GESTÃO, FINANÇAS E CONTABILIDADE

PRICE FORECASTING FOR FUTURE CONTRACTS ON AGRIBUSINESS THROUGH NEURAL NETWORK AND MULTIVARIATE SPECTRAL ANALYSIS

PREVISÃO DE PREÇOS PARA CONTRATOS FUTUROS AGROPECUÁRIOS ATRAVÉS DE REDES NEURAIS E ANÁLISE ESPECTRAL MULTIVARIADA

PRONÓSTICO DE PRECIOS PARA CONTRATOS DE FUTUROS AGRÍCOLAS A TRAVÉS DE REDES NEURONALES Y ANÁLISIS ESPECTRAL MULTIVARIANTE

DOI: 10.18028/2238-5320/rgfc.v6n3p98-124

Carlos Alberto Orge Pinheiro
Mestre em Modelagem Computacional pela Faculdade de Tecnologia Senai Cimatec
Professor Assistente da Universidade do Estado da Bahia (UNEB)
Endereço: Av. Silveira Martins, 2555 - Cabula
41.150-000 – Salvador-BA, Brasil
Email: carlos.orge@terra.com.br

Valter de Senna
Pós-Doutor em Probabilidade e Estatística pela University of Southampton
Professor da Faculdade de Tecnologia Senai Cimatec
Endereço: Av. Orlando Gomes, 1845 - Piatã
41.650-010 – Salvador-BA, Brasil
Email: vsenna@terra.com.br

Alberto Shigueru Matsumoto
Doutor em Administração pela Fundação Getúlio Vargas (FGV)
Professor do Instituto de Educação Superior de Brasília (IESB)
Endereço: SGAN, Quadra 609 - Módulo D - L2 Norte
70.830-404 – Brasilia-DF, Brasil
Email: alberto011052@gmail.com

ABSTRACT
This study aimed to compare the forecasting results from combining the two models, Multivariate Singular Spectrum Analysis (MSSA) and Artificial Neural Network (ANN), with the results obtained from classical forecasting and neural network models for prices of agricultural future contracts traded on BM&FBOVESPA. The forecasting results of the proposed combination, compared with those obtained from classical forecasting and neural network models showed the best performance for price forecasting. The use of the error measurements and predictive statistical test for the step-ahead confirm this. The research can help market professionals in the development and implementation of risk management policies due to the relevance of price forecasting as a planning tool, in addition to being useful in market behavior analysis in specifying the price trend of future contracts
Keywords: Multivariate spectral analysis. Neural networks. Forecasting. Future contracts.

Recebido em 14.01.2016. Revisado por pares em 18.02.2016. Reformulações em 23.02.2016 e 18.07.2016. Recomendado para publicação em 29.07.2016. Publicado em 13.09.2016

RESUMO
Este estudo teve como proposta comparar os resultados preditivos da combinação do modelo de Análise Espectral Singular Multivariada (AESM) e o modelo de Redes Neurais Artificiais (RNA) com os resultados obtidos por modelos clássicos de previsão e de redes neurais para preços dos contratos futuros agropecuários comercializados na BM&FBOVESPA. Os resultados preditivos da combinação proposta, quando comparados com os obtidos pelos modelos clássicos e de redes neurais apresentaram os melhores desempenhos para previsões dos preços. 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 de gestão de risco por conta da relevância da previsão dos preços como instrumento de planejamento além de ser útil na análise do comportamento do mercado ao especificar a tendência dos preços dos contratos futuros.
Palavras-chave: Análise espectral multivariada. Redes neurais. Previsão. Contratos futuros.

RESUMEN
Este estudio tuvo como objetivo comparar los resultados de predicción de la combinación del modelo de Análisis Espectral Singular Multivariado (AEMS) y el modelo de Redes Neuronales Artificiales (RNA) con los resultados obtenidos por los modelos de pronóstico clásicos y redes neuronales para los precios de los contratos de futuros agrícolas negociados en la BM&FBOVESPA. Los resultados predictivos de la combinación propuesta, en comparación con los obtenidos por los modelos clásicos y redes neuronales mostraron el mejor desempeño de las previsiones de los precios. El uso de mediciones de error y prueba estadística predictiva para el paso adelante, lo confirman. La investigación puede ayudar a los profesionales del mercado en la formulación y aplicación de las políticas de gestión de riesgos debido a la relevancia de la previsión de precio como una herramienta de planificación, además de ser útil en el análisis del comportamiento del mercado mediante la especificación de la tendencia de los precios de los contratos de futuros.
Palabras clave: Análisis espectral multivariante. Redes neuronales. Predicción. Contratos de futuros.

1. INTRODUCTION

The decisions made by producers, even before the harvesting of the crop, assume knowledge of price behaviors (RIBEIRO; SOSNOSKI; OLIVEIRA, 2010). Agricultural activities are characterized for presenting cyclical movements, being influenced by various market factors and presenting high volatility (OLIVEIRA; AGUIAR, 2003). This ends up being a hindrance to its predictability.

Thus, the main problem on the price forecasting of agricultural product resides in the seasonality that occurs as a result of climatic, market and situational factors, promoting income uncertainty for farmers, storage services, exporters and processors of these products (MARTINS; MARTINELLI, 2010).

In order to administer this seasonality, the future market allows the exchange of risks with speculators, presenting itself as a security instrument and price signaling. For Hull (1996) future contracts can be defined as commitments of buying or selling a determined asset, at a pre-established date, for a price that reflects the relationship of supply and demand at that time.

Thus, price forecasting of future contracts has increasingly been the object of interest of market professionals and scholars, since through it is possible to reduce the uncertainty in decision-making by those who trade in the market. Schwager (1995) explains that by forecasting is possible to assist those involved in agricultural markets as much as speculators seeking to obtain gains through price fluctuations in the short term, providing liquidity to the market.

Given the importance of forecasting future contracts quotations in the short term, research on prediction of their prices (BRESSAN, 2004; LIMA; GÓIS; ULISES, 2007; SOBREIRO *et. al*., 2008; LIMA *et al*., (2010); FERREIRA *et.al*., 2011; MIRANDA; CORONEL; VIEIRA, 2013; TIBULO; CARLI, 2014) alternate between empirical analysis that at times use classical models that decompose the time series for subsequent prediction or the neural network model, in which such decomposition cannot be performed.

In the study conducted by Lima *et al*. (2010) the authors highlight the importance of promoting further discussion on the combination of time-series decomposition models, with models for forecasting purposes, in order to improve the performance of the forecasts made.

From this, and on a short-term condition, the objective of this research is to perform price forecasts of agricultural future contracts traded on the Stock, Commodities and Future Exchange of São Paulo (BM&FBOVESPA) by combining the two models, MSSA and ANN, defined from now on as MSSA-ANN. The rationale for its use is due to the improved condition on forecasting performance.

