

Forecasting commodity prices in Brazil through hybrid SSA-complex seasonality models

Rafael Baptista Palazzi^{a,b*} , Paula Maçaira^c , Erick Meira^d , Marcelo Cabus Klotzle^a 

^aPontifícia Universidade Católica do Rio de Janeiro, IAG Escola de Negócios, Rio de Janeiro, RJ, Brasil

^bUniversity of Illinois at Urbana-Champaign, Agricultural and Consumer Economics, Urbana-Champaign, IL, USA

^cPontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, Rio de Janeiro, RJ, Brasil

^dFinanciadora de Estudos e Projetos, Rio de Janeiro, RJ, Brasil

*rafael_palazzi@prof.iag.puc-rio.br

Abstract

Paper aims: To predict monthly corn, soybean, and sugar spot prices in Brazil using hybrid forecasting techniques.

Originality: This study combines the Singular Spectrum Analysis with different forecasting methods.

Research method: This paper presents a set of hybrid forecasting approaches combining Singular Spectrum Analysis (SSA) with different univariate time series methods, ranging from complex seasonality methods to machine learning and autoregressive models to predict monthly corn, soybean, and sugar spot prices in Brazil. We carry out a range of out-of-sample forecasting experiments and use a comprehensive set of forecast evaluation metrics. We contrast the performance of the proposed approaches with that of a range of benchmark models.

Main findings: The results show that the proposed hybrid models present better performances, with the hybrid SSA-neural network approach providing the most competitive results in our sample.

Implications for theory and practice: Forecasting agricultural prices is of paramount importance to assist producers, farmers, and the industry in decision-making processes.

Keywords

Forecasting. Hybrid approaches. Singular spectrum analysis. Commodities.

How to cite this article: Palazzi, R. B., Maçaira, P., Meira, E., & Klotzle, M. C. (2023). Forecasting commodity prices in Brazil through hybrid SSA-complex seasonality models. *Production*, 33, e20220025. <https://doi.org/10.1590/0103-6513.20220025>

Received: Feb. 25, 2022; Accepted: Sept. 22, 2022.

1. Introduction

Reliable price forecasting for agricultural commodities is of paramount importance to assist farmers, producers, and the industry in their decision-making processes. In such contexts, having efficient mechanisms to predict price trends may support optimal input allocation, agricultural investments, and hedging decisions. However, forecasting agricultural commodities is challenging; besides the effects of supply and demand and seasonal patterns, exogenous factors such as climate changes, technology, exchange rates, and political and macroeconomic events may influence prices.

Several studies apply different filtering methods to overcome the complexity of commodity forecasting to remove noise in the time series to improve the prediction performance combined with forecasting methods. Safari & Davallou (2018) proposed a hybrid model combining the Kalman-filter algorithm with exponential smoothing, ARIMA, and NN. Xiong et al. (2018) combined season-trend decomposition and the ELM approach to forecast vegetable prices. Degiannakis et al. (2018) applied the SSA and Holt-Winters to predict implied volatility indices. All the studies mentioned earlier demonstrated the outperformance of the hybrid models.



Thus, motivated by the previous findings, we propose a hybrid model to predict spot prices for sugar, soybeans, and corn in the Brazilian market. Notably, we combine SSA with complex seasonality, machine learning algorithms, and autoregressive models to predict monthly commodities prices up to 2 and 3 steps ahead. We compare whether using the SSA decomposition method to isolate critical components before forecasting provides more competitive results than resorting to single benchmarks.

We contribute to the literature by including different approaches that consider seasonality, machine learning, and auto-regressive using the same forecasting algorithms in the reconstructing step of the SSA-denoised process. In addition, we perform a one-step ahead forecasting, refitting the model for each month in a twelve-month window. The results are of particular importance, considering Brazil's global role as the leading exporter of sugar and soybeans and the fourth-largest supplier of corn. Notably, the country accounted for 38% of the total sugar exported in 2020, followed by India with 11% (Companhia Nacional de Abastecimento, 2021). The country is also a leading soybean exporter, accounting for 37% of total world production. Although the country is not the leading corn exporter, it is one of the top four countries that provides 70% of corn supply worldwide. Corn production has increased over the years due to the "2a Safra" crop rotation. Therefore, agricultural commodities constitute a crucial driver of the Brazilian economy, corresponding to a large portion of the country's Gross Domestic Product (GDP) (Universidade de São Paulo, 2021).

The rest of the paper is organized as follows. Section 2 discusses the Literature on forecasting models using hybrid models. Section 3 describes the Singular Spectrum Analysis method and its implementation. Section 4 briefly overviews the forecasting benchmarks considered for comparison and describes the data and the setup to train and test the data. Section 5 presents and discusses the empirical results. Section 6 concludes and presents suggestions for future studies.

2. Literature review

The forecasting literature provides several approaches to capture the dynamics and complexity of commodities prices using different sorts of methods and combining them to give a more accurate prediction. For instance, Gibson & Schwartz (1990) and Schwartz (1997) described the behavior of commodity prices through differential stochastic models. The latter study applied the Kalman Filter to estimate the parameters of the unobserved state variables. Based on these previous findings, Ribeiro & Oliveira (2011) proposed a hybrid model for forecasting the prices of agricultural commodities considering the structure of the Schwartz (1997) stochastic process, which describes the price evolution following selected activation functions for Neural Networks (NNs), thus allowing the incorporation of nonlinear dynamics into the models.

Further studies showed the importance of the combination approaches to capture linear and nonlinear information of the time series. Xiong et al. (2015) proposed a combination of vector error correction (VEC) and multi-output Support Vector Regression (SVR) to forecast interval-valued agricultural commodity futures prices. Maia et al. (2008) introduced a new approach based on a hybrid ARIMA and ANN model. They showed that the proposed model outperformed the AR and ARIMA single approaches. Li et al. (2019) suggested two-hybrid methods for monthly crude oil price forecasting using a support vector machine (SVM) optimized by a genetic algorithm (GA) and a backpropagation neural network optimized by GA. The variational model decomposition with artificial intelligence provided more accurate results than all benchmark models. Other studies supported the advantages of using hybrid models to predict agricultural commodity prices (Fang et al., 2020; Paul & Garai, 2021; Wang et al., 2017).

