



TEXTOS PARA DISCUSSÃO

SECRETARIA DE PLANEJAMENTO, GOVERNANÇA E GESTÃO
SUBSECRETARIA DE PLANEJAMENTO
DEPARTAMENTO DE ECONOMIA E ESTATÍSTICA

N.º 8
EXCESSO DE ÓBITOS NO BRASIL E NAS SUAS REGIÕES

Bruno Paim e Marilyn Agranonik

EXCESSO DE ÓBITOS NO BRASIL E NAS SUAS REGIÕES

Bruno Paim
Marilyn Agranonik

Porto Alegre
Maio de 2023

GOVERNO DO ESTADO DO RIO GRANDE DO SUL

Governador: Eduardo Leite

Vice-Governador: Gabriel Vieira de Souza

SECRETARIA DE PLANEJAMENTO, GOVERNANÇA E GESTÃO

Secretária: Danielle Calazans

Secretário Adjunto: Bruno Silveira

SUBSECRETARIA DE PLANEJAMENTO

Subsecretária: Carolina Mór Scarparo

Subsecretário Adjunto: Alessandro Castilhos Martins

DEPARTAMENTO DE ECONOMIA E ESTATÍSTICA

Diretor: Pedro Tonon Zuanazzi

Diretor Adjunto: Rodrigo Daniel Feix

Divisão de Análise de Políticas Sociais: Tomás Pinheiro Fiori

Divisão de Análise Econômica: Martinho Roberto Lazzari

Divisão de Dados e Indicadores: Fernando Ioannides Lopes da Cruz

Divisão de Estudos de Atividades Produtivas: Sérgio Leusin Jr.

TEXTOS PARA DISCUSSÃO

Publicação seriada cujo objetivo é divulgar os estudos e as pesquisas em desenvolvimento no Departamento de Economia e Estatística, com vistas a fomentar o debate e oferecer subsídios à formulação e à avaliação de políticas públicas.

Textos para Discussão DEE/SPGG / Secretaria de Planejamento, Governança e Gestão, Departamento de Economia e Estatística. – Porto Alegre :
Secretaria de Planejamento, Governança e Gestão, 2020- .

1. Condições econômicas – Rio Grande do Sul. 2. Economia – Rio Grande do Sul. I. Rio Grande do Sul. Secretaria de Planejamento, Governança e Gestão. Departamento de Economia e Estatística.

CDU 338.1(816.5)

Bibliotecário responsável: João Vítor Ditter Wallauer – CRB 10/2016

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade do(s) autor(es), não exprimindo, necessariamente, o ponto de vista do Departamento de Economia e Estatística.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

<https://dee.rs.gov.br/textos-discussao>

Revisão técnica: Eduardo de Oliveira Correa e Fernando Ioannides Lopes da Cruz

Revisão de Língua Portuguesa e editoração: Susana Kerschner

Projeto gráfico: Vinicius Ximendes Lopes

COMO REFERENCIAR ESTE TRABALHO:

PAIM, B.; AGRANONIK, M. **Excesso de óbitos no Brasil e nas suas regiões**. Porto Alegre: Secretaria de Planejamento, Governança e Gestão, Departamento de Economia e Estatística, 2023. (Textos para discussão, n. 8).

Secretaria de Planejamento, Governança e Gestão/Departamento de Economia e Estatística (SPGG/DEE)
Av. Borges de Medeiros, 1501 - 10.º andar, Porto Alegre, RS — CEP 90119-900

EXCESSO DE ÓBITOS NO BRASIL E NAS SUAS REGIÕES*

Bruno Paim**

Marilyn Agranonik***

Resumo

O objetivo do presente trabalho é propor uma metodologia para cálculo do excesso de óbitos por causas naturais, causado direta ou indiretamente pela COVID-19, durante o período de março de 2020 a dezembro de 2021, no Brasil e em suas grandes regiões. Para tanto, utilizam-se dados mensais do Sistema de Informações sobre Mortalidade (SIM), fornecidos pelo Departamento de Informática do Sistema Único de Saúde (Datasus) do Ministério da Saúde. Após análise desses dados, propõe-se o cálculo da previsão de óbitos por causas naturais para as regiões com base em três modelos principais: ETS, SNAIVE e ARIMA, ou SARIMA, conforme identificação ou não de sazonalidade na localidade. A escolha do modelo procede com o menor erro absoluto médio (MAE), calculado dentro da amostra. Por fim, calcula-se o excesso de óbitos com base na diferença entre os óbitos observados e os esperados, para o Brasil e suas regiões.

Palavras-chave: excesso de óbitos; COVID-19; Brasil; regiões do Brasil

Abstract

The objective of the present work is to propose a methodology for calculating the excess of deaths from natural causes caused directly or indirectly by COVID-19 during the period from March 2020 to December 2021 in Brazil and Regions. The work uses monthly data from the Mortality Information System (SIM), provided by the Department of Informatics of the Unified Health System (DATASUS) of the Ministry of Health. After analyzing these data, it is proposed to calculate the prediction of deaths from natural causes for regions based on 2 main models: ETS, SNAIVE and ARIMA, or SARIMA, depending on whether seasonality is identified in the location. The choice of the best model proceeds with the lowest mean absolute error (MAE), calculated within the sample. Finally, it calculates the excess deaths based on the difference between observed and expected deaths, for Brazil and regions.

Keywords: excess deaths; COVID-19; Brazil; regions of Brazil

* Os autores agradecem aos integrantes do Grupo de Trabalho Saúde do Comitê de Dados para enfrentamento da pandemia de COVID-19 no Estado do Rio Grande do Sul pelas críticas e sugestões a uma versão preliminar do trabalho, sem implicá-los pelos erros e omissões por acaso ainda remanescentes.
Os dados podem ser consultados de forma interativa em <http://boletinscoronavirus.dee.rs.gov.br/>.

** Pesquisador em Economia do Departamento de Economia e Estatística da Secretaria de Planejamento, Governança e Gestão e doutorando em Economia do Desenvolvimento na Universidade Federal do Rio Grande do Sul.

*** Pesquisadora em Estatística do Departamento de Economia e Estatística da Secretaria de Planejamento, Governança e Gestão e Doutora em Saúde da Criança e do Adolescente na Universidade Federal do Rio Grande do Sul.

1 Introdução

Desde o início da pandemia de COVID-19, identificar óbitos específicos por essa doença tem sido um desafio, devido, principalmente, à dificuldade de testagem. Estudos têm indicado subnotificação tanto dos casos como dos óbitos (FRANÇA *et al.*, 2020). Estimativas precisas da mortalidade por COVID-19 são essenciais para a adequada vigilância da pandemia e para a consolidação dos impactos da doença sobre a saúde pública.

Devido à dificuldade de medir a magnitude da mortalidade por COVID-19, diversos autores têm sugerido a utilização do excesso de óbitos ocorridos em um determinado período, por ser um indicador objetivo e comparável para avaliar o impacto de pandemias sobre a mortalidade (PEARCE; LAWLOR; BRICKLEY, 2020; VIEIRA *et al.*, 2020; SETEL *et al.*, 2020; MARINHO *et al.*, 2020). Apesar de indicadores mais específicos sobre a doença serem imprescindíveis para um bom monitoramento da situação, o excesso de mortalidade representa um indicador relativamente simples e robusto, uma vez que é possível ser feita a contagem do número de óbitos que estão ocorrendo, independentemente da precisão da causa de morte (MARINHO *et al.*, 2020). Além disso, a avaliação do excesso de mortes permite que sejam estimados os efeitos direto e indireto da epidemia sobre a mortalidade, constituindo uma ferramenta importante para o processo de tomada de decisão em suas diferentes fases (SETEL *et al.*, 2020).

Para fins de comparação entre as regiões, é necessário considerar os diferenciais de mortalidade já existentes. Assim, o cálculo do excesso de óbitos permite avançar sobre alguns problemas enfrentados, na tentativa de atribuir o impacto da doença sobre a mortalidade e na comparação das taxas de mortalidade entre as regiões, que são:

- a) as dificuldades na identificação da doença, agravadas em momentos em que o atendimento hospitalar passava por momentos críticos;
- b) o fato de que óbitos podem ter ocorrido diretamente por outras causas, mas como decorrência do esgotamento da capacidade hospitalar provocado pela COVID-19;
- c) o diferencial de taxas de mortalidade por motivos diversos apresentados pelas regiões (como perfil demográfico ou mesmo condições climáticas), que afetam a análise dos impactos que poderiam ser atribuídos à COVID-19.

O objetivo deste estudo é estimar o excesso de mortes para o Brasil e suas regiões, de acordo com o município de residência do falecido, no período de janeiro de 2020 a dezembro de 2021. Com base nos dados do Sistema de Informações sobre Mortalidade (SIM) fornecidos pelo Departamento de Informática do Sistema Único de Saúde (Datasus) do Ministério da Saúde, estima-se o número de óbitos mensais esperado para os anos de 2020 e 2021.

2 Qualidade dos dados do SIM

2.1 Cobertura do SIM

A cobertura do SIM é um indicador que busca verificar se o Sistema de Informações está sendo notificado sobre os dados a que se refere, e é definida como a razão entre óbitos coletados pelo SIM e óbitos projetados pelo Instituto Brasileiro de Geografia e Estatística (IBGE). O último relatório em âmbito nacional foi publicado em 2013 e refere-se aos dados de 2000 a 2011. Nesse período, houve consistente avanço da cobertura, com o Brasil atingindo 96,1% em 2011 e os estados da Região Sul com cobertura próxima a 100% (BRASIL, 2013).

2.2 Proporção de óbitos por causas mal definidas (CMD)

Proporções elevadas de óbitos por causas mal definidas¹ (CMD) indicam problemas de acesso e qualidade de assistência médica recebida pela população, além de comprometerem a fidedignidade das estatísticas de mortalidade por causas (CUNHA; TEIXEIRA; FRANÇA, 2017). O Ministério da Saúde tem indicado como meta, desde 2008, a redução da proporção de óbitos por CMD para menos de 10% (BRASIL, 2008). Apenas a Região Norte apresentou proporção de óbitos por CMD superior à meta nos anos iniciais do período analisado (Figura 1). As demais regiões apresentaram valores inferiores a 7,7% ao longo de todo período.

Figura 1

Proporção de óbitos por causas mal definidas no total de óbitos das regiões do Brasil — 2011-21

Região												Total	Δ %		
	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021		2011-19	2019-20	2019-21
Centro-Oeste	4,4%	3,1%	2,6%	2,7%	3,0%	3,1%	2,9%	2,5%	3,0%	2,9%	2,3%	2,9%	-32,2%	-1,6%	-24,2%
Nordeste	7,7%	7,4%	7,3%	7,4%	7,2%	7,2%	6,4%	5,9%	6,2%	6,6%	5,9%	6,8%	-19,2%	+5,1%	-5,4%
Norte	11,0%	10,7%	8,9%	8,2%	7,9%	7,8%	7,1%	7,3%	8,2%	7,9%	6,3%	8,1%	-25,5%	-3,8%	-23,5%
Sudeste	6,6%	6,2%	5,8%	5,7%	5,6%	5,7%	5,7%	5,8%	5,8%	6,2%	5,7%	5,9%	-11,9%	+5,6%	-1,9%
Sul	4,5%	4,6%	4,1%	3,8%	3,4%	3,9%	3,5%	3,5%	3,5%	3,6%	3,1%	3,7%	-22,5%	+1,5%	-12,7%
Total	6,7%	6,3%	5,9%	5,8%	5,7%	5,8%	5,5%	5,4%	5,6%	5,8%	5,1%	5,7%	-17,0%	+4,5%	-7,5%

Fonte: Datasus (BRASIL, 2023).

Diferenças entre regiões no período pré-pandemia

A proporção de CMD no total de óbitos apresentou diferença entre as regiões brasileiras ao longo de todo período. Os maiores percentuais foram encontrados na Região Norte, seguida por Região Nordeste e Região Sudeste. Essas regiões, em 2019, contavam com 8,2%, 6,2% e 5,8% de óbitos por CMD. Por outro lado, as menores proporções de óbitos por CMD foram observadas nas Regiões Sul e Centro-Oeste, com 3,5% e 3,0% respectivamente. Para todas as regiões, foi observado decréscimo na proporção de óbitos por CMD entre 2011 e 2019.

Diferenças entre regiões no período de pandemia, em comparação com 2019

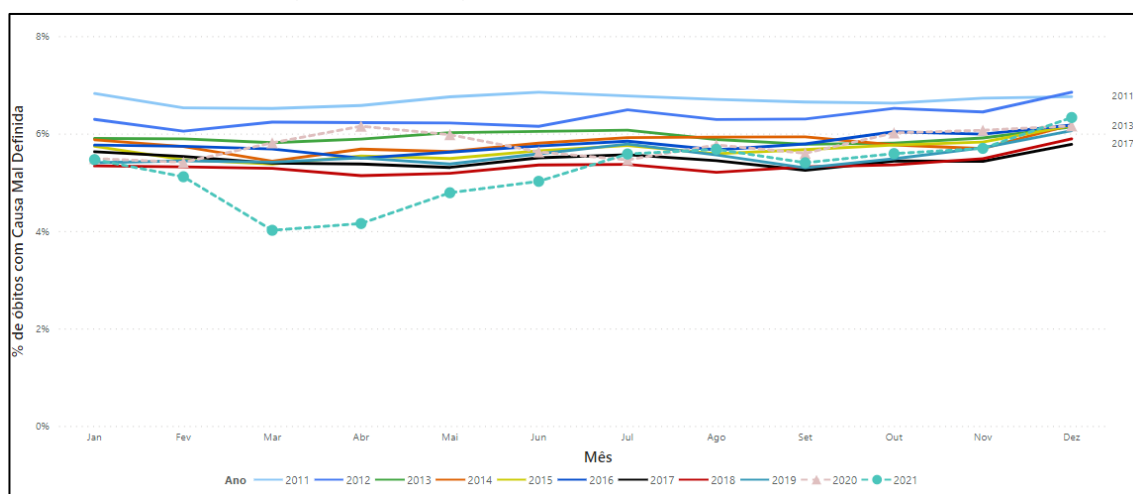
As Regiões Norte e Centro-Oeste mostraram uma pequena queda, seguida por uma queda maior no indicador. Na Região Nordeste, houve aumento entre 2019 e 2020 (5,1%). Já as Regiões Sudeste e Sul apresentaram os menores acréscimos nesses três anos. A ordem das regiões com menor proporção manteve-se a mesma ao longo de todo período avaliado (2011 a 2021).

A proporção mensal de óbitos por causas mal definidas indica um leve aumento nesse indicador nos meses finais de cada ano (Figura 2). Especificamente nos anos de pandemia, nota-se um comportamento diferenciado: (a) em 2020, houve aumento na proporção de óbitos por CMD no início da pandemia (março e abril), seguido por queda e novo aumento a partir de julho; (b) em 2021, nos meses com pico de casos de COVID-19, houve decréscimo da proporção de óbitos por CMD, atingindo o menor valor da série (4,0%), seguido por um aumento até o final do ano.

¹ Os óbitos por causas mal definidas (CMD) correspondem ao capítulo XVIII da Classificação Internacional de Doenças (CID-10).

Figura 2

Proporção de óbitos por causas mal definidas no Brasil — 2011-21



Fonte: Datasus (BRASIL, 2023).

3 Métodos

3.1 Base de dados

O estudo utiliza dados secundários de óbitos por causas naturais, de acordo com o município de residência do falecido. Os dados de mortalidade, no período de 2011 a 2021, foram obtidos do Sistema de Informações sobre Mortalidade do Departamento de Informática do SUS². O processo de coleta e pré-processamento dos dados utilizou o pacote de dados *microdatasus* (versão 2.1.0) (SALDANHA; BASTOS; BARCELLOS, 2019) desenvolvido para o *software* R (versão 4.2.2)³. Foram incluídas apenas as causas naturais⁴ nas análises, sendo excluídos os óbitos por causas externas (capítulo XX da Classificação Internacional de Doenças (CID-10), com códigos variando de V01 a Y98).

O SIM foi implantado em 1975 e seus dados provêm da Declaração de Óbito (DO), documento padronizado nacionalmente e distribuído pelo Ministério da Saúde. A DO inclui dados sociodemográficos do falecido e das condições e causas da ocorrência do óbito. Desde 1996, as causas de óbito baseiam-se na 10.^a Revisão da Classificação Internacional de Doenças (CID-10) (OMS, 1995), e os dados do SIM são disponibilizados no *site* do Datasus com, no máximo, dois anos de defasagem (BRASIL, 2011). As informações da DO são digitalizadas pelos municípios e transferidas para as secretarias de saúde estaduais, para que sejam feitas análises de consistência, correções e identificação de problemas na codificação das causas de morte. Após a validação dos estados, a base é transferida ao Ministério da Saúde, que, novamente, aplica uma validação final através de rotinas de controle da qualidade do dado, podendo solicitar esclarecimentos e até mesmo investigações de óbitos aos municípios e estados quando necessário.

² Disponível em: <https://datasus.saude.gov.br/>. Acesso em: 29 abr. 2023.

³ O **Apêndice II** apresenta as informações completas sobre a versão do R utilizada. Para maiores informações, entrar em contato com os autores.

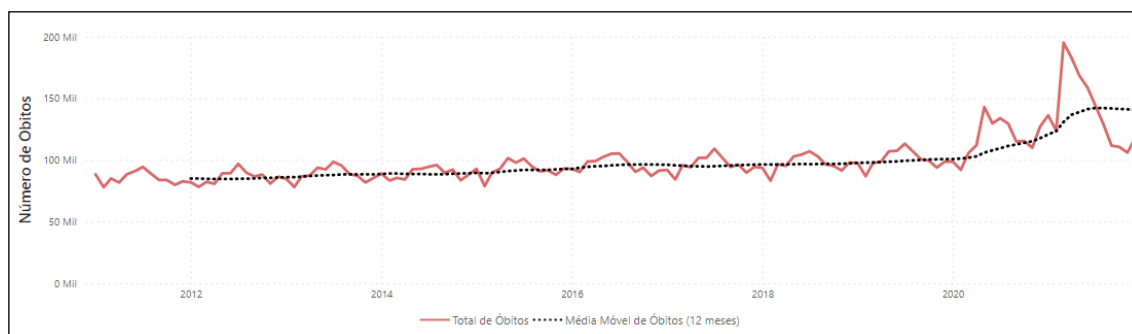
⁴ Embora seja possível utilizar também os óbitos por causas não naturais, violentos em sua maioria, para calcular o excesso de óbitos sob uma perspectiva mais ampla de impacto sobre a sociedade, por redução seja da criminalidade, seja de acidentes devido à menor circulação decorrente das restrições de mobilidade, preferiu-se manter a perspectiva estrita de impactos mais proximamente relacionados à saúde pública.

3.2 Óbitos por causas naturais no Brasil de 2011 a 2021

O número de óbitos por causas naturais no Brasil apresenta crescimento ao longo do período analisado. Tal comportamento é esperado tanto devido ao crescimento quanto ao gradual envelhecimento da população. A Figura 3 mostra o número de óbitos no País ao longo do tempo.

Figura 3

Número de óbitos por causas naturais e média móvel de 12 meses no Brasil — 2011-21

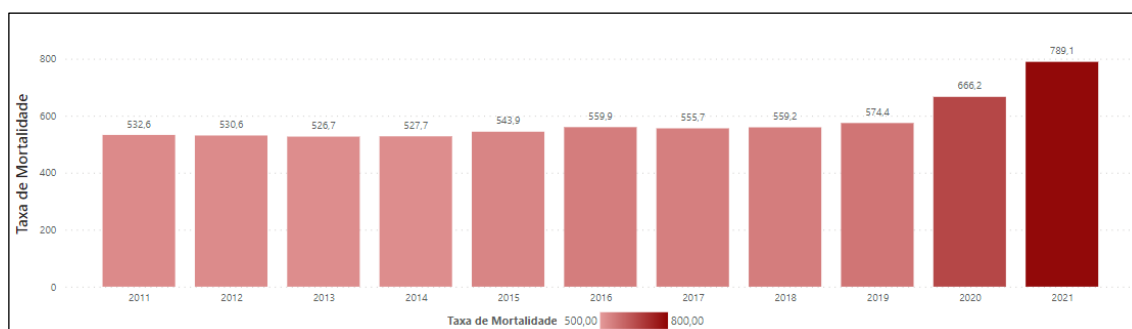


Fonte: Datasus (BRASIL, 2023).

A taxa de mortalidade por causas naturais apresenta menor variação ao longo do tempo, justamente por compensar o efeito do crescimento da população, embora também apresente leve aumento até 2019. Enquanto a média móvel de óbitos aumentou 17,8% de 2011 a 2019, a taxa de mortalidade variou 7,8% no mesmo período. Já a análise e a decomposição da elevação da mortalidade no total do País e nas suas regiões, para os anos de 2020 e 2021, são objeto deste estudo.

Figura 4

Taxa de mortalidade (por 100.000 habitantes) devido a causas naturais no Brasil — 2011-21



Fonte: Datasus (BRASIL, 2023).

