SEVEN publicações acadêmicas

ETHANOL IN BRAZIL: A PREDICTIVE PRICING APPROACH

ETANOL NO BRASIL: UMA ABORDAGEM PREDITIVA DE PREÇOS

ETANOL EN BRASIL: UN ENFOQUE DE PRECIOS PREDICTIVOS

https://doi.org/10.56238/sevened2025.029-058

Fabrício Bisset Silva de Brito[1], Leandro Brito Santos[2], Roberto Luiz Souza Monteiro[3], Everaldo Freitas Guedes[4]

ABSTRACT

The production of liquid biofuels has emerged as an alternative to the use of fossil fuels for transportation purposes. Ethanol is described as a renewable energy fuel. In 2021, Brazil produced 27.46% of the world's ethanol, ranking second among producing countries. The introduction of light commercial and passenger vehicles with Flex Fuel technology, which uses both Type C gasoline and hydrous ethanol, into the Brazilian market, began in March 2003, significantly impacting ethanol demand in the Brazilian consumer market. This study aims to compare the predictive accuracy of the ARFIMA, ARIMA, and Smoothed Exponential models for Brazilian ethanol prices over a four-year period.

Keywords: Ethanol. Commodities. Forecasting. Time Series.

RESUMO

A produção de biocombustíveis líquidos emergiu como alternativa ao uso de combustíveis fósseis para fins de transporte. O etanol é descrito como um combustível de energia renovável. O Brasil em 2021 produziu 27,46 % do etanol no mundo, ocupando a segunda posição entre os países produtores. A introdução no mercado brasileiro dos veículos comerciais ligeiros e de passageiros com a tecnologia Flex Fuel na qual, utiliza simultaneamente Gasolina tipo C ou Etanol hidratado. A partir de março 2003, impactou significativamente na demanda pelo etanol no mercado consumidor brasileiro. O presente trabalho propõe realizar comparação da acurácia preditiva entre os modelos matemáticos ARFIMA, ARIMA e Exponencial Suavizado dos preços do etanol brasileiro no período de 4 anos.

Palavras-chave: Etanol. Commodities. Previsão. Séries Temporais.

RESUMEN

La producción de biocombustibles líquidos ha surgido como una alternativa al uso de combustibles fósiles para el transporte. El etanol se describe como un combustible de energía renovable. En 2021, Brasil produjo el 27,46 % del etanol mundial, ocupando el

[1] Master's Degree in Computational Modeling and Industrial Technology. Universidade SENAI CIMATEC.
E-mail: fabissetc@gmail.com Orcid: https://orcid.org/0009-0004-0411-6331
[2] Dr. in Computational Modeling and Industrial Technology. Universidade SENAI CIMATEC.
E-mail: lsbrito@gmail.com Orcid: https://orcid.org/0000-0003-0132-4712
[3] Dr. in Knowledge Diffusion. Universidade Federal da Bahia (UFBA). E-mail: robertolsmonteiro@gmail.com
Orcid: https://orcid.org/0000-0002-3931-5953
[4] Dr. in Computational Modeling and Industrial Technology. Universidade SENAI CIMATEC.
E-mail: efgestatistico@gmail.com Orcid: https://orcid.org/0000-0002-2986-7367

segundo lugar entre los países productores. La introducción en el mercado brasileño de vehículos comerciales ligeros y de pasajeros con tecnología Flex Fuel, que utiliza tanto gasolina tipo C como etanol hidratado, comenzó en marzo de 2003, lo que impactó significativamente la demanda de etanol en el mercado brasileño. Este estudio busca comparar la precisión predictiva de los modelos ARFIMA, ARIMA y Exponencial Suavizado para los precios del etanol brasileño durante un período de cuatro años.

Palabras clave: Etanol. Materias Primas. Pronóstico. Series Temporales.

1 INTRODUCTION

The production of liquid biofuels has emerged as an alternative to the use of fossil fuels for transportation purposes [1]. Carbon emission mitigation, energy security and agricultural development are the main drivers of this type of bioenergy projects [1]. Ethanol is described as a renewable energy fuel that helps mitigate climate change, highlighting technological advances that could enable the production of not only traditional ethanol but also second-generation cellulosic ethanol from sugarcane bagasse [2].

