International Research Journal of Finance and Economics
ISSN 1450-2887 Issue 172 March, 2019
http://www.internationalresearchjournaloffinanceandeconomics.com

Formulation of Brazilian Sugar basis Forecasting using Time Series Models: Comparison between the Northeast and Southeast Spot and Ice Futures Markets

Felipe Araujo de Oliveira
Risk Management Analyst Management of Pricing Models at B3
Praça Antonio Prado, 48 – 4º. Andar – Centro
01010-010 – São Paulo/SP
Tel: (55 83) 99421-8868
E-mail: felipe.8.araujo@gmail.com

Márcio Rodrigues Bernardo
Researcher in Finance at COPPEAD
Universidade Federal do Rio de Janeiro COPPEAD
Rua Pascoal Lemme, 355 – Cidade Universitária
21941-918 – Rio de Janeiro/RJ
Tel: (55 21) 3938-9860
E-mail: marcio.bernardo@coppead.ufrj.br

Waldemar Antônio da Rocha de Souza
Professor at Faculdade de Economia, Administração e Contabilidade – FEAC
Universidade Federal de Alagoas
Campus Universitário 57072-900– Maceió/AL
Tel: (55 82) 99824-3899
E-mail: waldemar.souza@feac.ufal.br

Carlos Heitor Campani
Professor of Finance at COPPEAD – UFRJ & Research Associate at Edhec-Risk Institute
Universidade Federal do Rio de Janeiro, COPPEAD
Rua Pascoal Lemme, 355 – Cidade Universitária
21941-918 – Rio de Janeiro/RJ
Tel: (55 21) 3938-9860
E-mail: carlos.heitor@coppead.ufrj.br

Abstract

Sugar basis identification indicates a strategic tool for decision-making in Brazil. We formulate an accessible forecast using ARMA (p, q) time series models, comparing Brazilian sugar spot and ICE futures markets. We identify sugar basis in Alagoas and São Paulo with synchronous high and low primary movements, volatility, with a correlation coefficient of 0.70, illustrating a high and positive magnitude, and descriptive statistics for Alagoas that are higher than São Paulo. Alagoas shows seasonality over the year, whereas São Paulo, for nine months. Comparing forecasting errors between ARMA (p, q) and SARMAX (p, q) models, the monthly sugar basis model for Alagoas is SARMAX (1, 0). For São Paulo, the forecast model is ARMA (2, 0). In addition, the Alagoas sugar basis breakpoint month is 2008M09, and São Paulo breakpoint month is 2011M11. These breakpoints coincide with the identifiable sugar basis level and volatility trends in the examined period. Both forecast models are easy to use and implement, and are strategic informational inputs for more efficient allocative decisions by the Brazilian sugar supply chain.

Keywords: Brazilian sugar basis; Forecast; Time series; Breakpoints; ARMA (p, q) and SARMAX (p, q).
JEL Classification: C22, E27, Q14

1. Introduction

Brazil is the largest producer and exporter of sugar. In effect, Brazilian production predicts a 1.1 million ton rise, totaling a record of 40.2 million tons. Exports may increase by 1.1 million tons to a record 29.6 million based on larger exportable supplies despite China's limit of sugar imports from Brazil (USDA United States Department of Agriculture 2017).

Furthermore, to highlight price variation and shock and impact on Brazil and international sugar markets, in early 2011 international sugar prices reached a 30-year high of nearly US$ 37 per 50kg bag before falling 68%, to approximately US$ 12 per 50kg bag in late 2015. Recently, international sugar prices showed additional high volatility, reaching a peak of approximately US$ 25 per 50kg bag in mid-2016, and falling back to US$ 15 per 50kg bag in 2017, a 40% decrease (ICE, 2017).

To identify whether the sugar price spikes are temporary oscillations caused by a supply shock or a more permanent fundamental shift in global market dynamics, we need to evaluate the economic and policy factors driving production and trade in key global sugar markets. Since Brazil is the largest sugar producer and exporter, the underlying dynamics due to Brazil's exchange rates and ethanol's role in energy markets are clues to define global prices (USDA United States Department of Agriculture, 2010).

In this case, a strategic analytical variable to describe the underlying dynamics of the Brazilian sugar price is the basis. Basis is defined by the difference between spot and futures prices (Leuthold, Junkus, and Cordier 1989), which express, in particular, an idiosyncratic pattern. Also, it is linked to commodity portfolio allocation, risk premium, and convenience yields. Depending on the agent's position on the commodity supply chain, basis is the relevant variable, such as the US local elevators and exporters.

As such, an in-depth comprehension of the dynamics of sugar basis is relevant for the Brazilian sugar market agents. The sugar market agents need to identify the past and forecast the future trajectory of the basis. The Brazilian sugar basis trajectory can be influenced by several variables. Among others, seasonality expresses a major impact, as well as volatility, transportation, and storage costs, and convenience yield.

In addition, Brazilian sugar basis illustrates a significant role as a market sign of future macroeconomic expectations and production, trading, storage, hedging and arbitrage opportunities (Bailey and Chan 1993). The signs of the Brazilian sugar basis market can be forecasted using time series models (Szymanowska et al. 2014). In this regard, the constellation of autoregressive and moving average – ARMA (p, q) models, with the variety of inputs may compose robust forecasting techniques for the Brazilian sugar basis. We did not find in literature, research specifically aimed at analyzing the dynamics and forecasting the Brazilian sugar basis.

As such, the research objective of the study is to formulate a Brazilian sugar basis forecast using time series models, comparing between Northeast and Southeast spot and ICE futures markets. Survey questions are i. Identify the significant ARMA (p, q) models to forecast the Brazilian regional sugar basis; ii. Examine basis seasonality patterns to choose a best-fit SARMAX (p, q) model; and, iii.

Compare forecasting errors to estimate the most robust sugar basis time series model, indicating both basis series breakpoints. The results can be applied to increase the efficiency of allocation decisions throughout the Brazilian sugar supply chain.

Figlewski (1984) examined the basis and different sources of baseline risk of the S & P 500 index futures contract. By setting the baseline risk as the risk of unplanned future price variation by spot price variation, the author concluded that a non-systematic hedge of small portfolios and individual stocks was relevant and that the maturity of more than two months affected hedge efficiency.

