Universidade Federal da Paraíba
Centro de Informática
Programa de Pós-Graduação em Modelagem Matemática e Computacional

PREVISÃO E ANÁLISE DO ICMS NA PARAÍBA

Melquisedec Anselmo da Costa Azevêdo

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional, UFPB, da Universidade Federal da Paraíba, como parte dos requisitos necessários à obtenção do título de Mestre em Modelagem Matemática e Computacional.

Orientador: Tatiene Correia de Souza

João Pessoa
Julho de 2021

Ata da Sessão Pública de Defesa de Dissertação de Mestrado de **MELQUISEDEC ANSELMO DA COSTA AZEVÊDO**, candidato ao título de Mestre em Matemática Computacional, na área de Modelagem Matemática e Computacional, realizada no dia 29 de julho de 2021.

Aos vinte e nove dias do mês de julho do ano de dois mil e vinte um, às 13h, por meio de videoconferência, reuniram-se os membros da Banca Examinadora constituída para julgar o Trabalho Final do Sr. Melquisedec Anselmo da Costa Azevedo, vinculado à Universidade Federal da Paraíba sob a matrícula nº 20191000620, candidato ao grau de Mestre em "*Modelagem Matemática e Computacional",* na linha de pesquisa *"Modelagem Probabilística"*, do Programa de Pós-Graduação em Modelagem Matemática e Computacional. A comissão examinadora foi composta pelas professoras Tatiene Correia de Souza, Orientadora e Presidente da Banca; Tarciana Liberal Pereira de Araújo, Examinadora Interna ao Programa; e Ana Flávia Uzeda dos Santos Macambira, Examinadora Externa ao Programa. Dando início aos trabalhos, a Presidente da Banca cumprimentou os presentes, comunicou a finalidade da reunião e passou a palavra ao candidato para que fizesse, oralmente, a exposição do trabalho de dissertação intitulado *"Previsão e análise do ICMS da Paraíba"*. Concluída a exposição, o candidato foi arguido pela Banca Examinadora que emitiu o seguinte parecer: "***aprovado***". Do ocorrido, eu, Gean Paulo Pereira Maurício de Barros, Secretário do Programa de Pós-Graduação em Modelagem Matemática e Computacional (PPGMMC), lavrei a presente ata, que vai assinada por mim e pelos membros da Banca Examinadora.

João Pessoa, 29 de julho de 2021.

Gean Paulo Pereira Maurício de Barros
Secretário do PPGMMC
SIAPE 2326476

Prof.ª Dr.ª Tatiene Correia de Souza
Orientadora (PPGMMC)

Prof.ª Dr.ª Tarciana Liberal Pereira de Araújo
Examinadora Interna ao Programa (PPGMMC)

Prof.ª Dr.ª Ana Flávia Uzeda dos Santos Macambira
Examinadora Externa ao Programa (UFPB)

Catalogação na publicação
Seção de Catalogação e Classificação

A994p Azevêdo, Melquisedec Anselmo da Costa.
Previsão e análise do ICMS na Paraíba / Melquisedec Anselmo da Costa Azevêdo. - João Pessoa, 2021.
75 f. : il.

Orientação: Tatiene Correia de Souza.
Dissertação (Mestrado) - UFPB/CI.

1. ICMS. 2. Holt-Winters. 3. SARIMA. 4. Impostos - Paraíba. I. Souza, Tatiene Correia de. II. Título.

UFPB/BC CDU 336.22(043)

Elaborado por WALQUELINE DA SILVA ARAUJO - CRB-15/514

Este trabalho de pesquisa só foi possível através do apoio e suporte da minha esposa Marcilene Azevêdo. Dedico esta pesquisa a ela.

Agradecimentos

Gostaria de agradecer, primeiramente a Deus por ter me dado saúde e forças para encarar mais essa jornada em minha vida com êxito, apesar de todos os obstáculos que surgiram.

Aos meus pais, Antonio de Moura Azevêdo e Selma da Costa Azevêdo, minha esposa Marcilene Azevêdo e meus filhos Anthonny e Esther, por entender minha ausência e estarem sempre me apoiando em todas minhas decisões.

A minha orientadora Tatiene, a qual no momento mais difíceis da minha vida me aconselhou a não desistir me ajudando a continuar nessa jornada.

Aos meus amigos, Anderson, Nivaldo e Washington, que foram meus companheiros de estudo e ensinamento no PPGMMC.

Aos professores do Programa em Pós-graduação em Modelagem Matemática e Computacional da UFPB.

Aos membros da banca examinadora, pelas sugestões e correções.

"Todas as vitórias ocultam uma abdicação."

(Simone Beauvoir)

Resumo da Dissertação apresentada ao PPGMMC/CI/UFPB como parte dos requisitos necessários para a obtenção do grau de Mestre em Modelagem Matemática e Computacional (M.Sc.)

PREVISÃO E ANÁLISE DO ICMS NA PARAÍBA

Melquisedec Anselmo da Costa Azevêdo

Julho/2021

Orientador: Tatiene Correia de Souza

Programa: Modelagem Matemática e Computacional

Resumo

A busca por antecipar os fatos é bastante comum ao longo dos tempos, considerando algo como provável com base em indícios, sejam eles científicos ou por crenças populares. No âmbito econômico as previsões são necessárias para que se possa planejar as ações com antecedência e concluir sobre as principais intervenções e suas prováveis consequências. Se o orçamento for superestimado, acarretará em gastos acima do previsto o que poderá gerar um déficit ou contingenciamento, que é a redução temporária das despesas para atingir a meta fiscal. Porém se os recursos forem subestimado, o que pode dificultar realizações ações de urgência, e/ou de extrema importância. Desta forma a presente dissertação apresenta uma metodologia de modelagem, previsão e análise das arrecadações do Imposto sobre Operações Relativas à Circulação de Mercadorias e sobre Prestações de Serviços de Transporte Interestadual e Intermunicipal e de Comunicação do Estado da Paraíba (ICMS-PB), por representar mais de 80% da receita tributária do Estado. Foram coletados dados de janeiro de 1997 a abril de 2021, que é truncada em datas distintas gerando quatro séries para verificar se a dinâmica da série varia. Assim é utilizando, para as quatro série, os algoritmos de alisamento exponencial Holt-Winters com sazonalidade aditiva e multiplicativa, e modelos Box-Jenkins com os modelos sazonais auto-regressivos integrados de médias móveis (SARIMA) e o SARIMAX com a variável dummy referente a pandemia do COVID-19, a tendência e a sazonalidade como variáveis de regressivas. Comparando-as entre se e com os valores reais das arrecadações do ICMS da Paraíba. Finalmente, considerando o erro quadrático médio e erro total obtidos através da relação entre as arrecadações e previsões, selecionou-se os modelos que geraram as melhores previsões para cada série, exibindo o gráfico com os valoreis reais, as previsões e o intervalo de confiança de 95%, verificando quais as circunstâncias que os modelos melhor se adéqua para prever o ICMS da Paraíba.

Palavra-Chave: ICMS, Holt-Winters, SARIMA, Imposto-Paraíba.

Abstract of Dissertation presented to PPGMMC/CI/UFPB as a partial fulfillment of the requirements for the degree of Master of Mathematical and Computational Modeling (M.Sc.)

PREVISÃO E ANÁLISE DO ICMS NA PARAÍBA

Melquisedec Anselmo da Costa Azevêdo

July/2021

Advisor: Tatiene Correia de Souza

Program: Computational Mathematical Modelling

Abstract

The search to anticipate the facts is quite common throughout the ages, considering something as likely based on clues, be they scientific or popular beliefs. In the economic context, forecasts are necessary to plan actions in advance and conclude on the main interventions and their likely consequences, because if the budget is overestimated, will lead to over-spending, which may lead to a deficit or contingencing, which is the temporary reduction of expenditure to reach the fiscal target and if resources are underestimated, which may hinder urgent and/or extremely important actions. In this way the present dissertation presents a modeling methodology, forecasting and analysing the collection of the Transaction Tax on the Movement of Goods and on the Provision of Interstate and Inter-municipal Transport and Communication Services of the State of Paraíba (ICMS-PB)for representing more than 80% of the State's tax revenue. Data were collected from January 1997 to April 2021, which is truncated into distinct dates generating four series to verify if the dynamics of the series vary. So it is using, for the four series, the Holt-Winters exponential smoothing algorithms with additive and multiplicative seasonality, and Box-Jenkins models with the integrated seasonal auto-regressive models of moving averages (SARIMA) and SARIMAX with the variable dummy referent with the dummy variable referring to the COVID-19 pandemic, trend and seasonality as regressive variables. Comparing them between themselves and with the real values of the ICMS of Paraíba. Finally, considering the mean quadratic error and total error obtained through the relationship between collections and forecasts, the models that generated the best forecasts for each series were selected, displaying the graph with the real values, the forecasts and the 95% confidence interval, verifyng the circumstances that the models best fit to predict the ICMS of Paraíba.

Keywords: ICMS, Holt-Winters, SARIMA, Tax-Paraíba.

Sumário

Resumo viii

Abstract x

Lista de Figuras xiii

Lista de Tabelas xv

1 Introdução 1

2 Conceitos Básicos de Séries Temporais 4

2.1 Conceitos Básicos 4

2.1.1 Séries Temporais 5

2.1.2 Funções de Autocovariância e Autocorrelação 7

2.2 Algoritmos de Suavização Exponencial 7

2.2.1 Alisamento Exponencial Simples 8

2.2.2 Alisamento Exponencial Holt 9

2.2.3 Alisamento Exponencial Holt-Winters 9

2.3 Modelagem de Box-Jenkins 11

2.3.1 Modelos Auto-Regressivos 11

2.3.2 Modelos de Médias Móveis 12

2.3.3 Modelos Auto-Regressivos de Médias Móveis 12

2.3.4 Modelos Auto-Regressivos Integrados de Médias Móveis 13

2.3.5 Modelo Auto-Regressivo Integrado de Médias Móveis Sazonal 14

2.3.6 Modelo Auto-Regressivo Integrado de Médias Móveis Sazonal com Variáveis Exógenas 15

3 Análise da Série 16

3.1 Modelagem, Previsões e Resultados 18

3.1.1 Alisamento Exponencial de Holt-Winters 20

3.1.2 Modelagem de Box-Jenkins-SARIMA 22
3.1.3 Modelagem de Box-Jenkins com variáveis exógenas 24
3.1.4 Medidas de Precisão 27

4 Conclusões **32**

Referências Bibliográficas **34**

A Diagnóstico **36**

B Algoritmos **41**

Lista de Figuras

3.1 Valores do ICMS-PB deflacionado expressos em milhões de reais. . . 17
3.2 Decomposição da Série. 18
3.3 Função de autocorrelação da série do ICMS-PB. 19
3.4 Função de autocorrelação parcial da Série do ICMS-PB. 19
3.5 Gráfico das previsões para série ICMSPB-abr.19 com os limites superior e inferior com 95% de confiança, obtidos pelo modelo SARIMA$(0,1,1) \times (0,0,2)_{BIC}$ 28
3.6 Gráfico das previsões para série ICMSPB-out.19 com os limites superior e inferior com 95% de confiança, obtidos pelo HW com sazonalidade multiplicativa. 29
3.7 Gráfico das previsões para série ICMSPB-abr.20 com os limites superior e inferior com 95% de confiança, obtidos pelo modelo SARIMA$(0,1,1) \times (0,0,2)_{BIC}$. 30
3.8 Gráfico das previsões para série ICMSPB-out.20 com os limites superior e inferior com 95% de confiança, obtidos pelo HW com sazonalidade aditiva. 31

A.1 SARIMA$(1,1,1) \times (0,0,2)$. 36
A.2 SARIMA$(0,1,1) \times (0,0,2)$. 36
A.3 SARIMA$(1,1,1) \times (0,0,2)$. 37
A.4 Análise diagnóstica- SARIMA$(0,1,1) \times (0,0,2)$. 37
A.5 SARIMA$(1,1,1) \times (0,0,2)$. 37
A.6 SARIMA$(0,1,1) \times (0,0,2)$. 37
A.7 SARIMA$(1,1,1) \times (0,0,2)$. 37
A.8 SARIMA$(0,1,1) \times (0,0,2)$. 37
A.9 SARIMAX$(0,0,0) \times (2,0,0)$, reg.tendência. 38
A.10 ARIMAX$(0,1,1)$, reg.sazonal. 38
A.11 SARIMAX$(0,0,0) \times (2,0,0)$, reg.tendência. 38
A.12 ARIMAX$(0,1,1)$, reg.sazonal. 38
A.13 SARIMAX$(0,0,0) \times (2,0,0)$, reg.tendência. 39

A.14 ARIMAX$(0,1,1)$, reg.sazonal. 39
A.15 SARIMAX$(1,1,1)\times(0,0,2)$, reg.pandemia. 39
A.16 SARIMAX$(0,1,1)\times(0,0,2)$, reg.pandemia. 39
A.17 SARIMA$(0,0,0)\times(2,0,0)$, reg.tendência. 40
A.18 ARIMAX$(0,1,1)$, reg.sazonal. 40
A.19 SARIMAX$(1,1,1)\times(0,0,2)$, reg.pandemia. 40
A.20 SARIMAX$(0,1,1)\times(0,0,2)$, reg.pandemia. 40

Lista de Tabelas

3.1 Previsões obtidas para a série ICMSPB-abr.19 pelos algoritmos de Holt-Winters aditivo e multiplicativo 12 passos à frente. 20
3.2 Previsões obtidas para a série ICMSPB-out.19 pelos algoritmos de Holt-Winters aditivo e multiplicativo 12 passos à frente. 21
3.3 Previsões obtidas para a série ICMSPB-abr.20 pelos algoritmos de Holt-Winters aditivo e multiplicativo 12 passos à frente. 21
3.4 Previsões obtidas para a série ICMSPB-out.20 pelos algoritmos de Holt-Winters aditivo e multiplicativo 6 passos à frente. 22
3.5 Previsões obtidas para a série ICMSPB-abr.19 pelo modelo SARIMA$(1,1,1)\times(0,0,2)_{AIC}$ e SARIMA$(0,1,1)\times(0,0,2)_{BIC}$ 12 passos à frente. 23
3.6 Previsões obtidas para a série ICMSPB-out.19 pelo modelo SARIMA$(1,1,1)\times(0,0,2)_{AIC}$ e SARIMA$(0,1,1)\times(0,0,2)_{BIC}$ 12 passos à frente. 23
3.7 Previsões obtidas para a série ICMSPB-abr.20 pelo modelo SARIMA$(1,1,1)\times(0,0,2)_{AIC}$ e SARIMA$(0,1,1)\times(0,0,2)_{BIC}$ 12 passos à frente. 24
3.8 Previsões obtidas para a série ICMSPB-out.20 pelo modelo SARIMA$(1,1,1)\times(0,0,2)_{AIC}$ e SARIMA$(0,1,1)\times(0,0,2)_{BIC}$ 6 passos à frente. 24
3.9 Previsões obtidas para a série ICMSPB-abr.19 pelo modelo ARIMAX$(0,1,1)_{AIC,BIC}$ 12 passos à frente. 25
3.10 Previsões obtidas para a série ICMSPB-out.19 pelo modelo ARIMAX$(0,1,1)_{AIC,BIC}$ 12 passos à frente. 25
3.11 Previsões obtidas para a série ICMSPB-abr.20 12 passos à frente. . . 26
3.12 Previsões obtidas para a série ICMSPB-out.20 6 passos à frente. . . . 26
3.13 Medidas de precisão obtidas para a série ICMSPB-abr.19. 27
3.14 Medidas de precisão obtidas para a série ICMSPB-out.19. 28
3.15 Medidas de precisão obtidas para a série ICMSPB-abr.20. 29
3.16 Medidas de precisão obtidas para a série ICMSPB-out.20. 30

Capítulo 1

Introdução

A busca por antecipar os fatos é bastante comum ao longo dos tempos, em que considerar algo como provável com base em indícios, sejam eles científicos ou por crenças populares, atrelado ao cotidiano tornou-se uma ferramenta essencial no desenvolvimento de ações preventivas, permitindo o planejamento e tomadas de decisões corretas com base em argumentos concretos. As previsões vão desde algo simples como estimar a quantidade de salgados que cada convidado come em um aniversário, à intenções de voto dos eleitos de uma região. A principal e mais conhecida é a previsão do tempo que auxilia na agricultura, no abastecimento de água, geração de energia pelas usinas hidrelétricas, entre outras ações, são inúmeras situações que utiliza-se das previsões dos eventos, algumas sem mesmo perceber o que está sendo realizado.

No âmbito econômico as previsões são necessárias para que se possa planejar as ações com antecedência e concluir sobre as principais intervenções e suas prováveis consequências. Uma das formas de antecipar eventuais imprevistos é através do planejamento financeiro, que segundo Macedo jr (2013, p.41) "é o processo de gerenciar seu dinheiro com o objetivo de atingir a satisfação pessoal, permitindo que você controle sua situação financeira para atender necessidades e alcançar objetivos no decorrer da vida".

Em uma escala maior como em um Estado, o planejamento financeiro é formado por um sistema integrado de planejamento e orçamento, como o Plano Plurianual (PPA), a Lei de Diretrizes Orçamentárias (LDO) e a Lei Orçamentária Anual (LOA), que são reconhecidos na Constituição Federal, dentre outros que variam de acordo com os Estados. Um orçamento superestimado, acarreta em gastos acima do previstos o que poderá gerar um déficit ou contingenciamento, que é a redução temporária das despesas para atingir a meta fiscal e se os recursos forem subestimados, pode dificultar realizações e ações de urgência.

Para evitar essas situações, é necessário programar os gastos prevendo de forma

eficiente as receitas. Uma das principais captações de recursos no país é o imposto Estadual sobre circulação de mercadorias e prestação de serviços de transporte interestadual e intermunicipal e comunicação - ICMS. Este tributo estadual incide sobre a movimentação de produtos de diferentes setores, ao realizar-se uma operação ou compra, salvo operações com livros, jornais, periódicos e papeis destinados a impressões dos mesmos, o valor do ICMS está inserido no valor dos produtos, representando em 2019, 84,56% dos impostos arrecadados pelos Estados no Brasil.

Sendo de competência de cada Estado instituir esse imposto, conforme descrito pelo art. 155, I, b, da C.F./88 (na redação da EC nº 3/93) para operações relativas à circulação de mercadorias e sobre prestações de serviços de transporte interestadual e intermunicipal e de comunicação, ainda que as operações e as prestações se iniciem no exterior, sendo um imposto presente em todo território nacional porém com alíquotas diferentes dependendo de cada Estado.

