-02
		6			7,62E-03	3,80E-02
		12			1,11E-02	6,57E-02
			1,013E-03	1,746E-03		
CAFÉ	32 / 96					
		1			1,46E-04	2,93E-04
		3			6,94E-05	3,74E-04
		6			3,79E-05	4,86E-03
		12			1,13E-04	1,54E-02
			5,926E-03	6,252E-04		
SOJA	32 / 96					
		1			1,32E-02	5,62E-01
		3			1,79E-02	2,73E-01
		6			2,15E-02	5,40E-01
		12			2,98E-02	3,16E-01
			1,015E-03	6,568E-04		

Fonte: Dados obtidos pelos autores

Então, as previsões obtidas nos passos à frente h (1, 3, 6 e 12 semanas), foram confrontadas com o segundo grupo da amostra, composto das 12 últimas semanas. Para isso a avaliação deu-se através do uso do EQM, conforme:

$$EQM = \frac{1}{h}\sum_{j=k+1}^{T}\left(Y_j - \hat{Y}_j\right)$$

com Y_j representando o valor da série, $\hat{Y}_j$ o valor da previsão e h a quantidade de observações reservadas para avaliação.

Semelhante ao trabalho de Esquível (2012) quando o horizonte h aumenta a qualidade da previsão não apresenta bons resultados, conforme Tabela 3. Ainda na mesma Tabela é possível perceber que para as séries AÇUC (açúcar), ALGO (algodão), ARRO (arroz), CAFE (café) e SOJA (soja) os menores valores de EQMs foram obtidos para o modelo AESM que captura as estruturas que representam o comportamento mais abrangente ao levar em consideração os efeitos entre um conjunto M de séries.

5.1 Resultados para subséries ruído

Com a finalidade de avaliar se as subséries ruído, obtidas conforme descrito na seção 5, tanto para AESM como para AES, são estacionárias, foi realizado o teste de raiz unitária de Dickey-Fuller Aumentado. As Tabelas 4 e 5 mostram os resultados do teste. A hipótese nula de que as subséries possuem raiz unitária e, portanto, são não estacionárias, é rejeitada para as subséries ao nível de significância 5% uma vez que os valores do teste são menores do que o valor crítico.

Tabela 4 – Teste de raiz unitária em AESM

Subsérie ruído	Dickey Fulller	Valor Crítico 5%	p(valor)
AÇUC	-3,95	-3,43	0,01
ALGO	-7,64	-3,43	0,00
ARRO	-3,44	-3,43	0,03
CAFÉ	-5,50	-3,43	0,00
SOJA	-4,15	-3,43	0,01

Fonte: Dados obtidos pelos autores

Tabela 5 – Teste de raiz unitária em AES

Subsérie ruído	Dickey Fulller	Valor Crítico 5%	p(valor)
AÇUC	-5,43	-3,43	0,00
ALGO	-3,55	-3,43	0,02
ARRO	-11,80	-3,43	0,00
CAFE	-7,94	-3,43	0,00
SOJA	-6,78	-3,43	0,00

Fonte: Dados obtidos pelos autores

Também podemos interpretar o p(valor) que conduz a mesma conclusão uma vez que os mesmos, conforme Tabelas 4 e 5, são menores que o nível de significância prefixado. Assim, a probabilidade de existir equívocos ao rejeitar a hipótese de raiz unitária é menor do que o disposto (5%), logo se rejeita a hipótese. Assim, considerando que as séries temporais não apresentam raiz unitária as mesmas são consideradas estacionárias. Assim, suas propriedades estatísticas não sofrem modificação no tempo, percebendo, então, que a tendência e sazonalidade foram excluídas, o que torna as séries analisadas com características de ruído.

6. Considerações finais e sugestões

Levando em conta que a dinâmica dos preços agrícolas tem passado por mudanças econômicas e climáticas no tempo, é preciso ter certeza de que o método de previsão não é sensível a essas variações dinâmicas. Neste contexto, o modelo AESM pode ser considerado como aquele que não é sensível às quebras estruturais. A motivação pela utilização do método AESM como também do AES dá-se por causa da capacidade em lidar com séries estacionárias, bem como com séries não-estacionárias. Além disso, ao contrário dos métodos clássicos de previsão de séries temporais (que assumem normalidade e estacionariedade das séries), os métodos são não-paramétricos, não fazendo, portanto, suposições prévias sobre os dados observados.