In turn, Pennings and Meulenberg (1997) described a global risk reduction concept and a new measure of hedge efficiency, focusing on the hedging services offered by futures contracts. Their efficiency measure uses the basis risk, the market size and the distance between the hedging efficiency of the future contract and the optimal hedge. They concluded that the new measure created additional parameters for the efficient management of contracts on futures exchanges.

Frechette (2000) evaluated the demand for hedge operations by examining spatial-based risk. He concluded that the incremental value of offering local hedging instruments to mitigate the basis risk was greater when the demand for hedge was inelastic, and the operating costs were negligible. Also, the result was obtained with a wide basis, and local prices sufficiently detached from stock prices. Thus, hypothetical local future contracts would help to assess potential gains and identify welfare losses resulting from basis risk.

Analogously, Frechette (2001) analyzed the optimal hedge model in terms of operating costs of the hedge. It concluded that the basis and expected basis risk estimates depended on the structural model applied for forecasting, such as the simple, adaptive and rational expectations model, and there was no consensus in the literature on how to identify the most efficient model.

Briys, Chouhy, and Schlesinger (1993) examined the effects of basis risk on hedging strategies, noting that futures contracts were incomplete to provide partial hedging, subject to basis risk. They concluded that the basis risk could be divided into a pure noise effect, which could be analyzed as residual noise and an endogenous effect. Thus, although the basis risk would result in less efficient hedge rates, the hedger could adjust the volumes of the non-perfect hedge instrument.

Castelino (1992) points out that the different sources of risk and the size of the basis were fundamental when choosing the optimal hedge instrument. However, the author shows that once the basis convergence occurred, there was a tendency of the optimal hedge rate toward the unit, the closer it got to maturity.

Dark (2007) evaluated the magnitude of non-inclusion in the convergence specifications of the basis and long memory of volatility on dynamic optimal hedge rates. He concluded that the convergence of the basis became more relevant with longer-term maturity of the hedge. In addition, convergence was important for short-term hedges initiated close to maturity.

However, Garcia, Leuthold, and Sarhan (1984) analyzed the short-term basis risk, defined as the temporal variance of the basis random component for cattle and hogs. They pointed out that the basis risk was related to the factors that impacted the long-term pattern of the time series and the unforeseen price changes. There was little evidence that the basis risk varied with maturity or between markets, except for the specific swine market.

Similarly, several authors have researched basis prediction with multiple approaches. Leuthold and Peterson (1983) used a system of equations for basis, spot prices and futures of the swine market, identifying the relevance of structural components such as storage.

Jiang and Hayenga (1997) similarly applied a structural approach to forecast corn and soybean basis, comparing the results with simple historical averages and econometric models. They emphasized that complex models had better results in the short-term, but results lower than the simple averages in the long term.

Sanders and Manfredo (2006) compared time series models with simplified basis predictions as moving averages for the soybean complex. The more sophisticated models showed better results in the short term, but the forecasts weakened to distant horizons. Also, Tonsor, Dhuyetter, and Mintert (2004) analyzed the number of annual lags for historical averages, combined with the optimal level of updated information for basis forecasting in livestock markets.

Hatchett, Brorsen, and Anderson (2009) examined the optimal sizes of the historical average lags, pointing to a lower number of lags compared to other researches. They concluded that the reason for this might be the changes in the data structure.

Dhuyvetter and Kastens (1998) used historical averages, updated market information and a mix of both to predict the basis of wheat, soybeans, maize, and derivatives. Analogously, short-term forecasts were improved by using more complex models that included up-to-date information. However, historical averages have resulted in better long-term forecasts.

In turn, Taylor, Dhuyvetter, and Kastens (2004) highlighted the relevance of the updated information for short-term basis forecasts. However, they pointed out the reduction of the optimal size of the historical moving average compared to previous results. Recently, Sanders and Baker (2012) applied a soft transition regime (STAR) model to examine weekly forecasts of corn and soy basis. They concluded that the STAR model resulted in better short-term forecasts, although it lost efficiency for the simple moving average models in periods of high volatility.

In summary, the literature on commodity basis forecast is extensive. However, we did not identify specific research to forecast the sugar basis using time series models, in Brazil or other markets. The differential contribution of the article is the identification of time series forecasting models for the sugar basis of the main Brazilian producing regions, which are the Southeast and Northeast regions.

2. Methodology and Data

The methodological steps and data used in the research were inspired by Wayne C. Booth (2003).

2.1 ARMA Modeling for Basis Forecast

The forecasting of stationary time series applies ARMA (autoregressive-moving average) modeling. An ARMA (p, q) model composed of the sum of one autoregressive component, AR (p), and another of moving average, MA (q), is identified by Box, Jenkins, and Reinsel (2008a):

$$\tilde{z}_t = \alpha + \sum_{i=1}^{p} \phi_i \tilde{z}_{t-i} + \sum_{j=1}^{q} \theta_j a_{t-j} + a_t \qquad (1)$$

where $\tilde{z}_t$ is the dependent variable estimated by an ARMA (p, q) model; α is the intercept; ; $\tilde{z}_{t-i}$are autoregressive regressors; a_{t-j}, are moving average regressors; and ϕ_i and θ_j are the respective AR and MA coefficients.

A necessary condition to identify the model associated with the stochastic process of the time series is the stationarity of the ARMA (p, q) model. The stationarity of a time series indicates that the mean, the variance, and the autocorrelations can be approximated by sufficiently long-time averages based on a single set of observations.

The stationarity of a time series is usually tested by applying a unit root test, for example, Augmented Dickey-Fuller-ADF (Dickey and Fuller 1979) and the Phillips-Perron (1988). The hypothesis of the unit root can be used to identify the ARMA model (p, q) applicable to predictions, using the parsimony and predictive efficiency criteria (Box, Jenkins, and Reinsel 2008b). In this research, we tested the unit root of the Southeast and Northeast basis time series.

The ARMA (p, q) models were then selected for the basis in the regional sugar markets using the Akaike (AIC) and Schwarz (SBC), R2, and adjusted R2 criteria as the selection parameters (Enders 2010).

Next, we tested monthly seasonality of both basis, including the significant months in a SARMAX (p, q) model.

