

Previsão de preços através de redes neurais e análise espectral: evidências para o mercado futuro das commodities açúcar e soja

Recebimento dos originais: 05/02/2017
Aceitação para publicação: 30/12/2017

Carlos Alberto Orge Pinheiro

Mestre em Modelagem Computacional e Tecnologia Industrial
Centro Universitário Senai Cimatec e Universidade do Estado da Bahia – UNEB
Av. Orlando Gomes, 1845 – Piatã • CEP 41613-036 • Salvador, BA
E-mail: carlos.orge@terra.com.br

Valter de Senna

Doutor em Pesquisa Operacional
Centro Universitário Senai Cimatec
Av. Orlando Gomes, 1845 – Piatã • CEP 41613-036 • Salvador, BA
E-mail: valter.senna@gmail.com

Resumo

Este estudo teve como proposta explorar a combinação do modelo de Análise Espectral Singular (AES) com o modelo de Rede Neural Artificial (RNA). A escolha das commodities açúcar e soja deu-se por conta do destaque na pauta de exportações no ano de 2014. A combinação proposta demonstrou-se adequada, quando comparada com os modelos clássicos e de redes neurais, ao permitir melhores desempenhos das previsões dos preços. Os resultados obtidos no período fora da amostra, mediante o uso das medidas de erro e do teste estatístico preditivo para o passo à frente, confirmam isso. A pesquisa pode auxiliar os profissionais do mercado na formulação e aplicação de políticas direcionadas ao setor agrícola por conta da relevância da previsão dos preços como instrumento de planejamento bem como na análise do comportamento do mercado agrícola ao especificar a tendência dos preços das commodities no mercado futuro.

Palavras-Chave: Análise espectral. Rede neural. Mercados futuros.

1. Introdução

Nos últimos anos a agricultura no Brasil se desenvolveu e modernizou, destacando-se como atividade de alto valor econômico e social com tendência de crescimento acentuado (CAMPOS, 2007). O desenvolvimento dos mercados interno e externo contribuiu para o caráter dinâmico da agricultura, com a incorporação de tecnologias produtivas para satisfazer às exigências desses mercados. Para Oranje (2003) o melhor desempenho da agricultura pode ser obtido através da competitividade dos preços, explicada, pela opção dos importadores

substituírem o consumo de produtos, cujos preços se elevam, por aqueles que não sofrem majoração.

A análise de preços, como elemento do mecanismo de troca, reveste-se de importância para os participantes do mercado agrícola, sejam eles compradores, vendedores ou estejam em busca de proteção. Então, as decisões tomadas pelos produtores, mesmo antes da realização da colheita, presumem o conhecimento do comportamento dos preços (RIBEIRO, SOSNOSKI e OLIVEIRA, 2010). Além delas, as decisões referentes ao conjunto de produção e à adoção de políticas de financiamento como alternativa de assegurar preços, são baseadas em expectativas dos preços.

Como as atividades agrícolas são caracterizadas por apresentarem movimentos cíclicos, sofrendo influência de vários fatores do mercado e apresentado elevada volatilidade (Oliveira e Aguiar, 2003) isso acaba se constituindo num obstáculo para sua previsibilidade. Portanto, a previsão dos preços passa a ser uma das principais ferramentas para execução do planejamento e avaliação das atividades agrícolas uma vez que é fator de decisão dos investimentos.

O principal problema sobre a previsão dos preços dos produtos agrícolas reside na sazonalidade que ocorre em razão dos fatores climáticos, mercadológicos e conjunturais que promovem incerteza de renda para os agricultores, armazenadores, exportadores bem como processadores desses produtos (MARTINS e MARTINELLI, 2010). Assim, a previsão de preços dos produtos agrícolas tem sido cada vez mais objeto de interesse dos profissionais do mercado e dos acadêmicos uma vez que através dela é possível reduzir a incerteza no processo de decisão por parte daqueles que negociam no mercado futuro.

Schwager (1995) explica que através da previsão é possível auxiliar os envolvidos nos mercados de commodities (mercadoria padronizada de baixo valor agregado) agrícolas tanto como os *hedgers*, que lidam com transações físicas de produtos, buscando proteção de futuras oscilações de preços. Logo, os participantes do mercado que lidam com os produtos agrícolas necessitam de informações sobre os preços que são fundamentais para a rentabilidade de suas atividades.

Pesquisas realizadas sobre previsão de preços de commodities agrícolas (Bressan, 2004; Lima, Góis e Ulises, 2007; Sobreiro *et. al.*, 2008; Lima *et al*, 2010; Ferreira, Borenstein e Fischmann, 2011; Miranda, Coronel e Vieira, 2013; Tibulo e Carli, 2014) alternam-se entre análise empírica que ora fazem uso de modelos clássicos e de suavização exponencial que decompõem a série temporal ora utilizam o modelo RNA que não realiza tal decomposição.

No estudo realizado por Lima *et al.* (2010) sobre previsão para a commodity soja os autores argumentam sobre a importância da combinação de modelos de decomposição de séries temporais com modelos para fins de previsão, com o intuito de melhorar o desempenho das previsões realizadas. A partir disto, percebe-se a possibilidade de realizar a previsão dos preços das commodities agrícolas através do modelo RNA com a decomposição da série temporal realizada pelo modelo AES com o objetivo de promover uma maior discussão sobre a combinação de modelos de decomposição com modelos de previsão e verificar os desempenhos de previsão obtidos. Tal combinação irá ajudar os produtores brasileiros na formulação e na aplicação de políticas direcionadas ao setor agrícola por conta da relevância da previsão dos preços como instrumento de planejamento bem como na análise do comportamento do mercado futuro ao especificar a tendência dos preços das commodities.

O artigo foi organizado da seguinte forma. Na próxima seção é apresentada a fundamentação teórica de modelos utilizados em previsão e sobre estudos com commodities agrícolas no Brasil além do desempenho das previsões obtido nesses estudos. Na seção 3 são descritos a metodologia além dos modelos de previsão utilizados na pesquisa proposta. Na seção 4 os resultados dos testes de normalidade, linearidade, estacionaridade, testes preditivos e resultados empíricos são apresentados. As conclusões e sugestões para próximas pesquisas são expostos na seção 5.

2. Fundamentação Teórica

Na literatura de séries temporais é possível distinguir duas estratégias clássicas de modelagem. A primeira delas é considerada simples e refere-se aos modelos de suavização exponencial enquanto que a segunda caracteriza a metodologia de Box-Jenkins (ESQUIVEL, 2012). Os modelos de suavização exponencial, também definidos como modelos de alisamento ou amortecimento, são técnicas desenvolvidas para um fim específico e que não requerem fundamentação probabilística. Utilizam a ideia de distribuição de pesos no decorrer do período, com o objetivo de considerar ponderações variantes no tempo. Dentre os modelos de suavização exponencial o algoritmo sazonal Holt-Winters (HW) é indicado para séries temporais com padrão de comportamento mais complexo, que além de apresentarem sazonalidade também consideram tendência e ruído (MORETTIN e TOLOI, 2006).

