

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**

**PRISCILLA BASSETTO**

**ANÁLISE COMPARATIVA DE MODELOS LINEARES PARA PREVISÃO DO  
PREÇO DO AÇÚCAR**

**PONTA GROSSA**

**2022**

**PRISCILLA BASSETTO**

**ANÁLISE COMPARATIVA DE MODELOS LINEARES PARA PREVISÃO DO  
PREÇO DO AÇÚCAR**

**COMPARATIVE ANALYSIS OF LINEAR MODELS FOR SUGAR PRICE  
FORECAST**

Dissertação apresentada como requisito parcial à obtenção do título de mestre em Engenharia de Produção, do Programa de Pós-Graduação em Engenharia de Produção, da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador: Prof. Dr. Hugo Valadares Siqueira  
Coorientador: Prof. Dra. Yslene Rocha Kachba

**PONTA GROSSA**

**2022**



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Esta licença permite que outros remixem, adaptem e criem a partir do trabalho para fins não comerciais, desde que atribuam o devido crédito e que licenciem as novas criações sob termos idênticos. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.



**Ministério da Educação**  
**Universidade Tecnológica Federal do Paraná**  
**Campus Ponta Grossa**



PRISCILLA BASSETTO

## **ANÁLISE COMPARATIVA DE MODELOS LINEARES PARA PREVISÃO DO PREÇO DO AÇÚCAR**

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestra Em Engenharia De Produção da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Gestão Industrial.

Data de aprovação: 30 de Agosto de 2022

Dr. Hugo Valadares Siqueira, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Flavio Trojan, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Gilberto Reynoso Meza, Doutorado - Pontifícia Universidade Católica do Paraná (Pucpr)

Dr. Sergio Luiz Stevan Junior, Doutorado - Universidade Tecnológica Federal do Paraná

Dra. Yslene Rocha Kachba, Doutorado - Universidade Tecnológica Federal do Paraná

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 30/08/2022.

## **AGRADECIMENTOS**

Agradeço a todos aqueles que contribuíram direta e indiretamente para a realização deste trabalho, em especial:

A Deus, por nunca me deixar sozinha, sempre me acompanhando e abençoando.

Aos meus professores, por todo o ensinamento, orientação, apoio e confiança. Em especial ao meu orientador Prof. Dr. Hugo Valadares Siqueira e co-orientadora Prof. Dr. Yslene Rocha Kachba por toda confiança e dedicação depositadas nesse trabalho, contribuindo para a minha formação e amadurecimento enquanto pesquisadora.

Ao Programa de Pós-Graduação em Engenharia de Produção da UTFPR câmpus Ponta Grossa, que tem buscado sempre a melhoria do curso de pós-graduação.

Aos membros da banca, por aceitar colaborar com a avaliação do trabalho.

Aos meus pais, por sempre acreditarem em mim, por todo apoio, amor, dedicação e paciência, sem vocês, esse sonho não seria possível.

Aos meus amigos e colegas de mestrado, pelo compartilhamento de conhecimento, histórias e momentos de distrações. A todos, meus sinceros agradecimentos.

Obrigada

E ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa concedida durante o período de mestrado, processo nº 132985/2020-9.

## RESUMO

O processo de previsão de séries temporais baseia-se no pressuposto de que os valores futuros são previstos a partir de observações passadas e outras entradas. Trata-se de uma das mais antigas técnicas de análise preditiva e possui bases estatísticas profundas sendo amplamente utilizada nos ambientes organizacionais e de pesquisa. Uma abordagem comum para análise de séries temporais envolve avaliar várias técnicas de modelagem em um conjunto de dados e observar como elas explicam o comportamento do seu passado. Neste sentido, um problema de relevância para os dias atuais é compreender os movimentos nos preços mundiais do açúcar. Entender a dinâmica de preços entre os mercados doméstico e internacional é de grande importância para o planejamento estratégico do setor sucroalcooleiro. Dessa forma, o presente estudo tem como objetivo analisar modelos lineares de previsão de séries temporais para prever o preço do açúcar. Para isso, foram utilizados quatro bases de dados, Brasil, Estados Unidos, Mundo e União Europeia. Os dados foram aplicados usando 7 modelos, sendo eles, Autorregressivo (AR), Médias Móveis (MA), Autorregressivo e Médias Móveis (ARMA), Autorregressivo Integrado de Médias Móveis (ARIMA), Suavização exponencial simples (SES), Suavização exponencial com tendência (HOLT) e Suavização exponencial com tendência e sazonalidade (HOLT WINTERS) e ao final foram realizadas combinações de diferentes modelos baseado na média das saídas dos preditores. Para a presente pesquisa utilizou-se medidas quantitativas de acurácia para avaliar os modelos de previsão: Raiz Erro Quadrático Médio (RMSE) Erro Absoluto Médio (MAE) e Erro Percentual Absoluto Médio (MAPE). Pode-se observar que para previsões de 1 passo à frente os modelos da família ARIMA estimaram maior precisão o preço do açúcar. Já para as previsões de 3, 6 e 12 passos à frente a família HOLT WINTERS foram os modelos que apresentaram melhor desempenho.

**Palavras Chaves:** arima; holt winters; série temporal; ensemble; açúcar.

## ABSTRACT

The time series forecasting process is based on the assumption that future values are predicted from past observations and other inputs. It is one of the oldest predictive analysis techniques and has deep statistical bases and is widely used in organizational and research environments. A common approach to time series analysis involves evaluating various modeling techniques on a dataset and looking at how they explain past behavior. In this sense, a problem of relevance to the present day is to understand the movements in world sugar prices. Understanding the price dynamics between the domestic and international markets is of great importance for the strategic planning of the sugar and ethanol sector. Thus, the present study aims to analyze linear time series forecasting models to predict the price of sugar. For this, four databases were used, Brazil, United States, World and European Union. The data were applied using 7 models, namely, Autoregressive (AR), Moving Averages (MA), Autoregressive and Moving Averages (ARMA), Integrated Autoregressive Moving Averages (ARIMA), Simple Exponential Smoothing (SES), Exponential Smoothing with Trend (HOLT) and Exponential smoothing with trend and seasonality (HOLT WINTERS) and at the end, combinations of different models were performed based on the average of the predictors' outputs. For the present research, quantitative measurements of accuracy were used to evaluate the prediction models: Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percent Error (MAPE). It can be seen that for 1-step-ahead forecasts, the ARIMA family models estimated the sugar price more accurately. As for the 3, 6, and 12 steps forward predictions, the HOLT WINTERS family were the models that presented the best performance.

**Keywords:** arima; holt winter; time series; ensemble; sugar.

## LISTA DE FIGURAS

|  |    |
|--|----|
| <b>Figura 1 - Estrutura do trabalho</b> .....  | 20 |
| <b>Figura 2 - Técnicas de previsão de séries temporais</b> .....                               | 23 |
| <b>Figura 3 - Comportamento de uma série temporal em tendência, sazonalidade e ruído</b> ..... | 23 |
| <b>Figura 4 - Representação de uma série temporal como saída de um filtro linear</b><br>.....  | 29 |
| <b>Figura 5 - Modelo genérico de Ensemble</b> .....  | 34 |
| <b>Figura 6 - Etapas para construção de um modelo de previsão</b> .....                        | 37 |

## LISTA DE GRÁFICOS

|   |           |
|---|-----------|
| <b>Gráfico 1 - Série temporal preço do açúcar – Brasil .....</b>                            | <b>43</b> |
| <b>Gráfico 2 - Preço do açúcar Estados Unidos.....</b>                                      | <b>44</b> |
| <b>Gráfico 3 - Preço do açúcar Mundo .....</b>  | <b>44</b> |
| <b>Gráfico 4 - Preço do açúcar União Europeia.....</b>                                      | <b>44</b> |
| <b>Gráfico 5 - Sazonal Brasil .....</b>   | <b>46</b> |
| <b>Gráfico 6 – Sazonal Estados Unidos .....</b>   | <b>46</b> |
| <b>Gráfico 7 - Sazonal Mundo .....</b>  | <b>47</b> |
| <b>Gráfico 8 - Sazonal União Europeia.....</b>  | <b>47</b> |
| <b>Gráfico 9 - Subsérie sazonal Brasil .....</b>  | <b>48</b> |
| <b>Gráfico 10 - Subsérie sazonal Estados Unidos .....</b>                                   | <b>48</b> |
| <b>Gráfico 11 - Subsérie sazonal Mundo.....</b>   | <b>49</b> |
| <b>Gráfico 12 - Subsérie sazonal União Europeia .....</b>                                   | <b>49</b> |
| <b>Gráfico 13 - ACF e PACF Brasil.....</b>  | <b>50</b> |
| <b>Gráfico 14 - ACF e PACF Estados Unidos.....</b>  | <b>51</b> |
| <b>Gráfico 15 - ACF e PACF Mundo.....</b>   | <b>51</b> |
| <b>Gráfico 16 - ACF e PACF União Europeia .....</b>   | <b>51</b> |
| <b>Gráfico 17 - Conjunto de treinamento (vermelho) e teste (verde) Brasil .....</b>         | <b>52</b> |
| <b>Gráfico 18 - Conjunto de treinamento (vermelho) e teste (verde) Estados Unidos .....</b> | <b>52</b> |
| <b>Gráfico 19 - Conjunto de treinamento (vermelho) e teste (verde) Mundo .....</b>          | <b>53</b> |
| <b>Gráfico 20 - Conjunto de treinamento (vermelho) e teste (verde) União Europeia .....</b> | <b>53</b> |
| <b>Gráfico 21 - Modelo AR ajustado nos dados de treinamento – Brasil.....</b>               | <b>54</b> |
| <b>Gráfico 22 - Modelo MA ajustado nos dados de treinamento – Brasil.....</b>               | <b>54</b> |
| <b>Gráfico 23 - Modelo ARMA ajustado nos dados de treinamento – Brasil .....</b>            | <b>54</b> |
| <b>Gráfico 24 - Modelo ARIMA ajustado nos dados de treinamento – Brasil .....</b>           | <b>55</b> |
| <b>Gráfico 25 - Modelo SES ajustado nos dados de treinamento - Brasil.....</b>              | <b>55</b> |
| <b>Gráfico 26 - Modelo HOLT ajustado nos dados de treinamento – Brasil .....</b>            | <b>55</b> |
| <b>Gráfico 27 - Modelo HOLT WINTERS ajustado nos dados de treinamento – Brasil .....</b>    | <b>56</b> |



|  |    |
|--|----|
| <b>Gráfico 28 - Modelo AR ajustado nos dados de treinamento – Estados Unidos</b>           | 56 |
| .....  |    |
| <b>Gráfico 29 - Modelo MA ajustado nos dados de treinamento – Estados Unidos</b>           | 57 |
| .....  |    |
| <b>Gráfico 30 - Modelo ARMA ajustado nos dados de treinamento – Estados Unidos</b>         | 57 |
| .....  |    |
| <b>Gráfico 31 - Modelo ARIMA ajustado nos dados de treinamento – Estados Unidos</b>        | 57 |
| .....  |    |
| <b>Gráfico 32 - Modelo SES ajustado nos dados de treinamento – Estados Unidos</b>          | 58 |
| .....  |    |
| <b>Gráfico 33 - Modelo HOLT ajustado nos dados de treinamento – Estados Unidos</b>         | 58 |
| .....  |    |
| <b>Gráfico 34 - Modelo HOLT WINTERS ajustado nos dados de treinamento – Estados Unidos</b> | 58 |
| .....  |    |
| <b>Gráfico 35 - Modelo AR ajustado nos dados de treinamento – Mundo</b>                    | 59 |
| <b>Gráfico 36 - Modelo MA ajustado nos dados de treinamento – Mundo</b>                    | 59 |
| <b>Gráfico 37 - Modelo ARMA ajustado nos dados de treinamento – Mundo</b>                  | 59 |
| <b>Gráfico 38 - Modelo ARIMA ajustado nos dados de treinamento – Mundo</b>                 | 60 |
| <b>Gráfico 39 - Modelo SES ajustado nos dados de treinamento – Mundo</b>                   | 60 |
| <b>Gráfico 40 - Modelo HOLT ajustado nos dados de treinamento – Mundo</b>                  | 60 |
| <b>Gráfico 41 - Modelo HOLT WINTERS ajustado nos dados de treinamento – Mundo</b>          | 61 |
| .....  |    |
| <b>Gráfico 42 - Modelo AR ajustado nos dados de treinamento – União Europeia</b>           | 61 |
| <b>Gráfico 43 - Modelo MA ajustado nos dados de treinamento – União Europeia</b>           | 62 |
| <b>Gráfico 44 - Modelo ARMA ajustado nos dados de treinamento – União Europeia</b>         | 62 |
| .....  |    |
| <b>Gráfico 45 - Modelo ARIMA ajustado nos dados de treinamento – União Europeia</b>        | 62 |
| .....  |    |
| <b>Gráfico 46 - Modelo SES ajustado nos dados de treinamento – União Europeia</b>          | 63 |
| .....  |    |
| <b>Gráfico 47 - Modelo HOLT ajustado nos dados de treinamento – União Europeia</b>         | 63 |
| .....  |    |
| <b>Gráfico 48 - Modelo HOLT WINTERS ajustado nos dados de treinamento – União Europeia</b> | 63 |
| .....  |    |

|  |           |
|--|-----------|
| <b>Gráfico 49 - Previsão 1 passo à frente – Brasil.....</b>            | <b>66</b> |
| <b>Gráfico 50 - Previsão 1 passo à frente – Estados Unidos .....</b>   | <b>67</b> |
| <b>Gráfico 51 - Previsão 1 passo à frente – Mundo.....</b>             | <b>67</b> |
| <b>Gráfico 52 - Previsão 1 passo à frente – União Europeia .....</b>   | <b>68</b> |
| <b>Gráfico 53 - Previsão 3 passos à frente – Brasil.....</b>           | <b>70</b> |
| <b>Gráfico 54 - Previsão 3 passos à frente – Estados Unidos.....</b>   | <b>71</b> |
| <b>Gráfico 55 - Previsão 3 passos à frente – Mundo.....</b>            | <b>71</b> |
| <b>Gráfico 56 - Previsão 3 passos à frente – União Europeia .....</b>  | <b>72</b> |
| <b>Gráfico 57 - Previsão 6 passos à frente – Brasil.....</b>           | <b>74</b> |
| <b>Gráfico 58 - Previsão 6 passos à frente – Estados Unidos.....</b>   | <b>74</b> |
| <b>Gráfico 59 - Previsão 6 passos à frente – Mundo.....</b>            | <b>75</b> |
| <b>Gráfico 60 - Previsão 6 passos à frente – União Europeia .....</b>  | <b>75</b> |
| <b>Gráfico 61 - Previsão 12 passos à frente - Brasil .....</b>         | <b>77</b> |
| <b>Gráfico 62 - Previsão 12 passos à frente – Estados Unidos.....</b>  | <b>78</b> |
| <b>Gráfico 63 - Previsão 12 passos à frente – Mundo.....</b>           | <b>78</b> |
| <b>Gráfico 64 - Previsão 12 passos à frente – União Europeia .....</b> | <b>79</b> |

## LISTA DE TABELA

|   |           |
|---|-----------|
| <b>Tabela 1 - Descrição dos dados e conjunto de treinamento e teste .....</b>   | <b>42</b> |
| <b>Tabela 2 - Análise descritiva dos dados.....</b>                             | <b>42</b> |
| <b>Tabela 3 - Performance dos modelos – 1 passo à frente.....</b>               | <b>65</b> |
| <b>Tabela 4 - Ranking de desempenho das previsões 1 passo à frente .....</b>    | <b>65</b> |
| <b>Tabela 5 - Performance dos modelos – 3 passos à frente.....</b>              | <b>69</b> |
| <b>Tabela 6 - Ranking de desempenho das previsões 3 passos à frente .....</b>   | <b>69</b> |
| <b>Tabela 7 - Performance dos modelos – 6 passos à frente.....</b>              | <b>73</b> |
| <b>Tabela 8 - Ranking de desempenho das previsões 6 passos à frente .....</b>   | <b>73</b> |
| <b>Tabela 9 - Performance dos modelos – 12 passos à frente.....</b>             | <b>76</b> |
| <b>Tabela 10 - Ranking de desempenho das previsões 12 passos à frente .....</b> | <b>76</b> |

## LISTA DE SIGLAS

|       |   |
|-------|---|
| ACF   | Função de Autocorrelação                            |
| AR    | Autoregressivo                                      |
| AIC   | Critério de Informação de Akaike                    |
| AICc  | Critério de Informação de Akaike de Segunda Ordem   |
| ARMA  | Autoregressivo de Médias Móveis                     |
| ARIMA | Autoregressivo Integrado de Médias Móveis           |
| BIC   | Critério de Informação Bayesiano                    |
| KPSS  | Kwiatkowski Phillips Schmidt Shin                   |
| MA    | Médias Móveis                                       |
| MAE   | Erro Absoluto Médio                                 |
| MAPE  | Erro Percentual Absoluto Médio                      |
| PACF  | Função de Autocorrelação Parcial                    |
| RMSE  | Raiz Erro Quadrático Médio                          |
| SEH   | Suavização Exponencial com Tendência                |
| SES   | Suavização Exponencial Simples                      |
| SEHW  | Suavização Exponencial com Tendência e Sazonalidade |

## SUMÁRIO

|   |    |
|---|----|
| <b>1 INTRODUÇÃO</b> .....   | 14 |
| <b>1.1 Objetivos</b> .....  | 17 |
| 1.1.1 Objetivo geral.....   | 17 |
| 1.1.2 Objetivos específicos .....   | 17 |
| <b>1.2 Justificativa</b> .....  | 18 |
| <b>1.3 Estrutura da pesquisa</b> .....  | 19 |
| <b>2 REVISÃO DE LITERATURA</b> .....  | 21 |
| <b>2.1 Previsão de séries temporais</b> .....                                 | 21 |
| <b>2.2 Modelos lineares de previsão</b> .....                                 | 24 |
| <b>2.3 Família holt winters</b> .....   | 25 |
| 2.3.1 Suavização exponencial simples .....                                    | 25 |
| 2.3.2 Suavização exponencial com tendência (holt).....                        | 26 |
| 2.3.3 Suavização exponencial com tendência e sazonalidade (holt winters)..... | 27 |
| <b>2.4 Família arima</b> .....  | 28 |
| 2.4.1 Autorregressivo (AR).....   | 30 |
| 2.4.2 Médias móveis (MA) .....  | 31 |
| 2.4.3 Autorregressivo e médias móveis (ARMA) .....                            | 31 |
| 2.4.4 Autorregressivo integrado de médias móveis (ARIMA) .....                | 31 |
| <b>2.5 Ensemble</b> .....   | 33 |
| <b>3 METODOLOGIA</b> .....  | 35 |
| <b>3.1 Enquadramento metodológico</b> .....                                   | 35 |
| <b>3.2 Estruturação do modelo</b> .....                                       | 36 |
| 3.2.1 Base de dados .....   | 36 |
| 3.2.2 Pré-processamento dos dados .....                                       | 37 |
| 3.2.3 Previsão.....   | 39 |
| 3.2.4 Monitoramento e avaliação comparativa dos modelos .....                 | 41 |
| <b>4 RESULTADOS E DISCUSSÕES</b> .....  | 42 |
| <b>4.1 Análise dos dados</b> .....  | 43 |
| <b>4.2 Autocorrelação</b> .....   | 50 |
| <b>4.3 Ajuste do modelo</b> .....   | 52 |
| <b>4.4 Previsão</b> .....   | 64 |

|  |           |
|--|-----------|
| 4.4.1 Previsão 1 passo à frente .....            | 64        |
| 4.4.2 Previsão 3 passos à frente .....           | 68        |
| 4.4.3 Previsão 6 passos à frente .....           | 72        |
| 4.4.4 Previsão 12 passos à frente .....          | 75        |
| <b>4.5 Discussão .....</b>                       | <b>79</b> |
| <b>5 CONCLUSÃO .....</b>                         | <b>81</b> |
| <b>REFERÊNCIAS .....</b>                         | <b>83</b> |
| <b>APÊNDICE A – PARÂMETROS DOS MODELOS .....</b> | <b>91</b> |

## 1 INTRODUÇÃO

O açúcar é uma das mais antigas *commodities* comercializadas no mundo. A partir do século XV, o açúcar exerceu grande influência em diversas economias do mundo (GRIGGS, 2011; GUDOSHNIKOV *et al.*, 2004) e hoje representa uma importante *commodity* para as economias de diversos países e um dos produtos mais regulamentados nos mercados internacionais (BENEŠOVÁ *et al.*, 2015).

O mercado global do açúcar é liderado pelo Brasil, maior produtor e exportador do mundo. Em 2019, o país exportou US\$ 5,33 bilhões em açúcar bruto. No mesmo ano, este foi o 10º produto mais exportado do Brasil (OEC, 2019).

A expansão da produção de açúcar foi impulsionada por vários fatores, incluindo apoios governamentais, aumento da demanda de produção de etanol à base da cana-de-açúcar e recursos naturais vastos e adequados (OCDE/FAO, 2017).

A popularidade do etanol ganhou impulso em 2003 com a introdução dos veículos Flex, que combinam gasolina e etanol, aumentando consideravelmente as possibilidades de substituição entre tais combustíveis (BENTIVOGLIO, 2016). Conseqüentemente, o rápido crescimento da produção de etanol levantou preocupações sobre o impacto do etanol no nível de preços das *commodities* agrícolas, como o açúcar. Em 2014, o Brasil foi o maior produtor e consumidor mundial de etanol da cana-de-açúcar como combustível para transporte.

São os mercados internacionais que determinam quanto valem as *commodities* agrícolas. Além disso, os preços são regulados pela lei da oferta e demanda. Alguns estudos apontam as interdependências dos preços entre os mercados de alimentos, energia e biocombustíveis (HOCHMAN *et al.*, 2010; KRISTOUFEK *et al.*, 2012).

As usinas no Brasil geralmente têm a infraestrutura para alternar entre açúcar e etanol dependendo dos retornos de preços relativos, o que define a opção de produzir mais ou menos de cada um, com vistas a maximização dos lucros. A coordenadora de Inteligência de Mercado da Copersucar, Tainnah Ikegami, destaca que “Se a bolsa estiver disparada, mas a conversão em reais não gerar uma remuneração favorável, não necessariamente as usinas irão optar pelo açúcar” (COPERSUCAR, 2021).

As principais variáveis macroeconômicas que influenciam o mercado açucareiro são o crescimento populacional e o crescimento do PIB per capita

(MITCHELL 2004; CARMAN, 1982). Seu impacto é particularmente visível por meio de mudanças no lado da demanda do mercado. Por exemplo, no caso do PIB, em períodos de recessão, a demanda por fabricação de bebidas e alimentos tende a cair, exercendo baixa pressão sobre as cotações do açúcar (ISO, 2016).

Outro conjunto de variáveis que impulsionam os preços internacionais do açúcar diz respeito aos mercados financeiros. O mais importante é a flutuação cambial no valor da moeda brasileira (Real) em relação ao valor do dólar americano (US\$) (STEPHEN, 2013). Uma desvalorização tende a reduzir os custos de produção em dólares no Brasil, o que aumenta as exportações e afeta os preços mundiais. Além disso, como os preços mundiais em dólares americanos, uma desvalorização do Real incentiva os exportadores a embarcar maiores volumes para garantir maior retorno em moeda local.

O contrato *Sugar* N° 11 é a referência mundial para comercialização de açúcar bruto e está disponível na *The Intercontinental Exchange* (ICE). Os preços exibidos na *Trading Economics* são baseados em instrumentos financeiros de balcão (OTC) e contrato por diferença (CFD).

Os preços internacionais do açúcar são conhecidos por serem altamente oscilatórios (FAO, 2016). Isto decorre das características econômicas e físicas do mercado. O açúcar é conhecido por ser uma das *commodities* mais protegidas pelos governos, visto que procuram proteger os produtores dos preços baixos por meio mecanismos de proteção, como medidas de fronteira, nível de preço mínimo e subsídios (FAO, 2016).

A literatura sobre modelagem de mercados internacionais de açúcar e projeção de seus preços tende a usar especificações estruturais como modelos de equilíbrio geral (MENSBRUGGHE *et al.*, 2003) ou modelos de equilíbrio parcial (OECD, 2017; AMROUK.; HECKELEI, 2020; ADENÄUER *et al.*, 2004; e NOLTE *et al.*, 2012).

Outros modelos estudados abordam técnicas de séries temporais, como relatado em (STEPHEN, 2013; STEPHEN, 2015; CHANG *et al.*, 2018; RABOBANK, 2020; e RUMÁNKOVÁ *et al.*, 2019). Neste caso, as técnicas permitem o uso de dados em um nível de frequência muito mais alto (por exemplo, diário) do que modelos de equilíbrio parcial ou geral, permitindo tirar proveito de mais informações.

Lima *et al.* (2007) modelaram séries temporais para efeito de previsão com diferenciações inteira e fracionária, utilizando dados de preços futuros de *commodities*



agrícolas. Modelos de séries temporais do tipo ARMA/ARIMA (diferenciação inteira) foram estimados como termo de comparação com os modelos do tipo Autoregressivo Fracionário Integrado de Médias Móveis (*Autoregressive Fractionally Integrated Moving Average* – ARFIMA) (diferenciação fracionária). Na maioria dos casos os modelos ARFIMA mostraram melhor desempenho de previsão, apesar da pequena diferença em alguns casos.

Suresh e Priya (2011) propuseram prever a área canavieira, a produção e a produtividade, por meio, do ajuste de modelos Autoregressivo Integrado de Médias Móveis (*Auto Regressive Integrated Moving Average* - ARIMA). Os modelos ARIMA demonstraram bons resultados para a previsão da área de cana-de-açúcar, produção e produtividade de Tamilnadu para os anos subsequentes.

Bentivoglio *et al.* (2016) tiveram como objetivo explorar o impacto dos preços do etanol brasileiro sobre os preços do açúcar e da gasolina. Estas relações foram investigadas por meio de um modelo de Correção de Erro Vetorial (*Vector Error Correction Model* - VECM), suportado por testes de causalidade de Granger. Além disso, as Funções de Resposta ao Impulso (*Impulse Response Functions* - IRFs) e as Decomposições de Variância de Erro de Previsão (*Forecast Error Variance Decompositions* - FEVD) foram calculadas a fim de investigar as inter-relações dinâmicas dentro dessas séries. Os resultados evidenciaram que os preços do etanol são afetados tanto pelos preços dos alimentos quanto dos combustíveis, mas não há fortes evidências de que as mudanças nos preços do etanol tenham impacto nos preços dos alimentos (BENTIVOGLIO *et al.*, 2016).

Cheng (2017) utilizou as características sazonais para prever o preço do açúcar, usando método de decomposição sazonal. Em seguida, a sequência de ajuste sazonal foi prevista por suavização exponencial. Por fim, os fatores sazonais foram reabastecidos com a previsão de tendência e os resultados da previsão de preços do açúcar futuro foram obtidos.

David *et al.* (2017) analisaram as séries temporais dos preços das commodities soja, milho, café e açúcar, por meio da Transformação Fracionária de Fourier (*Fractional Fourier Transform* - FrFT) para encontrar padrões de tempo/frequência nos dados. Em uma segunda fase, aplicaram os modelos ARIMA e ARFIMA para a obtenção da composição dos preços à vista e previsões futuras. Por fim, ao comparar o desempenho dos modelos ARIMA e ARFIMA, observou-se que o ARFIMA obteve melhor desempenho.

Amrouk e Heckelei (2020) examinaram a importância relativa das variáveis externas para a previsão dos preços internacionais do açúcar. Eles combinaram uma técnica de média do modelo Bayesiano (BMA) para abordar a incerteza de especificação com uma análise fora da amostra para avaliar a previsibilidade de preços. Os resultados mostraram que as cotações mundiais do açúcar são influenciadas principalmente por sua própria dinâmica, mudanças nos preços internacionais de alimentos básicos, custos de produção de açúcar e variáveis macroeconômicas.

Devido à incerteza e as oscilações nos preços internacionais, torna-se importante realizar um estudo de previsão do preço do açúcar com o objetivo de minimizar os riscos futuros, apresentando uma estimativa mais precisa possível de suas tendências.

## **1.1 Objetivos**

### 1.1.1 Objetivo geral

O presente estudo tem como objetivo analisar as séries temporais para previsão do preço do açúcar e identificar o melhor modelo por meio de uma comparação de desempenho, utilizando modelos lineares e um ensemble baseado na média das respostas dos modelos.

### 1.1.2 Objetivos específicos

A presente pesquisa tem como objetivos específicos:

- Identificar literatura dos métodos de previsão de séries temporais que serão abordados no presente trabalho, a saber família Holt Winters, família ARIMA e *Ensemble*;
- Levantar e relacionar bases de dados Brasileira e mundial do açúcar;
- Realizar a previsão do preço do açúcar com 1, 3, 6 e 12 passos à frente utilizando granularidade mensal;
- Investigar os resultados obtidos de cada modelo de previsão;
- Analisar os resultados entre os métodos de previsão e o software utilizado;

- Encontrar o melhor modelo de previsão do preço do açúcar para o presente estudo, considerando os métodos abordados na pesquisa.

## 1.2 Justificativa

Dessa forma, entender a dinâmica de preços entre os mercados doméstico e internacional é de suma importância para o planejamento estratégico do setor sucroalcooleiro. O impacto no preço de uma mercadoria tão essencial tem um papel significativo na determinação da taxa de inflação, salários, vencimentos e várias políticas em uma economia. No caso de culturas comerciais como a cana-de-açúcar, o nível de produção afeta o custo da matéria-prima das indústrias usuárias e suas vantagens competitivas no mercado (PAGANI *et al.*, 2017).

Desde a década de 1960, os modelos lineares de previsão de séries temporais vêm alcançando novos avanços. Prever valores futuros é de suma importância em todas as ciências aplicadas e cada disciplina tem suas próprias características e necessidades típicas que devem ser levadas em consideração por pesquisadores que desenvolvem novos métodos de previsão.

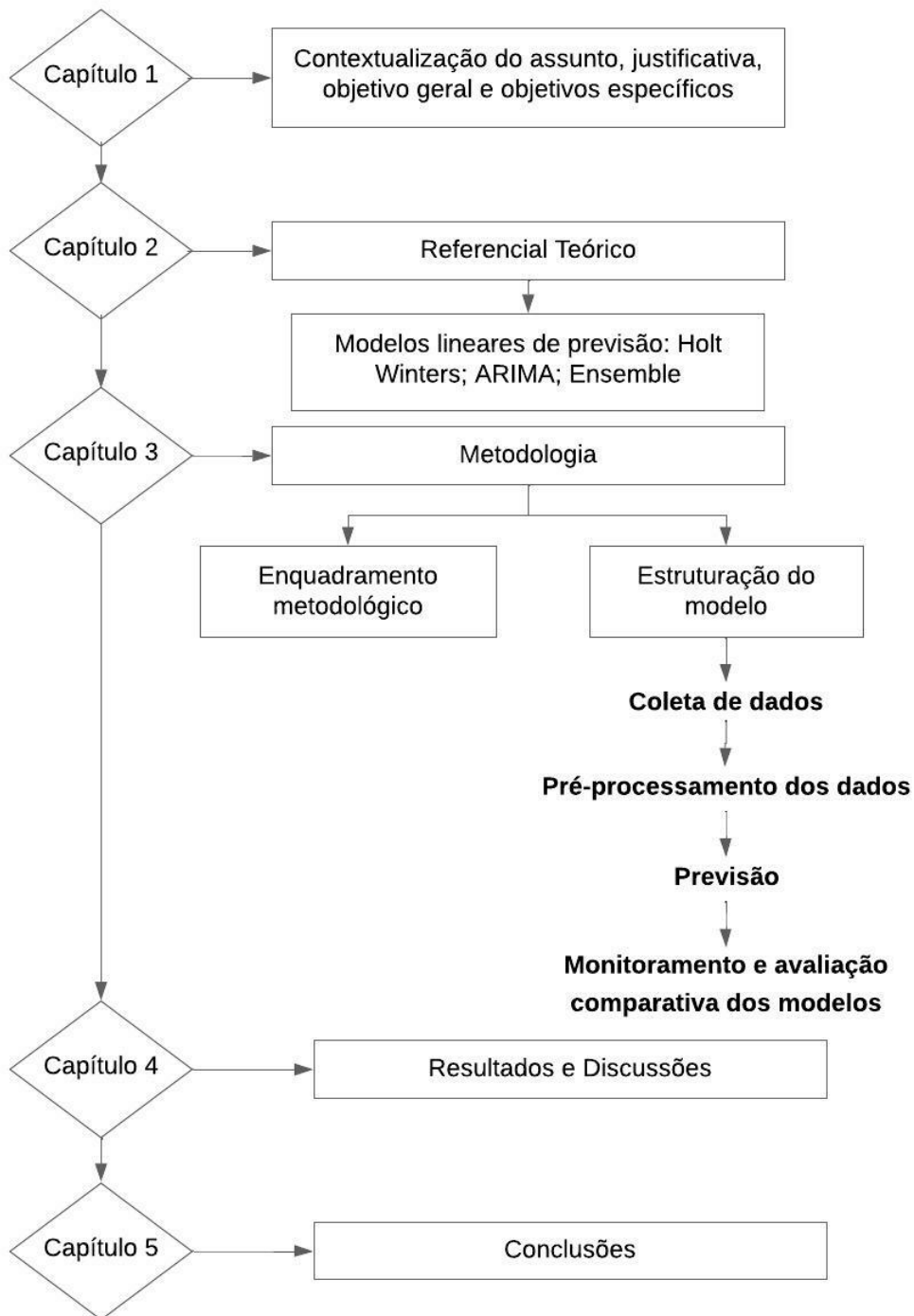
Neste sentido, compreender os movimentos nos preços mundiais do açúcar ajudam os participantes da cadeia de valor do açúcar a formular estratégias de investimento eficazes e prever os efeitos dos choques de mercado com mais precisão. Isto pode auxiliar na tomada de decisões de importação, exportação e fixação de preços (BAJPAI *et al.*, 2012). Assim, apresentar um estudo que auxilie de forma efetiva a sociedade, pesquisadores e cientistas que trabalham na área mostra-se relevante.