As séries históricas nesta pesquisa apresentam uma estrutura complexa e mudanças estruturais uma vez que não se ajustam à distribuição normal como também forte evidência de não-normalidade multivariada. Inicialmente a separação entre sinal e ruído para cada série temporal foi realizada. Com isso, a correlação ponderada obtida para cada série temporal, entre as subséries sinal e ruído, foi próxima à zero. Tal separação é fundamental uma vez que a subsérie sinal é utilizada no algoritmo de previsão para definição dos passos à frente. Os resultados da previsão para passos à frente foram favoráveis, no entanto, semelhante a outras pesquisas, quando o horizonte de previsão aumentou a qualidade da previsão não apresentou bom desempenho.

As correlações entre sinal e ruído para série de produtos agrícolas podem ser consideradas reduzidas de forma que as mesmas foram favoráveis à previsão dos preços nos passos à frente 1, 3, 6 e 12 semanas. As séries temporais AÇUC (açúcar), ALGO (algodão), ARRO (arroz), CAFE (café) e SOJA (soja) apresentaram os menores valores de EQMs para o método AESM. Isto ocorre uma vez que, ao contrário do método AES, este captura as estruturas que representam o comportamento mais abrangente ao levar em consideração os efeitos entre um conjunto M de séries.

A pesquisa realizada contribui para finanças à medida que agrega evidências favoráveis à generalidade da eficácia do método AESM aplicado no mercado agrícola. Do ponto de vista prático, os resultados obtidos podem auxiliar os profissionais do mercado na tomada de decisões de investimento e análise do mercado. Para próximas pesquisas sugere-se a utilização de outras bases de dados, a inclusão de outros produtos agrícolas, a adoção de outros períodos de análise e a inclusão de outras variáveis que possam aumentar o poder explicativo do método.

7. Referências

Box, G. E. P., & Jenkins, G. M. (1971). Time series analysis: Forecasting and control, *Operational Research Quarterly*, 22, 199-201.

Brockwell, P. J., & Davis R. A. (2002). *Introduction to Time Series and Forecasting*, 2nd edition. Springer, New York.

Cao, L. Y., & Soo, A. (1999). Nonlinear deterministic forecasting of daily dollar exchange rates, *International Journal of Forecasting*, 15, 421-430.

Climent, F. J., De Miguel, M. Del M., Olmeda, I. (2000). Linear and Non-Linear Dynamics Between Exchange Rates and Stock Markets Returns: An Application to the Financial Crises of Europe and Asia in the Ninetie. *Review of Financial Markets*, 5, 19-48.

ESQUIVEL, R.M. (2012). *Análise espectral singular: modelagens de séries temporais através de estudos comparativos usando diferentes estratégias de previsão*. Dissertação de mestrado, Faculdade de Tecnologia SENAI CIMATEC, Salvador, Ba, Brasil.

Ghodsi, M., Hassani, H., Sanei, S., & Hick, Y. (2009). The use of noise information for detecting temporomandibular disorder, *Biomedical Signal Processing and Control*, 4, 79-85.

Golyandina, N., Nekrutkin, V., & Zhigljavsky, A. (2001). Analysis of Time Series Structure: SSA and related techniques, Chapman & Hall/CRC, New York.

Golyandina, N. (2010). On the choice of parameters in Singular Spectrum Analysis and related subspace-based methods. Statistics and Its Interface, 3, 259-279.

Hassani, H., & Mahmoudvand, R. (2013). Multivariate singular spectrum analysis: a general view and new vector forescasting approach. *International Journal of Energy and Statistics*, 1, 55–83.

Hassani, H., Dionisio, A., & Ghodsi, M. (2010). The effect of noise reduction in measuring the linear and nonlinear dependency of financial markets, *Nonlinear Analysis: Real World Applications*, 11, 492-502.

Hsieh, D. A. (1991). Chaos and nonlinear Dynamics: Application to Financial Markets, *Journal of Finance*, 46, 1839-1877.

Keppenne, C. L., & M. Ghil. (1993). Adaptive filtering and prediction of noisy multivariate signals: An application to subannual variability in atmospheric angular momentum, *International Journal of Bifurcation and Chaos*, 3, 625–634.

Kumar U., & Jain V.K. (2010). Time Series Models (Grey-Markov, Grey Model with rolling mechanism and Singular Spectrum Analysis) to forecast Energy Consumption in India. *Energy*, 35, 1709-1716.
Plaut, G., & Vautard, R. (1994). Spells of low-frequency oscillations and weather regimes in the Northern Hemisphere, *Journal of the Atmospheric Sciences*, 51, 210-236.
Scheinkman, J., & LeBaron, B. (1989). Nonlinear Dynamics and Stock Returns, *Journal of Business*, 62, 311-337.