Entre os vários trabalhos a respeito do ICMS da literatura, Moço (2017) utiliza o modelo SARIMA e algoritmo Holt-Winters com sazonalidade aditiva e multiplicativa para estimar o ICMS do Rio de Janeiro. No Ceará, Silvestre (2013) utiliza os dados das arrecadações do ICMS do período de jan/00 a set/13 usando os modelos AR, MA e ARMA para prever os valores dos meses de jan/14 a mai/14. Castanho (2011) utiliza os modelos de Box-Jenkins e o alisamento Exponencial de Holt e Holt-Winters com sazonalidade aditiva e multiplicativa, ele analisa e prevê as arrecadações do ICMS do Espirito Santo, usando os dados de jan/00 a jan/09. O autor chegou à conclusão que o modelo de Holt-Winters é mais eficiente em previsões de períodos de curto prazo. Adicionalmente, Camargos (2008) analisa os tributos arrecadados no Estado de São Paulo, para a série do ICMS de jan/97 a jan/07, o modelo adotado é ARMA com variáveis dummies.

Na Paraíba, Duarte, Souza, e Girão (2014) preveem as arrecadações mensais do ICMS da Paraíba para o ano de 2013 utilizado os algoritmos Holt-Winters com sazonalidade aditiva e multiplicativa. Alem dos algoritmos Holt-Winters, Farias (2017) utilizou os modelos SARIMA e SARIMAX, considerando as arrecadações de jan/03 a ago/16 e previu a 5 passos a frente, ou seja até dez/16.

Desta forma a presente dissertação tem como objetivo realizar previsões e análise das arrecadações do ICMS no Estado da Paraíba, através dos modelos Box-Jenkins e do algoritmo de alisamento exponencial. A série temporal da arrecadação do ICMS da Paraíba correspondente do período de jan/97 a abril/21 foi extraída do site (https://www.confaz.fazenda.gov.br/boletim-de-arrecadacao-dos-tributos-estaduais) do Conselho Nacional de Política Fazendária (Confaz).

Afim de ter a mesma base de valores monetários será eliminado o efeito da inflação através do deflacionamento da série, ou seja, os valores arrecadados serão

corrigidos para retomar ao valor sem a incidência da inflação, o índice de correção utilizado é o Índice Geral de Preço - Disponibilidade Interna (IGP-DI), extraído do site (http://www14.fgv.br/fgvdados20/visualizaconsulta.aspx) do Instituto Brasileiro de Economia da Fundação Getúlio Vargas (FGV-IBE).

O IGP-DI é uma das versões do Índice Geral de Preços (IGP) que são medidos (ou calculados) durante todos os dias pelo FGV-IBE. Estes valores são formados por uma junção proporcional do IPA-DI (Índice de Preços por Atacado - Disponibilidade Interna), IPC-DI (Índice de Preços ao Consumidor - Disponibilidade Interna) e INCC-DI (Índice Nacional do Custo da Construção - Disponibilidade Interna).

O IGP-DI registra a inflação de preços que vão da produção de produtos agropecuários, industriais, alimentícios e bebidas, bens finais (combustíveis, máquinas e equipamentos, entre outros), bens intermediários (embalagens, suprimentos, entre outros), matérias-primas brutas até serviços finais, o que torna-o um importante índice inflacionário utilizado para reajustar as tarifas públicas, contratos de aluguel, planos e seguros entre outros valores.

O ambiente computacional utilizado para auxiliar nos cálculos é o `R`, pois é um software livre, ou seja, possibilita a criação de novas funções e modificar algumas já existentes, além de ser gratuito, pode-se fazer o download na plataforma: https://www.r-project.org/, para mais informações recomenda-se ler os manuais da linguagem presente em http://cran.r-project.org/manuals.html.

Esta dissertação apresenta uma metodologia de modelagem e previsão da arrecadação do Imposto sobre Operações Relativas à Circulação de Mercadorias e sobre Prestações de Serviços de Transporte Interestadual e Intermunicipal e de Comunicação do Estado da Paraíba (ICMS-PB). Para isto, esse trabalho está organizado em 6 capítulos.

No capítulo 2 são apresentados os conceitos básicos de séries temporais, os algoritmos de suavização exponencial e a modelagem de Box-Jenkins, bem como algumas definições. O Capítulo 3 traz uma análise da série do ICMS-PB, bem como a modelagem e as previsões em diferentes instantes da série, utilizando os algoritmos de alisamento exponencial de Holt-Winters com os efeitos sazonais aditivo e multiplicativo, assim como os modelos de Box-Jenkins, adicionalmente são apresentadas as medidas de precisão, dadas pelos erros quadrático médio-EQM e erros totais-ET para 1, 3, 6 e 12 passos à frente, obtidos nos modelos. Adicionalmente é selecionado o modelo que obteve as melhores medidas de precisão para cada série com base nos EQM e ET. O capítulo 4 apresenta as considerações finais sobre a análise dos resultados das principais modelagens realizadas para o ICMS da Paraíba.

Capítulo 2

Conceitos Básicos de Séries Temporais

Neste capítulo são introduzidos alguns conceitos básicos de sobre os algorítimos de suavização exponencial, bem como a modelagem de Box-Jenkins, que são utilizados como referência.

2.1 Conceitos Básicos

As séries temporais são conjuntos de observações de uma variável ao longo do tempo, sendo um importante instrumento na formulação de planos, ação e estratégias, possibilitando uma análise do que poderá ser feito, na identificação de crescimento/decrescimento dos dados, sua sazonalidade e além disso para prever observações futuras. Gomes (2003) elenca como principais objetivos para estudos das séries temporais, a saber, "investigar o mecanismo gerador da série, descrever seu comportamento através da construção de gráficos para verificação da existência de tendência e variações sazonais e fazer previsões de valores futuros".

Séries temporais podem ser definidas como contínuas, quando as observações são continuas no tempo, ou seja sendo y_t uma série temporal $y_t : t \in T$, $T = \{t : t_1 < t < t_2\}$ por exemplo do registro sobre o consumo de água na paraíba, ou discretas, quando as observações são feitas em tempos espaçados igualmente, ou seja, $y_t : t \in T$, $T = \{t_1, \cdots, t_n\}$, a exemplo dos valores diários da poluição na cidade de Itapororoca. Morettin e Toloi (2006) esclarecem que por muitas vezes uma serie temporal discreta pode ser obtida através das observações de uma série temporal contínua com tempos espaçados iguais, Δt. Assim, uma série contínua pode ser convertida e analisada como discreta, e tendo uma serie de N pontos no intervalo $[0, T]$, em que $N = \frac{T}{\Delta t}$.

E existe um número muito grande de modelos diferentes para descrever o comportamento de uma série. A formação destes modelos dependerá de vários fatores, tais como o comportamento dos dados e/ou do conhecimento antecipado do objetivo da análise. Os processos estocásticos, isto é, uma coleção de variáveis aleatórias utilizadas para estudar a evolução dos dados que são observados ao longo do tempo, controlados por leis probabilísticas, pode ser dividido em duas grandes classes de modelos, a saber: os estacionários e os não estacionários.

2.1.1 Séries Temporais

Uma série temporal $y_t,\ t \in \mathbb{N}$, é dita estacionária, de acordo com Morettin e Toloi (2006, p.4), "quando ela se desenvolve no tempo aleatoriamente ao redor de uma média constante, refletindo alguma forma de equilíbrio estável", sendo classificada em estritamente estacionaria ou de segunda ordem. Na série *estritamente estacionária* ou *fortemente estacionária* todas as distribuições de probabilidade conjuntas de dimensões finitas permanecem invariantes no tempo, em outros termos:

$$y_{t_1}, \ldots, y_{t_k} = y_{t_1+n}, \ldots, y_{t_k+n},$$

sendo que sua média e variância permanecem constantes na mudança do tempo, isto é,

$$\mathrm{E}[y_t] = \mu \ e \ \mathrm{Var}[y_t] = \sigma^2 \ , \ \forall t \ \epsilon \ \mathbb{N}.$$

, A série é *fracamente estacionária* ou de *segunda ordem* se a sua média é constante e sua função de autocovariância depende apenas da defasagem, ou seja:

$$\mathrm{E}[y_t] = \mu \ \ e \ \ \mathrm{Cov}[y_t, y_{t+n}] = \gamma_{t,t+n}.$$

Considerando uma série temporal $w_t : t = 1, 2, \ldots, n$ formada por variáveis aleatórias $w_1, w_2, \ldots, w_t$, sendo independentes e identicamente distribuídas *(i.i.d.)* e com uma média igual a zero, variância igual a σ^2 e com $\mathrm{Cor}(w_i, w_j) = 0 \ \forall \ i \neq j$, esse processo, estritamente estacionário, é definido por Cowpertwait e Metcalfe (2009) como ruído branco. BOX *et al.* (2015) afirmam que esse processo é um exemplo de fundamental importância de um processo estacionário, pois a independência resulta em w_t que são não correlacionados, ou seja, sua função de autocovariância é dada por:

$$\gamma_k = \mathrm{E}[w_t, w_{t+k}] = \begin{cases} \sigma_w^2 \ , \ k = 0 \\ 0 \ , \ k \neq 0 \end{cases}$$

utilizando apenas as propriedades de segunda ordem (fracamente estacionária), en-

tão uma sequência de variáveis aleatórias w_t, que não são correlacionadas, terá as mesmas características como acima citada, $\mathrm{E}(w_t) = 0$ e $\mathrm{Var}(w_t) = \sigma^2$ além de ter a mesma função de autocovariância γ_k, assim também pode ser chamado de ruído branco (BOX *et al.*, 2015).

Porém uma grande parte das séries encontradas apresentam algo de não estacionariedade, justificada em sua maioria, devido a quatro características a saber: sazonalidade, que é a presença de variações que ocorrem em intervalos regulares de tempo, tendência, que é o comportamento do movimento no desenvolver da série, o ciclo, refere-se aos movimentos recorrentes dos dados em torno da tendência, e por último o erro aleatório, que são movimentos não esperados, ou seja, não modeláveis. Um dos principais tipos de séries não estacionárias são os passeios aleatórios.

Cowpertwait e Metcalfe (2009) define o passeio aleatório como uma série temporal y_t, tal que:

$$y_t = y_{t-1} + w_t, \tag{2.1}$$

em que w_t representa o ruindo branco. Substituindo $y_{t-1} = y_{t-2} + w_{t-1}$, na equação (2.1) e continuando substituindo y_{t-2} e assim por diante, então resultará:

$$y_t = y_0 + w_t + w_{t-1} + w_{t-2} + \ldots \tag{2.2}$$

com y_0 tendendo a zero e com o tempo iniciando com t=1, a equação (2.2) não será infinita, ficando assim:

$$y_t = w_1 + w_2 + \ldots + w_t = \sum_{t=1}^{n} w_t$$

assim é fácil de visualizar que:

$$\mathrm{E}(y_t) = \sum_{t=1}^{n} \mathrm{E}(w_t) = 0$$

e,

$$\mathrm{Var}(y_t) = \sum_{t=1}^{t} \mathrm{Var}(w_t) = t\sigma_w^2$$

Como visto a variância depende de t, portanto o processo é não estacionário, porém a diferença de um passeio aleatório é estacionária, pois:

$$\Delta y_t = y_t - y_{t-1} = w_t \tag{2.3}$$

Os passeios aleatórios são muito utilizados para definir modelos de série temporal

mais complexos, e além de caracterizar diversos tipos de séries não estacionárias.

2.1.2 Funções de Autocovariância e Autocorrelação

Uma função bastante útil é a função de autocovariância (FACV), pois analisa a dependência entre as observações temporais, ajudando a descrever a estrutura de uma série temporal. Cryer e Chan (2008) a define como:

$$\gamma_{(t,t+h)} = \mathrm{Cov}(y_t, y_{t+h}) = \mathrm{E}[(y_t - \mathrm{E}(y_t))(y_{t+h} - \mathrm{E}(y_{t+h}))]$$

A FACV satisfaz as seguintes propriedades:

(i) $\gamma_0 > 0$;

(ii) $\gamma_{(-\tau)} = \gamma_{(\tau)}$;

(iii) $|\gamma_{(\tau)}| \leq \gamma_{(0)}$.

porém, como a autocovariância depende da unidade de medida adotada para medir y_t, é pertinente normalizar-las para que os valores obtidos sejam independentes da unidade medida, para isso divide-se a autocovariância pelo produto dos respectivos desvios padrão, o que da origem á função de autocorrelação (ACF)(HOFFMANN, 2016):

$$\rho(h) = \frac{\mathrm{Cov}(y_t, y_{t+h})}{\sigma_y^2} = \frac{\gamma(h)}{\gamma(0)}$$

e apresenta propriedades similarmente a γ_τ, menos para $\rho(0) = 1$. A função de autocorrelação mede o grau de correlação de uma série, em um dado instante, consigo mesma, em um instante de tempo posterior.

A função de autocorrelação parcial é uma ferramenta que explora o fato de que as autocorrelações parciais são zero além do lag p. Usada, assim, na identificação dos modelos para uma série temporal.

2.2 Algoritmos de Suavização Exponencial

Os algoritmos fazem previsões com base nos comportamentos das observações anteriores da série, porém, a principal dificuldade é identificar e separar esses comportamentos de possíveis ruídos. Um dos principais procedimentos é o algoritmo de alisamento (suavização) exponencial.

Estes algoritmos são classes de algoritmos de previsão, considerados bastante simples, nos quais os dados são ajustados para serem inseridos na série. Morettin e Toloi (2006, p.83) diz que "o proposito dos métodos é distinguir o padrão de qualquer

ruído que possa estar contido nas observações e então usar esse padrão para prever valores futuros da série".

2.2.1 Alisamento Exponencial Simples

O Alisamento Exponencial Simples (AES) é utilizado para séries que não têm tendência e efeitos sazonais no processo, ou caso tenham, foram identificados e removidos, podendo ser descrita como:

$$N_t = a_0 y_t + a_1 y_{t-1} + \ldots, \tag{2.4}$$

em que $N_t = \hat{y}_t$ é a estimativa para y_{t+1}, que é estimado através de uma média ponderada das observações anteriores, de uma série temporal $y_1, \ldots, y_t$ sem tendência e efeitos sazonais, e a_t são os pesos que decrescem exponencialmente à medida que regride,ou seja $a_0 > a_1 > \ldots$, definido por:

$$a_t = \alpha(1-\alpha)^t, \ t = 0, 1, \ldots, \tag{2.5}$$

desta forma, através de (2.4) e (2.5) tem-se:

$$N_t = \alpha y_t + \alpha(1-\alpha) y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} \ldots, \tag{2.6}$$

e ao colocar em evidência $(1-\alpha)$ em (2.6):

$$N_t = \alpha y_t + (1-\alpha)[\alpha y_{t-1} + \alpha(1-\alpha) y_{t-2} \cdots = \alpha y_t + (1-\alpha) N_{t-1}, \tag{2.7}$$

com $0 < \alpha < 1$, que é chamado por Ehlers (2007) de constante de alisamento por ser o valor que determina a suavização, pois quanto menor for α mais estáveis serão as previsões finais. Segundo Morettin e Toloi (2006) o principal problema é a precisão deste valor. Uma maneira de selecioná-lo é através da verificação visual, pois na medida que a série evolui suavemente faz sentido usar um peso alto para α e ao passo que a série evolui de forma errática faz sentido atribuir um valor pequeno à última observação. Definindo o erro $\epsilon_t, t = 1, 2, \ldots, n$, como $\epsilon_t = y_t - N_{t-1}$ substituindo-o na equação (2.7), a correção de erro fica:

$$N_t = N_{t-1} + \alpha \epsilon_t,$$

portanto, se $\epsilon_t = 0$, então $N_t = N_{t-1}$ assim a última previsão foi excelente, não precisando ser modificada. Do contrário a previsão é obtida adicionando à previsão

anterior o múltiplo do erro.

2.2.2 Alisamento Exponencial Holt

Como visto anteriormente o AES é ideal para séries que não apresentam tendência, porém quando aplicado a uma série que apresenta tendência produz previsões que subestimam ou superestimam constantemente os valoreis reais. A fim de impedir esse tipo de erro o algoritmo que pode ser aplicado é o Alisamento Exponencial Holt (AEH) cuja estimativa ocorre da seguinte forma:

$$N_t = \alpha y_t + (1-\alpha)(N_{t-1} + T_{t-1}),\ 0 < \alpha < 1$$

$$T_t = \beta(N_t - N_{t-1}) + (1-\beta)T_{t-1},\ 0 < \beta < 1,$$

em que N_t e $T_t, t = 1, \ldots, n$, são a estimativa de nível e a estimativa da tendência respectivamente, alem de α e β que são as constantes de suavização. A previsão para y_{t+1} é dada por:

$$\hat{y}_t = N_t + hT_t,\ h = 1, 2, \ldots$$

isto é, a previsão é igual ao valor da estimativa de nível mais o da tendencia multiplicada pelo número de passos h, que deseja-se prever. Morettin e Toloi (2006) afirmam que o procedimento de determinação das constantes de suavização é análogo ao do AES, porém ao invés de escolher apenas α que minimize a soma do erros quadráticos, será escolhido o valor do vetor (α, β) para que isto ocorra. E conforme Esquivel (2012), este algoritmo possui uma forma de correção dos erros, descrita assim:

$$N_t = N_{t-1} + T_{t-1} + \alpha\epsilon_t,$$

$$T_t = T_{t-1} + \alpha\beta\epsilon_t.$$

2.2.3 Alisamento Exponencial Holt-Winters

O Alisamento Exponencial Holt-Winters (AEHW) é utilizado em séries com tendência e com padrões sazonais. Existem dois tipos de procedimentos baseados na sazonalidade, a saber com efeito sazonal multiplicativo e aditivo, que são utilizados dependendo das características da série. Esses procedimentos baseiam-se em três equações com constantes de alisamento diferentes, proporcionais aos componentes do padrão da série: nível, tendência e sazonalidade (MORETTIN e TOLOI, 2006).

Considerando série sazonal com período de sazonalidade s, tem-se F_t como fator sazonal com variante que é multiplicativa, enquanto que a tendência continua aditiva (MORETTIN e TOLOI, 2006). A forma de recorrência permitida pelas três equações

de suavização são:

$$N_t = \alpha \frac{y_t}{F_{t-s}} + (1-\alpha)(N_{t-1} + T_{t-1}),\ 0 < \alpha < 1,$$

$$T_t = \beta(N_t - N_{t-1}) + (1-\beta)T_{t-1},\ 0 < \beta < 1,$$

$$F_t = \gamma \frac{y_t}{N_t} + (1-\gamma)F_{t-s},\ 0 < \gamma < 1,$$

em que $t = 1, \ldots, n$, α, β e γ são as constantes de suavização e ainda de acordo com Morettin e Toloi (2006) a previsão é definida por:

$$\hat{y}_t = (N_t + hT_t)F_{t+h-s},\ h = 1, 2, \ldots, s,$$

$$\hat{y}_t = (N_t + hT_t)F_{t+h-2s},\ h = s+1, s+2, \ldots, 2s.$$

$$\vdots$$

Gomes (2003) elenca como forma de correção dos erros para este algoritmo:

$$N_t = N_{t-1} + T_{t-1} + \alpha \frac{\epsilon_t}{F_{t-s}},$$

$$T_t = T_{t-1} + \alpha\beta \frac{\epsilon_t}{F_{t-s}},$$

$$F_t = F_{t-s} + \gamma(1-\alpha) \frac{\epsilon_t}{N_t},$$

em que $t = 1, 2, \ldots, n$.