Finally, to evaluate the predictive efficiency of the ARMA model (p, q), we calculated: i. MAE Sqr. Root, the square root of Mean Absolute Error; ii. MAE, the Mean Absolute Error; iii. PMAE, the Percent MAE; and, iv. The symmetric PMAE.

Lastly, we compared error results between the ARMA (p, q) and SARMAX (p, q) models, choosing the model with minimum forecasting errors.

From 1930's one of the important issues in setting accounting standards has been the all–inclusive concept of income measurement. A collection of papers related to the debate are present in Brief and Peasnell (1996). Although there has been a long debate on the all–inclusive concept, but little empirical studies have been conducted on the issue.

Rao and Walsh (1999) study the impact of applying the SFAF No. 130 to a sample of 103 Multinational firms from 11 industries for the 1997 fiscal year. The results indicate that the potential effect is that total comprehensive income is lower than the traditional net income number for a majority of firms studied. A majority of the firms are affected negatively by foreign currency translation adjustments.

Dhaliwal, Subramanyam, and Trezevant (1999) investigate the relative ability of comprehensive income and net income to summarize firm performance as reflected in stock returns. They find no evidence that comprehensive income is more strongly associated with returns/market value or better predicts future cash flows/income than net income. Their results do not support the claim that comprehensive income is a better measure of firm performance than net income. They raise questions about the appropriateness of items included in SFAS 130, comprehensive income, as well as, the need for mandating uniform comprehensive income disclosures for all industries.

Maines and McDaniel (2000) study the judgments of nonprofessional investors on different ways of disclosing comprehensive income, i.e., comprehensive income statement and owners' equity. They find that financial–statement format for presenting comprehensive income did not significantly affect nonprofessional investors' acquisition and evaluation of that information, but generally did significantly influence their information weighting and resulting performance judgments.

Cahan, Courtenay, Gronewoller and Upton (2000) study the usefulness of comprehensive income disclosures in a Statement of Changes in Equity (SCE) in New Zealand. Their results suggest that separation of revaluation increments and foreign currency translation adjustments in a SCE are unnecessary. To be exact, they find no evidence that the individual Other Comprehensive Income (OCI) items provide information that is incrementally value relevant above comprehensive income, and they find no evidence that the incremental value relevance of the OCI items relative to net income increased after the SCE was required.

Biddle and Choi (2003) investigate the relevance of comprehensive income for decision-making. Their results reveal that among income definitions, comprehensive income defined by Financial Accounting Standards Board Statement 130, dominates both traditional net income and fully comprehensive income in explaining equity returns, but that net income dominates the more comprehensive measures in explaining chief executive compensation.

Louis (2003) presents an economic analysis of the foreign translation adjustment as another comprehensive income item. He examines the association between change in firm value and the foreign translation adjustment for a sample of manufacturing firms. His study shows that, for firms in the manufacturing sector, the translation adjustment is associated with a loss of value instead of an increase in value.

Kanagaretnam, Mathieu and Shehata (2004) investigate usefulness of reporting comprehensive income in Canada. They examine the association between market value of equity/returns and the components of other comprehensive income to assess the information content of the new disclosures. They, also investigate the predictive ability of the aggregate comprehensive income relative to net income. They provide evidence that each of the four components of other comprehensive income is value relevant in explaining either the market value or the stock returns or both. They find, however, that net income is a better predictor for future firm's performance than aggregate comprehensive income.

Arab Mazar Yazdi and Radmehr (2003) by launching questionnaire ask the opinions of Iranian different financial information users and academics on each item of comprehensive income. They also studied the necessity of reporting such items in separate reports. Their findings indicate that from the respondents' points of view, disclosure of different items of comprehensive income is required in external reporting, but they find it unnecessary to report each item in a separate report.

Mojtahead Zaheh and Momeni (2003) using a questionnaire investigate the effects of comprehensive income statement on users' decision-making. They report that, users of financial information use some measures for management efficiency, investment returns and future cash flows prediction, in their decision-making process. Disclosure of comprehensive income paves the way for evaluation of those measures.

2.2. Data

The sugar monthly prices series were in US$ per 50 kg bag:

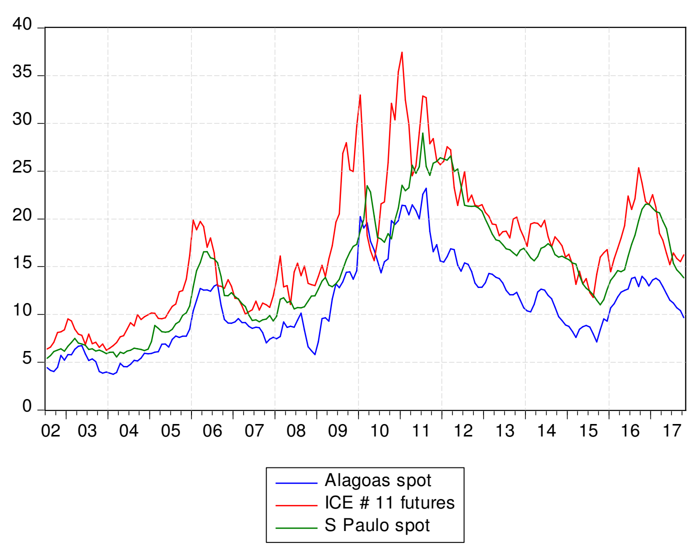
i. for the Northeast spot sugar market we used Alagoas Mercado Interno spot prices, applying a 0.54 multiply coefficient to approximate FOB prices since the internal market prices carry all direct and indirect tax burden; source: CEPEA (2016);
ii. For the Southeast spot sugar market we used São Paulo VHP spot prices; source: CEPEA (2016);
iii. For the futures prices, we used ICE sugar # 11 futures quotes; source: http://www.barchart.com/.

The period analyzed was from May 2002 to October 2017, totaling 186 observations, chosen in a non-probabilistic way, for accessibility and convenience. The choice of the period is justified by the accessibility of the data, as well as the inclusion of the subprime financial crisis in 2008, changing the levels and volatility of commodity prices (Commission, Directorate L. Economic analysis, and Analysis 2009).