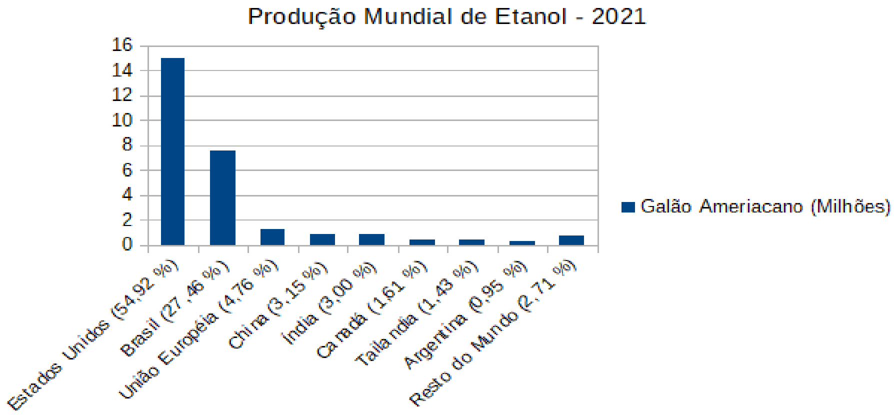
According to data from the Renewable Fuel Association (RFA), Brazil in 2021 produced 27.46% of the ethanol in the world, thus occupying the second position among producing countries (Figure 1) [3]. In Brazil, fuel ethanol is derived from sugarcane and is used pure or blended with gasoline in a mixture called gasohol (25% ethanol, 75% gasoline) [4].

The introduction and evolution in the Brazilian market of light commercial and passenger vehicles with Flex Fuel technology, which simultaneously uses type C gasoline or hydrous ethanol, as of March 2003, significantly impacted the demand for ethanol in the Brazilian consumer market and quantitatively the variation (the largest) of its price compared to type C gasoline [5]. Consequently, the introduction of Flex Fuel technology in Brazil, together with the increase in demand for sugar in the international market, has driven significant growth in the sugarcane industry in recent decades [6].

Also, according to data from the National Association of Motor Vehicles (ANFAVEA), in 2021 the number of registrations of motor vehicles and light commercial vehicles, with Flex Fuel technology engines, corresponded to 82.20% of the total vehicles registered in the Brazilian market (Figure 2) [7].

Figure 1

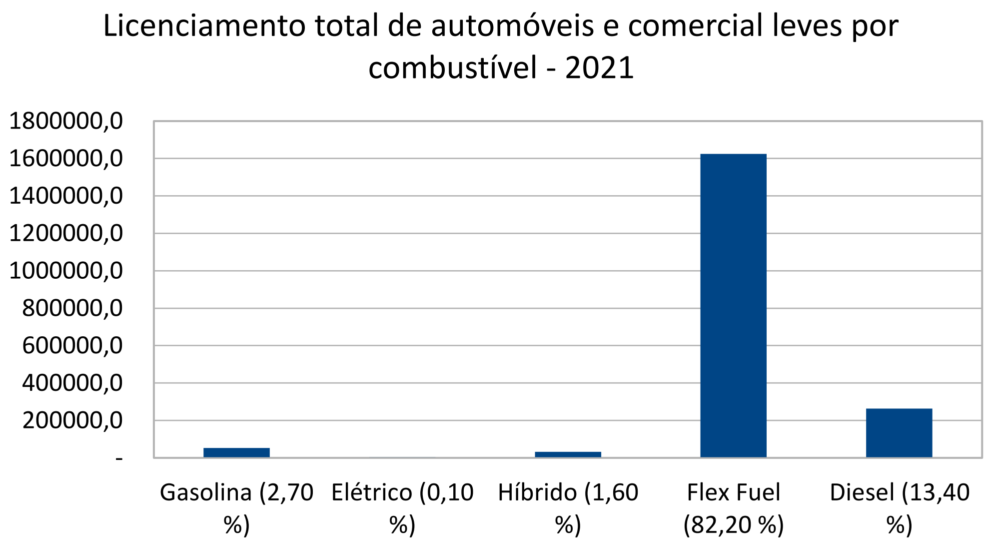
Ethanol production in the world in 2021



Source: RFA.

Figure 2

Full licensing of automobiles and light commercial vehicles by fuel in 2021



Source: ANFAVEA / RENAVAM

According to studies by [8] the volatility of ethanol prices in Brazil in consecutive years is mainly associated with the following factors: (i) amount of sugarcane production; (ii) percentage of sugarcane for ethanol production, i.e., for the production mix; (iii) consumer income; (iv) number of vehicles in the light commercial fleet; and (v) the price of gasoline in view of the compulsory blending of this type of ethanol in the sale of gasoline. To better understand the problem of volatility in Brazilian ethanol prices, [8] they applied several mathematical tools in their work, namely: Untrended Fluctuation Analysis (DFA), the exponents of Hurst (H) and Lyapunov (λ); as well as compare the accuracy between the Fractional Integrated Moving Average (ARFIMA) and Autoregressive Moving Average (ARIMA) models in a 365-day forecast of ethanol prices in a time series, presenting, as a result, an advantage of the ARFIMA model over ARIMA.

Thus, the present work proposes, specifically, to perform a similar comparison of the accuracy between the ARFIMA and ARIMA models, in a forecast for the next day (next point), in a time series of ethanol prices in Brazil; however, introducing a new element to the experiment performed by [8]: a Smoothed Exponential Mathematical Model.