Decomposition methods were also employed to improve model forecasting performance. For instance, Liu et al. (2022) proposed a hybrid NN approach to predict the copper price, using NN with Bayesian Optimization to search for the hyperparameters and Wavelet Transform to denoise the data and remove noise information. Wang & Li (2018) applied a NN forecasting model to the outputs of a Singular Spectrum Analysis (SSA) decomposition for corn, gold, and crude oil. In their proposal, SSA was first used to decompose the price series, and the smoothed commodity prices series were reapplied into the NN model (Huang et al., 2019). Overall, the SSA method decomposes the original series data into trend, oscillatory, and noise components. Meira et al. (2021) considered the use of Seasonal Trend Decomposition using Loess (STL) (Cleveland et al., 1990) to isolate key components of electricity supply time series before using exponential smoothing models, thus improving forecasting performance over several benchmarks. A similar process was adopted by Meira et al. (2022) to decompose key components of natural gas consumption time series. Athoillah et al. (2021) applied Support Vector Regression (SVR) to predict rainfall based on reconstructed series from an SSA decomposition, which was assumed to be free from noisy elements. Their results showed that the SVR with the filtered data outperformed the SVR model.

A few other studies applied the SSA and compared it with other forecasting models. For example, Hassani et al. (2009) used SSA, ARIMA, and Holt-Winters' techniques in a different window ($h=1,3,6$, and 12 months) for the industrial production sectors in German, French, and UK. The results were similar for short-term periods; however, SSA had a better performance at longer horizons. Hassani et al. (2015) analyzed the advantages of using SSA compared with ARIMA, exponential smoothing (ETS), and NN. The authors indicated the superiority of the SSA model over other single models in forecasting tourist arrivals into the US. Hassani et al. also found that SSA outperforms several single models such as ARIMA, ETS, NN, Trigonometric Box-Cox ARMA Trend Seasonal (TBATS), and Fractionalized ARIMA (ARFIMA) to predict tourism demand in European countries. Fathi et al. (2022) combined SSA and nonlinear autoregressive neural networks (NARNN) to predict 24 stock prices and compared them with ARIMA and NARNN. Their results also showed that SSA combined with NARNN provides better forecasting performance than ARIMA and NARNN.

3. Singular Spectrum Analysis (SSA)

This section briefly describes the methodology applied in the spot prices to “filter” the prices, removing the time series' noise (for more detailed information on the theory of SSA sees Golyandina et al., 2001).

SSA is a nonparametric method that can be used for time series analysis and forecasting and that does not require the classical assumptions over the normality of the residuals or the stationarity of the time series. Not only does this technique overcome the traditional approaches, but it also incorporates elements of classical time series analysis, multivariate statistics, and signal processing (Rodrigues et al., 2018).

SSA involves two stages of decomposing and reconstructing the time series, each one consisting of two steps:

3.1. Embedding

The embedding step involves mapping a one-dimensional time series $Y_N = (y_1, \dots, y_N)$ into the multidimensional series X_1, \dots, X_K with vectors $X_i = (x_i, \dots, x_{i+L-1})^T \in R^L, (1 \leq i \leq K)$, where $K=N-L+1$. The X_i vectors are called L -lagged vectors. L is an integer and describes the window length; its size should not exceed $N/2$. The single choice of L in the embedding step results in the trajectory matrix X formed by the sub-series X_i , which is also a Hankel matrix,

$$X = [x_1 \dots x_{K-1}; \dots; x_{L-1} \dots x_{N-1}] \tag{1}$$

L stands for the number of components that the time series is decomposed. The parameter is $1 < L < N$. Hassani (2007) pointed out that L should be large enough but not greater than $T/2$. Nonetheless, in case the time series has a seasonal component, for instance, it is recommendable that the window's length is proportional to the periodic component to improve the separability of this component.

3.2. Singular Value Decomposition (SVD)

The SVD of the trajectory matrix X denotes the sum of rank-one bi-orthogonal elementary matrices. Let $S = XX^T$ and $\lambda_1, \dots, \lambda_L$ be the eigenvalues of S ordered by their magnitude ($\lambda_1 \geq \dots \geq \lambda_L$). Let also U_1, \dots, U_L be the orthogonal eigenvectors of S related to the eigenvalues $v_i = \frac{X^T U_1}{\sqrt{\lambda_1}}$ ($i=1, \dots, d$). We can denote the SVD as the trajectory of the matrix X represented as follows:

$$X = X_1 + \dots + X_d \tag{2}$$

3.3 Grouping

The main goal of the grouping step is to distinguish the additive components of the time series in separable matrixes to identify the set of correlated components and to group them in order of the highest correlation group. Thus, the first step of stage 2 involves separating the elementary matrices X_i into several groups and then summing each group of matrices. The group of indices is defined by $I = \{i_1, \dots, i_p\}$; thus, the matrix X_I associated

with the group l can be defined as $X_J = X_{J1} + \dots + X_{Jp}$. The set of indices $J=1, \dots, d$ is divided into disjoint subsets I_1, \dots, I_m that correspond to the following representation (Hassani et al., 2015):

$$\mathbf{X} = X_{I_1} + \dots + X_{I_m} \tag{3}$$

The process of selecting the sets I_1, \dots, I_m is called eigentriple grouping, such that the contribution of group l to the X_J component is measured by the share of the corresponding eigenvalues: $\sum \lambda_i / \sum_{i=1}^d \lambda_i$.

3.4. Diagonal Averaging (DA)

DA consists of transforming each matrix I into an additive component of the original series. The transformation of the resulting matrices into the series occurs by applying the Hankelization (H) linear operator. This operator acts as an arbitrary matrix transforming it into a Hankel matrix - trajectory matrix - resulting in a time series.

3.5. Choosing the parameters L and r

The purpose of the SSA method is to filter the original data so that the time series forecasting models can be applied to the filtered observations, easing model estimation, and potentially reducing errors in the subsequent forecasts. The choice of the window length (L) and the reconstruction grouping (r) will be defined during the decomposition process. There are no strict rules for selecting the L parameter.

A data preprocessing technique combined with forecasting models has proven to improve prediction performance. However, Zhang et al. (2014) showed that several papers decompose the entire data series and then separate the components into calibration and validation sets. According to the authors, this procedure sends an amount of future information into the decomposition and reconstruction steps, using information from future values that have not been updated in the forecasting exercise. Such a procedure is called hindcast.

Therefore, to avoid this problem and ensure that no future information is taken to the filtered time series, we choose the L and r based on the best performance of the combination of the parameters obtained in the training step. In our study, we follow the procedure highlighted by Hassani et al. (2015). However, it is worth mentioning that to predict the parameters L and r , we apply the same forecasting models in the out-of-sample step rather than use the SSA algorithm described in the Hassani et al. (2015) procedure.