Then, in order to verify the performance improvement of the predictions made, the proposed research will conduct the comparison of the predictive performance of the MSSA-ANN model with the classical and neural networks models.

The article is organized as follows: in the first section the theoretical framework presents the models used in price forecasting studies for agricultural contracts in Brazil in addition to the performance of the forecasts obtained in these studies; in section 3 we describe the methodology in addition to the forecasting models used in the proposed research;

in section 4 we present the results of the tests for normality, multivariate normality, linearity, stationarity, predictive test and empirical results; the conclusions and suggestions for future research are set out in section 5.

2. THEORETICAL FRAMEWORK

In the literature on time series it is possible to distinguish two classical strategies of modeling. The first is considered to be simple and refers to the exponential smoothing models whereas the second characterizes the Box-Jenkins methodology. Exponential smoothing models, also defined as flattening or damped trend methods, are techniques developed for a specific purpose and that do not require probabilistic reasoning. They use the idea of weight distribution during the period, with the aim of considering weights variants in time. Among the exponential smoothing models, Holt-Winters seasonal (HW) algorithm is indicated for time series with more complex behavior pattern, which in addition to seasonality also consider trend and noise (MORETTIN; TOLOI, 2006).

By taking into account that through smoothing, a process of moving average is unintuitive to represent the behavior of a particular time series and, whereas the application of autoregressive models (defined by previous values and the addition of noise) is common in different areas of knowledge, we can use the autoregressive and moving average terms simultaneously when aiming at improvement. Thus, this combination characterizes the model defined by the literature as Autoregressive Moving Average Model (ARMA).

Though, if the series is non-stationary, a differentiation process is applied to it, that is, we take successive differentials from the original time series. By these means, the Autoregressive Integrated Moving Average (ARIMA) model is formed. This model is based on the construction of methods fitted in their probabilistic properties.

In some situations, time series may present periodic fluctuations, like the meteorological phenomena that, when evaluated on a quarterly basis, often have higher correlations when we use lags that are multiple of four, in accordance with the seasons of the year (ESQUIVEL, 2012). Thus, it is appropriate to consider a stochastic periodicity to evaluate the behavior of the time series. Therefore, when the ARIMA model takes seasonality into account it is then known in the literature as the Seasonal Autoregressive Integrated Moving Average (SARIMA).

In addition to these modeling strategies, research conducted with agricultural futures contracts (BRESSAN, 2004; SOBREIRO *et. al*., 2008; FERREIRA *et.al.*, 2011; MIRANDA; CORONEL; VIEIRA, 2013) make use of other strategy that does not require parameters of time series analysis. This is the Artificial Neural Network (ANN) model, in which by means of an automatic capture it approximates equations without having to deduct them. In addition to not requiring series parameters, the ANN model differs from the classical and exponential smoothing forecasting models for being a model that operates with learning algorithm. Such an algorithm seeks to imitate the interconnection structure of the human brain, with the purpose of incorporating the pattern of a time series behavior in order to efficiently provide, future values for this series (TURBAN, 1993).

The construction of the ANN model involves from the appropriate neural network modeling to the transformations used to transmit data to the network and the methods used to interpret the results. These aspects are given by: modeling, transformation and interpretation being fundamental in the use of the ANN model in order to perform price forecasting.

2.1 Forecasting performance of classic and neural networks models

When investigating the behavior of prices for soy bean contracts, through the ARIMA-GARCH model, Lima *et al*. (2010) explain that the price forecast results were particularly satisfactory. For the price forecasting of corn, Tibulo and Carli (2014) make use of the ARIMA models and the seasonal HW algorithm applied to the time series of average monthly price of corn in the region of Rio Grande do Sul. For the authors, the analysis results demonstrated that the additive seasonal HW algorithm, presented better results for the corn price forecasting when compared to the ARIMA model.

For the price forecasting of cattle, coffee and soy bean contracts, Bressan (2004) made use of the classical ARIMA model and the ANN model. The results presented gains in most of the analyzed contracts, indicating the potential use of the models as a decision tool in negotiations with an emphasis placed on operations based on forecasts made by the ARIMA model. In relation to the classical ARMA model and the ANN model used for price forecasting of future contracts for coffee, Miranda, Coronel and Vieira (2013) concluded that the ANN model demonstrated to be effective in predicting prices when evaluating its forecasting potentialities, since the predicted values were close to those observed.

In a research to evaluate the application of the ANN model Sobreiro *et al.* (2008) use prices of future contracts for sugar. The results show that the application obtained a significant approximation compared to the real quotations, which for the authors highlights the importance of the ANN model for price forecasting. Ferreira *et al*. (2011) also used the ANN model for price forecasting of futures contracts for soybeans, live cattle, corn and wheat. The data obtained in the research is evidence for the authors on the possibility of using neural networks as a pricing strategy.

These research display different methodologies to identify time series behavior patterns in analysis and forecasting. In assessing the time series behavior being studied, the classical models make use of the Box-Jenkins methodology or the idea of weight distribution during the period, with the aim of considering variant weights in time. Thus, the methodologies described in research consist of the systematic fitting of models to values of a time series so that the residues obtained in the decomposition represent noise. As the ANN model does not decompose the time series to perform the forecasting process, this ends up explaining the use of the MSSA model in the proposed research, in addition to its multivariate character.

3. METHODOLOGICAL PROCEDURES

The research to evaluate forecasting performances will compare the results from combining the MSSA and ANN models with those obtained by the seasonal HW algorithm and the SARIMA model, due to aspects of seasonality of time series, in addition to the results of the ANN model without the decomposition in time series being done.

In order to perform the decomposition by the MSSA model and the predictions of the models used in the study, we used the statistical package R. The same package is useful for statistical tests for normality, according to Anderson-Darling (AD) and Spiro-Wilk (SW) in addition to Doornik-Hansen-Omnibus (DHO) for multivariate normality. It was also useful in testing McLeod and Li (1983) and Tsay (1986) for linearity and Dickey-Fuller by Generalized Least Squares (DF-GLS) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) for time series stationarity.

As there may be error in forecasting, regardless of the model being used, it is usual to evaluate forecasting results by comparing the obtained values with the values of the original time series and determine its performance by some measure. Then, in the research, the predictions were confronted with the 8th week following the final week of the sample. To this end, the performance evaluation made use of the measure Mean Square Error (MSE) defined by:

$$EQM = \frac{1}{h}\sum_{N+1}^{N+h}\left(Y_j - \hat{Y}_j\right)^2 \qquad (01)$$

with Y_j representing the value of the original series, $\hat{Y}_j$ the value of forecast and h the amount of expected observations and reserved for evaluation. In addition to this measure, the research uses the methodology proposed by Goyal and Welch (2003), given by the differential between squared forecast errors accumulated from the best performance model and the squared forecast errors accumulated from the best subsequent performance model, considering the Cumulative Root Mean Square (CRMS) given by:

$$EQA = \sum_{N+1}^{N+h}\left(Y_j - \hat{Y}_j\right)^2 \qquad (02)$$

whenever this differential is positive, the best subsequent performance model surpasses the best performance one.