2 METHODOLOGY

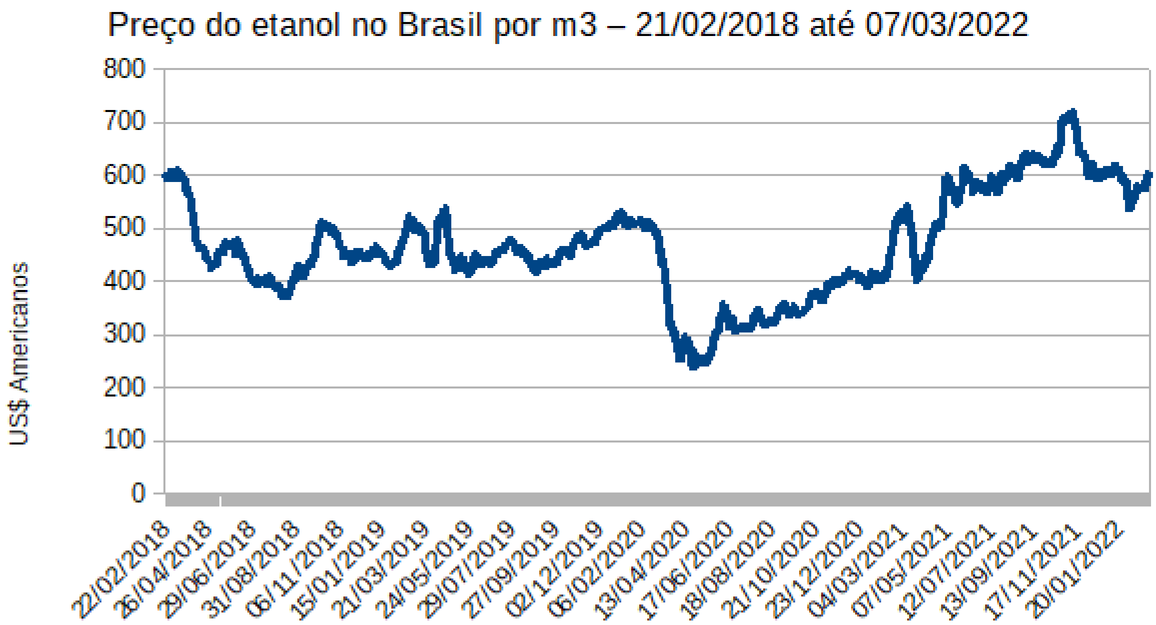
The methodology refers to the application of the mathematical models listed in the following sub-items in the database through the programming language and integrated development environment R [9].

2.1 DATABASE

The database used in the experiment is the time series of daily ethanol prices in US Dollars (US$), referring to the period from 02/21/2018 to 03/07/2022 (thousand points), provided by the Center for Advanced Studies in Applied Economics – CEPEA-Esalq / USP and available at [10]

Figure 3

Ethanol price in Brazil per m3 – 21/02/2018 to 07/03/2022



Source: CEPEA / ESALQ / USP

2.2 ARIMA MODEL

Proposed by [11], the ARIMA mathematical model is used in the analysis and prediction of time series and has as a form of ARIMA notation (*p*, *d*, *q*); where *p* represents the orders of autoregression, *d* represents the integration and *q* represents the order of the moving average.

To apply the ARIMA mathematical model in the time series under study, the *auto.arima ()* function available in the *R language forecast* library was used. According to the documentation of the language in question, the *auto.arima ()* function returns the best ARIMA model according to the AIC, AICc or BIC value. The function conducts a search on the possible model within the given order constraints [9].

2.3 ARFIMA MODEL

The ARFIMA model is a generalization of the ARIMA mathematical model and is characterized by the ability to model processes with long serial dependence[12].

To apply the ARFIMA mathematical model in the time series under study, the *arfima()* function available in the *forecast library* of the R language was used. According to the documentation of the language in question, an optimal ARFIMA model (p, d, q) is selected and estimated automatically [9].

2.4 SMOOTHED EXPONENTIAL MODEL

The design of Exponential Smoothing models involves attaching larger weights to the most recent observations, and the predictions are calculated using weighted averages where the weights decrease – exponentially – in the older observations in a time series [13].

To apply a Smoothed Exponential mathematical model in the time series under study, the ets*()* function available in the *forecast* library of the R language was used. The only argument required for the *ets()* function is the time series. An optimal Smoothed Exponential model is automatically chosen if it is not specified [9].

3 RESULTS AND DISCUSSION

The results were divided into (1) accuracy of the models used in the experiment and (2) prediction of the 1001st point per model.