Sabendo da importância desse produto não apenas na economia brasileira, mas também para os agentes da cadeia internacional, justifica-se a importância desta pesquisa, que busca analisar e prever o preço do açúcar em quatro bases de dados diferentes. Além disso, a literatura precisa de estudos que explorem a previsão por meio de combinações de modelos, assim, este trabalho utilizou *Ensemble* com o objetivo de melhorar ainda mais a previsão.

### **1.3 Estrutura da pesquisa**

O presente estudo está estruturado em cinco capítulos. No primeiro discutiu-se sobre a contextualização da pesquisa, justificativa e objetivos. No segundo capítulo, são apresentadas as teorias bases utilizadas no desenvolvimento deste trabalho e os conceitos fundamentais para o seu entendimento. O terceiro capítulo contempla a metodologia utilizada na elaboração do referencial teórico bem como para realização da pesquisa. No quarto capítulo, são apresentados os resultados e discussões dos métodos de previsão. E por fim, as conclusões, estão no quinto capítulo. A Figura 1 sumariza tal organização.

Figura 1 - Estrutura do trabalho



Fonte: Autoria própria (2022)

## 2 REVISÃO DE LITERATURA

Esta seção tem por objetivo apresentar as bases teóricas utilizadas para a fundamentação desta pesquisa. Para isso, foi realizada uma revisão de literatura em dissertações, livros e artigos nas bases de dados *Scopus*, *Science Direct* e *Web of Science*, as quais estão disponíveis no portal de periódicos da CAPES. Justifica-se o uso das três bases por criarem medidas amplamente utilizadas na avaliação da qualidade de periódicos. Ainda, buscou-se por literaturas que descrevessem os modelos estatísticos lineares de previsão utilizados, a saber: Família Holt Winters e ARIMA. Além destas, os Ensembles também foram investigados. Dessa forma, foi levado em consideração a disponibilidade de buscas por títulos, resumos, palavras-chave e expressões booleanas. As palavras-chave utilizadas estavam relacionadas aos eixos da pesquisa, a saber: previsão de séries temporais; preço do açúcar, modelos lineares. Ainda, não foi delimitado um horizonte temporal para buscas, pois ao delimitar o tema as buscas já filtravam os resultados.

### 2.1 Previsão de séries temporais

Como a presente pesquisa tem o propósito de analisar e prever o preço do açúcar, investigou-se os modelos de previsão de séries temporais. Nesta seção apresenta-se, especialmente, aqueles que serão utilizados na pesquisa, bem como os critérios para escolha e seleção dos mesmos.

De acordo com Reyhani e Moghadam (2011), os dados obtidos a partir da observação de um evento durante um período de tempo são muito úteis e de grande importância. Valor semanal de juros, preço de ações no mercado de ações, índices de preços mensais, valor de venda anual e etc. são observados na economia. Nas ciências biológicas, a atividade elétrica do coração é observada em durações de milissegundos. Em meteorologia, observa-se os graus máximos e mínimos de temperatura diária, índices anuais de precipitação e seca e velocidades do vento em diferentes horas. Todos esses casos estão lidando com dados de séries temporais.

Uma série temporal é uma ordem de números reais registrados em intervalos de tempo, como anual, mensal, semanal, diário ou até mesmo em milissegundos. Logo, a única variável independente na série é o tempo (REYHANI; MOGHADAM, 2011).

De acordo com Land (2015), previsão refere-se à formação de expectativas sobre estados ou processos futuros de entidades históricas específicas. A previsão envolve a estimativa ou cálculo de eventos futuros esperados ou desenvolvidos a partir de um modelo conceitual.

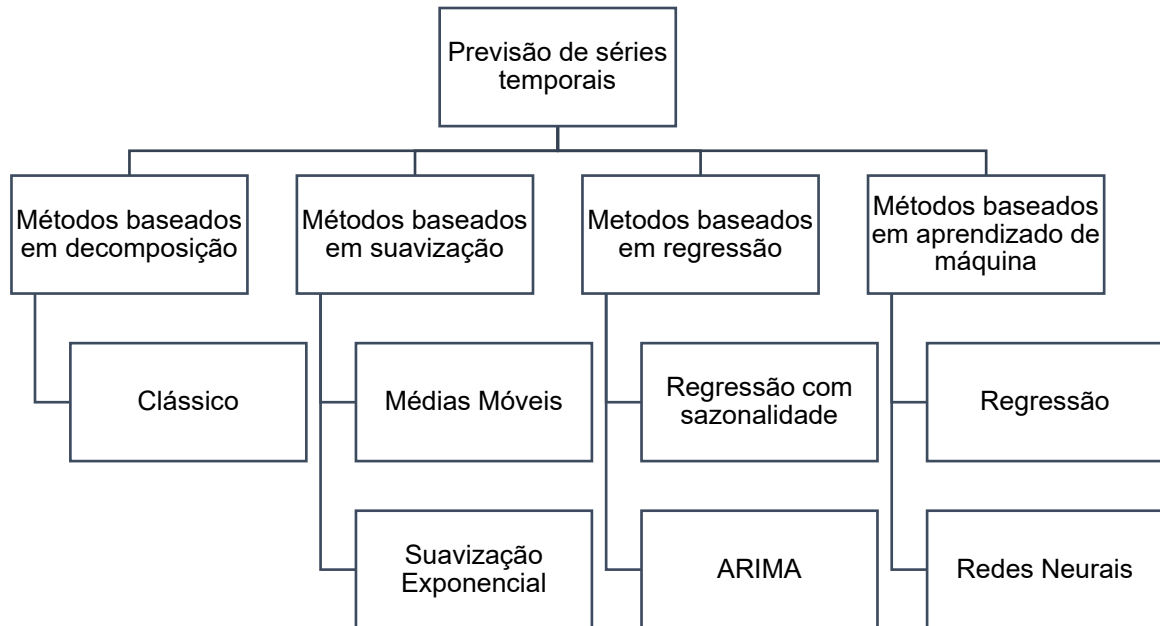
Assim, a previsão de série temporal é o processo de prever o valor futuro dos dados da série temporal com base em observações anteriores e outras entradas. A previsão de série temporal é uma das mais antigas conhecidas técnicas de análise preditiva. É amplamente utilizada em todos os ambientes organizacionais e possui bases estatísticas profundas (KOTU; DESHPANDE, 2019).

Esses métodos são baseados em leis e técnicas da ciência da estatística, que pode ser definida como o ramo da matemática que lida com a coleta, classificação, análise e interpretação de dados numéricos descrevendo um fenômeno. Os aspectos críticos das estatísticas se concentram na compreensão, descoberta e interpretação dos mecanismos que acompanham os dados do fenômeno sob estudo, derivando previsões e conclusões sobre sua evolução no futuro e sugerindo métodos e índices que podem garantir a validade das previsões (PROFILLIDIS; BOTZORIS, 2018).

Os modelos de séries temporais não visam explicar ou pesquisar a causa estrutural e efeito do problema, mas simplesmente, na detecção de seu padrão e evolução ao longo do tempo. Assim, quanto mais curto for o período de previsão, mais provável de se obter resultados estatisticamente válidos e precisos.

De acordo com Kotu e Deshpande (2019), a investigação de séries temporais pode ser amplamente dividida em modelagem descritivas, chamada de análise de série temporal, e modelagem preditiva, chamada de previsão de séries temporais. A previsão de séries temporais pode ser classificada em quatro grupos: Previsão com base em decomposição de séries temporais, técnicas baseadas em suavização, técnicas baseadas em regressão, e técnicas baseadas em aprendizado de máquina, conforme Figura 2.

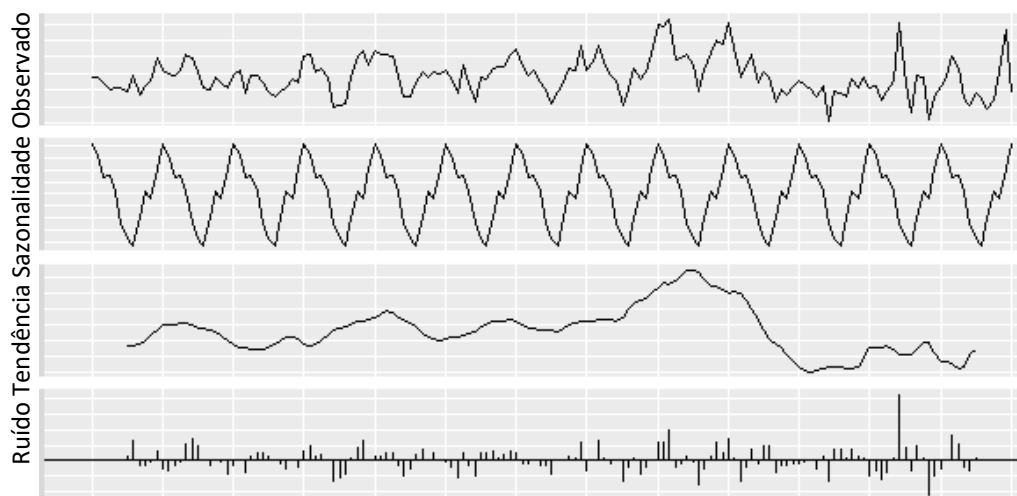
**Figura 2 - Técnicas de previsão de séries temporais**



**Fonte: Baseado em Kotu e Deshpande (2019)**

Assim, a decomposição de uma série temporal é o processo de desconstrução da série no número de comportamentos apresentados com cada um representando um fenômeno subjacente. A decomposição divide a série temporal em comportamento de tendência, sazonalidade e ruído, conforme Figura 3.

**Figura 3 - Comportamento de uma série temporal em tendência, sazonalidade e ruído**



**Fonte: Adaptado de Box et al. (2015)**



Os comportamentos de tendência e sazonalidade são previsíveis e chamados de sistemáticos, enquanto o ruído por definição é aleatório, denominado componente não sistemático. Antes de prever os próximos valores de uma série temporal, é importante compreender e identificar os componentes que descrevem a série (KOTU; DESHPANDE, 2019).

Na literatura existem inúmeros modelos de previsão, no entanto, para realizar o emprego do método correto é necessário analisar o comportamento da série que se deseja prever (MAKRIDAKIS *et al.*, 1998). Uma série temporal pode ser composta por um ou mais dos seguintes componentes:

a) **Tendência:** a série com tendência apresenta um comportamento não estacionário, apresentando aumento ou diminuição de longo prazo nos dados. Pode-se referir como uma “mudança de direção”, passando de uma tendência crescente para uma tendência decrescente. Além disso, não necessariamente é linear.

b) **Sazonalidade:** a componente sazonal acrescenta à padrões repetidos que aparecem continuamente. Ocorre quando uma série temporal é afetada por fatores sazonais, como a época do ano ou o dia da semana. A sazonalidade é sempre de frequência fixa e conhecida e pode ser dividida em horária, diária, semanal, mensal, trimestral, anual, estação, etc.

c) **Ciclo:** a componente cíclica também é não estacionária e representa oscilações durante longos períodos de tempo (anos ou décadas) em torno da curva do componente de tendência. Os dados exibem subidas e descidas que não são de frequência fixa. Essas flutuações geralmente são devidas a condições econômicas e muitas vezes estão relacionadas ao “ciclo de negócios”.

d) **Ruído:** qualquer comportamento que não seja representado por tendência, sazonalidade ou componente cíclico é considerado como ruído na série temporal. Tal componente é imprevisível, mas segue distribuição normal em casos ideais. Logo, todos os conjuntos de dados reais de uma série temporal terão ruído.

## **2.2 Modelos lineares de previsão**

Nessa seção serão apresentados os métodos lineares para modelagem e previsão de séries temporais univariadas abordados na presente pesquisa. Dessa forma, os modelos abordados são da família Holt Winters e família ARIMA. Além destes, são discutidos os Ensembles.

## 2.3 Família holt winters

A suavização exponencial foi proposta no final da década de 1950 (Brown, 1959; Holt, 1957; Winters, 1960) e motivou alguns dos métodos de previsão mais bem-sucedidos. As respostas produzidas usando estes métodos são médias ponderadas de observações anteriores, com os pesos decaindo exponencialmente à medida que as observações envelhecem. Em outras palavras, quanto mais recente a observação, maior o peso associado.

Grande parte dos métodos de previsão baseia-se na ideia de que observações passadas dispõe de informações sobre o padrão de comportamento da série temporal. Dessa forma, o objetivo é distinguir o padrão de qualquer ruído que possa estar contido nas observações e então usar esse padrão para prever valores futuros da série. Assim, técnicas específicas de previsão são propostas, como os modelos de suavização exponencial ao tentar tratar ambas as causas de flutuações em séries de tempo e das suavizações, que assumem que os valores extremos da série representam a aleatoriedade e, assim, por meio, da suavização desses extremos, pode-se identificar o padrão básico (MORETTIN; TOLOI, 2018).

Para Koehler *et al.* (2001), o método de previsão de suavização exponencial apresenta três variações: suavização exponencial simples (SES); suavização exponencial com tendência (modelo de HOLT); e suavização exponencial com tendência e sazonalidade (modelo de Winters, HOLT WINTERS).

### 2.3.1 Suavização exponencial simples

A Suavização exponencial simples (Simple Exponential Smoothing – SES) pressupõe que a série oscila em torno de um ciclo com base constante, ou seja, a partir do valor inicial, a base é corrigida a cada período, conforme novos dados são incorporados à série histórica. Além disso, esse modelo é adequado para séries temporais que não apresentam tendência e/ou sazonalidade. Por este motivo, a primeira estimativa é baseada na média de todos os dados históricos de uma série temporal. Para cada nova previsão é feito um ajuste entre o valor previsto e a série real anterior por meio de uma constante de suavização ( $\alpha$ ) (KOEHLER *et al.*, 2001). É sabido que em qualquer tipo de suavização exponencial, pesa-se os valores ou observações recentes mais fortemente do que os valores ou observações antigas. A

função de previsão no método de suavização exponencial simples é representada pela Equação 1 (MAKRIDAKIS *et al.*, 2008):

$$\text{Previsão: } \hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha)E_{t-1} \quad (1)$$

Em que:  $\alpha$  é a constante de suavização para o ciclo, que assume qualquer valor entre 0 e 1 ( $0 \leq \alpha \leq 1$ );  $Y_t$  é o valor da série no período  $t$ ;  $E_{t-1}$  é a base ao final do instante  $t - 1$ .

Quando alfa está mais próximo de 0, é considerado um aprendizado lento, pois o algoritmo está dando mais peso aos dados históricos. Se o valor de alfa estiver mais próximo de 1, ele é conhecido como aprendizado rápido, pois o algoritmo está dando mais peso às observações ou dados recentes. Daí pode-se dizer que as mudanças recentes nos dados vão causando um impacto maior na previsão.

### 2.3.2 Suavização exponencial com tendência (Holt)

Quando aplicado a uma série que apresenta tendência linear (positiva ou negativa), o método de SES fornece previsões que subestimam (ou superestimam) continuamente os valores reais. Dessa forma, com o intuito evitar esse erro foi desenvolvido o HOLT. Esse método é similar ao SES, no entanto, ao invés de suavizar apenas o ciclo, ele utiliza uma nova constante de suavização para modelar a tendência da série (MORETTIN; TOLOI, 2018).

Dessa forma, na Suavização exponencial com tendência (SEH) ou simplesmente HOLT, é adicionada uma segunda constante que reflete o crescimento da previsão da série de um período para outro (PELLEGRINI; FOGLIATTO, 2000). Neste caso, a tendência é a força de longo prazo ou direção geral de movimento em uma série temporal. Ela reflete as mudanças nos dados que ocorrem com o passar do tempo. Essa variável, da mesma forma que o ciclo, também será atualizada exponencialmente e aplicada no cálculo da previsão. Assim, o modelo é adequado para séries temporais que apresentam tendência, mas ausência de sazonalidade. A função de previsão no método é representada pelas Equações de 2 a 4 (MAKRIDAKIS *et al.*, 2008):

$$\text{Previsão: } \hat{y}_{t+n} = E_t + nT_t \quad (2)$$

$$\text{Ciclo: } E_t = \alpha Y_t + (1 - \alpha)(E_{t-1} + T_{t-1}) \quad (3)$$

$$\text{Tendência: } T_t = \beta(E_t - E_{t-1}) + (1 - \beta)T_{t-1} \quad (4)$$

Em que:  $\beta$  é a constante de suavização para a tendência, que assume qualquer valor entre 0 e 1 ( $0 \leq \beta \leq 1$ );  $E_t$  é a base no período  $t$ ;  $T_{t-1}$  é a tendência no instante  $t - 1$ ;  $n$  representa os períodos de tempo no futuro, em que  $n = 1, 2, 3, \dots, p$  e assim por diante.

### 2.3.3 Suavização exponencial com tendência e sazonalidade (Holt Winters)

Para séries temporais que apresentam um padrão de comportamento mais complexo, Winters (1960) desenvolveu um método com uma nova forma de suavização, o HOLT WINTERS, também conhecido como Suavização exponencial com tendência e sazonalidade (SEHW). Este envolve a estimativa de três coeficientes de suavização associados ao ciclo, tendência e fatores sazonais, ou seja, SES, SEH e SEHW (MORETTIN; TOLOI, 2018).

O método Holt Winters é uma extensão da suavização exponencial e aplicado a séries temporais univariadas (PAPASTEFANOPOULOS *et al.*, 2020). É adequado para previsões de curto prazo e usa a função de máxima verossimilhança para estimar os coeficientes de suavização (PAPASTEFANOPOULOS *et al.*, 2020), (AWAJAN *et al.*, 2018).

Este método não necessita de grande armazenamento de dados e é simples (AWAJAN *et al.*, 2018). Além disso, existem dois modelos de SEHW baseados na componente sazonal: aditivos, que se tem a amplitude da variação sazonal constante ao longo da série; multiplicativos, que é utilizado quando a amplitude da variação sazonal da série aumenta ou reduz ao longo do tempo (AWAJAN *et al.*, 2018).

Assim, o modelo aditivo é recomendado para previsões que apresentam dados com tendência e sazonalidade que não aumentam com o tempo (PAPASTEFANOPOULOS *et al.*, 2020). Matematicamente, o modelo aditivo é expresso na forma das Equações 5 a 8 (MAKRIDAKIS *et al.*, 2008):

$$\text{Previsão: } \hat{Y}_{t+n} = (E_t + nT_t) + s_{t+n-p} \quad (5)$$

$$\text{Ciclo: } E_t = \alpha(Y_t - s_{t-p}) + (1 - \alpha)(E_{t-1} + T_{t-1}) \quad (6)$$

$$\text{Tendência: } T_t = \beta(E_t - E_{t-1}) + (1 - \beta)T_{t-1} \quad (7)$$

$$\text{Sazonalidade: } S_t = \gamma(Y_t - E_t) + (1 - \gamma)s_{t-p} \quad (8)$$

Em que:  $E_t$ ,  $T_t$ ,  $S_t$ , representam respectivamente, as estimativas de ciclo, tendência e sazonalidade no tempo  $t$ . Já  $\alpha$ ,  $\beta$ ,  $\gamma$ , são as respectivas constantes de suavização,

podem assumir qualquer valores entre 0 e 1 ( $0 \leq \alpha \leq 1$ ;  $0 \leq \beta \leq 1$ ;  $0 \leq \gamma \leq 1$ ),  $p$  representa o período da sazonalidade e  $n$  denota o horizonte de previsão.

A versão multiplicativa é muito semelhante ao método aditivo, sendo recomendada para previsões que apresente aumento de sazonalidade com o tempo. No entanto, se uma série de dados contém alguns valores iguais a zero, o método SEHW multiplicativo não pode ser usado. Matematicamente este modelo é expresso pelas Equações 9 a 12 (MAKRIDAKIS *et al.*, 2008):

$$\text{Previsão: } \hat{Y}_{t+n} = (E_t + nT_t)S_{t+n-p} \quad (9)$$

$$\text{Ciclo: } E_t = \alpha \left( \frac{Y_t}{S_{t-p}} \right) + (1 - \alpha)(E_{t-1} + T_{t-1}) \quad (10)$$

$$\text{Tendência: } T_t = \beta(E_t - E_{t-1}) + (1 - \beta)T_{t-1} \quad (11)$$

$$\text{Sazonalidade: } S_t = \gamma \left( \frac{Y_t}{E_t} \right) + (1 - \gamma)s_{t-p} \quad (12)$$

As principais vantagens dos métodos da família Holt Winters são baixos custos computacionais, cálculos rápidos, simplicidade e robustez (TRATAR, 2015).

## 2.4 Família arima

Conhecidos como modelos Box e Jenkins, os métodos abordados de previsão de série temporal que compõe essa família são: Autorregressivo (*Auto Regressive - AR*), Médias Móveis (*Moving Average - MA*), Autorregressivo e Médias Móveis (*Auto Regressive Moving Average - ARMA*), e Autorregressivo Integrado de Médias Móveis (*Auto Regressive Integrated Moving Average - ARIMA*) (BOX *et al.*, 2015).

Os modelos ARIMA são um dos mais conhecidos e amplamente usados na previsão de séries temporais, sobretudo os que usam base linear (PAPASTEFANOPOULOS *et al.*, 2020). Nos modelos ARIMA, uma correlação linear é considerada entre as amostras temporais e são encontrados padrões de correlação entre as observações (PAPASTEFANOPOULOS *et al.*, 2020). Suas propriedades estatísticas e sua capacidade de implementar vários modelos de suavização exponencial contribuíram para sua popularidade e adoção generalizada (YULE, 1926; MCKENZIE, 1984).

Tais modelos tentam explorar as dependências lineares nas observações, a fim de extrair padrões locais, enquanto remove o ruído de alta frequência dos dados (ADHIKARI; AGRAWAL, 2013).

A modelagem ARIMA usa técnicas de correlação na geração de previsões. Pode ser usado para modelar padrões que não podem ser visíveis nos dados plotados. Além disso, os modelos podem ser misturados com variáveis exógenas, mas certifique-se de que o variáveis exógenas têm pontos de dados suficientes para cobrir o número adicional de períodos de previsão (GLANTZ; MUN, 2011).

Existem muitas razões pelas quais um modelo ARIMA é superior às séries temporais comuns, análise e regressões multivariadas. A descoberta comum na análise de séries temporais e regressão multivariada é que os erros residuais são correlacionados com seus próprios valores defasados. Esta correlação viola a suposição padrão da teoria de regressão que os distúrbios não estão correlacionados com outros distúrbios (GLANTZ; MUN, 2011).

Esses modelos são casos particulares de um modelo de filtro linear. Assim, supõe-se que a série temporal seja gerada por meio de um filtro linear (ou sistema linear), cuja entrada é um ruído branco, conforme Figura 4.

**Figura 4 - Representação de uma série temporal como saída de um filtro linear**



**Fonte: Box et al. (2015)**

Os modelos ARIMA são, portanto, estabelecidos pela combinação das operações autoregressiva (AR), que corresponde ao número do valor defasado do residual no modelo de previsão incondicional. Em essência, o modelo captura a variação histórica de dados reais para um modelo de previsão e usa essa variação ou residual para criar um modelo de previsão melhor; integração (I), corresponde ao número de diferenciações que as séries temporais a serem previstas passam para tornar os dados estacionários. Este elemento é responsável por quaisquer taxas de crescimento não linear existentes nos dados; e média móvel (MA), que é essencialmente a média móvel defasada de erros de previsão. Ao incorporar essa média móvel, o modelo em essência aprende com seus erros ou erros de previsão e corrige através de um cálculo de média móvel (BOX et al., 2015); (GLANTZ; MUN, 2011).

Para séries não estacionárias, aplica-se o modelo ARIMA, porém, é necessário utilizar algum tipo de transformação para estacionar os dados.

Modelos ARIMA  $(p, d, q)$  são obtidos tomando a diferença de séries de grau  $d$  e adicionando ao modelo ARMA  $(p, q)$  para o processo de estabilização. Nos modelos ARIMA  $(p, d, q)$ ,  $p$  é a ordem do modelo Autoregressivo (AR),  $q$  é a ordem do modelo de média móvel (MA) e  $d$  representa quantas diferenças são necessárias para tornar a série sem tendência. Assim, se série temporal apresente comportamento estacionário, isto é, séries que se desenvolvem no tempo ao redor de uma média constante, utiliza-se os modelos AR  $(p)$ , MA  $(q)$  ou ARMA  $(p, q)$  de forma direta (MORETTIN; TOLOI, 2018).

Outro benefício dos modelos ARIMA é sua capacidade de acomodar sistemas governados por dinâmicas que mudam ao longo do tempo, atualizando o modelo com base em eventos recentes para prever o estado futuro do sistema (KANE *et al.*, 2014).

#### 2.4.1 Autorregressivo (AR)

A estrutura de modelo mais simples e amplamente aplicada é o modelo AR (OSHODI *et al.*, 2017). Este estabelece que a observação da série temporal atual é uma combinação linear das observações anteriores excitado por um ruído branco.

No modelo AR, pressupõe que a previsão da série é dada pela combinação linear  $y_{t-1} + \dots + y_{t-p}$  dos termos passados da série, multiplicados pelos coeficientes livres do modelo  $\varphi_p$ , e adicionado o erro de especificação  $\varepsilon_t$ , que é associado ao erro da previsão, conforme Equação 13.

$$y_t = c + \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t \quad (13)$$

Em que:  $\varphi_1$  é o coeficiente de ordem  $p$ ,  $y_{t-1}$  são os valores passados da série que, linearmente combinados, juntamente com o erro  $\varepsilon_t$  ruído branco, e  $c$  representa o intercepto, que está relacionado à média ( $\mu$ ) da série e aos coeficientes de ordem  $\varphi_1, \varphi_2, \dots, \varphi_p$  pela Equação 14.

$$\mu = \frac{c}{1 - \varphi_1 - \varphi_2 - \dots - \varphi_p} \quad (14)$$

Nesse modelo, os valores dos coeficientes livres podem ser encontrados por equações lineares de forma determinística aplicando as equações de Yule-Walker,

garantindo o mínimo erro quadrático médio (BOX *et al.*, 1994). Tais equações definem os coeficientes baseando-se nas funções de autocorrelação.

#### 2.4.2 Médias móveis (MA)

No caso do modelo MA, o valor da série temporal atual é uma função de choques aleatórios passados que, na prática são tidos como os erros temporais das previsões. Assim, o modelo usa valores defasados do erro de previsão para melhorar a previsão atual, conforme Equação 15 (MOHD *et al.*, 2009).

$$\hat{Y}_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (15)$$

Em que:  $\mu$  é média da série temporal,  $\theta_1, \theta_2 \dots \theta_q$  são os coeficientes de ordem  $q$ .

Note que, diferentemente do modelo AR, não há fórmula fechada para cálculo dos coeficientes livres do modelo MA, sendo necessária a aplicação de alguma técnica de otimização para tal. A forma clássica lança mão do estimador de máxima verossimilhança (BOX *et al.*, 2015).

#### 2.4.3 Autorregressivo e médias móveis (ARMA)

Combinando AR e MA, obtém-se uma estrutura aumentada, conhecido como ARMA. O modelo pode ser usado para séries temporais apenas se a condição de estacionariedade for satisfeita, ou seja, se a média e a variância são constantes ao tempo (SHADAB *et al.* 2020). Assim, a previsão é definida de acordo com a Equação 16.

$$\hat{Y}_t = c + \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (16)$$

Similarmente ao modelo MA, não há fórmula fechada para cálculo dos coeficientes do modelo ARMA.

#### 2.4.4 Autorregressivo integrado de médias móveis (ARIMA)

O modelo ARIMA é escrito como ARIMA ( $p, d, q$ ), em que ' $p$ ' descreve a ordem do processo auto regressivo, ' $d$ ' define a ordem dos dados estacionários e ' $q$ ' representa a ordem do processo de média móvel (SHADAB *et al.*, 2019).



No caso de uma diferenciação de série temporal não estacionária o modelo ARIMA é usado, que é uma forma generalizada de todos os modelos apresentados anteriormente. Além disso, para previsões de curto prazo, o ARIMA pode fornecer resultados até mais precisos quando comparados a alguns dos métodos populares de aprendizado de máquina, como o perceptron multicamadas, vetor de suporte máquina e memória longa de curto prazo (MAKRIDAKIS *et al.*, 2018).

A metodologia ARIMA inclui três etapas iterativas: identificação do modelo, estimativa dos coeficientes e verificação. A ideia básica da identificação é que, se uma série temporal for gerada a partir de um processo ARIMA, ela deve ter algumas propriedades teóricas de autocorrelação. Combinando os padrões de autocorrelação empíricos com os teóricos, muitas vezes é possível identificar um ou vários modelos potenciais para a série temporal dada (BROCKWELL; DAVIS, 2016).

Box e Jenkins (1970) propuseram usar a função de autocorrelação (ACF) dos dados da amostra como ferramentas básicas para identificar a ordem do modelo ARIMA. Ainda nessa etapa, geralmente se faz necessário realizar a transformação dos dados para tornar a série temporal estacionária.

Quando a série temporal observada apresenta tendência e heterocedasticidade, a diferenciação e a transformação de potência são frequentemente aplicadas aos dados para remover a tendência e estabilizar a variância antes que um modelo ARIMA possa ser ajustado (BROCKWELL; DAVIS, 2016).

Uma vez que um modelo provisório é especificado, a estimativa dos coeficientes do modelo é direta. Os coeficientes são estimados de forma que uma medida geral de erros seja minimizada. Isso pode ser feito com um procedimento de otimização não linear.

Este modelo tem muitos exemplos práticos de predição de sucesso em literaturas em diferentes campos.

A primeira diferenciação ( $\nabla$ ) para retirada da tendência não estacionária  $Y_t$  da série temporal é obtida pela Equação 17.

$$\nabla^1 Y_t = Y_t - Y_{t-1} = Y'_t \quad (17)$$

Se a série ainda não está estacionária  $Y'_t$ , o processo de diferenciação é repetido por  $d$  vezes até ficar estacionário. Assim, para o processo de tomada de diferença é dada por:

$$\nabla^d Y_t = \nabla^{d-1} Y_t - \nabla^{d-1} Y_{t-1} \quad (18)$$

A expressão do modelo ARIMA ( $p, d, q$ ) pode ser definida como ARMA da Equação 16. Note que se  $d=0$  então o modelo ARIMA se torna ARMA (ALMASARWEH; ALWADI, 2018).

## 2.5 Ensemble

Uma estratégia muito comum consiste em combinar previsões de diversos modelos são os chamados *Ensemble*. Isso não é novidade e vem sendo explorado desde Bates e Granger (1969), com resultados bastante promissores. Estas abordagens são conhecidas como *ensemble learning* ou aprendizado em conjunto (Montero-Manso *et al.*, 2020, Moon *et al.*, 2020, Chen and Liu, 2020, Li *et al.*, 2020). A principal motivação para combinar modelos é que, na maioria dos casos, um conjunto pode obter melhor desempenho na previsão do que um modelo único (Bolón-Canedo; Alonso-Betanzos, 2019).

Sharkey (1999) mostrou que para a previsão de séries temporais, os melhores resultados são alcançados através da combinação de diferentes modelos de previsão ao invés da seleção daquele que apresenta melhor desempenho individual. Existem diversas formas para combinar preditores, dentre elas, a metodologia *Ensemble*. Assim, para o problema proposto, tal combinação foi investigada.

Esta abordagem por si só pode levar um melhor desempenho dos resultados, segundo Perrone e Cooper (1992). Sua vantagem é ser fácil de entender e implementar e não aumentar o erro esperado (BISHOP *et al.*, 1995).