3. Results and Discussion

Figure 1 shows Alagoas and São Paulo VHP spot prices and ICE # 11 futures sugar prices in levels:

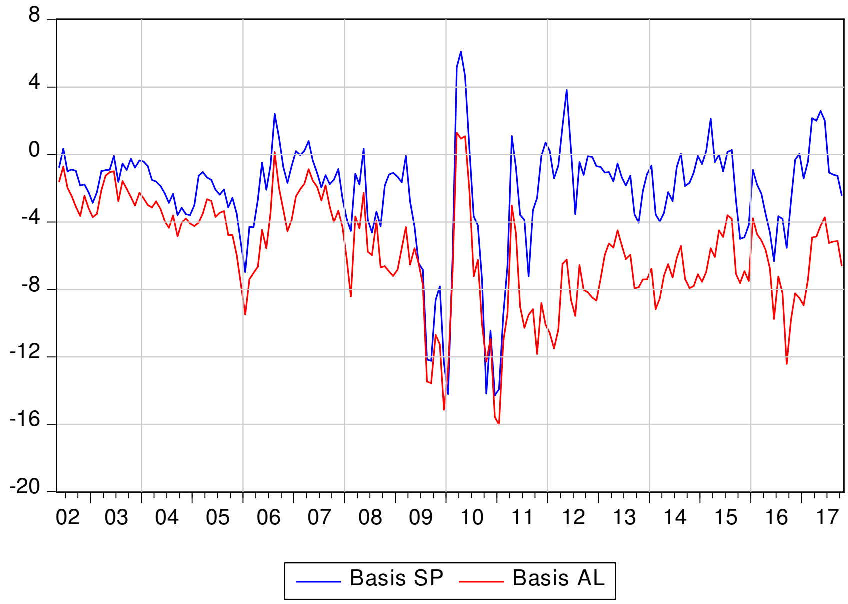
Figure 1: Sugar prices in levels. Alagoas and São Paulo VHP spot prices and ICE # 11 future prices. In US$ per 50kg bag. Period: 2002M05 to 2017M10



Analysis of Figure 1 shows absolute movements with approximate synchronization of Alagoas and São Paulo VHP sugar spot and ICE # sugar future prices. However, we can identify differences between sugar spot and future prices levels, which define the sugar basis dynamics. Therefore, sugar basis forecast is strategic for the supply chain of Brazilian sugar.

Next, we compared the dynamics of the sugar basis in Alagoas and São Paulo. Figure 2 shows the dynamics of the basis in the analyzed period. We identify that the high and low primary movements are synchronous. Thus, widening and narrowing in sugar basis in São Paulo and Alagoas are convergent. Furthermore, there is strong volatility between the years 2009 and 2011, coinciding with the peak sugar prices in 2011, and the fall with subsequent stabilization.

Figure 2: Basis in levels. Sugar prices in Alagoas and SP VHP versus ICE # 11 futures. In US$ per 50kg bag. Period: 2002M05 to 2017M10



In addition, the calculated correlation coefficient between the sugar basis of São Paulo and Alagoas is 0.70. The correlation registers a high positive magnitude, justifying the synchronized dynamics. One possible reason would be the symmetry of sugar pricing processes in Alagoas and São Paulo, for example, initial and final stocks, consumption, production, and common marketing drivers.

Table 1 shows the descriptive statistics of the sugar basis of São Paulo and Alagoas:

Table 1: Descriptive statistics. Monthly sugar basis of Alagoas and São Paulo VHP. Values in R$ per 50 kg bag. Period: 2002M05 to 2017M10.

Statistics	BASIS_AL	BASIS_SP
Average	-5.8731	-2.2835
Median	-5.5705	-1.6009
Maximum	1.2903	6.1100
Minimum	-16.0187	-14.2762
Standard-deviation	3.188443	3.2585
Asymmetry	-0.5624	-1.4632
Kurtosis	3.4362	6.7570
Jarque-Bera	11.1572	173.8725
Probability	0.0038	0.0000
Observations	184	184

The sugar basis of Alagoas registers the lowest mean, median, maximum and minimum values, standard deviation and kurtosis. However, the Alagoas basis illustrates the greater asymmetry and distance from normality. One possible explanation for the higher values would be the scale of production as well as the sea freight structure that depends on the supply and demand, higher in the sugar market of São Paulo, differently impacting the pricing of the Southeast and Northeast sugar markets.

In addition, the standard deviation, proxy for basis risk, registers the value of 3,188 and 3,259, for Alagoas and São Paulo, respectively. Thus, no significant difference between the sugar basis risks of the two regions is pointed out. Consequently, the use of ICE # 11 sugar futures contracts would define price risk mitigation strategies, incurring hedges in the basis risk identified in the results.

The next step tests the hypothesis of the existence of unit root in the sugar basis series in Alagoas and São Paulo, using the Phillips-Perron (PP) and Augmented Dickey-Fuller (ADF) models:

Table 2: Unit root tests (RU). Phillips-Perron (PP) and Augmented Dickey-Fuller (ADF) test. Model with constant. Sugar basis of Alagoas and São Paulo, level. Monthly values in US$ per 50 kg bag

Values	**PP**	**ADF**
	Basis AL	
t-Statistic	-4.5101*	-4.3380*
Prob.	***0.0003***	***0.0005***
	Basis SP	
t-Statistic	-4.7252*	-5.8229*
Prob.	***0.0001***	***0.0000***

The results of Table 2 suggest that the hypothesis of the existence of unit root at a significance level of 1% should be rejected. Therefore, the series of the sugar basis of São Paulo and Alagoas register stationarity, so we can apply the technique of time series forecasting, ARMA (Box, Jenkins, and Reinsel 2008b). Additionally, we used the monthly seasonality of the sugar basis identifying adjusted original ARMA models. As such, the additional forecast for Alagoas and São Paulo sugar basis models resulted in SARMAX - Seasonal ARMA with Exogenous Input.

Table 3 presents the monthly seasonality of the Alagoas and São Paulo sugar basis. Table 3 shows that the seasonality of the São Paulo sugar basis is statistically significant in the months of January-February and June-December. In addition, there is statistically significant seasonality of the Alagoas sugar basis in all months of the year, from January to December. In turn, the Alagoas sugar basis shows significant seasonal coefficients for all twelve months.