3.1 ACCURACY OF THE MODELS USED

From the application of the mathematical models ARFIMA, ARIMA and Exponential smoothed in the database object of this work, we obtained the accuracy averages of each model, namely: Mean Error (ME), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE), Mean Absolute Escalated Error (MASE) and Self-Correlation of Errors in Delay 1, as shown in Table 1.

Table 1

Measures of accuracy of prediction models

MODEL	ME	RMSE	MAE	MEP	MAP	MASE	ACF1
ARFIMA	0.01956952	6.386936	4.844744	-0.01488094	1.068626	0.9439653	-0.001305478
ARIMA (2,1,0)	0.0131833	6.393918	4.842453	0.001978034	1.067812	0.943519	1.43362e-05
EXP. SOFTENED	0.007449302	6.425317	4.868707	0.005221073	1.073333	0.9486344	-0.004736487

Source: experiment.

3.2 PREDICTION OF 1001ST POINT BY MODEL

Table 02 presents the prediction of the 1001st point by mathematical model used in the experiment, with confidence intervals of 80% and 95%.

Table 2

Prediction of the 1001st point by mathematical model

MODEL	POINT	FORECAST	LOW 80	HIGH 80	LOW 95	HIGH 95
ARIMA (2,1,0)	1001	610.8624	602.6559	619.0688	598.3117	623.4131
EXP. SMOOTHED	1001	611.8587	603.6036	620.1137	599.2337	624.4837
ARFIMA	1001	610.3783	602.1931	618.5635	597.8601	622.8965

Source: experiment.

4 CONCLUSION

In the present experiment, the statistical model that presented the best predictive accuracy for the prediction of ethanol prices – in the delimited time interval – was ARFIMA, taking into account the lower Root Mean Squared Error (RMSE) of this model in relation to the others, in a forecast of US$ 610.38 for the 1001st point; The inclusion of new mathematical models should be carried out in the experiment, aiming at greater accuracy in the daily forecast of ethanol prices.

ACKNOWLEDGMENTS

The authors would like to thank the Graduate Program in Computational Modeling and Industrial Technology at SENAI CIMATEC.

REFERENCES

Associação Nacional dos Fabricantes de Veículos Automotores. (2022). ANFAVEA. Retrieved March 15, 2022, from http://www.anfavea.com.br/

Benites-Lazaro, L. L., Mello-Théry, N. A., & Lahsen, M. (2017). Business storytelling about energy and climate change: The case of Brazil's ethanol industry. Energy Research & Social Science, 31, 77–85. https://doi.org/10.1016/j.erss.2017.06.007

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). Time series analysis: Forecasting and control (3rd ed.). Prentice Hall.

Centro de Estudos Avançados em Economia Aplicada. (2018). CEPEA. Retrieved March 15, 2022, from https://www.cepea.esalq.usp.br/br

David, S. A., Inácio, C. M. C., & Tenreiro Machado, J. A. (2018). Fractional dynamic behavior in ethanol prices series. Journal of Computational and Applied Mathematics, 339, 85–93. https://doi.org/10.1016/j.cam.2017.10.034

Dias de Oliveira, M. E., Vaughan, B. E., & Rykiel, E. J. (2005). Ethanol as fuel: Energy, carbon dioxide balances, and ecological footprint. BioScience, 55(7), 593–602. https://doi.org/10.1641/0006-3568(2005)055[0593:EAFECD]2.0.CO;2

El Montasser, G., Gupta, R., Martins, A. L., & Wanke, P. (2015). Are there multiple bubbles in the ethanol–gasoline price ratio of Brazil? Renewable and Sustainable Energy Reviews, 52, 19–23. https://doi.org/10.1016/j.rser.2015.07.097

Gomez, J. M. A., & Legey, L. F. L. (2015). An analysis of the impact of Flex-Fuel vehicles on fuel consumption in Brazil, applying cointegration and the Kalman Filter. Energy, 81, 696–705. https://doi.org/10.1016/j.energy.2015.01.003

Maçaira, P. M., Thomé, A. M. T., Oliveira, F. L. C., & Ferrer, A. L. C. (2015). Modelling and forecasting the residential electricity consumption in Brazil with Pegels exponential smoothing techniques. Procedia Computer Science, 55, 328–335. https://doi.org/10.1016/j.procs.2015.07.058

Morettin, P. A. (2008). Econometria financeira: Um curso em séries temporais financeiras (1st ed.). Blucher.

R Foundation. (2018). The Comprehensive R Archive Network. Retrieved July 15, 2018, from https://cran.r-project.org/

Renewable Fuels Association. (2022). Ethanolrfa. Retrieved March 15, 2022, from http://www.ethanolrfa.org/

Triana, C. A. R. (2011). Energetics of Brazilian ethanol: Comparison between assessment approaches. Energy Policy, 39(8), 4605–4613. https://doi.org/10.1016/j.enpol.2011.04.024