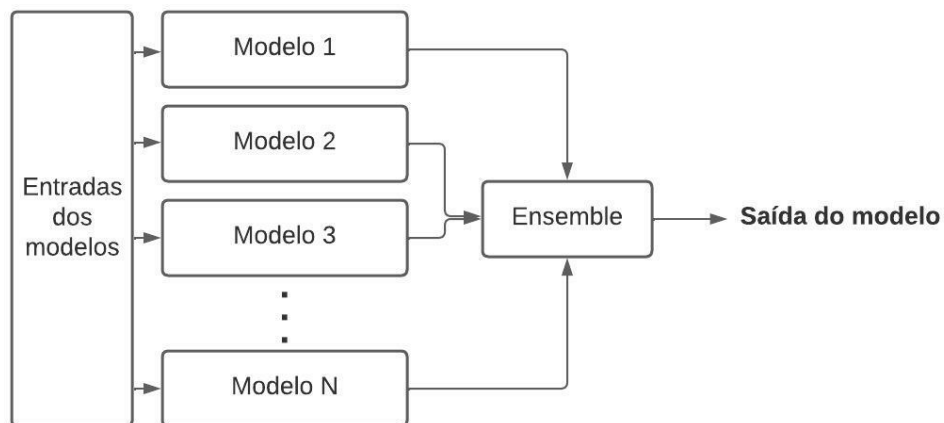
A seleção dos modelos do *ensemble* pode melhorar potencialmente a exatidão geral da previsão (ADHIKARIET *et al.*, 2015). Isso porque nem todos os modelos produzem boas previsões para séries temporais específicas. Em geral, o principal desafio na seleção é a escolha de uma heurística ou meta-heurística eficiente e rápida. Observa-se na literatura trabalhos que propõem abordagens de seleção para constituir *ensembles* de modo a minimizar o erro das previsões (Chen and Liu, 2020, Li *et al.*, 2020).

A estratégia de *ensemble* adotada neste estudo para combinar os modelos foi média simples. Esta soma as saídas (previsões) dos modelos e divide pela quantidade dos modelos para obter a saída final do *ensemble*.

Entre as alternativas de ensembles, o método de média simples é uma das formas mais comuns de combinar modelos apresentada na literatura (Kourentzes, Barrow e Crone (2014), Krishnamurti *et al.* (2000)). Neste sentido, o método de média simples assume que os modelos individuais contribuem igualmente no processo de combinação, isto é, um modelo com pouca acurácia contribui na combinação igual aos modelos muito acurados.

Na Figura 5 é apresentado um modelo genérico de *Ensemble*. É possível verificar que as saídas de todos os modelos que foram devidamente treinados para o problema, são combinadas a fim de melhorar a solução geral.

**Figura 5 - Modelo genérico de Ensemble**



**Fonte: Adaptado de Bates e Granger (1969)**

### **3 METODOLOGIA**

Esta seção tem por objetivo apresentar as etapas que compõem os procedimentos para a realização desta investigação. Logo, é abordado o enquadramento metodológico da pesquisa (Seção 3.1), e na sequência as etapas necessárias para realização da previsão do preço do açúcar (Seção 3.2).

#### **3.1 Enquadramento metodológico**

Para enquadramento metodológico da pesquisa utilizou-se as definições dadas por Gil (2018). Sendo assim, do ponto de vista dos objetivos, assume o perfil de pesquisa como exploratória, que está em concordância com outras fontes que darão base ao assunto abordado, como é o caso da pesquisa apresentada que se utiliza de bibliografias (literatura técnica). Assim, este estudo é uma hibridização dos conhecimentos já disponíveis na literatura com uma ampliação.

Do ponto de vista da sua natureza, este trabalho classifica-se como uma pesquisa aplicada, pois servirá para gerar conhecimentos para aplicação prática na solução de um problema específico, além de caracterizar-se como uma investigação concebida pelo interesse em adquirir novos conhecimentos. A pesquisa aplicada é a junção do conhecimento disponível e sua ampliação.

Em relação a abordagem do problema, esta caracterizou-se como quantitativa, pois faz uso de modelos preditivos exclusivamente matemáticos/estatísticos.

Do ponto de vista dos procedimentos técnicos, enquadra-se como bibliográfica e levantamento. Bibliográfica pois foi fundamentada na exploração de estudos científicos sobre métodos de previsão de preço. E levantamento, pois segue as seguintes etapas: definição do objetivo da pesquisa, definição da população e da amostra, coleta dos dados, processamento dos dados (tabulação), análise dos dados, apresentação e discussão dos resultados.

Quanto à área de ciência, o estudo é classificado como metodológico, pois apresenta uma metodologia que tem o objetivo de aperfeiçoar e sistematizar a

## 3.2 Estruturação do modelo

Com relação aos métodos de previsão, dentre os diversos modelos lineares, essa pesquisa se limitou as técnicas: AR, ARMA, ARIMA, SES, HOLT, HOLT WINTER e Ensemble de média simples. E como em todo trabalho científico, é interessante delimitar um sistema, empresa, ou segmento de mercado específico relacionado aos aspectos práticos da pesquisa, assim, o campo de aplicação limitou-se em realizar a previsão do preço do açúcar.

Para construção de um modelo de previsão, quatro fases são muito importantes: coleta de dados, pré-processamento, previsão e pós-processamento (KHALID; JAVAID, 2020).

### 3.2.1 Base de dados

Para a presente pesquisa, foram utilizadas quatro bases de dados, uma disponível no site do Centro de Estudos Avançados em Economia Aplicada da Escola de Agronomia Luiz de Queiroz (ESALQ) - Universidade de São Paulo (USP) (CEPEA), abordando dados do Brasil com um corte temporal mensal de Janeiro de 2002 a Janeiro de 2022, totalizando 241 amostras, e as outras três encontram-se no site Index Mundi, trazendo dados dos Estados Unidos e Mundo, compreendendo um período mensal de Março de 1992 a Março de 2022, totalizando 361 amostras e União Europeia com dados mensal de Março de 1997 a Março de 2022, com 301 amostras.

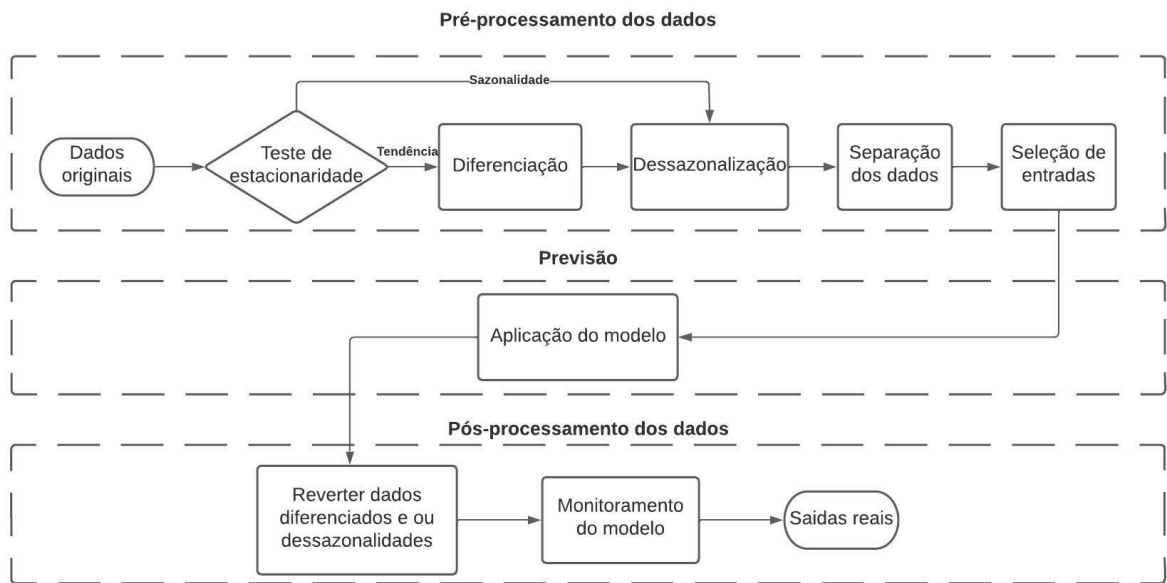
Os dados utilizados estão representados em dólar americano por quilograma (US\$/Kg). Os Indicadores de preços para o Mercado Internacional no banco de dados do CEPEA constituem-se em médias ponderadas mensais dos valores efetivamente comercializados por dois tipos de açúcar: VHP e Cristal. As médias mensais são calculadas utilizando-se como fator de ponderação os volumes negociados. A conversão em dólares americanos é feita pela taxa de câmbio média mensal, calculada a partir dos valores do dólar comercial, preço de venda, cotado diariamente às 16h30.

Ainda, a Index Mundi é uma base de dados que reúne fatos e estatísticas de várias fontes e os transforma em recursos visuais fáceis de usar, transformando dados brutos de todo o mundo em informações úteis para um público global. Capturam

estatísticas que estão espalhadas ou ocultas e as apresentam por meio de mapas, gráficos e tabelas.

De posse dos dados, deve-se seguir as fases de processamento seguinte (pré-processamento, previsão e pós-processamento) com o objetivo de identificar e desenvolver a técnica de previsão que melhor se adapte, conforme Figura 6.

**Figura 6 - Etapas para construção de um modelo de previsão**



**Fonte: Autoria própria (2022)**

### 3.2.2 Pré-processamento dos dados

Uma vez coletados e analisados os dados da série temporal, deve-se analisar pela técnica de previsão mais apropriada. Outros fatores que merecem destaque são: a) experiência passada com a aplicação de determinada técnica; b) período de planejamento para o qual se necessita da previsão (TUBINO, 2017).

Assim, inicialmente deve ser realizado uma análise estatística e gráfica da série temporal com o objetivo de identificar seu comportamento, que pode apresentar comportamento horizontal, tendência, sazonal e cíclico, conforme discutido na seção 2.1.

Os modelos de previsão abordados na presente pesquisa apresentam diferentes transformações no pré-processamento dos dados antes que se inicie o processo de previsão.

O método ARIMA, possui a fase de diferenciação (I) dos dados, de respectiva ordem ( $d$ ), e por isso, caso seja identificado que a série apresente comportamento de tendência, é necessário aplicar diferenciações para torná-la estacionária. Ainda, se for identificado que além de tendência a série é composta também por sazonalidade, aplica-se o processo de dessazonalização, retirando esse comportamento antes que os dados entrem no modelo preditivo (BOX *et al.*, 2015).

Em sequência, para verificar a estacionariedade da série, foi aplicado os testes de KPSS (Kwiatkowski Phillips Schmidt Shin) que usa uma regressão automática paramétrica, ignorando qualquer correlação serial na regressão do teste, e ADF (Dickey-Fuller Aumentado) que verifica a hipótese nula de um processo de raiz unitária contra a hipótese de um processo estacionário.

Ainda, a presença de sazonalidade pode ser verificada também por um teste não-paramétrico, teste de Friedman (MORETTIN E TOLOI, 2006), que também considera duas hipóteses:

$$H_0 = p - \text{valor} > 5\% \text{ Não há sazonalidade na série;}$$

$$H_1 = p - \text{valor} < 5\% \text{ Há sazonalidade na série.}$$

A estacionariedade da série temporal é verificada e, se esta não for considerada estacionária, os dados são diferenciados até que não haja padrões perceptíveis como tendência ou ciclicidade nos dados, tornando-a estacionária (SHADAB *et al.* 2020), conforme Equação 18.

Além disso, se também for identificado um componente sazonal, um processo estatístico conhecido como dessazonalização é aplicado nos dados (BALLINI, 2000), este é expresso na Equação 19.

$$z_t = \frac{x_t - \mu_m}{\sigma_m} \quad (19)$$

Em que:  $x_t$  é o dado que está sendo dessazonalizado,  $\mu_m$  é a média e  $\sigma_m$ , é o desvio padrão de cada valor de sazonalidade  $m$ . Por exemplo, para sazonalidade diária,  $m=7$ , mensal,  $m=12$ .

Para aplicação dos métodos de previsão, é necessário que os dados sejam divididos em dois grupos: treinamento e teste, em que os primeiros são usados para estimar quaisquer parâmetros de um método de previsão, enquanto e os dados de teste são usados para avaliar sua precisão. Habitualmente divide-se o conjunto de dados em proporções de treinamento que variam entre 70% a 90%, assim, o conjunto restante é separado para testar os parâmetros do modelo com dados não

utilizados pelo treinamento (SILVA; SPATTI; FLAUZINO, 2016). A fase de treinamento busca os parâmetros ótimos do modelo (JEREZ; KRISTJANPOLLER, 2020).

De acordo com Belotti (2019), é possível identificar se a generalização das soluções está com um erro aceitável para estimar o comportamento da série em situações reais, por meio, do processo de treinamento. Ainda, as amostras do conjunto de teste não podem ser utilizadas no treinamento.

Tendo os dados estacionários para os respectivos modelos que exigem essa condição e com os parâmetros definidos, segue-se para a etapa de previsão, discutida na seguinte seção.

### 3.2.3 Previsão

Definida a técnica de previsão e a aplicação dos dados passados para obtenção dos parâmetros necessários, podem-se obter as projeções futuras da série temporal. Quanto maior for o horizonte pretendido, menor a confiabilidade na previsão prevista (BOX *et al.*, 2015).

Para analisar as séries temporais, foi utilizado a linguagem de programação do software estatístico R e o pacote especializado de previsão (TEAM 2013), (HYNDMAN; KHANDAKAR, 2008).

Para acompanhamento das séries temporais, foram considerados os preços mensais, os quais são analisados pelos gráficos da Função de Autocorrelação (ACF) e da Função de Autocorrelação Parcial (PACF) (BOX *et al.*, 2015).

Para estimar os coeficientes da família ARIMA, selecionar a ordem correta dos modelos é de extrema importância (ZA'ER *et al.*, 2012), pois os desempenhos dos modelos são criticamente afetados pela seleção da ordem do modelo (ZHANG *et al.*, 2017; FENGA 2017). A identificação da ordem do modelo é uma etapa crucial no processo de estimativa dos coeficientes AR/ARMA precisos (GIANNAKIS; MENDEL, 1990).

Existem muitas técnicas para seleção dos coeficientes de ordem AR, ARMA e ARIMA, por exemplo métodos de máxima verossimilhança ou mínimos quadrados que fornecem o melhor ajuste aos dados. Para maior conhecimento, consultar Critério de Informação de Akaike (*Akaike Information Criterion - AIC*), Critério de Informação de Akaike de Segunda Ordem (*Second-order Akaike Information Criterion -*



AICc) (AKAIKE, 1970), Critério de Informação Bayesiano (*Bayesian Information Criterion* - BIC) (SCHWARZ, 1978) e máxima verossimilhança (MORETTIN; TOLOI, 2018).

Para os modelos da família Holt Winters, foi utilizado a função *ets* do pacote *forecast*, que é descrito da seguinte forma, (HYNDMAN, 2008):

- i) Para cada série, são aplicados todos os modelos e otimizados os parâmetros (parâmetros de suavização e variável de estado inicial);
- ii) O algoritmo seleciona o melhor modelo de acordo com o AIC, AICc e BIC;
- iii) São produzidas previsões pontuais utilizando o melhor modelo (com os parâmetros otimizados);

A seleção do método é baseada no reconhecimento de comportamentos chave da série temporal (tendência e sazonalidade) e na forma como estes entram no método de suavização (por exemplo, de forma aditiva, amortecida ou multiplicativa). O algoritmo faz a identificação da necessidade de amortecimento da série e usa o coeficiente de amortecimento *phi* para estimar de forma mais conservadora as tendências previstas, caso seja necessário.

Os métodos que incluem uma tendência amortecida provaram ser muito bem-sucedidos e são indiscutivelmente mais populares quando as previsões são necessárias automaticamente para muitas séries (Gardner & McKenzie, 1985).

Para os modelos da família ARIMA, os parâmetros foram calculados por meio da função *auto.arima*, que utiliza uma variação do algoritmo Hyndman-Khandakar (Hyndman; Khandakar, 2008), retornando a melhor configuração de acordo com o valor AIC, AICc ou BIC. A função realiza uma pesquisa sobre o modelo possível dentro das restrições de ordem fornecidas. A função também imprime os critérios de informação AIC, AICc e BIC, assim como o log da verossimilhança (*log likelihood*) e a variância ( $\sigma^2$ ), que é definida como o quadrado do desvio padrão. O algoritmo procede da seguinte forma:

i) O número de diferenças *d* é determinado usando sucessivas repetições o Teste de Kwiatkowski-Phillips-Schmidt-Shin (KPSS);

ii) Os valores de *p* e *q* são escolhidos através da minimização do AIC após a diferenciação dos dados *d* vezes. Ao invés de considerar cada combinação de *p* e *q*, o algoritmo utiliza um procedimento de pesquisa *stepwise* para percorrer o espaço do modelo, repetindo sucessivas vezes até verificar o menor AIC.

Neste estudo considerou-se o horizonte de 1, 3, 6 e 12 passos à frente.

### 3.2.4 Monitoramento e avaliação comparativa dos modelos

Após realizada a previsão, deve-se realizar o monitoramento e controle do modelo, de modo a possibilitar uma avaliação comparativa entre eles. Basicamente deve-se verificar se as métricas de erros adotadas são satisfeitas. A medida que as previsões forem sendo alcançadas pela série real, deve-se monitorar a extensão do erro entre a série real e a prevista para verificar se a técnica e os coeficientes empregados ainda são válidos. Além disso, várias estatísticas de diagnóstico e gráficos dos resíduos devem ser usados para examinar a qualidade do ajuste (BOX; JENKINS, 1970).

No entanto, se não for adequado, um novo modelo provisório deve ser identificado, que é novamente seguido pelas etapas de estimação dos parâmetros e verificação. As informações de diagnóstico podem ajudar a sugerir modelos alternativos (BOX; JENKINS, 1970).

Uma abordagem comum para análise de séries temporais envolve tentar várias técnicas de modelagem em um conjunto de dados e avaliar como elas explicam o comportamento passado da série temporal. Para a presente pesquisa utilizou-se de medidas quantitativas de acurácia, assim, para comparação das técnicas de previsão, as medidas foram limitadas para: Raiz Erro Quadrático Médio (RMSE – *Root Mean Squared Error*), Erro Absoluto Médio (MAE – *Mean Absolute Error*) e Erro Percentual Absoluto Médio (MAPE – *Mean Absolute Percentual Error*), conforme Equações 20, 21 e 22.

$$RMSE = \sqrt{\frac{1}{n} \sum_i^n (Y_i - \hat{Y}_i)^2} \quad (20)$$

$$MAE = \frac{1}{n} \sum_i^n |Y_i - \hat{Y}_i| \quad (21)$$

$$MAPE = \frac{1}{n} \sum_i^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \quad (22)$$

Em que:  $Y_i$  representa o valor real para a  $i$ -ésima observação na série temporal e  $\hat{Y}_i$  representa o valor previsto ou esperado para essa observação. Essas medidas estimam as diferenças entre os valores reais na série temporal e os valores previstos, ou ajustados, gerados pela técnica de previsão.

## 4 RESULTADOS E DISCUSSÕES

Na análise de séries temporais, um dado estacionário é uma suposição necessária para obter o modelo adequado. Portanto, foi feito o teste Ljung-Box para verificar se estes eram estacionários. Através dos resultados, notou-se que em todas as bases os dados não eram estacionários. Neste caso, constatou-se a presença de tendência, que foi corrigida através de diferenciação dentro do próprio modelo, sendo suficiente para tornar a série temporal estacionária. Este processo foi revertido após o processo de previsão.

Em seguida, foi feita a separação do banco de dados em dois grupos, treinamento, usado exclusivamente para a calibração do modelo, e teste, utilizado para avaliar o desempenho dos modelos estabelecidos, conforme Tabela 1.

**Tabela 1 - Descrição dos dados e conjunto de treinamento e teste**

| <b>Base</b>                  | <b>Amostra total</b> | <b>Treinamento</b> | <b>Teste</b> |
|------------------------------|----------------------|--------------------|--------------|
| CEPEA - Brasil               | 241                  | 193                | 48           |
| Index Mundi - Estados Unidos | 361                  | 289                | 72           |
| Index Mundi - Mundo          | 361                  | 289                | 72           |
| Index Mundi - União Europeia | 301                  | 241                | 60           |

**Fonte: Autoria própria (2022)**

Ainda, foi realizada uma análise descritiva das séries estudadas, conforme Tabela 2.

**Tabela 2 - Análise descritiva dos dados**

| <b>Base</b>                  | <b>Mínimo</b> | <b>Máximo</b> | <b>Média</b> | <b>Desvio Padrão</b> |
|------------------------------|---------------|---------------|--------------|----------------------|
| CEPEA - Brasil               | 0,1045        | 0,5794        | 0,2830       | 0,1085               |
| Index Mundi - Estados Unidos | 0,3800        | 0,8900        | 0,5384       | 0,1100               |
| Index Mundi - Mundo          | 0,1100        | 0,6500        | 0,2907       | 0,1129               |
| Index Mundi - União Europeia | 0,3400        | 0,7800        | 0,5063       | 0,1185               |

**Fonte: Autoria própria (2022)**

Observa-se uma similaridade no tocante ao desvio padrão, mas valores de mínimo com magnitudes até 3 vezes maior em alguns casos. O máximo e a média apresentam uma certa divergência no valor comparativo. Tais observações reforçam a necessidade de analisar-se modelos diferentes em cenários distintos.

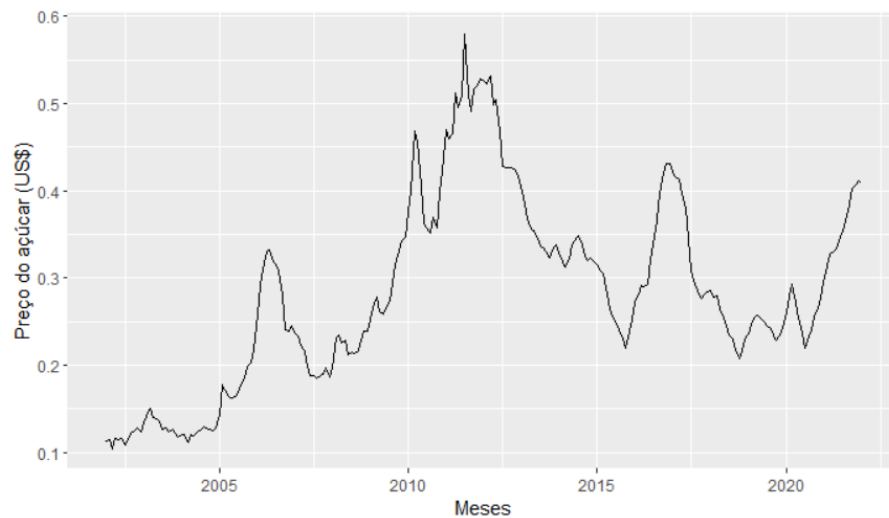
## 4.1 Análise dos dados

A primeira coisa a fazer em qualquer tarefa de análise de dados é plotar os dados, pois estes permitem que muitos recursos sejam visualizados, incluindo padrões, observações incomuns, mudanças ao longo do tempo e relações entre variáveis.

As características que são vistas nos gráficos devem então ser incorporadas, tanto quanto possível, aos métodos de previsão a serem usados. Assim, como o tipo de dados determina qual método de previsão usar, ele também determina quais gráficos são apropriados.

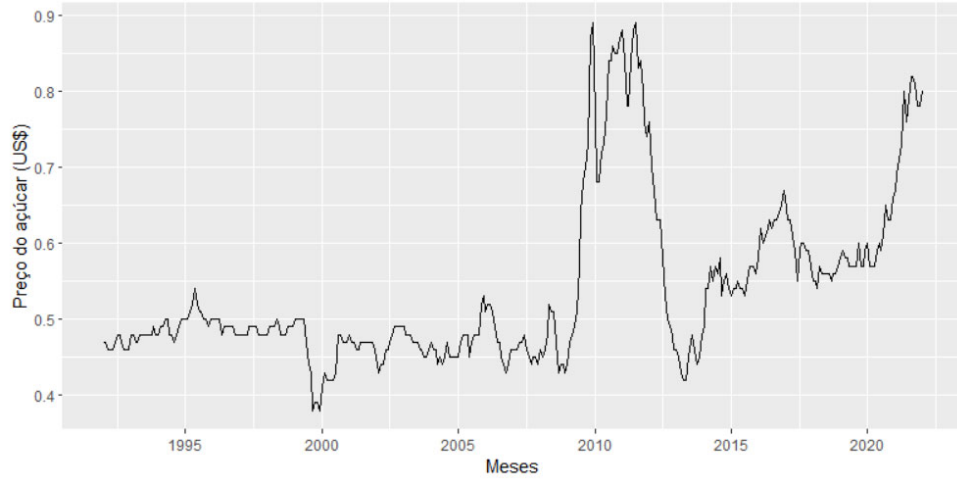
Para dados de séries temporais, as observações são plotadas em relação ao tempo de observação. A seguir, nos Gráficos 1 a 4 são apresentados os dados das 4 séries estudadas.

**Gráfico 1 - Série temporal preço do açúcar – Brasil**



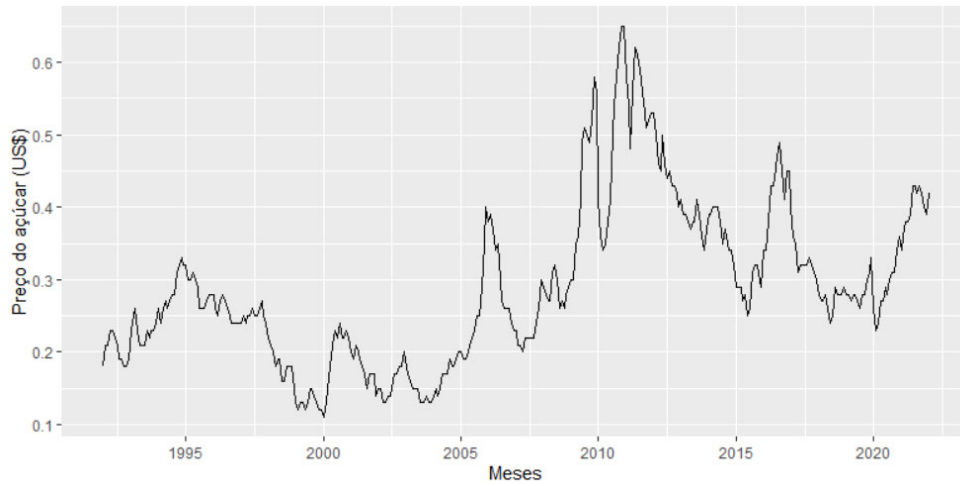
**Fonte: Autoria própria (2022)**

**Gráfico 2 - Preço do açúcar Estados Unidos**



**Fonte: Autoria própria (2022)**

**Gráfico 3 - Preço do açúcar Mundo**



**Fonte: Autoria própria (2022)**

**Gráfico 4 - Preço do açúcar União Europeia**



**Fonte: Autoria própria (2022)**

Analisando as séries, observou-se uma elevação considerável nos preços em 2000. Bentivoglio et al. (2016) explicam que esse aumento foi visível para *commodities* agrícolas. Ressalta-se que os preços permaneceram relativamente estáveis até o final de 2006, enquanto de 2007 a 2008 mais que dobraram, voltando a cair em 2009, atingindo os níveis de 2006.

No segundo semestre de 2010 os preços voltaram a subir, seguidos por uma ligeira queda em 2011. Há uma vasta literatura sobre as causas desse aumento (SCHUTTER, 2010; JACKS *et al.*, 2011; HUCHET-BOURDON, 2011; BALCOMBE, 2010). Vários estudos levaram alguns debates sobre o papel da especulação, aumentos nos preços da energia, mudanças na política de exportação, a desvalorização cambial do dólar e, principalmente, no caso das *commodities* alimentícias, o papel dos biocombustíveis.

Ainda, nas séries Brasil, Estados Unidos e Mundo observa-se um crescimento no preço desde 2020, marcado pelo alta do dólar frente ao real, queda das cotações na bolsa de Nova Iorque e baixa dos preços do petróleo (CONAB, 2020). O aumento do valor do açúcar no período anterior está positivamente associado ao aumento da cotação do açúcar no período atual, resultado que tende a ilustrar a dinâmica típica dos mercados de *commodities*.

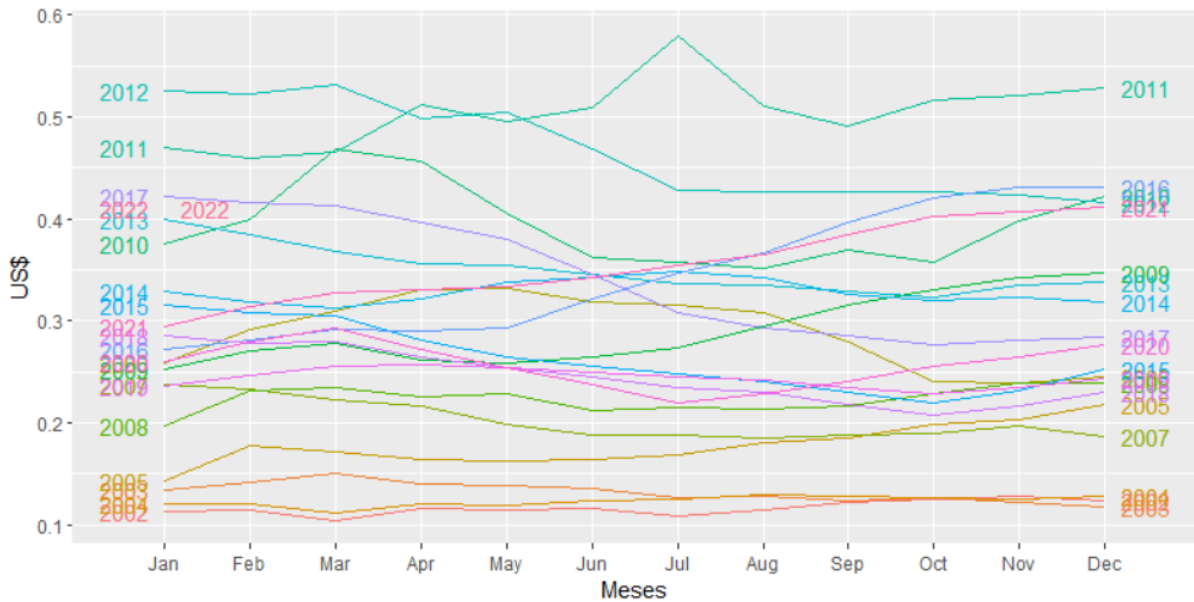
Antes de escolher um método de previsão, primeiro é preciso identificar os padrões nos dados e, em seguida, escolher um método capaz de capturá-los adequadamente.

Se as flutuações não são de frequência fixa, então elas são cíclicas; se a frequência é imutável e associada a algum aspecto do calendário, então o padrão é sazonal. Em geral, a duração média dos ciclos é maior do que a duração de um padrão sazonal, e as magnitudes dos ciclos tendem a ser mais variáveis do que as magnitudes dos padrões sazonais.

Assim, visualmente pode-se observar que as séries não apresentam um padrão sazonal definido. Nas séries Brasil, Estados Unidos e Mundo observou-se picos de alta e baixa no preço, e também um comportamento de tendência positiva indica uma linha crescente, pois há alteração na média e na variância, revelando que a série é não-estacionária. No entanto, analisando a União Europeia, é vista uma tendência de decréscimo.

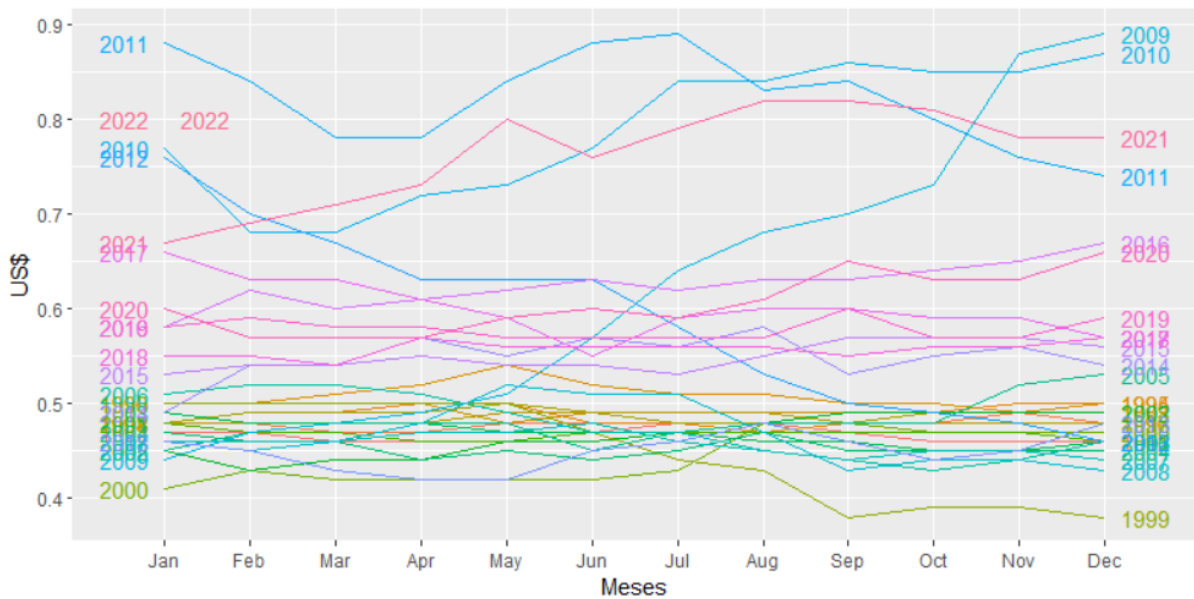
Um gráfico sazonal é semelhante a um gráfico de tempo, exceto que os dados são plotados em relação às “estações” individuais nas quais os dados foram observados, conforme Gráficos 5, 6, 7 e 8.