O procedimento de série sazonal multiplicativo, descrito anteriormente, pode ser modificado para tratar com situações em que o fator sazonal é aditivo (MORETTIN e TOLOI, 2006, p.97), e as estimativas são dadas por:

$$N_t = \alpha(y_t - F_{t-s}) + (1-\alpha)(N_{t-1} + T_{t-1}),\ 0 < \alpha < 1,$$

$$T_t = \beta(N_t - N_{t-1}) + (1-\beta)T_{t-1},\ 0 < \beta < 1,$$

$$F_t = \gamma y_t - N_t + (1-\gamma)F_{t-s},\ 0 < \gamma < 1.$$

As previsões dos valores futuros, h passos à frente, são obtidas partir das equações

$$\hat{y}_t = N_t + hT_t + F_{t+h-s},\ h = 1, 2, \ldots, s,$$

$$\hat{y}_t = N_t + hT_t + F_{t+h-2s},\ h = s+1, s+2, \ldots, 2s.$$

$$\vdots$$

Após operações matemáticas, a forma de correção dos erros é apresentada a seguir:

$$N_t = N_{t-1} + T_{t-1} + \alpha\epsilon_t,$$

$$T_t = T_{t-1} + \alpha\beta\epsilon_t,$$

$$F_t = F_{t-s} + \gamma(1-\alpha)\epsilon_t,$$

em que $t = 1, 2, \ldots, n$.

2.3 Modelagem de Box-Jenkins

Obter modelos apropriados para cada série temporal não é uma tarefa fácil, Cryer e Chan (2008) dizem que para diagnosticar um modelo, a principal preocupação é com a avaliação da qualidade do modelo que é especificado e estimado, verificando o quanto o modelo se ajusta aos dados, e assim o modelo poderá ser usado para prever valores futuros, se não for encontrada nenhuma inadequação, caso contrário, será escolhido outro.

O desenvolvimento de um modelo, segundo BOX *et al.* (2015), passa por uma abordagem iterativa para descrever a estrutura de dependência em uma série temporal observada. De maneira geral, o modelo é melhor obtido por um procedimento iterativo de três estágios, no qual a princípio é escolhido uma classe de modelos com base em critérios específicos a cada tipo de série (identificação), após isso os parâmetros são estimados e ajustados (estimação), e assim avaliados se são adequados (diagnóstico), caso o modelo escolhido não seja o adequado, será reiniciado todo o ciclo.

Entre os principais e mais usados modelos estão os auto-regressivos, de médias moveis, os modelos auto-regressivos integrados e de médias móveis.

2.3.1 Modelos Auto-Regressivos

O modelo auto-regressivo de ordem p denotado por AR(p) é definido por:

$$y_t = c + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \cdots + \Phi_p y_{t-p} + w_t, t = 1, \ldots, n$$

sendo $\Phi_1, \Phi_2, \ldots \Phi_p$, parâmetros auto-regressivos de ordem p e w_t o ruído branco, podendo ser reescrito de outra forma:

$$y_t = c + \Phi_1 B y_t + \Phi_2 B^2 y_t + \cdots + \Phi_p B^p y_t + w_t$$

e colocando y_t em evidência,

$$(1 - \Phi_1 B - \Phi_2 B^2 - \cdots - \Phi_p B^p) y_t = c + w_t,$$

e assim,

$$\Phi(B) y_t = c + w_t, \ t = 1, \ldots, n$$

em que, $\Phi(B) = 1 - \Phi_1 B - \Phi_2 B^2 \cdots - \Phi_p B^p$ é o polinômio operador de defasagens de ordem p.

2.3.2 Modelos de Médias Móveis

O processo de média móvel de ordem q, MA(q) é uma combinação linear do termo de ruído branco recente e os q termos de ruído branco anteriores (COWPERTWAIT e METCALFE, 2009), definido por Morettin e Toloi (2006) como:

$$y_t = \mu + w_t + \Theta_1 w_{t-1} + \cdots + \Theta_q w_{t-q}, t = 1, 2, \ldots, n$$

em que w_t é o ruído branco com média zero e variância σ_w^2. Podendo ser reescrito

$$y_t = \mu + w_t + \Theta_1 B w_t + \cdots + \Theta_q B^q w_t, \ t = 1, 2, \ldots, n.$$

Colocando-se w_t em evidência, tem-se:

$$y_t = \mu + (1 + \Theta_1 B + \cdots + \Theta_q B^q) w_t, \ t = 1, 2, \ldots, n$$

e assim obtêm-se:

$$y_t = \mu + \Theta(B) w_t, \ t = 1, 2, \ldots, n$$

em que $\Theta(B) = \Theta_1 B + \cdots + \Theta_q B^q$ é o polinômio de médias móveis de ordem q.

2.3.3 Modelos Auto-Regressivos de Médias Móveis

Como sugerido pelo nome, este modelo é a junção do auto-regressivo AR(p) com o de médias móveis MA(q), obtendo o modelo misto denotado por ARMA(p, q)

definido da seguinte forma:

$$y_t = c+\Phi_1 y_{t-1}+\Phi_2 y_{t-2}+\cdots+\Phi_p y_{t-p}+w_t+\Theta_1 w_{t-1}+\cdots+\Theta_q w_{t-q}, t = 1, \ldots, n \quad (2.8)$$

em que os $\Phi's$ e $\Theta's$ são os parâmetros auto-regressivos e de médias móveis respectivamente, assim a equação (2.8) pode ser escrita como:

$$\Phi(B)y_t = c + \Theta(B)w_t, \ t = 1, 2, \ldots, n \quad (2.9)$$

cujo $\Phi(B)$ e $\Theta(B)$ são os polinômios de AR(p) e de MA(q), respectivamente.

2.3.4 Modelos Auto-Regressivos Integrados de Médias Móveis

Os modelos AR(p), MA(q) e ARMA(p, q), segundo Morettin e Toloi (2006) são apropriados para descrever series estacionárias. Entretanto, muitas séries encontradas não são estacionárias, como visto anteriormente, devido principalmente a existência de tendência e sazonalidades contidas nas séries. Brockwell e Davis (2016) afirmam que a generalização da classe ARMA(p, q), incorpora uma grande quantidade de séries não estacionárias, quando diferenciadas finitamente muitas vezes, fornece o processos de modelos auto-regressivos integrados de médias móveis - ARIMA(p, d, q), em que p é a ordem do auto-regressivo, d é a ordem da diferença e q é a ordem de médias móveis. Assim sendo, y_t uma série não estacionaria,porém se aplicar d diferenças aos dados, tem-se

$$y_t - y_{t-1} = (1 - B)y_t = \Delta y_t, \ t = 1, 2, \ldots, n.$$

Neste caso tem-se d=1 (1ª ordem), porém generalizando a integração de ordem d

$$(1 - B)^d y_t = \Delta^d y_t,$$

assim a equação (2.9) pode ser escrita

$$\Phi(B)[\Delta^d y_t - c] = \Theta(B)w_t, \ t = 1, 2, \ldots, n$$

Contudo, na maioria dos casos usa-se valores pequenos para d, sendo $d = 1$ ou $d = 2$, que correspondem a dois casos comuns de não estacionaridade homogênea, a saber, séries não estacionárias quanto ao nível e quanto a inclinação, respectivamente. (MORETTIN e TOLOI, 2006)

2.3.5 Modelo Auto-Regressivo Integrado de Médias Móveis Sazonal

Conforme visto anteriormente, uma das características de uma série não estacionaria é o componente sazonal. O modelo SARIMA é formado pela inclusão de termos sazonais nos modelos ARIMA(p,d,q).

A maneira propícia de diferenciar dados com padrão sazonal, é tomando as diferenças no período sazonal $\nabla^d = (1-B)^d$, (s é o período sazonal), ou seja, para dados mensais a primeira diferença sazonal e dada por:

$$\nabla_{12} y_t = (1-B)^{12} y_t = y_t - y_{t-12},$$

tendo uma variabilidade menor do que a primeira diferença não sazonal da equação (2.3), ficando mais fácil de identificar e estimar. A D-ésima diferença sazonal é então denotada por $\nabla_s^D = (1-B^s)^D$, em que D indica o número dessa diferenças, (EHLERS, 2007).

O modelo auto-regressivo integrado de médias móveis sazonal é denotado por SARIMA $(p,d,q) \times (P,D,Q)$, em que as letras maiúsculas representam a parte sazonal do modelo enquanto letras minúsculas representam a parte não sazonal do modelo. Porém, para obtenção "d"e "D"deve-se utilizar conjuntamente, ou seja, $\nabla^d \nabla_s^D y_t$ de forma a tornar estacionária, e a obtenção das ordens "p", "P", "q"e "Q"é feita como nos modelos não sazonais (FISCHER, 1982).

Segundo Morettin (2011) essa classe de modelo é dada por:

$$\Phi(B)\Phi(B^s)\nabla^d \nabla_s^D y_t = \Theta(B)\Theta(B^s) w_t$$

em que,
$\Phi(B) = 1 - \Phi_1 B - \cdots - \Phi_p B^p$, é o operador auto-regressivo estacionário de ordem p;
$\Phi(B^s) = 1 - \Phi_1 B^s - \cdots - \Phi_P B^{sP}$, é o operador auto-regressivo sazonal de ordem P, estacionário;
$\Theta(B) = 1 - \Theta_1 B - \cdots - \Theta_q B^q$, é o operador de médias móveis invertível de ordem q;
$\Theta(B^s) = 1 - \Theta_1 B^s - \cdots - \Theta_Q B^{sQ}$, é o operador de médias móveis sazonal de ordem Q, invertível.

2.3.6 Modelo Auto-Regressivo Integrado de Médias Móveis Sazonal com Variáveis Exógenas

A letra X acrescida em SARIMAX representa as variáveis exógenas, estendendo o modelo SARIMA ao combiná-lo com uma estrutura de regressão, que pode ser a tendência, sazonalidade ou uma variável dummy, entre outras.

SARIMAX é um modelo de regressão linear que utiliza um processo do modelo SARIMA $(p, d, q) \times (P, D, Q)$. O modelo SARIMAX requer não apenas os argumentos p, d e q, mas também requer outro argumento que é a periodicidade do ciclo sazonal dos argumentos, em que a modelagem é dada por:

$$\Phi(B)\Phi(B^s)(1-B)^d(1-B)^D y_t = X\beta + \Theta(B)\Theta(B^s)w_t$$

Capítulo 3

Análise da Série

Neste capítulo são analisados e explicitados os métodos e funções empregadas para obtenção das previsões dos valores futuros do ICMS da Paraíba, onde será utilizado o software `R` para obter os resultados. As previsões serão feitas através do algoritmo de suavização exponencial, mais precisamente Holt-Winters aditivo e multiplicativo. Também será utilizada a modelagem Box-Jenkins, com o objetivo de compará-las e assim determinar qual o melhor método de previsão dos valores futuros da arrecadação do ICMS da Paraíba. A base de dados para análise do ICMS e do IGP-DI, referem-se ao período de janeiro de 1997 a abril de 2021, obtidos no Confaz (https://www.confaz.fazenda.gov.br/boletim-de-arrecadacao-dos-tributos-estaduais) e no FGV-IBE (http://www14.fgv.br/fgvdados20/visualizaconsulta.aspx), respectivamente.

Inicialmente deve-se analisar os dados obtidos a fim de verificar se há presença de sazonalidade, tendência e estacionariedade. Existem vários testes que podem auxiliar nesta etapa, a exemplo do teste de Cox-Stuart (COX e STUART, 1955) e o teste de Wald-Wolfowitz (WALD e WOLFOWITZ, 1948) utilizados para verificar a existência de tendência e os testes de Kruskal-Wallis (KRUSKAL e WALLIS, 1952) e o de Friedman (FRIEDMAN, 1937) para a sazonalidade. Porém um importante instrumento de analise é o gráfico temporal da série, em que na Figura 3.1 nota-se indícios de tendência, não-estacionariedade e sazonalidade. Vale salientar que a série encontra-se em milhões de reais e foi deflacionada usando o IGP-DI a preço constante de abril de 2021.

Para não haver duvidas sobre a existência de tendência, realizamos o teste de Cox-Stuart que consiste em agrupar os dados em pares $(y_{1+c}, y_1), (y_{2+c}, y_2), \ldots, (y_{N+c}, y_N)$, tal que N é o número de observações da série e c=$\frac{N}{2}$ para N par, ou c=$\frac{N+1}{2}$ caso N seja ímpar, e calculando as diferenças entre esses pares, atribuindo $-$ para valores negativos e $+$ para os positivos, anulando entre se. Assim, Temos como hipótese nula, $H_0 : P(y_i < y_{i+c}) = P(y_i > y_{i+c})$:

não existe tendência nas observações, e $H_1 : P(y_i < y_{i+c}) \neq P(y_i > y_{i+c})$: existe tendência nas observações. Verificando os resultados ao nível de significância de 5%, com p-valor $< 2,2x10^{-}16$, a hipótese nula deve ser rejeitada, apontando a presença de tendência.

Sobre a existência de sazonalidade, verificamos através da realização do teste de Kruskal-Wallis que tem como hipótese nula, o H_0: não existe sazonalidade e a hipótese alternativa, o H_1: existe sazonalidade. Obtemos o p-valor $= 0.489 > 0,05$. Então ao nível de 5% de significância, rejeitamos H_0. Logo a série possui sazonalidade.

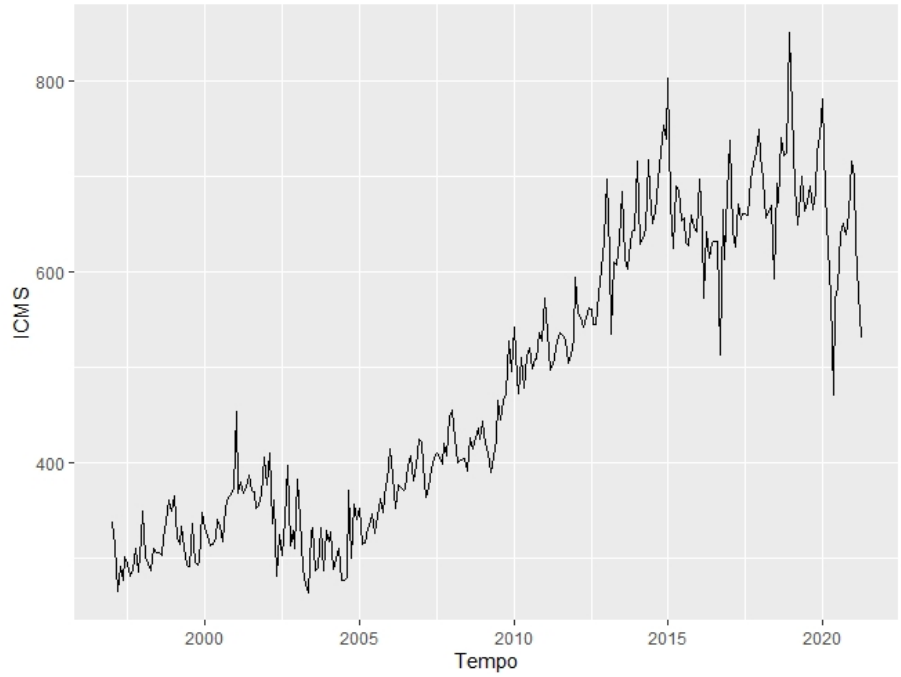

Figura 3.1: Valores do ICMS-PB deflacionado expressos em milhões de reais.

Essas características da série podem ser vistas mais claramente na Figura 3.2. A partir desta figura pode-se verificar a existência de tendência e sazonalidade. O segundo gráfico indica uma tendência crescente, e o terceiro gráfico evidencia a ocorrência do principal pico no primeiro mês de cada ano, já o quarto gráfico mostra que entre os anos de 2005 e 2010 os resíduos são mais próximos de zero, e um valor que se destaca é o resíduo de maio de 2020.

A série tem como valor médio de arrecadações, durante o período analisado, de R$ 482,1 milhões, tendo a menor arrecadação em maio de 2003, R$ 263,7 milhões e em dezembro de 2018 o Estado obteve sua maior arrecadação, aproximadamente R$ 851 milhões, por fim o desvio padrão da série que é de aproximadamente R$ 151,78 milhões.

A Figura 3.3 mostra a função de autocorrelação da série. É possível verificar um

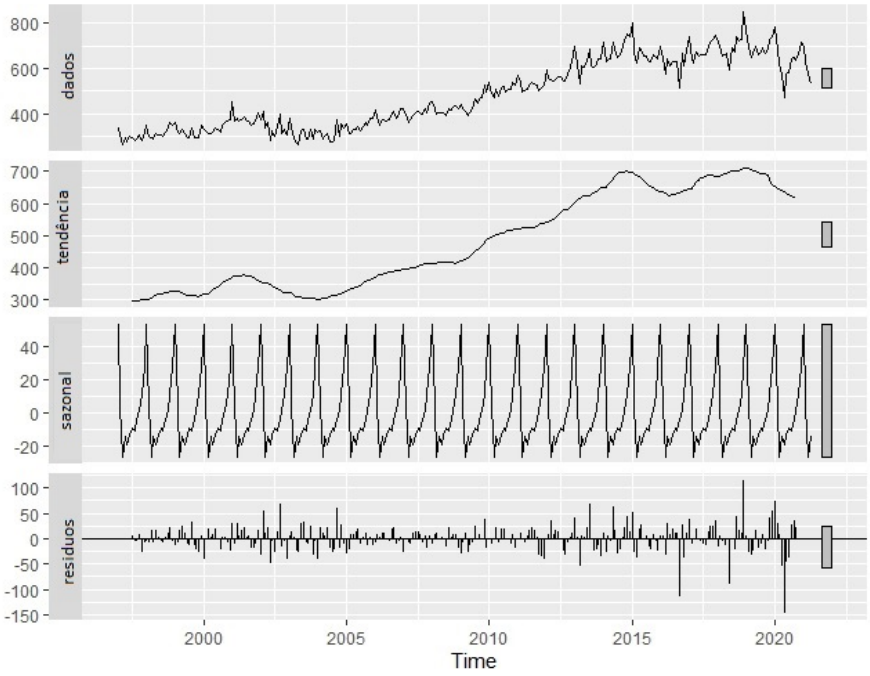

Figura 3.2: Decomposição da Série.

decaimento lento da autocorrelação não havendo um ponto q, ordem de um processo MA(q), que a partir dele as autocorrelações amostrais permanecerão no intervalo assintótico $\left\{\frac{-2}{\sqrt{N}}, \frac{2}{\sqrt{N}}\right\}$, sendo N a quantidade de observações, revelando uma não-estacionariedade da série, que pode ser comprovada através da função `adf.test` do `R`. A Figura 3.4 mostra a PAC, onde é possível observar não haver um padrão, não sendo possível obter as ordens dos modelos.