Levando em conta que, através da suavização, um processo de médias móveis é pouco intuitivo para representar o comportamento de uma particular série temporal e, considerando

que a aplicação dos modelos autorregressivos (definido por valores anteriores e adição do ruído) é comum em diferentes áreas do conhecimento, pode-se utilizar os termos autorregressivos e de médias móveis simultaneamente com o objetivo de melhoria. Assim, tal combinação caracteriza o modelo definido pela literatura como *Autoregressive Moving Average Model* (ARMA).

Outra possibilidade é tornar a série temporal estacionária através de um processo de diferenciação, ou seja, tomar diferenças sucessivas da série temporal original. Com isso, é dada a formação do modelo *Autoregressive Integrated Moving Average* (ARIMA). Esse modelo fundamenta-se na construção de métodos ajustados em suas propriedades probabilísticas.

Em algumas situações as séries temporais podem apresentar flutuações periódicas a exemplo dos fenômenos meteorológicos que, quando avaliados trimestralmente costumam apresentar correlações mais elevadas quando são usadas defasagens múltiplas de quatro, conforme as estações do ano (ESQUIVEL, 2012). Com isso, é adequado considerar uma periodicidade estocástica para avaliar o comportamento da série temporal. Portanto, quando o modelo ARIMA leva em consideração essa sazonalidade ele passa a ser conhecido como *Seasonal Autoregressive Integrated Moving Average* (SARIMA).

Além dessas estratégias de modelagem, pesquisas realizadas com commodities (Bressan, 2004; Sobreiro *et. al.*, 2008; Ferreira, Borenstein e Fischmann, 2011; Miranda, Coronel e Vieira, 2013) fazem uso de outra estratégia que não requer parâmetros da série temporal em análise. Trata-se do modelo RNA que através de um captador automático aproxima equações sem que seja necessário deduzi-las. Além de não requerer parâmetros da série o modelo RNA se diferencia dos modelos clássicos e de suavização exponencial de previsão por ser um modelo que opera com algoritmo de aprendizado. Tal algoritmo busca imitar a estrutura de interconexões do cérebro humano, com o intuito de incorporar o padrão de comportamento de uma série temporal de modo a prever, da maneira eficiente, valores futuros dessa série temporal (TURBAN, 1993).

A construção do modelo RNA envolve desde a modelagem adequada da rede neural até as transformações utilizadas para transmitir os dados a rede e aos métodos utilizados para interpretar os resultados obtidos. Esses aspectos são dados pela: modelagem, transformação e interpretação sendo fundamentais na utilização do modelo RNA para realização das previsões dos preços.

Ao investigarem o comportamento dos preços da commodity soja, através do modelo ARIMA-GARCH, Lima *et al.* (2010) com base na análise dos dados da saca de 60 quilos, explicam que os resultados de previsão dos preços foram particularmente satisfatórios. Para a previsão dos preços da commodity milho, Tibulo e Carli (2014) fazem uso dos modelos ARIMA e do algoritmo sazonal HW aplicados à série temporal do preço médio mensal do milho na região do Rio Grande do Sul. Para os autores a análise dos resultados demonstrou que o algoritmo sazonal HW aditivo apresentou melhores resultados para previsões do preço do milho em comparação ao modelo ARIMA.

Para previsão dos preços das commodities boi gordo, café e soja, Bressan (2004) faz uso do modelo clássico ARIMA e do modelo RNA. Os resultados apresentaram ganhos na maioria dos contratos analisados no mercado futuro, indicando o potencial de utilização dos modelos como ferramenta de decisão em negociações com destaque para operações fundamentadas nas previsões realizadas pelo modelo ARIMA. Em relação ao modelo clássico ARMA e o modelo RNA utilizados para previsão de preços do café, Miranda, Coronel e Vieira (2013) ao avaliarem suas potencialidades de previsão concluíram que o modelo RNA demonstrou-se eficaz na previsão do preço do café, uma vez que os preços previstos foram próximos aos observados.

Em pesquisa para avaliar a aplicação do modelo RNA Sobreiro *et al.* (2008) utilizam os preços da commodity açúcar. Os resultados obtidos demonstram que a aplicação obteve uma aproximação significativa comparada às cotações reais, o que para os autores evidencia a importância do modelo RNA como alternativo na estimação dos preços. Para Ferreira, Borenstein e Fischmann (2011) as previsões dos preços das quatro commodities específicas: soja, boi gordo, milho e trigo foram realizadas utilizando-se o modelo RNA. Os dados obtidos na pesquisa evidenciam para os autores a possibilidade do uso de redes neurais como uma estratégia de precificação.

Essas pesquisas apresentam diferentes metodologias para identificar padrões de comportamento da série temporal em análise e realizar previsões. Ao analisarem o comportamento da série temporal em estudo os modelos clássicos fazem uso da metodologia de Box-Jenkins ou da ideia de distribuição de pesos no decorrer do período, com o objetivo de considerar ponderações variantes no tempo.

As metodologias descritas nas pesquisas consistem no ajuste sistemático dos modelos aos valores de uma série temporal de forma que os resíduos obtidos na decomposição representem o ruído. Como o modelo RNA não faz a decomposição da série temporal para

realização do processo de previsão, isto acaba justificando a utilização de algum modelo de decomposição. Para a pesquisa proposta o modelo AES foi utilizado para decomposição uma vez que o mesmo captura a estrutura da série temporal.

3. Metodologia e Dados

A metodologia utilizada fundamenta-se na análise e na construção de modelo de previsão de séries temporais. A pesquisa pode ser classificada como descritiva, quantitativa, bibliográfica e experimental. Descritiva quanto a seus objetivos, voltados à descrição e à análise dos dados para amostras das séries temporais dos preços das commodities açúcar e soja. Quantitativa devido à abordagem do problema, uma vez que os recursos utilizados para análise dos dados são baseados em métodos estatísticos. Também pode ser classificada como bibliográfica, pois realiza um levantamento de trabalhos que abordam previsão de preços para o mercado futuro de commodities agrícolas. Finalmente, é experimental pela utilização de métodos estatísticos de precisão para obtenção de uma resposta ao problema de pesquisa apresentado.

Para aplicação da combinação proposta entre os modelos a escolha das commodities açúcar e soja não está somente no crescimento do seu volume de exportações nos últimos 5 anos, conforme dados disponibilizados pelo Ministério da Agricultura, Pecuária e Abastecimento (MAPA) mas, por desempenhar papel significativo na pauta das exportações brasileiras bem como participação do volume de contratos negociados no mercado futuro. Assim, em 2014 do total de exportações 42% correspondem aos produtos agrícolas. Relatórios do MAPA indicam que as commodities açúcar e soja responderam por 13,72% e 34,97%, respectivamente, desse total.

Os preços que compõem as séries temporais foram obtidos no banco de dados da Escola Superior de Agricultura Luiz de Queiroz (ESALQ) e correspondem ao período de 13 de janeiro de 2012 a 20 de dezembro de 2013, cuja periodicidade é semanal totalizando 96 semanas. As séries temporais na pesquisa são utilizadas na forma logarítmica.

Para realização da decomposição pelo modelo AES e das previsões dos modelos utilizados na pesquisa usou-se o pacote estatístico R. O mesmo pacote foi útil para aplicação dos testes estatísticos para a normalidade, conforme Anderson-Darling (AD) e Shapiro-Wilk (SW). Também foi útil nos testes de McLeod e Li (1983) e Tsay (1986) para linearidade e de

Dickey-Fuller Aumentado (DFA) e Kwiatkowski-Phillips-Schmidt-Shin (KPSS) para a estacionariedade das séries temporais.