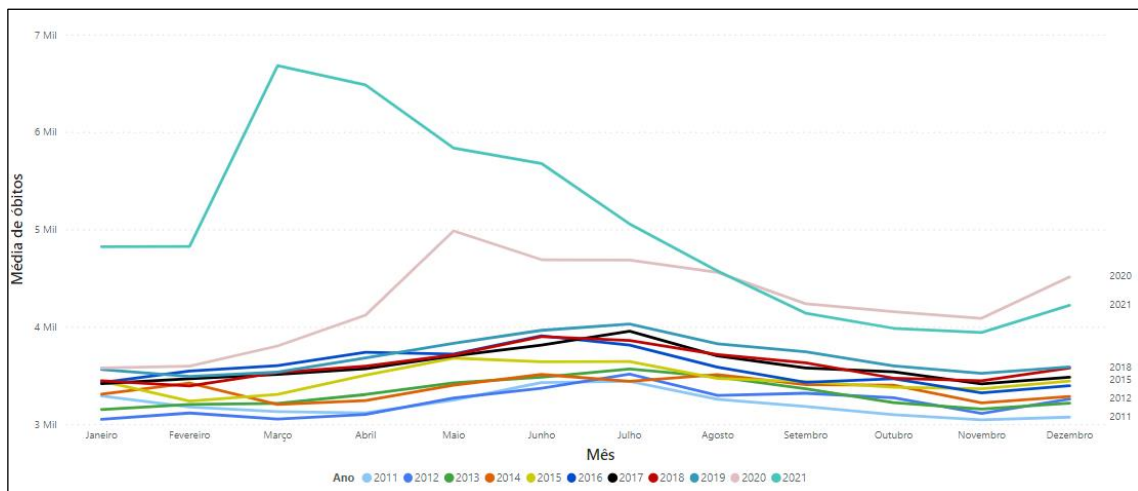
Projeções da população (IBGE, 2022).

Diversos trabalhos (MARINHO *et al.*, 2020; SANTOS *et al.*, 2021; ORELLANA *et al.*, 2021) têm utilizado a data da semana epidemiológica do óbito para fins de modelagem e cálculo do excesso de óbitos. No presente trabalho, optou-se pelos dados mensais, dividindo-se o total de óbitos em cada mês pelo número de dias, para se chegar a uma média mensal de óbitos diários. Essa padronização, além de facilitar a comparação direta ao longo dos meses, também acomoda o tratamento dos anos bissextos, como foi o caso de 2012, 2016 e 2020.

A visualização gráfica dos dados mensais de óbitos por causas naturais já indica que, no total para o Brasil, há presença de sazonalidade que eleva a mortalidade nos meses de inverno (Figura 5). Além disso, percebe-se que as linhas anuais são ascendentes, o que indica novamente uma tendência de crescimento dos óbitos ao longo do tempo.

Figura 5

Média mensal de óbitos diários no Brasil — 2011-21



Fonte: Datasus (BRASIL, 2023).

3.3 Modelos

A seguir, apresentam-se os quatro modelos utilizados tanto para analisar as séries como para projetar o número de óbitos esperados no Brasil e nas regiões, nos anos de 2020 e 2021.

ARIMA e SARIMA

Seguindo a metodologia de Box-Jenkins (BOX; JENKINS; REINSEL, 2008), tentou-se gerar um modelo Sazonal Autorregressivo Integrado de Médias Móveis — *Seasonal Autoregressive Integrated Moving Average* (SARIMA). O modelo ARIMA, sem sazonalidade, pode ser especificado por:

$$\phi_p(L)(1-L)^d Z_t = \theta_q(L)\varepsilon_t + \mu \quad (1)$$

Em que:

L = operador de defasagens;

ϕ_p = vetor dos coeficientes autorregressivos de ordem p ;

θ_q = vetor dos coeficientes de médias móveis de ordem q ;

d = número de diferenças aplicadas para tornar o modelo estacionário;

μ = intercepto;

ε_t = erro;

Z_t = série de dados.

Nesse caso, a metodologia empregada encontra os parâmetros p , d e q que melhor se ajustam aos dados. A denominação de cada modelo é dada por: ARIMA (p, d, q).

O modelo SARIMA (BOX; JENKINS; REINSEL, 2008, p. 353) inclui a possibilidade de **sazonalidade(s)** em sua especificação, adicionando os componentes (P, D, Q) nas defasagens que caracterizam a sazonalidade, passando a ser definido como:

$$\phi_P(L^s)\phi_p(L)(1-L)^d(1-L^s)^D Z_t = \theta_Q(L^s)\theta_q(L)\varepsilon_t + \mu \quad (2)$$

Em que:

L^s = operador de defasagens sazonais;

φ_p = vetor dos coeficientes autorregressivos sazonais de ordem P ;

θ_Q = vetor dos coeficientes de médias móveis sazonais de ordem Q ;

D = número de diferenças sazonais aplicadas para tornar o modelo estacionário;

s = periodicidade do componente sazonal.

Para denominação do modelo, usa-se a seguinte notação: SARIMA $(p, d, q) (P, D, Q)_s$.

A abordagem proposta por Box-Jenkins presentemente adotada propõe as seguintes etapas para especificação do modelo:

- diferenciar z_t em d e/ou D até atingir a estacionariedade da série. Neste trabalho, utilizaram-se os testes Dickey-Fuller Aumentado (ADF) e Kwiatkowski-Phillips-Schmidt-Shin (KPSS) para indicar a estacionariedade;
- inspecionar as Funções de Autocorrelação (ACF) e Autocorrelação Parcial (PACF) para escolher um modelo que se ajuste ao comportamento observado;
- estimar os coeficientes do modelo;
- analisar os resíduos do modelo, buscando a simplificação, sem perder poder explicativo. Neste trabalho, procedeu-se com teste Ljung-Box para analisar se os resíduos se comportavam como ruído-branco e optou-se pelo menor Critério de Informação Akaike Corrigido (AICc), que ajusta o AIC a amostras menores. É válido salientar que ambos os critérios, AIC e AICc, só podem ser utilizados para comparar modelos de níveis de integração d e D idênticos. Da mesma forma, só podem ser utilizados para comparação de uma mesma classe de modelos.

Decomposição STL

Na decomposição STL (CLEVELAND *et al.*, 1990), do inglês *Seasonal and Trend Decomposition using Loess*, considera-se o seguinte modelo:

$$y_t = T_t + S_t + R_t \quad (3)$$

Em que:

y_t = variável observada;

T_t = componente tendencial;

S_t = componente sazonal;

R_t = resíduo.

A partir da equação acima, é possível medir a intensidade do efeito da tendência e da sazonalidade sobre a variável observada. Para tanto, utiliza-se a relação da variância do resíduo com a variância de cada componente da seguinte forma⁵:

A intensidade do componente tendencial é dada por

$$I_T = \max \left(0 \quad ; \quad 1 - \frac{\text{Var}(R_t)}{\text{Var}(T_t + R_t)} \right) \quad (4),$$

⁵ Kang, Hyndman e Smith-Miles (2017) apresentam uma versão inicial dessa discussão. Em Hyndman e Athanasopoulos (2021), tem-se a elaboração aqui reproduzida.

de modo que, caso a intensidade do componente tendencial seja alta, a variância da série dessazonalizada ($Var(T_t + R_t)$) será maior em relação à variância do resíduo da série livre de sazonalidade e tendência ($Var(R_t)$), resultando em uma razão relativamente pequena. No caso oposto, em que o componente tendencial é baixo, as variâncias devem aproximar-se, ocasionando uma razão que resulta próxima à unidade, que por consequência reduz o máximo possível da função.

Simetricamente, a intensidade do componente **sazonal** é calculada por meio da relação entre a variância do resíduo da série livre de sazonalidade e tendência $Var(R_t)$ e a variância da série com a tendência retirada $Var(S_t + R_t)$. Com isso, a função que tem a base inferior limitada em zero e o máximo alcançado pelo segundo componente tende a reduzir quando a relação se aproxima de 1 e aumentar quando a razão

$$I_S = \max \left(0 ; 1 - \frac{Var(R_t)}{Var(S_t + R_t)} \right) \quad (5)$$

Embora a decomposição STL não seja utilizada como um modelo para previsão neste trabalho, ela irá auxiliar na caracterização das séries analisadas.

ETS

O modelo ETS, também denominado de Holt-Winters⁶, é um método de suavização exponencial para previsão de séries temporais para dados univariados. Tem por objetivo decompor a série em tendência, sazonalidade e resíduo, com o suplemento de permitir especificar individualmente a forma de interação de cada componente, se **aditivo**, **multiplicativo** ou **ausente**⁷. Considera que uma previsão é uma soma ponderada de observações passadas e usa um peso exponencialmente decrescente dessas observações. Ou seja, quanto mais recente a observação, maior será seu peso no modelo preditivo.

SNAIVE

O modelo de previsão SNAIVE, denominado Seasonal Naïve, consiste em repetir o mesmo valor do último período observado. Ao considerar a periodização utilizada, mensal, por exemplo, preserva-se a possível sazonalidade da série.

4 Especificação do Modelo

4.1 Análise dos dados e Teste de Estacionariedade

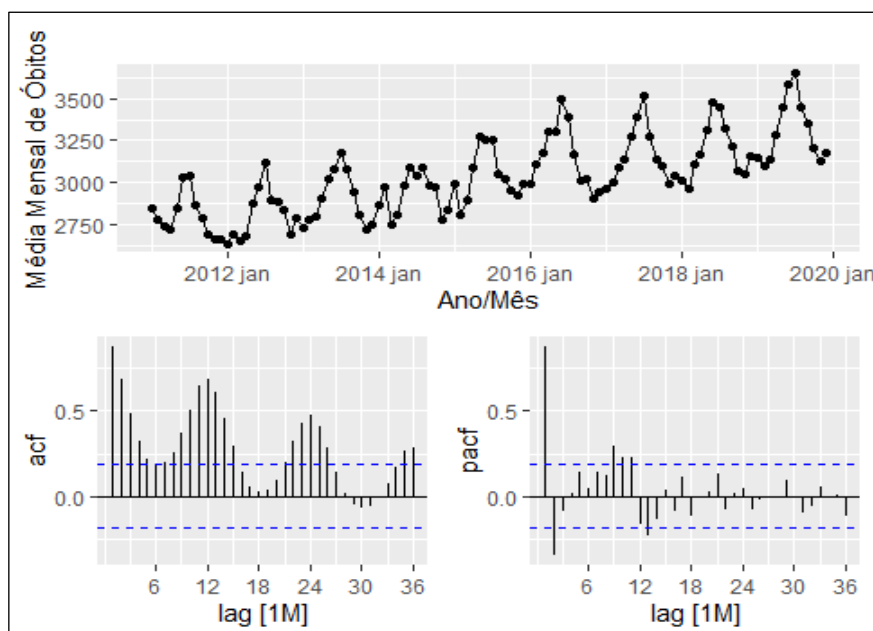
Utilizou-se a análise da série de média mensal de óbitos diários no Brasil de 2011 a 2019 para projetar os óbitos esperados em 2020 e 2021. Na Figura 6, percebem-se tanto o aumento do número de óbitos ao longo do tempo (parte superior) como a presença de maior autocorrelação em torno da defasagem 12 (parte inferior esquerda, ACF).

⁶ Holt (1957, republicado em 2004) e Winters (1960) desenvolveram respectivamente os métodos para tratamento da **tendência** e **sazonalidade** a partir da metodologia de amortecimento exponencial simples.

⁷ No caso específico da função ETS, do pacote *fable* (versão 0.3.2) para o *software* R, procede-se com uma identificação das características específicas da série analisada para apontar o modelo com menor critério AICc.

Figura 6

Média mensal de óbitos diários, Função Autocorrelação (ACF) e Função de Autocorrelação Parcial (PACF) para o Brasil — 2011-19

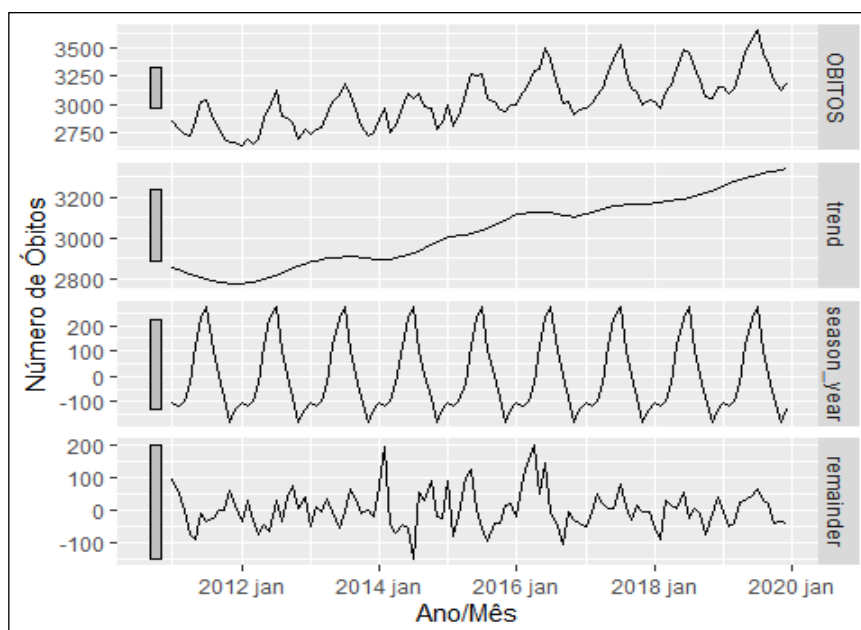


Fonte dos dados brutos: Datasus (BRASIL, 2023).

Como mostra a Figura 7, a decomposição STL apresenta a presença nítida de uma tendência (*trend*) e de uma flutuação sazonal anual (*season_year*). Já a Figura 8 mostra como essa elevação sazonal ocorre nos meses de inverno no País.

Figura 7

Decomposição STL do número de óbitos ocorridos no Brasil — 2011-19

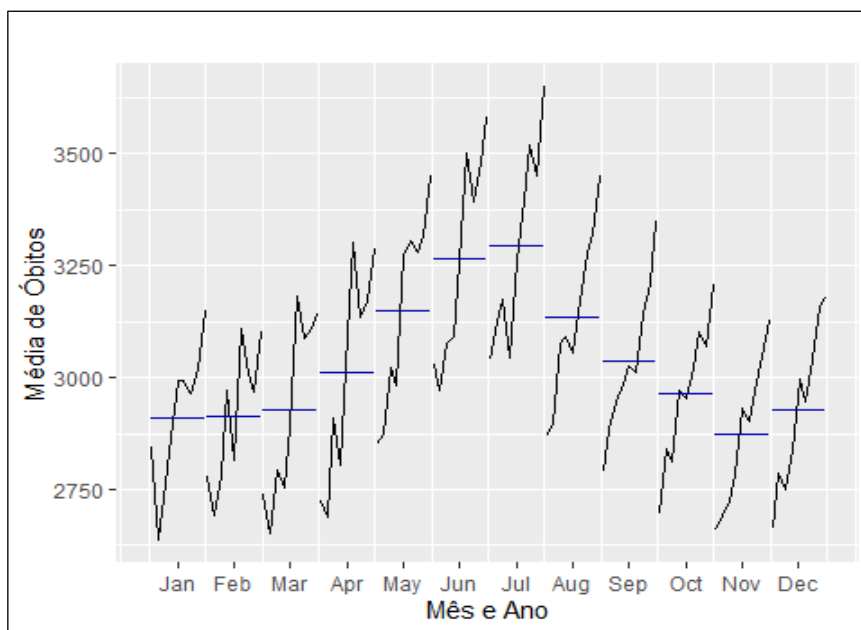


Fonte dos dados brutos: Datasus (BRASIL, 2023).

Nota: O gráfico apresenta escalas diferentes para cada elemento.

Figura 8

Série anual de óbitos mensais no Brasil — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

A presença de tendência na série indica a sua não estacionariedade, confirmada pelos testes Dickey-Fuller Aumentado⁸ e Kwiatkowski-Phillips-Schmidt-Shin (KPSS)⁹.

⁸ No ADF, a hipótese nula (H_0) é de que **há** raiz unitária, ou seja, que a série **não** é estacionária. A estatística do teste resultou em 0,0963, que, comparada com os valores críticos -2,58 a 1% de significância, indica que não é possível rejeitar a hipótese nula, ou seja, não é possível rejeitar a hipótese de que a série não é estacionária.

⁹ No KPSS, a hipótese nula (H_0) é de que **não** há raiz unitária, ou seja, que a série **é** estacionária. A estatística do teste foi calculada em 1,564154 e sinaliza que o p-valor é inferior a 1%, rejeitando-se a hipótese nula de que a série é estacionária.

Quadro 1

Teste Dickey-Fuller Aumentado (ADF) para o Brasil — 2011-19

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -242.948  -80.769   4.269   96.674  199.525
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1  0.0003429   0.0035613   0.096   0.923
##
## Residual standard error: 112 on 106 degrees of freedom
## Multiple R-squared:  8.745e-05, Adjusted R-squared: -0.009346
## F-statistic: 0.009271 on 1 and 106 DF,  p-value: 0.9235
##
##
## Value of test-statistic is: 0.0963
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Quadro 2

Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
para o Brasil — 2011-19

	x
kpss_stat	0.5000000
kpss_pvalue	0.0416667

Segundo Hyndman e Athanasopoulos (2021), quando é observada a presença de uma sazonalidade na série, realiza-se inicialmente a primeira diferença para a parcela sazonal ($D = 1$). Seguindo esse procedimento, obtém-se a estacionariedade da série, conforme testes ADF e KPSS¹⁰.

¹⁰ Neste caso, o teste ADF apresentou valor estatístico abaixo do valor crítico a 1%, rejeitando a hipótese nula de que há raiz unitária. No teste KPSS, o p-valor chegou a 0,1, podendo ser inclusive superior, indicando que a série é estacionária.

Quadro 3
Teste Dickey-Fuller Aumentado (ADF) em 1.^a diferença sazonal (D = 1)
para o Brasil — 2011-19

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##   Min     1Q   Median     3Q    Max
## -233.48 -34.01  34.35  78.99 293.57
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## z.lag.1 -0.39073    0.07982  -4.895 4.07e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 94.93 on 94 degrees of freedom
## Multiple R-squared:  0.2031, Adjusted R-squared:  0.1946
## F-statistic: 23.96 on 1 and 94 DF, p-value: 4.068e-06
##
##
## Value of test-statistic is: -4.895
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.6 -1.95 -1.61
```

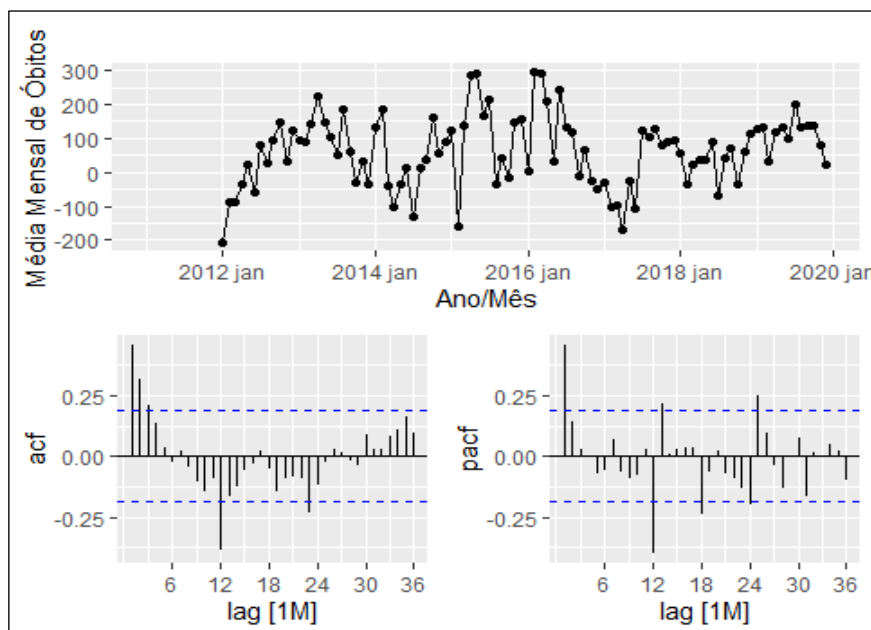
Quadro 4
Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS) em
1.^a diferença sazonal (D = 1) para o Brasil — 2011-19

	x
kpss_stat	0.1123108
kpss_pvalue	0.1000000

Com a primeira diferença sazonal, obtêm-se os seguintes resíduos (Figura 9):

Figura 9

Análise dos resíduos da média diária de óbitos em diferença sazonal ($D = 1$), Função Autocorrelação (ACF) e Função de Autocorrelação Parcial (PACF) dos óbitos no Brasil — 2011-19



4.2 Modelagem

Após aplicar a primeira diferença sazonal (defasagem 12), o comportamento da série aparenta ter eliminado a tendência apresentada anteriormente, como indicado pela estacionariedade testada no ADF e no KPSS. Na parcela não sazonal, a partir da análise da Função de Autocorrelação (FAC), observa-se uma redução gradual enquanto na Função de Autocorrelação (FACP) há somente o primeiro pico, indicando um AR(1). Para a parcela sazonal, nas defasagens múltiplas de 12, observa-se que a FACP apresenta queda gradual, enquanto a FAC apresenta a primeira defasagem significativa e a vigésima terceira. Nesse caso, sugere-se o modelo SARIMA(1,0,0)(0,1,1)₁₂, com deslocamento. Também se utilizou a **função ARIMA**, do pacote *fable* (HYNDMAN; ATHANASOPOULOS, 2021) que indicou o modelo SARIMA(1,0,0)(2,1,0)₁₂, com deslocamento¹¹.

Adicionaram-se também a estimativa utilizando a decomposição STL, já apresentada, e os modelos ETS e SNAIVE, para fins de comparação. O Quadro 5 lista os critérios de informação cada um dos modelos:

Quadro 5

Critérios de informação dos modelos utilizados
para o Brasil — 2011-19

Model name	Orders
auto_arima	<ARIMA(1,0,0)(2,1,0)[12] w/ drift>
arima100011	<ARIMA(1,0,0)(0,1,1)[12] w/ drift>
ets	<ETS(A,N,A)>
stl	
snaive	

¹¹ Pela **função ARIMA**, é possível realizar a busca pelo modelo mais ajustado procedendo ou não com um algoritmo que otimiza essa busca (chamado de *STEP*). Ativando ou não esse dispositivo, chegou-se ao mesmo resultado SARIMA(1,0,0)(2,1,0)₁₂.

Quadro 6

Análise dos resíduos de treinamento, ordenados por erro absoluto médio (MAE), para o Brasil — 2011-19

.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
arima100011	Training	-1.412	60.456	45.054	-0.097	1.485	0.449	0.498	-0.028
auto_arima	Training	-0.590	71.024	51.729	-0.074	1.705	0.516	0.585	-0.046
ets	Training	4.845	65.630	50.957	0.119	1.688	0.508	0.541	0.038
stl	Training	4.407	60.053	44.721	0.108	1.471	0.446	0.495	0.346
snaive	Training	62.315	121.401	100.282	1.958	3.256	1.000	1.000	0.454

Quadro 7

Análise dos critérios de informação dos modelos efetivos, ordenados por Critério de Informação Akaike Corrigido (AICc), para o Brasil — 2011-19

.model	sigma2	log_lik	AIC	AICc	BIC
arima100011	4244.399	-548.988	1105.976	1106.416	1116.234
auto_arima	5921.743	-553.980	1117.960	1118.627	1130.782
ets	4948.837	-704.711	1439.422	1444.639	1479.654

Sob todos os critérios, o modelo SARIMA(1,0,0)(0,1,1)₁₂ (ARIMA100011) apresentou os melhores resultados, com os menores valores de resíduos, bem como os menores AIC, AICc e BIC¹². No entanto, ao se analisarem os coeficientes do modelo, percebe-se que o coeficiente da parcela atribuída à média móvel sazonal (SMA(1)) está próximo à unidade, o que pode causar a instabilidade do modelo, como também pode ser observado no teste de raiz unitária na Figura 10.

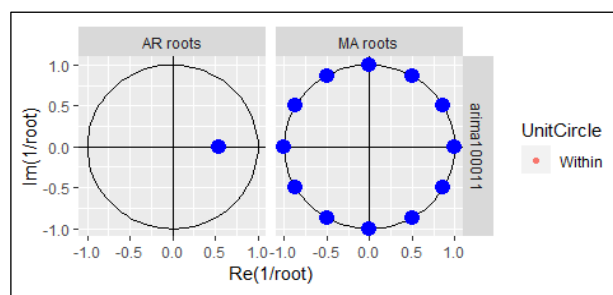
Quadro 8

Especificação do modelo ARIMA(1,0,0)(0,1,1)₁₂ para o Brasil — 2011-19