3.1 Modelagem, Previsões e Resultados

Nesta seção são analisados alguns modelos que podem representar adequadamente a série do ICMS-PB, com os quais são feitas previsões. Serão utilizadas duas medidas de precisão, a saber:

Erro Quadrático Médio - EQM dado por,

$$\frac{1}{N}\sum_{t=1}^{N}(\epsilon_t)^2 \ ;$$

Erro Total - ET definido por,

$$\sum_{t=1}^{N}(\epsilon_t),$$

em que $\epsilon_t = y_t - \hat{y}_t$.

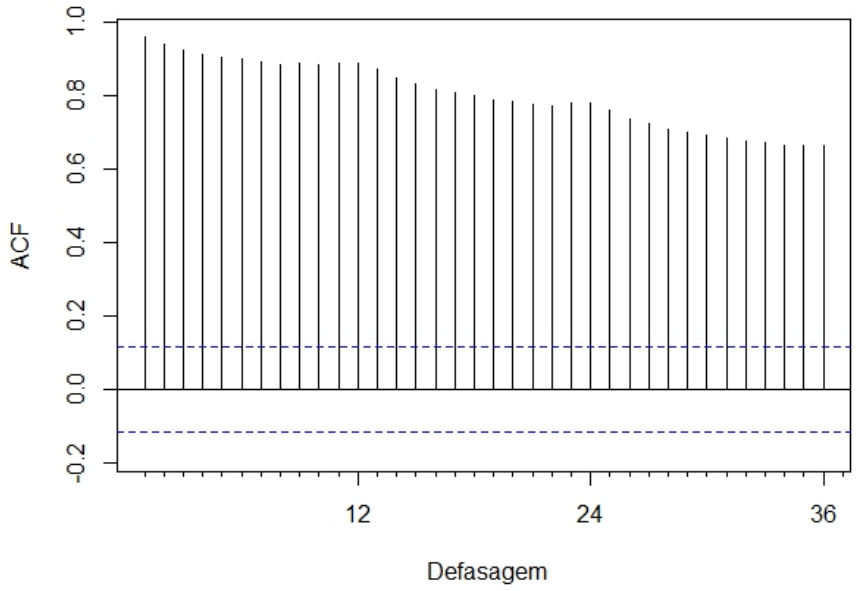

Figura 3.3: Função de autocorrelação da série do ICMS-PB.

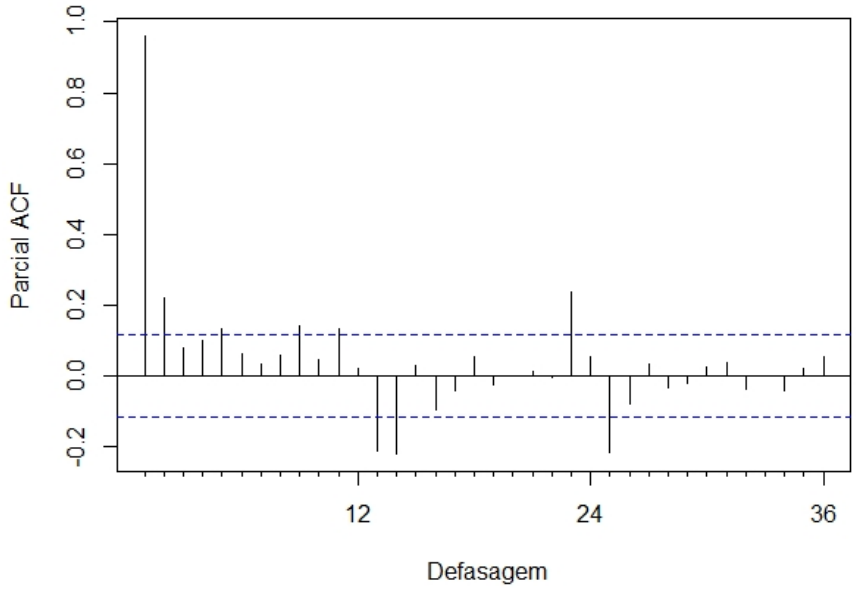

Figura 3.4: Função de autocorrelação parcial da Série do ICMS-PB.

Para avaliar melhor as previsões da série em diferentes instantes de tempo, ela será particionada em datas distintas, o que pode colaborar para verificar se a dinâmica da série varia. Desta forma serão modeladas a série do ICMS-PB contendo valores do período de janeiro de 1997 à:

- abril de 2019, denotada ICMSPB-abr.19;
- outubro de 2019, denotada ICMSPB-out.19;
- abril de 2020, denotada ICMSPB-abr.20;
- outubro de 2020, denotada ICMSPB-out.20;

possuindo, respectivamente, 268, 274, 280 e 286 observações, em que serão considerados para cada partição as previsões 1, 3, 6, e 12 passos à frente.

3.1.1 Alisamento Exponencial de Holt-Winters

O método utilizado nesta seção é o da classe de algoritmos de alisamento exponencial, mais precisamente o alisamento exponencial de Holt-Winters (Seção 2.2.3), devido a presença de sazonalidade, conforme verificada nas Figuras 3.1 e 3.2. Entretanto como não se sabe se a natureza do efeito sazonal, se o mesmo é aditivo ou multiplicativo, serão empregados os dois algoritmos, com a finalidade de comparar qual previsão (sazonalidade aditiva ou multiplicativa) melhor representa os dados.

Tabela 3.1: Previsões obtidas para a série ICMSPB-abr.19 pelos algoritmos de Holt-Winters aditivo e multiplicativo 12 passos à frente.

Mês	Arrecadação Real	HW aditivo	HW multiplicativo
Mai 2019	700.39	692.29	686.34
Jun 2019	663.79	684.02	679.75
Jul 2019	673.21	706.81	700.32
Ago 2019	689.84	694.03	685.99
Set 2019	665.61	701.06	692.99
Out 2019	684.81	723.15	717.27
Nov 2019	726.15	724.66	720.78
Dez 2019	741.68	755.31	752.94
Jan 2020	781.26	765.84	778.94
Fev 2020	681.21	704.35	711.16
Mar 2020	624.67	674.08	672.97
Abr 2020	573.19	705.45	703.04

A Tabela 3.1 mostra as previsões obtidas para a série ICMSPB-abr.19 pelos algoritmos de Holt-Winters com sazonalidade aditiva e multiplicativa 12 passos à frente, na qual verifica-se que a maioria das previsões superestimaram as arrecadações da série, exceto no mês de agosto para HW multiplicativo e nos meses de maio, novembro e janeiro, para os dois modelos.

As previsões obtidas para a série ICMSPB-out.19 através dos algoritmos de Holt-Winters aditivo e multiplicativo 12 passos à frente, estão contidas na Tabela 3.2. É possível notar que grande parte das previsões geradas pelos métodos sazonal aditivo

Tabela 3.2: Previsões obtidas para a série ICMSPB-out.19 pelos algoritmos de Holt-Winters aditivo e multiplicativo 12 passos à frente.

Mês	Arrecadação Real	HW aditivo	HW multiplicativo
Nov 2019	726.15	698.07	700.62
Dez 2019	741.68	728.35	732.00
Jan 2020	781.26	739.17	757.07
Fev 2020	681.21	677.87	691.31
Mar 2020	624.67	647.80	654.28
Abr 2020	573.19	679.13	683.67
Mai 2020	470.95	687.55	688.74
Jun 2020	572.74	674.78	677.36
Jul 2020	583.31	696.58	697.49
Ago 2020	640.53	689.70	688.44
Set 2020	650.21	691.45	690.44
Out 2020	639.31	714.28	715.17

e multiplicativo superestimaram os valores da série, menos as previsões advindas dos meses de novembro, dezembro de 2019 e janeiro de 2020, além do mês de fevereiro para HW multiplicativo. É importante destacar que as previsões referentes ao mês de maio, apresentam valores muito distantes dos valores reais arrecadados.

Tabela 3.3: Previsões obtidas para a série ICMSPB-abr.20 pelos algoritmos de Holt-Winters aditivo e multiplicativo 12 passos à frente.

Mês	Arrecadação Real	HW aditivo	HW multiplicativo
Mai 2020	470.95	652.12	645.14
Jun 2020	572.74	638.07	635.46
Jul 2020	583.31	661.03	653.67
Ago 2020	640.53	654.37	645.19
Set 2020	650.21	655.75	647.73
Out 2020	639.31	679.07	670.62
Nov 2020	661.55	687.87	679.46
Dez 2020	715.74	715.76	706.30
Jan 2021	699.83	729.53	733.68
Fev 2021	628.81	659.58	665.58
Mar 2021	578.17	625.67	628.38
Abr 2021	530.87	646.14	647.02

A Tabela 3.3 exibe as previsões obtidas para a série ICMSPB-abr.20 através dos algoritmos de Holt-Winters com sazonalidade aditiva e multiplicativa, na qual as previsões obtidas superestimaram os valores reais de arredações, as exceções são nos meses de setembro e dezembro de 2020 no algoritmo HW com sazonalidade multiplicativa.

A Tabela 3.4 mostra as previsões originadas através dos algoritmos de Holt-Winters com sazonalidade aditiva e multiplicativa com 6 passos à frente para a

Tabela 3.4: Previsões obtidas para a série ICMSPB-out.20 pelos algoritmos de Holt-Winters aditivo e multiplicativo 6 passos à frente.

Mês	Arrecadação Real	HW aditivo	HW multiplicativo
Nov 2020	661.55	651.28	649.09
Dez 2020	715.74	677.27	674.12
Jan 2021	699.83	691.95	704.04
Fev 2021	628.81	623.69	639.92
Mar 2021	578.17	591.42	604.85
Abr 2021	530.87	613.17	622.39

série ICMSPB-out.20, na qual pode-se verificar que metade das previsões obtidas, subestimaram as arrecadações reais.

3.1.2 Modelagem de Box-Jenkins-SARIMA

Nesta etapa é utilizada a modelagem de Box-Jenkins e em virtude das característica da série principalmente a sazonalidade o modelo selecionado é o SARIMA. Utilizando as mesmas séries vistas anteriormente no modelo Holt-Winters, e através da função `auto.arima` da biblioteca `forecast` do `R`, encontramos o modelo mais adequado para cada série, através de critérios para seleção de modelos baseados na função de verossimilhança. Assim são escolhidos dois critérios para seleção desses modelos, o critério de Informação de Akaike (AIC)(AKAIKE, 1978) e o critério Bayesiano de Schwarz (BIC)(AKAIKE, 1978). Apos os ajustes dos modelos é utilizada a função `tsdiag` para traçar um diagnóstico de correlação dos resíduos através dos gráficos do ACF dos resíduos, da estacionaridade do resíduo e do p-valor para o teste de Ljung-Box (LJUNG e BOX, 1978). Todos os modelos selecionados, para cada série, foram bem ajustados como pode ser visto no Apêndice A.

Os modelos obtidos pelos menores valores dos critérios AIC e BIC, para a série ICMSPB-abr.19 são SARIMA$(1,1,1)\times(0,0,2)$ e SARIMA$(0,1,1)\times(0,0,2)$, respectivamente. Os valores das previsões são observado na Tabela 3.5 na qual verifica-se que ambos os modelos superestimaram os valores arrecadados, exceto nos meses de maio, novembro de 2019 e janeiro de 2020 para os dois modelos, além do mês de agosto de 2019 para o modelo SARIMA$(0,1,1)\times(0,0,2)$.

Para a série ICMSPB-out.19 os modelos selecionados foram SARIMA$(1,1,1)\times(0,0,2)$ através do menor valor do AIC e SARIMA$(0,1,1)\times(0,0,2)$ com o menor valor do BIC. As previsões obtidas por esses modelos estão apresentadas na Tabela 3.6, na qual observa-se que os valores das previsões nos dois modelos são aproximados, com os obtidos pelo SARIMA$(0,1,1)\times(0,0,2)$ menores em todas as observações, ficando mais próximas das arrecadações reais.

A Tabela 3.7 exibe as previsões 12 passos à frente obtidas por meio dos mo-

Tabela 3.5: Previsões obtidas para a série ICMSPB-abr.19 pelo modelo SARIMA$(1,1,1) \times (0,0,2)_{AIC}$ e SARIMA$(0,1,1) \times (0,0,2)_{BIC}$ 12 passos à frente.

Mês	Arrecadação Real	SARIMA $(1,1,1) \times (0,0,2)$	SARIMA $(0,1,1) \times (0,0,2)$
Mai 2019	700.39	689.98	686.46
Jun 2019	663.79	675.75	667.55
Jul 2019	673.21	697.56	689.83
Ago 2019	689.84	693.99	686.08
Set 2019	665.61	732.94	727.82
Out 2019	684.81	714.52	707.51
Nov 2019	726.15	725.38	719.25
Dez 2019	741.68	750.52	745.24
Jan 2020	781.26	719.74	712.69
Fev 2020	681.21	704.14	696.53
Mar 2020	624.67	689.33	681.03
Abr 2020	573.19	688.09	679.59

Tabela 3.6: Previsões obtidas para a série ICMSPB-out.19 pelo modelo SARIMA$(1,1,1) \times (0,0,2)_{AIC}$ e SARIMA$(0,1,1) \times (0,0,2)_{BIC}$ 12 passos à frente.

Mês	Arrecadação Real	SARIMA $(1,1,1) \times (0,0,2)$	SARIMA $(0,1,1) \times (0,0,2)$
Nov 2019	726.15	695.85	692.70
Dez 2019	741.68	719.75	717.12
Jan 2020	781.26	691.50	687.27
Fev 2020	681.21	676.83	672.11
Mar 2020	624.67	663.03	657.74
Abr 2020	573.19	661.99	656.44
Mai 2020	470.95	673.81	668.94
Jun 2020	572.74	647.07	640.74
Jul 2020	583.31	670.37	665.06
Ago 2020	640.53	669.08	663.74
Set 2020	650.21	679.96	675.03
Out 2020	639.31	673.46	668.07

delos SARIMA$(1,1,1) \times (0,0,2)$ e SARIMA$(0,1,1) \times (0,0,2)$, considerados pelo AIC e BIC, respectivamente, para a série ICMSPB-abr.20. Na qual mostra que as previsões obtidas pelo SARIMA$(1,1,1) \times (0,0,2)$ em sua maioria superestimam os valores reais de arrecadação, entretanto a maior parte dos valores obtidos no modelo SARIMA$(0,1,1) \times (0,0,2)$ subestimaram as arrecadações reais.

A Tabela 3.8 contém as previsões obtidas para a série ICMSPB-out.20 geradas pelos modelos SARIMA$(1,1,1) \times (0,0,2)$ e SARIMA$(0,1,1) \times (0,0,2)$ 6 passos à frente. Verifica-se que as previsões superestimaram os valores da série nos meses de março e abril de 2021.

Tabela 3.7: Previsões obtidas para a série ICMSPB-abr.20 pelo modelo SARIMA$(1,1,1) \times (0,0,2)_{AIC}$ e SARIMA$(0,1,1) \times (0,0,2)_{BIC}$ 12 passos à frente.

Mês	Arrecadação Real	SARIMA $(1,1,1) \times (0,0,2)$	SARIMA $(0,1,1) \times (0,0,2)$
Mai 2020	470.95	634.85	636.29
Jun 2020	572.74	616.67	607.19
Jul 2020	583.31	641.61	631.46
Ago 2020	640.53	641.04	630.65
Set 2020	650.21	650.10	639.27
Out 2020	639.31	644.69	633.85
Nov 2020	661.55	654.58	644.21
Dez 2020	715.74	680.10	670.79
Jan 2021	699.83	674.78	666.08
Fev 2021	628.81	638.87	627.94
Mar 2021	578.17	619.43	607.20
Abr 2021	530.87	611.38	598.66

Tabela 3.8: Previsões obtidas para a série ICMSPB-out.20 pelo modelo SARIMA$(1,1,1) \times (0,0,2)_{AIC}$ e SARIMA$(0,1,1) \times (0,0,2)_{BIC}$ 6 passos à frente.

Mês	Arrecadação Real	SARIMA $(1,1,1) \times (0,0,2)$	SARIMA $(0,1,1) \times (0,0,2)$
Nov 2020	661.55	639.65	642.12
Dez 2020	715.74	661.51	668.94
Jan 2021	699.83	655.25	661.96
Fev 2021	628.81	621.59	625.90
Mar 2021	578.17	603.55	606.08
Abr 2021	530.87	595.99	598.60

3.1.3 Modelagem de Box-Jenkins com variáveis exógenas

Nesta etapa são obtidas as previsões das séries descritas anteriormente, através dos modelos ARIMAX, quando utilizada a sazonalidade como variável exógena, e SARIMAX, quando temos como variáveis exógenas a tendência, e uma variável dummy. Obtidos a parti dos critérios AIC e BIC.

Entretanto os modelos obtidos com a tendência como variável exógenas não estão bem ajustados, de acordo o teste de ljung-box e ACF dos resíduos, como mostra os gráficos no Apêndice A, assim não realizamos as previsões para estes modelos. Os modelos com a variável dummy foram selecionados afim verificar se as observações encontram-se com algum tipo de alteração durante os meses referentes a pandemia que iniciou-se em mar/20 até nossa ultima observação, abr/21. Assim foram utilizados apenas para as séries ICMSPB-abr.20 e ICMSPB-out.20, as quais estas observações estão presentes.

O modelo obtido, através dos critérios AIC e BIC, para a série ICMSPB-abr.19,

usando a sazonalidade como variável de regressão é o ARIMAX$(0,1,1)$. Os valores das previsões são observados na Tabela 3.9, na qual verifica-se que as previsões subestimaram valores arrecadados nos meses de julho, agosto, outubro, novembro, dezembro de 2019 e janeiro de 2020.

Tabela 3.9: Previsões obtidas para a série ICMSPB-abr.19 pelo modelo ARIMAX$(0,1,1)_{AIC,BIC}$ 12 passos à frente.

Mês	Arrecadação Real	ARIMAX$(0,1,1)$
Mai 2019	700.39	740.95
Jun 2019	663.79	696.76
Jul 2019	673.21	667.33
Ago 2019	689.84	676.89
Set 2019	665.61	674.83
Out 2019	684.81	680.50
Nov 2019	726.15	684.22
Dez 2019	741.68	682.71
Jan 2020	781.26	688.69
Fev 2020	681.21	694.45
Mar 2020	624.67	703.96
Abr 2020	573.19	720.15

Para a série ICMSPB-out.19 o modelo selecionado usando a sazonalidade como variável exógena é o ARIMAX$(0,1,1)$ considerando os dois critérios, AIC e BIC. As previsões obtidas por esse modelo estão apresentadas na Tabela 3.6, na qual observa-se que os valores das previsões, subestimaram os valores reais nos meses de dezembro, janeiro e fevereiro.

Tabela 3.10: Previsões obtidas para a série ICMSPB-out.19 pelo modelo ARIMAX$(0,1,1)_{AIC,BIC}$ 12 passos à frente.