**Gráfico 5 - Sazonal Brasil**



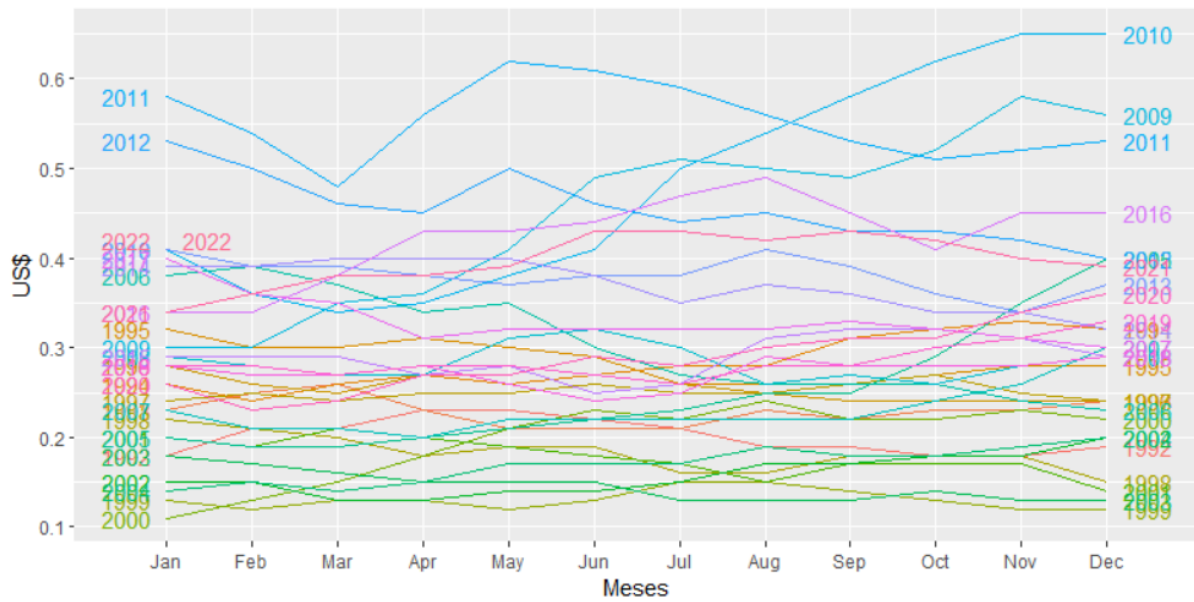
Fonte: Aatoria própria (2022)

**Gráfico 6 – Sazonal Estados Unidos**



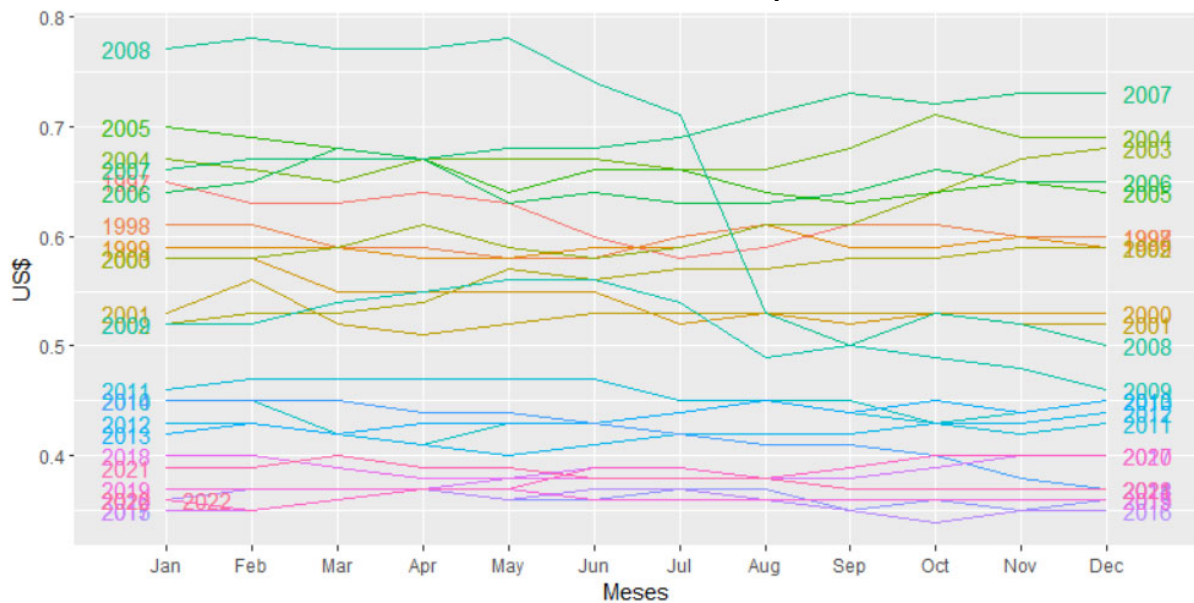
Fonte: Aatoria própria (2022)

Gráfico 7 - Sazonal Mundo



Fonte: Autoria própria (2022)

Gráfico 8 - Sazonal União Europeia



Fonte: Autoria própria (2022)

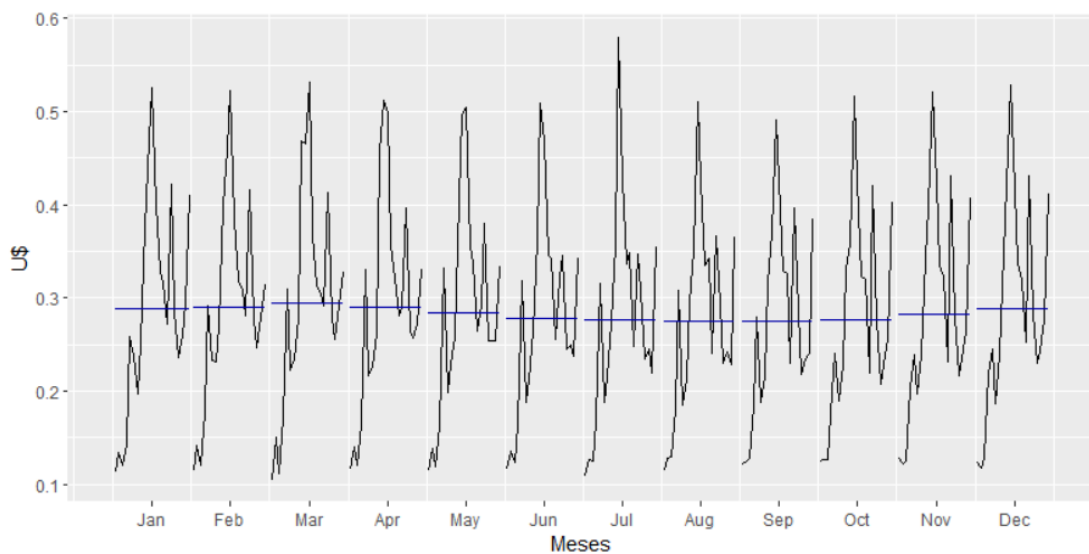
Estes são exatamente os mesmos dados mostrados anteriormente, mas agora plotados a cada temporada de forma sobreposta. Um gráfico deste tipo permite que o padrão sazonal subjacente seja visto com mais clareza e é especialmente útil na identificação de anos em que o padrão muda, para granularidade mensal.



Nesse caso, pode-se observar que não existe um padrão de repetição sazonal entre os anos. Ou seja, os dados não apresentam comportamento sazonal definido.

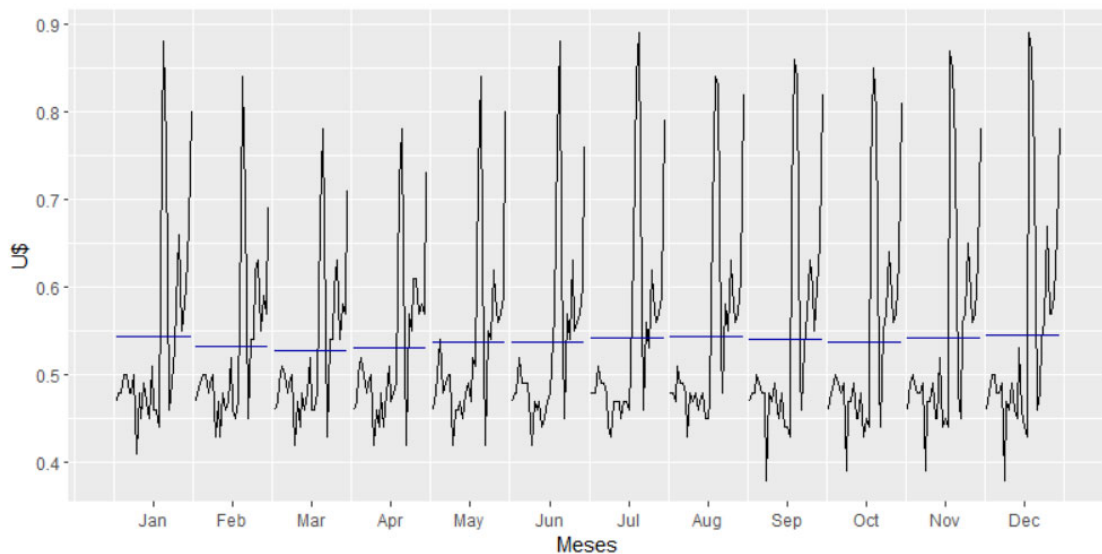
Ainda observando a sazonalidade das séries temporais, foi desenvolvido um gráfico alternativo que enfatiza os padrões sazonais, em que os dados para cada estação são representados por parcelas de subséries sazonais, conforme Gráficos 9, 10, 11 e 12.

**Gráfico 9 - Subsérie sazonal Brasil**

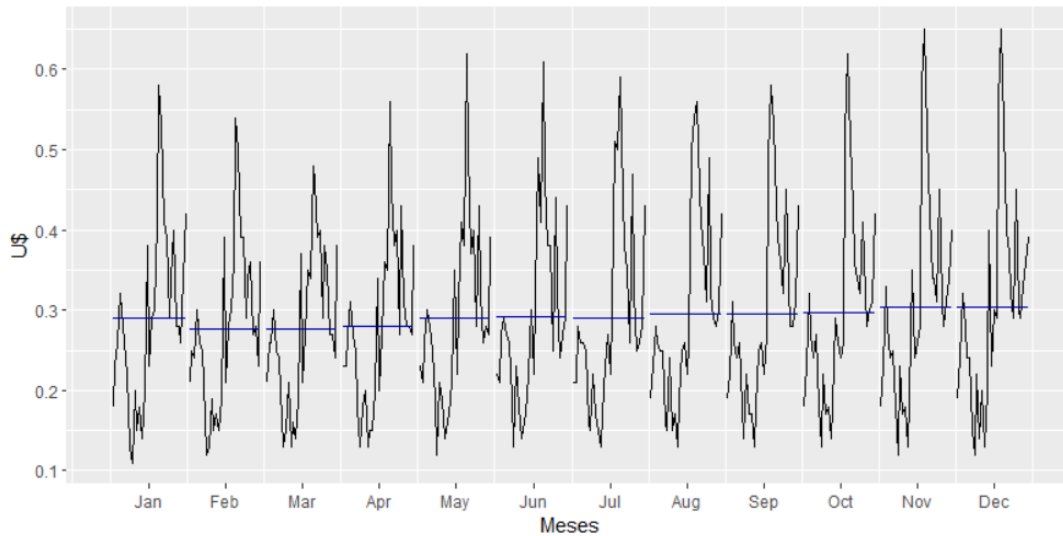


Fonte: Autoria própria (2022)

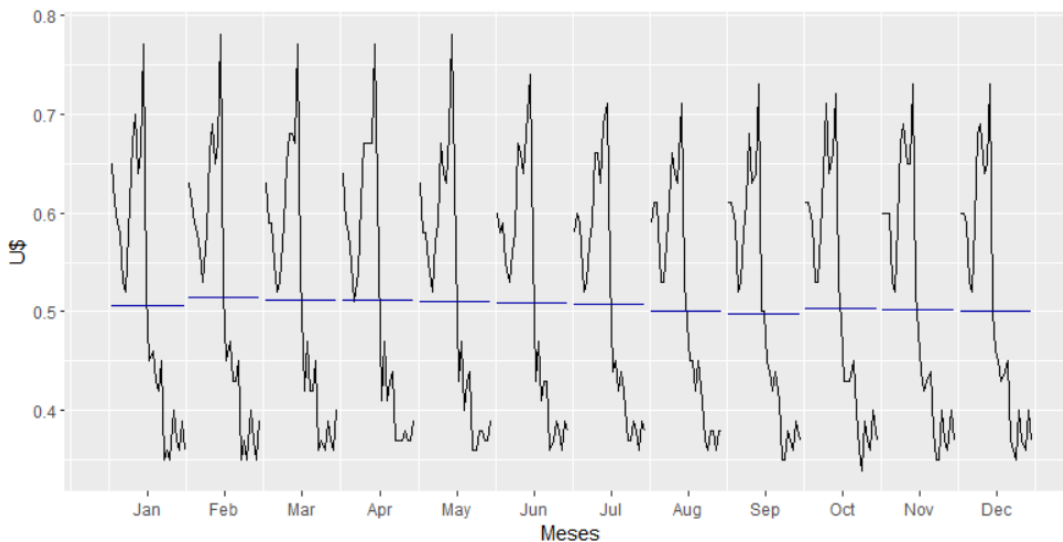
**Gráfico 10 - Subsérie sazonal Estados Unidos**



Fonte: Autoria própria (2022)

**Gráfico 11 - Subsérie sazonal Mundo**

Fonte: Autoria própria (2022)

**Gráfico 12 - Subsérie sazonal União Europeia**

Fonte: Autoria própria (2022)

As linhas horizontais indicam as médias para cada mês. Esta forma de gráfico permite que o padrão sazonal subjacente seja visto claramente e também mostra as mudanças na sazonalidade ao longo do tempo. É especialmente útil na identificação de mudanças em determinadas estações do ano, em alguns casos, essa é a maneira mais útil de visualizar as mudanças sazonais.

Os gráficos discutidos até agora são úteis para visualizar séries temporais individuais e explorar as relações entre elas.

## 4.2 Autocorrelação

A autocorrelação mede a relação linear entre os valores defasados de uma série temporal. Os coeficientes são plotados para mostrar a função de autocorrelação ou ACF.

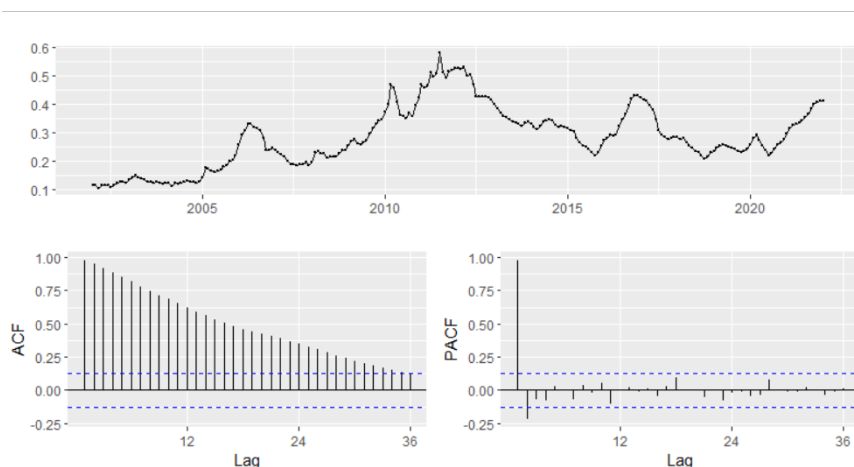
Quando os dados têm uma tendência, as autocorrelações para pequenas defasagens tendem a ser grandes e positivas porque as observações próximas no tempo também estão próximas em tamanho. Assim, a ACF destas séries tendem a ter valores positivos que diminuem lentamente à medida que as defasagens aumentam.

Quando os dados são sazonais, as autocorrelações serão maiores para os atrasos sazonais (em múltiplos da frequência sazonal) do que para os outros. Quando os dados apresentam tendência e sazonalidade, é perceptível uma combinação desses efeitos.

Já a definição do coeficiente de correlação parcial tem relação direta com modelos auto-regressivos. De fato, O coeficiente de autocorrelação parcial é o último coeficiente de um modelo AR ( $p$ ), ajustado para uma série temporal. Este também contribui na observação dos atrasos mais relevantes no processo de previsão.

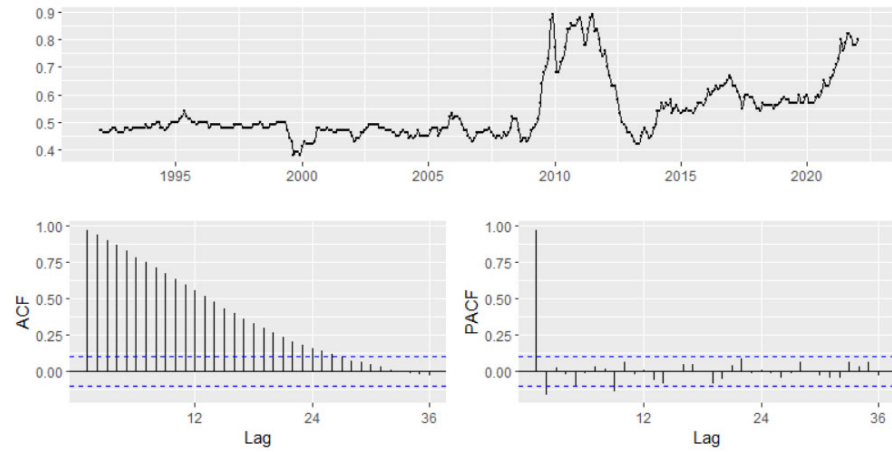
Os gráficos 13, 14, 15 e 16 referem-se a autocorrelação e autocorrelação parcial do preço do açúcar.

**Gráfico 13 - ACF e PACF Brasil**



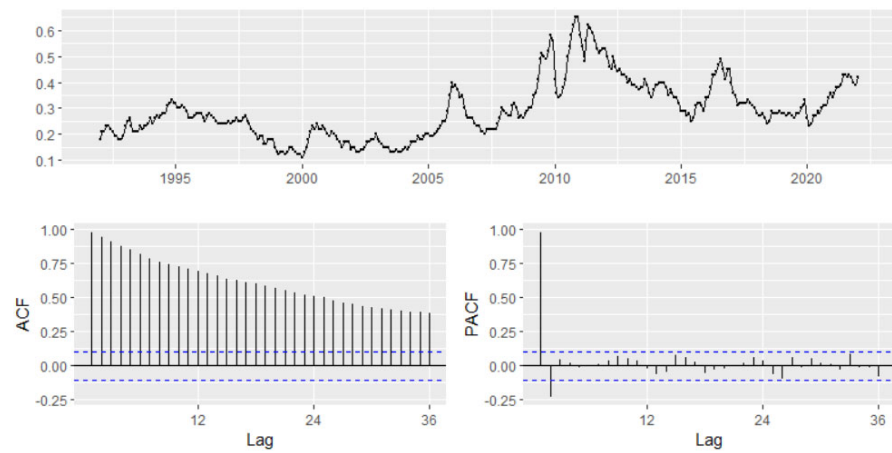
**Fonte: Autoria própria (2022)**

**Gráfico 14 - ACF e PACF Estados Unidos**



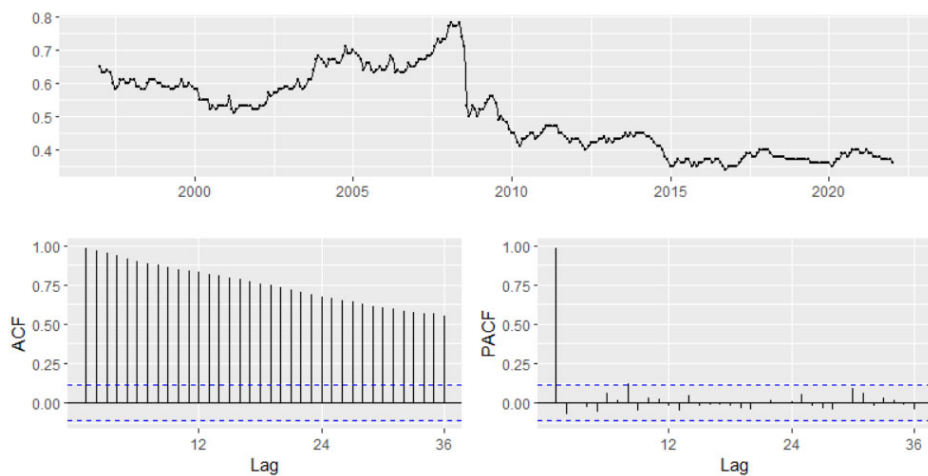
**Fonte: Autoria própria (2022)**

**Gráfico 15 - ACF e PACF Mundo**



**Fonte: Autoria própria (2022)**

**Gráfico 16 - ACF e PACF União Europeia**



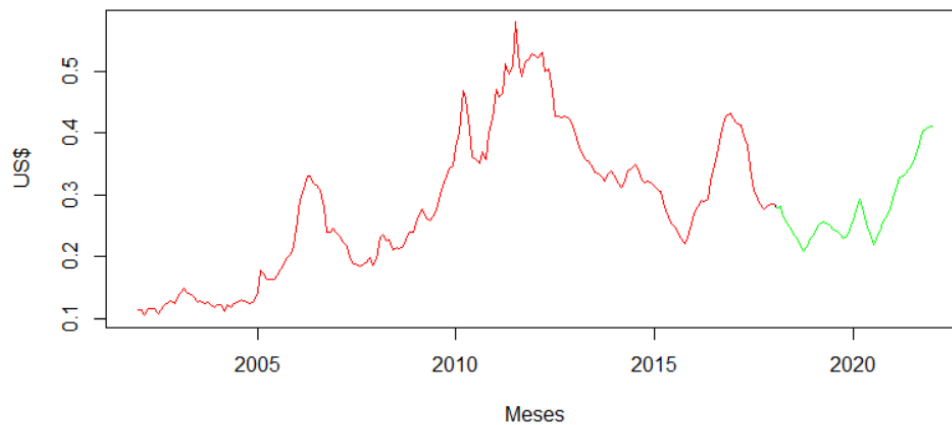
**Fonte: Autoria própria (2022)**

Essas figuras indicam que as séries possuem certa dependência, que pode ser notada pelo seu decaimento lento nos Gráficos ACF. Essas características também apontam para a não estacionariedade da série.

### 4.3 Ajuste do modelo

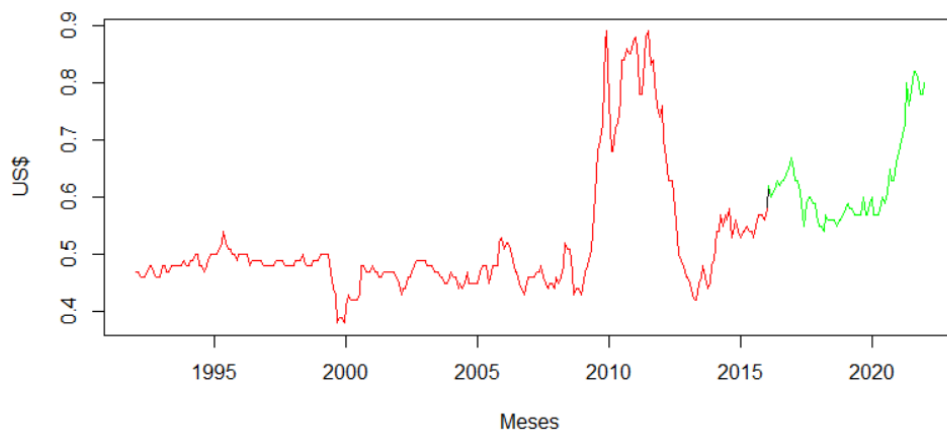
Idealmente, o conjunto de teste deve ser pelo menos tão grande quanto o horizonte de previsão máximo necessário. Assim, nos Gráficos 17, 18, 19 e 20 pode-se observar essa separação dos dados, sendo 80% dos dados treinamento e 20% teste.

**Gráfico 17 - Conjunto de treinamento (vermelho) e teste (verde) Brasil**

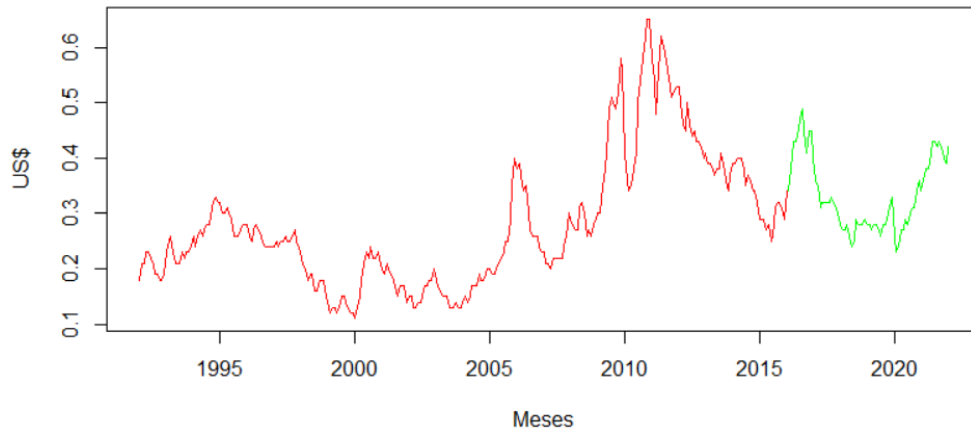


**Fonte: Autoria própria (2022)**

**Gráfico 18 - Conjunto de treinamento (vermelho) e teste (verde) Estados Unidos**



**Fonte: Autoria própria (2022)**

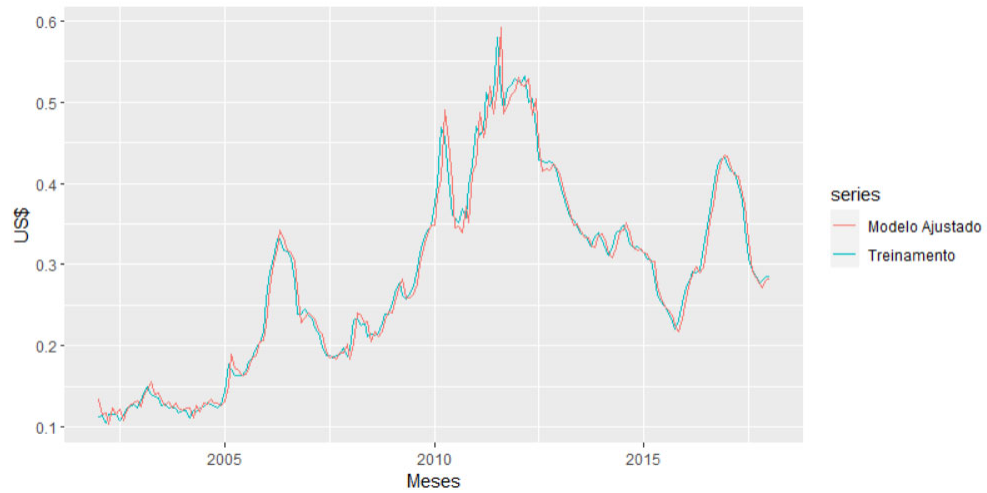
**Gráfico 19 - Conjunto de treinamento (vermelho) e teste (verde) Mundo**

**Fonte: Autoria própria (2022)**

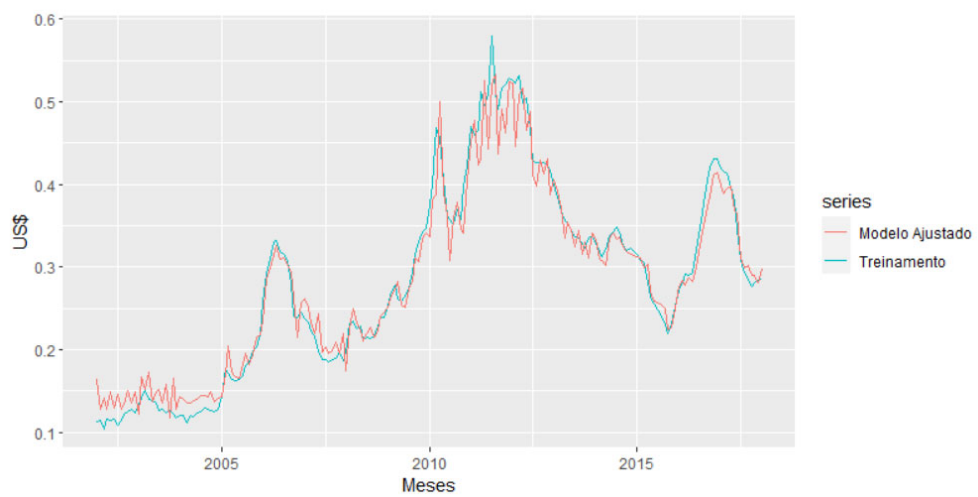
**Gráfico 20 - Conjunto de treinamento (vermelho) e teste (verde) União Europeia**

**Fonte: Autoria própria (2022)**

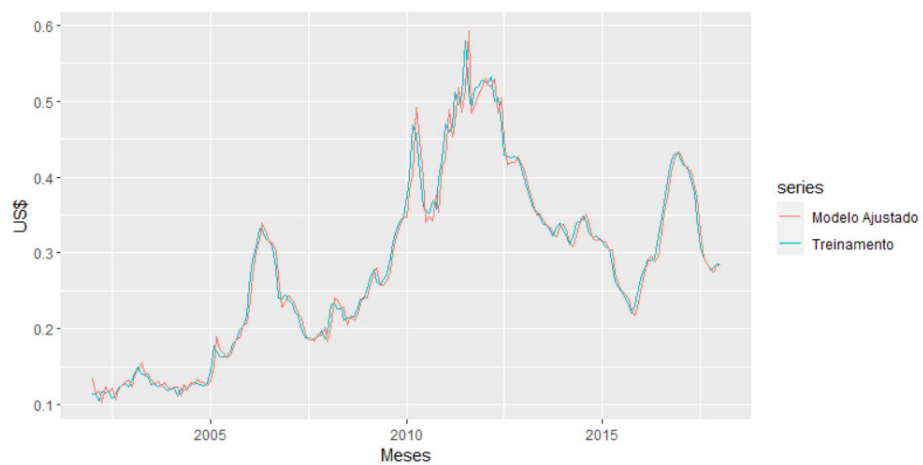
Em sequência, os parâmetros foram calculados conforme discutido na seção 3.2.3 e os melhores ajustes do modelo encontram-se nos Apêndice A. A seguir, serão ilustrados os modelos ajustados para as quatro séries temporais em análise. Nos Gráficos 21 ao 27 são apresentados os resultados do treinamento comparando-se a resposta dos modelos AR, ARMA, ARIMA, SES, HOLT e HOLT WINTER e os valores observados para a série Brasil.