When considering two forecasts of a time series Y_t, and defining e_{it} and e_{jt} as the respective forecast errors, an analysis of the losses associated with each of these forecasts must be made through the statistical test proposed by Diebold and Mariano (1995), defined in this research as DM test, which makes use of a loss function to measure forecast error, i.e., the loss is calculated from actual and forecasted values of the variable in question. Thus, the test verifies that the differential loss is not significant between the two forecasts made.

3.1 Model used in the study

3.1.1 MSSA model

The MSSA model was initially used in atmospheric data. To this end, most of the time series were extracted from variables associated to the climate and represented by localities or regions on a map (KEPPENNE; GHIL, 1993; PLAUT; VAUTARD, 1994). The multivariate aspect of the model seeks a better separation between signal and noise preventing information from getting lost on the time series relationships amongst themselves.

Similar to the spectral model of univariate analysis, the MSSA model is defined in two stages: The stage of decomposition is given by two steps: incorporation and singular value decomposition. The incorporation can be considered as a mapping that transfers a M set of one-dimensional time series $Y_{t_i}^{(i)} = \left(y_1^{(i)},...,y_{N_i}^{(i)}\right)$, with $i = 1,...,\mathrm{M}$, for a multidimensional matrix $\left[X_1^{(i)},...,X_{K_i}^{(i)}\right]$ with vectors $X_j^{(i)} = \left(y_j^{(i)},...,y_{j+L_i+1}^{(i)}\right)^T \in R^{L_i}$, where $K_i = N_i - L_i + 1$, with a L window length.

Vectors $X_j^{(i)}$ are termed as lagged vectors. The matrix $X^{(i)}$ is a Hankel matrix which is characterized by constant inputs along diagonals parallel to the secondary diagonal. . In this step, given a M set of time series, with $t = 1,...,N$, being defined as trajectory matrices $X^{(i)}$, for $i = 1,...,\mathrm{M}$ in each time series $Y_{t_i}^{(i)}$, considering that the trajectory matrix is a sequence of the lagged vectors. The result of the incorporation step, as described by Hassani and Mahmoudvand (2013) it is the formation of a block of trajectories matrices X_V, according to:

$$X_V = \begin{bmatrix} X^{(1)} \\ \vdots \\ X^{(\mathrm{M})} \end{bmatrix} \tag{03}$$

In the second step, defined as the singular-value decomposition (SVD), the decomposition of the matrix $X_V X_V^T$ is performed obtaining a sum of elementary matrices. Thus, the eigenvalue of $X_V X_V^T$ is denoted by $\lambda_{V_1},...,\lambda_{V_{\mathrm{M}\times L}}$ in a descending order of magnitude $\left(\lambda_{V_1} \geq ... \geq \lambda_{V_{\mathrm{M}\times L}} \geq 0\right)$ and by the $U_{V_1},...,U_{V_{\mathrm{M}\times L}}$ orthogonal eigenvectors. The matrix $X_V X_V^T$ is given according to:

$$X_V X_V^T = \begin{bmatrix} X^{(1)}X^{(1)T} & X^{(1)}X^{(2)T} & \cdots & X^{(1)}X^{(\mathrm{M})T} \\ X^{(2)}X^{(1)T} & X^{(2)}X^{(2)T} & \cdots & X^{(2)}X^{(\mathrm{M})T} \\ \vdots & \vdots & \ddots & \vdots \\ X^{(\mathrm{M})}X^{(1)T} & X^{(\mathrm{M})}X^{(2)T} & \cdots & X^{(\mathrm{M})}X^{(\mathrm{M})T} \end{bmatrix} \tag{04}$$

The structure in (4) is similar to the matrix of variance-covariance obtained in the classical literature of multivariate statistical analysis and the matrix $X^{(i)}X^{(i)T}$ $X^{(i)}X^{(i)T}$ is the same used by the univariate model for a single time series $Y_{t_i}^{(i)}$. The singular-value decomposition step is represented by:

$$X_V = H_{V_1} + ... + H_{V_D} \tag{05}$$

where $H_{V_i} = \sqrt{\lambda_{V_i}} U_{V_i} V_{V_i}^T$ the elementary block matrix, $V_{V_i} = X_V^T U_{V_i} \big/ \sqrt{\lambda_{V_i}}$, the group $\sqrt{\lambda_{V_i}}, U_{V_i}, V_{Vi}$ as eigentriples and V_D the rank of the block matrix that corresponds to the number of not null eigenvectors.

In the reconstruction stage the clustering step for the MSSA model corresponds in dividing the blocks of elementary matrices $H_{V_1},\ldots,H_{V_D}$ into disjoined groups by adding them up in each group (HASSANI; MAHMOUDVAND, 2013). Thus, the unfolding of the set of indexes $J=\{1,\ldots,D\}$ into disjoined subgroups $I_1,\ldots,I_M$ corresponding to the representation:

$$X_V = H_{I_1} + \ldots + H_{I_M} \tag{06}$$

where $H_{I_1},\ldots,H_{I_M}$ are defined as a block resulting matrices.

Thus, as a simple case that presents the signal (trend and seasonality) and noise components of the time series, two groups of indices are used, according to $I_1=\{1,\ldots,a\}$ and $I_2=\{a+1,\ldots,D\}$, the first and last group associated with component signal and noise respectively, with a an integer greater than 1.

The tool which assists in the separation of the components is the graph *w*-cumulative correlation. Its methodology considers the definition of the *w*-correlation $C(f)$ cumulative values (PATTERSON *et al.*, 2011). Thus, the *w*-correlation $C(1)$ is defined with the first eigentriple group as part of the signal subseries $Y_{N_i}^{(s)}$ and the remaining eigentriple groups for the formation of the noise subseries $Y_{N_i}^{(n)}$. The *w*-correlation $C(2)$ is defined with the first and second groups of eigentriples as part of the signal subseries $Y_{N_i}^{(s)}$ and the remaining groups for the formation of the noise subseries $Y_{N_i}^{(n)}$ and so forth.