Let us consider a time series $Y = (y_{(1,1)} + \dots + y_{(1,S)}, \dots, y_{(N,S)})$ with NS observations and a seasonal period of length S . Further, we split the data into two parts: a training set $Y = (y_{(1,1)} + \dots + y_{(1,S)}, \dots, y_{(K,S)})$ with size $T = KS$, and a test set $Y = (y_{(K+1,1)} + \dots + y_{(K+1,S)}, \dots, y_{(N,S)})$ with $h = (N - K - 2)S$ observations.

Then, the following steps are conducted for $L = S, 2S, \dots, \frac{KS}{2}$:

- i. construct the trajectory matrix $\mathbf{X} = [X_1 + \dots + X_K]$ using the training set;
- ii. obtain the SVD by calculating $\mathbf{X}\mathbf{X}^T$;
- iii. for $m = 1, \dots, L$, group the first components $\mathbf{X} = X_1 + \dots + X_m$, then apply the diagonal average in \mathbf{X} to obtain the filtered series \tilde{Y} . Choose the forecasting model (ETS, ARIMA, NN, NNETAR, ELM, and Fourier in our study) and apply it to \tilde{Y} to obtain the h steps ahead forecasts $\tilde{Y} = (\tilde{y}_{(1,1)} + \dots + \tilde{y}_{(1,S)}, \dots, \tilde{y}_{(N,S)})$. Finally, compute the Mean Absolute Percentage Error (MAPE) when comparing the forecast \tilde{Y} with the validation set. The L and r vector parameters will be selected as being the ones presenting the lowest MAPE, which describes the optimal decomposition and reconstruction choice for the SSA approach.

The MAPE is defined as follows:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{\hat{y}_t} \right| \tag{4}$$

where n is the forecast horizon (number of forecasting steps ahead), y_t represents the actual (observed) prices of the commodity, and \hat{y}_t comprises the forecasts generated by each forecasting model. Figure 1. depicts the main steps involved in the forecasting process.

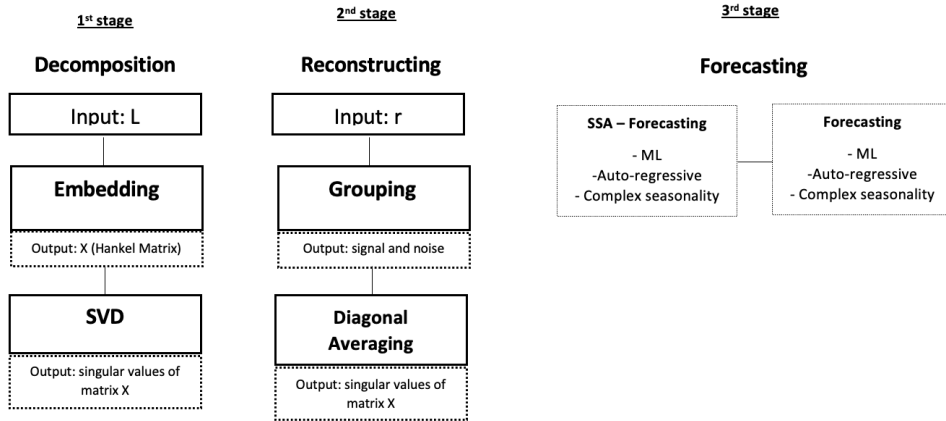


Figure 1. Singular spectrum analysis (SSA) process (Sanei & Hassani, 2015) and the forecasting scheme.

4. Forecasting benchmarks

4.1. Exponential smoothing methods

We apply three forecasting models to account for seasonal patterns in the commodities prices. First, we employ the Exponential Smoothing (ETS) algorithm that considers error, trend, and seasonal components by selecting the best exponential smoothing model from several combinations. The best-fit model is selected by default by minimizing the Akaike Information Criterion with corrections (AICc) (Sugiura, 1978). This automatic algorithm can be implemented through the *forecast* package for the R software (Hyndman et al., 2022).

In addition, we consider the additive Holt-Winters' method, which can be written as:

$$\hat{y}_{t+1} = L_t + kT_t + S_{t+k-m} \quad (5)$$

where L_t represents the level component, T_t is the trend, and S_t the seasonality. These, in turn, can be represented as follows:

$$L_t = \alpha(y_t - S_{t-m}) + (1 - \alpha)(L_{t-1} + T_{t-1}),$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1},$$

$$S_t = \gamma(y_t - L_t) + (1 - \gamma)S_{t-m} \quad (6)$$

where α , β e γ are the smooth hyperparameters corresponding to the level, trend, and seasonality components. The *forecast* package in R also provides the implementation of the additive Holt-Winters' forecasting method that is obtained by minimizing the BIC.

4.2. Seasonal Autoregressive, Integrated, Moving Average (SARIMA) formulations

Two SARIMA formulations are considered for comparison. The popular automatic version of the (Box & Jenkins, 1970) – ARIMA model is provided in the forecast R package. The `auto.arima()` function automatically selects the number of differences d , and p and q are estimated by minimizing the AICc, p , and q , respectively, referring to the orders of the autoregressive and moving-average parts of the selected SARIMA model. The algorithm allows for selecting three different unit-root tests, the KPSS (Kwiatkowski et al., 1992) test, the Augmented Dickey-Fuller (ADF) test, and the Phillips-Perron (PP)(Phillips & Perron, 1988).

Lastly, to deal with possible multiple seasonal patterns in the time series, we also consider the extension of the traditional, univariate SARIMA formulation to a dynamic harmonic regression framework. This consists of adding Fourier terms in the SARIMA formulation as regressors (exogenous factors), thus representing other possible seasonal patterns that may be present in the involved time series (Hyndman & Athanasopoulos, 2021).

4.3. Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANNs) models have become popular in many areas of applied sciences since the development of data availability and computing power. A neural network is organized into several layers: input, output, and hidden layers. In the hidden layer, the neurons assume estimated values (weights). Each signal is multiplied by a weight, which indicates the influence on the unit's output. The number of hidden layers is equivalent to linear regressions, and the coefficients attached to the predictors are called weights. The weights are selected through an algorithm that minimizes a forecast error measure. One of the advantages of the ANN models is their non-parametric nature and their capability to generate output with higher classification accuracies than traditional statistical classifiers (Kavzoglu & Mather, 2003).

An intermediate layer is added, so each layer of the node receives inputs from the previous layers in a process called a multilayer feed-forward network using a weighted linear combination (Hyndman & Athanasopoulos, 2021),

$$z_j = b_j + \sum_{i=1}^n w_{i,j}x_i \tag{7}$$

The Equation 7 is modified using a nonlinear function such as a sigmoid, where $s(z) = \frac{1}{1 + e^{-z}}$.