Mês	Arrecadação Real	ARIMAX$(0,1,1)$
Nov 2019	726.15	731.41
Dez 2019	741.68	687.39
Jan 2020	781.26	658.08
Fev 2020	681.21	667.60
Mar 2020	624.67	665.55
Abr 2020	573.19	671.19
Mai 2020	470.95	674.91
Jun 2020	572.74	673.39
Jul 2020	583.31	679.35
Ago 2020	640.53	685.09
Set 2020	650.21	694.56
Out 2020	639.31	710.68

A Tabela 3.11 apresenta as previsões 12 passos à frente para a série ICMSPB-abr.20, obtidas por meio do modelo ARIMAX$(0,1,1)$, com a sazonalidade como va-

riável exógena, considerando os dois critérios, e com a variável dummy pandemia, os modelos selecionados são SARIMAX$(1,1,1)\times(0,0,2)$ e SARIMAX$(0,1,1)\times(0,0,2)$ usando os critérios AIC e BIC, respectivamente. É possível verificar que todas as previsões obtidas pelo modelo com a dummy como variável regressiva superestimaram os valores reais arrecadados, porém as previsões do modelo ARIMAX referentes os meses de agosto a janeiro subestimaram as arrecadações reais entre todas as previsões.

Tabela 3.11: Previsões obtidas para a série ICMSPB-abr.20 12 passos à frente.

Mês	Arrecadação Real	ARIMAX $(0,1,1)$	SARIMAX $(1,1,1)\times(0,0,2)$	SARIMAX $(0,1,1)\times(0,0,2)$
Mai 2020	470.95	693.28	688.93	695.17
Jun 2020	572.74	646.81	664.91	667.00
Jul 2020	583.31	615.87	688.54	691.13
Ago 2020	640.53	625.92	687.42	689.94
Set 2020	650.21	623.75	697.77	700.51
Out 2020	639.31	629.71	691.59	694.00
Nov 2020	661.55	633.63	701.14	703.88
Dez 2020	715.74	632.03	726.63	730.23
Jan 2021	699.83	638.32	719.99	723.62
Fev 2021	628.81	644.38	685.60	687.87
Mar 2021	578.17	654.38	684.46	687.73
Abr 2021	530.87	671.40	677.11	680.17

A Tabela 3.12 contém as previsões obtidas para a série ICMSPB-out.20 geradas pelos modelos SARIMAX$(1,1,1)\times(0,0,2)$ e SARIMAX$(0,1,1)\times(0,0,2)$ com a variável dummy. considerando os critérios AIC e BIC, respectivamente, e o modelo ARIMAX$(0,1,1)$ com a sazonalidade como variável regressiva para ambos os critérios. Verifica-se que as previsões superestimaram os valores da série exceto no modelo ARIMAX$(0,1,1)$ nos meses de dezembro, janeiro e fevereiro.

Tabela 3.12: Previsões obtidas para a série ICMSPB-out.20 6 passos à frente.

Mês	Arrecadação Real	ARIMAX $(0,1,1)$	SARIMAX $(1,1,1)\times(0,0,2)$	SARIMAX $(0,1,1)\times(0,0,2)$
Nov 2020	661.55	686.55	731.03	745.41
Dez 2020	715.74	639.23	752.75	771.82
Jan 2021	699.83	607.71	745.07	763.36
Fev 2021	628.81	617.95	712.96	729.59
Mar 2021	578.17	615.74	714.46	732.50
Abr 2021	530.87	621.81	707.74	725.87

3.1.4 Medidas de Precisão

Com base nas previsões obtidas esta etapa apresenta a seleção dos modelos para cada série, utilizando como critério os modelos que têm os menores valores absolutos nas medidas de precisão, de maneira geral, obtidos através das previsões, ou seja, os menores EQM e ET relativos aos valores das previsões.

Tabela 3.13: Medidas de precisão obtidas para a série ICMSPB-abr.19.

	passos à frente			
Medidas de precisão - EQM	1	3	6	12
HW Aditivo	65.55	534.49	724.61	2103.55
HW Multiplicativo	197.37	395.53	500.88	1938.02
SARIMA$(1,1,1)\times(0,0,2)_{AIC}$	108.40	281.41	1046.28	2337.45
SARIMA$(0,1,1)\times(0,0,2)_{BIC}$	193.88	161.36	813.89	2031.27
ARIMAX$_{Saz}(0,1,1)_{AIC,BIC}$	1645.95	922.54	506.45	3741.77
	passos à frente			
Medidas de precisão - ET	1	3	6	12
HW Aditivo	8.09	−45.73	−123.71	−325.23
HW Multiplicativo	14.04	−29.01	−85.01	−296.67
SARIMA$(1,1,1)\times(0,0,2)_{AIC}$	10.41	−25.89	−127.08	−276.12
SARIMA$(0,1,1)\times(0,0,2)_{BIC}$	13.92	−6.45	−87.60	−193.75
ARIMAX$_{Saz}(0,1,1)_{AIC,BIC}$	−40.57	−67.66	−59.63	−105.66

Desta forma, tem-se como modelo selecionado para a série ICMSPB-abr.19, o modelo SARIMA$(0,1,1)\times(0,0,2)$ é melhor em 3 passos à frente nos erros EQM e ET, e o segundo melhor para 12 passos à frente com o menor erro total, como exibido na Tabela 3.13. As doze arrecadações reais estão entre o limite superior e inferior do intervalo de confiança de 95% para a previsão para esta série, como mostra a Figura 3.5.

A Figura 3.6 apresenta o gráfico da série ICMSPB-out.19 com as previsões e os limites de confiança de 95% obtido através do algoritmo HW com efeito sazonal multiplicativo, pois nas medidas de precisão, presentes na Tabela3.14, para 3 passos à frente é o mais adequado e para 1 passo é o segundo com menores valores, em ambas medidas. No entanto cinco arrecadações, abril, maio, junho, julho e outubro, estão fora dos limite de 95% de confiança das previsões obtidas para esta série.

Para série ICMSPB-abr.20 o modelo que obteve os menores valores nas medidas de precisão, como verificado na Tabela3.15, é o modelo SARIMA$(0,1,1)\times(0,0,2)$ obtido pelo critério BIC, apresentando os menores erros em 3, 6 e 12 passos a frente. Os valores reais das arrecadações permaneceram dentro dos limites do intervalo com 95% de confiança, a exceção é apenas o mês de maio de 2020, presente no gráfico da Figura 3.7.

O algoritmo HW com sazonalidade aditiva é o que obteve as melhores medidas de precisão para série ICMSPB-out.20, podendo ser verificado na Tabela 3.16. O

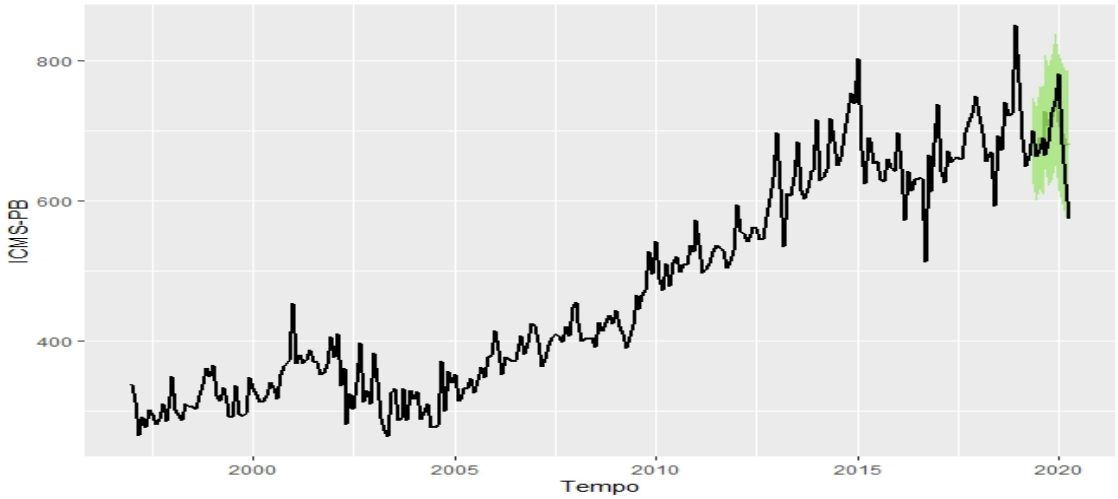

Figura 3.5: Gráfico das previsões para série ICMSPB-abr.19 com os limites superior e inferior com 95% de confiança, obtidos pelo modelo SARIMA$(0,1,1)\times(0,0,2)_{BIC}$

Tabela 3.14: Medidas de precisão obtidas para a série ICMSPB-out.19.

	passos à frente			
Medidas de precisão - EQM	1	3	6	12
HW Aditivo	788.33	912.59	2417.75	7043.79
HW Multiplicativo	651.94	443.58	2419.20	7966.83
SARIMA$(1,1,1)\times(0,0,2)_{AIC}$	918.42	3152.14	3138.50	6329.61
SARIMA$(0,1,1)\times(0,0,2)_{BIC}$	1118.71	3518.91	3110.53	5929.50
ARIMAX$_{Saz}(0,1,1)_{AIC,BIC}$	27.64	6048.4	4934.41	8301.16
	passos à frente			
Medidas de precisão - ET	1	3	6	12
HW Aditivo	28.07	83.49	−42.23	−639.54
HW Multiplicativo	25.53	59.39	−90.78	−691.38
SARIMA$(1,1,1)\times(0,0,2)_{AIC}$	30.30	141.99	19.22	−437.47
SARIMA$(0,1,1)\times(0,0,2)_{BIC}$	33.45	151.99	44.79	−379.74
ARIMAX$_{Saz}(0,1,1)_{AIC,BIC}$	−5.25	172.19	46.91	−514.06

gráfico da Figura 3.8 exibe que os valores reais arrecadados estão compreendidos no intervalo de confiança 95% das previsões exceto o mês de abril de 2021.

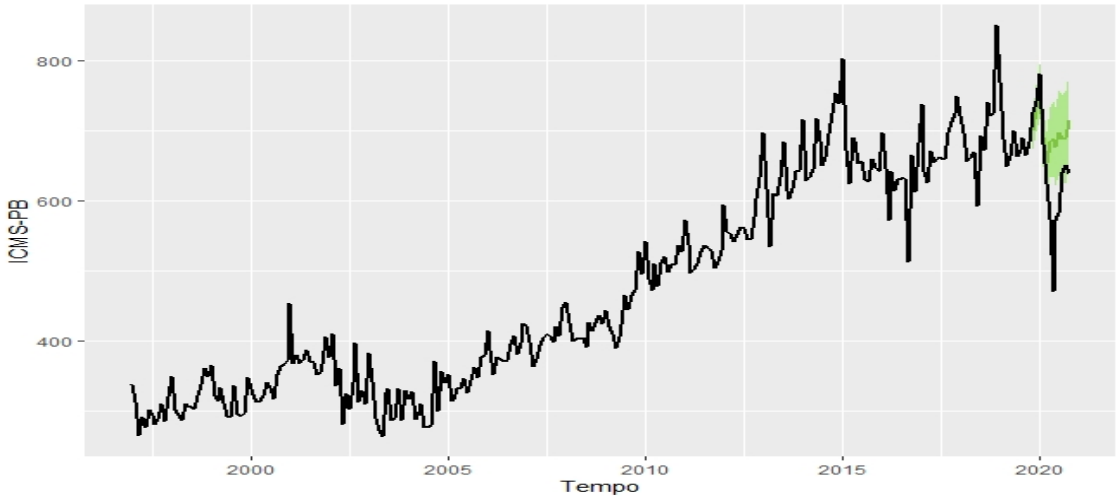

Figura 3.6: Gráfico das previsões para série ICMSPB-out.19 com os limites superior e inferior com 95% de confiança, obtidos pelo HW com sazonalidade multiplicativa.

Tabela 3.15: Medidas de precisão obtidas para a série ICMSPB-abr.20.

	passos à frente			
Medidas de precisão - EQM	1	3	6	12
HW Aditivo	32825.11	14378.50	7489.76	5250.30
HW Multiplicativo	30343.13	13076.40	6706.17	4929.69
SARIMA$(1,1,1)\times(0,0,2)_{AIC}$	26863.18	10730.76	5370.24	3537.80
SARIMA$(0,1,1)\times(0,0,2)_{BIC}$	27338.44	10281.23	5181.80	3332.55
ARIMAX$_{Saz}(0,1,1)_{AIC,BIC}$	49432.01	18660.27	9497.60	7862.86
SARIMAX$_{pand}(1,1,1)\times(0,0,2)_{AIC}$	47520.08	22363.98	12381.24	9357.84
SARIMAX$_{pand}(0,1,1)\times(0,0,2)_{BIC}$	50277.16	23596.89	13125.80	9925.93
	passos à frente			
Medidas de precisão - ET	1	3	6	12
HW Aditivo	−181.17	−324.24	−383.38	−632.96
HW Multiplicativo	−174.19	−307.28	−340.78	−586.22
SARIMA$(1,1,1)\times(0,0,2)_{AIC}$	−163.89	−266.13	−271.92	−336.09
SARIMA$(0,1,1)\times(0,0,2)_{BIC}$	−165.34	−247.94	−221.66	−221.56
ARIMAX$_{Saz}(0,1,1)_{AIC,BIC}$	−222.33	−328.98	−278.33	−337.54
SARIMAX$_{pand}(1,1,1)\times(0,0,2)_{AIC}$	−217.99	−415.40	−562.15	−942.14
SARIMAX$_{pand}(0,1,1)\times(0,0,2)_{BIC}$	−224.22	−426.32	−580.73	−979.31

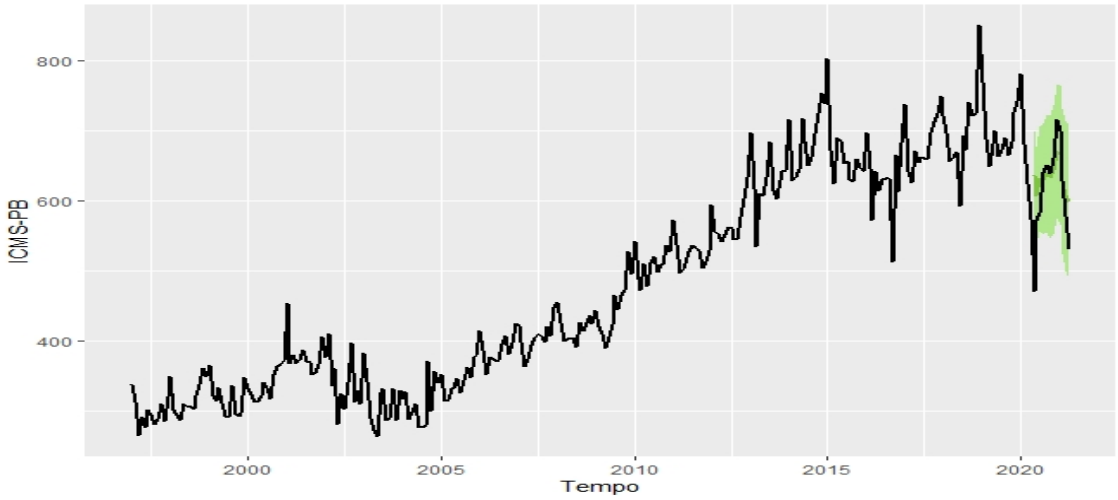

Figura 3.7: Gráfico das previsões para série ICMSPB-abr.20 com os limites superior e inferior com 95% de confiança, obtidos pelo modelo SARIMA$(0,1,1)\times(0,0,2)_{BIC}$.

Tabela 3.16: Medidas de precisão obtidas para a série ICMSPB-out.20.

	passos à frente		
Medidas de precisão - EQM	1	3	6
HW Aditivo	105.53	549.28	1437.13
HW Multiplicativo	155.35	635.08	1852.69
SARIMA$(1,1,1)\times(0,0,2)_{AIC}$	479.80	1802.54	1724.10
SARIMA$(0,1,1)\times(0,0,2)_{BIC}$	377.46	1333.81	1562.74
ARIMAX$_{Saz}(0,1,1)_{AIC,BIC}$	625.17	4988.04	4127.35
SARIMAX$_{pand}(1,1,1)\times(0,0,2)_{AIC}$	4827.63	2748.32	10864.91
SARIMAX$_{pand}(0,1,1)\times(0,0,2)_{BIC}$	7033.45	4738.30	14369.68
	passos à frente		
Medidas de precisão - ET	1	3	6
HW Aditivo	10.27	56.63	−33.81
HW Multiplicativo	12.46	49.87	−79.44
SARIMA$(1,1,1)\times(0,0,2)_{AIC}$	21.90	120.71	37.42
SARIMA$(0,1,1)\times(0,0,2)_{BIC}$	19.43	104.09	22.71
ARIMAX$_{Saz}(0,1,1)_{AIC,BIC}$	−25.00	143.62	25.95
SARIMAX$_{pand}(1,1,1)\times(0,0,2)_{AIC}$	−69.48	−151.74	−549.07
SARIMAX$_{pand}(0,1,1)\times(0,0,2)_{BIC}$	−83.86	−203.47	−653.60

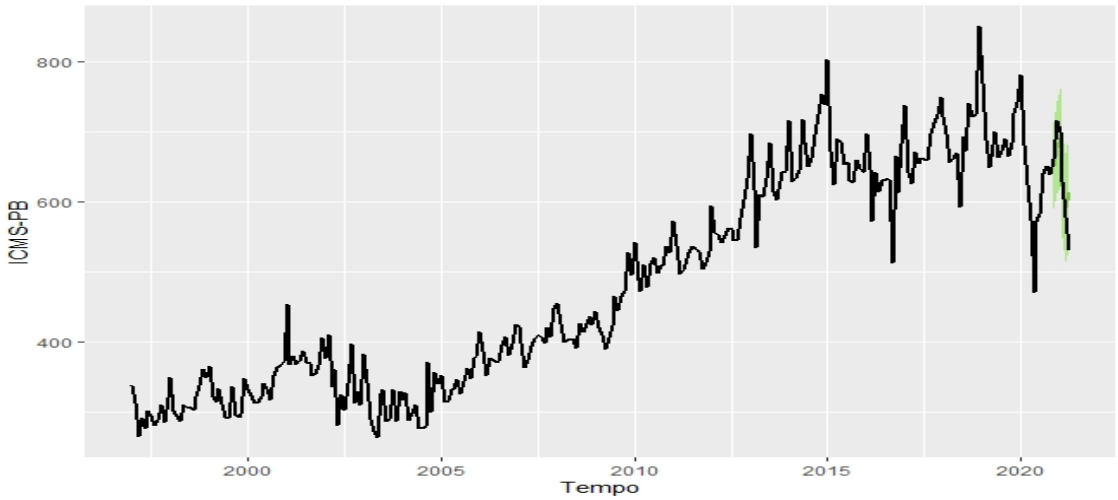

Figura 3.8: Gráfico das previsões para série ICMSPB-out.20 com os limites superior e inferior com 95% de confiança, obtidos pelo HW com sazonalidade aditiva.