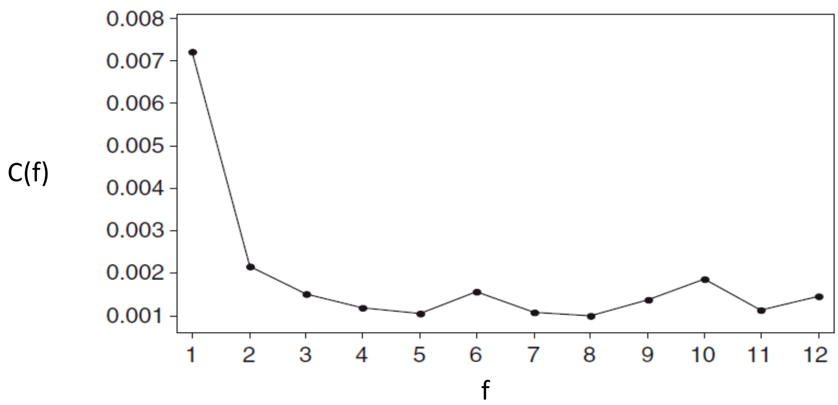


Figure 1 – Cumulative correlations for the first eigentriple
Source: Patterson *et al.*, (2011)

These *w*- cumulative correlations are plotted on a graph, as shown in Figure 1. Thus, the existence of the time series structure is indicated by local maximum and minimum. A typical pattern is a decline to the *w*-cumulative correlations and that corresponding to a separation of signal and noise components. Based on this, we notice in Figure 1 that the signal subseries $Y_{N_i}^{(s)}$ will be given by the 1-5eigentriples groups and, the noise subseries $Y_{N_i}^{(n)}$ by the groups 6-12, since $C(6)$ indicates a change of this decline.

3.1.2 Seasonal HW algorithm

Exponential smoothing models, also defined as flattening or damped trend methods, are developed for a specific purpose and that do not require probabilistic reasoning. They use the idea of weight distribution during the period, with the aim of considering weights variants in time. The incorporation of seasonality in the seasonal HW algorithm can be performed by two distinct approaches, dependent on the seasonal pattern identified in the series: multiplicative and additive seasonality. When considering the multiplicative seasonality, Morettin and Toloi (2006) explain that the time series can be defined by:

$$Y_t = N_t S_t + m_t + \varepsilon_t \qquad (07)$$

with N_t the level of the series, S_t the seasonal factor, m_t the trend component, ε_t the random error at time t and $t = 1,\ldots,N$.

The form of recurrence for the multiplicative approach, in this research, is given by HW_m, with the multiplicative seasonal factor represented by the equations involving the three smoothing constant, α, β and γ, according to:

$$\begin{cases} N_t = \alpha \dfrac{yt}{S_{t-s}} + (1-\alpha)(N_{t-1} + m_{t-1}) \\ m_t = \beta(N_t - N_{t-1}) + (1-\beta)m_{t-1} \\ S_t = \gamma \dfrac{yt}{N_t} + (1-\gamma)S_{t-s} \end{cases} \qquad (08)$$

in which $0 < \alpha < 1$, $0 < \beta < 1$ and $0 < \gamma < 1$ are the conditions of the model smoothing constant and s represents the number of observations.

The forecasts for future values take into account the steps ahead h, thus, in each equation the seasonal factor considers the corresponding period, according to the following formulas:

$$\begin{cases} \widehat{Y}_t(h) = (N_t + hm_t)S_{t+h-s}; \quad h = 1,2,3,\ldots,s \\ \widehat{Y}_t(h) = (N_t + hm_t)S_{t+h-2s}; \quad h = s+1, s+2, s+3,\ldots,2s \\ \widehat{Y}_t(h) = (N_t + hm_t)S_{t+h-3s}; \quad h = 2s+1, 2s+2, 2s+3,\ldots,3s \\ \vdots \end{cases} \tag{09}$$

For the multiplicative seasonal approach the correction of errors e_t is given:

$$\begin{cases} N_t = N_{t-1} + m_{t-1} + \alpha \dfrac{e_t}{F_{t-s}} \\ m_t = m_{t-1} + \alpha\beta \dfrac{e_t}{F_{t-s}} \\ S_t = S_{t-s} + \gamma(1-\alpha)\dfrac{e_t}{N_t} \end{cases} \tag{10}$$

The other focus of the method, given in this research by HW_a, is applied when the series features additive seasonality. Thus, for Morettin and Toloi (2006), by taking the additive seasonal factor, the time series is represented by the sum of all the components according to:

$$Y_t = N_t + m_t + S_t + \varepsilon_t \tag{11}$$

In the additive seasonality the form of recurrence is given by the formulas:

$$\begin{cases} N_t = \alpha(Y_t - S_{t-s}) + (1-\alpha)(N_{t-1} + m_{t-1}) \\ m_t = \beta(N_t - N_{t-1}) + (1-\beta)m_{t-1} \\ S_t = \gamma(Y_t - N_t) + (1-\gamma)S_{t-s} \end{cases} \tag{12}$$

with the same conditions of the smoothing constants of the model for the multiplicative approach as well as s representing the number of observations.

The future values are predicted through the formulas:

$$\begin{cases} \widehat{Y}_t(h) = (N_t + hm_t) + S_{t+h-s}; \quad h = 1,2,3,\ldots,s \\ \widehat{Y}_t(h) = (N_t + hm_t) + S_{t+h-2s}; \quad h = s+1, s+2, s+3,\ldots,2s \\ \widehat{Y}_t(h) = (N_t + hm_t) + S_{t+h-3s}; \quad h = 2s+1, 2s+2, 2s+3,\ldots,3s \\ \vdots \end{cases} \tag{13}$$

The procedure of correction of errors for this type of seasonality is given by:

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

3.1.3 SARIMA Model

In some situations it is important to consider the stochastic seasonality to explain the seasonal behavior of the time series. In this manner, the recommendation is for one of the variations of the ARIMA model to be used. This model is the SARIMA. For Box and Jenkins (1976) the general model represented by ARIMA (*p,d,q*)x(P,D,Q) can be defined as:

$$\phi(B)\Phi_P(B^s)\left[\Delta^d \Delta_s^D - \mu\right]Y_t = \theta(B)\Theta_Q(B^s)\varepsilon_t \tag{15}$$

with $\phi(B)$ the autoregressive operator, Φ_P the stationary seasonal autoregressive polynomial of order P, Δ the differential operator, μ the expected value of the series, $\theta(B)$ the moving average operator, Θ_Q the invertible polynomial of seasonal moving averages of order Q and ε_t a random error.

The stationary seasonal autoregressive polynomial of order P is given by:

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

The invertible polynomial of seasonal moving averages of order Q is given by:

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

with the seasonal differential operator of order D represented by:

$$\Delta_s^D = (1 - B^s)^D Y_t \tag{18}$$

In which, in general, the first seasonal differentiation $\Delta_s^D = (1-B^s)^D Y_t = Y_t - Y_{t-s}$ is able to exclude the seasonality of the time series (ESQUIVEL, 2012).

3.1.4 ANN model

The ANN model is adaptable to the time series and differs from classical forecasting models for being a non-parametric model and for involving learning algorithms (LIMA *et al.*, 2010). Simply put, a neural network is a computational structure based on a biological process inspired by the human brain architecture.