We compare three different ANN approaches to forecast commodity prices. First, we consider the multilayer perceptron (MLP), which comprises a system of interconnected neurons that map input data sets onto a set of outputs, representing a nonlinear mapping. The MLP makes no prior assumptions regarding its distribution (Gardner & Dorling, 1998). We select the algorithm from the `nnfor` package using the `mlp()` function (Crone & Kourentzes, 2010; Fildes & Allen, 2015). The algorithm employs a supervised learning method called backpropagation for training the networks. Second, we compare the `mlp()` with the `nnetar()` function in the `forecast` package that fits a NN to a nonlinear auto-regressive model. The `nnetar()` function uses a simulation (default 1000) to derive the prediction intervals using the `npaths` argument (Hyndman & Athanasopoulos, 2021). Finally, we compare these approaches with the extreme learning machine (ELM). According to Ding et al. (2015), ELM is a single hidden layer feedforward neural network based on a learning process that needs only a single iteration. ELM has the ability to approximate a complex nonlinear function and solve the issues that traditional parameter learning methods cannot accomplish (Ding et al., 2015). Thus, we apply the `elm()` function from the `nnfor` package. The function is a training method that neither requires iterative tuning nor setting the hyperparameters; it can be faster than other methods and overcomes over-fitting problems (Huang et al., 2006).

4.4. Experimental setup

The training and test set data were used to choose the best L and r parameters, described in section 2.5. We apply the same forecasting models to find the L and r parameters used to reconstruct the denoised series using the Training set from November 2013 to November 2020 and then testing using two steps ahead. After we find the optimal parameters, we input them into the SSA equation to generate the new reconstructed time series. In the next step, we use the training 2 set to train the model by applying the denoised series to predict the remaining observation with the benchmark forecasting models and compare it with the test 2 set. In this second step, we train the new data series from November 2013 to December 2020 and January 2021. The test set consists of the following 2 and 3 months. The training and test for both procedures are depicted in Figure 2.

This study considers the monthly spot prices of corn, soybean, and sugar in the Brazilian commodity market depicted in Figure 3. We retrieved the database from CEPEA/Esalq website (Universidade de São Paulo, 2022) from November 2013 to March 2021.

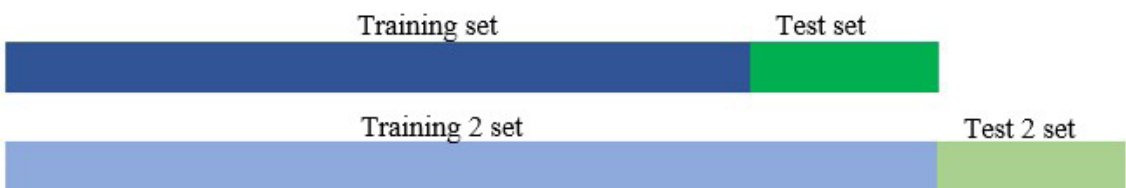
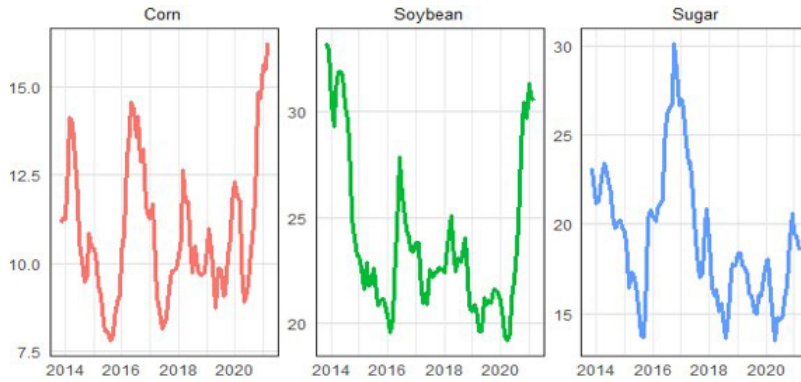


Figure 2. Training and test set to perform the forecasting models.



Source: CEPEA/ESALQ.

Figure 3. Monthly spot prices of corn (US\$/60kg bag), soybean (US\$/60kg bag), and sugar (US\$/50kg bag). Source: Universidade de São Paulo (2022).

5. Empirical results

5.1. Forecasting using 2 and 3 months ahead

This section compares the accuracy of the Hybrid approaches, which consist of reconstructing the denoised series using SSA and applying seven different models (ARIMA, ELM, ETS, Fourier, HW, NN, NNETAR), with the accuracy of the seven different models when directly applied to the original series, i.e., with no SSA involved. We examine which forecasting approach performs the best in terms of MAPE (the lower, the better). We use 2 and 3 months in advance to forecast the commodities prices for the test set and the validation set. Figure 4 and 5 show the MAPE values of the different forecasting methods involved when forecasting 2 and 3 months ahead. The bars in blue correspond to MAPE values observed for the hybrid approaches, while MAPE values for the single methods are presented in orange bars.

Overall, the hybrid approach outperforms the Single models, particularly when forecasting up to two months ahead. Prominent results arose from the combined use of SSA with selected machine learning approaches – see, for instance, the MAPE values of the SSA-NN method when forecasting sugar prices two months in advance, and the MAPE values of SSA-NN, SSA-NNETAR, and the SSA-ELM when forecasting soybean prices. The hybrid SSA-exponential smoothing method also presented competitive results, outperforming several time series methods in most cases. In turn, the Hybrid SSA-ARIMA and the single ARIMA models were the least competitive overall, particularly when forecasting corn and soybean prices, as depicted by their considerably higher values of MAPE.

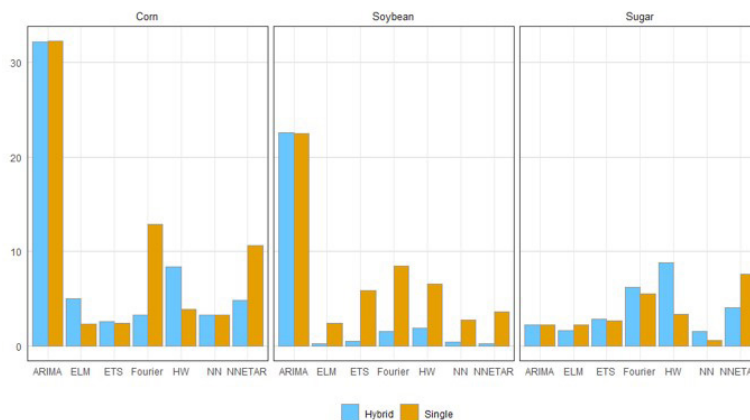


Figure 4. Forecast errors (MAPE) comparison between the hybrid and the Single model with $h=2$.