Capítulo 4

Conclusões

O objetivo dessa dissertação foi prever e analisar as arrecadações mensais do ICMS da Paraíba, usando os dados de janeiro de 1997 a abril de 2021. Para alcançar o objetivo, foi utilizado os algoritmos de alisamento exponencial Holt-Winters com sazonalidade aditiva e multiplicativa, e modelos Box-Jenkins SARIMA, SARIMAX com a tendência e uma variável dummy referente a pandemia do COVID-19 como variáveis exógenas, e para a utilização da sazonalidade como variável regressiva utilizamos o modelo ARIMAX.

Após identificar os modelos, estimar os parâmetros, e diagnosticar através de avaliação dos resíduos, foram obtidas previsões e feitas comparações com os valores reais de arrecadações.

Considerando o erro quadrático médio e erro total obtidos através da relação entre as arrecadações e previsões, selecionaram os modelos que geraram as melhores previsões para cada série, assim observou-se que algoritmo de alisamento exponencial Holt-Winters com sazonalidade multiplicativa obteve as melhores previsões para a série ICMSPB-out.19, com valores previstos bem próximos dos encontrados pelo algoritmo de alisamento exponencial Holt-Winters com sazonalidade aditiva. Porém para a série ICMSPB-out.20, o algoritmo de alisamento exponencial Holt-Winters com sazonalidade aditiva foi o mais adequado seguido pelo HW com sazonalidade multiplicativa.

Entretanto o modelo que alcançou as melhores previsões para as séries ICMSPB-abr.19 e ICMSPB-abr.20 foi o SARIMA$(0,1,1) \times (0,0,2)$, sendo assim o mais adequado para realizar as previsões para o ICMS da Paraíba.

Diante do exposto, a previsão da arrecadação do ICMS da Paraíba pode ser realizada pelo modelo SARIMA$(0,1,1) \times (0,0,2)$, o qual poderá contribuir com o Estado na estimação das receitas e fixação das despesas para os anos seguintes. Porém é visto que para previsões de períodos curto, como 1 e 3 passos a frente, o algoritmo de alisamento exponencial Holt-Winters tem valores mais próximos dos

valoreis reais, quando comparados com os modelos Box-Jenkins.

Referências Bibliográficas

AKAIKE, H., 1978, "A Baysian Analysis of The Minimum AIC Procedure", .

BOX, G. E. P., JENKINS, G. M., REINSEL, G. C., et al., 2015, *Forecasting and Control*. Wiley.

BROCKWELL, P. J., DAVIS, R. A., 2016, *Introduction to Time Series and Forecasting*. N. 4. Springer.

CAMARGOS, A. A. B., 2008, *Modelos de Previsão da Arrecadação Tributária do Estado de São paulo:ICMS, IPVA, ITCMD e Taxas*. Tese de Mestrado.

CASTANHO, B. J. D. S., 2011, *Modelos para previsão de receitas tributárias: o ICMS do Estado do Espírito Santo*. Tese de Mestrado.

CONFAZ. "Imposto sobre a circulação de mercadorias(ICMS)-Paraíba". Disponível em: <https://www.confaz.fazenda.gov.br/boletim-de-arrecadacao-dos-tributos-estaduais>. acesso em: 13 de Junho de 2021.

COWPERTWAIT, P. S., METCALFE, A. V., 2009, *Introductory Time Series with R*. Springer-Verlag.

COX, D. R., STUART, A., 1955, "Some quick tests for trend in location and dispersion", v. 42, pp. 80–95.

CRYER, J. D., CHAN, K.-S., 2008, *Time Series Analysis: With Applications in R*, v. 2. Springer-Verlag.

DUARTE, F. C. D. L., SOUZA, M. F. D., GIRÃO, L. F. D. A. P., 2014, "Previsão da Arrecadação do ICMS no Estado da Paraíba", v. 7, pp. 235–250.

EHLERS, R. S., 2007, "Análise de Séries Temporais", Disponível em: <http://www.each.usp.br/rvicente/AnaliseDeSeriesTemporais.pdf>. acesso em: 08 Novembro 2020.

ESQUIVEL, R. D. M., 2012, "Análise Espectral Singular: Modelagem de séries temporais através de estudos comparativos usando diferentes estratégias de previsão", .

FARIAS, J. F., 2017, *Uma aplicação de séries temporais na arrecadação do ICMS do Estado da Paraíba*. Tese de Mestrado.

FGV-IBE, I. B. D. E. "Série historica do Índice Geral de Preços – Disponibilidade Interna (IGP-DI)". Disponível em: <http://www14.fgv.br/fgvdados20/visualizaconsulta.aspx>. acesso em: 13 de junho de 2021.

FISCHER, S., 1982, *Séries univariantes de tempo - metodologia de Box e Jenkins*. N. 4. Fundação de Economia e Estatística.

FRIEDMAN, M., 1937, "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance", v. 32, pp. 675–701.

GOMES, A. D. S., 2003, *Modelagem e Previsão da Arrecadação do Imposto de Renda no Brasil*. Tese de Mestrado.

HOFFMANN, R., 2016, *Análise de regressão: uma introdução à econometria*.

KRUSKAL, W. H., WALLIS, W. A., 1952, "Use of Ranks in One-Criterion Variance Analysis", v. 47, pp. 80–95.

LJUNG, G. M., BOX, G. E. P., 1978, "On a measure of lack of fit in time serie models", p. 297–303.

MACEDO JR, J. S., 2013, *A Árvore do Dinheiro*. Editora Insular.

MOÇO, J. R. C., 2017, *Estimação de um modelo de previsão do ICMS do Estado do Rio de Janeiro*. Tese de Mestrado.

MORETTIN, P. A., 2011, *Econometria Financeira: Um Curso de Séries Temporais Financeiras*. Edgard Blucher.

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

SILVESTRE, C. D., 2013, *Análise do ICMS do estado do Ceará no período de 2000 a 2013*. Tese de Mestrado.

WALD, A., WOLFOWITZ, J., 1948, "Optimum character of the sequential probabilidade ratio test", v. 19, pp. 326–339.

Apêndice A

Diagnóstico

Este apêndice contém os gráficos da ACF dos resíduos, da estacionariedade dos resíduos e do p valor para o teste de Ljung-Box, utilizados para análise diagnóstica dos resíduos, para os modelos escolhidos através dos critérios AIC e BIC para cada série.

Série ICMSPB-abr.19

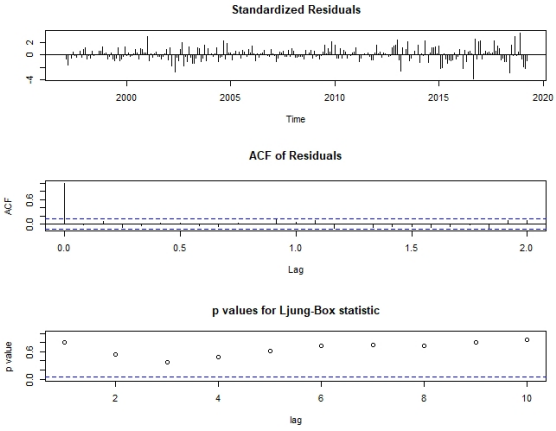

Figura A.1: SARIMA$(1,1,1) \times (0,0,2)$.

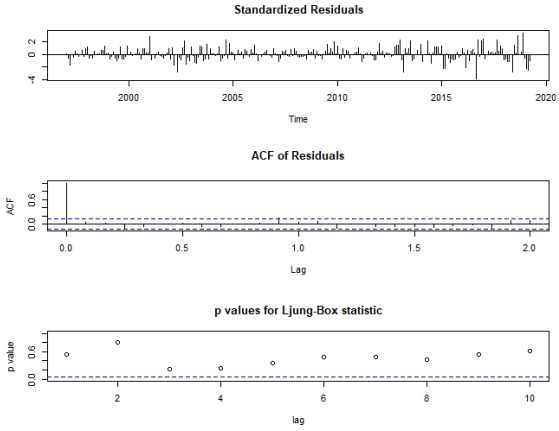

Figura A.2: SARIMA$(0,1,1) \times (0,0,2)$.

Série ICMSPB-out.19

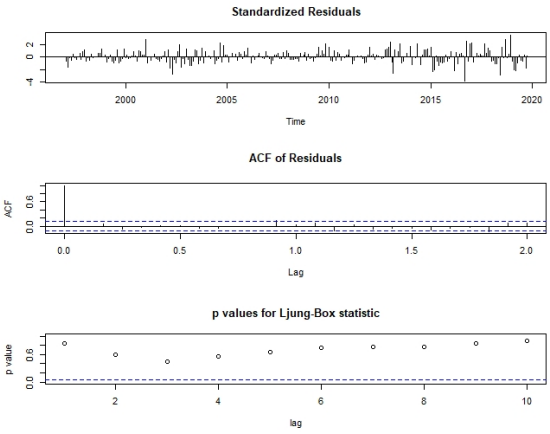

Figura A.3: SARIMA$(1,1,1) \times (0,0,2)$.

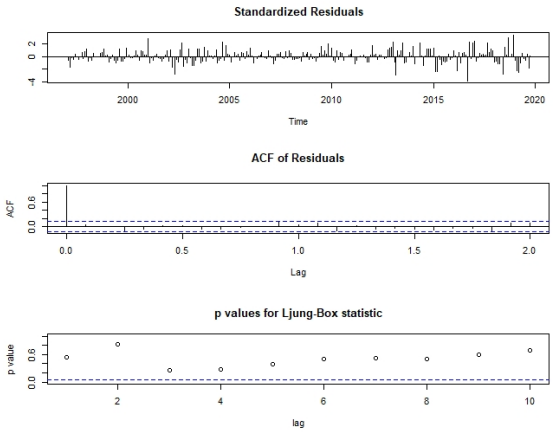

Figura A.4: Análise diagnóstica-SARIMA$(0,1,1) \times (0,0,2)$.

Série ICMSPB-abr.20

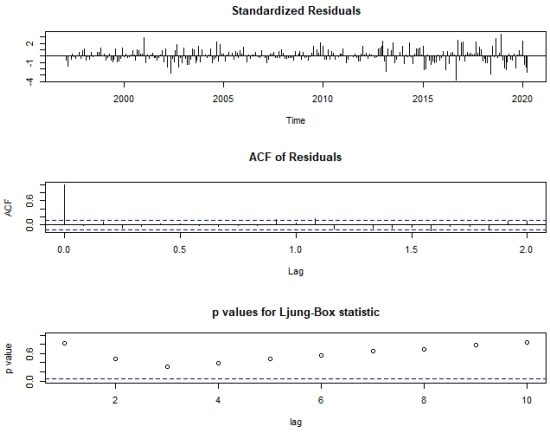

Figura A.5: SARIMA$(1,1,1) \times (0,0,2)$.

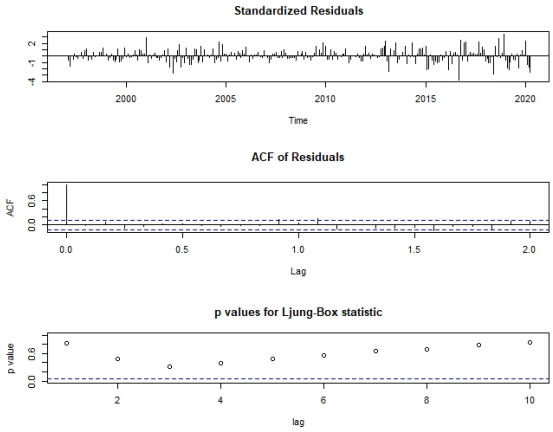

Figura A.6: SARIMA$(0,1,1) \times (0,0,2)$.

Série ICMSPB-out.20

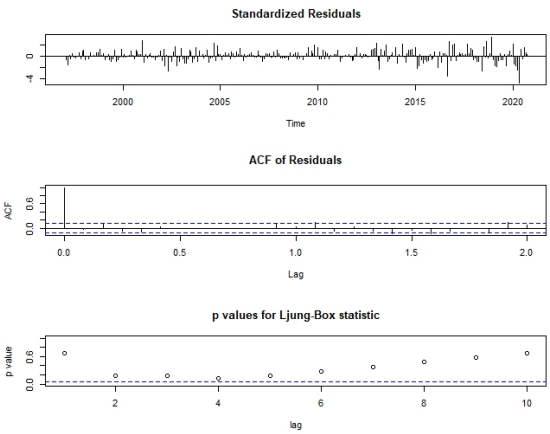

Figura A.7: SARIMA$(1,1,1) \times (0,0,2)$.

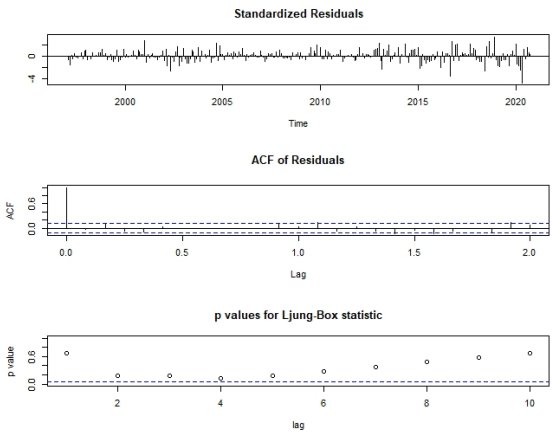

Figura A.8: SARIMA$(0,1,1) \times (0,0,2)$.

Série ICMSPB-abr.19

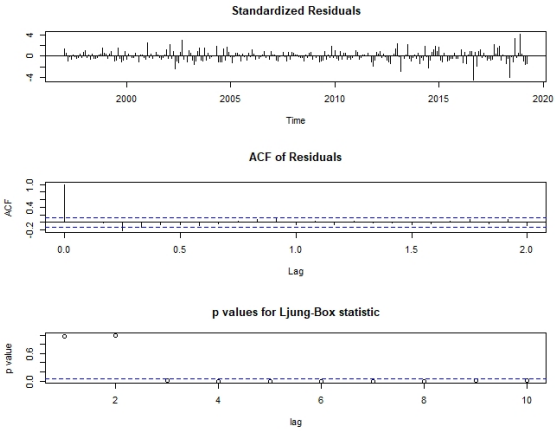

Figura A.9: SARIMAX$(0,0,0) \times (2,0,0)$, reg.tendência.

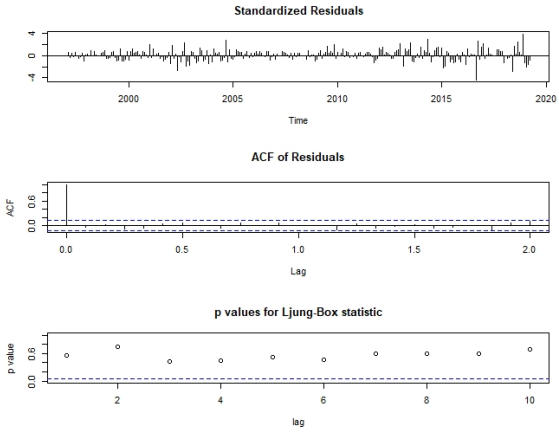

Figura A.10: ARIMAX$(0,1,1)$, reg.sazonal.

Série ICMSPB-out.19

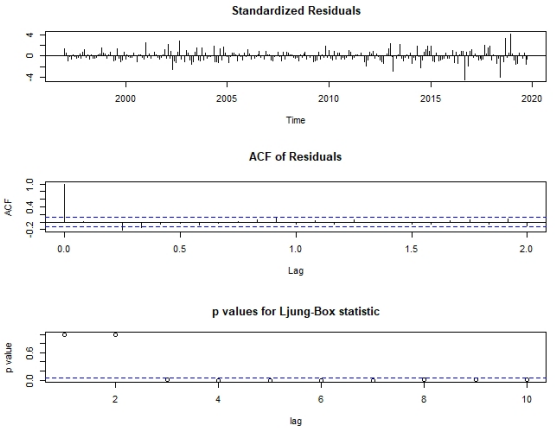

Figura A.11: SARIMAX$(0,0,0) \times (2,0,0)$, reg.tendência.

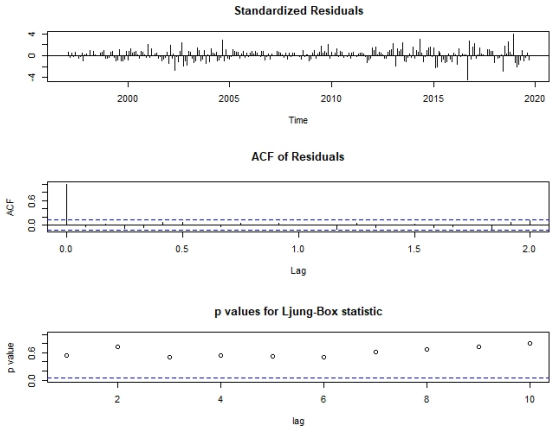

Figura A.12: ARIMAX$(0,1,1)$, reg.sazonal.

Série ICMSPB-abr.20

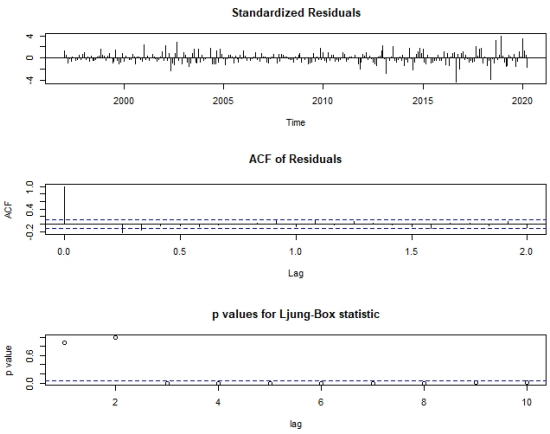

Figura A.13: SARIMAX$(0,0,0)\times(2,0,0)$, reg.tendência.

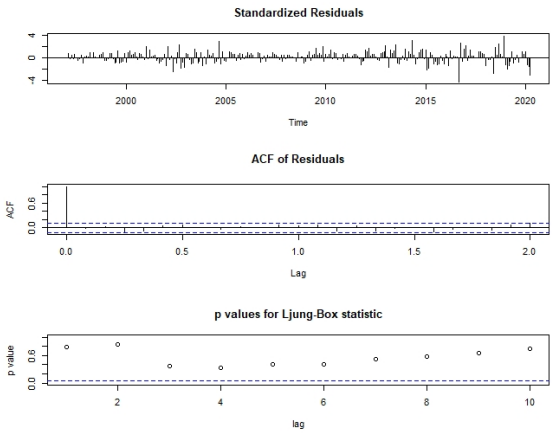

Figura A.14: ARIMAX$(0,1,1)$, reg.sazonal.