Table 3: Seasonality test of the Alagoas and São Paulo sugar basis. Monthly values

Month	**Basis SP**	**Basis AL**
January	-3.3614*	-7.2851*
February	-2.5521*	-6.5250*
March	-0.9330	-4.8957*
April	-0.5623	-4.2020*
May	-0.4085	-3.8170*
June	-2.0247**	-5.2815*
July	-2.5124*	-5.4936*
August	-2.3774*	-5.0645*
September	-3.2982*	-6.7091*
October	-3.3323*	-6.9786*
November	-2.6897*	-6.6625*
December	-3.1911*	-7.5121*

Next, the ARMA model for the São Paulo sugar basis is identified, shown in Table 4:

Table 4: Characterization of the ARMA model (2, 0) to forecast the monthly sugar basis. Spot prices for São Paulo VHP and future prices of ICE

	Coefficient	Standard Errors	Test-t	Prob.
C	-2.2797	0.6560	-3.4752	0.0006
AR(1)	1.0009	0.0547	18.2974	0.0000
AR(2)	-0.2727	0.0453	-6.0251	0.0000
R^2	0.6508	**Akaike info criteria**		4.1925
Adjusted R^2	0.6450	**Schwarz info criteria**		4.2624

Table 4 shows an ARMA (2, 0), with all coefficients statistically significant at 1%. A strong and positive magnitude of the AR (1) is illustrated, underlining the autoregressivity of the series, which is counterbalanced by the negative values of the AR (2) and the constant. Also, high values of adjusted R2 and R2 are recorded, around 0.650. Thus, the goodness of fit of the model is high. The next step defines the SARMAX model of the São Paulo sugar basis, Table 5.

The exogenous intervention analyzed the statistical significance of the elements of the original ARMA model (2, 0), including the significant monthly seasonality, indicated in Table 3. After the inclusion of the seasonality, the constant was excluded. Table 5 shows the statistical magnitudes and significance of the SARMAX (2, 0) elements. Specifically, approximately 5% improvement of R2 and R2 adjusted, approximate values of 0.680, of the SARMAX (2, 0) results relative to the ARMA (2, 0) is reported.

Table 5: Characterization of the SARMAX model (2, 0) to forecast the monthly sugar basis. Spot prices for São Paulo VHP and future prices of ICE

	Coefficient	Standard Errors	Test-t	Prob.
January	-2.7925	0.7920	-3.5257	0.0005
February	-1.7579	0.5236	-3.3569	0.0010
June	-1.7164	0.6319	-2.7161	0.0073
July	-2.3576	0.8465	-2.7850	0.0060
August	-2.2549	0.8330	-2.7069	0.0075
September	-3.1752	0.8528	-3.7235	0.0003
October	-3.1756	0.8704	-3.6487	0.0003
November	-2.4243	0.9711	-2.4963	0.0135
December	-2.8074	0.9485	-2.9597	0.0035
AR(1)	0.9948	0.0666	14.933	0.0000
AR(2)	-0.2421	0.0569	-4.2521	0.0000
R^2	0.6928	**Akaike info criteria**		4.1514
Adjusted R^2	0.6731	**Schwarz info criteria**		4.3611

Table 6 compares the forecast errors between the ARMA (2, 0) and SARMAX (2, 0) models of the São Paulo sugar basis:

Table 6: Comparison of forecast errors between the ARMA (2, 0) and SARMAX (2, 0) models of the São Paulo sugar basis. Forecast within the sample. Period 2015M11 and 2017M10, totaling 24 observations

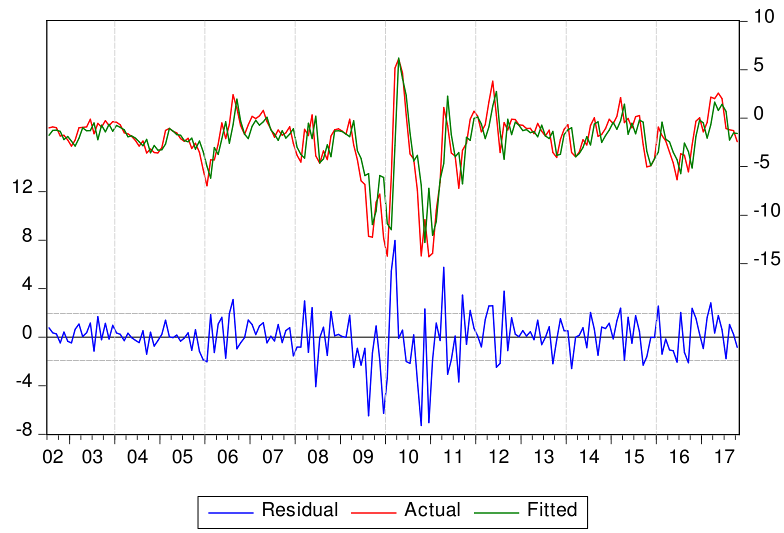
Model	MAE Square Root[1]	MAE[2]	PMAE[3]	Theil Coefficient		Symmetric PMAE[4]
				Inequality	U2	
ARMA (2, 0)[a]	2.440	1.9325	299.8847	0.4285	0.7506	87.6089
SARMAX (2, 0)[b]	2.942	1.9632	337.5735	0.4131	0.9982	105.5178
Difference (b/a)	20.6%	1.6%	12.6%	-3.6%	33.0%	20.4%

The values of the MAE Square Root, MAE, PMAE, Theil's Inequality, and U2 Coefficients, and the Symmetric PMAE of the ARMA (2, 0) and SARMAX (2, 0) models of the São Paulo sugar basis are shown in Table 6. SARMAX (2, 0) errors are higher, except for Theil's Inequality Coefficient. However, the adjusted R2 and R2 values of SARMAX (2.0) are five percent higher than those of ARMA (2, 0).

However, using the criterion of use of the model with minimum errors, the application of the ARMA (2, 0) model was used to forecast the São Paulo sugar basis.