**Gráfico 21 - Modelo AR ajustado nos dados de treinamento – Brasil**

Fonte: Autoria própria (2022)

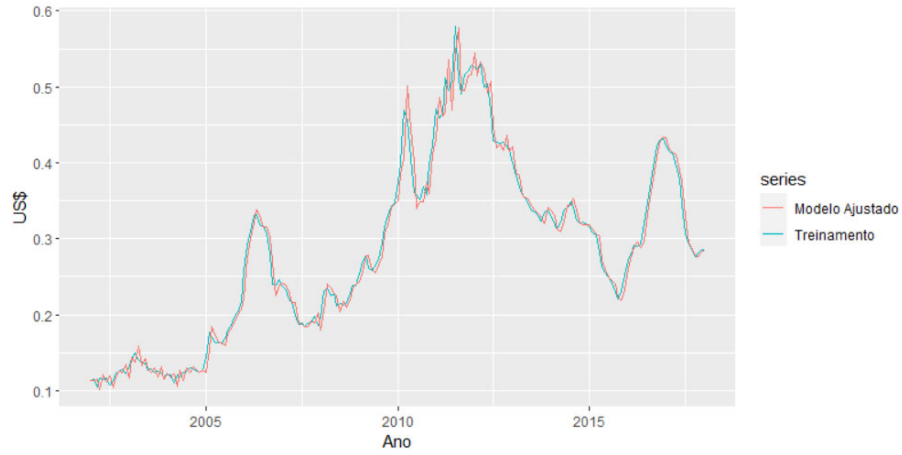
**Gráfico 22 - Modelo MA ajustado nos dados de treinamento – Brasil**

Fonte: Autoria própria (2022)

**Gráfico 23 - Modelo ARMA ajustado nos dados de treinamento – Brasil**

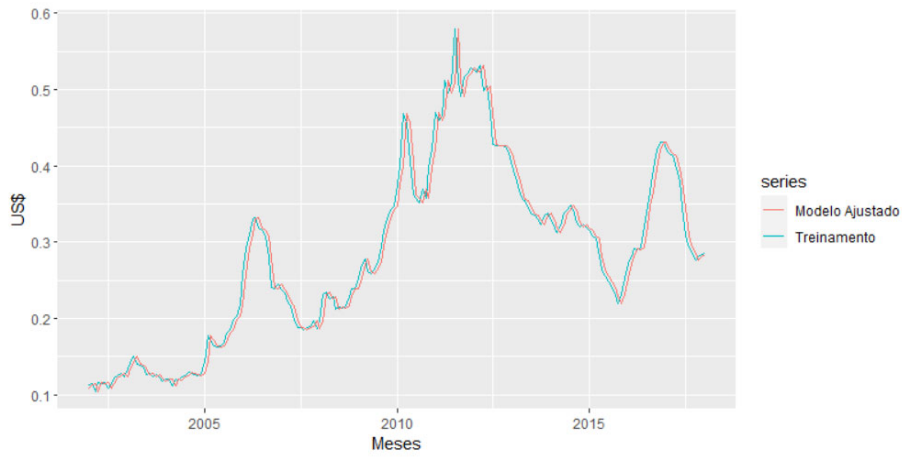
Fonte: Autoria própria (2022)

**Gráfico 24 - Modelo ARIMA ajustado nos dados de treinamento – Brasil**



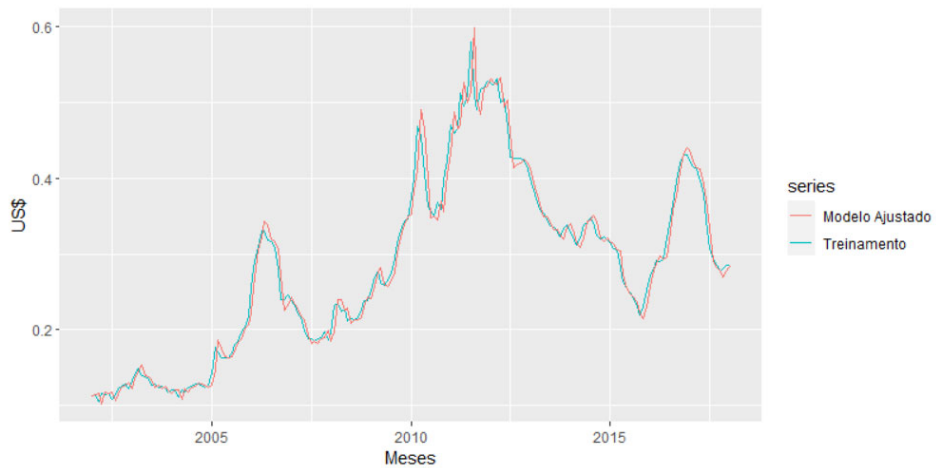
**Fonte: Autoria própria (2022)**

**Gráfico 25 - Modelo SES ajustado nos dados de treinamento - Brasil**



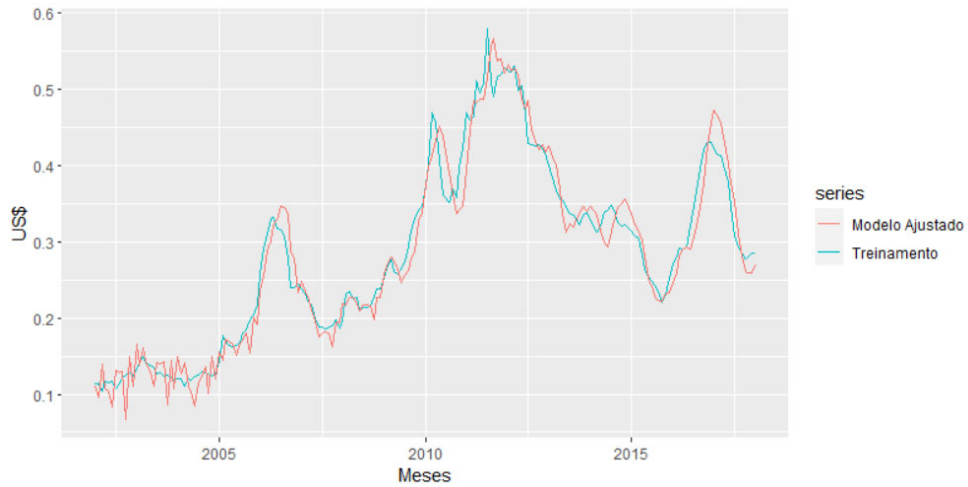
**Fonte: Autoria própria (2022)**

**Gráfico 26 - Modelo HOLT ajustado nos dados de treinamento – Brasil**



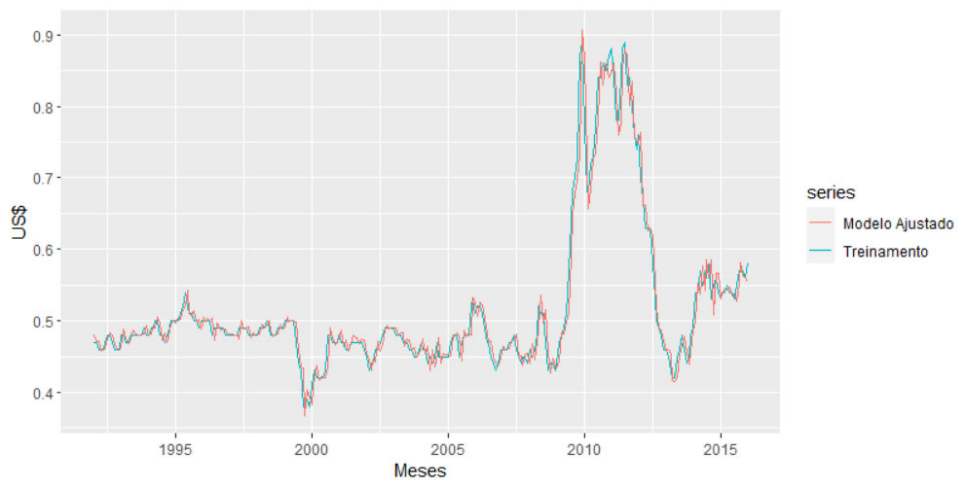
**Fonte: Autoria própria (2022)**



**Gráfico 27 - Modelo HOLT WINTERS ajustado nos dados de treinamento – Brasil**

**Fonte: Autoria própria (2022)**

Em seguida, foi realizado os ajustes dos modelos AR, ARMA, ARIMA, SES, HOLT e HOLT WINTER para a série Estados Unidos, conforme Gráficos 28, 29, 30, 31, 32, 33 e 34.

**Gráfico 28 - Modelo AR ajustado nos dados de treinamento – Estados Unidos**

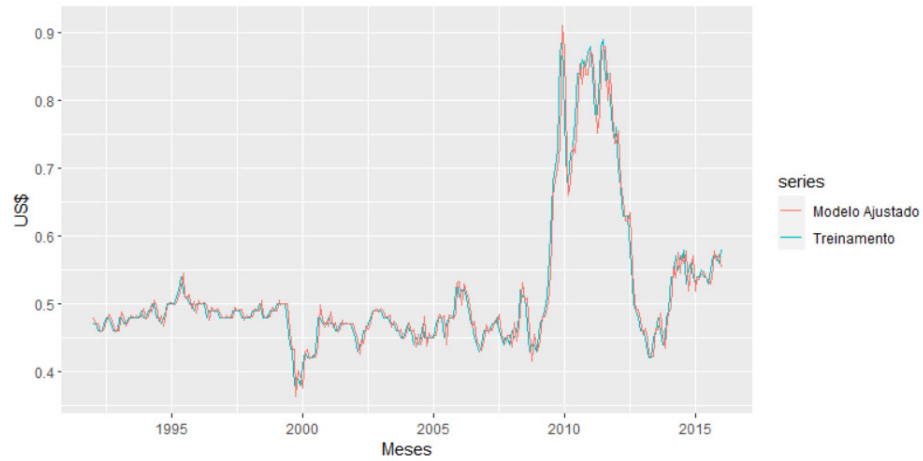
**Fonte: Autoria própria (2022)**

**Gráfico 29 - Modelo MA ajustado nos dados de treinamento – Estados Unidos**



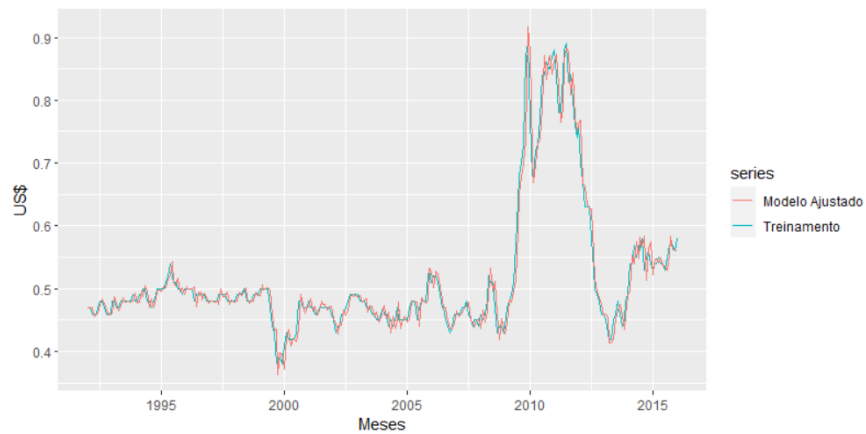
**Fonte: Autoria própria (2022)**

**Gráfico 30 - Modelo ARMA ajustado nos dados de treinamento – Estados Unidos**

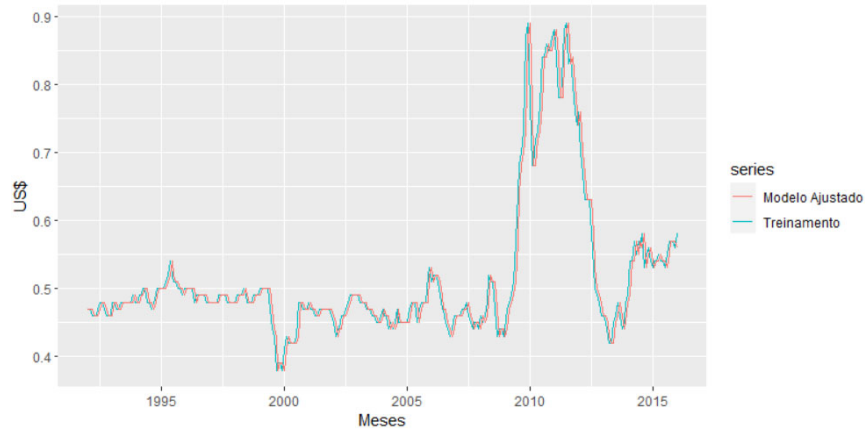


**Fonte: Autoria própria (2022)**

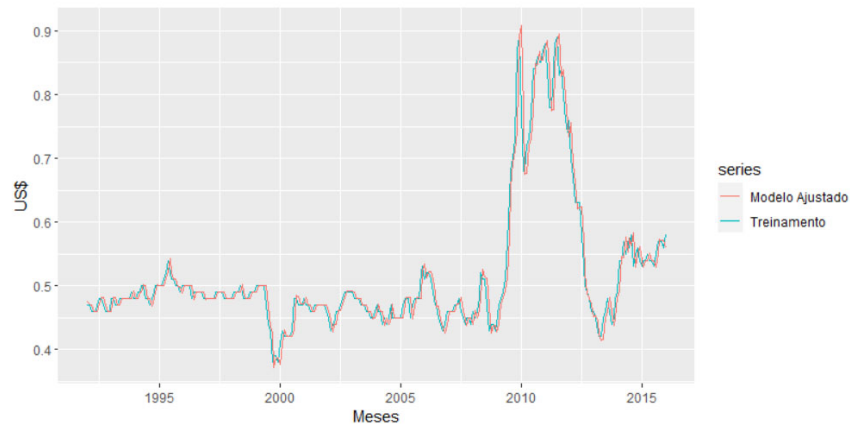
**Gráfico 31 - Modelo ARIMA ajustado nos dados de treinamento – Estados Unidos**



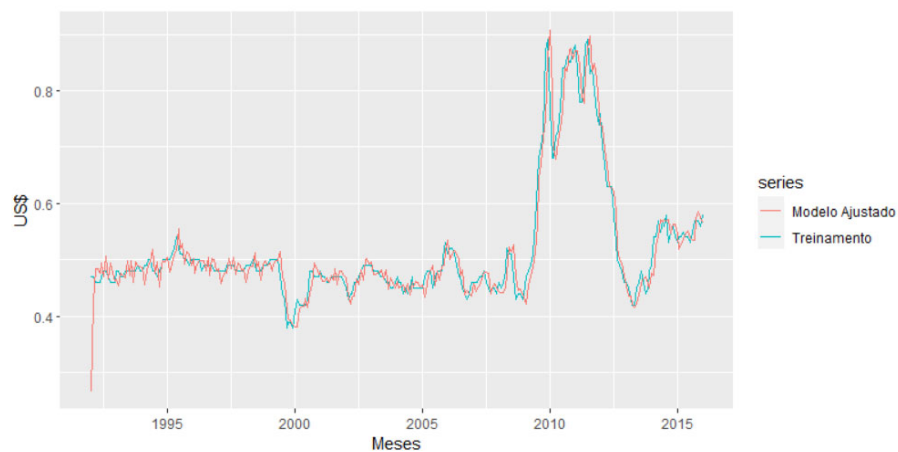
**Fonte: Autoria própria (2022)**

**Gráfico 32 - Modelo SES ajustado nos dados de treinamento – Estados Unidos**

Fonte: Autoria própria (2022)

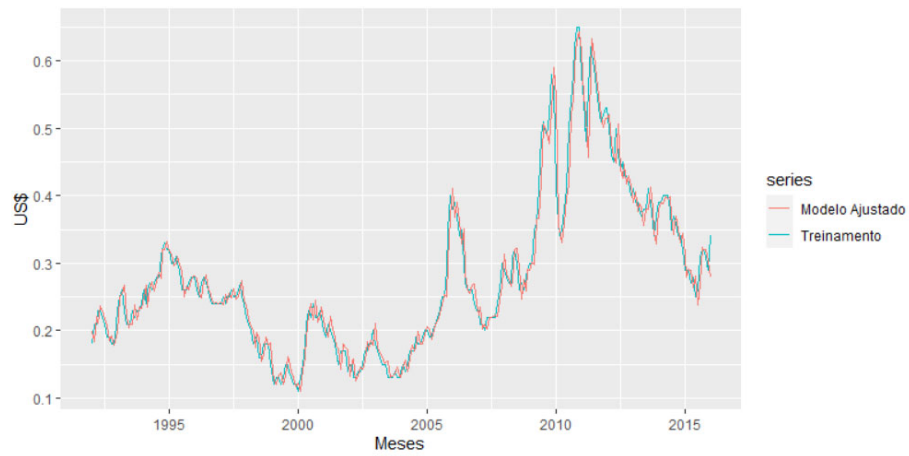
**Gráfico 33 - Modelo HOLT ajustado nos dados de treinamento – Estados Unidos**

Fonte: Autoria própria (2022)

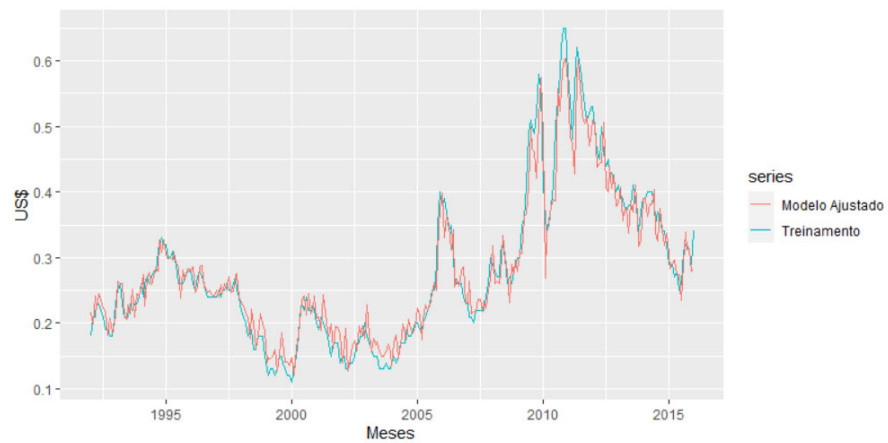
**Gráfico 34 - Modelo HOLT WINTERS ajustado nos dados de treinamento – Estados Unidos**

Fonte: Autoria própria (2022)

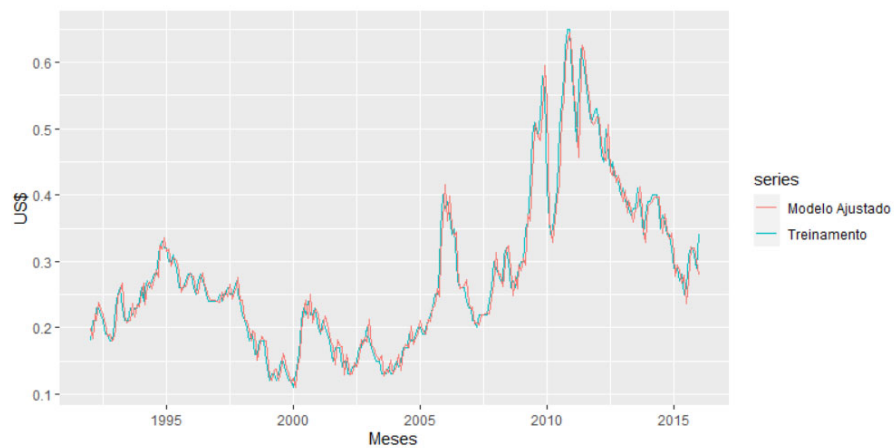
Calculados os parâmetros dos modelos AR, ARMA, ARIMA, SES, HOLT e HOLT WINTER para a série Mundo, pode-se observar a modelagem conforme Gráficos 35, 36, 37, 38, 39, 40 e 41.

**Gráfico 35 - Modelo AR ajustado nos dados de treinamento – Mundo**

Fonte: Autoria própria (2022)

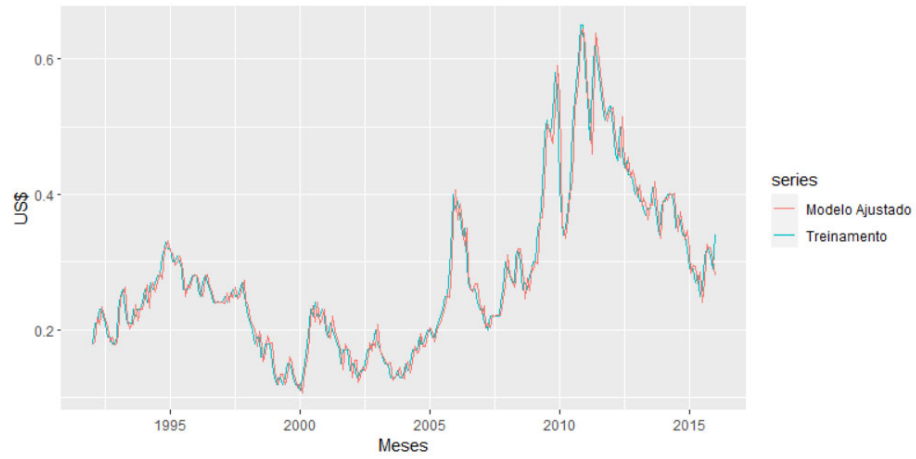
**Gráfico 36 - Modelo MA ajustado nos dados de treinamento – Mundo**

Fonte: Autoria própria (2022)

**Gráfico 37 - Modelo ARMA ajustado nos dados de treinamento – Mundo**

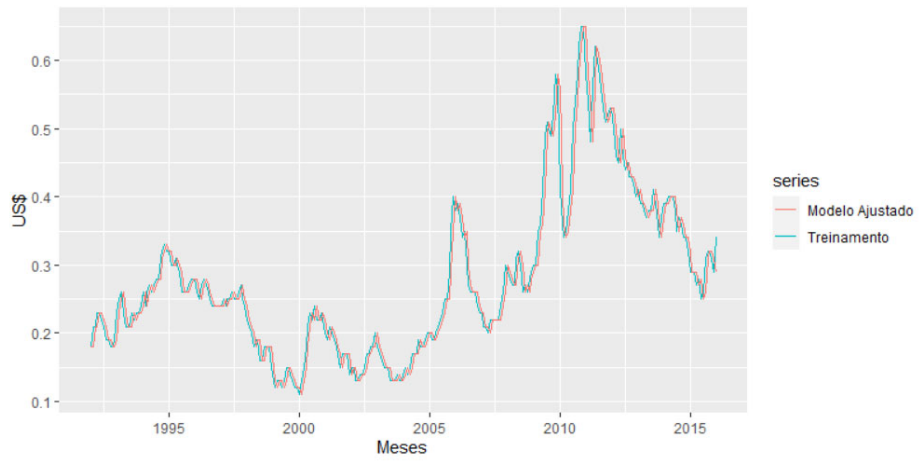
Fonte: Autoria própria (2022)

**Gráfico 38 - Modelo ARIMA ajustado nos dados de treinamento – Mundo**



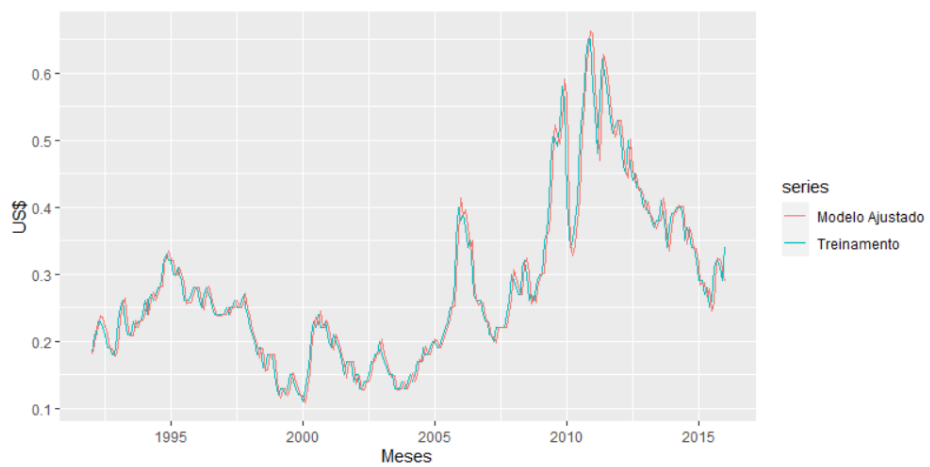
**Fonte: Autoria própria (2022)**

**Gráfico 39 - Modelo SES ajustado nos dados de treinamento – Mundo**

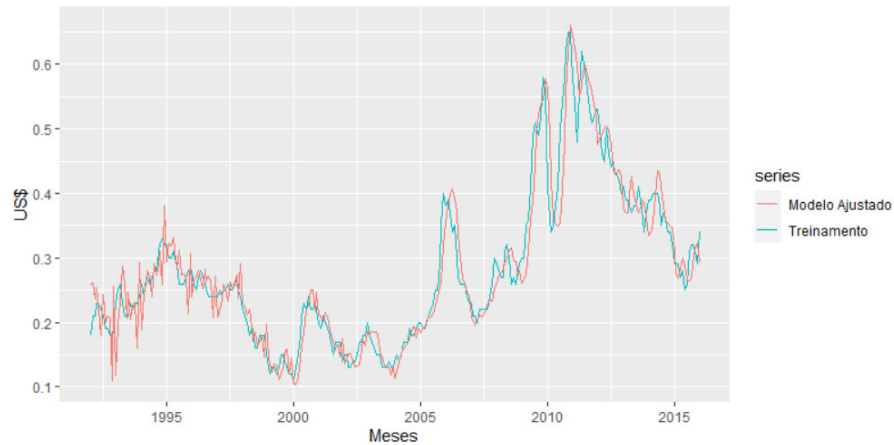


**Fonte: Autoria própria (2022)**

**Gráfico 40 - Modelo HOLT ajustado nos dados de treinamento – Mundo**



**Fonte: Autoria própria (2022)**

**Gráfico 41 - Modelo HOLT WINTERS ajustado nos dados de treinamento – Mundo**

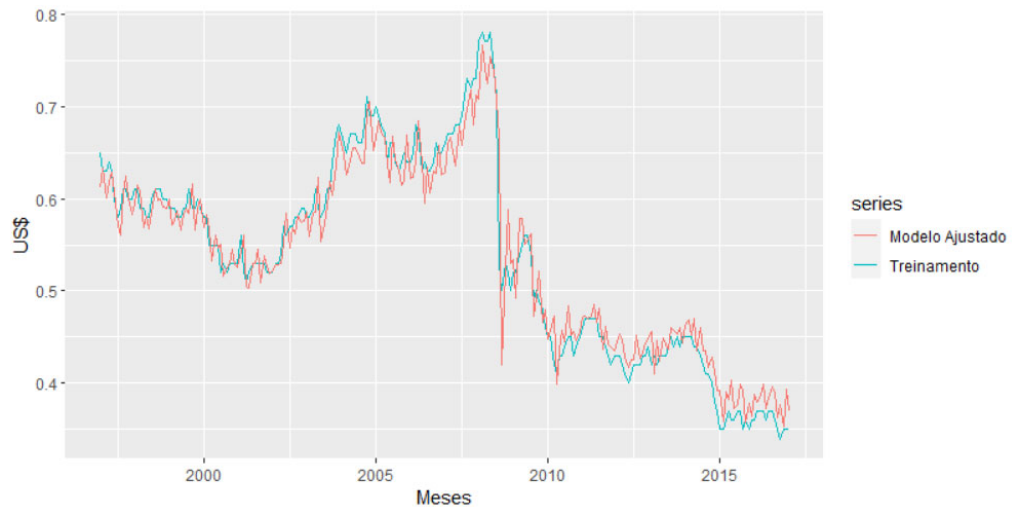
Fonte: Autoria própria (2022)

E por fim, foram calculados os parâmetros dos modelos AR, ARMA, ARIMA, SES, HOLT e HOLT WINTER para a série União Europeia, ilustrados nos Gráficos 42, 43, 44, 45, 46, 47 e 48.

**Gráfico 42 - Modelo AR ajustado nos dados de treinamento – União Europeia**

Fonte: Autoria própria (2022)

**Gráfico 43 - Modelo MA ajustado nos dados de treinamento – União Europeiaia**



**Fonte: Autoria própria (2022)**

**Gráfico 44 - Modelo ARMA ajustado nos dados de treinamento – União Europeiaia**



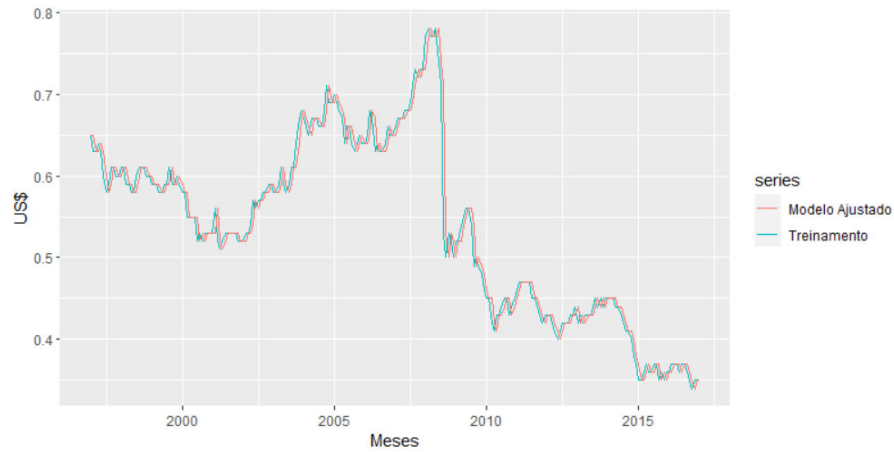
**Fonte: Autoria própria (2022)**

**Gráfico 45 - Modelo ARIMA ajustado nos dados de treinamento – União Europeiaia**



**Fonte: Autoria própria (2022)**

**Gráfico 46 - Modelo SES ajustado nos dados de treinamento – União Europeia**



**Fonte: Autoria própria (2022)**

**Gráfico 47 - Modelo HOLT ajustado nos dados de treinamento – União Europeia**



**Fonte: Autoria própria (2022)**

**Gráfico 48 - Modelo HOLT WINTERS ajustado nos dados de treinamento – União Europeia**



**Fonte: Autoria própria (2022)**



Um modelo que se ajusta bem aos dados de treinamento não necessariamente fará uma boa previsão. Um ajuste adequado sempre pode ser obtido usando um modelo com parâmetros suficientes. O treinamento excessivo de um modelo aos dados é tão ruim quanto não identificar um padrão sistemático.

Em geral, visualmente pode-se observar que os modelos se ajustaram bem ao conjunto de treinamento, captando os padrões das séries. Na análise do Brasil, o HOLT WINTERS foi o que apresentou pior ajuste. Isso pode ser explicado por não apresentar comportamento sazonal. Já nas séries Estados Unidos, Mundo e União Europeia, o ajuste que apresentou pior desempenho foi o MA.

#### **4.4 Previsão**

Definidas as técnicas de previsão e a aplicação dos dados passados para obtenção dos parâmetros necessários, podem-se obter as projeções futuras da série temporal. A seguir, é apresentado os resultados computacionais obtidos nesta investigação.

##### **4.4.1 Previsão 1 passo à frente**

Com os parâmetros definidos, as previsões 1 passo à frente ( $P = 1$ ) para as séries do preço do açúcar foram realizadas. Na Tabela 3 são apresentados os resultados das métricas RMSE, MAE e MAPE.

Tabela 3 - Performance dos modelos – 1 passo à frente

| MODELOS      | Métricas de desempenho | Horizonte de Previsão 1 passo à frente |                |               |                |
|--------------|------------------------|--|----------------|---------------|----------------|
|              |                        | Brasil                                 | Estados Unidos | Mundo         | União Europeia |
| AR           | RMSE                   | 0,0097                                 | 0,0217         | 0,0218        | 0,0078         |
|              | MAE                    | 0,0080                                 | 0,0167         | 0,0167        | 0,0057         |
|              | MAPE                   | 3,0057                                 | 2,6101         | 5,0052        | 1,5233         |
| MA           | RMSE                   | 0,0116                                 | 0,0324         | 0,0282        | 0,0116         |
|              | MAE                    | 0,0089                                 | 0,0243         | 0,0202        | 0,0090         |
|              | MAPE                   | 3,0611                                 | 3,7334         | 5,9314        | 3,1252         |
| ARMA         | RMSE                   | 0,0094                                 | 0,0219         | 0,0229        | 0,0076         |
|              | MAE                    | 0,0078                                 | 0,0165         | 0,0176        | 0,0057         |
|              | MAPE                   | 2,9065                                 | 2,5756         | 5,3102        | 1,5079         |
| ARIMA        | RMSE                   | <b>0,0091</b>                          | 0,0214         | 0,0219        | 0,0070         |
|              | MAE                    | <b>0,0073</b>                          | 0,0161         | 0,017         | 0,0050         |
|              | MAPE                   | <b>2,7376</b>                          | 2,5317         | 5,0886        | 1,3270         |
| SES          | RMSE                   | 0,0116                                 | 0,0195         | 0,0222        | 0,0067         |
|              | MAE                    | 0,0101                                 | <b>0,0142</b>  | 0,0164        | <b>0,0042</b>  |
|              | MAPE                   | 3,7453                                 | <b>2,2168</b>  | 4,9279        | <b>1,0993</b>  |
| HOLT         | RMSE                   | 0,0098                                 | 0,0195         | <b>0,0214</b> | <b>0,0065</b>  |
|              | MAE                    | 0,0082                                 | 0,0144         | <b>0,0161</b> | 0,0046         |
|              | MAPE                   | 3,0920                                 | 2,2558         | <b>4,8823</b> | 1,2112         |
| HOLT WINTERS | RMSE                   | 0,0140                                 | <b>0,0182</b>  | 0,023         | 0,0077         |
|              | MAE                    | 0,0110                                 | 0,0147         | 0,0184        | 0,0061         |
|              | MAPE                   | 3,9481                                 | 2,3333         | 5,6212        | 1,6220         |
| ENSEMBLE     | RMSE                   | 0,0094                                 | 0,0214         | 0,0215        | 0,0067         |
|              | MAE                    | 0,0079                                 | 0,0161         | 0,0164        | 0,0047         |
|              | MAPE                   | 2,9221                                 | 2,5231         | 4,9537        | 1,2528         |

Fonte: Autoria própria (2022)

Em seguida um ranking de desempenho foi elaborado na Tabela 4, indicando o modelo ganhador para cada caso.