As described by Pasquotto (2010) each artificial neuron functions as a unit with autonomy whose objective is to convert an input signal into another output signal. Because neurons work in network, the intensity of these signals is amplified or damped according to the parameters that are assigned to synapses, also defined by synaptic weights or simply weights.

Artificial neurons are grouped into three types of layers: the input layer, the intermediate or hidden layer and output layer. For Haykin (2001), the neurons of different layers are connected by synapses which, in turn, are associated with weights or relative importance of each neuron of a layer with the neuron of a subsequent layer.

The artificial neuron model represented by Figure 2 is given by several elements. The artificial neuron elements described in Figure 1 are represented by: m which indicates the number of neuron input signals; x_j the *j*-th neuron input signal; w_{gj} the weight associated with the *j*-th input signal in the neuron g ; b the threshold of each neuron also referred to as bias; v_g a weighted combination of input signals and the bias of the *g*-th neuron and $\varphi(.)$ as the activation function of the *g*-th neuron.

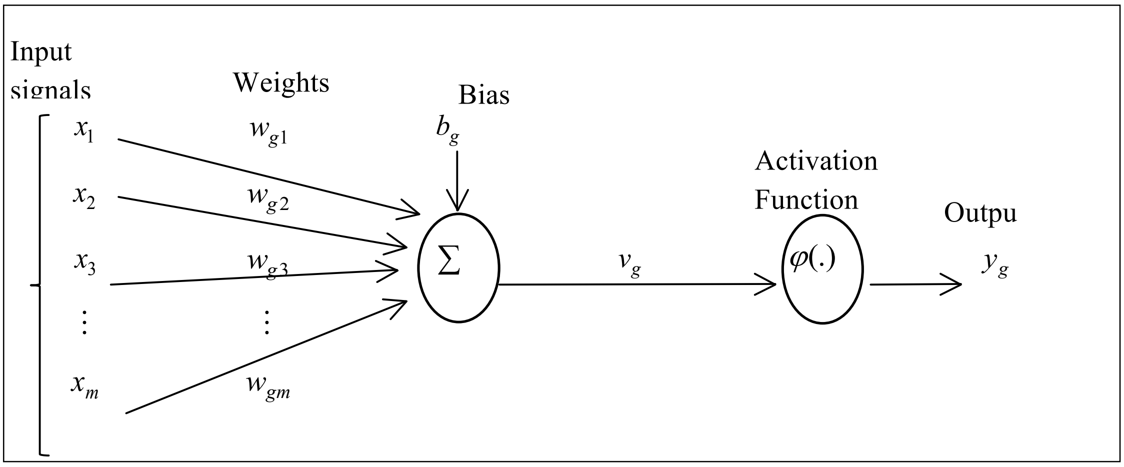


Figure 2 – Artificial Neuron
Source: Haykin (2001)

The bias has the effect of increasing or reducing the input of the activation function as its signal changes from positive to negative. With a small adjustment, Pasquotto (2010) explains that it is possible to replace the bias b_g for a fixed input $x_0 = 1$ so that the bias becomes a new synaptic weight $w_{g0} = b_g$.

Thus, we can mathematically describe neuron g by:

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

and

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

with v_g defined as the induced local field or activation potential and the activation function defining the output $y_g(t)$ of the *g*-th neuron at time *t*.

The activation function can be linear or nonlinear according to convenience. Amongst the functions most commonly used in the literature Pasquotto (2010) cites: i) threshold function, which is a discontinuous and binary function; ii) sigmoid function, which is a continuous S-shaped function, varying from 0 to 1 and iii) hyperbolic tangent function, which is a continuous and differentiable function in all its points, which makes it into a convenient function for the training algorithm.

The architecture for the ANN model goes through changes according to its purpose. The way the neurons are distributed in network is related to the learning algorithm being used. The classification given in the literature considers how the processing occurs in the neural network as well as how the neurons are arranged in layers. For Haykin (2001), classifications in single layer, multilayer, fed forward and recurrent are amongst some of the many possible ways to present the architecture of the neural network.

Another important aspect concerns the type of training which consists of adjusting the parameters of the network interactively. Pasquotto (2010) explains that there are two neural networks learning patterns: supervised and unsupervised.

Supervised learning works by indicating at the network output the correct answer for each situation. There is a set of input data presented to the neural network as examples, which generate a network output which is compared with the expected output, thus obtaining the corresponding error. Considering the neuron g at the output of a network at time t, the corresponding error e_g will be defined by:

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

with $d_g(t)$ as the neuron's desired response signal g at time t and $y_g(t)$ as the output signal of the neuron g at time t.

The error is used as an interactive parameter of fitted weights whose intention is to gradually reduce the error to a minimum acceptable value. The backpropagation algorithm is widely used for supervised learning (LIMA ET. AL., 2010).

Though it, we can browse the error function at the network output looking for a minimum point. The synaptic weights may alter after two stages are covered: : i) forward propagation and ii) backpropagation. . In the first stage, the signal is propagated along the network, starting from the first layer until it generates the error in the last layer. In the second stage, the error is corrected layer by layer, changing the weights in the reverse direction.

In the backpropagation algorithm, the input-output pairs are presented, each of them to the neural network, and there are two ways to apply the correction of errors. In the former, defined as the incremental change, the change of weights is performed whenever a new input-output pair is presented to the network generating an error that is corrected individually shortly after each pair being subjected to the network.

Unlike the former way, in the change by batch, all the *n* pairs are presented to the neural network generating an error corresponding to all the batch, and only after this, the updating of weights is performed, and it may involve the repetitive representation of the same group of pairs often times.

Unsupervised learning however, is characterized by the absence of a correct answer at the output of the neural network. That is, as there are no vectors of desired responses, there are no comparisons to provide errors. In this learning situation, the neural network is provided with conditions for the implementation of a measure regardless of the task that must be learned and the free network parameters are optimized in relation to this measure. For this type of learning Haykin (2001) explains that there are two ways of conducting it: by reinforcement and in a self-organized manner.

3.2 Data

The data used in this research were the prices of futures contracts for coffee, ethanol, cattle, corn and soybeans traded on BM&FBOVESPA obtained from the exchange database for the period January 1, 2005 to October 3, 2014, totaling 3,562 daily records. The use of contract prices closer nearest to maturity is justified since they were the most active in negotiating volume in the period being analyzed. Then, the time series was converted into a weekly basis for a total of 509 weeks.

4. REDULTS ANALYSIS

4.1 Statistical tests applied to the data

The series of prices for contracts traded on BM&FBOVESPA are identified in the research as presented: ICF (coffee), ETH (ethanol), BGI (live cattle), CCM (corn) and SOJ (soybean) following the code suggested by the stock exchange. To test whether the sample data comes from a population with a specific distribution AD and SW tests are used. In the research the two tests allow a comprehensive view whether data follow a normal distribution. As can be seen from the *p*-value results displayed in Table 1, we can reject the null hypothesis. With this the time series are not normally distributed.