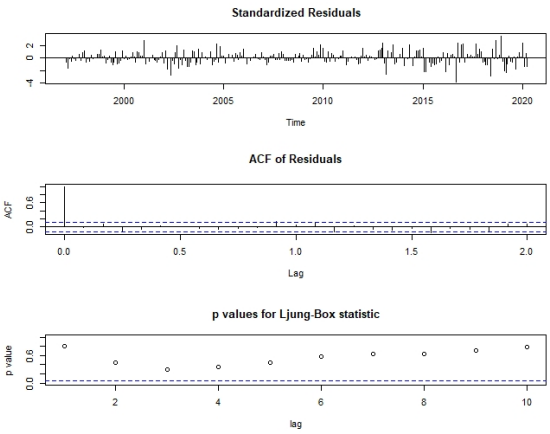

Figura A.15: SARIMAX$(1,1,1)\times(0,0,2)$, reg.pandemia.

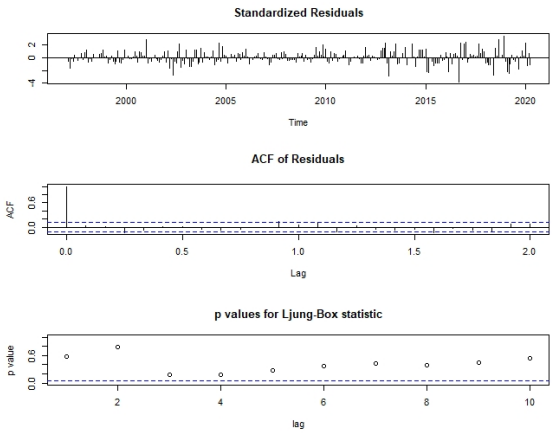

Figura A.16: SARIMAX$(0,1,1)\times(0,0,2)$, reg.pandemia.

Série ICMSPB-out.20

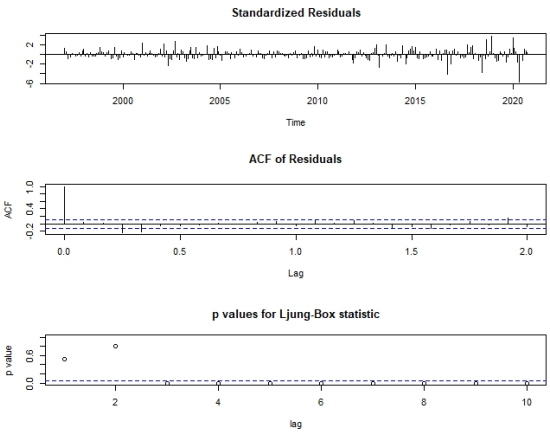

Figura A.17: SARIMA$(0,0,0) \times (2,0,0)$, reg.tendência.

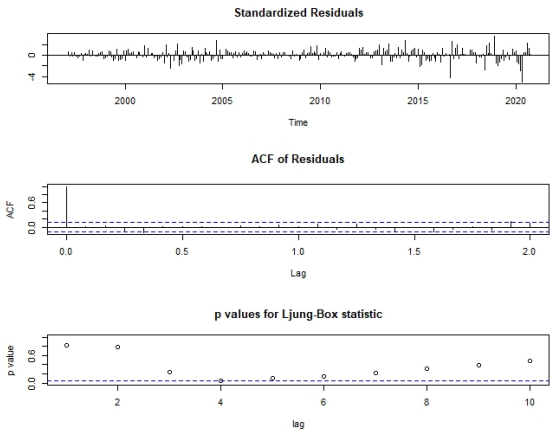

Figura A.18: ARIMAX$(0,1,1)$, reg.sazonal.

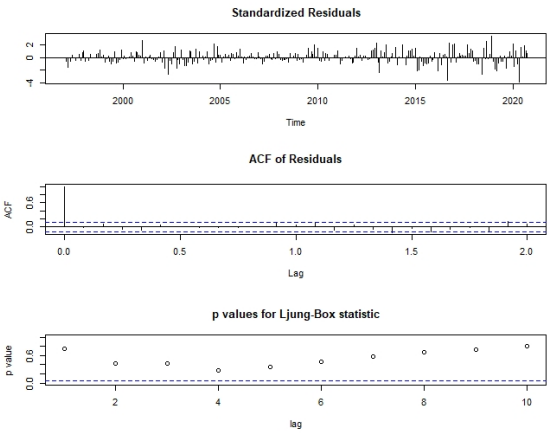

Figura A.19: SARIMAX$(1,1,1)\times(0,0,2)$, reg.pandemia.

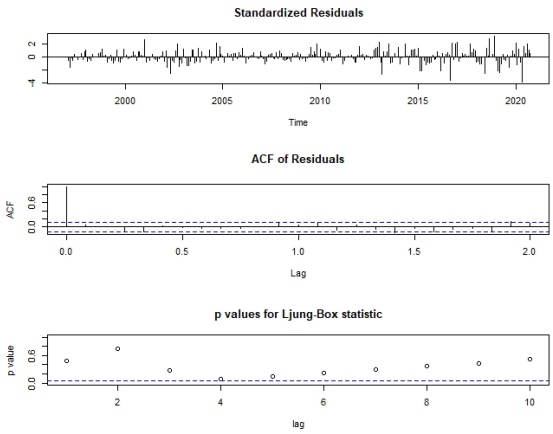

Figura A.20: SARIMAX$(0,1,1)\times(0,0,2)$, reg.pandemia.

Apêndice B

Algoritmos

Aqui são apresentados os algorítimos utilizados para obtenção dos dados, gráficos e tabelas contidos nesse trabalho.