Como as previsões podem apresentar erros, independente do modelo adotado, costuma-se avaliar o resultado das previsões ao comparar os valores obtidos com os valores da série temporal original e, determinar o seu desempenho através de alguma medida. Então, na pesquisa, as previsões serão confrontadas com a 12^a semana subsequente a semana final da amostra. Para isso, a avaliação de desempenho faz uso da medida Erro Quadrático Médio (EQM) definida por:

$$EQM = \frac{1}{h} \sum_{N+1}^{N+h} (Y_j - \hat{Y}_j)^2 \quad (1)$$

com Y_j representando o valor da série original, \hat{Y}_j o valor da previsão e h a quantidade de observações previstas e reservadas para avaliação. Além dessa medida, a pesquisa faz uso da metodologia proposta por Goyal e Welch (2003), dada pela diferença entre os erros de previsão ao quadrado acumulados do modelo de melhor desempenho e os erros de previsão ao quadrado acumulados do modelo de melhor desempenho subsequente, considerando o Erro Quadrático Acumulado (EQA) dado por:

$$EQA = \sum_{N+1}^{N+h} (Y_j - \hat{Y}_j)^2 \quad (2)$$

Então, sempre que esta diferença for positiva, o modelo de melhor desempenho subsequente supera o de melhor desempenho.

Ao considerar duas previsões de uma série temporal Y_t , e definindo e_{it} e e_{jt} como os respectivos erros de previsão, uma análise das perdas associadas a cada uma dessas previsões deve ser feita através do teste estatístico proposto por Diebold e Mariano (1995), definido nesta pesquisa como teste DM, que faz uso de uma função perda para medir o erro de previsão, ou seja, a perda é calculada a partir de valores realizados e previstos da variável em questão. Com isso, o teste verifica se a perda diferencial não é significativa entre as previsões realizadas.

3.1 Modelo utilizados na pesquisa

3.1.1 Modelo AES

O modelo incorpora os elementos de análise de séries temporais clássicas, estatística multivariada, geometria multivariada, sistemas dinâmicos e processamento de sinais (GOLYANDINA, NEKRUTKIN e ZHIGLJAVSKY, 2001). Em sua apresentação consiste em dois estágios complementares: decomposição e reconstrução em que ambos incluem dois passos separados.

3.1.1.1 Estágio da decomposição

3.1.1.1.1 Passo incorporação

Podemos ser considerado como um mapeamento que transfere uma série temporal $Y_t = (y_1, \dots, y_N)$, considerando $t = 1, \dots, N$ para a série multidimensional X_1, \dots, X_K com vetores $X_i = (y_i, \dots, y_{i+L-1})^T \in R^L$, onde $K = N - L + 1$ e os vetores X_i são definidos como vetores defasados. O único parâmetro da incorporação é o comprimento da janela L , um número inteiro dado por $L \leq N/2$ para atender a condição ótima. O resultado do passo incorporação é a definição da matriz trajetória $X = [X_1, \dots, X_K] = (x_{ij})_{i,j=1}^{L,K}$, de forma que a mesma é uma matriz Hankel por apresentar suas entradas constantes ao longo das diagonais paralelas à diagonal secundária.

3.1.1.1.2 Passo decomposição em valor singular (DVS)

Neste passo a decomposição da matriz trajetória é obtida pela soma de matrizes elementares. Assim, denota-se por $\lambda_1, \dots, \lambda_L$ os autovalores de XX^T em ordem decrescente de magnitude ($\lambda_1 \geq \dots \geq \lambda_L \geq 0$) e por U_1, \dots, U_L os autovetores ortogonais. Ao estabelecer que $V_i = X^T U_i / \sqrt{\lambda_i}$, a DVS da matriz trajetória pode ser escrita como:

$$X = E_1 + \dots + E_D \quad (3)$$

onde $E_i = \sqrt{\lambda_i} U_i V_i^T$ representa uma matriz de posto unitário ou comumente matriz elementar, o conjunto $\sqrt{\lambda_i}, U_i, V_i$ como autotripla e D o posto da matriz que corresponde ao número de autovalores não nulos.

3.1.1.2 Estágio da reconstrução

3.1.1.2.1 Passo agrupamento

Corresponde em dividir as matrizes elementares em grupos somando-as dentro de cada grupo (sinal e ruído). Ao deixar que $I = \{i_1, \dots, i_p\}$ seja um grupo de índices i_1, \dots, i_p , então, a matriz E_I que corresponde ao grupo I é definida por $E_I = E_{i_1}, \dots, E_{i_p}$. Assim, o desdobramento do conjunto de índices $J = \{1, \dots, D\}$ em subconjuntos disjuntos I_1, \dots, I_M corresponde a representação:

$$X = E_{I_1} + \dots + E_{I_M} \quad (4)$$

onde E_{I_1}, \dots, E_{I_M} são definidas como matrizes resultantes.

Em (4) tem-se uma nova decomposição de matrizes que é denominada como decomposição agrupada.

3.1.1.2.1.1 Separabilidade

Considerando que a série temporal Y_t pode ser representada pela soma de duas subséries $Y_t = Y_t^{(1)} + Y_t^{(2)}$ representando sinal e ruído, respectivamente, a separabilidade delas implica que as componentes obtidas pela DVS da matriz trajetória X podem ser agrupadas 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 $Y_t^{(2)}$ (GOLYANDINA, NEKRUTKIN e ZHIGLJAVSKY, 2001).

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) não ocorre separabilidade exata, mas, tão somente separabilidade aproximada. A qualidade da separabilidade é avaliada pela medida denominada correlação ponderada ρ_{ij}^w 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 expressão:

$$\rho_{12}^w = \frac{\langle Y_t^{(1)}, Y_t^{(2)} \rangle_w}{\|Y_t^{(1)}\|_w \|Y_t^{(2)}\|_w} \quad (5)$$

com $\|Y_t^{(i)}\|_w = \sqrt{\langle Y_t^{(i)}, Y_t^{(i)} \rangle_w}$ representando a norma da i -ésima subsérie e $\langle Y_t^{(i)}, Y_t^{(j)} \rangle_w$ o produto interno entre um par de subséries, considerando que o ponderador w é definido por $w = \min\{k, L, N - k\}$, com k representando a fração entre a periodicidade da série e frequência dos dados. Por exemplo, se a série temporal apresenta uma periodicidade anual e a frequência dos dados da série é trimestral k assume o valor 4.

A expressão (5) pode apresentar valores entre 0 e 1. Desta maneira, ao considerar $\rho_{12}^w = 0$ tem-se que as subséries $Y_t^{(1)}$ e $Y_t^{(2)}$ são separáveis e de outra forma, quando $\rho_{12}^w = 1$ isto significa que as subséries $Y_t^{(1)}$ e $Y_t^{(2)}$ não são tão bem separáveis, ou seja, devem ser reunidas num mesmo grupo.