Table 1 – Tests of normality AD, SW and *p*-value

	ICF	ETH	BGI	CCM	SOJ
Number of Observations	509	509	509	509	509
Shapiro-Wilk	0.91	0.90	0.92	0.91	0.91
p-value	0.00	0.00	0.00	0.01	0.00
Anderson-Darling	2.61	1.47	3.21	1.73	1.82
p-value	0.00	0.00	0.00	0.00	0.00

Source: Research data

In order to evaluate the aspect of normality of the data group, we used the DHO test, which is a multivariate normality test applied between pairs formed by time series whose null hypothesis is that data follow a normal multivariate distribution. The *p*-value results presented in Table 2 indicate that there is strong evidence of multivariate normality between the formed pairs. In the research using the DHO test is justified to the knowledge of the time series characteristics since the MSSA model does not presuppose the hypothesis of normality of the data.

Table 2 – Test of multivariate normality, DHO and *p*-value

	ICF	ETH	BGI	CCM	SOJ
COFF.		24.03	47.09	27.03	37.44
p-value		0.00	0.00	0.00	0.00
ETHA.			26.05	24.29	16.84
p-value			0.00	0.07	0.00

CATT.	46.59	24.15
p-value	0.00	0.00
CORN		21.53
p-value		0.00

Source: Research data

The non-linearity test by Tsay (1986) and McLeod and Li (1983) are also applied to the time series. For the first test data is filtered by an AR model. Table 3 presents the results for the sample data based on lags of 5 and 10 weeks. Whereas the two tests have as null hypothesis that the data follow a linear behavior, based on the *p*-values displayed it is possible to reject the null hypothesis and conclude therefore, that the time series are not linear.

Table 3 – *p*-value for Tsay and McLeod's test.

	Lags	ICF	ETH	BGI	CCM	SOJ
Number of Observations		509	509	509	509	509
Tsay	5					
p-value		0.01	0.00	0.00	0.00	0.00
Tsay	10					
p-value		0.00	0.00	0.01	0.00	0.00
McLeod	5					
p-value		0.00	0.00	0.00	0.01	0.00
McLeod	10					
p-value		0.01	0.00	0.00	0.00	0.00

Source: Research data

Finally, we performed the DF-GLS and KPSS tests to evaluate the stationarity of the time series. Table 4 shows the results of the two tests. For the first test, the null hypothesis is that time series have unit roots, and therefore are not stationary. On the second test the null hypothesis is that time series do not have root unit, and therefore are stationary. Thus, the tests confirm that the time series are not stationary. In summary, the time series used in the research are not stationary, non-linear and, do not present multivariate normality.

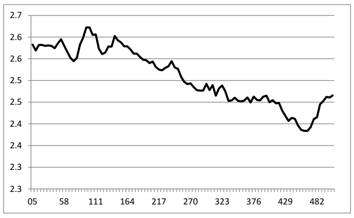
Table 4 –DF-GLS and KPSS test for the set of series

Time series	DF-GLS	5% Critical Value	KPSS	5% Critical Value
ICF	-1.57	-2.89	0.17	0.146
ETH	-1.25	-2.89	0.21	0.146
BGI	-2.13	-2.89	0.18	0.146
CCM	-2.15	-2.89	0.24	0.146
SOJ	-1.48	-2.89	0.16	0.146

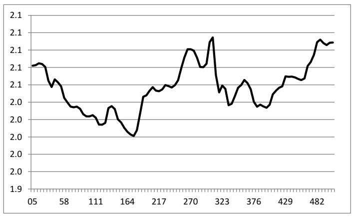
Source: Research data

Figure 3 presents the time series behavior of each futures contract prices in the period from 1 January 2005, to 3 October 2014, for a total of 509 weeks. Through it we can verify the seasonality aspect of prices for each contract.

(a) Time series ICF and ETH

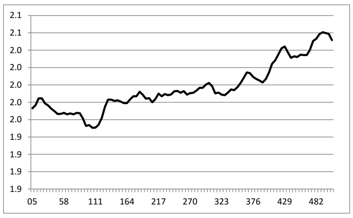


week

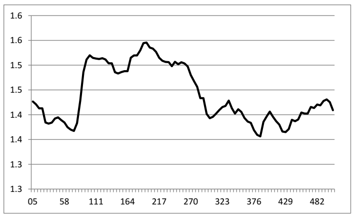


week

(b) Time series BGI and CCM

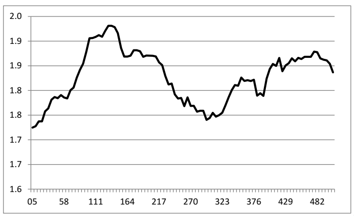


week



week

(c) Time series SOJ



week

Figure 3 – Weekly price behavior of contracts
Source: Research data

4.2 Empirical applications

When performing the forecasts through the HWa and HWm models used in the research, the choice of the smoothing parameter values is defined by the R package based on the minimization of the sum of squares of the one-step-ahead forecasts errors for subsequent application of routine forecast.

In relation to the SARIMA model, the orders $(p,d,q)(P.D,Q)$ were defined based on the minimization of values for Akaike Information Criterion (AIC) through routine of the R package. That is, the routine is supposed to identify the most suitable model for each series by identifying orders $(p,d,q)(P.D,Q)$ by minimizing the values for AIC, considering d and D, the degree of differentiation, and the degree of seasonal differentiation, respectively.

To verify the presence of seasonal unit root the DHF (Dickey, Hazsa, Fuller) test was applied, whose null hypothesis is the presence of a unit root. Through Table 5 we identify that the series have seasonal root (not stationary).

Table 5 –DHF test for the set of series

Time series	DHF	5% Critical Value
ICF	-1.45	-3.74
ETH	-1.53	-3.74
BGI	-2.33	-3.74
CCM	-2.76	-3.74
SOJ	-1.97	-3.74

Source: Research data

In the MSSA-ANN model we used as the window length L equal to 85 weeks since $L=(1+N)/(1+M)$ defines the optimal window size (HASSANI; MAHMOUDVAND, 2013). Then, after the decomposition and removal of noise the time series was submitted to the predictive process of the ANN model. In the application of the ANN model the R package made use of the neural network with one layer comprised of some neurons, assuming a maximum number of iterations equal to 100,000 with supervised learning and application of backpropagation algorithm.

Of the 509 weeks available for each time series the first 420 were used as input and the remaining ones as training. Since the package allows changing the number of neurons, we used models with some neurons. After the training, the model with 7 neurons was the one that had the lowest value for AIC, thus being selected as the most suitable one. As the research also proposes the application of the ANN without combining it with another model this same routine is applicable in the time series without the decomposition and noise elimination being done.