Algorítimo para Análise e Previsão do ICMS-PB

```
library(expsmooth)
library(fma)
library(forecast)
library(tseries)
library(fpp)
library (ggplot2)

dados = read.table("igpdi_icms_confaz_atualizado.txt", header=T, stringsAsFactors =

igpdi= dados[,2]
icmspb= dados[,3]/1000 # formatando os dados para Milhões

tam_igpdi = length(igpdi) # "tamanho" quantidade de dados do igpdi
tam_icmspb = length(icmspb) # "tamanho" quantidade de dados do icms

# Aqui dividimos todas as observações da série
# por seu ultimo valor. Dessa forma, a série deflacionada
# fica a precos constantes da ultima observação.
```

```
igpdi_d = igpdi/igpdi[tam_igpdi]
igpdi_d

# Agora dividimos os valores da arrecadados do ICMS
# pelo IGP-DI a fim de obter uma serie livre de
# dinamina inflacionaria
icmspb_d = icmspb/igpdi_d

# Aqui vamos excluir as 6, 12, 18 e 24 observações

icmspb_d_6=icmspb_d[-c(287:292)]
length(icmspb_d_6)
# valores reais
h.1_6= icmspb_d[287]
h.3_6= icmspb_d[287:289]
h.6_6= icmspb_d[287:292]

icmspb_d_12=icmspb_d[-c(281:292)]
length(icmspb_d_12)
h.1_12= icmspb_d[281]
h.3_12= icmspb_d[281:283]
h.6_12= icmspb_d[281:286]
h.12_12= icmspb_d[281:292]

icmspb_d_18=icmspb_d[-c(275:292)]
length(icmspb_d_18)
h.1_18= icmspb_d[275]
h.3_18= icmspb_d[275:277]
h.6_18= icmspb_d[275:280]
h.12_18= icmspb_d[275:286]

icmspb_d_24=icmspb_d[-c(269:292)]
```

```
length(icmspb_d_24)
h.1_24= icmspb_d[269]
h.3_24= icmspb_d[269:271]
h.6_24= icmspb_d[269:274]
h.12_24= icmspb_d[269:280]


# Aqui formatamos os dados como uma série temporal
icmspb_d = ts(icmspb_d,start=c(1997,1),frequency=12)
length(icmspb_d)

icmspb_out.20 = ts(icmspb_d_6,start=c(1997,1),frequency=12)
length(icmspb_out.20)
icmspb_out.20

icmspb_abr.20 = ts(icmspb_d_12,start=c(1997,1),frequency=12)
length(icmspb_abr.20)
icmspb_abr.20

icmspb_out.19 = ts(icmspb_d_18,start=c(1997,1),frequency=12)
length(icmspb_out.19)
icmspb_out.19

icmspb_abr.19 = ts(icmspb_d_24,start=c(1997,1),frequency=12)
length(icmspb_abr.19)
icmspb_abr.19

# diferenciação
ndiffs(icmspb_d) # d=1, ou seja, para tornar a série estacionária precisamos tomar
nsdiffs(icmspb_d) # D=0

# Estatística descritiva da série
```

```
summary(icmspb_d)
sd(icmspb_d)# desvio padrão

# Estatística descritiva da série diferenciada
mean(diff(icmspb_d))
median(diff(icmspb_d))
sqrt(var(diff(icmspb_d)))
min(diff(icmspb_d))
max(diff(icmspb_d))

# Gráficos
autoplot(icmspb_d, main = "", xlab="Tempo", ylab="ICMS")
autoplot(decompose(icmspb_d),main="")

# função de autocorrelação amostral e parcial
acf(icmspb_d , lag.max =36,main='')
pacf (icmspb_d , lag.max =36,main='')

################ Alisamento exponencial  ######################

######### efeito sazonal aditivo ############

############ serie truncada abr.19 ######################

hw.abr.19 = HoltWinters (icmspb_abr.19, seasonal = "additive")
hw.abr.19
p.ae.1_24 = predict (hw.abr.19, 1)
p.ae.3_24 = predict (hw.abr.19, 3)
p.ae.6_24 = predict (hw.abr.19, 6)
p.ae.12_24 = predict (hw.abr.19, 12)

## erros para a serie truncada abr.19 ##
```

```
accuracy(p.ae.1_24,  h.1_24)
accuracy(p.ae.3_24,  h.3_24)
accuracy(p.ae.6_24,  h.6_24)
accuracy(p.ae.12_24,  h.12_24)


############ serie truncada out.19 ####################

hw.out.19 = HoltWinters (icmspb_out.19)
hw.out.19
p.ae.1_18 = predict (hw.out.19, 1)
p.ae.3_18 = predict (hw.out.19, 3)
p.ae.6_18 = predict (hw.out.19, 6)
p.ae.12_18 = predict (hw.out.19, 12)

## erros para a serie truncada out.19 ##
accuracy(p.ae.1_18,  h.1_18)
accuracy(p.ae.3_18,  h.3_18)
accuracy(p.ae.6_18,  h.6_18)
accuracy(p.ae.12_18,  h.12_18)


############ serie truncada abr.20 ###################

hw.abr.20 = HoltWinters (icmspb_abr.20)
hw.abr.20
p.ae.1_12 = predict (hw.abr.20, 1)
p.ae.3_12 = predict (hw.abr.20, 3)
p.ae.6_12 = predict (hw.abr.20, 6)
p.ae.12_12 = predict (hw.abr.20, 12)

## erros para a serie truncada abr.20 ##
```

```
accuracy(p.ae.1_12,  h.1_12)
accuracy(p.ae.3_12,  h.3_12)
accuracy(p.ae.6_12,  h.6_12)
accuracy(p.ae.12_12,  h.12_12)


########### serie truncada out.20 ##################

hw.out.20 = HoltWinters (icmspb_out.20)
hw.out.20
p.ae.1_6 = predict (hw.out.20, 1)
p.ae.3_6 = predict (hw.out.20, 3)
p.ae.6_6 = predict (hw.out.20, 6)

## erros para a serie truncada out.20 ##
accuracy(p.ae.1_6,  h.1_6)
accuracy(p.ae.3_6,  h.3_6)
accuracy(p.ae.6_6,  h.6_6)

#####################################################
######### Efeito Sazonal Multiplicativo ############
#####################################################

########### serie truncada abr.19 ######################

hwm.abr.19 = HoltWinters (icmspb_abr.19,seasonal = "multiplicative")
hwm.abr.19
p.me.1_24 = predict (hwm.abr.19, 1)
p.me.3_24 = predict (hwm.abr.19, 3)
p.me.6_24 = predict (hwm.abr.19, 6)
p.me.12_24 = predict (hwm.abr.19, 12)
```

```
## erros  ##
accuracy(p.me.1_24,  h.1_24)
accuracy(p.me.3_24,  h.3_24)
accuracy(p.me.6_24,  h.6_24)
accuracy(p.me.12_24,  h.12_24)


############ serie truncada out.19 ####################

hwm.out.19 = HoltWinters (icmspb_out.19, seasonal = "multiplicative")
hwm.out.19
p.me.1_18 = predict (hwm.out.19, 1)
p.me.3_18 = predict (hwm.out.19, 3)
p.me.6_18 = predict (hwm.out.19, 6)
p.me.12_18 = predict (hwm.out.19, 12)

## erros  ##
accuracy(p.me.1_18,  h.1_18)
accuracy(p.me.3_18,  h.3_18)
accuracy(p.me.6_18,  h.6_18)
accuracy(p.me.12_18,  h.12_18)

############ serie truncada abr.20 ###################

hwm.abr.20 = HoltWinters (icmspb_abr.20, seasonal = "multiplicative")
hwm.abr.20
p.me.1_12 = predict (hwm.abr.20, 1)
p.me.3_12 = predict (hwm.abr.20, 3)
p.me.6_12 = predict (hwm.abr.20, 6)
p.me.12_12 = predict (hwm.abr.20, 12)

# erros  ##
```

```
accuracy(p.me.1_12,  h.1_12)
accuracy(p.me.3_12,  h.3_12)
accuracy(p.me.6_12,  h.6_12)
accuracy(p.me.12_12,  h.12_12)

########### serie truncada out.20 ##################

hwm.out.20 = HoltWinters (icmspb_out.20, seasonal = "multiplicative")
hwm.out.20
p.me.1_6 = predict (hwm.out.20, 1)
p.me.3_6 = predict (hwm.out.20, 3)
p.me.6_6 = predict (hwm.out.20, 6)

## erros  ##
accuracy(p.me.1_6,  h.1_6)
accuracy(p.me.3_6,  h.3_6)
accuracy(p.me.6_6,  h.6_6)

############## BOX-JENKINS ################
#### selecao de modelos considerando as series truncadas ####

auto.arima(icmspb_abr.19, ic="aic") # seleção do modelo para a serie icmspb_abr.19
auto.arima(icmspb_abr.19, ic="bic") # seleção do modelo para a serie icmspb_abr.19

auto.arima(icmspb_out.19, ic="aic")# seleção do modelo para a serie icmspb_out.19 p
auto.arima(icmspb_out.19, ic="bic")# seleção do modelo para a serie icmspb_out.19 p

auto.arima(icmspb_abr.20, ic="aic") # seleção do modelo para a serie icmspb_abr.20
auto.arima(icmspb_abr.20, ic="bic")# seleção do modelo para a serie icmspb_abr.20 p

auto.arima(icmspb_out.20, ic="aic")# seleção do modelo para a serie icmspb_out.20 p
auto.arima(icmspb_out.20, ic="bic")# seleção do modelo para a serie icmspb_out.20 p
```

```
########### serie truncada abr.19#####################

ajuste.1_abr.19 = Arima (icmspb_abr.19 , order =c(1,1,1),seasonal = list(order=c(0,
tsdiag(ajuste.1_abr.19)

ajuste.2_abr.19 = Arima (icmspb_abr.19 , order =c(0,1,1),seasonal = list(order=c(0,
tsdiag(ajuste.2_abr.19)

# valores previstos para ajuste 1
p1.1_24=forecast(ajuste.1_abr.19, h=1)
p1.3_24=forecast(ajuste.1_abr.19, h=3)
p1.6_24=forecast(ajuste.1_abr.19, h=6)
p1.12_24=forecast(ajuste.1_abr.19, h=12)

# valores previstos para ajuste 2
p2.1_24=forecast(ajuste.2_abr.19, h=1)
p2.3_24=forecast(ajuste.2_abr.19, h=3)
p2.6_24=forecast(ajuste.2_abr.19, h=6)
p2.12_24=forecast(ajuste.2_abr.19, h=12)


p1.1_24=as.numeric(p1.1_24$mean)
p1.3_24=as.numeric(p1.3_24$mean)
p1.6_24=as.numeric(p1.6_24$mean)
p1.12_24=as.numeric(p1.12_24$mean)

p2.1_24=as.numeric(p2.1_24$mean)
p2.3_24=as.numeric(p2.3_24$mean)
p2.6_24=as.numeric(p2.6_24$mean)
p2.12_24=as.numeric(p2.12_24$mean)
```

```
## erros para previsões ajuste 1 ##
accuracy(p1.1_24,  h.1_24)
accuracy(p1.3_24,  h.3_24)
accuracy(p1.6_24,  h.6_24)
accuracy(p1.12_24,  h.12_24)

## erros para previsões ajuste 2 ##
accuracy(p2.1_24,  h.1_24)
accuracy(p2.3_24,  h.3_24)
accuracy(p2.6_24,  h.6_24)
accuracy(p2.12_24,  h.12_24)

########### serie truncada out.19  #############

ajuste.1_out.19 = Arima (icmspb_out.19 , order =c(1,1,1),seasonal = list(order=c(0,
tsdiag(ajuste.1_out.19)

ajuste.2_out.19 = Arima (icmspb_out.19 , order =c(0,1,1),seasonal = list(order=c(0,
tsdiag(ajuste.2_out.19)

# valores previstos para ajuste 1
p1.1_18=forecast(ajuste.1_out.19, h=1)
p1.3_18=forecast(ajuste.1_out.19, h=3)
p1.6_18=forecast(ajuste.1_out.19, h=6)
p1.12_18=forecast(ajuste.1_out.19, h=12)

# valores previstos para ajuste 2
p2.1_18=forecast(ajuste.2_out.19, h=1)
p2.3_18=forecast(ajuste.2_out.19, h=3)
p2.6_18=forecast(ajuste.2_out.19, h=6)
```

```
p2.12_18=forecast(ajuste.2_out.19, h=12)


p1.1_18=as.numeric(p1.1_18$mean)
p1.3_18=as.numeric(p1.3_18$mean)
p1.6_18=as.numeric(p1.6_18$mean)
p1.12_18=as.numeric(p1.12_18$mean)

p2.1_18=as.numeric(p2.1_18$mean)
p2.3_18=as.numeric(p2.3_18$mean)
p2.6_18=as.numeric(p2.6_18$mean)
p2.12_18=as.numeric(p2.12_18$mean)

## erros para previsões ajuste 1 ##
accuracy(p1.1_18,  h.1_18)
accuracy(p1.3_18,  h.3_18)
accuracy(p1.6_18,  h.6_18)
accuracy(p1.12_18,  h.12_18)

## erros para previsões ajuste 2 ##
accuracy(p2.1_18,  h.1_18)
accuracy(p2.3_18,  h.3_18)
accuracy(p2.6_18,  h.6_18)
accuracy(p2.12_18,  h.12_18)


########### serie truncada abr.20 #################

ajuste.1_abr.20 = Arima (icmspb_abr.20 , order =c(1,1,1),seasonal = list(order=c(0,
tsdiag(ajuste.1_abr.20)

ajuste.2_abr.20 = Arima (icmspb_abr.20 , order =c(0,1,1),seasonal = list(order=c(0,
```

```
tsdiag(ajuste.1_abr.20)

# valores previstos para ajuste 1
p1.1_12=forecast(ajuste.1_abr.20, h=1)
p1.3_12=forecast(ajuste.1_abr.20, h=3)
p1.6_12=forecast(ajuste.1_abr.20, h=6)
p1.12_12=forecast(ajuste.1_abr.20, h=12)

# valores previstos para ajuste 2
p2.1_12=forecast(ajuste.2_abr.20, h=1)
p2.3_12=forecast(ajuste.2_abr.20, h=3)
p2.6_12=forecast(ajuste.2_abr.20, h=6)
p2.12_12=forecast(ajuste.2_abr.20, h=12)



p1.1_12=as.numeric(p1.1_12$mean)
p1.3_12=as.numeric(p1.3_12$mean)
p1.6_12=as.numeric(p1.6_12$mean)
p1.12_12=as.numeric(p1.12_12$mean)

p2.1_12=as.numeric(p2.1_12$mean)
p2.3_12=as.numeric(p2.3_12$mean)
p2.6_12=as.numeric(p2.6_12$mean)
p2.12_12=as.numeric(p2.12_12$mean)

## erros para ajuste 1 ##
accuracy(p1.1_12,  h.1_12)
accuracy(p1.3_12,  h.3_12)
accuracy(p1.6_12,  h.6_12)
accuracy(p1.12_12,  h.12_12)

## erros para ajuste 2 ##
```

```
accuracy(p2.1_12,  h.1_12)
accuracy(p2.3_12,  h.3_12)
accuracy(p2.6_12,  h.6_12)
accuracy(p2.12_12,  h.12_12)


########### serie truncada out.20  #################

ajuste.1_out.20 = Arima (icmspb_out.20 , order =c(1,1,1),seasonal = list(order=c(0,
tsdiag(ajuste.1_out.20)

ajuste.2_out.20 = Arima (icmspb_out.20 , order =c(0,1,1),seasonal = list(order=c(0,
tsdiag(ajuste.1_out.20)

# valores previstos para ajuste 1
p1.1_6=forecast(ajuste.1_out.20, h=1)
p1.3_6=forecast(ajuste.1_out.20, h=3)
p1.6_6=forecast(ajuste.1_out.20, h=6)

# valores previstos para ajuste 2
p2.1_6=forecast(ajuste.2_out.20, h=1)
p2.3_6=forecast(ajuste.2_out.20, h=3)
p2.6_6=forecast(ajuste.2_out.20, h=6)

p1.1_6=as.numeric(p1.1_6$mean)
p1.3_6=as.numeric(p1.3_6$mean)
p1.6_6=as.numeric(p1.6_6$mean)

p2.1_6=as.numeric(p2.1_6$mean)
p2.3_6=as.numeric(p2.3_6$mean)
p2.6_6=as.numeric(p2.6_6$mean)
```

```
## erros para ajuste 1 ##
accuracy(p1.1_6,  h.1_6)
accuracy(p1.3_6,  h.3_6)
accuracy(p1.6_6,  h.6_6)

## erros para ajuste 2 ##
accuracy(p2.1_6,  h.1_6)
accuracy(p2.3_6,  h.3_6)
accuracy(p2.6_6,  h.6_6)

###########################################
############SARIMAX#########################
###########################################

#Variavel Dummy
pandemia = c(rep(0,278),rep(1,14)) # variavel dummy  que recebe 1 a partir de março
str(pandemia)
pandemia12=pandemia[-c(281:292)]
pandemia6=pandemia[-c(287:292)]
#tendencia estimada
STL=stl(icmspb_d,s.window = 'periodic') #extraindo tendecia da série
t=STL$time.series[,2]
str(t)
tendencia.abr.19 = t[-c(269:292)] # truncando a tendencia para a serie icmspb_abr.1
tendencia.out.19 = t[-c(275:292)] # truncando a tendencia para a serie icmspb_out.1
tendencia.abr.20 = t[-c(281:292)] # truncando a tendencia para a serie icmspb_abr.2
tendencia.out.20 = t[-c(287:292)] # truncando a tendencia para a serie icmspb_out.2

#sazonalidade estimada
s=STL$time.series[,1]#extraindo a sazonalidade
str(s)
sazonal.abr.19 = s[-c(269:292)] # truncando a sazonalidade para a serie icmspb_abr.
```

```
sazonal.out.19 = s[-c(275:292)] # truncando a sazonalidade para a serie icmspb_out.
sazonal.abr.20 = s[-c(281:292)] # truncando a sazonalidade para a serie icmspb_abr.
sazonal.out.20 = s[-c(287:292)] # truncando a sazonalidade para a serie icmspb_out.

#### selecao de modelos considerando as series truncadas ####

####### seleção do modelo para a serie icmspb_abr.19#######

auto.arima(icmspb_abr.19, ic="aic",xreg = tendencia.abr.19 ) #  pelo criterio AIC-
auto.arima(icmspb_abr.19, ic="bic",xreg = tendencia.abr.19 ) #  pelo criterio BIC-
auto.arima(icmspb_abr.19, ic="aic",xreg = sazonal.abr.19 )   #  pelo criterio AIC-
auto.arima(icmspb_abr.19, ic="bic",xreg = sazonal.abr.19 )   #  pelo criterio BIC-

####### seleção do modelo para a serie icmspb_out.19#######
auto.arima(icmspb_out.19, ic="aic",xreg = tendencia.out.19) #  pelo criterio AIC- S
auto.arima(icmspb_out.19, ic="bic",xreg = tendencia.out.19) #  pelo criterio BIC- S
auto.arima(icmspb_out.19, ic="aic",xreg = sazonal.out.19)   #  pelo criterio AIC- A
auto.arima(icmspb_out.19, ic="bic",xreg = sazonal.out.19)   #  pelo criterio BIC- A

####### seleção do modelo para a serie icmspb_abr.20#######
auto.arima(icmspb_abr.20, ic="aic",xreg = tendencia.abr.20) #  pelo criterio AIC- S
auto.arima(icmspb_abr.20, ic="bic",xreg = tendencia.abr.20) #  pelo criterio BIC- S
auto.arima(icmspb_abr.20, ic="aic",xreg = sazonal.abr.20)   #  pelo criterio AIC- A
auto.arima(icmspb_abr.20, ic="bic",xreg = sazonal.abr.20)   #  pelo criterio BIC- A
auto.arima(icmspb_abr.20, ic="aic",xreg = pandemia12)       #  pelo criterio AIC- S
auto.arima(icmspb_abr.20, ic="bic",xreg = pandemia12)       #  pelo criterio BIC- S

####### seleção do modelo para a serie icmspb_out.20#######
auto.arima(icmspb_out.20, ic="aic",xreg = tendencia.out.20) #  pelo criterio AIC- S
auto.arima(icmspb_out.20, ic="bic",xreg = tendencia.out.20) #  pelo criterio BIC- S
auto.arima(icmspb_out.20, ic="aic",xreg = sazonal.out.20)   #  pelo criterio AIC- A
auto.arima(icmspb_out.20, ic="bic",xreg = sazonal.out.20)   #  pelo criterio AIC- A
```

```
auto.arima(icmspb_out.20, ic="aic",xreg = pandemia6)        #  pelo criterio AIC- S
auto.arima(icmspb_out.20, ic="bic",xreg = pandemia6)        #  pelo criterio BIC- S

############################################################
########### ajuste serie truncada abr.19####################

ajusteAIC.ten_abr.19 = Arima (icmspb_abr.19 , order =c(0,0,0),seasonal = list(order
tsdiag(ajusteAIC.ten_abr.19) # modemo não bem estimado

ajuste.saz_abr.19 = Arima (icmspb_abr.19 , order =c(0,1,1),xreg = sazonal.abr.19)##
tsdiag(ajuste.saz_abr.19)


# valores previstos para ajuste.saz_abr.19
ps.1_24=forecast(ajuste.saz_abr.19,xreg = sazonal.abr.19, h=1)
ps.3_24=forecast(ajuste.saz_abr.19,xreg = sazonal.abr.19, h=3)
ps.6_24=forecast(ajuste.saz_abr.19,xreg = sazonal.abr.19, h=6)
ps.12_24=forecast(ajuste.saz_abr.19,xreg = sazonal.abr.19,h=12)

ps.1_24=as.numeric(ps.1_24$mean[1])
ps.3_24=as.numeric(ps.3_24$mean[1:3])
ps.6_24=as.numeric(ps.6_24$mean[1:6])
ps.12_24=as.numeric(ps.12_24$mean[1:12])


## erros para previsões ajuste.saz_abr.19 ##
accuracy(ps.1_24,  h.1_24)
accuracy(ps.3_24,  h.3_24)
accuracy(ps.6_24,  h.6_24)
accuracy(ps.12_24, h.12_24)

############################################################
```

```
########### ajuste serie truncada out.19#####################
ajuste.ten_out.19 = Arima (icmspb_out.19 , order =c(0,0,0),seasonal = list(order=c(
tsdiag(ajuste.ten_out.19)# modemo não bem estimado

ajuste.saz_out.19 = Arima (icmspb_out.19 , order =c(0,1,1),xreg = sazonal.out.19)##
tsdiag(ajuste.saz_out.19)

# valores previstos para ajuste.saz_out.19
ps.1_18=forecast(ajuste.saz_out.19, h=1,xreg = sazonal.out.19)
ps.3_18=forecast(ajuste.saz_out.19, h=3,xreg = sazonal.out.19)
ps.6_18=forecast(ajuste.saz_out.19, h=6,xreg = sazonal.out.19)
ps.12_18=forecast(ajuste.saz_out.19, h=12,xreg = sazonal.out.19)

ps.1_18=as.numeric(ps.1_18$mean[1])
ps.3_18=as.numeric(ps.3_18$mean[1:3])
ps.6_18=as.numeric(ps.6_18$mean[1:6])
ps.12_18=as.numeric(ps.12_18$mean[1:12])

## erros para previsões ajustex_out.19 ##
accuracy(ps.1_18,  h.1_18)
accuracy(ps.3_18,  h.3_18)
accuracy(ps.6_18,  h.6_18)
accuracy(ps.12_18,  h.12_18)

############################################################
########### ajuste serie truncada abr.20#####################
ajuste.ten_abr.20 = Arima (icmspb_abr.20 , order =c(0,0,0),seasonal = list(order=c(
tsdiag(ajuste.ten_abr.20)# modemo não bem estimado

ajuste.saz_abr.20 = Arima (icmspb_abr.20 , order =c(0,1,1),xreg = sazonal.abr.20)##
tsdiag(ajuste.saz_abr.20)
```

```
ajusteAIC.pan_abr.20 = Arima (icmspb_abr.20 , order =c(1,1,1),seasonal = list(order
tsdiag(ajusteAIC.pan_abr.20)

ajusteBIC.pan_abr.20 = Arima (icmspb_abr.20 , order =c(0,1,1),seasonal = list(order
tsdiag(ajusteBIC.pan_abr.20)


# valores previstos para ajuste.saz_abr.20
ps.1_12=forecast(ajuste.saz_abr.20 ,h=1,xreg = sazonal.abr.20)
ps.3_12=forecast(ajuste.saz_abr.20, h=3,xreg = sazonal.abr.20)
ps.6_12=forecast(ajuste.saz_abr.20, h=6,xreg = sazonal.abr.20)
ps.12_12=forecast(ajuste.saz_abr.20,h=12,xreg = sazonal.abr.20)

# valores previstos para ajusteAIC.pan_abr.20
ppAIC.1_12=forecast(ajusteAIC.pan_abr.20 ,h=1,xreg = pandemia12)
ppAIC.3_12=forecast(ajusteAIC.pan_abr.20, h=3,xreg = pandemia12)
ppAIC.6_12=forecast(ajusteAIC.pan_abr.20, h=6,xreg = pandemia12)
ppAIC.12_12=forecast(ajusteAIC.pan_abr.20, h=12,xreg = pandemia12)

# valores previstos para ajusteBIC.pan_abr.20
ppBIC.1_12=forecast(ajusteBIC.pan_abr.20 ,h=1,xreg = pandemia12)
ppBIC.3_12=forecast(ajusteBIC.pan_abr.20, h=3,xreg = pandemia12)
ppBIC.6_12=forecast(ajusteBIC.pan_abr.20, h=6,xreg = pandemia12)
ppBIC.12_12=forecast(ajusteBIC.pan_abr.20, h=12,xreg = pandemia12)


ps.1_12=as.numeric(ps.1_12$mean[1])
ps.3_12=as.numeric(ps.3_12$mean[1:3])
ps.6_12=as.numeric(ps.6_12$mean[1:6])
ps.12_12=as.numeric(ps.12_12$mean[1:12])

ppAIC.1_12=as.numeric(ppAIC.1_12$mean[1])
```

```
ppAIC.3_12=as.numeric(ppAIC.3_12$mean[1:3])
ppAIC.6_12=as.numeric(ppAIC.6_12$mean[1:6])
ppAIC.12_12=as.numeric(ppAIC.12_12$mean[1:12])

ppBIC.1_12=as.numeric(ppBIC.1_12$mean[1])
ppBIC.3_12=as.numeric(ppBIC.3_12$mean[1:3])
ppBIC.6_12=as.numeric(ppBIC.6_12$mean[1:6])
ppBIC.12_12=as.numeric(ppBIC.12_12$mean[1:12])


## erros para previsões ajuste.saz_abr.20 ##
accuracy(ps.1_12,  h.1_12)
accuracy(ps.3_12,  h.3_12)
accuracy(ps.6_12,  h.6_12)
accuracy(ps.12_12,  h.12_12)

## erros para previsões ajusteAIC.pan_abr.20 ##
accuracy(ppAIC.1_12,  h.1_12)
accuracy(ppAIC.3_12,  h.3_12)
accuracy(ppAIC.6_12,  h.6_12)
accuracy(ppAIC.12_12,  h.12_12)

## erros para previsões ajusteBIC.pan_abr.20 ##
accuracy(ppBIC.1_12,  h.1_12)
accuracy(ppBIC.3_12,  h.3_12)
accuracy(ppBIC.6_12,  h.6_12)
accuracy(ppBIC.12_12,  h.12_12)

#############################################################
########### ajuste serie truncada out.20#####################
ajuste.ten_out.20 = Arima (icmspb_out.20 , order =c(0,0,0),seasonal = list(order=c(
tsdiag(ajuste.ten_out.20)# modemo não bem estimado
```

```
ajuste.saz_out.20 = Arima (icmspb_out.20 , order =c(0,1,1),xreg = sazonal.out.20)##
tsdiag(ajuste.saz_out.20)

ajusteAIC.pan_out.20 = Arima (icmspb_out.20 , order =c(1,1,1),seasonal = list(order
tsdiag(ajusteAIC.pan_out.20)

ajusteBIC.pan_out.20 = Arima (icmspb_out.20 , order =c(0,1,1),seasonal = list(order
tsdiag(ajusteBIC.pan_out.20)

# valores previstos para ajuste.saz_out.20
pz.1_6=forecast(ajuste.saz_out.20, h=1,xreg = sazonal.out.20)
pz.3_6=forecast(ajuste.saz_out.20, h=3,xreg = sazonal.out.20)
pz.6_6=forecast(ajuste.saz_out.20, h=6,xreg = sazonal.out.20)

# valores previstos para ajusteAIC.pan_out.20
ppAIC.1_6=forecast(ajusteAIC.pan_out.20, h=1,xreg = pandemia6)
ppAIC.3_6=forecast(ajusteAIC.pan_out.20, h=3,xreg = pandemia6)
ppAIC.6_6=forecast(ajusteAIC.pan_out.20, h=6,xreg = pandemia6)

# valores previstos para ajusteBIC.pan_out.20
ppBIC.1_6=forecast(ajusteBIC.pan_out.20, h=1,xreg = pandemia6)
ppBIC.3_6=forecast(ajusteBIC.pan_out.20, h=3,xreg = pandemia6)
ppBIC.6_6=forecast(ajusteBIC.pan_out.20, h=6,xreg = pandemia6)

pz.1_6=as.numeric(pz.1_6$mean[1])
pz.3_6=as.numeric(pz.3_6$mean[1:3])
pz.6_6=as.numeric(pz.6_6$mean[1:6])

ppAIC.1_6=as.numeric(ppAIC.1_6$mean[1])
ppAIC.3_6=as.numeric(ppAIC.3_6$mean[1:3])
ppAIC.6_6=as.numeric(ppAIC.6_6$mean[1:6])
```

```
ppBIC.1_6=as.numeric(ppBIC.1_6$mean[1])
ppBIC.3_6=as.numeric(ppBIC.3_6$mean[1:3])
ppBIC.6_6=as.numeric(ppBIC.6_6$mean[1:6])

## erros para previsões ajuste.saz_out.20 ##
accuracy(pz.1_6,  h.1_6)
accuracy(pz.3_6,  h.3_6)
accuracy(pz.6_6,  h.6_6)

## erros para previsões ajusteAIC.pan_out.20 ##
accuracy(ppAIC.1_6,  h.1_6)
accuracy(ppAIC.3_6,  h.3_6)
accuracy(ppAIC.6_6,  h.6_6)

## erros para previsões ajusteBIC.pan_out.20 ##
accuracy(ppBIC.1_6,  h.1_6)
accuracy(ppBIC.3_6,  h.3_6)
accuracy(ppBIC.6_6,  h.6_6)

############## Gráfico das séries selecionadas com os limites superior e inferior 95

##### para série icmspb_abr.19 #############

autoplot(icmspb_abr.20,xlab = "Tempo",ylab = "ICMS-PB")+
autolayer(forecast(ajuste.2_abr.19, h=12,c(95)),col="#6eeb34",lwd=0.75)+ ### SARIMA
autolayer(icmspb_abr.20,colour =0,lwd=0.71)


##### para série icmspb_out.19 #############

autoplot(icmspb_out.20,xlab = "Tempo",ylab = "ICMS-PB")+
```

```
autolayer(forecast(hwm.out.19, h=12,c(95)),col="#6eeb34",lwd=0.75,lty=1)+ ### HW Mu
autolayer(icmspb_out.20,colour =0,lwd=0.71)

##### para série icmspb_abr.20 ##############

autoplot(icmspb_d,xlab = "Tempo",ylab = "ICMS-PB")+
autolayer(forecast(ajuste.2_abr.20, h=12,c(95)),col="#6eeb34",lwd=0.75)+ ### SARIMA
autolayer(icmspb_d,colour =0,lwd=0.71)

##### para série icmspb_out.20 ##############

autoplot(icmspb_d,xlab = "Tempo",ylab = "ICMS-PB")+
autolayer(forecast(hw.out.20, h=6,c(95)),col="#6eeb34",lwd=0.75)+ ### HW Aditivo
autolayer(icmspb_d,colour =0,lwd=0.71)
```