As such, Figure 3 maps the results of the ARMA (2, 0) forecast model and the dynamics of the residues:

Figure 3: Results of ARMA (2, 0) forecast model and residues dynamics. Sugar basis for São Paulo. Forecast within the sample. Period 2015M11 and 2017M10, totaling 24 observations. Source: Research data



The analysis of Figure 3 indicates the fluctuation band of the ARMA (2, 0) model except between 2009 and 2011, coinciding the peak and fall in sugar prices. Thus, in this period it is possible to see an increase in the basis volatility, with convergence after the stabilization of sugar prices. Thus, the residuals converge within the previous fluctuation band, indicating the applicability for the São Paulo sugar basis using the ARMA (2, 0) model. In relation to the Alagoas sugar basis, the ARMA (1, 1) model is identified, Table 7:

Table 7: Characterization of the ARMA model (1, 1) to predict the monthly sugar basis. Prices for São Paulo VHP and future prices of ICE

Coefficient	Standard Errors	Test-t	Prob.	Coefficient
C	-5.8203	0.7020	-8.2914	0.0000
AR(1)	0.7614	0.0526	14.480	0.0000
MA(1)	0.1947	0.0734	2.6516	0.0087
R^2	0.6862	**Akaike info criteria**		4.0424
Adjusted R^2	0.6801	**Schwarz info criteria**		4.1123

Table 7 defines an ARMA (1, 1) model, with all coefficients statistically significant at 1%. A high positive magnitude of the AR (1) term is highlighted, indicating the strong autoregressivity of the Alagoas basis series. Unlike the ARMA model for the São Paulo sugar basis, a MA (1) linear error component is expressed. Also, high values of R2 and adjusted R2 are recorded, around 0.68. Thus, the quality of fit degree of the ARMA (1, 1) for the Alagoas sugar basis time series model is high.

The next step defines the SARMAX model of the Alagoas sugar basis, Table 8:

Table 8: Characterization of the SARMAX model (1, 0) model to predict the monthly sugar basis

	Coefficient	Standard Errors	Test-t	Prob.
January	-7.174	1.0726	-6.6885	0.0000
February	-6.3982	1.0325	-6.1970	0.0000
March	-4.7450	0.9483	-5.0087	0.0000
April	-4.0326	0.9666	-4.1719	0.0000
May	-3.6192	1.0891	-3.3231	0.0011
June	-5.0495	1.0259	-4.9218	0.0000
July	-5.4561	0.9433	-5.7843	0.0000
August	-5.0219	0.8941	-5.6168	0.0000
September	-6.6600	0.9324	-7.1431	0.0000
October	-6.9216	1.0078	-6.8679	0.0000
November	-6.5728	1.1600	-5.6664	0.0000
December	-7.4130	1.1143	-6.6524	0.0000
AR(1)	0.8429	0.0438	19.2245	0.0000
R^2	0.7516	**Akaike info criteria**		3.9178
Adjusted R^2	0.7326	**Schwarz info criteria**		4.1624

Thus, the exogenous intervention examined the statistical significance of the elements of the original ARMA (1, 1) model including the significant monthly seasonality, Table 3. After the inclusion of seasonality's, the constant and the MA (1) term were excluded. Table 8 shows the magnitudes and statistical significance of the elements of the resulting model, SARMAX (1, 0). A There was an approximate 9% improvement in the adjusted R2 and R2, approximate values of 0.740, of the SARMAX (1, 0) results relative to the ARMA (1, 1) are reported.

Table 9 compares the forecast errors between the ARMA (1, 1) and SARMAX (1, 0) models of the Alagoas sugar basis:

Table 9: Comparison of prediction errors between the ARMA (1, 1) and SARMAX (1, 0) models of the Alagoas sugar basis. Forecast within the sample. Period 2015M11 and 2017M10, totaling 24 observations

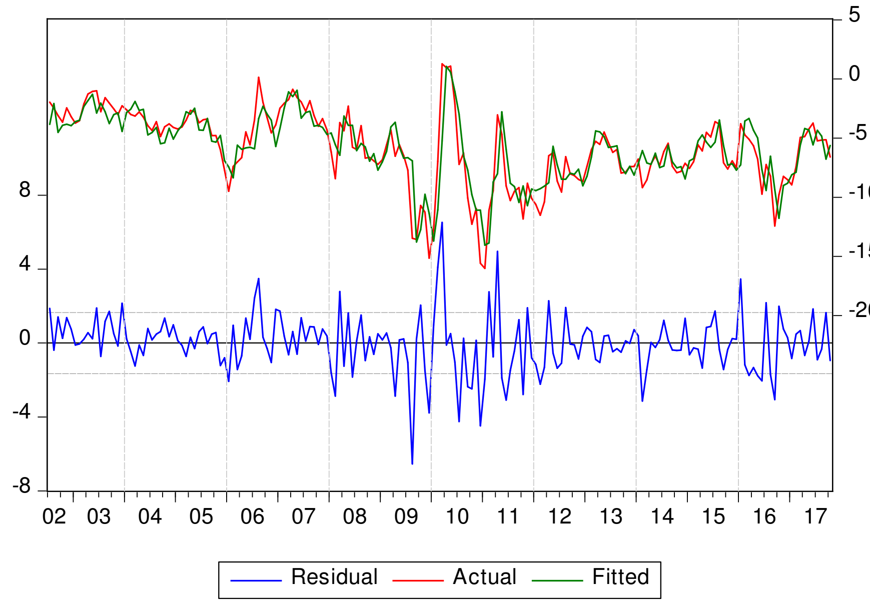
Model	MAE Square Root[1]	MAE[2]	PMAE[3]	Theil Coefficient		Symmetric PMAE[4]
				Inequality	U2	
ARMA (1, 1)[a]	2.3025	1.8137	26.6747	0.1755	1.3112	27.1766
SARMAX (1, 0)[b]	2.1831	1.6211	23.9929	0.1669	1.3091	24.6708
Difference (b/a)	-5.2%	-10.6%	-10.1%	-4.9%	-0.2%	-9.2%

Table 9 records the values of the MAE Square Root, MAE, PMAE, Theil's Inequality and U2 Coefficients, and the Symmetric PMAE of the ARMA (1, 1) and SARMAX (1, 0) forecasting models for the Alagoas sugar basis. In line with the adjusted R2 and R2, lower forecast errors are shown for the SARMAX (1, 0) model compared to the ARMA (1, 1) model.