Similar to research carried out by Lima, Góis and Ulises (2007), Sobreiro *et al*. (2008) and Lima *et al*. (2010) who used short time periods, according to 13, 17 and 10 days, respectively for agricultural price forecasting, the proposed research made use of steps ahead of 4 and 8 weeks to forecast. This period is considered as out-of-sample period and meets the short-term objectives for those who negotiate and guarantee liquidity to the market.

Table 6 – Forecasting performance through MSE

	Parameters			MSE					
Series	L	(p,d,q)(P,D,Q)	h	MSSA-ANN	HWa	HWm	SARIMA	ANN	MD
ICF	85	(0,1,0)(0,1,1)	4	1.2-E04	1.8E-03	1.8E-03	6.2E-03	1.3E-03	MSSA-ANN
			8	1.3-E04	1.9E-03	1.9E-03	6.4E-03	1.4E-03	MSSA-ANN
ETH	85	(0,1,1)(0,1,1)	4	2.4E-04	2.6E-04	2.6E-04	2.5E-04	2.9E-04	MSSA-ANN
			8	2.4E-04	2.5E-04	2.5E-04	2.5E-04	2.9E-04	MSSA-ANN
BGI	85	(1,1,0)(0,1,1)	4	4.0E-04	5.0E-04	5.0E-04	1.5E-04	2.1E-04	SARIMA
			8	4.1E-04	5.3E-04	5.3E-04	1.4E-04	2.0E-04	SARIMA
CCM	85	(1,1,0)(0,1,1)	4	5.1E-04	7.1E-04	7.1E-04	5.7E-04	6.4E-04	MSSA-ANN
			8	5.2E-04	7.3E-04	7.3E-04	5.6E-04	6.4E-04	MSSA-ANN
SOJ	85	(1,1,0)(0,1,1)	4	1.7E-05	1.5E-03	1.5E-03	8.6E-05	1.9E-04	MSSA-ANN
			8	1.6E-05	1.3E-03	1.3E-03	8.5E-05	1.3E-04	MSSA-ANN

Source: Research data

In addition to the window length L for MSSA and the parameters used by the SARIMA model, we can observe in Table 6 based on the MSE measure that there is no difference to the forecasts made between the seasonal HW algorithm with additive and multiplicative seasonality. The forecasts obtained through the MSSA-ANN when compared with forecasts obtained by the seasonal HWa and HWm algorithm, by the SARIMA and ANN model correspond to the best performance (MD in Table 6), due to the lower values for the measure of error described in (02). The exception for this is given in the steps-ahead h (4 and

8 weeks) for the BGI times series whose best forecasting performances are given to the SARIMA model.

In Table 7, which evaluates the performance of models through the differential between the accumulated squared forecasting errors of the best performance model (shaded column in Table 7) and the accumulated squared forecast errors of the best subsequent performance model, we notice that the model of best forecasting performance exceeds the best subsequent performance one since the differentials presented are negative.

Table 7 – Forecasting performance through CRMS differential

Series	h	MSSA-ANN	ANN	Differential
ICF	4	1.6E-03	1.4E-02	-1.2E-02
	8	1.3E-03	1.5E-02	-1.4E-02
		MSSA-ANN	SARIMA	
ETH	4	5.4E-05	1.7E-03	-1.6E-03
	8	5.3E-06	1.9E-03	-1.9E-03
		SARIMA	ANN	
BGI	4	7.0E-05	4.0E-04	-3.3E-04
	8	7.0E-06	4.9E-04	-4.9E-04
		MSSA-ANN	SARIMA	
CCM	4	2.5E-05	7.5E-04	-7.3E-04
	8	1.3E-06	8.5E-04	-8.5E-04
		MSSA-ANN	SARIMA	
SOJ	4	6.3E-05	4.3E-03	-4.2E-03
	8	8.0E-05	1.3E-03	-1.2E-03

Source: Research data

Then, in order to evaluate whether the difference between the MSEs of the model of best performance with the model of best subsequent performance is statistically significant, we applied the DM test. The *p*-value results exposed in Table 8 indicate that, for the models being compared, the null hypothesis which claims that the differential between the measured error is zero can be rejected for the time series ICF, ETH, CCM and SOJ. In relation to the BGI time series the same hypothesis may also be rejected, indicating that there is superiority of the results obtained through the model SARIMA.

Table 8 –Diebold-Mariano test and models compared

Series	h	DM	*p*-value	Models
ICF	4	5.87	0.00	
	8	5.61	0.00	MSSA-ANN, ANN

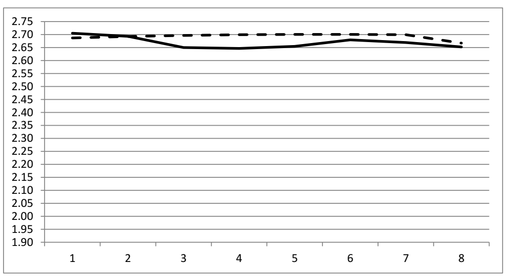
ETH	4	9.13	0.00	
	8	9.01	0.00	MSSA-ANN, SARIMA
BGI	4	5.00	0.00	
	8	5.79	0.00	SARIMA, ANN
CCM	4	5.79	0.00	
	8	5.23	0.00	MSSA-ANN, SARIMA
SOJ	4	6.21	0.00	
	8	5.57	0.00	MSSA-ANN, SARIMA

Source: Research data

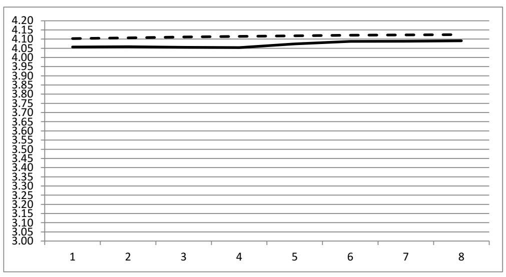
We can conclude, based on the DM statistical test, through the performance of the forecasts made for the steps-ahead *h* (4 and 8 weeks), that the MSSA-ANN added in the analysis period evidence supporting its application in the agricultural futures market since the differential between the measures of error being equal zero it may be rejected. This indicates, therefore, that the best forecast performance obtained by MSSA-ANN statistically exceed that model of best subsequent performance.

In Figure 4 we present, through the graphics of prices for weekly period, the behaviors of the original time series (solid lines) and forecasted time series (dashed lines) obtained through the MSSA-ANN, since it presented the best forecasting performance. The graphics were developed in the period from 10 October 2014, to 28 November 2014, out-of-sample. We can observe that the MSSA-ANN, in this period, was able to detect the trend of prices, except for the BGI time series.