Existem algumas ferramentas usadas para auxiliar na separação das componentes da série temporal. Entre elas, o gráfico dos valores singulares $\sqrt{\lambda_i}$. A ferramenta baseia-se em dispor em ordem decrescente de magnitude os valores singulares. Assim, na análise o valor singular de maior magnitude representa a componente tendência e os valores singulares de mesma magnitude a periodicidade, enquanto que, os valores singulares lentamente decrescentes o ruído. Dessa maneira, as componentes são reunidas em seus respectivos grupos.

3.1.1.2.2 Passo média diagonal

$$\tilde{y}_t = \begin{cases} \frac{1}{u-1} \sum_{l=1}^{u-1} y_{l,u-l} & 2 \leq u \leq L-1 \\ \frac{1}{L} \sum_{l=1}^L y_{l,u-l} & L \leq u \leq K+1 \\ \frac{1}{K+L-u+1} \sum_{l=u-K}^L y_{l,u-l} & K+2 \leq u \leq K+L \end{cases} \quad (6)$$

em que u representa a soma dos índices da matriz X de dimensão $(L \times K)$.

Considerando a matriz X de dimensão $(L \times K)$, x_{ij} representando um elemento de tal matriz e cada termo resultante da média diagonal, definido por \tilde{y}_t , com $t = 1, \dots, N$ é obtido conforme dado acima.

Assim, ao aplicar a média diagonal a todos os componentes na expressão (4) obtém-se a expansão dada por $X = \tilde{E}_{I_1} + \dots + \tilde{E}_{I_M}$ com $\tilde{E}_{I_1} \dots \tilde{E}_{I_M}$ representando as matrizes na forma de

uma matriz de Hankel. O agrupamento adequado, responsável por uma decomposição em que a expansão apresenta matrizes de Hankel, leva a uma separabilidade aproximada.

3.1.2 Algoritmo sazonal HW

A incorporação da sazonalidade no algoritmo sazonal HW pode ser realizada através de dois enfoques distintos, dependentes do padrão de sazonalidade identificado na série: sazonalidade multiplicativa e aditiva. Ao considerar a sazonalidade multiplicativa, Morettin e Toloí (2006) explicam que série temporal pode ser definida por:

$$Y_t = N_t S_t + m_t + \varepsilon_t \quad (7)$$

com N_t o nível da série, S_t o fator sazonal, m_t o componente de tendência, ε_t o erro aleatório no instante t e $t = 1, \dots, N$.

A forma de recorrência para o enfoque multiplicativo, nesta pesquisa, é dado por HW_m , com o fator de sazonalidade multiplicativo representado pelas equações que envolvem as três constantes de suavização, α , β e γ , conforme:

$$\begin{cases} N_t = \alpha \frac{y_t}{S_{t-s}} + (1-\alpha)(N_{t-1} + m_{t-1}) \\ m_t = \beta(N_t - N_{t-1}) + (1-\beta)m_{t-1} \\ S_t = \gamma \frac{y_t}{N_t} + (1-\gamma)S_{t-s} \end{cases} \quad (8)$$

em que $0 < \alpha < 1$, $0 < \beta < 1$ e $0 < \gamma < 1$ são as condições das constantes de suavização do modelo e s representa o número de observações.

As previsões para os valores futuros levam em conta os passos à frente h , assim, em cada equação o fator sazonal considera o período correspondente, de acordo com as equações abaixo:

$$\left\{ \begin{array}{l} \hat{Y}_t(h) = (N_t + hm_t)S_{t+h-s}; \quad h = 1, 2, 3, \dots, s \\ \hat{Y}_t(h) = (N_t + hm_t)S_{t+h-2s}; \quad h = s+1, s+2, s+3, \dots, 2s \\ \hat{Y}_t(h) = (N_t + hm_t)S_{t+h-3s}; \quad h = 2s+1, 2s+2, 2s+3, \dots, 3s \\ \text{M} \end{array} \right. \quad (9)$$

Para o enfoque sazonalidade multiplicativa a correção dos erros e_t é dada:

$$\left\{ \begin{array}{l} N_t = N_{t-1} + m_{t-1} + \alpha \frac{e_t}{F_{t-s}} \\ m_t = m_{t-1} + \alpha\beta \frac{e_t}{F_{t-s}} \\ S_t = S_{t-s} + \gamma(1-\alpha) \frac{e_t}{N_t} \end{array} \right. \quad (10)$$

O outro enfoque do método, dado nesta pesquisa por HW_a, é aplicado quando a série apresenta sazonalidade da forma aditiva. Assim, para Morettin e Tolo (2006) ao tomar por base o fator sazonal aditivo a série temporal fica representada pela soma de todos os componentes conforme:

$$Y_t = N_t + m_t + S_t + \varepsilon_t \quad (11)$$

Na sazonalidade aditiva a forma de recorrência é dada pelas equações:

$$\left\{ \begin{array}{l} N_t = \alpha(Y_t - S_{t-s}) + (1-\alpha)(N_{t-1} + m_{t-1}) \\ m_t = \beta(N_t - N_{t-1}) + (1-\beta)m_{t-1} \\ S_t = \gamma(Y_t - N_t) + (1-\gamma)S_{t-s} \end{array} \right. \quad (12)$$

com as mesmas condições das constantes de suavização do modelo para o enfoque multiplicativo bem como s representando o número de observações.

Os valores futuros são previstos através das equações:

$$\left\{ \begin{array}{l} \hat{Y}_t(h) = (N_t + hm_t) + S_{t+h-s}; \quad h = 1, 2, 3, \dots, s \\ \hat{Y}_t(h) = (N_t + hm_t) + S_{t+h-2s}; \quad h = s+1, s+2, s+3, \dots, 2s \\ \hat{Y}_t(h) = (N_t + hm_t) + S_{t+h-3s}; \quad h = 2s+1, 2s+2, 2s+3, \dots, 3s \\ \text{M} \end{array} \right. \quad (13)$$

O procedimento de correção dos erros para este tipo de sazonalidade passa a ser dado por:

$$\begin{cases} N_t = N_{t-1} + m_{t-1} + \alpha e_t \\ m_t = m_{t-1} + \alpha \beta e_t \\ S_t = S_{t-s} + \gamma(1-\alpha)e_t \end{cases} \quad (14)$$

3.1.3 Modelo SARIMA

Em algumas situações é importante considerar a sazonalidade estocástica para explicar o comportamento sazonal da série temporal. Desta maneira, a recomendação é que uma das variações do modelo ARIMA seja utilizado. Trata-se do modelo ARIMA sazonal multiplicativo ou modelo SARIMA. Para Box e Jenkins (1976) o modelo geral representado por ARIMA $(p,d,q) \times (P,D,Q)$ pode ser definido:

$$\phi(B)\Phi_P(B^s)[\Delta^d \Delta_s^D - \mu]Y_t = \theta(B)\Theta_Q(B^s)\varepsilon_t \quad (15)$$

com $\phi(B)$ o operador autoregressivo, Φ_P o polinômio estacionário autoregressivo sazonal de ordem P , Δ o operador diferença, μ o valor esperado da série, $\theta(B)$ o operador de médias móveis, Θ_Q o polinômio invertível de médias móveis sazonal de ordem Q e ε_t um erro aleatório.