```
## Series: OBITOS
## Model: ARIMA(1,0,0)(0,1,1)[12] w/ drift
##
## Coefficients:
##   ar1   sma1 constant
##  0.5444 -0.9997 28.7079
## s.e. 0.0895 0.8669 2.3239
##
## sigma^2 estimated as 4244: log likelihood=-548.99
## AIC=1105.98 AICc=1106.42 BIC=1116.23
```

Figura 10

Inverso das raízes unitárias do modelo SARIMA(1,0,0)(0,1,1)₁₂ com deslocamento para o Brasil — 2011-19



¹² Salienta-se que os modelos SARIMA e ETS não são comparáveis por meio desses critérios, pois não pertencem à mesma classe de modelos.

Quadro 9

Teste Ljung-Box (lag = 36) do modelo
ARIMA(1,0,0)(0,1,1)₁₂ com três graus de
liberdade para o Brasil — 2011-19

.model	lb_stat	lb_pvalue
arima100011	27.39554	0.7421259

Desse modo, prosseguiu-se o exercício com o modelo sugerido pela função ARIMA do pacote *fable*, acima relatado, que resultou em um SARIMA (1,0,0)(2,1,0)₁₂ com deslocamento. A partir desse modelo, projetou-se o número de óbitos para 2020 e 2021, para cálculo posterior do excesso de óbitos no Brasil nesse período. Abaixo, segue a descrição completa dos coeficientes, bem como os resíduos obtidos a partir do modelo, os testes de raiz unitária e da aplicação do teste Ljung-Box com 36 defasagens¹³.

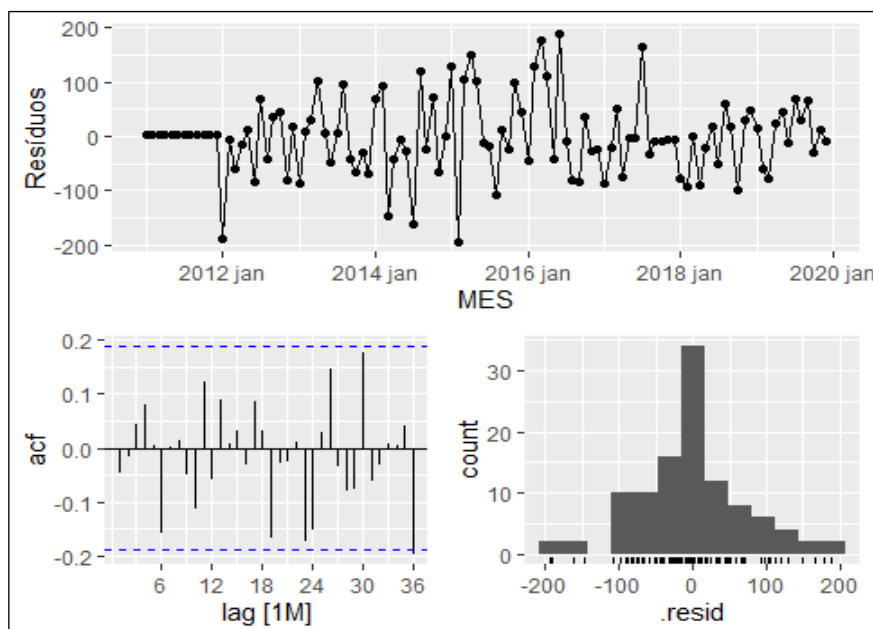
Quadro 10

Especificação do modelo ARIMA(1,0,0)(2,1,0)₁₂ para o Brasil — 2011-19