Therefore, if we apply the criterion of use of the minimum errors the SARMAX (1, 0) model can be chosen to forecast the Alagoas sugar basis.

Figure 4 below shows the forecasting results and residuals of the SARMAX (1, 0) model for the Alagoas sugar basis:

Figure 4: Results of the SARMA (1, 0) forecast model and residues dynamics. Alagoas sugar basis. Forecast within the sample. Period 2015M11 and 2017M10, totaling 24 observations



The analysis of Figure 4 shows the residual bandwidth of the SARMAX (1, 0) model, which is uniformly distributed except between 2009 and 2011, coinciding with the peak and fall in sugar prices. The result is analogous to Figure 2, for the ARMA (2, 0) model residues for the São Paulo sugar basis forecast model. Thus, in that period the increase in basis volatility was pointed out, with convergence after the stabilization of sugar prices. Therefore, the residuals converge again within the previous fluctuation band, pointing to the forecasting applicability of the Alagoas sugar basis using the SARMAX (1, 0) model.

Finally, by applying the Quandt-Andrews test for structural breaks, we identify the dates of the monthly breaks of the sugar basis series of Alagoas and São Paulo, Table 10:

Table 10: Quandt-Andrews test for unknown structural breaks1. Dates of the breaks of the sugar basis series of Alagoas and São Paulo. Period: 2002M05 to 2017M10.

Sugar basis	Structural break month	Quandt-Andrews statistics	Value	p-value
Alagoas	2008M09	Maximum LR F-statistic	95.5403	0.0000
		Maximum Wald F-statistic	95.5403	0.0000
São Paulo	2011M11	Maximum LR F-statistic	11.0470	0.0162
		Maximum Wald F-statistic	11.0470	0.0162

The two dates indicated, 2008M09 and 2011M11, the structural breaks in the sugar basis series of Alagoas and São Paulo respectively. They coincide with the increase in sugar price volatility. The increase in volatility of the residues dynamics is expressed in Figures 3 and 4.

In sum, the forecast of São Paulo and Alagoas spot and ICE # 11 sugar monthly basis show different models classified by the error minimization criteria. As such, for the São Paulo sugar monthly basis forecast, we can apply an ARMA (2, 0) model. For the Alagoas sugar monthly basis forecast, we can use a SARMAX (1, 0) model. The ARMA and SARMAX models formulate a strategic mechanism for assessing the Brazilian sugar basis relative to ICE sugar # 11 futures. In addition, both models are easy to analyze and implement, resulting in an improved informational input for a thorough understanding of basis relationships and patterns over time and space.

To conclude, the Brazilian sugar agents can use ICE sugar # 11 calculating local sugar basis, resulting in efficient allocative decisions. To illustrate, a forecasted weak basis estimates that a Brazilian sugar agent would benefit from using an open-basis contract, or a hedge using ICE # 11 futures waiting for the basis to improve over time. On the other hand, a forecasted strong sugar basis, which is more probable to weaken by delivery time, suggests using a deferred delivery contract or a basis contract to lock in the forecasted strong basis value. Hence, sugar basis dynamics forecast using the ARMA framework is strategic, resulting in lower cost-benefit alternatives increasing the efficiency of production, storage, commercialization, and hedging decisions.

4. Summary and Conclusions

Sugar basis identification indicates a strategic tool for efficient decision making for the sugar supply chain agents of Brazil. However, we did not find in the literature, aimed at analyzing the dynamics and forecast the Brazilian sugar basis. Therefore, this study formulates a Brazilian sugar basis forecast using time series models, comparing Northeast and Southeast spot and ICE futures markets.

Specifically, we identify the significant ARMA (p, q) models to forecast the monthly Brazilian regional sugar basis. Also, we examine basis seasonality patterns to choose a best-fit SARMAX (p, q) model, comparing the forecasting errors to estimate the most robust sugar basis time series model, besides indicating both basis series breakpoints.

We identify that both sugar basis in Alagoas and São Paulo show synchronous primary movements and high volatility. In addition, the correlation coefficient between the sugar basis, 0.70, illustrates a high and positive magnitude, confirming the cross-effects between both cities' sugar basis. Descriptive statistics for Alagoas sugar basis are higher than those for São Paulo.

Both Alagoas and São Paulo sugar basis are stationary series, I (0), which can be forecast using the ARMA (p, q) model. Also, both cities' sugar basis illustrates strong monthly seasonality with different periods. Specifically, Alagoas sugar basis shows seasonality for all twelve months, whereas São Paulo basis, for nine months, January and February, and June through December. The knowledge of the sugar basis seasonality patterns can enhance efficient decision-making in the Brazilian sugar supply chain.

After comparing sugar basis in sample forecasting errors between ARMA (p, q) and SARMAX (p, q) models, Alagoas monthly sugar basis robust forecast model is a SARMAX (1, 0). For the São Paulo monthly sugar basis, the robust forecast model is an ARMA (2, 0). In addition, the Alagoas sugar basis breakpoint month is 2008M09, and the São Paulo sugar basis breakpoint month is 2011M11. The breakpoint months coincide with the identifiable sugar basis level and volatility trends abnormal trajectory in the examined period.

Therefore, the ARMA (p, q) time series model describes a strategic management tool to predict future values of the Brazilian sugar basis, in the Southeast and Northeast markets, which is easy to use and implement. Sugar basis defines multiple applications regarding production, trading, storage, expected risk and hedging for the Brazilian sugar supply chain. Future research could analyze basis dynamics between the monthly breakpoints, or a shorter timeframe, e.g., weekly basis. The application of non-linear forecasting models for sugar basis forecasting, such as the GARCH family, and the analysis of the Brazilian sugar basis risk are relevant research issues.

References

[1] Bailey, Warren, and K. C. Chan. 1993. "Macroeconomic Influences and the Variability of the Commodity Futures Basis." The Journal of Finance 48 (2). Wiley/Blackwell (10.1111): 555–73. https://doi.org/10.1111/j.1540-6261.1993.tb04727.x.

[2] Booth, Wayne C. 2003. The Craft of Research. Linear Algebra and Its Applications. Vol. 422. https://doi.org/10.1016/j.laa.2006.10.019.