O polinômio estacionário autoregressivo sazonal de ordem P é dado por:

$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps} \quad (16)$$

O polinômio invertível de médias móveis sazonal de ordem Q é dado por:

$$\Theta_Q(B^s) = 1 - \theta_1 B^s - \theta_2 B^{2s} - \dots - \theta_Q B^{Qs} \quad (17)$$

O operador diferença sazonal de ordem D representado por:

$$\Delta_s^D = (1 - B^s)^D Y_t \quad (18)$$

em que de modo geral, a primeira diferenciação sazonal $\Delta_s^D = (1 - B^s)^D Y_t = Y_t - Y_{t-s}$ consegue excluir a sazonalidade da série temporal (ESQUIVEL, 2012).

3.1.4 Modelo RNA

O modelo é adaptável às séries temporais e se diferencia dos modelos clássicos de previsão por ser um modelo não paramétrico e por envolver algoritmos de aprendizado

(LIMA *ET. AL.*, 2010). De forma simplificada, uma rede neural é uma estrutura computacional baseada num processo biológico inspirado na arquitetura do cérebro humano.

Conforme descrito por Pasquotto (2010) cada neurônio artificial funciona como uma unidade com autonomia cujo objetivo é o de converter um sinal de entrada em outro sinal de saída. Como os neurônios estão atuando em rede a intensidade desses sinais é amplificada ou amortecida de acordo com os parâmetros que são atribuídos às sinapses, também definidos por pesos sinápticos ou apenas pesos.

As características do modelo RNA podem ser resumidas por suas atividades, que são: i) processar a informação localmente; ii) refletir a força de cada conexão através dos pesos das conexões entre os neurônios; iii) extrair conhecimento dos dados disponíveis em um processo de aprendizagem; iv) processar em paralelo ao contrário de ser somente sequencial e v) capturar a essência de um processo podendo reproduzi-lo de maneira confiável (Pasquotto, 2010).

3.1.4.1 O modelo de neurônio artificial

Os neurônios artificiais se agrupam em três tipos de camadas: a camada de entrada, a camada intermediária ou oculta e a camada de saída. Para Haykin (2001), os neurônios entre diferentes camadas são conectados por sinapses que, por sua vez, estão associadas a pesos ou importâncias relativas de cada neurônio de uma camada com o neurônio de uma camada subsequente. O modelo de neurônio artificial representado pela Figura 1 é dado por diversos elementos.

Os elementos do neurônio artificial descritos na Figura 1 são representados por: m que indica o número de sinais de entrada do neurônio; x_j o j -ésimo sinal de entrada do neurônio; w_{gj} o peso associado com o j -ésimo sinal de entrada no neurônio g ; b o limiar de cada neurônio também denominado de bias; v_g uma combinação ponderada de sinais de entrada e do bias do g -ésimo neurônio e $\varphi(.)$ como função de ativação do g -ésimo neurônio.

O bias apresenta o efeito de aumentar ou reduzir a entrada da função de ativação conforme seu sinal muda de positivo para negativo. Com um pequeno ajuste, Pasquotto (2010) explica que é possível realizar a substituição do bias b_g para uma entrada fixa $x_0 = 1$ de forma que o bias torna-se um novo peso sináptico $w_{g0} = b_g$.

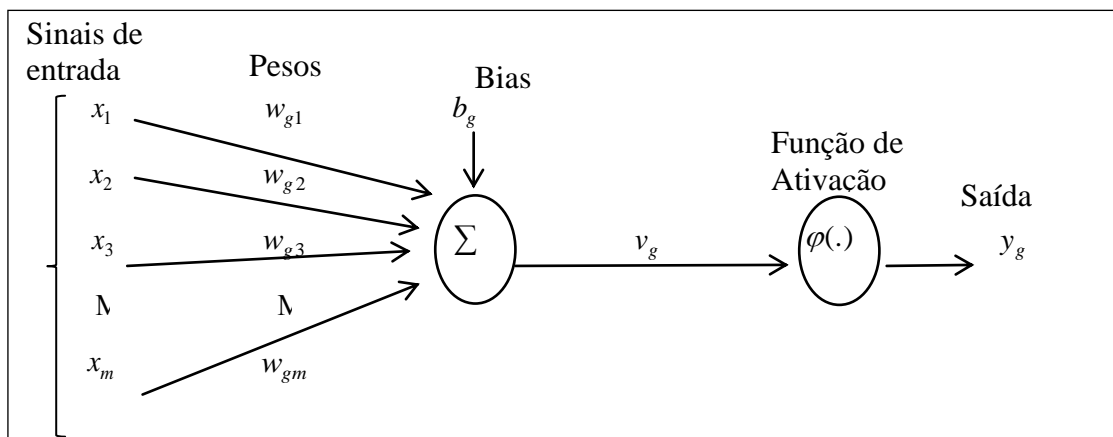


Figura 1 : Elementos do neurônio artificial

Fonte: Elaborada pelo autor considerando Haykin (2001)

Com isso, pode-se descrever matematicamente o neurônio g por:

$$v_g = \sum_{j=0}^m x_j w_{gj} \quad (19)$$

e

$$y_g = \varphi(v_g(t)) \quad (20)$$

com v_g definido como campo local induzido ou potencial de ativação e a função de ativação definindo a saída $y_g(t)$ do g -ésimo neurônio no instante t .

Dentre as funções mais utilizadas na literatura Pasquotto (2010) cita: i) função de limiar, que é uma função descontínua e binária; ii) função sigmoide, que é uma função contínua em forma de S, variando de 0 a 1 e iii) função tangente hiperbólica, que é uma função contínua e derivável em todos os seus pontos, o que a torna uma função conveniente para o algoritmo de treinamento.

3.1.4.2 Arquitetura das redes neurais

A arquitetura para o modelo RNA sofre mudanças de acordo com sua finalidade. A maneira como os neurônios estão distribuídos em rede está relacionada com o algoritmo de aprendizagem utilizado. A classificação dada na literatura considera a maneira como o processamento ocorre na rede neural bem como os neurônios estão dispostos em camadas. Para Haykin (2001) as classificações expostas abaixo constituem algumas maneiras dentre muitas possibilidades de apresentar a arquitetura da rede neural.

3.1.4.2.1 Camada única

Para este tipo de arquitetura os neurônios são dispostos em paralelo em uma camada única. Nas entradas deste tipo de rede neural existem nós que não são neurônios, não existindo computação do sinal, de forma que o processamento tem prosseguimento apenas numa camada de onde emergem as saídas da rede.

3.1.4.2.2 Multicamadas

Nesta arquitetura existem camadas posicionadas entre os nós de entrada e a camada responsável por gerar as saídas da rede. As camadas escondidas, também definidas como camadas ocultas ou intermediárias propagam os sinais até que cheguem na saída da rede neural.

3.1.4.2.3 Alimentadas à frente

Este tipo de arquitetura de rede tem o processamento num único sentido, ou seja, da entrada até a saída, com a ausência de realimentação. Considerando a representação no sentido horizontal, este tipo de processamento acontece com partida das entradas situadas à esquerda, movendo-se para as saídas localizadas à direita e, fundamentalmente sem retornos.

3.1.4.2.4 Recorrentes

Existe neste tipo de arquitetura pelo menos um laço de realimentação. Com isso, a saída de pelo menos um dos neurônios é reintroduzida em algum ponto anterior da rede configurando recorrência de processamento. Quando a realimentação ocorre no próprio neurônio que a originou este tipo de operação é definido como auto realimentação (HAYKIN, 2001).