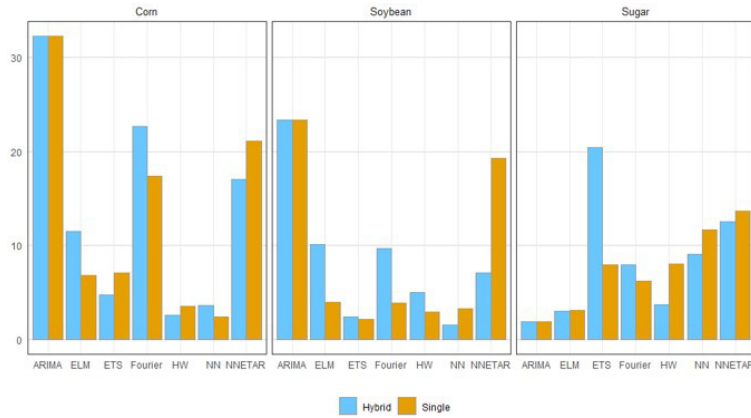


Figure 5. Forecast errors (MAPE) comparison between the hybrid and the Single model with h=3.

Moreover, the forecasting models for soybean prices had a better performance than the other commodities. Noteworthy, the corn price had a steep spike in the last observations while sugar prices declined. For robustness purposes, we also considered a forecast evaluation using the Mean Average Error (MAE) and the Mean Absolute Scale Error (MASE) methods to measure the forecast accuracy. These metrics are computed according to the formulae depicted in Equation 6. Tables 1 and 2 illustrate the forecast errors using these metrics (MAE, MAPE, and MASE) (Hyndman & Koehler, 2006). Overall, the results in terms of MAE and MASE followed, to a greater extent, those observed in terms of MAPE.

Table 1. MAE, MAPE, and MASE values when forecasting two steps ahead (h=2).

	Corn			Soybean			Sugar		
	MAE	MAPE	MASE	MAE	MAPE	MASE	MAE	MAPE	MASE
<i>Hybrid approaches</i>									
SSA-ETS	0.40	2.56	0.17	0.16	0.51	0.05	0.52	2.79	0.12
SSA-ARIMA	5.11	32.20	2.19	6.91	22.59	2.37	0.42	2.22	0.10
SSA-HW	1.33	8.40	0.56	0.57	1.87	0.18	1.65	8.80	0.39
SSA-NN	0.53	3.24	0.23	0.13	0.42	0.04	0.29	1.55	0.07
SSA-NNETAR	0.76	4.85	0.33	0.06	0.21	0.02	0.76	4.03	0.18
SSA-ELM	0.78	4.98	0.33	0.08	0.25	0.02	0.31	1.65	0.07
SSA-Fourier	0.51	3.23	0.22	0.48	1.57	0.16	1.15	6.17	0.27
<i>Single methods</i>									
ETS	0.38	2.36	0.16	1.78	5.82	0.51	0.50	2.67	0.12
ARIMA	5.11	32.21	2.15	6.86	22.44	1.96	0.42	2.22	0.10
HW	0.62	3.91	0.26	2.01	6.58	0.58	0.63	3.31	0.15
NN	0.51	3.24	0.22	0.83	2.73	0.24	0.12	0.60	0.03
NNETAR	1.70	10.63	0.72	1.10	3.59	0.31	1.45	7.60	0.34
ELM	0.36	2.27	0.15	0.73	2.38	0.21	0.43	2.22	0.10
Fourier	2.05	12.90	0.86	2.59	8.48	0.74	1.04	5.50	0.24

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

$$MAE_{naive}^{in-sample} = \frac{1}{N-1} \sum_{t=1}^n |y_t - y_{t-1}|$$

$$MASE = \frac{MAE}{MAE_{naive}^{in-sample}} \quad (8)$$

Table 2. Performance evaluation for Hybrid and Single models when forecasting three months ahead (h=3).

	Corn			Soybean			Sugar		
	MAE	MAPE	MASE	MAE	MAPE	MASE	MAE	MAPE	MASE
<i>Hybrid approaches</i>									
SSA-ETS	0.75	4.76	0.32	0.72	2.36	0.22	3.88	20.43	0.89
SSA-ARIMA	5.09	32.23	2.17	7.2	23.35	2.12	0.36	1.9	0.08
SSA-HW	0.4	2.57	0.18	1.53	5	0.46	0.7	3.68	0.16
SSA-NN	0.58	3.61	0.25	0.47	1.55	0.14	1.72	9.06	0.4
SSA-NNETAR	2.7	17.03	1.15	2.17	7.08	0.67	2.37	12.54	0.55
SSA-ELM	1.82	11.52	0.78	3.1	10.07	0.91	0.57	2.98	0.13
SSA-Fourier	3.59	22.69	1.53	2.95	9.63	0.87	1.51	7.96	0.35
<i>Single methods</i>									
ETS	1.12	7.06	0.47	0.66	2.14	0.2	1.51	7.96	0.35
ARIMA	5.09	32.23	2.15	7.2	23.35	2.11	0.36	1.91	0.08
HW	0.56	3.53	0.24	0.89	2.9	0.26	1.52	8.01	0.35
NN	0.38	2.41	0.16	1	3.23	0.29	2.23	11.68	0.51
NNETAR	3.35	21.14	1.42	5.92	19.24	1.74	2.57	13.61	0.59
ELM	1.08	6.79	0.46	1.22	3.96	0.36	0.59	3.1	0.14
Fourier	2.74	17.35	1.16	1.19	3.86	0.35	1.17	6.25	0.27

Hassani et al. (2015) found similar results. The authors compared the SSA forecast with ETS, ARIMA, and NNETAR. They showed that SSA performed better than the alternative approaches providing a significant advantage in forecasting tourist arrivals in the US. Also, Lima et al. (2010) compared the ARIMA-GARCH and the NN forecasting models for the Brazilian soybean log returns. However, they applied the wavelet decomposition to filter the time series and reapply into the econometric models. Results showed that using wavelet decomposition improved the forecasts, and the NN presented better forecasts than the ARIMA-GARCH. Ribeiro & Oliveira (2011) also found that using NN can reduce forecast errors.