[3] Box, George E. P., Gwilym M. Jenkins, and Gregory C. Reinsel. 2008a. Time Series Analysis : Forecasting and Control. John Wiley.

[4] ———. 2008b. Time Series Analysis. Hoboken, NJ: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118619193.

[5] Briys, Eric, Michel Crouhy, and Harris Schlesinger. 1993. "Optimal Hedging in a Futures Market with Background Noise and Basis Risk." European Economic Review 37 (5). North-Holland: 949–60. https://doi.org/10.1016/0014-2921(93)90103-H.

[6] Castelino, Mark G. 1992. "Hedge Effectiveness: Basis Risk and Minimum-variance Hedging." Journal of Futures Markets 12 (2). Wiley-Blackwell: 187–201. https://doi.org/10.1002/fut.3990120207.

[7] CEPEA. 2016. "Centro de Estudos Avançados Em Economia Aplicada." Imagenet Tecnologia. 2016. https://www.cepea.esalq.usp.br/br.

[8] Commission, European, perspectives and evaluations Directorate L. Economic analysis, and L.5. Agricultural trade policy Analysis. 2009. "Historical Price Volatility." https://ec.europa.eu/agriculture/trade-analysis_en.

[9] Dark, Jonathan. 2007. "Basis Convergence and Long Memory in Volatility When Dynamic Hedging with Futures." Journal of Financial and Quantitative Analysis 42 (04). Cambridge University Press: 1021. https://doi.org/10.1017/S0022109000003483.

[10] Dhuyvetter, Kevin C, and Terry L Kastens. 1998. "Forecasting Crop Basis : Practical Alternatives." http://www.farmdoc.uiuc.edu/nccc134].

[11] Dickey, David A., and Wayne A. Fuller. 1979. "Distribution of the Estimators for Autoregressive Time Series With a Unit Root." Journal of the American Statistical Association 74 (366). Taylor & Francis, Ltd.American Statistical Association: 427. https://doi.org/10.2307/2286348.

[12] Enders, Walter. 2010. Applied Econometric Time Series. Wiley.

[13] Figlewski, Stephen. 1984. "Hedging Performance and Basis Risk in Stock Index Futures." The Journal of Finance 39 (3). Wiley American Finance Association: 657. https://doi.org/10.2307/2327924.

[14] Frechette, Darren L. 2000. "The Demand for Hedging and the Value of Hedging Opportunities." American Journal of Agricultural Economics 82 (4). Oxford University PressAgricultural & Applied Economics Association: 897–907. https://doi.org/10.1111/0002-9092.00089.

[15] ———. 2001. "The Demand for Hedging with Futures and Options." Journal of Futures Markets 21 (8). Wiley-Blackwell: 693–712. https://doi.org/10.1002/fut.1801.

[16] Garcia, Philip, R.M. Leuthold, and M.E. Sarhan. 1984. "Basis Risk: Measurement and Analysis of Basis Fluctuations for Selected Livestock Markets." Agricultural & Applied Economics Association 66 (4). Oxford University Press: 499–504. https://doi.org/10.2307/1240929.

[17] Hatchett, Robert B., B. Wade Brorsen, and Kim B. Anderson. 2009. "Optimal Length of Moving Average to Forecast Futures Basis." Journal of Agricultural and Resource Economics 35. Western Agricultural Economics Association: 18–33. https://doi.org/10.2307/23243034.

[18] Jiang, Bingrong, and Marvin Hayenga. 1997. "Corn and Soybean Basis Behavior and Forecasting : Fundamental and Alternative Approaches by Bingrong Jiang and Marvin Hayenga," no. June 2003. http://lib.dr.iastate.edu/rtd.

[19] Leuthold, Raymond M., Joan C. Junkus, and Jean E. Cordier. 1989. The Theory and Practice of Futures Markets. Stipes Pub. L.L.C.

[20] Leuthold, Raymond M., and Paul E. Peterson. 1983. “The Cash-Futures Price Spread for Live Hogs.” North Central Journal of Agricultural Economics 5 (1). Oxford University PressAgricultural & Applied Economics Association: 25. https://doi.org/10.2307/1349189.

[21] Pennings, Joost M.E., and Matthew T.G. Meulenberg. 1997. “Hedging Efficiency: A Futures Exchange Management Approach.” Journal of Futures Markets 17 (5). Wiley-Blackwell: 599–615. https://doi.org/10.1002/(SICI)1096-9934(199708)17:5<599::AID-FUT5>3.0.CO;2-A.

[22] Phillips, Peter C. B., and Pierre Perron. 1988. “Testing for a Unit Root in Time Series Regressions.” Biometrika 75 (2). Oxford University PressBiometrika Trust: 335–46. https://doi.org/10.2307/2336182.

[23] Sanders, Daniel J, and Timothy G Baker. 2012. “Forecasting Corn and Soybean Basis Using Regime-Switching Models By.” https://pdfs.semanticscholar.org/bbbb/05e258fcc31703c65c2c70f308eed05941b1.pdf.

[24] Sanders, Dwight R., and Mark R. Manfredo. 2006. “Forecasting Basis Levels in the Soybean Complex: A Comparison of Time Series Methods.” Journal of Agricultural and Applied Economics 38 (03). Cambridge University Press: 513–23. https://doi.org/10.1017/S1074070800022586.

[25] Szymanowska, Marta, Frans De Roon, Theo Nijman, and Rob Van Den Goorbergh. 2014. “An Anatomy of Commodity Futures Risk Premia.” Journal of Finance 69 (1). Wiley/Blackwell (10.1111): 453–82. https://doi.org/10.1111/jofi.12096.

[26] Tonsor, Glynn T., Kevin C. Dhuyvetter, and James R. Mintert. 2004. “Improving Cattle Basis Forecasting.” Journal of Agricultural and Resource Economics 29 (2). Western Agricultural Economics Association: 228–41. https://ideas.repec.org/a/ags/jlaare/31115.html.

[27] USDA United States Department of Agriculture. 2017. “Record Global Production Spurs Record Consumption Sugar Overview.” Sugar: World Markets and Trade. https://public.govdelivery.com/accounts/USDAFAS/subscriber/new.