3.1.4.3 Tipos de treinamento

O treinamento para RNA consiste em ajustar os parâmetros da rede de maneira interativa requerendo uma sequência de eventos, conforme: i) estímulo da rede neural pelo ambiente; ii) modificações em seus pesos por conta dos estímulos dados e iii) resposta da rede

de uma forma diferente ao ambiente em razão das modificações. Para as redes neurais existem dois padrões de aprendizado (PASQUOTTO, 2010).

3.1.4.3.1 Aprendizado supervisionado

Este tipo de aprendizado funciona indicando na saída da rede a resposta correta para cada situação. Existe um conjunto de dados de entrada apresentados à rede neural como exemplos que geram uma saída de rede que é comparada com a saída esperada, obtendo-se desta forma o erro correspondente.

Considerando o neurônio g na saída de uma rede no instante t , o erro correspondente e_g será definido por:

$$e_g = d_g(t) - y_g(t) \quad (21)$$

com $d_g(t)$ como sinal de resposta desejada do neurônio g no instante t e $y_g(t)$ como sinal de saída do neurônio g no instante t . O erro é usado como parâmetro iterativo de ajuste de pesos cuja intenção é a de reduzir progressivamente o erro até um valor mínimo aceitável. O algoritmo de retropropagação é muito utilizado para o aprendizado supervisionado (LIMA ET. AL., 2010).

3.1.4.3.1.1 Algoritmo básico de retropropagação

O algoritmo de retropropagação percorre a função de erro na saída da rede em busca de um ponto mínimo. Os pesos sinápticos podem ficar alterados após serem percorridas duas etapas: i) propagação à frente e ii) retropropagação. Na primeira etapa o sinal é propagado ao longo da rede iniciando pela primeira camada até gerar o erro na última camada. Na segunda etapa, o erro é corrigido camada a camada, alterando os pesos no sentido inverso.

No algoritmo de retropropagação os pares de entrada-saída são apresentados, cada uma deles para a rede neural, existindo duas maneiras de se aplicar a correção dos erros. Na primeira, definida como alteração incremental, a alteração nos pesos é realizada sempre que um novo par entrada-saída é apresentado à rede gerando um erro que é corrigido individualmente logo após cada par ser submetido à rede.

Ao contrário da primeira maneira, na alteração por lote, todos os n pares são apresentados à rede neural gerando-se um erro correspondente a todo lote e, só depois disto, é

realizada a atualização dos pesos, podendo envolver a apresentação repetitiva do mesmo conjunto de pares por muitas vezes.

3.1.4.3.2 Aprendizado não supervisionado

Tem como característica a ausência na saída da rede neural de uma resposta correta. Ou seja, como não existem vetores de respostas desejadas não existem comparações para fornecer erros. Nesta situação de aprendizado são fornecidas a rede neural condições para realização de uma medida independentemente da tarefa que deve ser aprendida e os parâmetros livres da rede são otimizados em relação a esta medida. Para este tipo de aprendizado, Haykin (2001) explica que existem duas maneiras de conduzi-lo: por reforço e de forma auto-organizada.

4. Resultados e Discussões

4.1 Testes estatísticos aplicados sobre os dados

As séries temporais preços das commodities agrícolas são identificadas na pesquisa conforme: AÇUC e SOJA. Para testar se a amostra de dados tem origem de uma população com uma distribuição específica os testes AD e SW são usados. Na pesquisa os dois testes permitem uma visão abrangente dos resultados. Como pode ser visto, a partir dos resultados dispostos na Tabela 1, para um nível de 1% de significância, as séries temporais AÇUC e SOJA não estão distribuídas normalmente.

Tabela 1 : Teste de normalidade AD, SW e *p*-valor

	AÇUC	SOJA
Número de Observações	96	96
Shapiro-Wilk	0.96	0.95
<i>p</i> -valor	0.00	0.00
Anderson-Darling	1.05	1.87
<i>p</i> -valor	0.00	0.00

Fonte: Elaborado pelos autores

Os testes de não linearidade de Tsay (1986) e de McLeod e Li (1983) também são aplicados na série temporal. Para o primeiro teste, antes de sua aplicação, os dados são **Custos e @gronegocio on line** - v. 13, n. 4, Out/Dez - 2017. ISSN 1808-2882
www.custoseagronegocioonline.com.br

filtrados por um modelo AR. A Tabela 2 apresenta os resultados para os dados da amostra com base em *lags* de 5 e 10 semanas. Para um nível de 1% de significância as séries temporais podem ser consideradas lineares.

Tabela 2 : p -valor para testes Tsay e McLeod

	<i>Lags</i>	AÇUC	SOJA
Número de Observações		96	96
Tsay <i>p</i> -valor	5	0.29	0.60
Tsay <i>p</i> -valor	10	0.77	0.57
McLeod <i>p</i> -valor	5	0.85	0.80
McLeod <i>p</i> -valor	10	0.90	0.90

Fonte: Elaborado pelos autores

Por fim, foram realizados os testes DFA e KPSS para avaliar a estacionariedade das séries temporais. A Tabela 3 apresenta os resultados dos dois testes. Para o primeiro teste a hipótese nula é que as séries temporais possuem raiz unitária e, portanto, não são estacionárias. Em relação ao segundo teste a hipótese nula é que as séries temporais não possuem raiz unitária e, portanto, são estacionárias. Desta forma, os testes confirmam à sua maneira, que as séries temporais AÇUC e SOJA não são estacionárias. Em resumo, as séries temporais utilizadas na pesquisa não estão distribuídas normalmente, são estacionárias e lineares.

Tabela 3 : Teste DFA e KPSS para série simulada e conjuntos de séries reais

Séries temporais	Valor		Valor	
	DFA	Crítico 1%	KPSS	Crítico 1%
AÇUC	-2,39	-4,05	0,26	0,21
SOJA	-2,28	-4,05	0,27	0,21

Fonte: Elaborado pelos autores

A Figura 2 apresenta, respectivamente, o comportamento das séries temporais das commodities AÇUC e SOJA no período de 13 de janeiro de 2012 a 20 de dezembro de 2013. Com base nela é possível perceber o comportamento das séries.