Besides the forecasting error measures presented in Tables 1. and 2., we also test whether the difference between the methods is (statistically) significant. Hence, we apply the Nemenyi test (Demšar, 2006) that ranks the performance method for each data series (see the results in Appendix A). Overall, the ETS and machine learning models were revealed to have the best performance in most of the results. However, the results comparing the hybrid models were mixed. For example, the soybean results showed that ELM was the best, followed by hybrid machine learning techniques. For the corn results, the best was ETS, followed by ELM and the SSA-NNETAR. The Nemenyi test for sugar demonstrates the best performance of both ETS, single, and hybrid. Nevertheless, the hybrid models, on average, especially the machine learning methods, outperformed the other models, which emphasizes our previous findings.

5.2. One-step ahead forecasting for the best model

In this step, we compare the best models in the previous section by forecasting one step ahead with re-estimation to ensure whether the hybrid models outperform the single approach. We start analyzing the NNETAR model for soybean. Figure 6 shows the forecasting results for a 12-month window. The forecast error measured by the MAPE shows a significant difference between the hybrid model (SSA_NNETAR: MAPE=0.25) and the single method (NNETAR: MAPE=1.55); thus, the results corroborate our previous analysis of the power prediction of the hybrid model for the soybean.

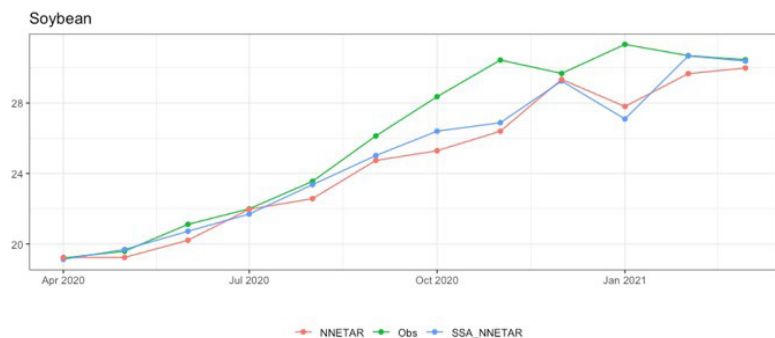


Figure 6. Forecasting one step ahead with re-estimating soybean price over 12 months using the hybrid and single NNETAR model.

Next, we proceed with the same approach by forecasting with a one-step ahead estimation over a 12-month window for the corn spot price using the ETS forecasting model in Figure 7. The results also indicate an advantage of the hybrid model over the single method. We compute the MAPE to measure the forecast error, so for the hybrid approach, the MAPE was 2.21, while the single model was 3.35.



Figure 7. Forecasting one step ahead with re-estimating corn price over 12 months using the hybrid and single NNETAR model.

Lastly, we apply the one-step forecasting for the sugar price using the ELM forecast approach in Figure 8. In comparing the hybrid and the single model, the latter performed better. We computed the MAPE to measure the forecast error, which showed a forecast error of 2.80 for the ELM hybrid and 5.20 for the ELM model.

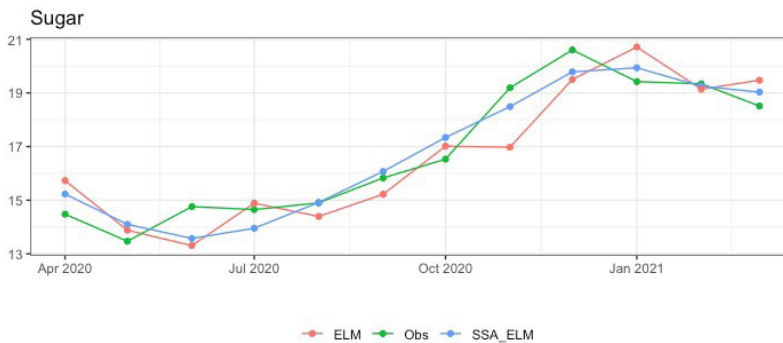


Figure 8. Forecasting one step ahead with re-estimating sugar price over 12 months using the hybrid and single NNETAR model.

6. Concluding remarks

Forecasting models play an important role in estimating future price behavior in commodity markets. Producers and farmers can benefit from accurate models to support input allocation, while commodity traders may sharpen their optimal hedging and investment decisions. In Brazil, having reliable spot price forecasts of selected agricultural commodities, such as corn, soybean, and sugar, is of utmost importance since these commodities constitute an important driver of the country's economy.

The present study contributes to the related literature by proposing a range of hybrid approaches applied to Brazilian agricultural commodities price forecasting. The methods first consider the Singular Spectrum Analysis (SSA) technique to decompose the original time series into their key components: trend, oscillatory, and noise components. Then, the SSA reconstructs the denoised series, further fed into several econometric and machine learning models. To compare whether filtering the time series with the SSA reduces forecasting errors. It is worth noting that the main characteristic of agricultural commodities prices is their seasonal component. Therefore, we consider several seasonal models – ETS, Fourier, and SARIMA formulations – in the hybrid strategies. We also include the combined use of SSA with selected machine learning algorithms, particularly ELM, NNETAR, and NN.

In most cases, the hybrid approaches, i.e., those that combine the SSA technique with the selected econometric or machine learning methods – outperform those that do not consider the previous filtering with SSA. Among single models, the machine learning algorithms presented the best performances overall.

A possible avenue for future research is to explore alternative decomposition schemes for time series filtering besides SSA, such as wavelet decomposition. In addition, further studies could extend the forecasting investigation using a multivariate framework.