```
## Series: OBITOS
## Model: ARIMA(1,0,0)(2,1,0)[12] w/ drift
##
## Coefficients:
##      ar1      sar1      sar2 constant
##      0.5281 -0.6323 -0.2902  55.6602
## s.e. 0.0912  0.0990  0.0977   8.2836
##
## sigma^2 estimated as 5922: log likelihood=-553.98
## AIC=1117.96 AICc=1118.63 BIC=1130.78
```

Figura 11

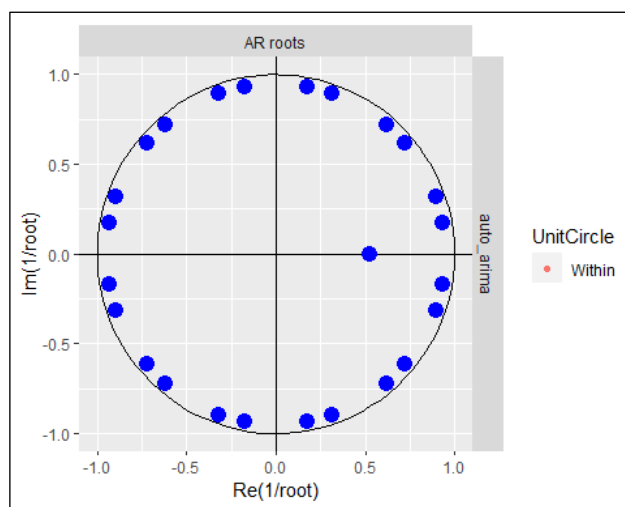
Análise dos resíduos do modelo SARIMA (1,0,0)(2,1,0)₁₂ com deslocamento, Função Autocorrelação (ACF) e histograma — 2011-19



¹³ No teste Ljung-Box, a hipótese nula é de que os resíduos são independentemente distribuídos. O p-valor do teste chegou a 0,17, indicando não ser possível rejeitar a hipótese nula a um nível de significância de 5% e considerando 36 lags para análise e 4 graus de liberdade (3 coeficientes e 1 constante).

Figura 12

Inverso das raízes unitárias do modelo SARIMA (1,0,0)(2,1,0)₁₂ com deslocamento para o Brasil — 2011-19



Quadro 11

Teste Ljung-Box (lag = 36) do modelo ARIMA(1,0,0)(2,1,0)₁₂ com quatro graus de liberdade para o Brasil — 2011-19

.model	lb_stat	lb_pvalue
auto_arima	39.41	0.1723166

4.3 Teste de previsão e escolha do modelo

Uma vez que o objetivo principal deste trabalho é encontrar a melhor previsão para o número de óbitos para o período subsequente, procedeu-se com o exercício de previsão dentro da amostra, visando a comparar todos os modelos estimados, segundo os critérios de análise dos erros. Para tanto, a amostra foi reduzida para o período de 2011 a 2017, com previsão para os dois anos subsequentes, portanto, antes dos efeitos da propagação por COVID-19 no País. Com isso, foi possível comparar as previsões de cada modelo com o número de óbitos observado em 2018 e 2019 (Quadro 12).

Quadro 12

Modelos de treinamento para o Brasil — 2011-17

Model name	Orders
arima100210	<ARIMA(1,0,0)(2,1,0)[12] w/ drift>
auto_arima	<ARIMA(1,0,0)(2,1,0)[12] w/ drift>
auto_arima_STEP	<ARIMA(1,0,0)(2,1,0)[12] w/ drift>
ets	<ETS(M,N,A)>
stl	
snaive	

O modelo ARIMA100210 representa o modelo especificado na seção anterior, porém com coeficientes diferentes, já que a amostra para fins de treinamento exclui os dados de 2018 e 2019¹⁴. O modelo

¹⁴ É interessante perceber que, nesse caso, o modelo selecionado pelo algoritmo da função ARIMA do pacote *fable*, quando se ativou o dispositivo de busca por etapa (*auto_arima_STEP*) foi novamente idêntico ao modelo selecionado sem esse dispositivo.

encontrado pelo algoritmo foi o mesmo encontrado na seção anterior, com a série completa de 2011 a 2019, do tipo SARIMA(1,0,0)(2,1,0)₁₂. O interessante desse exercício é que se demonstra certa robustez na determinação do modelo quando, mesmo retirando a parte final da série (anos de 2018 e 2019), não há alteração na especificação do modelo, embora exista diferença nos coeficientes estimados. Como os três procedimentos utilizados selecionaram o mesmo modelo, manteve-se apenas o ARIMA100210 para fins descritivos.

Os resultados estão apresentados no Quadro 13, ordenados pelo tipo de avaliação e pelo erro médio absoluto.

Quadro 13

Análise dos resíduos de treinamento ordenados por erro absoluto médio (MAE) para o Brasil — 2011-17

.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
stl	Training	5.727	65.233	48.362	0.135	1.604	0.458	0.507	0.363
arima100210	Training	-1.184	75.120	54.099	-0.103	1.807	0.512	0.584	-0.045
ets	Training	4.562	68.964	54.980	0.108	1.839	0.520	0.536	0.037
snaive	Training	58.909	128.673	105.688	1.879	3.490	1.000	1.000	0.459
arima100210	Test	-33.327	73.619	59.275	-1.078	1.851	0.561	0.572	0.562
ets	Test	81.132	116.776	97.544	2.400	2.948	0.923	0.908	0.693
snaive	Test	88.719	115.121	100.246	2.689	3.048	0.949	0.895	0.537

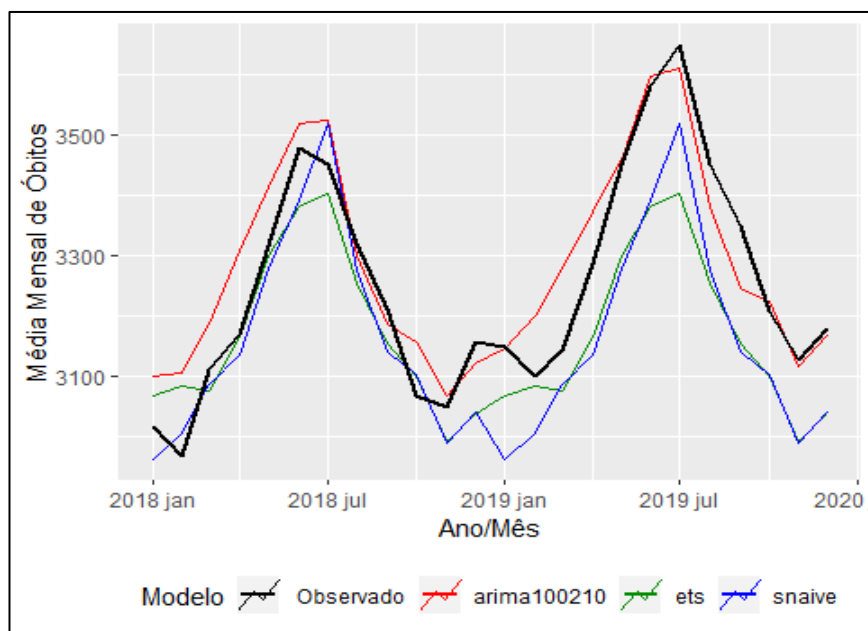
Os resultados denominados como *Training* indicam o ajuste dentro da amostra, ou seja, é uma análise dos resíduos do modelo e mostram seu poder explicativo. Com isso, é possível perceber que o modelo com melhor ajuste é o **STL**, seguido do **SARIMA**, do **ETS** e, por fim, do **SNAIVE**. No entanto, é válido salientar que o melhor ajuste não necessariamente indica a melhor capacidade de previsão, o que pode ser percebido nos resultados sob a denominação *Test*¹⁵, que resultam da diferença entre os óbitos previstos em cada modelo e o efetivamente observado no País em 2018 e 2019. Nesse caso, o menor MAE¹⁶ é do modelo ARIMA100210. A Figura 13 mostra a comparação entre os modelos SARIMA, ETS e SNAIVE, favorecendo o primeiro na comparação com o observado nesses dois anos. Percebe-se também como a distância entre o observado e o estimado pelo modelo ETS se amplia no segundo ano.

¹⁵ Observa-se que, nesses resultados, não foi utilizado o modelo STL, que não se propõe a ser um modelo para previsão, sendo simplesmente uma alternativa descritiva ao modelo, que, como mostrado anteriormente, decompõe a série em três partes: sazonalidade, tendência e resíduo.

¹⁶ Se observado o critério de menor erro quadrático médio (RMSE, da sigla em inglês *Root Mean Squared Error*), o mesmo modelo obteve menor valor.

Figura 13

Previsão dentro da amostra, utilizando três modelos, para a média mensal de óbitos no Brasil — 2018-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

5 Estimativas do excesso de óbitos para o Brasil

Com base no modelo escolhido, foi possível projetar os óbitos para o Brasil nos anos de 2020 e 2021, mantendo-se o intervalo de confiança (95%) indicado pelo próprio modelo. Com isso, é possível calcular o excesso de óbitos máximo e mínimo para o País. O excesso de óbitos é, portanto, calculado como:

$$\text{Excesso de Óbitos} = \text{Óbitos Observados} - \text{Óbitos Esperados}_{\text{pontual}}$$

$$\text{Excesso de Óbitos}_{\text{max}} = \text{Óbitos Observados} - \text{Óbitos Esperados}_{\text{min}}$$

$$\text{Excesso de Óbitos}_{\text{min}} = \text{Óbitos Observados} - \text{Óbitos Esperados}_{\text{max}}$$

Da mesma forma, procedeu-se ao cálculo do excesso proporcional de óbitos, medida que permite, de uma forma mais intuitiva, comparar quantos óbitos ocorreram acima do previsto. Para tanto, utilizou-se o cálculo a seguir:

$$\text{Excesso Proporcional de Óbitos} = \frac{\text{Óbitos Observados} - \text{Óbitos Esperados}_{\text{pontual}}}{\text{Óbitos Esperados}_{\text{pontual}}} \cdot 100\%$$

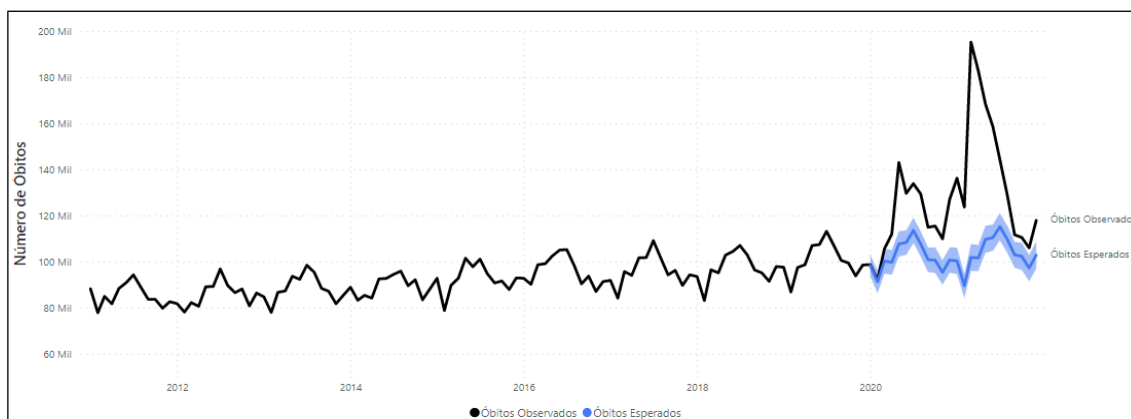
$$\text{Excesso Proporcional de Óbitos}_{\text{max}} = \frac{\text{Óbitos Observados} - \text{Óbitos Esperados}_{\text{min}}}{\text{Óbitos Esperados}_{\text{min}}} \cdot 100\%$$

$$\text{Excesso Proporcional de Óbitos}_{\text{min}} = \frac{\text{Óbitos Observados} - \text{Óbitos Esperados}_{\text{max}}}{\text{Óbitos Esperados}_{\text{max}}} \cdot 100\%$$

Na Figura 14, é possível perceber que os meses de janeiro e fevereiro de 2020 apresentam proximidade com o esperado, o que indica um bom ajuste do modelo, visto que os óbitos por COVID-19 iniciam período de crescimento somente em março no Brasil.

Figura 14

Total mensal de óbitos observados em 2011-21 e esperados em 2020-21 no Brasil



Fonte dos dados brutos: Datasus (BRASIL, 2023).

A Figura 15 mostra que, somente no mês de março de 2021, ocorreram, no País, cerca de 93 mil óbitos a mais do que o que seria esperado para aquele mês, ou seja, morreu cerca de o dobro do que foi previsto para o período. No acumulado de todo o período de março de 2020 e dezembro de 2021, houve, no País, 27,5% (cerca de 625 mil óbitos) a mais de óbitos do que seria esperado para o período, o intervalo variando entre 20,9% e 34,8%, quando considerado 95% de confiança.

Figura 15

Óbitos esperados, óbitos observados, excesso de óbitos e excesso proporcional de óbitos no Brasil — 2020-21

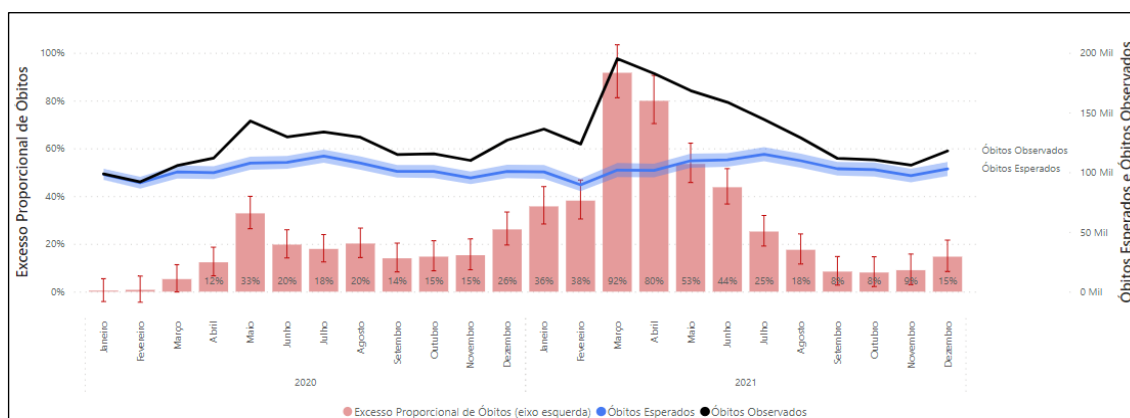
Ano	Óbitos Observados	Óbitos Esperados	Excesso de Óbitos	Excesso Proporcional de Óbitos	Excesso Proporcional de Óbitos (MIN)	Excesso Proporcional de Óbitos (MAX)
2020	1.220.153	1.034.382	185.771	18,0%	12,1%	24,5%
Março	105.513	100.240	5.273	5,3%	-0,2%	11,3%
Abril	111.892	99.644	12.248	12,3%	6,6%	18,6%
Mai	142.924	107.633	35.291	32,8%	26,3%	39,9%
Junho	129.593	108.257	21.336	19,7%	14,1%	25,9%
Julho	133.786	113.490	20.296	17,9%	12,4%	23,9%
Agosto	129.331	107.640	21.691	20,2%	14,3%	26,6%
Setembro	114.875	100.797	14.078	14,0%	8,2%	20,3%
Outubro	115.411	100.652	14.759	14,7%	8,7%	21,3%
Novembro	109.883	95.332	14.551	15,3%	9,2%	22,1%
Dezembro	126.945	100.697	26.248	26,1%	19,5%	33,4%
2021	1.683.327	1.243.190	440.137	35,4%	28,3%	43,4%
Janeiro	136.141	100.315	35.826	35,7%	28,3%	44,0%
Fevereiro	123.657	89.529	34.128	38,1%	30,4%	46,8%
Março	195.151	101.821	93.330	91,7%	81,2%	103,4%
Abril	182.652	101.522	81.130	79,9%	70,4%	90,6%
Mai	168.289	109.644	58.645	53,5%	45,7%	62,2%
Junho	158.629	110.367	48.262	43,7%	36,7%	51,5%
Julho	144.098	115.124	28.974	25,2%	19,1%	31,9%
Agosto	128.877	109.669	19.208	17,5%	11,5%	24,2%
Setembro	111.584	102.938	8.646	8,4%	2,7%	14,7%
Outubro	110.473	102.274	8.199	8,0%	2,2%	14,6%
Novembro	105.953	97.214	8.739	9,0%	3,0%	15,8%
Dezembro	117.823	102.773	15.050	14,6%	8,5%	21,6%
Total	2.903.480	2.277.572	625.908	27,5%	20,9%	34,8%

Fonte dos dados brutos: Datasus (BRASIL, 2023).

Na Figura 16, mostra-se como, somente em janeiro, fevereiro e março de 2020, a ausência de excesso de óbitos ficou dentro do intervalo de confiança da previsão realizada. Em todos os demais meses, a estimativa de excesso proporcional de óbitos ficou acima de 5%, com intervalo de confiança superando o zero absoluto a partir de abril, de forma que se pode dizer, com maior certeza, que houve excesso de óbitos no País desde abril de 2020 até o final de 2021.

Figura 16

Óbitos observados, óbitos esperados e excesso proporcional de óbitos no Brasil — 2020-21



Fonte dos dados brutos: Datasus (BRASIL, 2023).

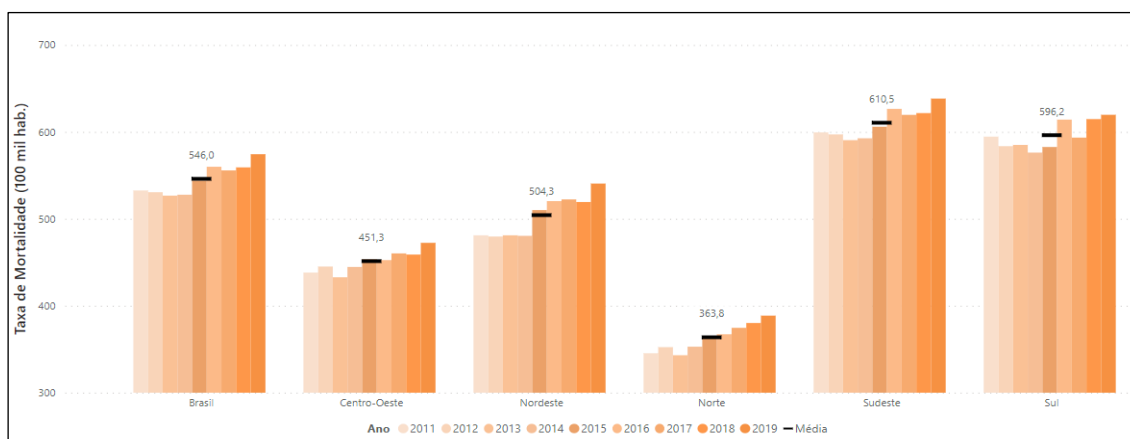
Nesta seção, desenvolveu-se a análise para o País como um todo. No entanto, é válido salientar que as diferenças regionais podem ter influência importante na interpretação desta análise. A seção seguinte, ao calcular e analisar o excesso de óbitos regionalmente, permite uma avaliação mais precisa da situação em cada localidade do Brasil.

6 Estimativa do excesso de óbitos para as regiões

Como observado anteriormente, as diferenças demográficas regionais repercutem também sobre a mortalidade por causas naturais, de modo que a Região Sudeste, que possuía a maior mortalidade média entre os anos de 2011 e 2019, apresenta uma taxa cerca de 65% superior à da Região Norte, a menor taxa no período entre as regiões (Figura 17).

Figura 17

Taxa de mortalidade por causas naturais, por regiões, do Brasil — 2011-2019



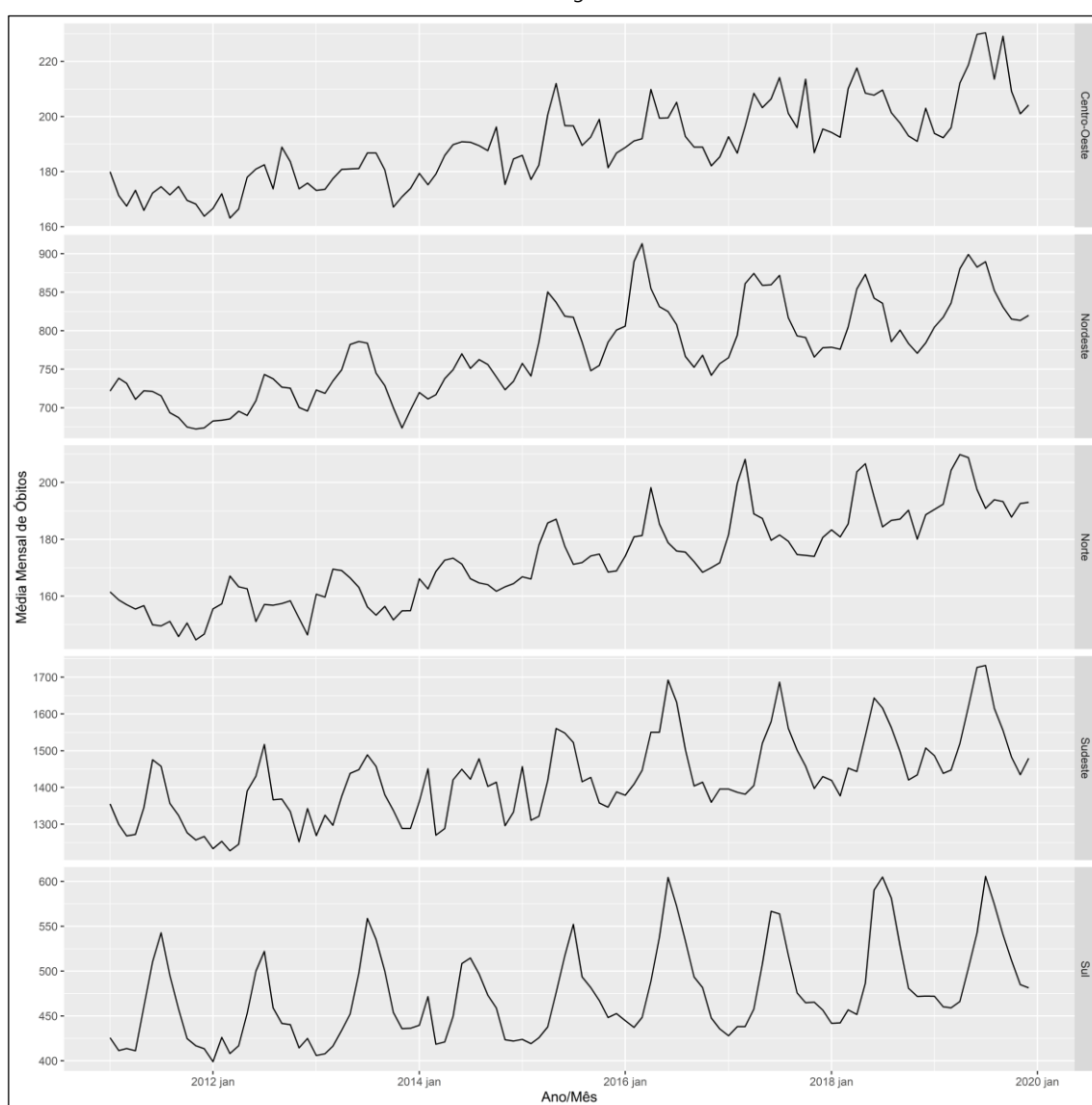
Fonte: Datasus (BRASIL, 2023).
Projeções da população (IBGE, 2022).

A evolução do número de óbitos nas regiões também difere na magnitude da tendência, no grau de intensidade da sazonalidade e no período do ano em que ocorre o aumento sazonal.

Na Figura 18, mostra-se o comportamento da média diária de óbitos mensais por região. Comparando o total anual de óbitos, a Região Norte apresentou um aumento de 29% no período de 2011 a 2019. Nesse mesmo período, a elevação na Região Sul foi de 13,4% e na Sudeste de 16,2%. Considerar essas diferenças regionais tem impacto no cálculo do excesso de óbitos, pois, conforme a localidade do óbito, ele pode ser considerado excesso ou não, visto que a probabilidade de que haveria, de qualquer forma, aumento em alguns estados é maior do que em outros, caso perdure esse movimento heterogêneo entre os estados.

Figura 18

Média mensal de óbitos nas regiões do Brasil — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

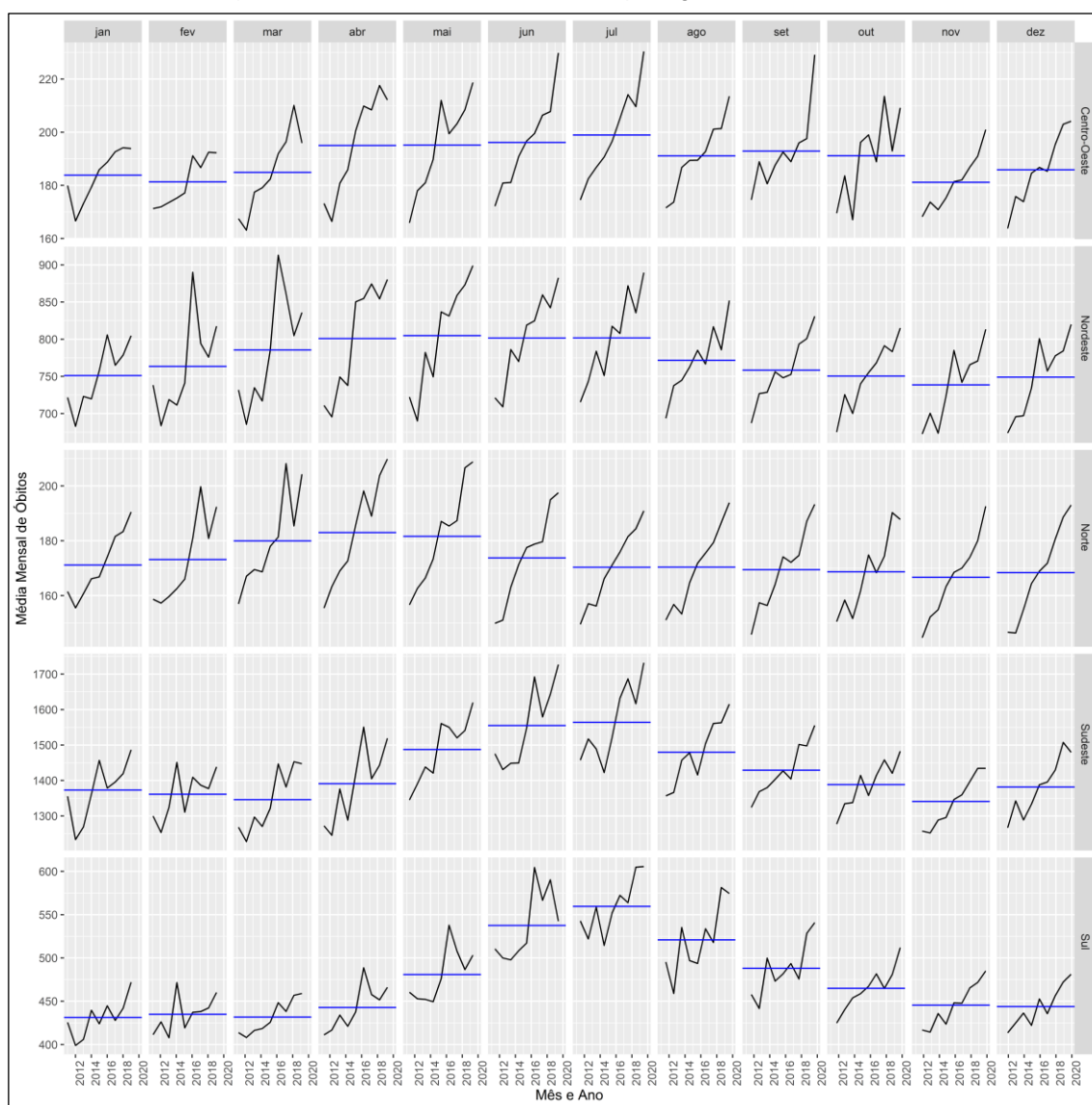
Nota: O gráfico apresenta escalas diferentes para cada região.

Na Figura 19, percebe-se que, na Região Nordeste e, principalmente, na Norte, a sazonalidade se dá nos meses de março, abril e maio. Diferentemente das demais regiões, que apresentam aumento, em intensidades diferentes, nos meses de inverno no País, o que, dada a maior participação no total de óbitos

do País, acaba por predominar também no comportamento nacional, como visto na seção anterior. Essa diferença no período da sazonalidade mostra também a vantagem de se calcular o excesso de óbitos especificamente para cada região. Também é possível perceber que a intensidade da sazonalidade é maior nas Regiões Sul e Sudeste, de forma que o impacto sobre o excesso também deve incorporar essa diferença regional.

Figura 19

Comportamento sazonal do número de óbitos, por região, no Brasil — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

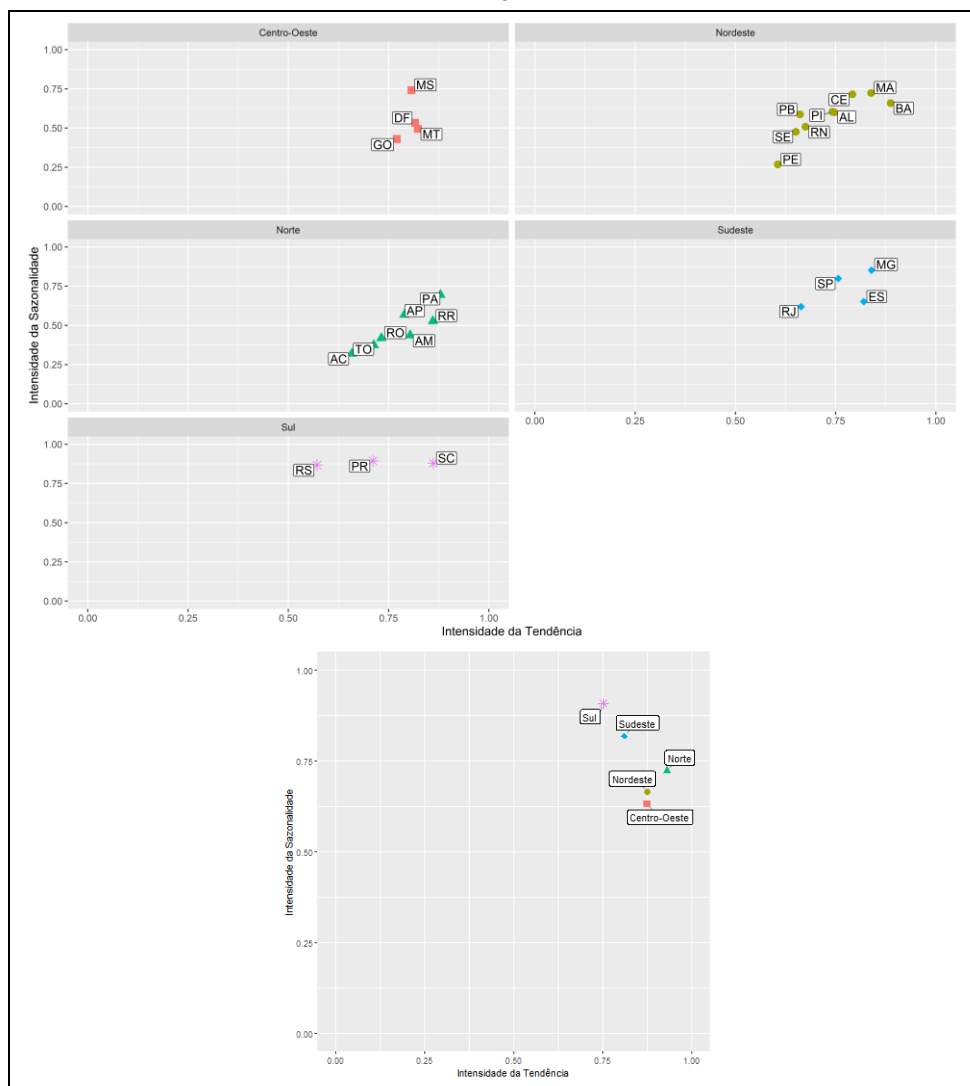
Nota: O gráfico apresenta escalas diferentes para cada região.

Consolida-se a informação das duas últimas figuras em uma só (Figura 20) por meio do cálculo da intensidade da tendência e da sazonalidade a partir da decomposição STL para cada unidade da Federação, agregadas por região, conforme anteriormente apresentado na seção 3.3. Pode-se perceber que o comportamento dos óbitos no País ao longo dos anos, percebido como tendência, e a dinâmica dentro de um ano, percebida como sazonalidade, é uma característica regional. A Figura 20 mostra que a Região Sul apresenta a maior intensidade do componente sazonal, enquanto os estados da Região Norte apresentam a maior intensidade do componente de tendência. Não se pretende, aqui, discutir especificamente os condicionan-

tes dessas características, mas apenas apontá-los, para mostrar a importância de se buscar modelos regionais que melhor se adaptem ao comportamento específico de cada localidade.

Figura 20

Intensidade da sazonalidade e da tendência do número de óbitos por causas naturais, por unidade da Federação e região, no Brasil — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Identificadas as diferenças acima, procedeu-se ao mesmo processo descrito anteriormente para o caso nacional, a fim de estimar a previsão de óbitos esperados e, posteriormente, calcular o excesso de óbitos regional:

- estimar modelos ETS, ARIMA ou SARIMA e SNAIVE, com base no período de 2011-17, gerando as previsões para os anos de 2018 e 2019;
- escolher a classe de modelos que minimiza o MAE por região, ou seja, escolher a classe de modelo cuja previsão mais se ajustou ao observado nos dados de 2018 e 2019;
- reestimar as especificações do modelo escolhido na etapa anterior, agora considerando a série de 2011 a 2019;
- gerar as previsões para os anos de 2020 e 2021;
- calcular o excesso de óbitos, no ponto, máximo e mínimo, por região.