Tabela 4 - Ranking de desempenho das previsões 1 passo à frente

| MODELOS      | Horizonte de Previsão 1 passo à frente |     |      |                |     |      |       |     |      |                |     |      | Ranking |
|--------------|--|-----|------|----------------|-----|------|-------|-----|------|----------------|-----|------|---------|
|              | Brasil                                 |     |      | Estados Unidos |     |      | Mundo |     |      | União Europeia |     |      |         |
|              | RMSE                                   | MAE | MAPE | RMSE           | MAE | MAPE | RMSE  | MAE | MAPE | RMSE           | MAE | MAPE |         |
| AR           |  |     |      |                |     |      |       |     |      |                |     |      | 0       |
| MA           |  |     |      |                |     |      |       |     |      |                |     |      | 0       |
| ARMA         |  |     |      |                |     |      |       |     |      |                |     |      | 0       |
| ARIMA        | 1                                      | 1   | 1    |                |     |      |       |     |      |                |     |      | 3       |
| SES          |  |     |      |                | 1   | 1    |       |     |      |                | 1   | 1    | 4       |
| HOLT         |  |     |      |                |     |      | 1     | 1   | 1    | 1              |     |      | 4       |
| HOLT WINTERS |  |     |      | 1              |     |      |       |     |      |                |     |      | 1       |
| ENSEMBLE     |  |     |      |                |     |      |       |     |      |                |     |      | 0       |

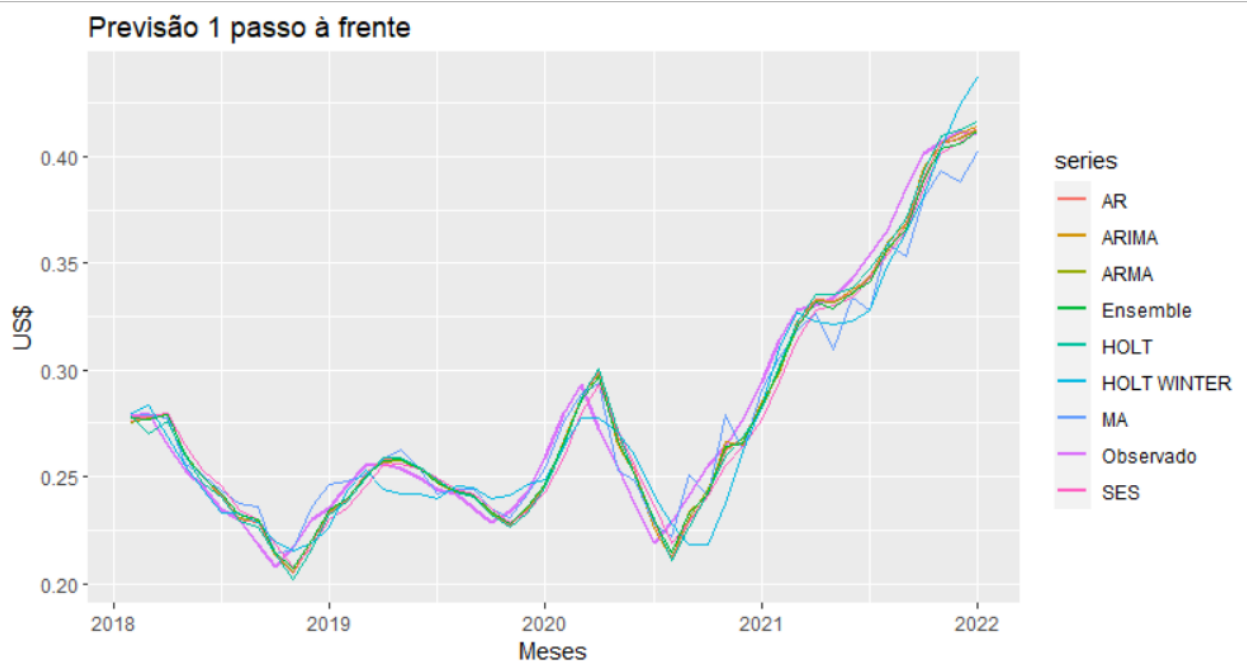
Fonte: Autoria própria (2022)

Os valores em negrito representam os melhores erros alcançados. Assim, o modelo ARIMA desempenhou melhor resultado para a série Brasil para todas as medidas de erro analisadas, seguido do HOLT WINTERS para as séries Estados Unidos e União Europeia, e HOLT para a série Mundo. Ainda, o modelo MA apresentou maior erro para as séries Estados Unidos, Mundo e União Europeia, enquanto na série Brasil o pior foi o HOLT WINTERS.

Uma vez que o MAE corresponde ao erro absoluto médio e o MAPE é uma média percentual simples, ambos não fazem distinção de erros pequenos e erros grandes, como o caso do RMSE que os eleva os erros ao quadrado. Por essa razão é aconselhável o uso deste último (ARMSTRONG; COLLOPY, 1992). Assim, considerando o erro RMSE do primeiro colocado em termos percentual para as séries Brasil, Estados Unidos, Mundo e União Europeia foi de 35%, 39,91%, 24,11% e 63,79%, respectivamente.

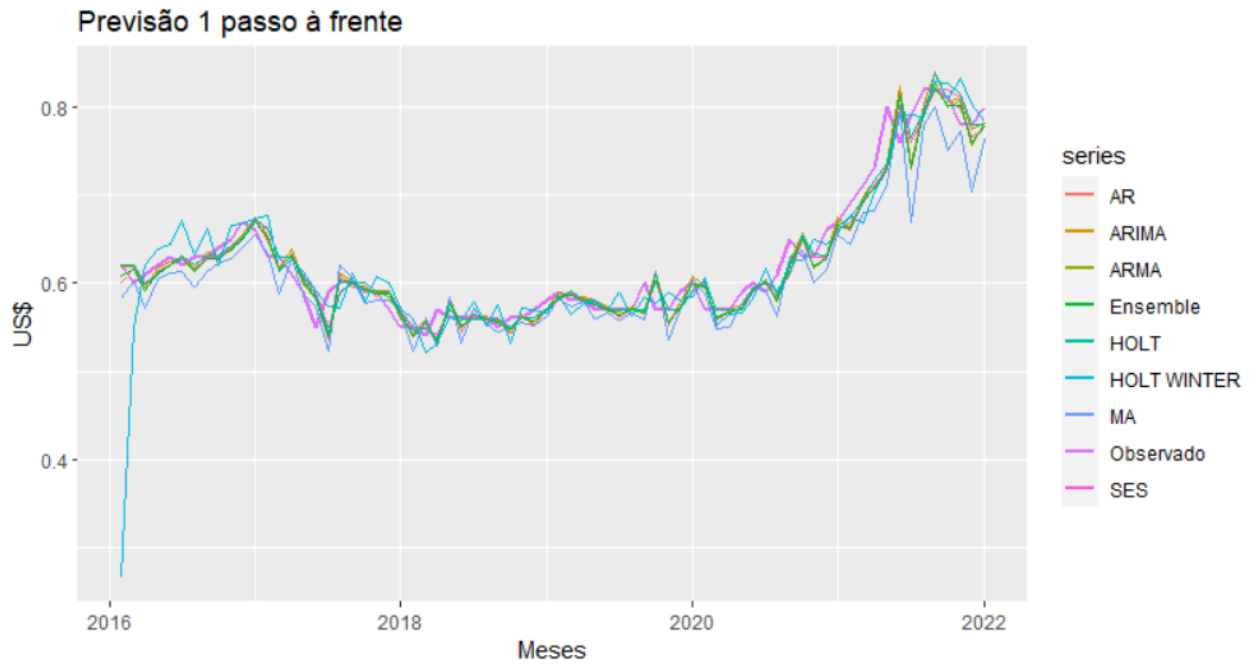
De acordo com a Tabela 4, houve um empate dos métodos que desempenharam melhores resultados para previsão de um passo à frente, sendo eles SES e HOLT. Os Gráficos 49, 50, 51 e 52 apresentam as previsões realizadas pelos modelos e os dados reais separados para teste.

**Gráfico 49 - Previsão 1 passo à frente – Brasil**



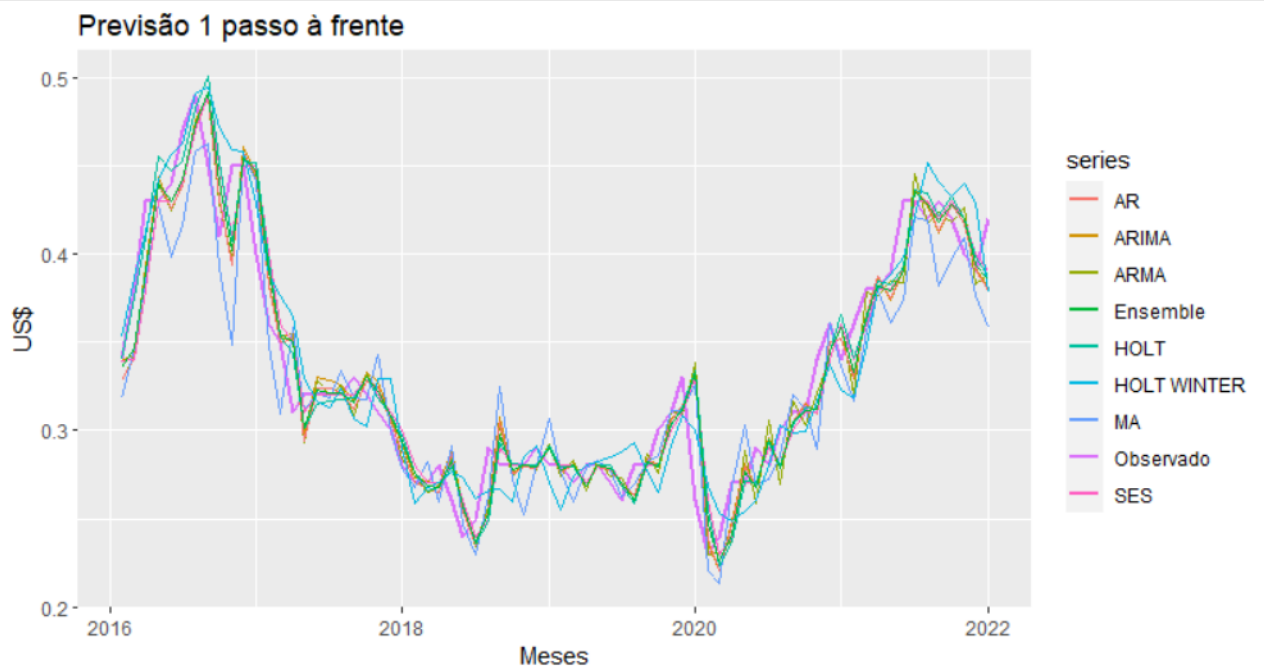
Fonte: Autoria própria (2022)

Gráfico 50 - Previsão 1 passo à frente – Estados Unidos



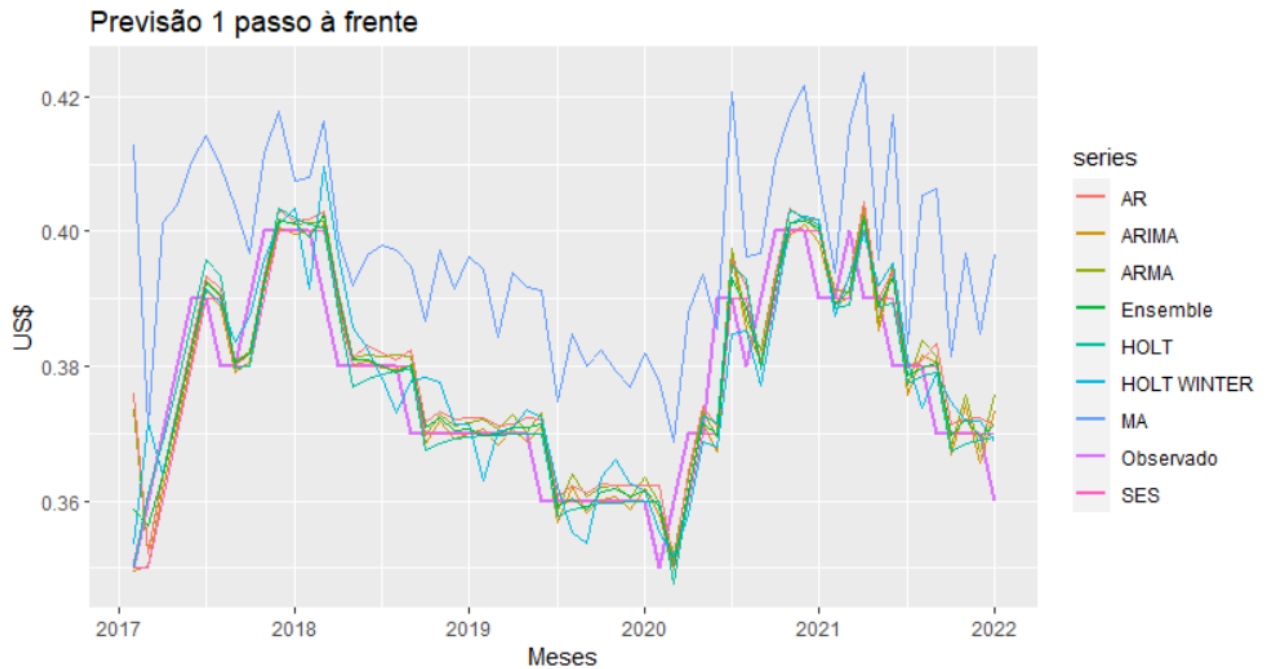
Fonte: A autoria própria (2022)

Gráfico 51 - Previsão 1 passo à frente – Mundo



Fonte: A autoria própria (2022)

Gráfico 52 - Previsão 1 passo à frente – União Europeia



Fonte: Autoria própria (2022)

Ao analisar as Figuras e as Tabelas acima observa-se que os valores previstos um passo à frente e os dados reais estão próximos, apresentando uma estimativa aceitável da amostra real. No entanto, o modelo MA apresentou resultados ruins, considerando-se inadequado para descrever o conjunto de dados em consideração.

#### 4.4.2 Previsão 3 passos à frente

Dando sequência, fez-se também previsões para as mesmas séries considerando o horizonte de 3 passos à frente, tendo seus resultados comparados pelas métricas RMSE, MAE e MAPE, conforme Tabela 5 e ranking da Tabela 6.

Tabela 5 - Performance dos modelos – 3 passos à frente

| MODELOS      | Métricas de desempenho | Horizonte de Previsão 3 passos à frente |                |               |                |
|--------------|------------------------|---|----------------|---------------|----------------|
|              |                        | Brasil                                  | Estados Unidos | Mundo         | União Europeia |
| AR           | RMSE                   | 0,0288                                  | 0,0385         | 0,0392        | 0,0143         |
|              | MAE                    | 0,0245                                  | 0,0301         | 0,0314        | 0,0125         |
|              | MAPE                   | 9,0476                                  | 4,6673         | 9,5963        | 3,3067         |
| MA           | RMSE                   | 0,0419                                  | 0,0884         | 0,0662        | 0,0412         |
|              | MAE                    | 0,0318                                  | 0,0688         | 0,0501        | 0,0327         |
|              | MAPE                   | 10,6105                                 | 10,2472        | 13,8679       | 11,1739        |
| ARMA         | RMSE                   | 0,0284                                  | 0,0377         | 0,0399        | 0,0131         |
|              | MAE                    | 0,0244                                  | 0,0286         | 0,032         | 0,0113         |
|              | MAPE                   | 8,9010                                  | 4,401          | 9,7764        | 2,9927         |
| ARIMA        | RMSE                   | 0,0280                                  | 0,0352         | 0,0395        | 0,0121         |
|              | MAE                    | 0,0236                                  | 0,028          | 0,0317        | 0,0091         |
|              | MAPE                   | 8,7035                                  | 4,3978         | 9,8551        | 2,3814         |
| SES          | RMSE                   | 0,0306                                  | 0,033          | 0,0393        | 0,0123         |
|              | MAE                    | 0,0272                                  | 0,0254         | 0,0301        | <b>0,0088</b>  |
|              | MAPE                   | 9,8753                                  | 3,9384         | 9,2734        | <b>2,2977</b>  |
| HOLT         | RMSE                   | 0,0286                                  | 0,0323         | 0,0384        | <b>0,0117</b>  |
|              | MAE                    | 0,0235                                  | 0,0254         | 0,0289        | 0,0091         |
|              | MAPE                   | 8,7418                                  | 3,9891         | 9,1124        | 2,3866         |
| HOLT WINTERS | RMSE                   | <b>0,0230</b>                           | <b>0,0313</b>  | <b>0,0342</b> | 0,0122         |
|              | MAE                    | <b>0,0185</b>                           | <b>0,0249</b>  | <b>0,028</b>  | 0,0100         |
|              | MAPE                   | <b>6,6050</b>                           | <b>3,9083</b>  | <b>8,5407</b> | 2,6375         |
| ENSEMBLE     | RMSE                   | 0,027                                   | 0,0337         | 0,0355        | 0,0117         |
|              | MAE                    | 0,0232                                  | 0,0263         | 0,0282        | 0,0094         |
|              | MAPE                   | 8,4855                                  | 4,0956         | 8,8289        | 2,4716         |

Fonte: Autoria própria (2022)

Tabela 6 - Ranking de desempenho das previsões 3 passos à frente

| MODELOS      | Horizonte de Previsão 3 passos à frente |     |      |                |     |      |       |     |      |                |     |      | Ranking |
|--------------|---|-----|------|----------------|-----|------|-------|-----|------|----------------|-----|------|---------|
|              | Brasil                                  |     |      | Estados Unidos |     |      | Mundo |     |      | União Europeia |     |      |         |
|              | RMSE                                    | MAE | MAPE | RMSE           | MAE | MAPE | RMSE  | MAE | MAPE | RMSE           | MAE | MAPE |         |
| AR           |   |     |      |                |     |      |       |     |      |                |     |      | 0       |
| MA           |   |     |      |                |     |      |       |     |      |                |     |      | 0       |
| ARMA         |   |     |      |                |     |      |       |     |      |                |     |      | 0       |
| ARIMA        |   |     |      |                |     |      |       |     |      |                |     |      | 0       |
| SES          |   |     |      |                |     |      |       |     |      |                | 1   | 1    | 2       |
| HOLT         |   |     |      |                |     |      |       |     |      | 1              |     |      | 1       |
| HOLT WINTERS | 1                                       | 1   | 1    | 1              | 1   | 1    | 1     | 1   | 1    |                |     |      | 9       |
| ENSEMBLE     |   |     |      |                |     |      |       |     |      | 1              |     |      | 1       |

Fonte: Autoria própria (2022)

Pela análise da Tabela 4, o HOLT WINTERS prevê com mais precisão nas três primeiras séries, para todas as medidas de erro analisadas. Já a série da União

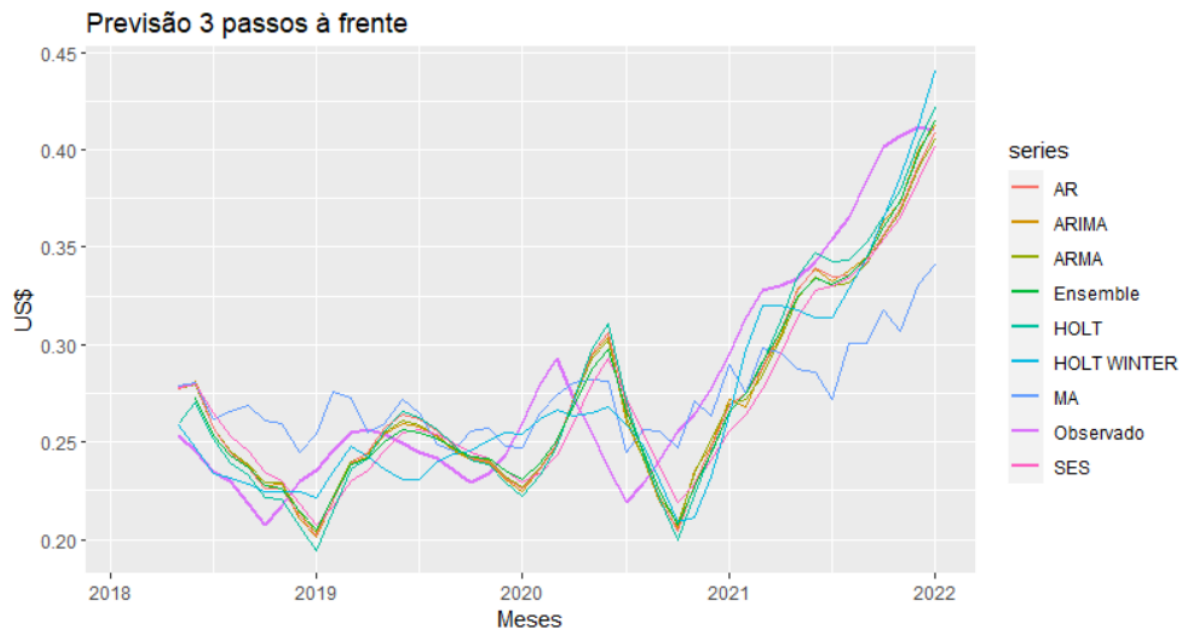
Europeia apresentou menor RMSE com o modelo HOLT, porém menores valores de MAPE e MAE calculados por SES.

Já os piores resultados foram calculados pelo modelo HOLT WINTERS para a série Brasil. Analisando o desempenho do modelo MA, este alcançou os piores resultados para as três séries restantes avaliadas.

Em geral, o ranking que aponta os melhores resultados de previsão para 3 passos à frente é obtido por meio do modelo HOLT WINTERS.

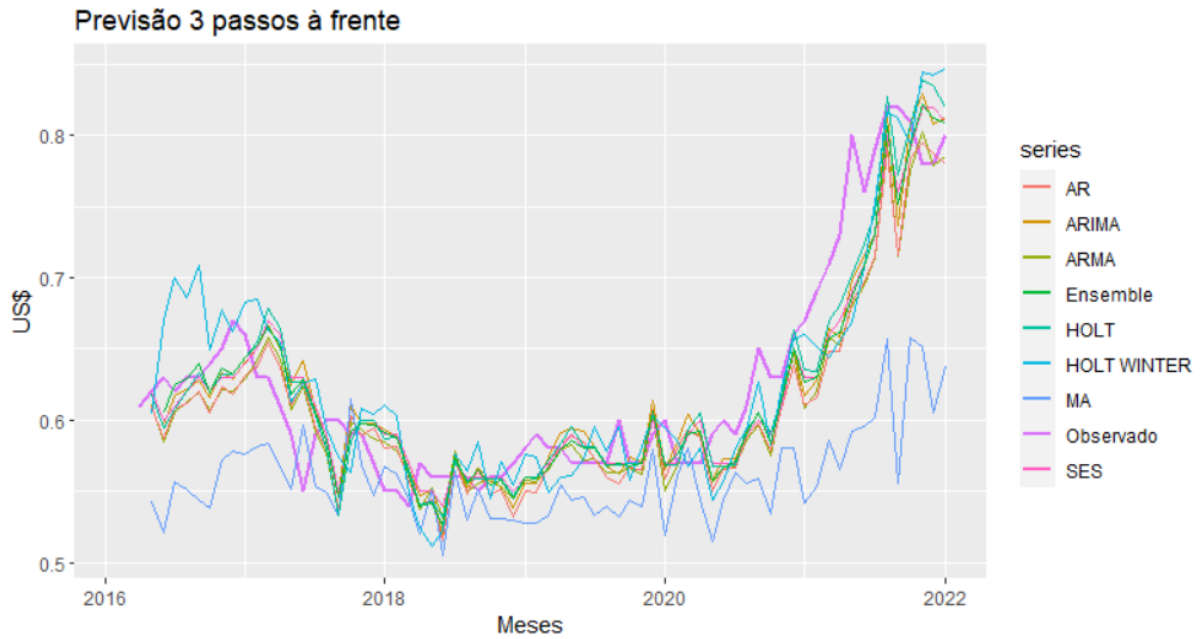
Os Gráficos 53, 54, 55 e 56 apresentam as previsões realizadas pelos modelos e os dados reais observados no teste.

**Gráfico 53 - Previsão 3 passos à frente – Brasil**



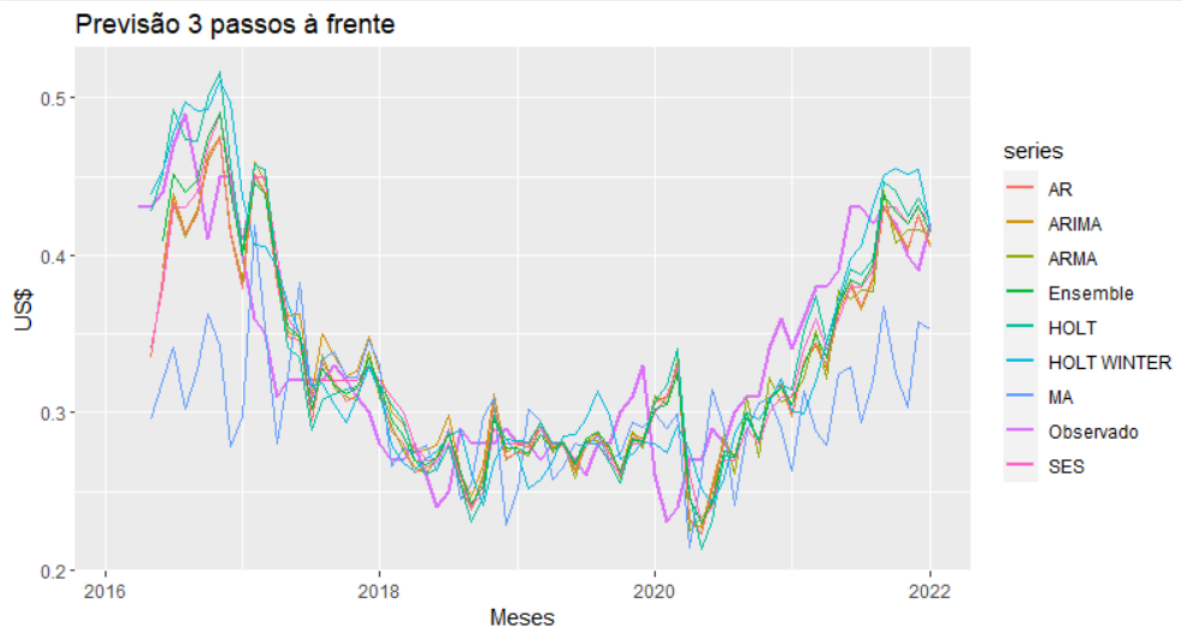
Fonte: Autoria própria (2022)

Gráfico 54 - Previsão 3 passos à frente – Estados Unidos



Fonte: Autoria própria (2022)

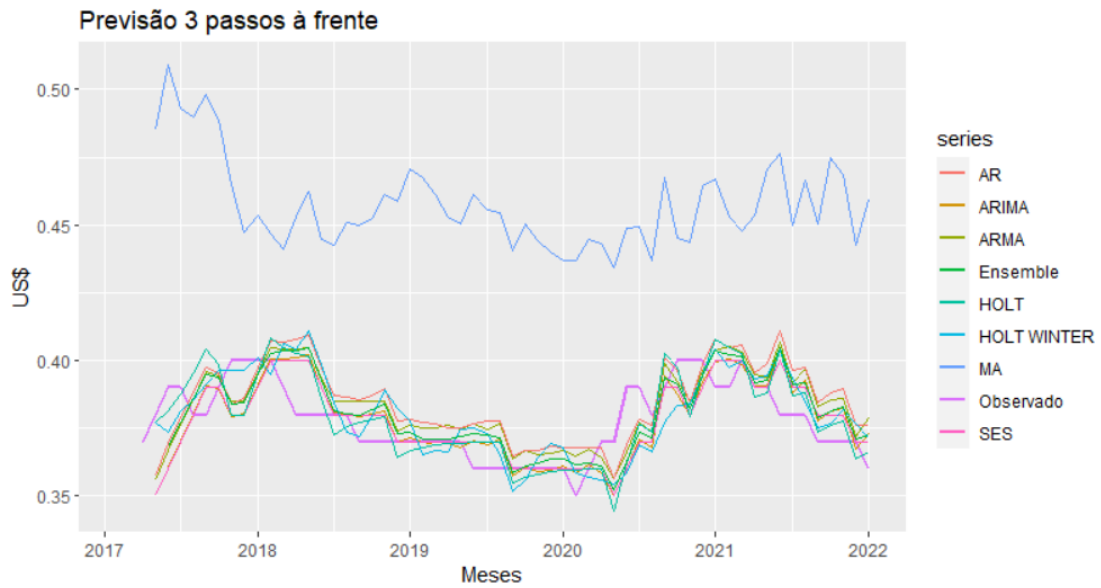
Gráfico 55 - Previsão 3 passos à frente – Mundo



Fonte - Autoria própria (2022)



**Gráfico 56 - Previsão 3 passos à frente – União Europeiaia**



**Fonte: Autoria própria (2022)**

#### 4.4.3 Previsão 6 passos à frente

Dando continuidade as previsões, foi estimada para as mesmas séries a previsão considerando o horizonte de 6 passos à frente, tendo seus resultados comparados pelas métricas RMSE, MAE e MAPE, conforme Tabela 6 e ranking da Tabela 7.

Tabela 7 - Performance dos modelos – 6 passos à frente

| Modelos      | Métricas de desempenho | Horizonte de Previsão 6 passos à frente |                |                |                |
|--------------|------------------------|---|----------------|----------------|----------------|
|              |                        | Brasil                                  | Estados Unidos | Mundo          | União Europeia |
| AR           | RMSE                   | 0,0482                                  | 0,0619         | 0,0528         | 0,0225         |
|              | MAE                    | 0,0419                                  | 0,0455         | 0,0422         | 0,0197         |
|              | MAPE                   | 14,7345                                 | 6,7992         | 12,4497        | 5,2604         |
| MA           | RMSE                   | 0,0605                                  | 0,1149         | 0,0753         | 0,0590         |
|              | MAE                    | 0,0446                                  | 0,0875         | 0,0524         | 0,0458         |
|              | MAPE                   | 14,5544                                 | 12,8775        | 14,0549        | 15,3656        |
| ARMA         | RMSE                   | 0,0480                                  | 0,0623         | 0,0531         | 0,0192         |
|              | MAE                    | 0,0414                                  | 0,0455         | 0,0424         | 0,0168         |
|              | MAPE                   | 14,3334                                 | 6,7903         | 12,4918        | 4,4841         |
| ARIMA        | RMSE                   | 0,0475                                  | 0,0534         | 0,0546         | 0,0169         |
|              | MAE                    | 0,0409                                  | 0,0399         | 0,0451         | 0,0143         |
|              | MAPE                   | 14,2922                                 | 6,0875         | 13,7818        | 3,7642         |
| SES          | RMSE                   | 0,0482                                  | 0,0524         | 0,0566         | 0,0167         |
|              | MAE                    | 0,0417                                  | 0,0374         | 0,045          | 0,0141         |
|              | MAPE                   | 14,4319                                 | 5,6773         | 13,5976        | 3,6991         |
| HOLT         | RMSE                   | 0,0481                                  | 0,0504         | 0,0585         | <b>0,0163</b>  |
|              | MAE                    | 0,0412                                  | 0,0377         | 0,0465         | <b>0,0134</b>  |
|              | MAPE                   | 14,7348                                 | 5,8123         | 14,2291        | <b>3,5202</b>  |
| HOLT WINTERS | RMSE                   | <b>0,0352</b>                           | <b>0,0484</b>  | 0,0517         | 0,0171         |
|              | MAE                    | <b>0,0265</b>                           | <b>0,0359</b>  | <b>0,0386</b>  | 0,0140         |
|              | MAPE                   | <b>8,7030</b>                           | <b>5,4848</b>  | <b>11,2546</b> | 3,6832         |
| ENSEMBLE     | RMSE                   | 0,0449                                  | 0,0534         | <b>0,0497</b>  | 0,0169         |
|              | MAE                    | 0,0385                                  | 0,0381         | 0,0396         | 0,0144         |
|              | MAPE                   | 13,2845                                 | 5,7630         | 12,0773        | 3,8160         |

Fonte: Autoria própria (2022)

Tabela 8 - Ranking de desempenho das previsões 6 passos à frente

| MODELOS      | Horizonte de Previsão 3 passos à frente |     |      |                |     |      |       |     |      |                |     |      | Ranking |
|--------------|---|-----|------|----------------|-----|------|-------|-----|------|----------------|-----|------|---------|
|              | Brasil                                  |     |      | Estados Unidos |     |      | Mundo |     |      | União Europeia |     |      |         |
|              | RMSE                                    | MAE | MAPE | RMSE           | MAE | MAPE | RMSE  | MAE | MAPE | RMSE           | MAE | MAPE |         |
| AR           |   |     |      |                |     |      |       |     |      |                |     |      | 0       |
| MA           |   |     |      |                |     |      |       |     |      |                |     |      | 0       |
| ARMA         |   |     |      |                |     |      |       |     |      |                |     |      | 0       |
| ARIMA        |   |     |      |                |     |      |       |     |      |                |     |      | 0       |
| SES          |   |     |      |                |     |      |       |     |      |                |     |      | 0       |
| HOLT         |   |     |      |                |     |      |       |     |      | 1              | 1   | 1    | 3       |
| HOLT WINTERS | 1                                       | 1   | 1    | 1              | 1   | 1    |       | 1   | 1    |                |     |      | 8       |
| ENSEMBLE     |   |     |      |                |     |      | 1     |     |      |                |     |      | 1       |

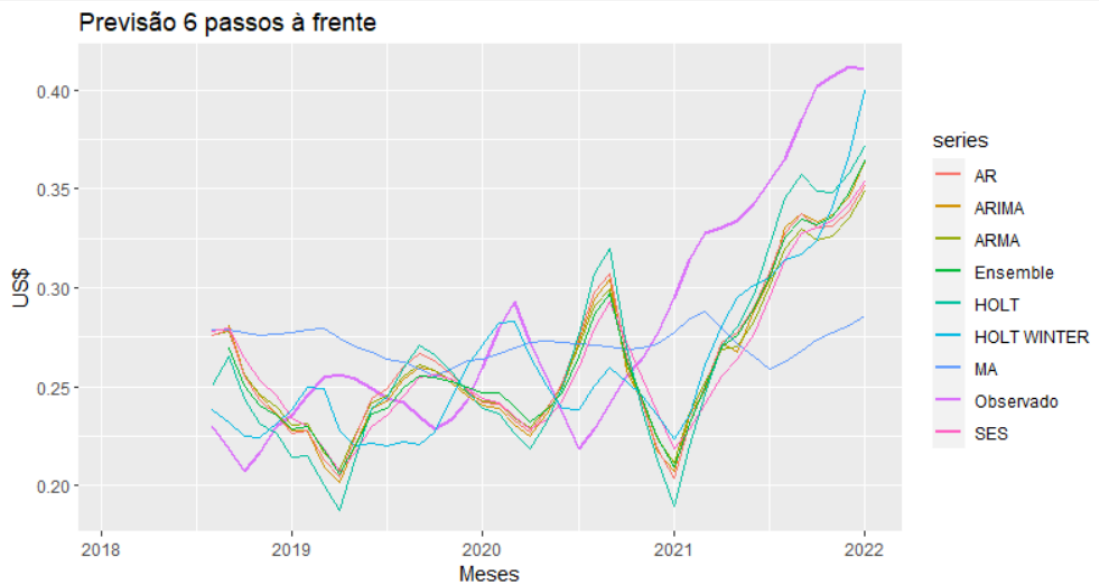
Fonte: Autoria própria (2022)

Os resultados da Tabela 7 apresentam que a melhor previsão realizada foi obtida pela modelo HOLT WINTERS, seguido do HOLT, para todas as medidas de erro analisadas. Isto fica evidente na Tabela 7. Ainda, avaliando os piores resultados,

observou-se que o MA indicou um ajuste ruim para todas as séries em análise, ou seja, os maiores erros.