(a) Time series ICF and ETH

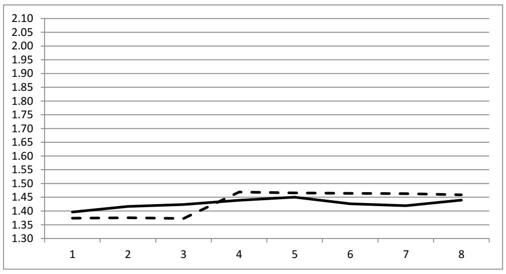


week

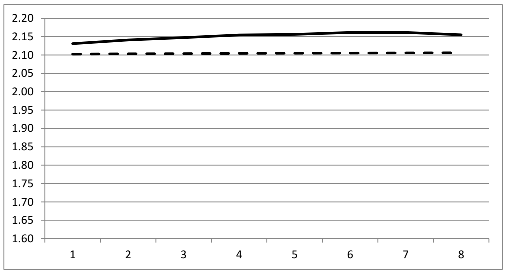


week

(b) Time series CCM and BGI

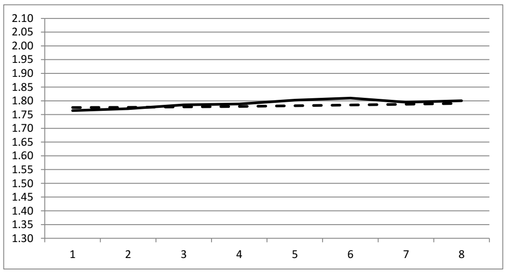


week



week

(c) Time series SOJ



week

Figure 4 –Behavior of the original and expected prices of contracts

Source: Research data

5. CONCLUSIONS

Price forecasting of agricultural future contracts, as the exchange mechanism element, is of particular importance for those involved in the agricultural market as it is for speculators. Thus, the decisions made by these participants in the short term, assume the knowledge of price behavior.

Research made on price forecasting of agricultural future contracts are provided over the empirical analysis which alternate between using classical models that decompose the time series or the application of neural network model. An alternative provided in the literature is the combination of time-series decomposition models with models for forecasting purposes.

In this context, the objective of the research is combining the use of the MSSA model with the ANN in order to perform price forecasting of agricultural future contracts traded on the BMF&BOVESPA, in the sort-term. The justification for its application is due to the improved condition of forecasting performance.

In addition, in choosing the MSSA model for the research to decompose the time series we sought to capture the time series structures representing the comprehensive behavior taking into account the effects of the combination, given the character of the multivariate model.

In the research conducted by Lima *et al*. (2010) to investigate the behavior of prices of soybean contracts based on the ARIMA-GARCH / ANN model, the forecast results were favorable to the ARIMA-GARCH models. Other than that, the results indicated forecasting superiority for MSSA-ANN.

Regarding soybeans and corn, the study by Ferreira *et al*. (2011) highlighted the possibility of using the ANN model as a pricing strategy due to the favorable results. Also regarding the same contracts, the study developed by Lima, Góis and Ulises (2007) indicated that the autoregressive model showed a better forecasting power. These results were not confirmed since the predictive performance showed superiority for MSSA-ANN in relation to the autoregressive and neural network models.

In evaluating the forecasts for the coffee future contract prices, Miranda, Coronel and Vieira (2013) concluded that the ANN model when compared to the ARMA model proved to be effective for price forecasting, since the predicted prices were close to the observed ones. For coffee prices the forecasting performance results were favorable to MSSA-ANN. The only exception for forecasting superiority over the MSSA-ANN was for the live cattle future contracts since the SARIMA model presented better results.

Therefore, in the context of the research and except for the prices of live cattle futures contracts, empirical results demonstrate superiority for MSSA-ANN, when compared to the HWa, HWm, SARIMA and ANN models, by allowing a greater number of better forecasting performances. The results obtained in the out-of-sample period, through the use of the MSE and CRMS error measures in addition to the MD test for steps-ahead (4 and 8 weeks) confirm this.

Generally speaking, MSSA-ANN exceeded, in terms of statistical losses, the model which presented the best subsequent performance. Thus, in the research proposing the combination between the MSSA and ANN models has contributed in the discussion regarding the combination of time series decomposition models with models for forecasting purposes.

From a practical point of view the results constitute an alternative to assist in the formulation and implementation of risk management policies, due to the relevance of price forecasting and market behavior analysis when specifying the trend of prices of agricultural future contracts.

For future research we recommend the use of other databases, for example, prices of future contracts in international markets, the inclusion of other agricultural products, the adoption of other periods of analysis and the use of other variables that may increase the explanatory power for MSSA-ANN given its multivariate character.

REFERENCES

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

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

DIEBOLD, F., MARIANO, R. Comparing Predictive Accuracy. **Journal of Business and Economic Statistics**, v. 13, n. 3, p. 253-265, 1995.

ESQUIVEL, R. M. **Análise espectral singular: modelagens de séries temporais através de estudos comparativos usando diferentes estratégias de previsão**. 161f. 2012. Dissertação (Mestrado em Modelagem Computacional e Tecnologia Industrial) – Faculdade de Tecnologia SENAI CIMATEC, Salvador, 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, n. 4, p. 6-26, 2011.

GOYAL, A.; WELCH, I. Predicting the equity premium with dividend ratios. **Management Science**, v. 49, n. 5, 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, n. 1, p. 55-83, 2013.

HULL, J. **Introdução aos mercados futuros e de opções**. 2 ed. São Paulo: BM&F, 1996.

KEPPENNE, C.; GHIL, M. Adaptive filtering and prediction of noisy multivariate signals: An application to subannual variability in atmospheric angular momentum. **International Journal of Bifurcation and Chaos**, v. 3, n. 3, p. 625-634, 1993.

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, n. 3, 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, n. 2, 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. 1, n. 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, n. 4, p. 269-273, 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**, n. 38, p. 66-98, 2013.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. São Paulo: Blucher, 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, 4, 2003, Ribeirão Preto/SP. **Anais Eletrônicos**. Ribeirão Preto, 2003.

PASQUOTTO, J. L. D. **Previsão de séries temporais no varejo brasileiro: Uma investigação comparativa da aplicação de redes neurais recorrentes de Elman**. 2010. 189f. Dissertação (Mestrado em Administração) – Faculdade de Economia, Administração e Contabilidade de São Paulo, Universidade de São Paulo, São Paulo, 2010.

PATTERSON, K.; HASSANI, H.; HERAVI, S.; ZHIGLJAVSKY, A. Multivariate singular spectrum analysis for forecasting revisions to real-time data. **Journal of Applied Statistics**, v. 38, n. 10, p. 2183-2211, 2011.

PLAUT, G.; VAUTARD, R. Spells of low-frequency oscillations and weather regimes in the northern hemisphere. **Journal Atmospheric Sciences,** v. 51, n. 2, p. 210–236, 1994.

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, n. 4, 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**, v. 6, n. 1, p. 36-53, 2008.

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

TSAY, R. Non-linearity tests for time series. **Biometrika**, v. 73, n. 2, p. 461-466, 1986.

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