Abaixo, segue especificação dos modelos escolhidos, bem como a série com as previsões e o observado para cada região. No **Apêndice I**, podem ser encontradas todas as etapas descritas acima, discriminadas para cada região:

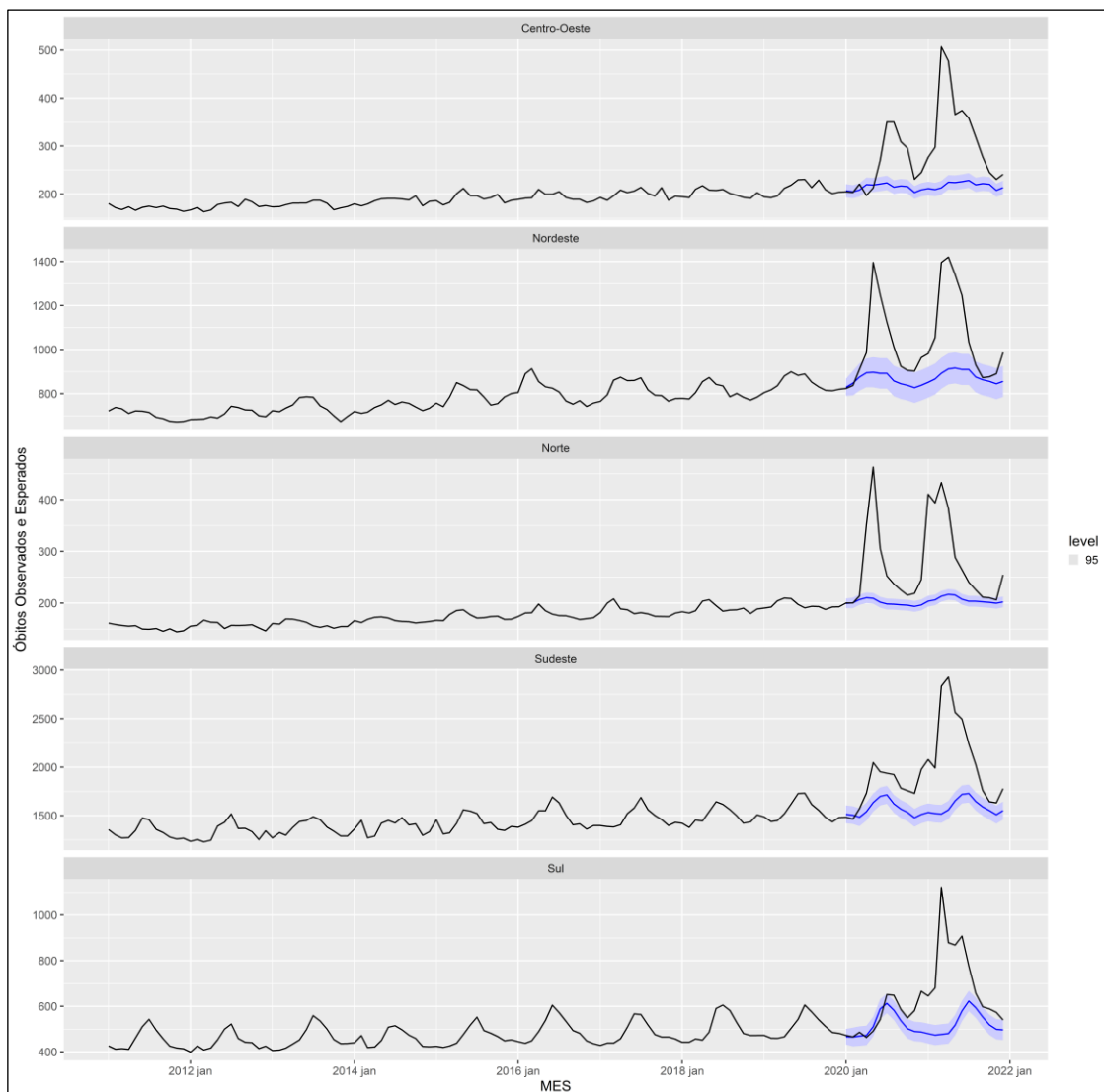
Quadro 14

Especificação dos modelos escolhidos para as regiões do Brasil

REGIAO	Orders
Nordeste	<ARIMA(0,0,5)(0,1,1)[12] w/ drift>
Norte	<ARIMA(3,0,0)(0,1,2)[12] w/ drift>
Sudeste	<ARIMA(0,0,0)(2,1,1)[12] w/ drift>
Sul	<ARIMA(0,0,2)(2,1,0)[12] w/ drift>
Centro-Oeste	<ETS(M,A,M)>

Figura 21

Óbitos observados e esperados, por região, no Brasil — 2011-21



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Nota: O gráfico apresenta escalas diferentes para cada região.

Com essas estimativas, foi possível calcular o excesso de óbitos por região no Brasil, no total do período entre março 2020 e dezembro de 2021, conforme apresentado na Figura 22.

Figura 22

Óbitos observados, esperados, excesso de óbitos e excesso proporcional de óbitos, por região, no Brasil — mar./20-dez./21

Ano	Óbitos Observados	Óbitos Esperados	Excesso de Óbitos	Excesso Proporcional de Óbitos	Excesso Proporcional de Óbitos (MIN)	Excesso Proporcional de Óbitos (MAX)
2020	1.220.153	1.035.874	184.279	17,8%	10,36%	26,29%
Centro-Oeste	82.152	65.833	16.319	24,8%	17,00%	33,70%
Nordeste	317.515	264.905	52.610	19,9%	11,14%	30,06%
Norte	83.506	61.425	22.081	35,9%	29,29%	43,33%
Sudeste	563.591	483.039	80.552	16,7%	10,22%	23,93%
Sul	173.389	160.672	12.717	7,9%	-0,21%	17,48%
2021	1.683.327	1.248.033	435.294	34,9%	26,29%	44,72%
Centro-Oeste	120.836	79.713	41.123	51,6%	42,12%	62,42%
Nordeste	396.240	320.997	75.243	23,4%	14,32%	34,14%
Norte	106.950	75.310	31.640	42,0%	35,23%	49,52%
Sudeste	790.487	580.561	209.926	36,2%	28,63%	44,63%
Sul	268.814	191.452	77.362	40,4%	29,47%	53,36%
Total	2.903.480	2.283.907	619.573	27,1%	19,07%	36,36%

Fonte dos dados brutos: Datasus (BRASIL, 2023).

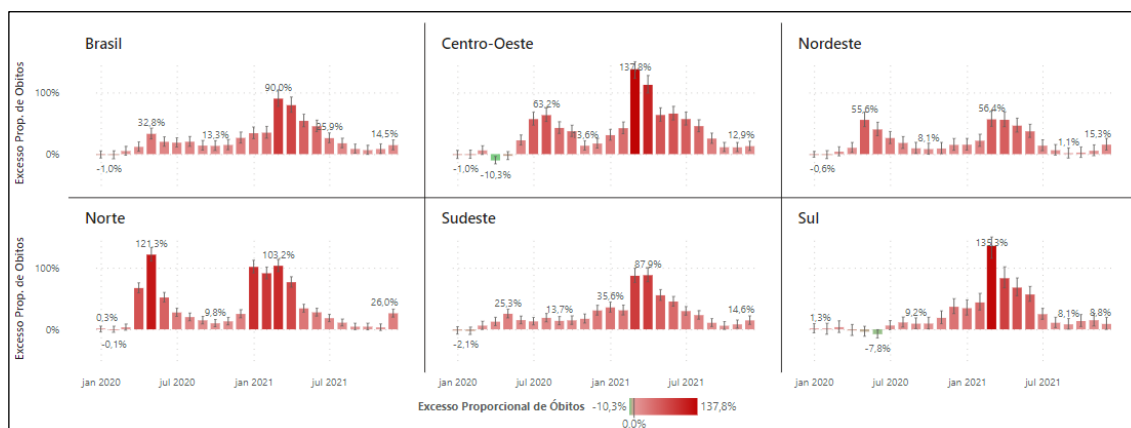
Nota: Para 2020, considerou-se o período de março a dezembro.

Em 2020, o maior excesso de óbitos foi observado na Região Norte, com 35,9%, seguido pelas Regiões Centro-Oeste (24,8%), Nordeste (19,9%), Sudeste (16,7%) e Sul (7,9%). Em 2021, o excesso de óbitos aumentou para todas as regiões, com destaque para a Região Sul, que teve um excesso aproximadamente cinco vezes maior do que em 2020 (40,4% *versus* 7,9%).

A diferença regional no total dos anos de 2020 e 2021 pode ser observada na dinâmica mensal do excesso proporcional de óbitos, na Figura 23. Nas Regiões Norte, Nordeste e Centro-Oeste, houve um ciclo de elevação pronunciado em 2020. Na Região Sudeste, a elevação de 2020 chegou a, no máximo, 25,3% em maio. A Região Sul, por sua vez, manteve as taxas baixas, chegando a apresentar excesso proporcional de óbitos negativos em junho de 2020, significando que houve menos óbitos do que o esperado para o mês. Em 2021, no entanto, a Região Sul apresentou grande elevação, principalmente no mês de março.

Figura 23

Excesso proporcional de óbitos acumulados no mês no Brasil e em suas regiões — jan./2020-dez./2021



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Por fim, é válido considerar algumas questões relevantes sobre a comparação dos resultados obtidos nas duas últimas seções: enquanto, nesta seção, a estimativa para o Brasil resulta da soma do excesso calculado para cada região, na seção anterior efetuou-se uma estimativa direta para o total de óbitos do País. O fato de essas duas formas de estimação resultarem em números bastante próximos (625.908 *versus* 619.573, com diferença de 1% somente) indica que os modelos individualizados para as regiões decompõem de forma satisfatória o comportamento agregado nacional.

7 Considerações finais

O trabalho buscou encontrar o melhor modelo para estimar o excesso de óbitos no Brasil e em suas regiões, nos anos de 2020 e 2021. Para tanto, desenhou-se uma metodologia para escolher o modelo que melhor se ajustava ao período e à regionalização analisada, a fim de se obter uma estimativa do número de óbitos esperados para os anos de 2020 e 2021, com base nos óbitos observados durante o período de 2011 a 2019. A partir desses óbitos esperados, calculou-se o excesso de óbitos.

Foram utilizados os modelos ARIMA, SARIMA, ETS, STL e SNAIVE. Para o Brasil, SARIMA mostrou-se superior aos demais. Para a estimativa nacional, o modelo sugerido mostrou-se robusto, uma vez que não houve alteração em sua especificação, mesmo retirando a parte final da série (anos de 2018 e 2019), embora diferentes coeficientes tenham sido estimados. Além disso, as estimativas geradas para os primeiros meses de 2020 (período pré-pandemia) corroboram com a hipótese de um bom ajuste do modelo, visto que os valores observados foram semelhantes aos esperados.

Mostrou-se que esse período da pandemia levou a óbito cerca de 625 mil vidas a mais do que seria esperado para o País. Desde março de 2020, houve acúmulo de excesso de óbitos no Brasil em todos os meses. O momento mais crítico foi entre os meses de março e abril de 2021, em que houve quase o dobro de óbitos esperados para o período. No agregado do período observado, quando considerado o modelo para Brasil apresentado na seção 5, o excesso proporcional de óbitos no País chegou a 27,5%, sendo cerca de 70% atribuídos ao excesso de no ano de 2021.

Na avaliação das regiões brasileiras, foram observadas diferenças quanto à tendência e à sazonalidade de cada região. A Região Sul apresentou a maior intensidade do componente sazonal, enquanto os estados da Região Norte apresentam a maior intensidade do componente de tendência. Devido a essas diferenças, modelos específicos foram estimados para cada região. Modelos SARIMA foram estimados para as Regiões Norte, Nordeste, Sudeste e Sul, enquanto, para a Região Centro-Oeste, foi utilizado um modelo ETS.

A partir da análise das regiões, foi possível perceber que o cálculo do agregado nacional acomoda diferenças regionais importantes. Embora todas as regiões tenham apresentado excesso de óbitos ao longo do período, esse número varia de 21,8% na Região Nordeste até 39,5% na Região Centro-Oeste, sendo que todas apresentaram maior acúmulo de excesso de óbitos em 2021. Enquanto a Região Sul apresentou o menor excesso proporcional de óbitos no ano de 2020, com 7,9%, a Nordeste apresentou a menor taxa em 2021, com 23,4% e preponderando também o melhor desempenho no agregado dos dois anos.

Referências

- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time Series Analysis: Forecasting and Control**. 4th Edition. Oxford: Wiley, 2008. Disponível em: <https://doi.org/10.1002/9781118619193>. Acesso em: 23 fev. 2023.
- BRASIL. Ministério da Saúde. **Datasus**: informações de saúde (Tabnet): estatísticas vitais. Brasília, DF: Ministério da Saúde, 2023. Disponível em: <https://datasus.saude.gov.br/mortalidade-desde-1996-pela-cid-10>. Acesso em: 29 abr. 2023.
- BRASIL. Ministério da Saúde. **Manual de Instruções para o Preenchimento da Declaração de Óbito**. Brasília: Ministério da Saúde, 2011. Disponível em: http://svs.aids.gov.br/download/manuais/Manual_Instr_Preench_DO_2011_jan.pdf. Acesso em: 23 fev. 2023.
- BRASIL. Ministério da Saúde. **Manual para investigação do óbito com causa mal definida**. Brasília: Ministério da Saúde, 2008. Disponível em: https://bvsms.saude.gov.br/bvs/publicacoes/manual_investigacao_obito.pdf. Acesso em: 23 fev. 2023.
- BRASIL. Ministério da Saúde. **Painel Coronavírus**. [Brasília, DF], 2023a. Disponível em: <https://covid.saude.gov.br/>. Acesso em: 23 fev. 2023.
- BRASIL. Ministério da Saúde. Coordenação Geral de Informações e Análise Epidemiológica. **Sistema de Informações sobre Mortalidade — SIM**. Consolidação da base de dados 2011 / Ministério da Saúde, Secretaria de Vigilância em Saúde. Brasília: Ministério da Saúde, 2013. Disponível em: http://tabnet.datasus.gov.br/cgi/sim/Consolida_Sim_2011.pdf. Acesso em: 23 fev. 2023.
- CLEVELAND, R. B. *et al.* STL: A seasonal-trend decomposition procedure based on loess. **Journal of Official Statistics**, v. 6, n. 1, p. 3-73, Mar., 1990. Disponível em: <http://bit.ly/stl1990>. Acesso em: 23 fev. 2023.
- CUNHA, C. C.; TEIXEIRA, R.; FRANÇA, E. Avaliação da investigação de óbitos por causas mal definidas no Brasil em 2010. **Epidemiologia e Serviços de Saúde**, v. 26, n. 1, p. 19-30, jan.-mar., 2017. Disponível em: http://scielo.iec.gov.br/scielo.php?script=sci_arttext&pid=S1679-49742017000100019#:~:text=Resultados%3A,nos%20estados%20como%20nas%20capitais. Acesso em: 23 fev. 2023.
- FRANÇA, E. B. *et al.* Deaths due to COVID-19 in Brazil: how many are there and which are being identified? **Revista Brasileira de Epidemiologia**, v. 23, 7p., jun., 2020. Disponível em: <https://www.scielo.br/j/rbepid/a/75zrygtRM8GMdgKYhTLfmpH/?lang=en&format=pdf>. Acesso em: 23 fev. 2023.
- HOLT, C. C. Forecasting seasonals and trends by exponentially weighted averages, **International Journal of Forecasting**, v. 20, n. 1, p. -10, Jan.-Mar., 2004. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0169207003001134>. Acesso em: 23 fev. 2023.
- HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. 3rd edition. Melbourne: OTexts, 2021. Disponível em: [OTexts.com/fpp3](https://otexts.com/fpp3). Acesso em: 13 ago. 2022.
- IBGE. **Projeções da população**. Brasília, DF: IBGE, 2022. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/populacao/9109-projecao-da-populacao.html?=&t=resultados>. Acesso em: 23 fev. 2022.

KANG, Y.; HYNDMAN, R. J.; SMITH-MILES, K. Visualising forecasting algorithm performance using time series instance spaces. **International Journal of Forecasting**, v. 33, n. 2, p. 345-358, Apr.-Jun., 2017. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0169207016301030>. Acesso em: 23 fev. 2023.

MARINHO, F. *et al.* **Aumento das mortes no Brasil, Regiões, Estados e Capitais em tempos de COVID-19: excesso de óbitos por causas naturais que não deveria ter ocorrido.** Vital Strategies, 2020. (Nota Técnica). Disponível em: https://www.vitalstrategies.org/wp-content/uploads/RMS_ExcessMortality_BR_Report-Portuguese.pdf. Acesso em: 23 fev. 2023.

ORELLANA, J. D. Y. *et al.* Excesso de mortes durante a pandemia de COVID-19: subnotificação e desigualdades regionais no Brasil. **Cadernos de Saúde Pública**, v. 37, n. 1, fev., 2021. Disponível em: <https://doi.org/10.1590/0102-311X00259120>. Acesso em: 15 ago. 2022.

ORGANIZAÇÃO MUNDIAL DA SAÚDE — OMS. Classificação estatística internacional de doenças e problemas relacionados à saúde. 10ª edição. São Paulo, 1995.

PEARCE, N.; LAWLOR, D. A.; BRICKLEY, E. B. Comparisons between countries are essential for the control of COVID-19. **International Journal of Epidemiology**, v. 49, n. 4, p. 1.059-62, Aug., 2020. Disponível em: <https://academic.oup.com/ije/article/49/4/1059/5864919>. Acesso em: 23 fev. 2023.

SALDANHA, R. de F.; BASTOS, R. R.; BARCELLOS, C. Microdatasus: pacote para download e pré-processamento de microdados do Departamento de Informática do SUS (DATASUS). **Cadernos de Saúde Pública**, v. 35, n. 9, set., 2019. Disponível em: <https://www.scielo.br/j/csp/a/gdJXqcrW5PPDHX8rwPDYL7F/>. Acesso em: 23 fev. 2023.

SANTOS, A. M. dos *et al.* Excess deaths from all causes and by COVID-19 in Brazil in 2020. **Revista de Saúde Pública**, v. 55, out., 2021. Disponível em: <https://doi.org/10.11606/s1518-8787.2021055004137>. Acesso em: 23 fev. 2023.

SETTEL, P. *et al.* Mortality surveillance during the COVID-19 pandemic. **Bulletin of the World Health Organization**, v. 98, n. 6, p. 374, Jun., 2020. Disponível em: <https://apps.who.int/iris/handle/10665/332363?locale-attribute=pt&>. Acesso em: 23 fev. 2023.

VIEIRA, A. *et al.* Rapid estimation of excess mortality during the COVID-19 pandemic in Portugal: beyond reported deaths. **Journal of Epidemiology and Global Health**, v. 10, n.3, p. 209-13, Sep., 2020. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7509098/>. Acesso em: 23 fev. 2023.

WINTERS, P. R. Forecasting sales by exponentially weighted moving averages. **Management Science**, v. 6, n. 3, p. 324-342, Apr., 1960. Disponível em: <https://www.jstor.org/stable/2627346>. Acesso em: 23 fev. 2023.

Apêndice I

Resultados por região

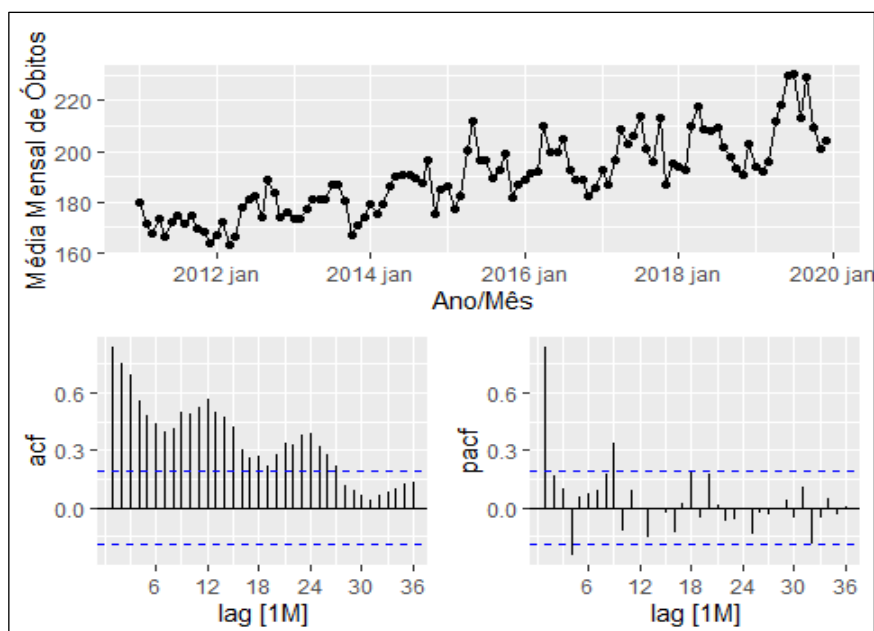
As figuras e os quadros abaixo mostram as etapas reproduzidas para cada região do Brasil, conforme apresentado ao longo do texto.

Região Centro-Oeste

Análise dos dados e Teste de Estacionariedade

Figura A.1

Média mensal de óbitos diários, Função Autocorrelação (ACF) e Função de Autocorrelação Parcial (PACF) para a Região Centro-Oeste — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

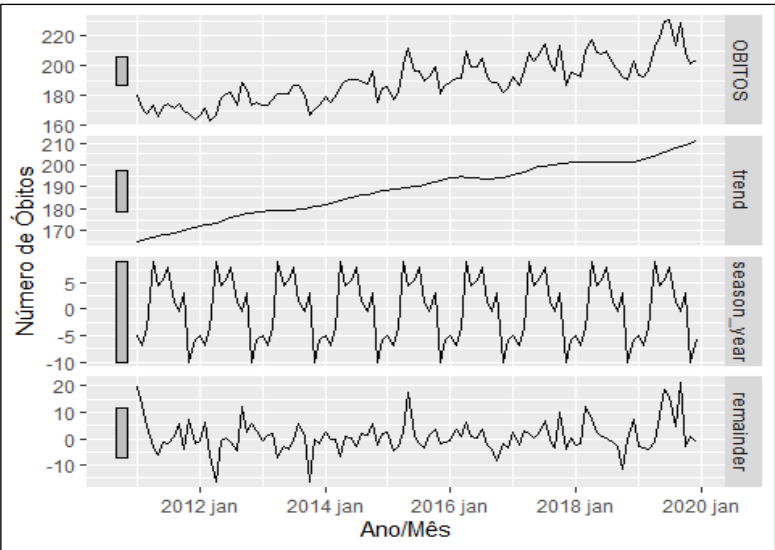
Quadro A.1
Teste Dickey-Fuller Aumentado (ADF) para a Região Centro-Oeste — 2011-19

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -26.6877  -5.1239   0.7634   5.3811  18.2125
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## z.lag.1 0.0001789  0.0043922   0.041   0.968
##
## Residual standard error: 8.643 on 106 degrees of freedom
## Multiple R-squared:  1.565e-05, Adjusted R-squared:  -0.009418
## F-statistic: 0.001659 on 1 and 106 DF,  p-value: 0.9676
##
##
## Value of test-statistic is: 0.0407
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Quadro A.2
Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
para a Região Centro-Oeste — 2011-19

	x
kpss_stat	1.562126
kpss_pvalue	0.010000

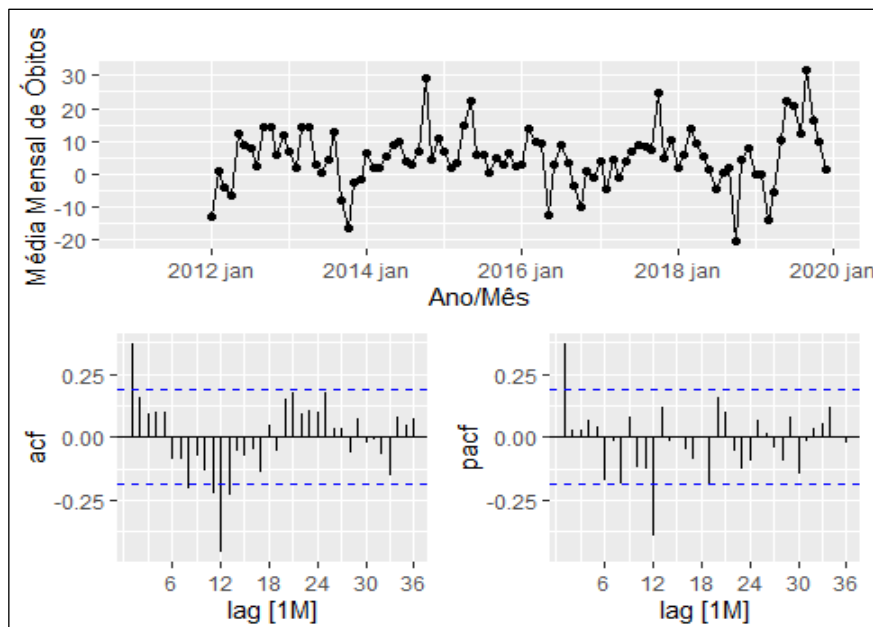
Figura A.2
Decomposição STL do número de óbitos para a Região Centro-Oeste — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Figura A.3

Análise dos resíduos da média diária de óbitos em 1.ª diferença sazonal ($D = 1$), Função Autocorrelação (ACF) e Função de Autocorrelação Parcial (PACF) para a Região Centro-Oeste — 2011-19



Quadro A.3

Teste Dickey-Fuller Aumentado (ADF) em 1.ª diferença sazonal ($D = 1$) para a Região Centro-Oeste — 2011-19

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##   Min     1Q  Median     3Q    Max
## -21.417 -1.713   2.232   6.577  25.373
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## z.lag.1 -0.46804    0.08623  -5.428 4.44e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.509 on 94 degrees of freedom
## Multiple R-squared:  0.2386, Adjusted R-squared:  0.2305
## F-statistic: 29.46 on 1 and 94 DF, p-value: 4.435e-07
##
##
## Value of test-statistic is: -5.4279
##
## Critical values for test statistics:
##   1pct  5pct 10pct
## tau1 -2.6 -1.95 -1.61
```