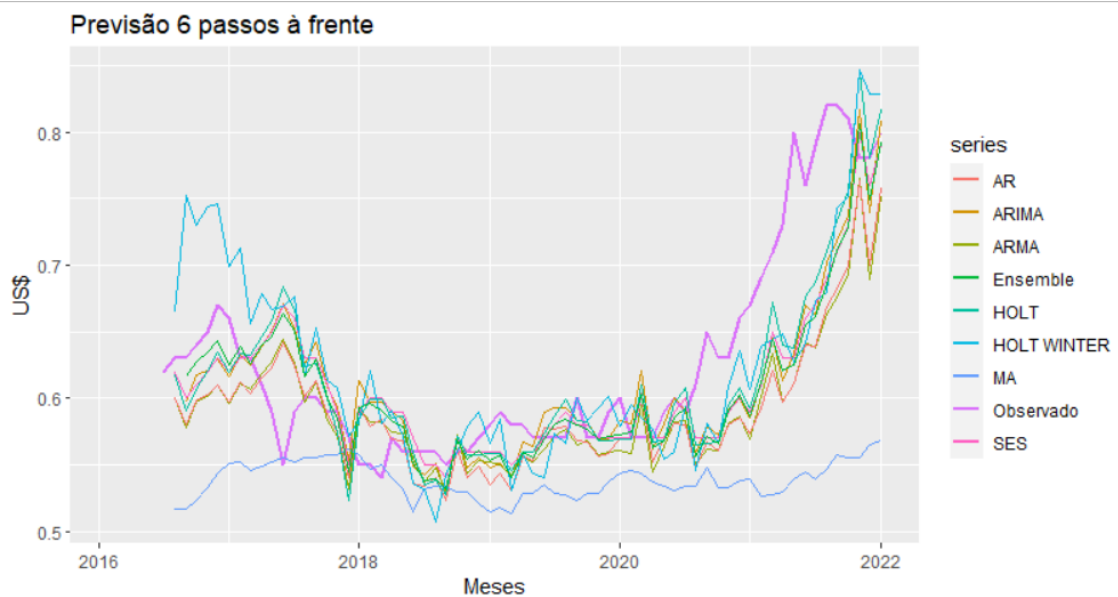
Os Gráficos 57, 58, 59 e 60 apresentam as previsões realizadas pelos modelos e os dados reais separados para teste.

**Gráfico 57 - Previsão 6 passos à frente – Brasil**



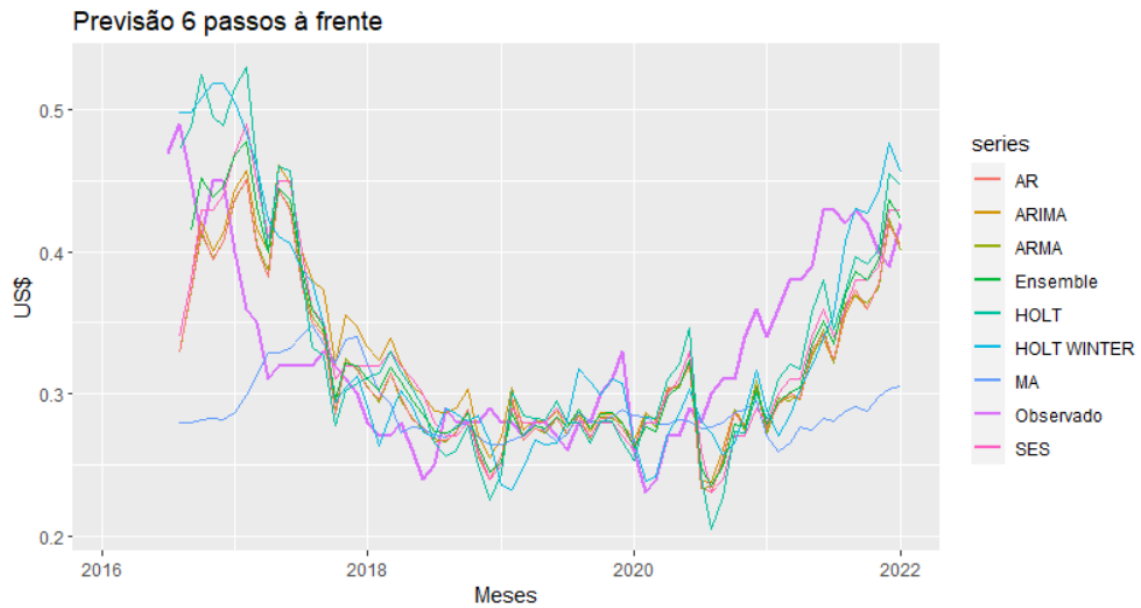
Fonte: Autoria própria (2022)

**Gráfico 58 - Previsão 6 passos à frente – Estados Unidos**



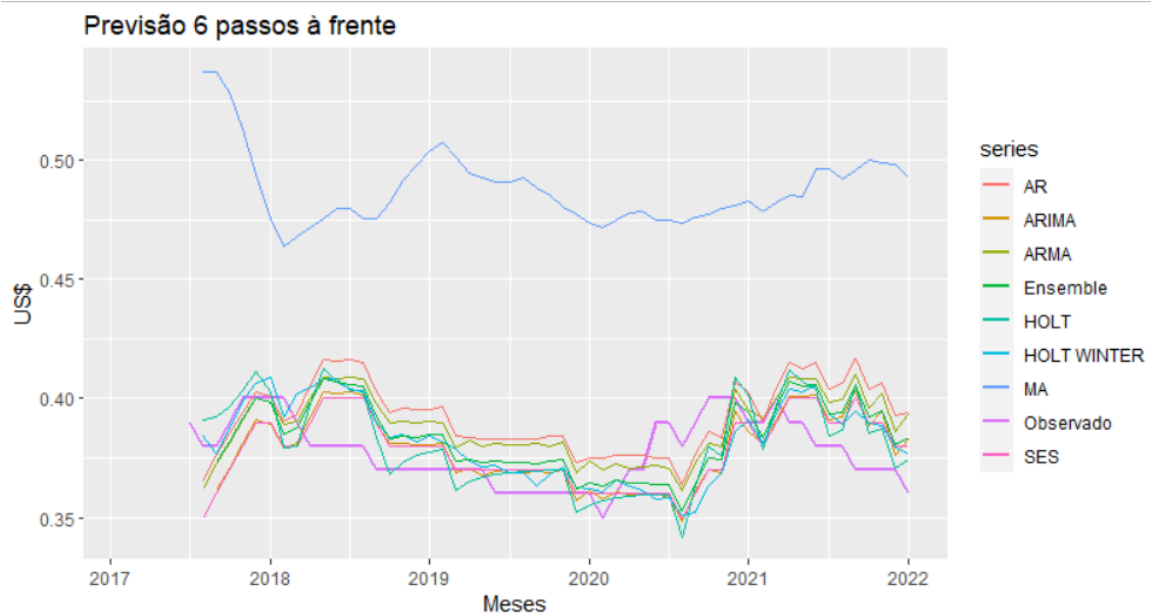
Fonte: Autoria própria (2022)

Gráfico 59 - Previsão 6 passos à frente – Mundo



Fonte: Autoria própria (2022)

Gráfico 60 - Previsão 6 passos à frente – União Europeia



Fonte: Autoria própria (2022)

#### 4.4.4 Previsão 12 passos à frente

Por fim, fez-se também previsões para as mesmas séries considerando o horizonte de 12 passos à frente, tendo seus resultados comparados pelas mesmas métricas de previsão citadas anteriormente, conforme Tabela 9 e ranking dos melhores preditores na Tabela 10.

Tabela 9 - Performance dos modelos – 12 passos à frente

| MODELOS      | Métricas de desempenho | Horizonte de Previsão 12 passos à frente |                |                |                |
|--------------|------------------------|--|----------------|----------------|----------------|
|              |                        | Brasil                                   | Estados Unidos | Mundo          | União Europeia |
| AR           | RMSE                   | 0,0640                                   | 0,0953         | 0,0665         | 0,0371         |
|              | MAE                    | <b>0,0427</b>                            | 0,0661         | 0,0514         | 0,0322         |
|              | MAPE                   | <b>12,6162</b>                           | 9,5818         | 15,2953        | 8,7024         |
| MA           | RMSE                   | <b>0,0619</b>                            | 0,1148         | <b>0,0599</b>  | 0,0596         |
|              | MAE                    | 0,0436                                   | 0,0847         | <b>0,0417</b>  | 0,0443         |
|              | MAPE                   | 13,1439                                  | 12,3741        | <b>11,8938</b> | 13,8121        |
| ARMA         | RMSE                   | 0,0653                                   | 0,0964         | 0,067          | 0,0305         |
|              | MAE                    | 0,0440                                   | 0,0665         | 0,0518         | 0,0271         |
|              | MAPE                   | 12,9706                                  | 9,6096         | 15,4253        | 7,3149         |
| ARIMA        | RMSE                   | 0,0673                                   | 0,0847         | 0,0789         | 0,0241         |
|              | MAE                    | 0,0469                                   | 0,0622         | 0,0639         | 0,0207         |
|              | MAPE                   | 14,0268                                  | 9,2696         | 19,7316        | 5,4656         |
| SES          | RMSE                   | 0,0679                                   | 0,084          | 0,0801         | <b>0,0240</b>  |
|              | MAE                    | 0,0479                                   | 0,0612         | 0,0625         | <b>0,0204</b>  |
|              | MAPE                   | 14,3587                                  | <b>9,0885</b>  | 19,1304        | <b>5,3825</b>  |
| HOLT         | RMSE                   | 0,0677                                   | <b>0,0816</b>  | 0,0956         | 0,0268         |
|              | MAE                    | 0,0463                                   | <b>0,0612</b>  | 0,0707         | 0,0223         |
|              | MAPE                   | 14,2025                                  | 9,2001         | 21,6037        | 5,8808         |
| HOLT WINTERS | RMSE                   | 0,0676                                   | 0,0817         | 0,0904         | 0,0263         |
|              | MAE                    | 0,0494                                   | 0,0612         | 0,0675         | 0,0235         |
|              | MAPE                   | 15,4168                                  | 9,1935         | 20,4923        | 6,2210         |
| ENSEMBLE     | RMSE                   | 0,0645                                   | 0,0874         | 0,0707         | 0,0260         |
|              | MAE                    | 0,0438                                   | 0,0643         | 0,0552         | 0,0239         |
|              | MAPE                   | 12,9582                                  | 9,5277         | 16,5797        | 6,3758         |

Fonte: Autoria própria (2022)

Tabela 10 - Ranking de desempenho das previsões 12 passos à frente

| MODELOS      | Horizonte de Previsão 3 passos à frente |     |      |                |     |      |       |     |      |                |     |      | Ranking |
|--------------|---|-----|------|----------------|-----|------|-------|-----|------|----------------|-----|------|---------|
|              | Brasil                                  |     |      | Estados Unidos |     |      | Mundo |     |      | União Europeia |     |      |         |
|              | RMSE                                    | MAE | MAPE | RMSE           | MAE | MAPE | RMSE  | MAE | MAPE | RMSE           | MAE | MAPE |         |
| AR           |   | 1   | 1    |                |     |      |       |     |      |                |     |      | 2       |
| MA           | 1                                       |     |      |                |     |      | 1     | 1   | 1    |                |     |      | 4       |
| ARMA         |   |     |      |                |     |      |       |     |      |                |     |      | 0       |
| ARIMA        |   |     |      |                |     |      |       |     |      |                |     |      | 0       |
| SES          |   |     |      |                | 1   | 1    |       |     |      | 1              | 1   | 1    | 5       |
| HOLT         |   |     |      |                | 1   |      |       |     |      |                |     |      | 1       |
| HOLT WINTERS |   |     |      | 1              | 1   |      |       |     |      |                |     |      | 2       |
| ENSEMBLE     |   |     |      |                |     |      |       |     |      |                |     |      | 0       |

Fonte: Autoria própria (2022)

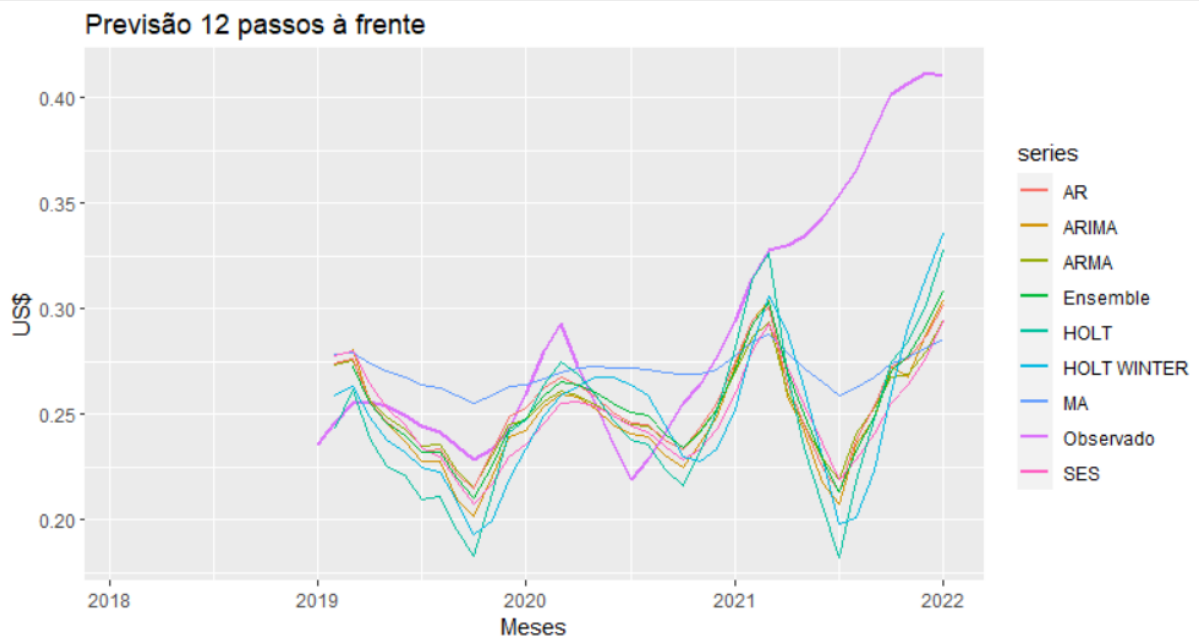
Analisando a previsão 12 passos à frente, observou-se que o método SES apresentou melhor desempenho no ranking geral. No entanto, os erros se mostraram

muito alto, pois quanto maior o horizonte de previsão, mais incerteza estará associada à previsão.

Como o caso do RMSE que os eleva os erros ao quadrado, o modelo MA desempenhou melhor ajuste para a séries Brasil e Mundo. Já para a série Estados Unidos o modelo HOLT apresentou melhores resultados, seguido do SES para a série União Europeia.

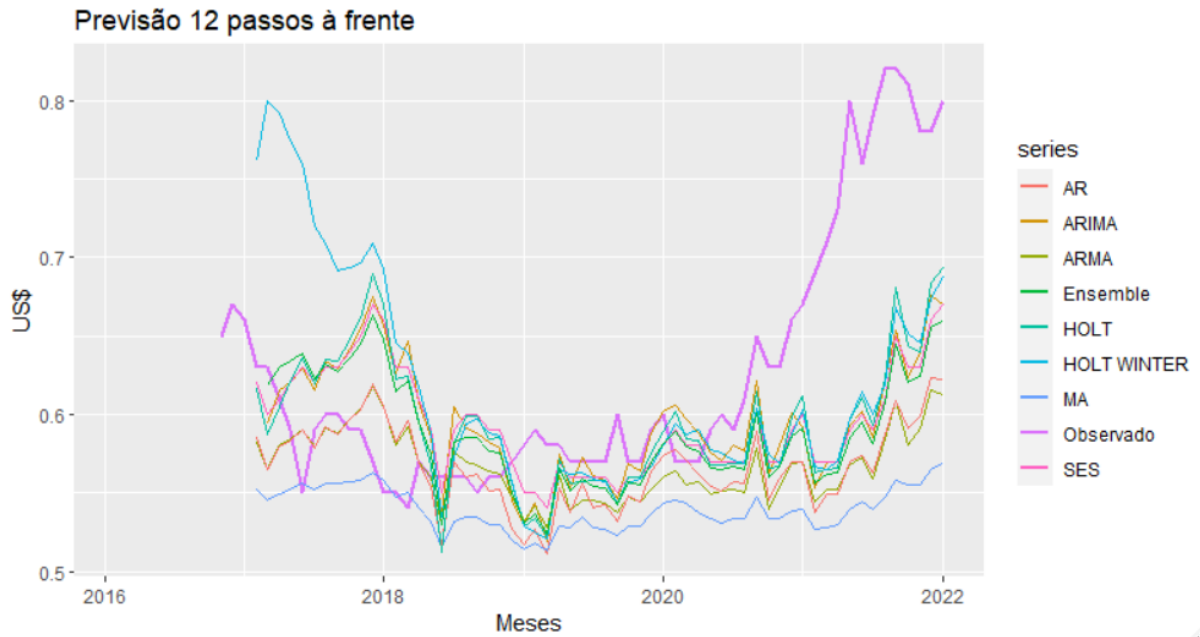
Os Gráficos 61, 62, 63 e 64 trazem as previsões realizadas pelos modelos e os dados reais separados para teste.

**Gráfico 61 - Previsão 12 passos à frente - Brasil**



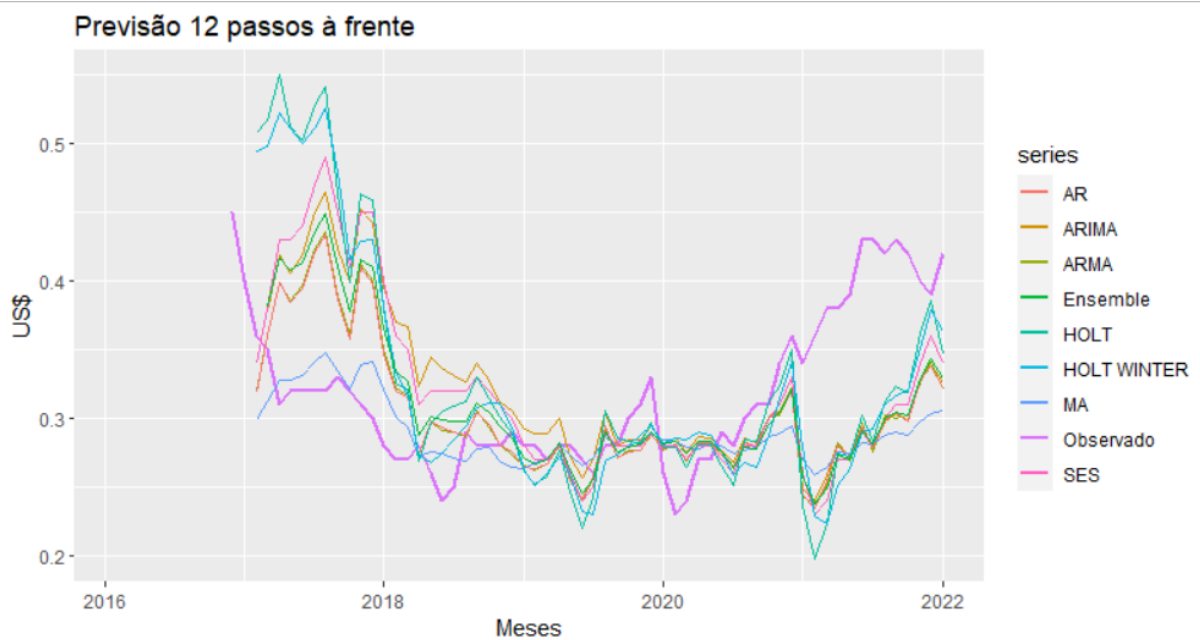
Fonte: Autoria própria (2022)

Gráfico 62 - Previsão 12 passos à frente – Estados Unidos



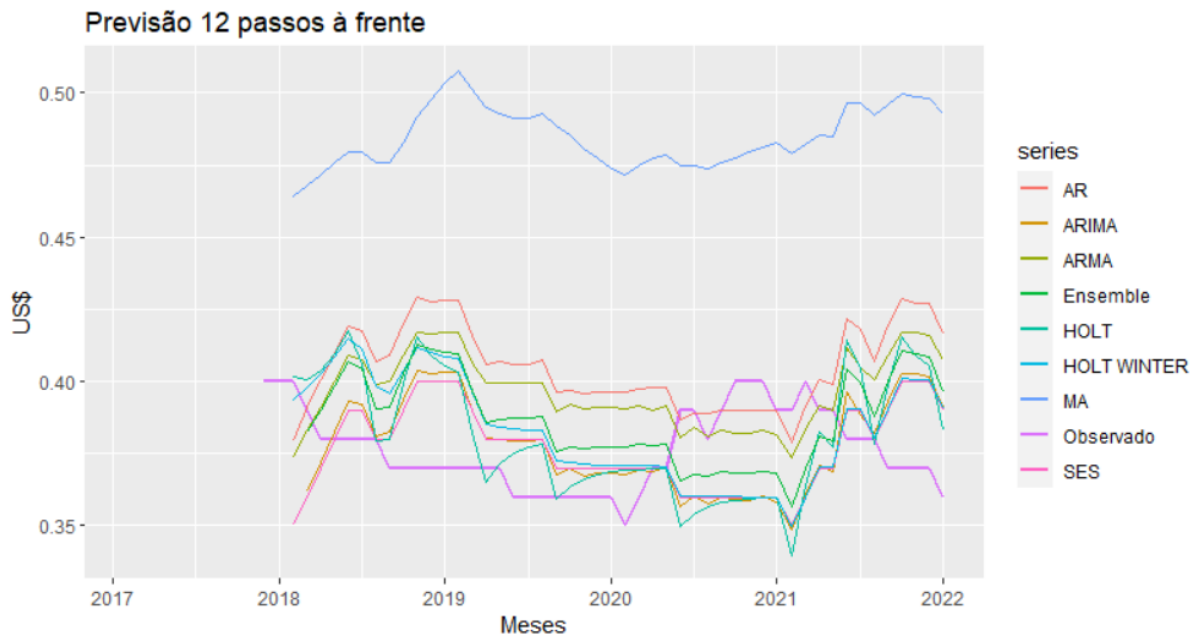
Fonte: Autoria própria (2022)

Gráfico 63 - Previsão 12 passos à frente – Mundo



Fonte: Autoria própria (2022)

Gráfico 64 - Previsão 12 passos à frente – União Europeia



Fonte: Autoria própria (2022)

#### 4.5 Discussão

Em geral, os *ensembles* apresentam resultados satisfatórios, em alguns casos desempenhando melhores previsões quando comparado aos modelos individuais, demonstrando o potencial dessa abordagem. Segundo Adhikariet *et al.* (2015), a correta seleção dos modelos do *ensemble* pode melhorar potencialmente a exatidão geral da previsão. Isso porque nem todos os modelos produzem boas previsões para séries temporais específicas. O principal desafio na seleção de modelos é a escolha de uma heurística ou meta-heurística eficiente e rápida.

Também pode-se observar que quanto mais à frente prever, mais incerteza estará associada e, mais amplos serão os intervalos de previsão (BOX; JENKIS, 2015). Isso ocorre porque as correlações entre as amostras subsequentes são menores do que no caso de dados temporalmente consecutivos (SILVA *et al.*, 2019).

Dessa forma, pode-se observar que para previsão de 1 passo à frente os modelos desempenharam um bom ajuste. No entanto, aumentando o horizonte de previsão percebe-se que os erros também aumentam.

Avaliando os ajustes e modelos em geral, observou-se que o modelo MA é considerado inadequado para descrever o conjunto de dados em análises. Até



visualmente percebe-se que os gráficos produzem resultados ruins, de modo que os erros de previsão comprovam isso.

Uma desvantagem do modelo SES é a dificuldade em realizar previsões a longo prazo. Ao mesmo tempo a facilidade para atualização das previsões pode ser vista como uma vantagem.

Como modelar efetivamente essas séries e como melhorar a qualidade das previsões ainda são questões que podem ser investigadas, visto que nenhum modelo de previsão é exatamente perfeito, visto que existem várias externas não previstas.

O objetivo deste trabalho foi comparar o desempenho dos 7 modelos de previsão e *Ensemble* de média simples quando aplicados ao preço do açúcar. Assim, os resultados mostraram que o desempenho geral dos modelos avaliados via RMSE, MAE e MAPE evidenciaram que para previsões de 1, 3, 6 e 12 passos os melhores preditores foram HOLT, HOLT WINTERS, HOLT WINTERS e SES, respectivamente.

## 5 CONCLUSÃO

A análise e previsão de séries temporais é uma área de pesquisa ativa nas últimas décadas e seu desenvolvimento é fundamental para muitos processos de decisão. Dessa forma, o presente estudo buscou comparar o desempenho de previsão do preço do açúcar por meio dos modelos lineares e *ensemble*, utilizando o ambiente de programação RStudio.

Foram discutidos e aplicados os modelos Autorregressivo (AR), Médias Móveis (MA), Autorregressivo e Médias Móveis (ARMA), Autorregressivo Integrado de Médias Móveis (ARIMA), Suavização exponencial simples (SES), Suavização exponencial com tendência (SEH ou HOLT), Suavização exponencial com tendência e sazonalidade (SEHW ou HOLT WINTERS) e um *Ensemble* baseado na média. Previsões univariadas de 1, 3, 6 e 12 passos à frente foram abordadas.

Os resultados mostraram que o desempenho geral dos modelos HOLT WINTERS e ARIMA avaliados via RMSE, MAE e MAPE são bastante semelhantes nas previsões no horizonte de um passo à frente. No entanto, ao analisar horizontes maiores, como 3, 6 e 12 passos esse cenário muda, apresentando erros expressivamente grande. Pode-se observar que o modelo MA apresentou resultados inadequados para previsão de múltiplos passos. Por fim, nota-se que os modelos que apresentaram melhores desempenhos pertencem a família HOLT WINTERS.

Tanto as previsões de uma etapa quanto as de várias etapas são importantes para facilitar um planejamento e uma tomada de decisão. Eles simulam o ambiente de previsão do mundo real no qual os dados precisam ser projetados para períodos curtos e longos

Além disso, analisando o estado da arte, pode-se observar que a geração de projeções robustas para os preços do açúcar pode apoiar o planejamento e o investimento setorial, além de fornecer uma avaliação baseada em evidências da natureza, nível e duração das intervenções no mercado, como aquelas que envolvem compras de ações domésticas de açúcar e vendas no mercado aberto.

Em geral, sabe-se que, nenhuma técnica de modelagem de previsão é exatamente completa. Dessa forma, para pesquisas futuras recomenda-se explorar modelos de previsão não lineares, além das especificações lineares exploradas neste estudo. Combinando diferentes métodos, o problema de seleção de modelos pode ser amenizado com menos esforço computacional. Ainda, as séries temporais do mundo

real raramente são lineares ou não lineares puras. Assim, a combinação de diversos modelos previsões é uma tendência nas atuais linhas de pesquisas.

## REFERÊNCIAS

ADENÄUER, M.; LOUHICHI, K.; FRAHAN, B. H. D.; WITZKE, H. P. Impact of the “everything but arms” initiative on the EU sugar sub-sector. **In International Conference on Policy Modelling**, n. 330600001, EcoMod2004, Paris, v. 30, 2004.

ADHIKARI, R.; AGRAWAL, R. K. An introductory study on time series modeling and forecasting. **arXiv preprint arXiv:1302.6613**, 2013.

ADHIKARI, R.; VERMA, G; KHANDELWAL, I. A model ranking based selective ensemble approach for timeseries forecasting. **Procedia Computer Science** 48: 14 – 21, 2015.

AKAIKE, H. Statistical predictor identification. **Annals of the institute of Statistical Mathematics**, v. 22, n. 1, p. 203-217, 1970.

ALMASARWEH, M.; ALWADI, S. ARIMA model in predicting banking stock market data. **Modern Applied Science**, v. 12, n. 11, p. 309, 2018.

AMROUK, E. M.; HECKELEI, T. Forecasting International Sugar Prices: A Bayesian Model Average Analysis. **Sugar Tech**, v. 22, p. 552–562, 2020.

ARMSTRONG, J. S.; COLLOPY, F. Error measures for generalizing about forecasting methods: Empirical comparisons. **International journal of forecasting**, v. 8, n. 1, p. 69-80, 1992.

AWAJAN, A. M.; ISMAIL, M. T.; AL WADI, S. Improving forecasting accuracy for stock Market data using EMD-HW bagging. **PloS one**, v. 13, n. 7, p. e0199582, 2018.

BAJPAI, P.K., PRIYA, K.; MALIK, M. Selection of Appropriate Growth Model for Prediction of Sugarcane Area, Production and Productivity of India. **Sugar Tech** 14, 188–191, 2012.

BALCOMBE, K.; RAPSOMANIKIS, G. Bayesian estimation of nonlinear vector error correction models: the case of sugar-ethanol-oil nexus in Brazil. **Am. J. Agric. Econ.** 2008, 90, 658–668.

BALCOMBE, K. The nature and determinants of volatility in agricultural prices: An empirical study from 1962-2008. **MARKET REVIEW**, p. 2, 2010.

BALLINI, R. **Análise e Previsão de Vazões Utilizando Séries Temporais, Redes Neurais e Redes Neurais Nebulosas**. Tese de Doutorado, FEEC-Unicamp, Brasil, 2000.

BANDEIRA, S. G.; ALCALÁ, S. G. S.; VITA, R. O. BARBOSA, T. M. G. A. Comparison of selection and combination strategies for demand forecasting methods. **Production** 30: e20200009, 2020.