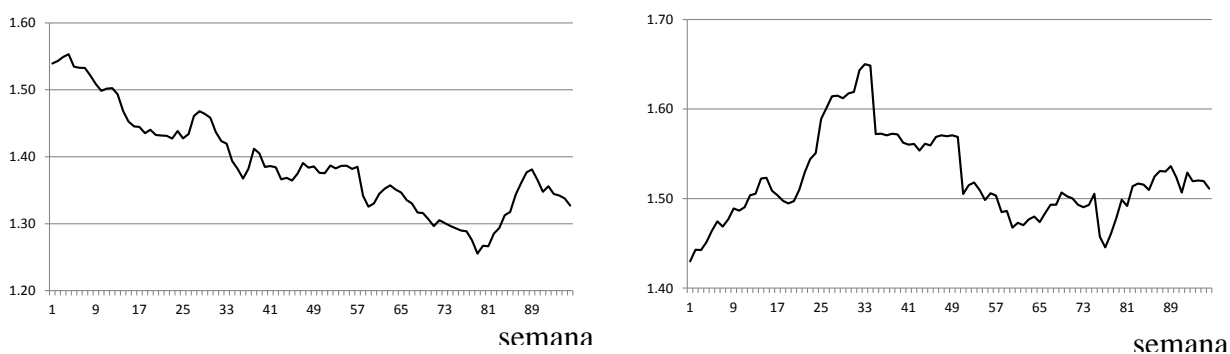


Figura 2 : Comportamento semanal dos preços das commodities açúcar e soja

Fonte: Elaborado pelos autores

4.2 Aplicações empíricas

Para definição dos resultados os modelo HWa e HWm utilizados na pesquisa, tem a escolha dos valores de parâmetros de alisamento definida pelo pacote R com base na minimização da soma dos quadrados dos erros de previsão um passo à frente para posterior aplicação da rotina de previsão utilizada nesta pesquisa. Em relação ao modelo SARIMA, as ordens foram definidas com base na minimização dos valores para *Akaike Information Criterion* (AIC) através de rotina própria do pacote R. No modelo AES foi utilizado como comprimento da janela 48 semanas uma vez que este valor define o tamanho ótimo de janela e o agrupamento feito com base no gráfico dos valores singulares. Por fim, para o modelo RNA o pacote R faz uso de redes neurais com 1 camada composta de 7 neurônios, admitindo um número máximo de iterações igual a 100.000 com aprendizado supervisionado e aplicação do algoritmo de retropropagação.

Além dos parâmetros utilizados pelos modelos AES e SARIMA, observa-se na Tabela 4 com base na medida EQM não existir diferenças significativas para as previsões realizadas entre o algoritmo sazonal HW com sazonalidade aditiva e multiplicativa. As previsões obtidas pela combinação proposta (AES-RNA) quando comparada com a previsão obtida pelo algoritmo sazonal HWa e HWm, pelo modelo SARIMA e o modelo RNA correspondem ao melhor desempenho (MD na Tabela 4), por conta dos menores valores para a medida de erro descrita em (1).

Tabela 4 : Desempenho de previsão pelo EQM

Série	Parâmetros			EQM					MD
	L	(p,d,q)(P,D,Q)	h	RNA	AESM-RNA	HWa	HWm	SARIMA	
AÇUC	16	(1,1,0)(1,0,0)	12	9.8E-03	9.2E-04	1.2E-03	1.3E-03	1.3E-03	AESM-RNA
SOJA	48	(0,1,0)(1,0,0)	12	1.6E-02	2.0E-04	1.6E-03	1.7E-03	4.9E-03	AESM-RNA

Fonte: Elaborado pelos autores

Tabela 5 : Desempenho de previsão pela Diferença do EQA

Série	h	EQM		Diferença
		AESM-RNA	HWa	
AÇUC	12	1.1E-02	1.4E-02	-3.4E-03
SOJA	12	1.9E-02	2.4E-02	-5.3E-03

Fonte: Elaborado pelos autores

Na Tabela 5, que avalia o desempenho dos modelos através da diferença entre os erros de previsão ao quadrado acumulados do modelo de melhor desempenho (coluna sombreada da Tabela 5) e os erros de previsão ao quadrado acumulados do modelo de melhor desempenho subsequente, conclui-se que o modelo de melhor desempenho de previsão supera o de melhor desempenho subsequente uma vez que as diferenças apresentadas são negativas.

Em seguida, com o objetivo de avaliar se a diferença entre os EQMs do modelo de melhor desempenho com o modelo de melhor desempenho subsequente é estatisticamente significativa, aplicou-se o teste DM. Os resultados expostos na Tabela 6 indicam que, para os modelos comparados, a hipótese nula de que a diferença entre as medidas de erro é zero pode ser rejeitada ao nível estatístico 1% de significância para as séries temporais.

Tabela 6 : Teste Diebold-Mariano e modelos comparados

Série	h	DM	p-value	Modelos
AÇUC	12	5.83	0.00	AESM-RNA, HWa
SOJA	12	5.75	0.00	AESM-RNA, HWa

Fonte: Elaborado pelos autores

Pode-se concluir, com base no teste estatístico DM, através do desempenho das previsões realizadas para o passo à frente h (12 semanas), que o modelo proposto AESM-RNA agrega evidências favoráveis a sua aplicação no mercado de commodities agrícola.

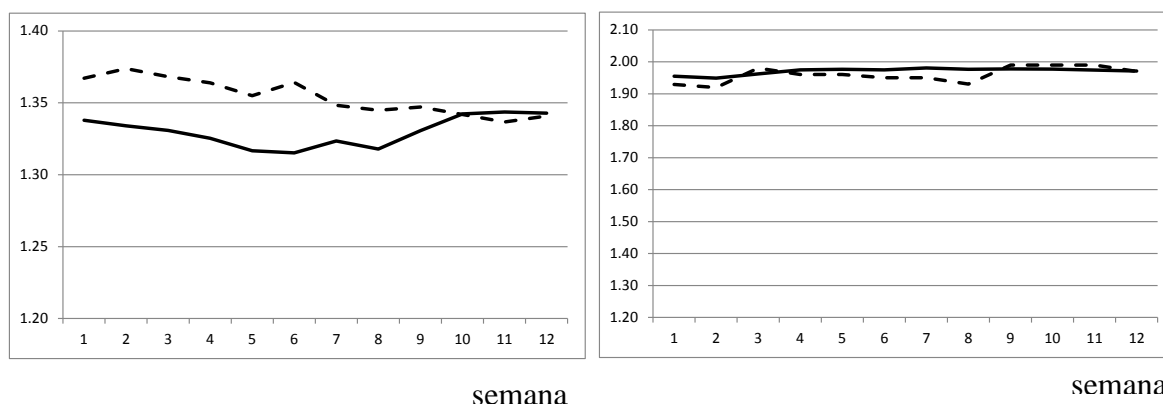


Figura 3 : Previsão para os preços das commodities destinadas à exportação

Fonte: Elaborado pelos autores

Na Figura 3 são apresentados, através do gráfico dos preços das commodities AÇUCAR e SOJA no período semanal, o comportamento das séries temporais originais (linhas sólidas) e séries temporais previstas (linhas tracejadas) obtidas pela combinação proposta AES-RNA. O gráfico foi elaborado no período de 27 de dezembro de 2012 a 14 de março de 2014, fora da amostra. Observa-se que o modelo AES-RNA, no período em análise, conseguiu detectar a tendência dos preços das commodities agrícolas.

5. Conclusões e Sugestões

A análise dos preços das commodities agrícolas, como elemento do mecanismo de troca, reveste-se de singular importância para os participantes do mercado, sejam eles

compradores, vendedores ou estejam à procura de proteção. Assim, as decisões tomadas por estes participantes presumem o conhecimento do comportamento dos preços.

Pelo fato de que algumas pesquisas realizadas sobre previsão de preços de commodities agrícolas serem dadas entre a análise empírica que se alterna entre o uso de modelos clássicos e de suavização exponencial que decompõem a série temporal ou pela aplicação do modelo de redes neurais que não realiza tal decomposição, as pesquisas que fazem uso do modelo de previsão com base em redes neurais acabam não sendo estruturadas para realização de previsões de preços da commodity agrícola com base na decomposição da série temporal.