Quadro A.4

Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
em 1.ª diferença sazonal (D = 1) para a
Região Centro-Oeste — 2011-19

	x
kpss_stat	1.148444
kpss_pvalue	0.010000

Teste de previsão e escolha do modelo

Quadro A.5

Descrição dos modelos de treinamento (2011-2017) e efetivo (2011-2019) para a Região Centro-Oeste

REGIAO	.model	fit_train_2011-2017	fit_2011-2019
Centro-Oeste	auto_arima_STEP	<ARIMA(1,1,1)(1,0,2)[12]>	<ARIMA(1,1,1)(1,0,2)[12]>
Centro-Oeste	auto_arima	<ARIMA(1,1,1)(1,0,2)[12]>	<ARIMA(1,1,1)(1,0,2)[12]>
Centro-Oeste	ets	<ETS(M,A,M)>	<ETS(M,A,M)>
Centro-Oeste	stl		
Centro-Oeste	snaive		

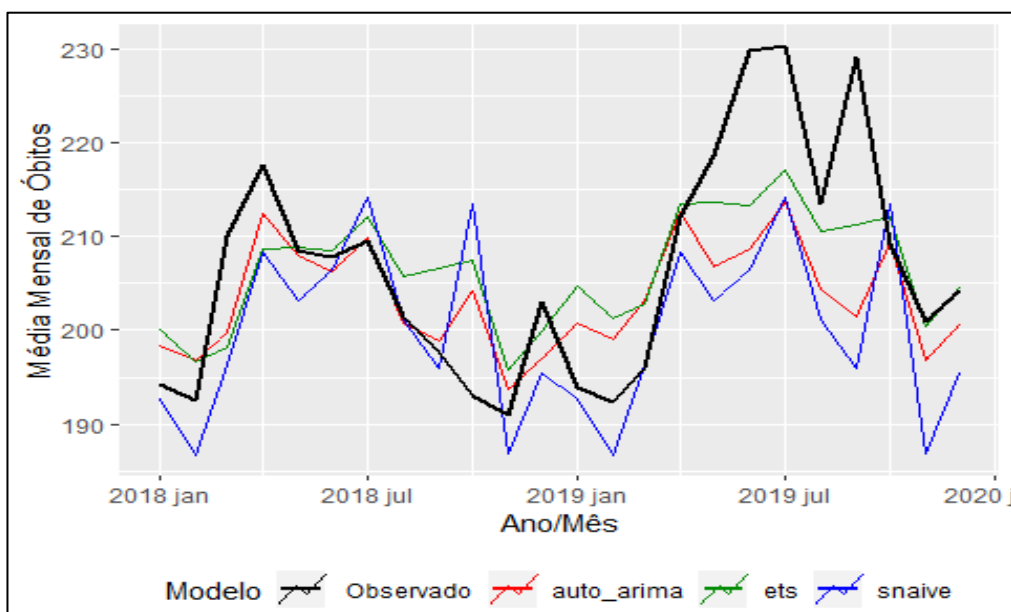
Quadro A.6

Análise dos resíduos de treinamento, ordenados por erro absoluto médio (MAE), para a Região Centro-Oeste — 2011-17)

REGIAO	.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
Centro-Oeste	stl	Training	-0.042	5.176	3.389	-0.088	1.852	0.466	0.562	0.123
Centro-Oeste	ets	Training	-0.126	5.418	4.022	-0.167	2.179	0.553	0.588	0.251
Centro-Oeste	auto_arima_STEP	Training	0.748	6.060	4.857	0.310	2.600	0.667	0.657	-0.007
Centro-Oeste	auto_arima	Training	0.748	6.060	4.857	0.310	2.600	0.667	0.657	-0.007
Centro-Oeste	snaive	Training	4.841	9.219	7.276	2.482	3.854	1.000	1.000	0.278
Centro-Oeste	ets	Test	0.100	8.369	6.584	-0.154	3.157	0.905	0.908	0.466
Centro-Oeste	auto_arima	Test	3.041	9.679	6.853	1.256	3.222	0.942	1.050	0.444
Centro-Oeste	auto_arima_STEP	Test	3.041	9.679	6.853	1.256	3.222	0.942	1.050	0.444
Centro-Oeste	snaive	Test	6.449	12.093	8.938	2.946	4.205	1.228	1.312	0.219

Figura A.4

Previsão de treinamento e observado para a média mensal de óbitos na Região Centro-Oeste — 2018-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Quadro A.7

Descrição dos modelos para a Região Centro-Oeste — amostra 2011-2019

Model name	Orders
auto_arima	<ARIMA(1,1,1)(1,0,2)[12]>
auto_arima_STEP	<ARIMA(1,1,1)(1,0,2)[12]>
ets	<ETS(M,A,M)>
snaive	

Quadro A.8

Análise dos resíduos dos modelos efetivos, ordenados por erro absoluto médio (MAE), para a Região Centro-Oeste — 2011-19

.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
stl	Training	0.896	6.208	4.185	0.363	2.185	0.540	0.616	0.247
ets	Training	-0.149	6.014	4.453	-0.189	2.330	0.575	0.597	0.329
auto_arima_STEP	Training	0.940	6.699	5.192	0.381	2.695	0.670	0.665	-0.036
auto_arima	Training	0.940	6.699	5.192	0.381	2.695	0.670	0.665	-0.036
snaive	Training	4.979	10.072	7.747	2.463	3.969	1.000	1.000	0.369

Quadro A.9

Análise dos critérios de informação dos modelos efetivos, ordenados por Critério de Informação Akaike Corrigido (AICc), para a Região Centro-Oeste — 2011-2019

.model	sigma2	log_lik	AIC	AICc	BIC
auto_arima_STEP	47.515	-359.543	731.085	731.925	747.122
auto_arima	47.515	-359.543	731.085	731.925	747.122
ets	0.001	-445.312	924.623	931.423	970.220

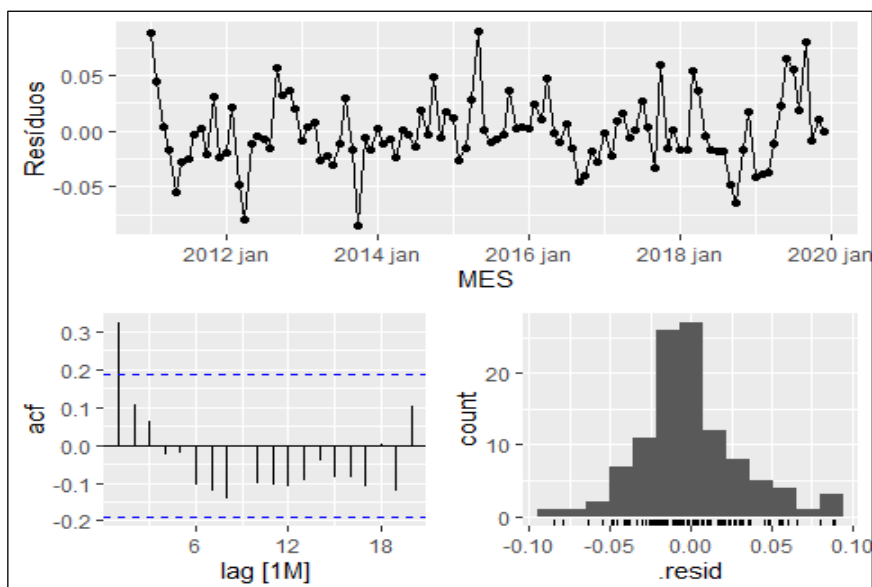
Quadro A.10

Especificação do modelo escolhido para a Região Centro-Oeste — 2011-19

```
## Series: OBITOS
## Model: ETS(M,A,M)
## Smoothing parameters:
##   alpha = 0.008066977
##   beta  = 0.0001000885
##   gamma = 0.0001012244
##
## Initial states:
##   l[0]  b[0]  s[0]  s[-1]  s[-2]  s[-3]  s[-4]  s[-5]  s[-6]  s[-7]  s[-8]
##   s[-9] s[-10] s[-11]
## 168.5595 0.3965099 0.9685227 0.9438655 1.003688 1.011575 1.00168 1.044
## 968 1.035514 1.028224 1.033742 0.9826015 0.9670896 0.9785296
##
## sigma^2: 0.0012
##
##   AIC  AICc  BIC
## 924.6234 931.4234 970.2196
```

Figura A.5

Resíduos do modelo escolhido para a Região Centro-Oeste — 2011-19



Quadro A.11

Teste Ljung-Box (lag = 36) com sete graus de liberdade
para a Região Centro-Oeste — 2011-19

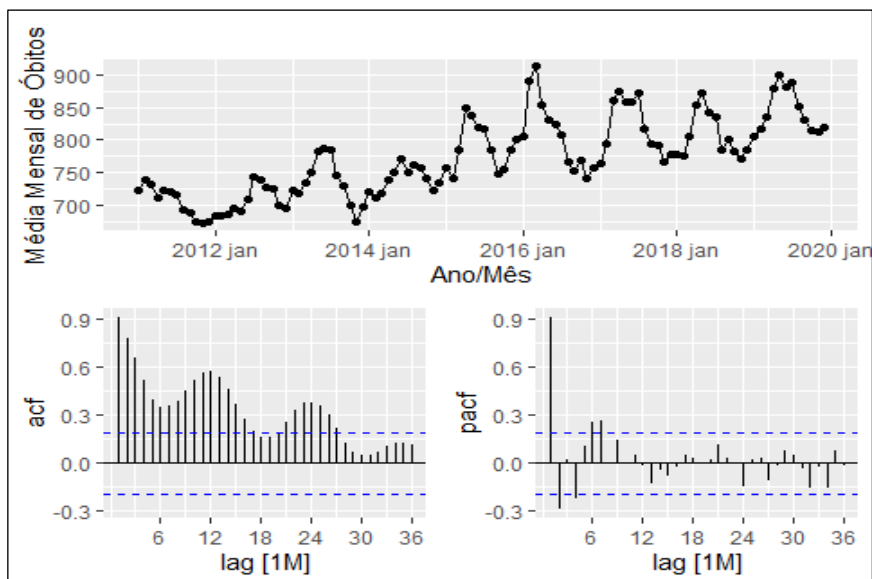
REGIAO	.model	lb_stat	lb_pvalue
Centro-Oeste	ets	49.44338	0.0001582

Região Nordeste

Análise dos dados e Teste de Estacionariedade

Figura A.6

Média mensal de óbitos diários, Função Autocorrelação (ACF) e Função de Autocorrelação Parcial (PACF) para a Região Nordeste — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Quadro A.12

Teste Dickey-Fuller Aumentado (ADF) para a Região Nordeste — 2011-19

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -58.958 -16.587  -1.457  14.928  83.509
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1  0.0006863   0.0030461   0.225   0.822
##
## Residual standard error: 24.42 on 106 degrees of freedom
## Multiple R-squared:  0.0004786, Adjusted R-squared:  -0.008951
## F-statistic: 0.05076 on 1 and 106 DF, p-value: 0.8222
##
##
## Value of test-statistic is: 0.2253
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

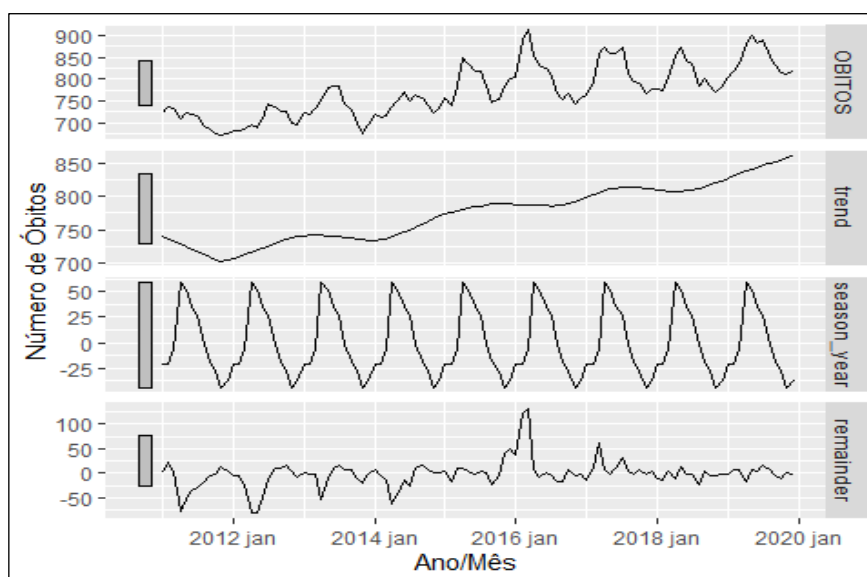
Quadro A.13

Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
para a Região Nordeste — 2011-19

	x
kpss_stat	5.037041
kpss_pvalue	0.010000

Figura A.7

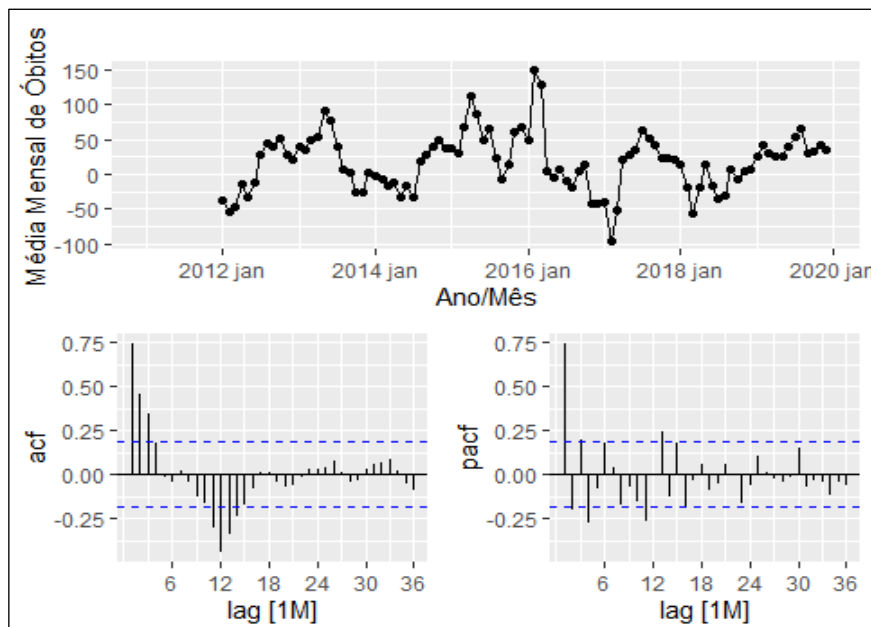
Decomposição STL do número de óbitos para a Região Nordeste — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Figura A.8

Análise dos resíduos da média diária de óbitos em 1.^a diferença sazonal (D = 1), Função Autocorrelação (ACF) e Função de Autocorrelação Parcial (PACF) para a Região Nordeste — 2011-19



Quadro A.14

Teste Dickey-Fuller Aumentado (ADF) em 1.^a diferença sazonal (D = 1) para a Região Nordeste — 2011-19

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##   Min     1Q   Median     3Q      Max
## -96.622 -9.393  3.319 20.456 110.794
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## z.lag.1 -0.21071    0.06323  -3.332 0.00123 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 27.54 on 94 degrees of freedom
## Multiple R-squared:  0.1057, Adjusted R-squared:  0.09614
## F-statistic: 11.1 on 1 and 94 DF, p-value: 0.001232
##
##
## Value of test-statistic is: -3.3324
##
## Critical values for test statistics:
##   1pct  5pct 10pct
## tau1 -2.6 -1.95 -1.61
```

Quadro A.15

Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
em 1.ª diferença sazonal (D = 1) para a
Região Nordeste — 2011-19

	x
kpss_stat	0.6719363
kpss_pvalue	0.0160967

Teste de previsão e escolha do modelo

Quadro A.16

Descrição dos modelos de treinamento (2011-2017) e efetivo (2011-2019) para a Região Nordeste

REGIAO	.model	fit_train_2011-2017	fit_2011-2019
Nordeste	auto_arima_STEP	<ARIMA(2,1,3)>	<ARIMA(1,0,2)(0,1,1)[12] w/ drift>
Nordeste	auto_arima	<ARIMA(5,1,1)>	<ARIMA(0,0,5)(0,1,1)[12] w/ drift>
Nordeste	ets	<ETS(M,N,N)>	<ETS(M,N,A)>
Nordeste	stl		
Nordeste	snaive		

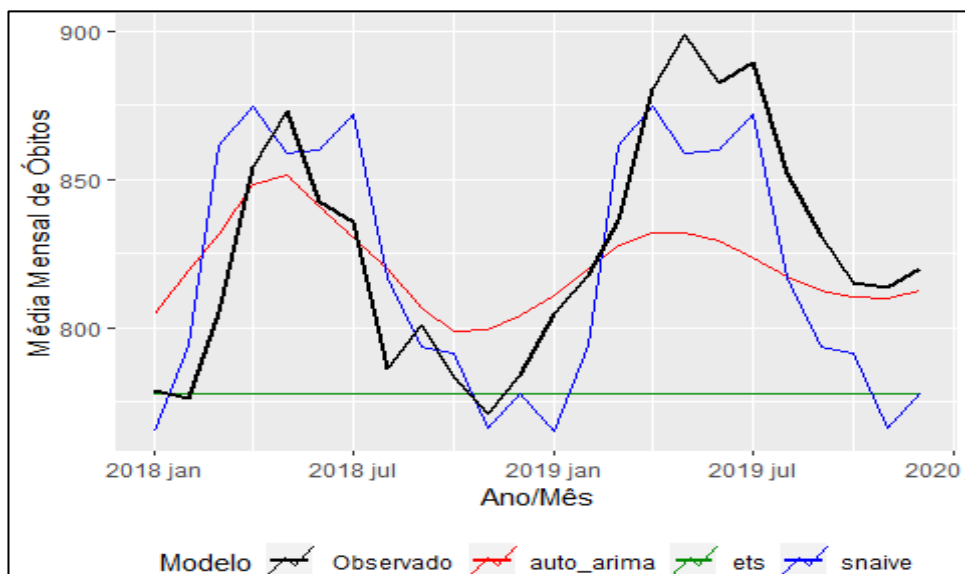
Quadro A.17

Análise dos resíduos de treinamento, ordenados por erro absoluto médio (MAE), para a Região Nordeste — 2011-17

REGIAO	.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
Nordeste	auto_arima	Training	2.530	19.096	14.431	0.283	1.882	0.378	0.399	-0.025
Nordeste	auto_arima_STEP	Training	0.467	21.008	15.988	0.047	2.075	0.419	0.439	-0.012
Nordeste	stl	Training	2.531	24.129	16.080	0.210	2.062	0.422	0.504	0.615
Nordeste	ets	Training	0.676	24.280	18.286	0.043	2.364	0.480	0.507	0.250
Nordeste	snaive	Training	18.976	47.909	38.132	2.295	4.885	1.000	1.000	0.742
Nordeste	auto_arima	Test	5.818	30.448	23.086	0.552	2.754	0.605	0.636	0.764
Nordeste	snaive	Test	7.106	28.546	24.804	0.825	2.997	0.650	0.596	0.543
Nordeste	auto_arima_STEP	Test	43.228	57.117	44.415	5.038	5.191	1.165	1.192	0.775
Nordeste	ets	Test	48.489	61.598	49.235	5.672	5.768	1.291	1.286	0.769

Figura A.9

Previsão de treinamento e observado para a média mensal de óbitos na Região Nordeste — 2018-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Quadro A.18

Descrição dos modelos para a Região Nordeste — amostra 2011-19

Model name	Orders
auto_arima	<ARIMA(0,0,5)(0,1,1)[12] w/ drift>
auto_arima_STEP	<ARIMA(1,0,2)(0,1,1)[12] w/ drift>
ets	<ETS(M,N,A)>
snaive	

Quadro A.19

Análise dos resíduos dos modelos efetivos, ordenados por erro absoluto médio (MAE), para a Região Nordeste — 2011-19

.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
auto_arima	Training	0.133	17.892	12.360	-0.026	1.570	0.346	0.401	-0.028
auto_arima_STEP	Training	0.105	18.399	12.786	-0.041	1.629	0.358	0.413	-0.001
ets	Training	0.945	19.603	14.199	0.089	1.820	0.397	0.440	0.089
stl	Training	-1.427	28.350	16.509	-0.315	2.138	0.462	0.636	0.611
snaive	Training	17.471	44.601	35.736	2.098	4.521	1.000	1.000	0.743

Quadro A.20

Análise dos critérios de informação dos modelos efetivos, ordenados por Critério de Informação Akaike Corrigido (AICc), para a Região Nordeste — 2011-2019

.model	sigma2	log_lik	AIC	AICc	BIC
auto_arima	388.481	-426.071	868.142	869.797	888.657
auto_arima_STEP	401.764	-429.753	871.505	872.449	886.892
ets	0.001	-571.455	1172.909	1178.127	1213.141

Quadro A.21

Especificação do modelo escolhido para a Região Nordeste — 2011-19