References

- Athoillah, I., Wigena, A. H., & Wijayanto, H. (2021). Hybrid modeling of singular spectrum analysis and support vector regression for rainfall prediction. *Journal of Physics: Conference Series*, 1863(1), 12054. <http://dx.doi.org/10.1088/1742-6596/1863/1/012054>.
- Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: forecasting and control*. San Francisco: Holden-Day.
- Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: a seasonal-trend decomposition procedure based on loess (with discussion). *Journal of Official Statistics*, 6(1), 3-73.
- Companhia Nacional de Abastecimento – CONAB. (2021). *Informações agropecuárias*. Retrieved in 202, February 25, from <https://www.conab.gov.br/info-agro>
- Crone, S. F., & Kourentzes, N. (2010). Feature selection for time series prediction: a combined filter and wrapper approach for neural networks. *Neurocomputing*, 73(10-12), 1923-1936. <http://dx.doi.org/10.1016/j.neucom.2010.01.017>.
- Degiannakis, S., Filis, G., & Hassani, H. (2018). Forecasting global stock market implied volatility indices. *Journal of Empirical Finance*, 46, 111-129. <http://dx.doi.org/10.1016/j.jempfin.2017.12.008>.
- Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7, 1-30.
- Ding, S., Zhao, H., Zhang, Y., Xu, X., & Nie, R. (2015). Extreme learning machine: algorithm, theory and applications. *Artificial Intelligence Review*, 44(1), 103-115. <http://dx.doi.org/10.1007/s10462-013-9405-z>.
- Fang, Y., Guan, B., Wu, S., & Heravi, S. (2020). Optimal forecast combination based on ensemble empirical mode decomposition for agricultural commodity futures prices. *Journal of Forecasting*, 39(6), 877-886. <http://dx.doi.org/10.1002/for.2665>.
- Fathi, A. Y., El-Khodary, I. A., & Saafan, M. (2022). Integrating singular spectrum analysis and nonlinear autoregressive neural network for stock price forecasting. *IAES International Journal of Artificial Intelligence*, 11(3), 851. <http://dx.doi.org/10.11591/ijai.v11.i3.pp851-858>.
- Fildes, R., & Allen, P. G. (2015). Forecasting. In R. W. Griffin (Ed.), *Management*. Oxford: Oxford University Press. <http://dx.doi.org/10.1093/obo/9780199846740-0064>.
- Gardner, M. W., & Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron): a review of applications in the atmospheric sciences. *Atmospheric Environment*, 32(14), 2627-2636. [http://dx.doi.org/10.1016/S1352-2310\(97\)00447-0](http://dx.doi.org/10.1016/S1352-2310(97)00447-0).
- Gibson, R., & Schwartz, E. S. (1990). Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance*, 45(3), 959-976. <http://dx.doi.org/10.1111/j.1540-6261.1990.tb05114.x>.
- Golyandina, N., Nekrutkin, V., & Zhigljavsky, A. A. (2001). *Analysis of time series structure SSA and related techniques*. London: Chapman & Hall.
- Hassani, H. (2007). Singular spectrum analysis: methodology and comparison. *Journal of Data Science : JDS*, 5(2), 239-257. [http://dx.doi.org/10.6339/JDS.2007.05\(2\).396](http://dx.doi.org/10.6339/JDS.2007.05(2).396).
- Hassani, H., Heravi, S., & Zhigljavsky, A. (2009). Forecasting European industrial production with singular spectrum analysis. *International Journal of Forecasting*, 25(1), 103-118. <http://dx.doi.org/10.1016/j.ijforecast.2008.09.007>.
- Hassani, H., Webster, A., Silva, E. S., & Heravi, S. (2015). Forecasting U.S. tourist arrivals using optimal singular spectrum analysis. *Tourism Management*, 46, 322-335. <http://dx.doi.org/10.1016/j.tourman.2014.07.004>.
- Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K. (2006). Extreme learning machine: theory and applications. *Neurocomputing*, 70(1), 489-501. <http://dx.doi.org/10.1016/j.neucom.2005.12.126>.
- Huang, X., Maçaira, P. M., Hassani, H., Cyrino Oliveira, F. L., & Dhesi, G. (2019). Hydrological natural inflow and climate variables: time and frequency causality analysis. *Physica A*, 516, 480-495. <http://dx.doi.org/10.1016/j.physa.2018.09.079>.
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: principles and practice* (3rd ed.). Melbourne: OTexts.
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679-688. <http://dx.doi.org/10.1016/j.ijforecast.2006.03.001>.
- Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., Yasmeeen, F., Garza, F., Girolimetto, D., Ihaka, R., Reid, D., Shaub, D., Tang, Y., Wang, X., & Zhou, Z. (2022). *forecast: forecasting functions for time series and linear models*. Retrieved in 202, February 25, from <https://cran.r-project.org/package=forecast>
- Kavzoglu, T., & Mather, P. M. (2003). The use of backpropagating artificial neural networks in land cover classification. *International Journal of Remote Sensing*, 24(23), 4907-4938. <http://dx.doi.org/10.1080/0143116031000114851>.
- Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1), 159-178. [http://dx.doi.org/10.1016/0304-4076\(92\)90104-Y](http://dx.doi.org/10.1016/0304-4076(92)90104-Y).
- Li, J., Zhu, S., & Wu, Q. (2019). Monthly crude oil spot price forecasting using variational mode decomposition. *Energy Economics*, 83, 240-253. <http://dx.doi.org/10.1016/j.eneco.2019.07.009>.
- Lima, F. G., Kimura, H., Assaf Neto, A., & Perera, L. C. J. (2010). Previsão de preços de commodities com modelos ARIMA-GARCH e redes neurais com ondaletas: velhas tecnologias - novos resultados. *Revista de Administração*, 45(2), 188-202. <https://doi.org/10.1590/S0080-21072010000200008>.
- Liu, K., Cheng, J., & Yi, J. (2022). Copper price forecasted by hybrid neural network with Bayesian Optimization and wavelet transform. *Resources Policy*, 75, 102520. <http://dx.doi.org/10.1016/j.resourpol.2021.102520>.
- Maia, A. L. S., Carvalho, F. A. T., & Ludermit, T. B. (2008). Forecasting models for interval-valued time series. *Neurocomputing*, 71(16-18), 3344-3352. <http://dx.doi.org/10.1016/j.neucom.2008.02.022>.