- BELOTTI, J. T. **Previsão de vazões afluentes utilizando redes neurais artificiais e ensembles**. 2019. Dissertação (Mestrado). Programa de Pós-Graduação em Ciência da Computação, Universidade Tecnológica Federal do Paraná, Ponta Grossa, 2019.
- BENEŠOVÁ, I.; ŘEZBOVÁ, H.; SMUTKA, L.; TOMŠÍK, K.; LAPUTKOVÁ, A. European sugar market – impact of quota system. **Acta Univ. Agric. Silvic. Mendelianae Brunensis**, v. 63, p. 1825-1838, 2015.
- BENTIVOGLIO, D.; FINCO, A.; BACCHI, M. R. P. Interdependencies between biofuel, fuel and food prices: The case of the Brazilian ethanol market. **Energies**, v. 9, n. 6, p. 464, 2016.
- BISHOP, Christopher M et al. **Neural networks for pattern recognition**. [S.l.]: Oxford university press, 1995.
- BOLÓN-CANEDO, V.; ALONSO-BETANZOS, A. Ensembles for feature selection: A review and future trends. **Information Fusion** 52: 1 – 12, 2019.
- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Times series analysis: forecasting and control**. 3ª Ed. San Francisco: Holden-Day, 1994.
- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C.; JUNG, G; M. **Time series analysis: forecasting and control**. John Wiley & Sons, 2015.
- BOX, G.; JENKINS, G. **Time Series Analysis Forecasting and Control**. Holden Day, San Francisco, Califórnia, 1970; John Wiley & Sons: Hoboken, NJ, EUA, 2015.
- BROCKWELI, P. J.; DAVIS, R. A. Introduction to time series and forecasting. **Springer**, 2016.
- BROWN, R. G. **Statistical forecasting for inventory control**. McGraw/Hill, 1959.
- CARMAN, H. F. A trend projection of high fructose corn syrup substitution for sugar. **American Journal of Agricultural Economics** 64: 625–633, 1982.
- CHANG, CL.; MCALEER, M.; WANG, YA. Modelling volatility spillovers for bio-ethanol, sugarcane and corn spot and futures prices. **Renewable and Sustainable Energy Reviews**, n. 81, p. 1002–1018, 2018.
- CHEN, C. AND LIU, H. Medium-term wind power forecasting based on multi-resolution multi-learner ensemble and adaptive model selection. **Energy Conversion and Management** 206: 112492, 2020.
- CHENG, S. Short-term Forecast Model of Sugar Futures Price Based on Seasonal Decomposition. In: **Fifth Symposium of Risk Analysis and Risk Management in Western China (WRARM 2017)**. Atlantis Press, 2017.

DAVID, S. A.; MACHADO, J. A. T.; TREVISAN, L. R.; INÁCIO JR., C. M.C.; LOPES, A. M. Dynamics of commodities prices: Integer and fractional models. **Fundamenta Informatica e**, v. 151, n. 1-4, p. 389-408, 2017.

SCHUTTER, O. Food commodities speculation and food price crises: Regulation to reduce the risks of price volatility. **United Nations Special Rapporteur on the Right to Food Briefing Note**, v. 2, p. 1-14, 2010.

FENGA, L. Bootstrap order determination for ARMA models: A comparison between different model selection criteria. **Journal of Probability and Statistics**, v. 2017, 2017.

FOOD AND AGRICULTURE ORGANIZATION (FAO). **Medium-term prospects for RAMHOT products**. Trade and Markets Division. Rome: FAO, 2016.

FOOD AND AGRICULTURE ORGANIZATION (FAO). **FAO agricultural outlook 2017–2026**. Paris: OECD Publishing, 2017.

GARDNER, E. S.; MCKENZIE, E. Forecasting trends in time series. **Management Science**, 31(10), 1237–1246, 1985.

GIANNAKIS, G. B.; MENDEL, J. M. Cumulant-based order determination of non-Gaussian ARMA models. **IEEE transactions on acoustics, speech, and signal processing**, v. 38, n. 8, p. 1411-1423, 1990.

GIL, A. C. **Como elaborar projetos de pesquisa**. 6. ed. São Paulo: Atlas, 2018.

GLANTZ, M.; MUN, J. **Chapter 8 - Projections and Risk Assessment**. Credit engineering for bankers. Academic Press. 2 ed, p. 185-236, 2011.

GRIGGS, P. D. **Global Industry, Local Innovation: the History of Cane Sugar Production in Australia**. Bern, Switerland, p. 1820-1995, 2011.

GUDOSHNIKOV, S.; JOLLY, L.; SPENCE, D. **The World Sugar Market**. Elsevier, 2004.

HOCHMAN, Gal; SEXTON, Steven; ZILBERMAN, David. Food and biofuel in a global environment. In: *Handbook of bioenergy economics and policy*. Springer, New York, NY, 2010. p. 267-286.

HOLT, C. E. **Forecasting seasonals and trends by exponentially weighted averages** (O.N.R. Memorandum No. 52). Carnegie Institute of Technology, Pittsburgh USA, 1957.

HUCHET-BOURDON, M. Agricultural commodity Price volatility: An Overview. In *OECD Food, Agriculture and Fisheries Working Papers*; Organisation for Economic Co-operation and Development (OECD) Publishing: Paris, France, 2011; p. 52.

HYNDMAN, R. J.; KHANDAKAR, Y. Automatic time series forecasting: the forecast package for R. **Journal of statistical software**, v. 27, n. 1, p. 1-22, 2008.

IHAKA, R.; R. **Past and Future History**. Statistics Department, University of Auckland, 1998.

INTERNATIONAL SUGAR ORGANIZATION (ISO). **Industrial and direct sugar consumption: an international survey**. London: ISO, 2016.

JACKS, D. S.; O'ROURKE, K. H.; WILLIAMSON, J. G. Commodity price volatility and world market integration since 1700. **Review of Economics and Statistics**, v. 93, n. 3, p. 800-813, 2011.

JEREZ, T.; KRISTJANPOLLER, W. Effects of the validation set on stock returns forecasting. **Expert Systems with Applications**, v. 150, p. 113271, 2020.

KANE, M. J.; PRICE, N.; SCOTCH, M.; RABINOWITZ, P. Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. **BMC bioinformatics**, v. 15, n. 1, p. 1-9, 2014.

KHALID, R; JAVAID, N. A survey on hyper parameters optimization algorithms of forecasting models in smart grid. **Sustainable Cities and Society**, p. 102275, 2020.

KOEHLER, A. B.; SNYDER, R. D.; ORD, J. K. Forecasting Models and Prediction Intervals for the Multiplicative Holt Winters Method. **International Journal of Forecasting**, v. 17, n. 2, p. 269-286, 2001.

KOTU, V.; DESHPANDE, B. Time series forecasting. **Data Science; Elsevier: Amsterdam, The Netherlands**, p. 395-445, 2019.

KOURENTZES, Nikolaos; BARROW, Devon K.; CRONE, Sven F. Neural network ensemble operators for time series forecasting. **Expert Systems with Applications**, v. 41, n. 9, p. 4235-4244, 2014.

KRISHNAMURTI, Tiruvalam Natarajan et al. Multimodel ensemble forecasts for weather and seasonal climate. **Journal of Climate**, v. 13, n. 23, p. 4196-4216, 2000.

KRISTOUFEK, L.; JANDA, K.; ZILBERMAN, D. Mutual Responsiveness of Biofuels, Fuels and Food Prices. CAMA Working Paper 38. 2012. Disponível em: <http://cama.crawford.anu.edu.au/pdf/working-papers/2012/382012.pdf>. Acesso em 10 de maio de 2022.

LAND, K. C. Forecasting. International Encyclopedia of the Social & Behavioral Sciences (Second Edition), **Elsevier**, 316–323, 2015.

LI, F.; ZHANG, L.; CHEN, B.; GAO, D.; CHENG, Y.; ZHANG, X.; YANG, Y.; GAO, K.; HUANG, Z. (2020). An optimal stacking ensemble for remaining useful life estimation of systems under multi-operating conditions. **IEEE Access** 8: 31854–31868, 2020.

LIMA, R. C.; GÓIS, M. R.; ULISES, C. Previsão de preços futuros de commodities agrícolas com diferenciações inteira e fracionária, e erros heteroscedásticos. **Revista de Economia e Sociologia Rural**, v. 45, n. 3, p. 621-644, 2007.

MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; HYNDMAN, R. J. **Forecasting: methods and applications**. 3ª ed., New York: John Wiley & Sons, 1998.

MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; HYNDMAN, R. J. **Forecasting methods and applications**. John Wiley & Sons, 2008.

MAKRIDAKIS, S.; SPILIOTIS, E.; ASSIMAKOPOULOS, V. Statistical and Machine Learning forecasting methods: Concerns and ways forward. **PLoS one**, v. 13, n. 3, p. e0194889, 2018.

MCKENZIE, E. D. General exponential smoothing and the equivalent ARMA process. **Journal of Forecasting**, v. 3, n. 3, p. 333-344, 1984.

MENSBRUGGHE, D. V. D.; BEGHIN, J. C.; MITCHELL, D. **Modeling tariff rate quotas in a global context**: The case of sugarmarkets in OECD countries. n. 03, -wp343. Center for Agricultural and Rural Development (CARD) Publications. Iowa State University, 2003.

MITCHELL, D. **Sugar policies opportunity for change. Policy, Research working paper, WPS 3222**. Washington, DC: World Bank, 2004.

MONTERO-MANSO, P., ATHANASOPOULOS, G., HYNDMAN, R. J.; TALAGALA, T. S. Feature-based forecast model averaging. **International Journal of Forecasting** 36(1): 86 – 92, 2020.

MOON, J., JUNG, S., REW, J., RHO, S. AND HWANG, E. Combination of short-term load forecasting models based on a stacking ensemble approach. **Energy and Buildings** 216: 109921, 2020.

MORANDAT, F; HILL, B.; OSVALD, L.; VITEK, J. **Evaluating the design of the R language: objects and functions for data analysis**. Purdue University, 2012.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. Ed. Blucher, 2006.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais: modelos lineares univariados**. São Paulo: Blucher, 3. ed., v. 1, 2018.

NOLTE, S.; BUYSSE, J.; HUYLENBROECK, G. V. Modelling the effects of an abolition of the EU sugar quota on internal prices, production and imports. **European review of agricultural economics**, v. 39, n. 1, p. 75-94, 2012.

OBSERVATORY OF ECONOMIC COMPLEXITY (OEC). **Raw Sugar**. 2019. Disponível em: <https://oec.world/en/profile/hs92/41701/>. Acesso em 25 de abril de 2021.



OSHODI, O. S.; EJOHWOMU, O. A.; FAMAKIN, I. O.; CORTEZ, P. Comparing univariate techniques for tender price index forecasting: Box-Jenkins and neural network model. **Construction Economics and Building**, v. 17, p. 109–123, 2017.

PAGANI, V.; STELLA, T.; GUARNERI, T.; FINOTTO, G.; BERG, M. V. D.; MARIN, F. R.; ACUTIS, M.; CONFALONIERI, R. Forecasting sugarcane yields using agro-climatic indicators and Canegro model: A case study in the main production region in Brazil. **Agricultural Systems**, v. 154, p. 45-52, 2017.

PAPASTEFANOPOULOS, V.; LINARDATOS, P.; KOTSIANTIS, S. Covid-19: A comparison of time series methods to forecast percentage of active cases per population. **Applied Sciences**, v. 10, n. 11, p. 3880, 2020.

PELLEGRINI, F. R.; FOGLIATTO, F. Estudo comparativo entre modelos de Winters e de Box-Jenkins para a previsão de demanda sazonal. **Revista Produto & Produção**, v. 4, p. 72-85, 2000.

PERRONE, Michael P; COOPER, Leon N. **When networks disagree: Ensemble methods for hybrid neural networks**. [S.l.], 1992.

PROFILLIDIS, V. A.; BOTZORIS, G. N. **Modeling of transport demand: Analyzing, calculating, and forecasting transport demand**. Elsevier, 2018.

RABOBANK. **Sugar quarterly Q4 2020**, London: RaboResearch, 2020. Disponível em:  
[https://research.rabobank.com/far/en/documents/105146\\_2889676\\_Rabobank\\_Sugar%20Quarterly%20Q4%202020.pdf](https://research.rabobank.com/far/en/documents/105146_2889676_Rabobank_Sugar%20Quarterly%20Q4%202020.pdf). Acesso em 01 de junho de 2021.

RAPSOMANIKIS, G.; HALLAM, D. **Threshold cointegration in the sugar-ethanol-oil price system in Brazil: evidence from nonlinear vector error correction models**. FAO commodity and trade policy research working paper 22. 2006. Available em: [http://sharing.mywordsolution.com/xtringfiles/240\\_thresholdVectorErrorCorrection.pdf](http://sharing.mywordsolution.com/xtringfiles/240_thresholdVectorErrorCorrection.pdf). Acesso em 10 de maio de 2022.

REYHANI, R.; MOGHADAM, A. M. E. A heuristic method for forecasting chaotic time series based on economic variables. In: **2011 Sixth International Conference on Digital Information Management**. IEEE, 2011. p. 300-304.

RUMÁNKOVÁ, L.; SMUTKA, L.; MAITAH, M.; BENEŠOVÁ, I. The interrelationship between sugar prices at the main world sugar commodities markets. **Sugar Tech**, v. 21, n. 6, p. 853-861, 2019.

SCHWARZ, G. Estimating the dimension of a model. **Annals of statistics**, v. 6, n. 2, p. 461-464, 1978.

SERRA, T. Volatility spillover between food and energy market: A semiparametric approach. **Energy Econ**. 2010, 33, 1155–1164.

SERRA, T.; ZILBERMAN, D.; GIL, J. Price volatility in ethanol markets. *Eur. Rev. Agric. Econ.* 2011, 38, 259–280.

SHADAB, A.; SAID, S.; AHMAD, S. Box–Jenkins multiplicative ARIMA modeling for prediction of solar radiation: a case study. *International Journal of Energy and Water Resources*, v. 3, n. 4, p. 305-318, 2019.

SHADAB, A.; AHMAD, S.; SAID, S. Spatial forecasting of solar radiation using ARIMA model. *Remote Sensing Applications: Society and Environment*, v. 20, p. 100427, 2020.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais Para Engenharia e Ciências Aplicadas: Fundamentos Teóricos e Aspectos Práticos**. 2 ed. São Paulo: Artliber, 2016.

SILVA, N.; SIQUEIRA, I.; OKIDA, S.; JR. STEVAN, S. L.; SIQUEIRA, H. Neural networks for predicting prices of sugarcane derivatives. *Sugar Tech*, v. 21, n. 3, p. 514-523, 2019.

STEPHEN, H. **World raw sugar prices: The influence of Brazilian costs of production and world surplus/deficit measures**. Sugar and Sweeteners SSS-M-297-01. United States Department of Agriculture, Economic Research Service, Washington, DC, 2013.

STEPHEN, H. **Projecting world raw sugar prices**. Sugar and Sweeteners SSS-M-317-01. United States Department of Agriculture, Economic Research Service, Washington, DC, 2015.

SURESH, K.; PRIYA, S. K. Forecasting sugarcane yield of tamilnadu using arima models. *Sugar Tech*, v. 13, n. 1, p. 23–26, 2011.

TEAM, R. Core et al. R: **A language and environment for statistical computing**. 2013.

TRATAR, L. F. Forecasting method for noisy demand. *International journal of production economics*, v. 161, p. 64-73, 2015.

TUBINO, D. F. **Planejamento e controle da produção: teoria e prática**. 3. ed. São Paulo: Atlas, 2017.

WICHARD, J. D. An adaptive forecasting strategy with hybrid ensemble models, 2016 **International Joint Conference on Neural Networks (IJCNN), IEEE, Vancou-ver, BC, CA**, pp. 1495–1498, 2016.

WINTERS, P. R. Forecasting sales by exponentially weighted moving averages. *Management science*, v. 6, n. 3, p. 324-342, 1960.

WORLD BANK. 2018. **World Bank commodities price data** (The Pink Sheet), September 2018. Washington, DC: World Bank.

YULE, G. U. Why do we sometimes get nonsense-correlations between Time-Series? A study in sampling and the nature of time-series. **Journal of the royal statistical society**, v. 89, n. 1, p. 1-63, 1926.

ZA'ER, S. AH.; OTHMAN, M.K. A.; ADNAN, M. AS.; MAHA, I. Z.; MOHAMMAD, S. S. ARMA model order and parameter estimation using genetic algorithms. **Mathematical and Computer Modelling of Dynamical Systems**, v. 18, n. 2, p. 201-221, 2012.

ZHANG, Y.; LIU, B.; JI, X.; HUANG, D. Classification of EEG signals based on autoregressive model and wavelet packet decomposition. **Neural Processing Letters**, v. 45, n. 2, p. 365-378, 2017.

## **APÊNDICE A – PARÂMETROS DOS MODELOS**

## Ajuste do modelo SES - Brasil

```
ETS(A,N,N)
Call:
ets(y = training, model = "ZNN", damped = NULL, alpha = NULL,
    Call:
      phi = NULL, additive.only = TRUE, lambda = TRUE, biasadj = TRUE,
    Call:
      restrict = TRUE, allow.multiplicative.trend = TRUE)
Box-Cox transformation: lambda= 1
Smoothing parameters:
  alpha = 0.9999
Initial states:
  l = -0.8922
sigma: 0.0178
      AIC      AICc      BIC
-536.1681 -536.0411 -526.3800
```

## Ajuste do modelo HOLT - Brasil

```
ETS(A,Ad,N)
Call:
ets(y = training, model = "ZZN", damped = TRUE, alpha = NULL,
    Call:
      phi = NULL, additive.only = TRUE, lambda = TRUE, biasadj = TRUE,
    Call:
      restrict = TRUE, allow.multiplicative.trend = TRUE)
Box-Cox transformation: lambda= 1
Smoothing parameters:
  alpha = 0.9999
  beta  = 0.2831
  phi   = 0.8
Initial states:
  l = -0.8883
  b = 0.0018
sigma: 0.0171
      AIC      AICc      BIC
-547.5744 -547.1228 -527.9983
```

## Ajuste do modelo HOLT WINTERS – Brasil

```
ETS(A,Ad,A)
Call:
ets(y = training, model = "ZZA", damped = TRUE, alpha = NULL,
Call:
  phi = NULL, additive.only = TRUE, lambda = TRUE, biasadj = TRUE,
Call:
  restrict = TRUE, allow.multiplicative.trend = TRUE)
Box-Cox transformation: lambda= 1
Smoothing parameters:
alpha = 0.2113
beta = 0.2112
gamma = 0.181
phi = 0.8065
Initial states:
l = -0.9177
b = 0.0086
s = -0.025 0.0132 -0.0484 0.0112 0.0053 0.007
      -0.0257 -1e-04 0.0074 0.0326 7e-04 0.0219
sigma: 0.0279
      AIC      AICc      BIC
-347.0469 -343.1159 -288.3185
```

## Ajuste do modelo AR – Brasil

```
Series: training
ARIMA(4,0,0) with non-zero mean
Coefficients:
      ar1      ar2      ar3      ar4      mean
      1.2843 -0.3221  0.1939 -0.1742  0.2649
s.e.  0.0706  0.1159  0.1157  0.0707  0.0583
sigma^2 = 0.0002786: log likelihood = 516.54
AIC=-1021.08  AICc=-1020.63  BIC=-1001.5
```

## Ajuste do modelo MA - Brasil

```
Series: training
ARIMA(0,0,4)(0,0,1)[12] with non-zero mean
Coefficients:
      ma1      ma2      ma3      ma4      sma1      mean
      1.6665  1.6467  1.4212  0.7897  0.3884  0.2792
s.e.  0.0605  0.0926  0.0701  0.0397  0.0638  0.0141
sigma^2 = 0.0005062: log likelihood = 458
AIC=-902  AICc=-901.4  BIC=-879.16
```

## Ajuste do modelo ARMA – Brasil

```
Series: training
ARIMA(1,0,4) with non-zero mean
Coefficients:
      ar1      ma1      ma2      ma3      ma4      mean
      0.9737  0.3319  0.0834  0.1802  0.0586  0.2603
s.e.  0.0161  0.0767  0.0819  0.0679  0.0727  0.0640
sigma^2 = 0.0002804: log likelihood = 516.45
AIC=-1018.9  AICc=-1018.29  BIC=-996.06
```

### Ajuste do modelo ARIMA - Brasil

Series: training  
ARIMA(1,1,4)

Coefficients:

|      | ar1     | ma1    | ma2    | ma3    | ma4    |
|------|---------|--------|--------|--------|--------|
|      | -0.9239 | 1.3366 | 0.4038 | 0.2308 | 0.2615 |
| s.e. | 0.0560  | 0.0859 | 0.1260 | 0.1196 | 0.0644 |

sigma<sup>2</sup> = 0.000258: log likelihood = 522.81  
AIC=-1033.62 AICc=-1033.16 BIC=-1014.07

### Ajuste do modelo SES – Estados Unidos

ETS(A,N,N)

Call:  
ets(y = training, model = "ZNN", damped = NULL, alpha = NULL,

Call:  
phi = NULL, additive.only = TRUE, lambda = TRUE, biasadj = TRUE,

Call:  
restrict = TRUE, allow.multiplicative.trend = TRUE)

Box-Cox transformation: lambda= 1

Smoothing parameters:  
alpha = 0.9999

Initial states:  
l = -0.53

sigma: 0.0221

| AIC       | AICc      | BIC       |
|-----------|-----------|-----------|
| -561.5703 | -561.4861 | -550.5710 |

### Ajuste do modelo HOLT – Estados Unidos

ETS(A,Ad,N)

Call:  
ets(y = training, model = "ZZN", damped = TRUE, alpha = NULL,

Call:  
phi = NULL, additive.only = TRUE, lambda = TRUE, biasadj = TRUE,

Call:  
restrict = TRUE, allow.multiplicative.trend = TRUE)

Box-Cox transformation: lambda= 1

Smoothing parameters:  
alpha = 0.9999  
beta = 0.1001  
phi = 0.8607

Initial states:  
l = -0.5266  
b = 0.0017

sigma: 0.022

| AIC       | AICc      | BIC       |
|-----------|-----------|-----------|
| -562.7427 | -562.4448 | -540.7442 |

## Ajuste do modelo HOLT WINTERS – Estados Unidos

```
ETS(A,Ad,A)
Call:
ets(y = training, model = "ZZA", damped = TRUE, alpha = NULL,
    Call:
      phi = NULL, additive.only = TRUE, lambda = TRUE, biasadj = TRUE,
    Call:
      restrict = TRUE, allow.multiplicative.trend = TRUE)
Box-Cox transformation: lambda= 1
Smoothing parameters:
alpha = 0.7662
beta = 0.0737
gamma = 0.0307
phi = 0.8813
Initial states:
l = -0.7541
b = 0.0123
s = 0.0112 0.0141 0.0075 -0.0131 0.0117 -0.0079
      0.0059 -0.0127 -0.0091 -0.0084 -0.0092 0.0101
sigma: 0.0288
      AIC      AICC      BIC
-393.5977 -391.0644 -327.6020
```

## Ajuste do modelo AR – Estados Unidos

```
Series: training
ARIMA(3,0,0)(2,0,0)[12] with non-zero mean
Coefficients:
      ar1      ar2      ar3      sar1      sar2      mean
      1.3293 -0.4858  0.1274  0.1632 -0.1561  0.5181
s.e.  0.0585  0.0937  0.0584  0.0586  0.0594  0.0381
sigma^2 = 0.000415: log likelihood = 716.13
AIC=-1418.26  AICC=-1417.86  BIC=-1392.6
```

## Ajuste do modelo MA – Estados Unidos

```
Series: training
ARIMA(0,0,4)(0,0,1)[12] with non-zero mean
Coefficients:
      ma1      ma2      ma3      ma4      sma1      mean
      1.5234  1.4905  0.9994  0.4530  0.3954  0.5170
s.e.  0.0537  0.0792  0.0794  0.0527  0.0518  0.0113
sigma^2 = 0.0006757: log likelihood = 645.4
AIC=-1276.81  AICC=-1276.41  BIC=-1251.14
```

## Ajuste do modelo ARMA – Estados Unidos

```
Series: training
ARIMA(2,0,3) with non-zero mean
Coefficients:
      ar1      ar2      ma1      ma2      ma3      mean
      0.0937  0.8335  1.2478  0.3908  0.0030  0.5188
s.e.  0.1341  0.1271  0.1520  0.1207  0.0681  0.0403
sigma^2 = 0.0004261: log likelihood = 712.67
AIC=-1411.33  AICC=-1410.93  BIC=-1385.67
```



## Ajuste do modelo ARIMA – Estados Unidos

```
Series: training
ARIMA(0,1,3)(2,0,0)[12]

Coefficients:
      ma1      ma2      ma3      sar1      sar2
    0.3701  0.0183 -0.1608  0.1448 -0.1614
s.e.  0.0579  0.0587  0.0572  0.0581  0.0591

sigma^2 = 0.0004134:  log likelihood = 715.27
AIC=-1418.54  AICc=-1418.25  BIC=-1396.57
```

## Ajuste do modelo SES – Mundo

```
ETS(A,N,N)

Call:
ets(y = training, model = "ZNN", damped = NULL, alpha = NULL,

Call:
  phi = NULL, additive.only = TRUE, lambda = TRUE, biasadj = TRUE,

Call:
  restrict = TRUE, allow.multiplicative.trend = TRUE)

Box-Cox transformation: lambda= 1

Smoothing parameters:
  alpha = 0.9999

Initial states:
  l = -0.82

sigma: 0.0234

      AIC      AICC      BIC
-529.0681 -528.9839 -518.0688
```

## Ajuste do modelo HOLT – Mundo

```
ETS(A,Ad,N)

Call:
ets(y = training, model = "ZZN", damped = TRUE, alpha = NULL,

Call:
  phi = NULL, additive.only = TRUE, lambda = TRUE, biasadj = TRUE,

Call:
  restrict = TRUE, allow.multiplicative.trend = TRUE)

Box-Cox transformation: lambda= 1

Smoothing parameters:
  alpha = 0.9999
  beta  = 0.1402
  phi   = 0.8

Initial states:
  l = -0.8377
  b = 0.0218

sigma: 0.0233

      AIC      AICC      BIC
-529.2889 -528.9911 -507.2904
```

### Ajuste do modelo HOLT WINTERS – Mundo

```
ETS(A,Ad,A)
Call:
ets(y = training, model = "ZZA", damped = TRUE, alpha = NULL,
    Call:
      phi = NULL, additive.only = TRUE, lambda = TRUE, biasadj = TRUE,
    Call:
      restrict = TRUE, allow.multiplicative.trend = TRUE)
Box-Cox transformation: lambda= 1
Smoothing parameters:
  alpha = 0.3458
  beta  = 0.1336
  gamma = 0.1348
  phi   = 0.8002
Initial states:
  l = -0.7117
  b = 0.0118
  s = 0.0597 -0.0603 0.0206 0.012 -0.0062 0.0222
      -0.0289 0.0109 0.0248 -0.0087 -0.0082 -0.0379
sigma: 0.0381
      AIC      AICc      BIC
-233.2941 -230.7607 -167.2984
```

### Ajuste do modelo AR – Mundo

```
Series: training
ARIMA(2,0,0)(1,0,0)[12] with non-zero mean
Coefficients:
      ar1      ar2      sar1      mean
      1.2949 -0.3273  0.1622  0.2788
s.e.  0.0563  0.0562  0.0609  0.0437
sigma^2 = 0.000475: log likelihood = 695.79
AIC=-1381.58  AICc=-1381.37  BIC=-1363.25
```

### Ajuste do modelo MA – Mundo

```
Series: training
ARIMA(0,0,4)(0,0,1)[12] with non-zero mean
Coefficients:
      ma1      ma2      ma3      ma4      sma1      mean
      1.4836  1.5175  1.1062  0.4839  0.3709  0.2794
s.e.  0.0534  0.0773  0.0762  0.0504  0.0522  0.0122
sigma^2 = 0.0007758: log likelihood = 625.52
AIC=-1237.04  AICc=-1236.65  BIC=-1211.38
```

### Ajuste do modelo ARMA – Mundo

```
Series: training
ARIMA(3,0,1)(1,0,0)[12] with non-zero mean

Coefficients:
      ar1      ar2      ar3      ma1      sar1      mean
 0.3509  0.8603 -0.2747  0.9806  0.1748  0.2789
s.e.  0.0603  0.0467  0.0591  0.0170  0.0617  0.0442

sigma^2 = 0.0004666: log likelihood = 699.18
AIC=-1384.36  AICc=-1383.96  BIC=-1358.7
```

### Ajuste do modelo ARIMA – Mundo

```
Series: training
ARIMA(2,1,1)(1,0,0)[12]

Coefficients:
      ar1      ar2      ma1      sar1
 1.2305 -0.3414 -0.9171  0.1502
s.e.  0.1013  0.0559  0.0966  0.0644

sigma^2 = 0.0004766: log likelihood = 694.58
AIC=-1379.16  AICc=-1378.95  BIC=-1360.84
```

### Ajuste do modelo SES – União Europeia

```
ETS(A,N,N)

Call:
ets(y = training, model = "ZNN", damped = NULL, alpha = NULL,

Call:
phi = NULL, additive.only = TRUE, lambda = TRUE, biasadj = TRUE,

Call:
restrict = TRUE, allow.multiplicative.trend = TRUE)

Box-Cox transformation: lambda= 1

Smoothing parameters:
alpha = 0.9999

Initial states:
l = -0.3498

sigma: 0.018

      AIC      AICc      BIC
-611.2789 -611.1777 -600.8245
```

## Ajuste do modelo HOLT – União Europeia

```
ETS(A,Ad,N)
Call:
ets(y = training, model = "ZZN", damped = TRUE, alpha = NULL,
    Call:
      phi = NULL, additive.only = TRUE, lambda = TRUE, biasadj = TRUE,
    Call:
      restrict = TRUE, allow.multiplicative.trend = TRUE)
Box-Cox transformation: lambda= 1
Smoothing parameters:
  alpha = 0.9999
  beta  = 0.2673
  phi   = 0.8
Initial states:
  l = -0.3179
  b = -0.0287
sigma: 0.0182
      AIC      AICc      BIC
-602.3713 -602.0124 -581.4626
```

## Ajuste do modelo HOLT WINTERS – União Europeia

```
ETS(A,Ad,A)
Call:
ets(y = training, model = "ZZA", damped = TRUE, alpha = NULL,
    Call:
      phi = NULL, additive.only = TRUE, lambda = TRUE, biasadj = TRUE,
    Call:
      restrict = TRUE, allow.multiplicative.trend = TRUE)
Box-Cox transformation: lambda= 1
Smoothing parameters:
  alpha = 0.6972
  beta  = 0.0045
  gamma = 0.3006
  phi   = 0.8642
Initial states:
  l = -0.3425
  b = -0.0464
  s = 0.0133 0.0255 0.0058 -0.0082 -0.0184 -0.0381
      0.0026 -0.0242 -0.0062 0.0083 0.0164 0.0232
sigma: 0.0237
      AIC      AICc      BIC
-462.7160 -459.6349 -399.9897
```

## Ajuste do modelo AR – União Europeia

```
Series: training
ARIMA(2,0,0)(1,0,0)[12] with non-zero mean

Coefficients:
      ar1      ar2      sar1      mean
    1.1654 -0.1813  0.1172  0.5280
s.e.  0.0636  0.0641  0.0654  0.0673

sigma^2 = 0.0003122: log likelihood = 630.69
AIC=-1251.38  AICc=-1251.13  BIC=-1233.96
```

### Ajuste do modelo MA – União Europeia

Series: training  
 ARIMA(0,0,4)(0,0,1)[12] with non-zero mean

Coefficients:

|      | ma1    | ma2    | ma3    | ma4    | sma1   | mean   |
|------|--------|--------|--------|--------|--------|--------|
|      | 1.5343 | 1.4941 | 1.1321 | 0.5676 | 0.4804 | 0.5370 |
| s.e. | 0.0553 | 0.0840 | 0.0772 | 0.0523 | 0.0517 | 0.0126 |

sigma<sup>2</sup> = 0.0005722: log likelihood = 557.76  
 AIC=-1101.52 AICc=-1101.04 BIC=-1077.13

### Ajuste do modelo ARMA – União Europeia

Series: training  
 ARIMA(3,0,2) with non-zero mean

Coefficients:

|      | ar1     | ar2    | ar3    | ma1    | ma2    | mean   |
|------|---------|--------|--------|--------|--------|--------|
|      | -0.5805 | 0.9178 | 0.6222 | 1.7692 | 0.8576 | 0.5186 |
| s.e. | 0.1429  | 0.0276 | 0.1496 | 0.1067 | 0.1112 | 0.0780 |

sigma<sup>2</sup> = 0.0003002: log likelihood = 636.15  
 AIC=-1258.31 AICc=-1257.83 BIC=-1233.92

### Ajuste do modelo ARIMA – União Europeia

Series: training  
 ARIMA(2,1,2)(1,0,0)[12]

Coefficients:

|      | ar1     | ar2     | ma1    | ma2    | sar1   |
|------|---------|---------|--------|--------|--------|
|      | -1.5517 | -0.6189 | 1.7517 | 0.8438 | 0.0617 |
| s.e. | 0.1638  | 0.1671  | 0.1245 | 0.1252 | 0.0689 |

sigma<sup>2</sup> = 0.0002994: log likelihood = 635.26  
 AIC=-1258.52 AICc=-1258.16 BIC=-1237.63