```
## Series: OBITOS
## Model: ARIMA(0,0,5)(0,1,1)[12] w/ drift
##
## Coefficients:
##      ma1  ma2  ma3  ma4  ma5  sma1 constant
##      1.0311 0.6817 0.5429 0.5147 0.3023 -0.8166 17.6298
## s.e. 0.0997 0.1526 0.1687 0.1391 0.1151 0.1616 2.8800
##
## sigma^2 estimated as 388.5: log likelihood=-426.07
## AIC=868.14 AICc=869.8 BIC=888.66
```

Figura A.10

Resíduos do modelo escolhido para a Região Nordeste — 2011-19

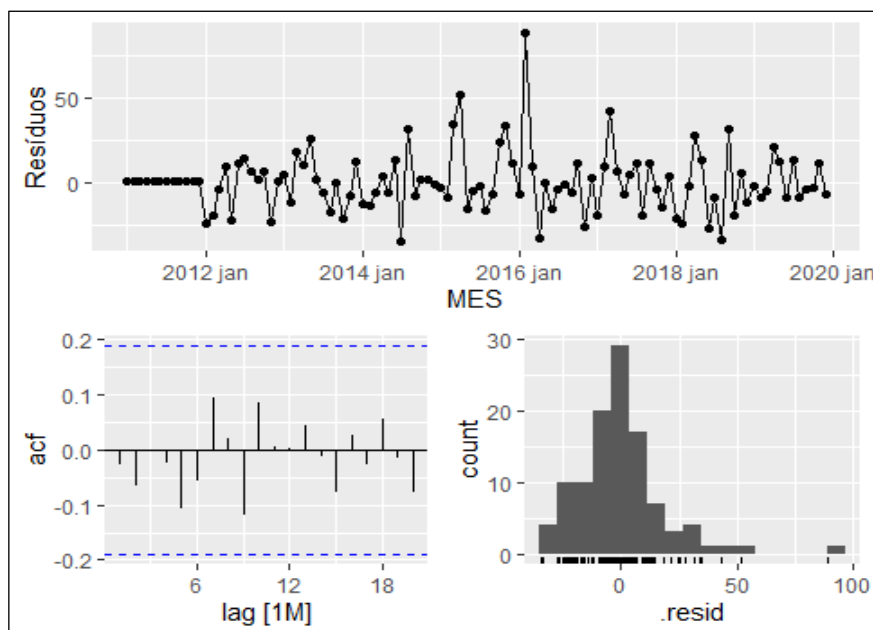
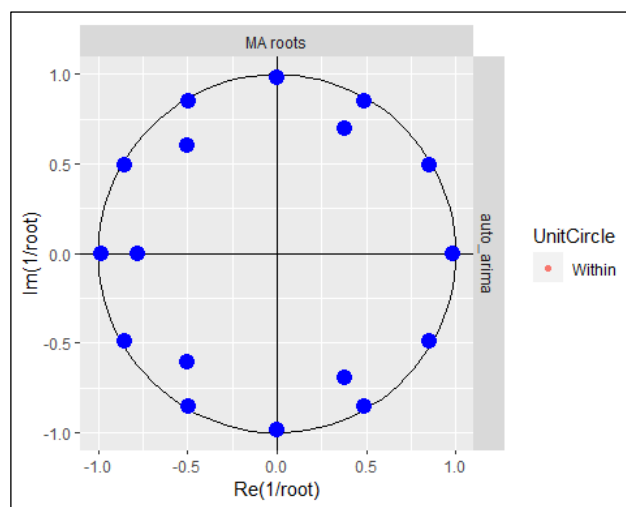


Figura A.11

Inverso das raízes unitárias do modelo escolhido para a Região Nordeste — período 2011-19



Quadro A.22

Teste Ljung-Box (lag = 36) com sete graus de liberdade para a Região Nordeste — 2011-19

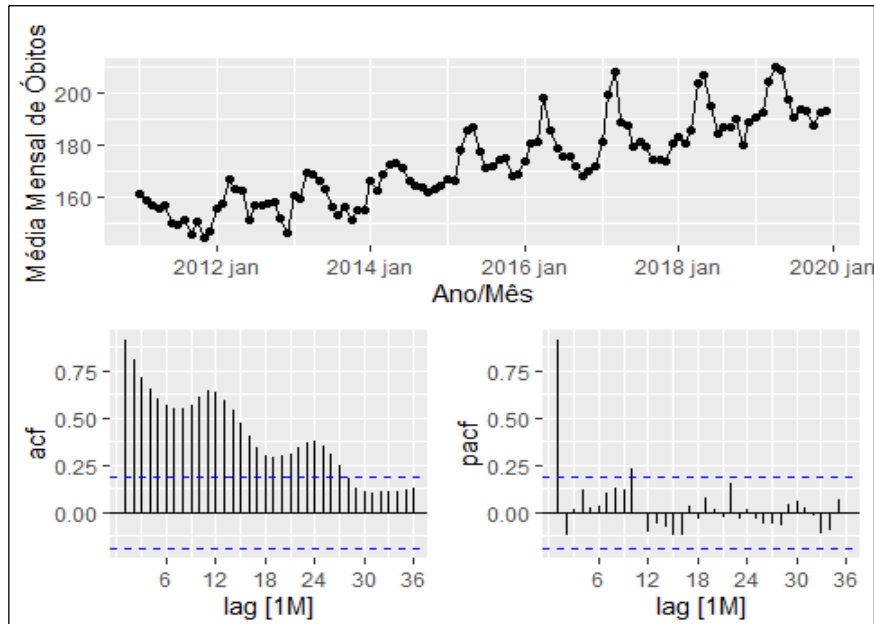
REGIAO	.model	lb_stat	lb_pvalue
Nordeste	auto_arima	21.73084	0.8311619

Região Norte

Análise dos dados e Teste de Estacionariedade

Figura A.12

Média mensal de óbitos diários, Função Autocorrelação (ACF) e Função de Autocorrelação Parcial (PACF) para a Região Norte — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Quadro A.23

Teste Dickey-Fuller Aumentado (ADF) para a Região Norte — 2011-19

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##   Min     1Q   Median     3Q    Max
## -19.379  -3.528   0.228   2.882  18.154
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## z.lag.1  0.001040   0.003616   0.288   0.774
##
## Residual standard error: 6.492 on 106 degrees of freedom
## Multiple R-squared:  0.0007794, Adjusted R-squared:  -0.008647
## F-statistic: 0.08268 on 1 and 106 DF, p-value: 0.7743
##
##
## Value of test-statistic is: 0.2875
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

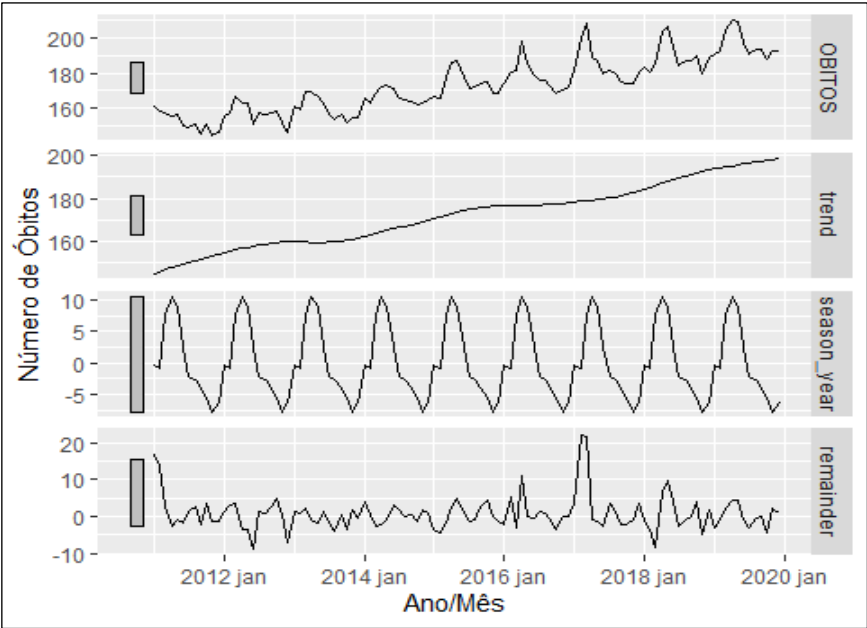
Quadro A.24

Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
para a Região Norte — 2011-19

	x
kpss_stat	1.467297
kpss_pvalue	0.010000

Figura A.13

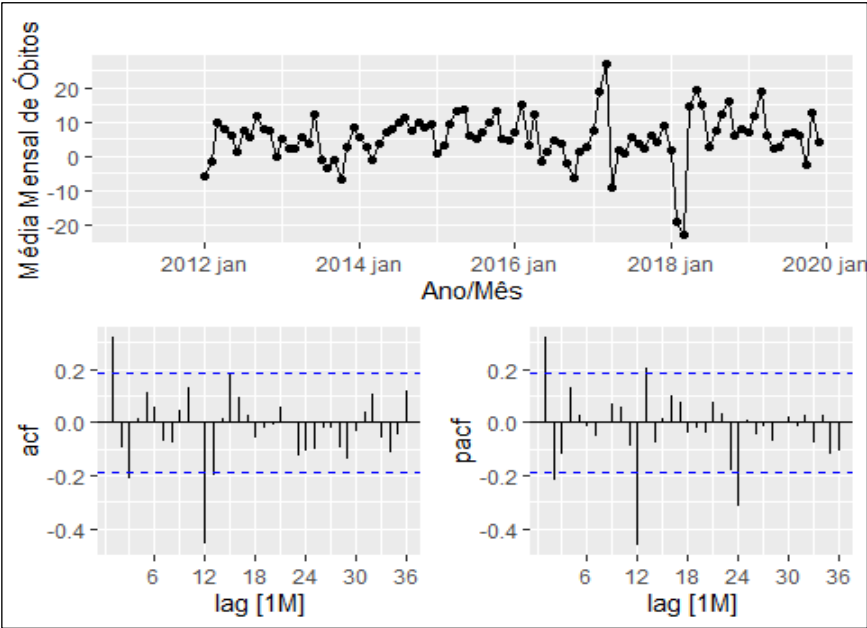
Decomposição STL do número de óbitos para a Região Norte — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Figura A.14

Análise dos resíduos da média diária de óbitos em 1.^a diferença sazonal (D = 1),
Função Autocorrelação (ACF) e Função de Autocorrelação Parcial (PACF)
para a Região Norte — 2011-2019



Quadro A.25

Teste Dickey-Fuller Aumentado (ADF) em 1.^a diferença sazonal (D = 1) para a Região Norte — 2011-19

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -24.8334  -0.4613   2.5504   5.8854  28.0319
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1 -0.41735    0.08369  -4.987 2.8e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.32 on 94 degrees of freedom
## Multiple R-squared:  0.2092, Adjusted R-squared:  0.2008
## F-statistic: 24.87 on 1 and 94 DF, p-value: 2.8e-06
##
##
## Value of test-statistic is: -4.9868
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.6 -1.95 -1.61
```

Quadro A.26

Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS) em 1.^a diferença sazonal (D=1) para a Região Norte — 2011-19

	x
kpss_stat	1.482702
kpss_pvalue	0.010000

Teste de previsão e escolha do modelo

Quadro A.27

Descrição dos modelos de treinamento (2011-2017) e efetivo (2011-2019) para a Região Norte

REGIAO	.model	fit_train_2011-2017	fit_2011-2019
Norte	auto_arima_STEP	<ARIMA(1,0,0)(0,1,1)[12] w/ drift>	<ARIMA(2,0,2)(0,1,2)[12] w/ drift>
Norte	auto_arima	<ARIMA(1,0,0)(0,1,1)[12] w/ drift>	<ARIMA(3,0,0)(0,1,2)[12] w/ drift>
Norte	ets	<ETS(M,N,A)>	<ETS(M,A,M)>
Norte	stl		
Norte	snaive		

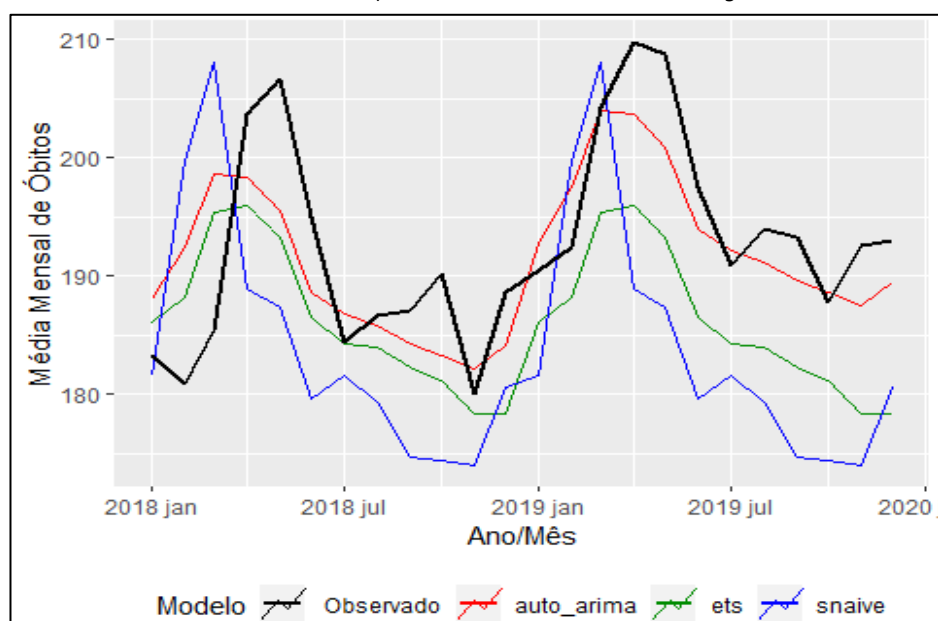
Quadro A.28

Análise dos resíduos de treinamento, ordenados por erro absoluto médio (MAE), para a Região Norte — 2011-17

REGIAO	.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
Norte	stl	Training	1.054	4.895	2.778	0.539	1.614	0.433	0.619	0.359
Norte	auto_arima_STEP	Training	-0.115	4.524	3.024	-0.143	1.748	0.471	0.572	0.040
Norte	auto_arima	Training	-0.115	4.524	3.024	-0.143	1.748	0.471	0.572	0.040
Norte	ets	Training	0.409	4.801	3.605	0.195	2.114	0.562	0.608	0.048
Norte	snaive	Training	5.311	7.903	6.419	3.049	3.718	1.000	1.000	0.265
Norte	auto_arima	Test	1.142	5.825	4.756	0.508	2.463	0.741	0.737	0.387
Norte	auto_arima_STEP	Test	1.142	5.825	4.756	0.508	2.463	0.741	0.737	0.387
Norte	ets	Test	6.629	9.330	8.301	3.339	4.250	1.293	1.181	0.519
Norte	snaive	Test	8.624	14.371	13.018	4.346	6.709	2.028	1.818	0.469

Figura A.15

Previsão de treinamento e observado para a média mensal de óbitos na Região Norte — 2018-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Quadro A.29

Descrição dos modelos para a Região Norte — amostra 2011-19

Model name	Orders
auto_arima	<ARIMA(3,0,0)(0,1,2)[12] w/ drift>
auto_arima_STEP	<ARIMA(2,0,2)(0,1,2)[12] w/ drift>
ets	<ETS(M,A,M)>
snaive	

Quadro A.30

Análise dos resíduos dos modelos efetivos, ordenados por erro absoluto médio (MAE), para a Região Norte — 2011-19

.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
stl	Training	0.791	4.787	3.016	0.391	1.716	0.417	0.536	0.341
auto_arima_STEP	Training	-0.219	4.324	3.160	-0.203	1.786	0.436	0.484	-0.011
auto_arima	Training	-0.246	4.294	3.175	-0.219	1.795	0.439	0.480	0.005
ets	Training	-0.057	4.643	3.592	-0.132	2.049	0.496	0.519	0.314
snaive	Training	5.492	8.939	7.241	3.045	4.046	1.000	1.000	0.319

Quadro A.31

Análise dos critérios de informação dos modelos efetivos, ordenados
por Critério de Informação Akaike Corrigido (AICc),
para a Região Norte — 2011-2019

.model	sigma2	log_lik	AIC	AICc	BIC
auto_arima	22.131	-291.156	596.312	597.585	614.262
auto_arima_STEP	22.688	-290.584	597.168	598.823	617.682
ets	0.001	-415.673	865.347	872.147	910.943

Quadro A.32

Especificação do modelo escolhido para a Região Norte — 2011-19

```
## Series: OBITOS
## Model: ARIMA(3,0,0)(0,1,2)[12] w/ drift
##
## Coefficients:
##      ar1      ar2      ar3      sma1      sma2      constant
##  0.4078 -0.1190 -0.1724 -0.9891  0.1073  4.9046
## s.e. 0.1026  0.1129  0.1079  0.2190  0.1561  0.1587
##
## sigma^2 estimated as 22.13: log likelihood=-291.16
## AIC=596.31  AICc=597.58  BIC=614.26
```

Figura A.16

Resíduos do modelo escolhido para a Região Norte — 2011-19

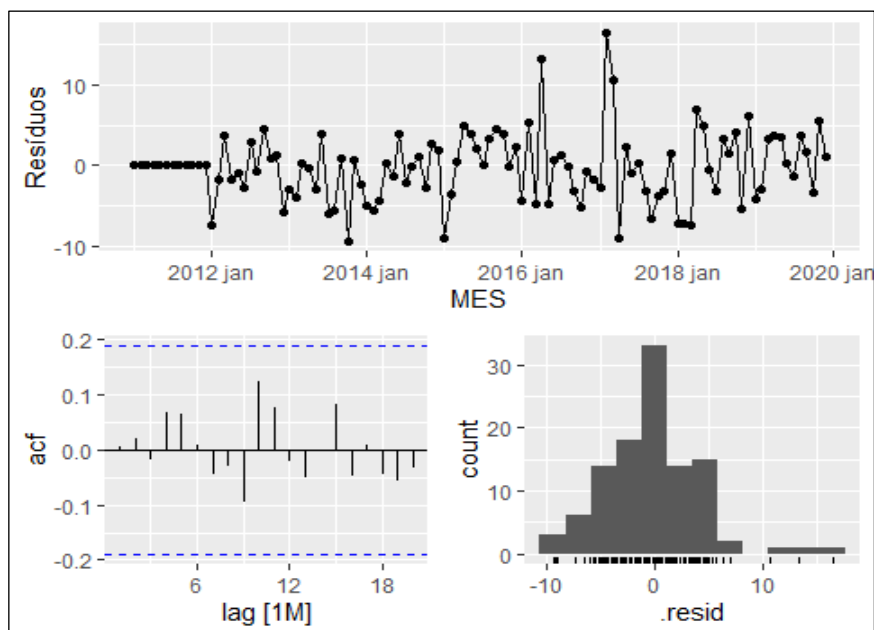
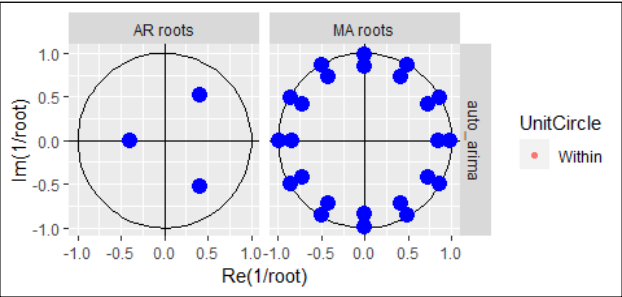