- Meira, E., Cyrino Oliveira, F. L., & de Menezes, L. M. (2021). Point and interval forecasting of electricity supply via pruned ensembles. *Energy*, 232, 121009. <http://dx.doi.org/10.1016/j.energy.2021.121009>.
- Meira, E., Cyrino Oliveira, F. L., & de Menezes, L. M. (2022). Forecasting natural gas consumption using Bagging and modified regularization techniques. *Energy Economics*, 106, 105760. <http://dx.doi.org/10.1016/j.eneco.2021.105760>.
- Paul, R. K., & Garai, S. (2021). Performance comparison of wavelets-based machine learning technique for forecasting agricultural commodity prices. *Soft Computing*, 25(20), 12857-12873. <http://dx.doi.org/10.1007/s00500-021-06087-4>.
- Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335-346. <http://dx.doi.org/10.1093/biomet/75.2.335>.
- Ribeiro, C. O., & Oliveira, S. M. (2011). A hybrid commodity price-forecasting model applied to the sugar-alcohol sector. *The Australian Journal of Agricultural and Resource Economics*, 55(2), 180-198. <http://dx.doi.org/10.1111/j.1467-8489.2011.00534.x>.
- Rodrigues, P. C., Tuy, P. G. S. E., & Mahmoudvand, R. (2018). Randomized singular spectrum analysis for long time series. *Journal of Statistical Computation and Simulation*, 88(10), 1921-1935. <http://dx.doi.org/10.1080/00949655.2018.1462810>.
- Safari, A., & Davallou, M. (2018). Oil price forecasting using a hybrid model. *Energy*, 148, 49-58. <http://dx.doi.org/10.1016/j.energy.2018.01.007>.
- Sanei, S., & Hassani, H. (2015). *Singular spectrum analysis of biomedical signals*. Boca Raton: CRC Press. <http://dx.doi.org/10.1201/b19140>.
- Schwartz, E. S. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, 52(3), 923-973. <http://dx.doi.org/10.1111/j.1540-6261.1997.tb02721.x>.
- Sugiura, N. (1978). Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics. Theory and Methods*, 7(1), 13-26. <http://dx.doi.org/10.1080/03610927808827599>.
- Universidade de São Paulo – USP, Centro de Estudos Avançados em Economia Aplicada – CEPEA. (2021). *PIB do agronegócio brasileiro*. Retrieved in 202, February 25, from <https://www.cepea.esalq.usp.br/br/pib-do-agronegocio-brasileiro.aspx>
- Universidade de São Paulo – USP, Centro de Estudos Avançados em Economia Aplicada – CEPEA. (2022). *Price index*. Retrieved in 202, February 25, from <http://www.cepea.esalq.usp.br/>
- Wang, D., Yue, C., Wei, S., & Lv, J. (2017). Performance analysis of four decomposition-ensemble models for one-day-ahead agricultural commodity futures price forecasting. *Algorithms*, 10(3), 108. <http://dx.doi.org/10.3390/a10030108>.
- Wang, J., & Li, X. (2018). A combined neural network model for commodity price forecasting with SSA. *Soft Computing*, 22(16), 5323-5333. <http://dx.doi.org/10.1007/s00500-018-3023-2>.
- Xiong, T., Li, C., & Bao, Y. (2018). Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: evidence from the vegetable market in China. *Neurocomputing*, 275, 2831-2844. <http://dx.doi.org/10.1016/j.neucom.2017.11.053>.
- Xiong, T., Li, C., Bao, Y., Hu, Z., & Zhang, L. (2015). A combination method for interval forecasting of agricultural commodity futures prices. *Knowledge-Based Systems*, 77, 92-102. <http://dx.doi.org/10.1016/j.knsys.2015.01.002>.
- Zhang, W., Su, Z., Zhang, H., Zhao, Y., & Zhao, Z. (2014). Hybrid wind speed forecasting model study based on SSA and intelligent optimized algorithm. *Abstract and Applied Analysis*, 2014, 693205.

Appendix A. Robustness test using statistical comparisons of classifiers

Figure 1A, 2A, and 3A show that the models are represented by lines that fall inside the gray area, indicating that they are not statistically different in performance from each other.

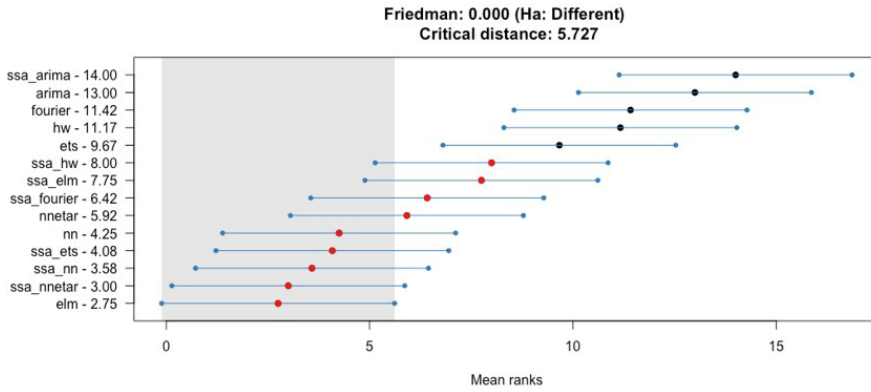


Figure 1A. Nemenyi test results at 95% confidence levels for the soybean data series. The hierarchical forecasting models are classified as bottom-up according to the MASE mean rank.

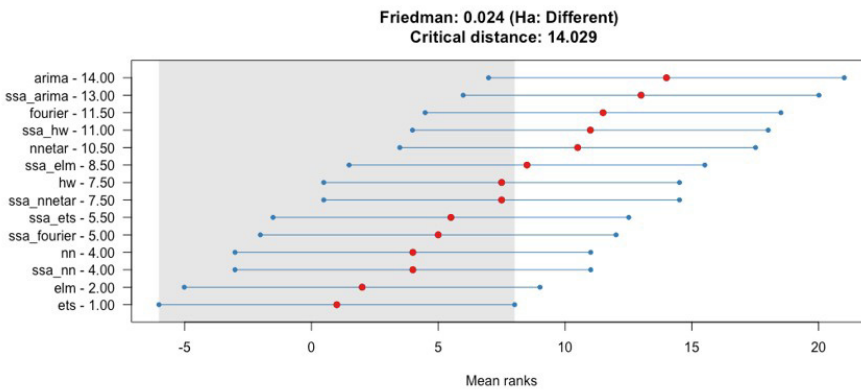


Figure 2A. Nemenyi test results at 95% confidence levels for the corn data series. The hierarchical forecasting models are classified as bottom-up according to the MASE mean rank.

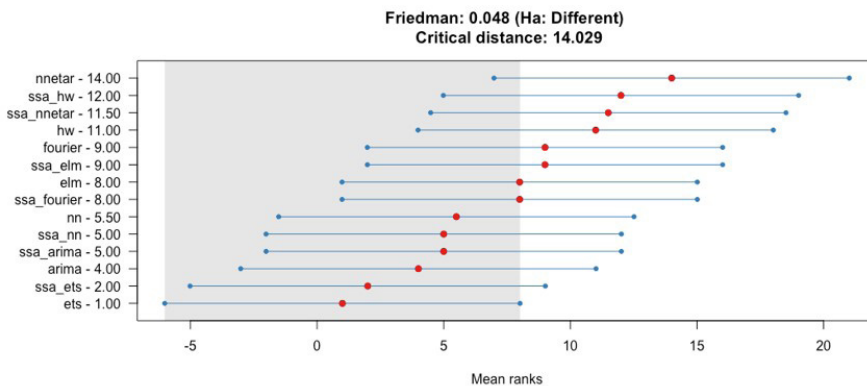


Figure 3A. Nemenyi test results at 95% confidence levels for the sugar data series. The hierarchical forecasting models are classified as bottom-up according to the MASE mean rank.