Como a dinâmica das séries temporais de commodities agrícolas apresentam mudanças no tempo, por conta da sazonalidade, é preciso ter certeza de que o modelo de previsão não seja sensível a essas mudanças. A motivação pelo uso dos modelos AES com RNA visa promover uma maior discussão sobre a combinação de modelos de decomposição com modelos de previsão. Além disso, a pesquisa ao escolher o modelo AES para decomposição das séries temporais busca capturar as estruturas das séries temporais que representam o comportamento abrangente da série temporal.

Os resultados dos testes estatísticos aplicados à série temporal da commodity agrícola soja indicam que a série não é estacionária e é linear. No contexto da pesquisa, os resultados empíricos demonstram superioridade da combinação proposta AESM-RNA, quando comparado com os modelos HWa, HWm, SARIMA e RNA, ao permitir melhor desempenho da previsão. Os resultados obtidos no período fora da amostra, mediante o uso das medidas de erro EQM e EQA além do teste MD das previsões para o passo à frente h (12 semanas) confirmam isso. Ou seja, de modo geral, o modelo proposto AES-RNA superou, em termos de perdas estatísticas, o modelo que apresentou o melhor desempenho subsequente.

A pesquisa ao propor a combinação do modelo AES para decomposição da série temporal com o modelo RNA para previsão agregou evidências favoráveis ao seu uso por conta da eficácia do modelo proposto AES-RNA aplicado no mercado de commodities.

Do ponto de vista prático, os resultados obtidos podem ajudar os profissionais do mercado na formulação e na aplicação de políticas direcionadas ao setor agrícola por conta da relevância da previsão dos preços como instrumento de planejamento bem como na análise do comportamento do mercado futuro ao especificar a tendência dos preços de suas commodities agrícolas.

Para próximas pesquisas sugere-se a utilização de outras bases de dados, a inclusão de outras commodities, a adoção de outros períodos de análise e o uso de outras variáveis que possam aumentar o poder explicativo da combinação proposta AES-RNA.

6. Referências

BOX, G. E. P.; JENKINS, G. M. *Time series analysis forecasting and control*. San Francisco: H. Day, 1976.

BRESSAN, A. A. Tomada de decisão em futuros agropecuários com modelos de previsão de séries temporais. *Revista de Administração Eletrônica*, v. 3, p. 1-20, 2004.

CAMPOS, K. C. Análise da volatilidade de preços de produtos agropecuários no Brasil, *Revista de Economia e Agronegócio*, v. 5, p. 303-328, 2007.

DIEBOLD, F.; MARIANO, R. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, v. 13, p. 253-265, 1995.

ESQUIVEL, R. M. *Análise espectral singular: modelagens de séries temporais através de estudos comparativos usando diferentes estratégias de previsão*. Dissertação (Mestrado), Programa de Pós Graduação em Modelagem Computacional, Faculdade de Tecnologia SENAI CIMATEC, Salvador, BA, Brasil, 2012.

FERREIRA, L.; MOURA, G. L.; BORENSTEIN, D.; FISCHMANN, A. A. Utilização de redes neurais artificiais como estratégia de previsão de preços no contexto de agronegócio. *Revista de Administração e Inovação*, v. 8, p. 6-26, 2011.

GOLYANDINA, N.; NEKRUTKIN, V.; ZHIGLJAVSKY, A. *Analysis of time series structure: SSA and related techniques*. New York: Chapman & Hall/CRC, 2001.

GOYAL, A.; WELCH, I. Predicting the equity premium with dividend ratios. *Management Science*, v. 49, p. 639-654, 2003.

HAYKIN, S. *Redes neurais: princípios e práticas*. 2 ed. Porto Alegre: Bookman, 2001.

HASSANI, H.; MAHMOUDVAND, R. Multivariate singular spectrum analysis: a general view and new vector forecasting approach. *International Journal of Energy and Statistics*, v. 1, p. 55-83, 2013.

LIMA, R.C.; GÓIS, M. R.; ULISES, C. Previsão de preços futuros de Commodities Agrícolas com diferenciações inteira e fracionária, e erros heteroscedásticos. *Revista de Economia e Sociologia Rural*, v. 45, p. 621-644, 2007.

LIMA, F. G.; KIMURA, H.; ASSAF NETO, A.; PERERA, L. C. J. Previsão de preços de commodities com modelos ARIMA-GARCH e redes neurais com ondaletas: velhas tecnologias – novos resultados. *Revista de Administração*, v. 45, p. 188-202, 2010.

MARTINS, T.M.; MARTINELLI, D. P. Ciclos e previsão cíclica dos preços de commodities: um modelo indicador antecedente para commodity açúcar. *Revista de Administração, Contabilidade e Economia*, v. 2, p. 2-12, 2010.

MCLEOD, A. I.; LI, W. K. Diagnostic checking ARMA time series models using squared residual autocorrelations, *Journal of Time Series Analysis*. v. 4, p. 169-176, 1983.

MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO. *Projeções do agronegócio 2013/14 a 2025/25*. Assessoria de Gestão Estratégica. Brasília, 2013.

MIRANDA, A. P.; CORONEL, D. A.; VIEIRA, K. M. Previsão do mercado futuro do café arábica utilizando redes neurais e métodos econométricos. *Revista Estudos do CEPE*, v. 38, p. 66-98, 2013.

MORETTIN, P.A.; TOLOI, C.M.C. *Análise de séries temporais*. São Paulo: Blücher, 2006.

OLIVEIRA, V. A.; AGUIAR, D.R. Determinantes do desempenho dos contratos futuros de commodities agropecuários no Brasil. In: CONGRESSO INTERNACIONAL DE ECONOMIA E GESTÃO DE REDES AGROALIMENTARES. 2003, Ribeirão Preto. *Anais...* Ribeirão Preto, 2003.

ORANJE, M. *Competitividade das frutas brasileiras no comércio internacional*. Dissertação de Mestrado, Universidade Federal de Viçosa, Viçosa, MG, Brasil, 2003.

PASQUOTTO, J. L. D. *Previsão de séries temporais no varejo brasileiro: Uma investigação comparativa da aplicação de redes neurais recorrentes de Elman*. Dissertação (Mestrado). Pós Graduação em Administração, Faculdade de Economia, Administração e Contabilidade de São Paulo, Universidade de São Paulo, São Paulo, 2010.

RIBEIRO, C. O.; SOSNOSKI, A. A. K.; OLIVEIRA, S. M. Um modelo hierárquico para previsão de preços de commodities agrícolas, *Revista Produção On-line*. v. 10, p. 719-733, 2010.

SCHWAGER, J. D. *Fundamental analysis*. New York: John Wiley & Sons, 1995.

SOBREIRO, V. A.; ARAÚJO, P. H. S. L.; MENDONÇA, M. M.; NAGANO, M. S. Uma estimação do valor da commodity de açúcar usando redes neurais artificiais. *Revista P&D em Engenharia de Produção*, p. 36-53, 2008.

TIBULO, C.; CARLI, V. Previsão do preço do milho através de séries temporais, *Scientia Plena*, v. 10, p. 2-10, 2014.

TSAY, R. Non-linearity tests for time series, *Biometrika*, v. 73, p. 461-466, 1986.

TURBAN, E. *Decision Support and Expert Systems: Management Support Systems*. New York: MacMillan, 1993.