Figura A.17
Inverso das raízes unitárias do modelo escolhido
para a Região Norte — 2011-19

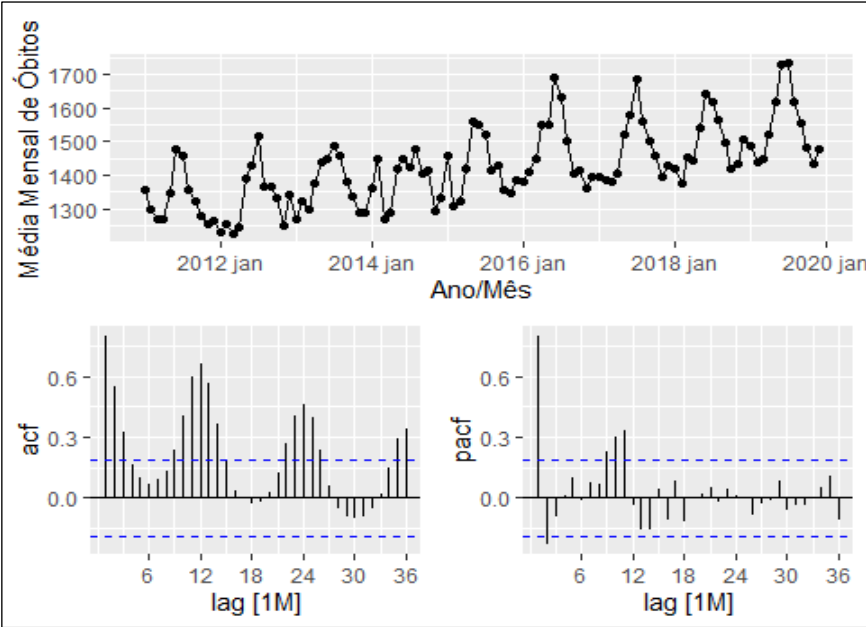


Quadro A.33
Teste Ljung-Box (lag = 36) com sete graus de
liberdade para a Região Norte — 2011-19

REGIAO	.model	lb_stat	lb_pvalue
Norte	auto_arima	26.27631	0.6609565

Região Sudeste
Análise dos dados e Teste de Estacionariedade

Figura A.18
Média mensal de óbitos diários, Função Autocorrelação (ACF) e Função de Autocorrelação
Parcial (PACF) para a Região Sudeste — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

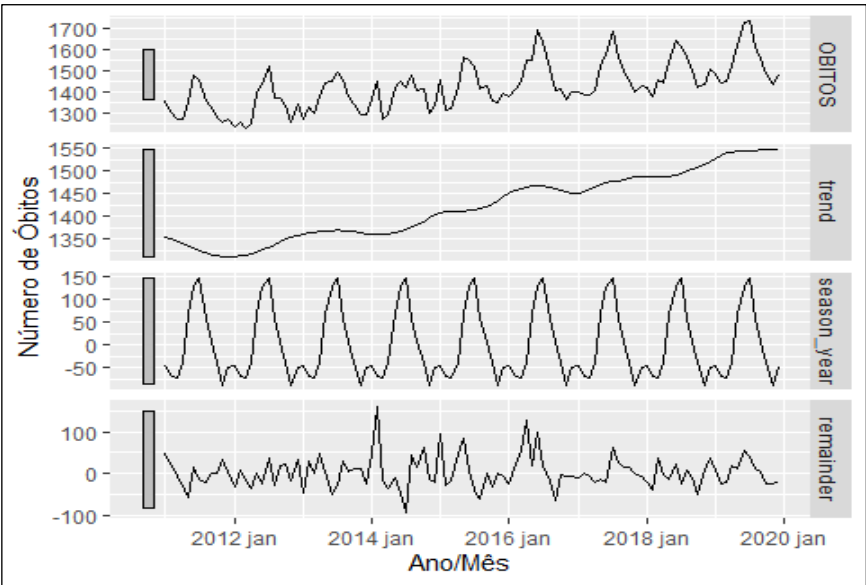
Quadro A.34
Teste Dickey-Fuller Aumentado (ADF) para a Região Sudeste — 2011-19

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -179.938  -47.475   0.648   50.741  145.437
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1 -0.000443    0.004853  -0.091   0.927
##
## Residual standard error: 71.72 on 106 degrees of freedom
## Multiple R-squared:  7.858e-05, Adjusted R-squared:  -0.009355
## F-statistic: 0.00833 on 1 and 106 DF,  p-value: 0.9275
##
##
## Value of test-statistic is: -0.0913
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Quadro A.35
Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
para a Região Sudeste — 2011-19

	x
kpss_stat	5.13878
kpss_pvalue	0.01000

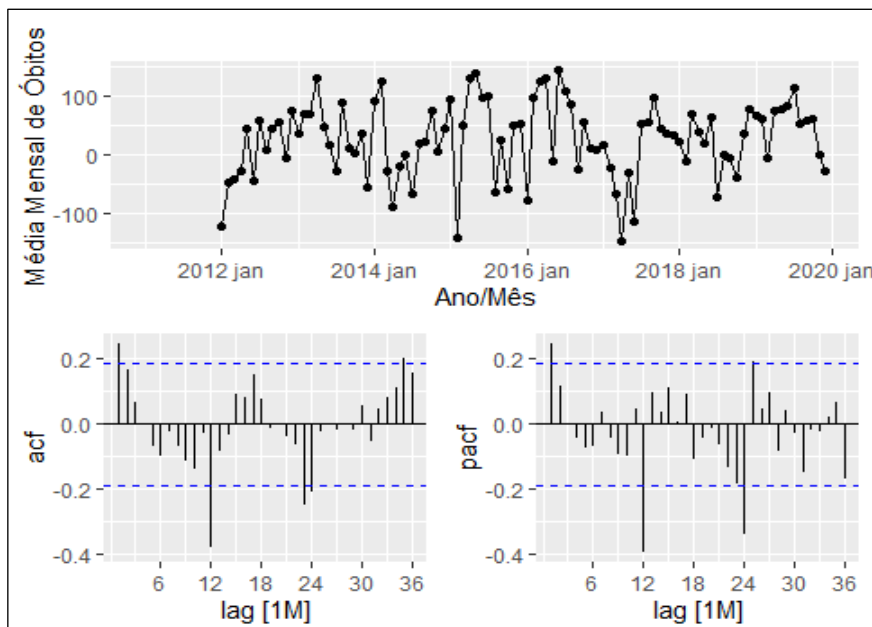
Figura A.19
Decomposição STL do número de óbitos para a Região Sudeste — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Figura A.20

Análise dos resíduos da média diária de óbitos em 1.^a diferença sazonal (D = 1), Função Autocorrelação (ACF) e Função de Autocorrelação Parcial (PACF) para a Região Sudeste — 2011-19



Quadro A.36

Teste Dickey-Fuller Aumentado (ADF) em 1.^a diferença sazonal (D = 1) para a Região Sudeste — 2011-19

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##   Min     1Q  Median     3Q    Max
## -175.60 -21.57  24.08  60.52 147.89
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## z.lag.1 -0.62936    0.09408  -6.69 1.59e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 63.75 on 94 degrees of freedom
## Multiple R-squared:  0.3225, Adjusted R-squared:  0.3153
## F-statistic: 44.75 on 1 and 94 DF, p-value: 1.594e-09
##
##
## Value of test-statistic is: -6.6898
##
## Critical values for test statistics:
##   1pct  5pct 10pct
## tau1 -2.6 -1.95 -1.61
```

Quadro A.37

Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
em 1.^a diferença sazonal (D = 1) para a
Região Sudeste — 2011-19

	x
kpss_stat	1.177154
kpss_pvalue	0.010000

Teste de previsão e escolha do modelo

Quadro A.38

Descrição dos modelos de treinamento (2011-2017) e efetivo (2011-2019) para a Região Sudeste

REGIAO	.model	fit_train_2011-2017	fit_2011-2019
Sudeste	auto_arima_STEP	<ARIMA(1,0,0)(2,1,0)[12] w/ drift>	<ARIMA(0,0,0)(2,1,1)[12] w/ drift>
Sudeste	auto_arima	<ARIMA(1,0,0)(2,1,0)[12] w/ drift>	<ARIMA(0,0,0)(2,1,1)[12] w/ drift>
Sudeste	ets	<ETS(M,N,A)>	<ETS(A,N,A)>
Sudeste	stl		
Sudeste	snaive		

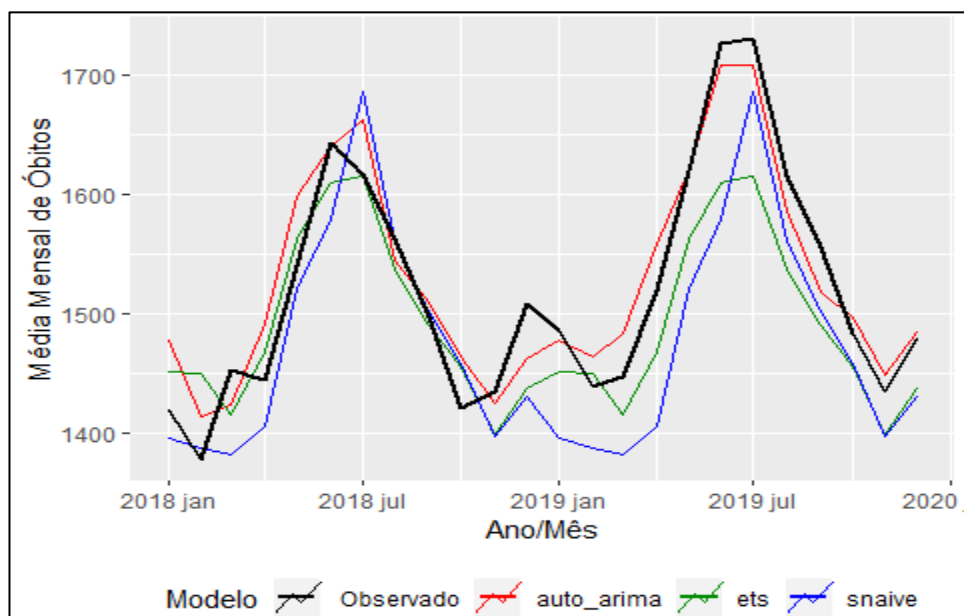
Quadro A.39

Análise dos resíduos de treinamento, ordenados por erro absoluto médio (MAE), para a Região Sudeste — 2011-17

REGIAO	.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
Sudeste	stl	Training	7.712	44.042	28.962	0.469	2.018	0.477	0.604	0.208
Sudeste	ets	Training	3.451	45.817	34.909	0.148	2.473	0.575	0.628	0.069
Sudeste	auto_arima_STEP	Training	-1.313	49.700	36.245	-0.213	2.565	0.597	0.681	-0.020
Sudeste	auto_arima	Training	-1.313	49.700	36.245	-0.213	2.565	0.597	0.681	-0.020
Sudeste	snaive	Training	24.267	72.950	60.712	1.591	4.271	1.000	1.000	0.235
Sudeste	auto_arima	Test	-8.919	32.240	27.439	-0.642	1.830	0.452	0.442	0.195
Sudeste	auto_arima_STEP	Test	-8.919	32.240	27.439	-0.642	1.830	0.452	0.442	0.195
Sudeste	ets	Test	28.001	53.083	44.417	1.723	2.878	0.732	0.728	0.515
Sudeste	snaive	Test	43.635	63.830	53.784	2.827	3.492	0.886	0.875	0.409

Figura A.21

Previsão de treinamento e observado para a média mensal de óbitos na Região Sudeste — 2018-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Quadro A.40

Descrição dos modelos para a Região Sudeste — amostra 2011-19

Model name	Orders
auto_arima	<ARIMA(0,0,0)(2,1,1)[12] w/ drift>
auto_arima_STEP	<ARIMA(0,0,0)(2,1,1)[12] w/ drift>
ets	<ETS(A,N,A)>
snaive	

Quadro A.41

Análise dos resíduos dos modelos efetivos, ordenados por erro absoluto médio (MAE), para a Região Sudeste — 2011-19

.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
stl	Training	3.206	38.294	27.754	0.160	1.929	0.483	0.553	0.150
auto_arima_STEP	Training	-2.372	43.361	31.614	-0.284	2.204	0.550	0.626	0.263
auto_arima	Training	-2.372	43.361	31.614	-0.284	2.204	0.550	0.626	0.263
ets	Training	3.419	43.645	33.444	0.159	2.341	0.582	0.631	0.065
snaive	Training	26.889	69.221	57.456	1.749	3.973	1.000	1.000	0.246

Quadro A.42

Análise dos critérios de informação dos modelos efetivos, ordenados por Critério de Informação Akaike Corrigido (AICc), para a Região Sudeste — 2011-19

.model	sigma2	log_lik	AIC	AICc	BIC
auto_arima_STEP	2207.205	-512.149	1034.297	1034.964	1047.119
auto_arima	2207.205	-512.149	1034.297	1034.964	1047.119
ets	2188.625	-660.653	1351.307	1356.524	1391.539

Quadro A.43

Especificação do modelo escolhido para a Região Sudeste — 2011-19

Series: OBITOS
Model: ARIMA(0,0,0)(2,1,1)[12] w/ drift
##
Coefficients:
sar1 sar2 sma1 constant
-0.0877 -0.2178 -0.7827 36.4036
s.e. 0.1632 0.1379 0.2453 2.0758
##
sigma^2 estimated as 2207: log likelihood=-512.15
AIC=1034.3 AICc=1034.96 BIC=1047.12

Figura A.22

Resíduos do modelo escolhido para a Região Sudeste — 2011-19

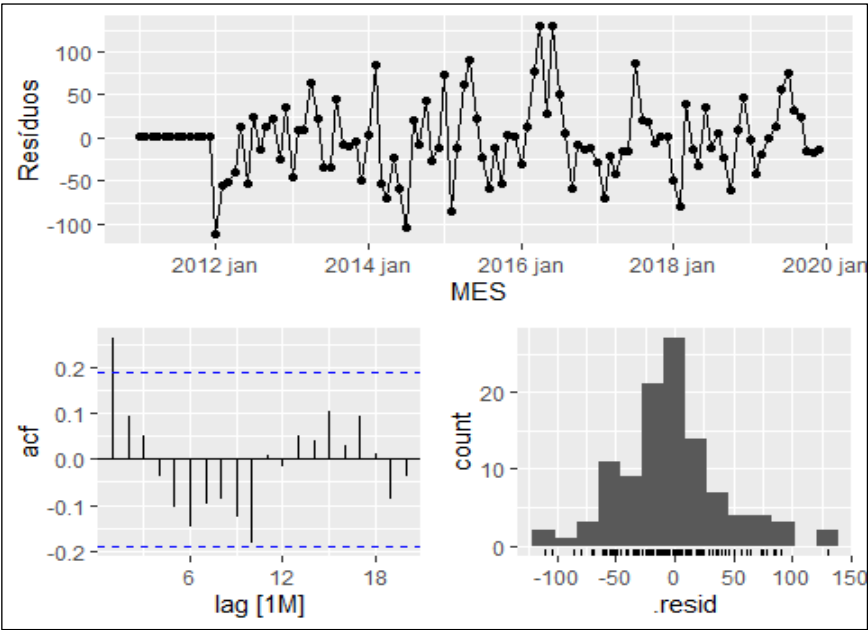
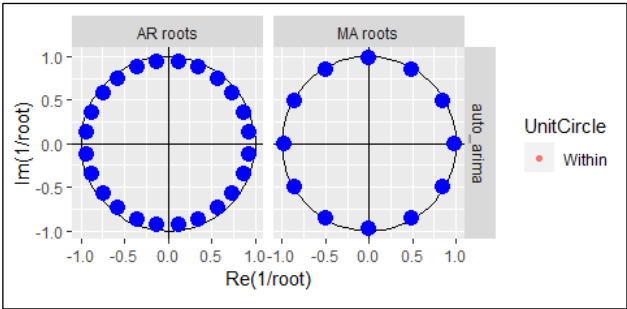


Figura A.23

Inverso das raízes unitárias do modelo escolhido para a Região Sudeste — 2011-19



Quadro A.44

Teste Ljung-Box (lag = 36) com sete graus de liberdade para a Região Sudeste — 2011-19

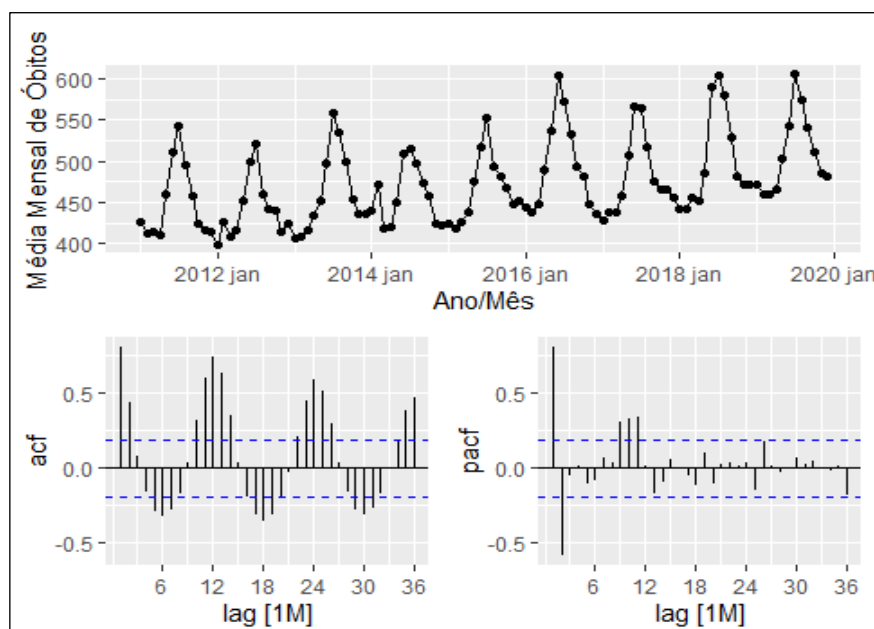
REGIAO	.model	lb_stat	lb_pvalue
Sudeste	auto_arima	37.45734	0.2329465

Região Sul

Análise dos dados e Teste de Estacionariedade

Figura A.24

Média mensal de óbitos diários, Função Autocorrelação (ACF) e Função de Autocorrelação Parcial (PACF) para a Região Sul — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Quadro A.45

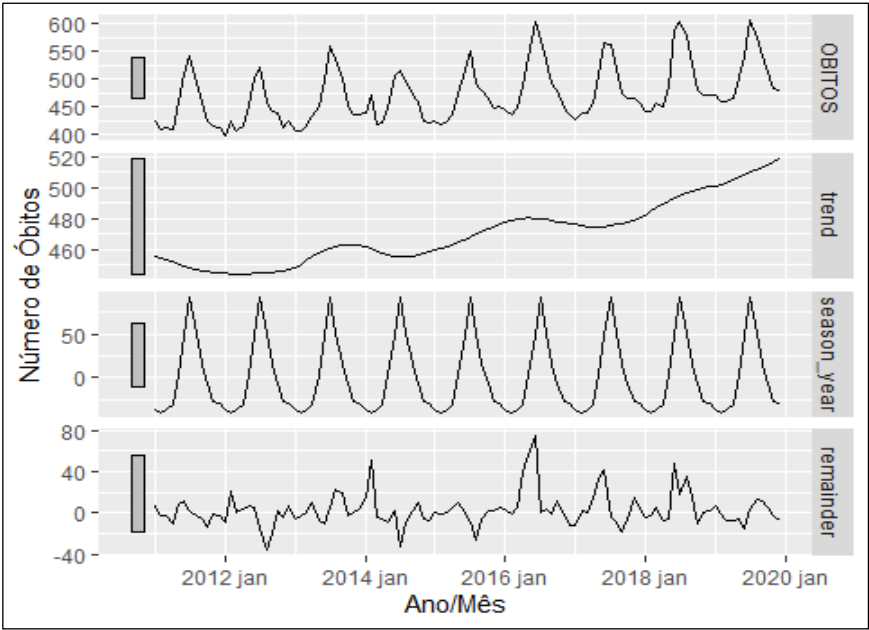
Teste Dickey-Fuller Aumentado (ADF) para a Região Sul — 2011-19

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##   Min     1Q   Median     3Q    Max
## -62.404 -18.019  -0.985  18.421 104.407
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## z.lag.1 -0.001142   0.006415  -0.178   0.859
##
## Residual standard error: 31.59 on 106 degrees of freedom
## Multiple R-squared:  0.0002989, Adjusted R-squared:  -0.009132
## F-statistic: 0.03169 on 1 and 106 DF, p-value: 0.859
##
##
## Value of test-statistic is: -0.178
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

Quadro A.46
Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
para a Região Sul — 2011-19

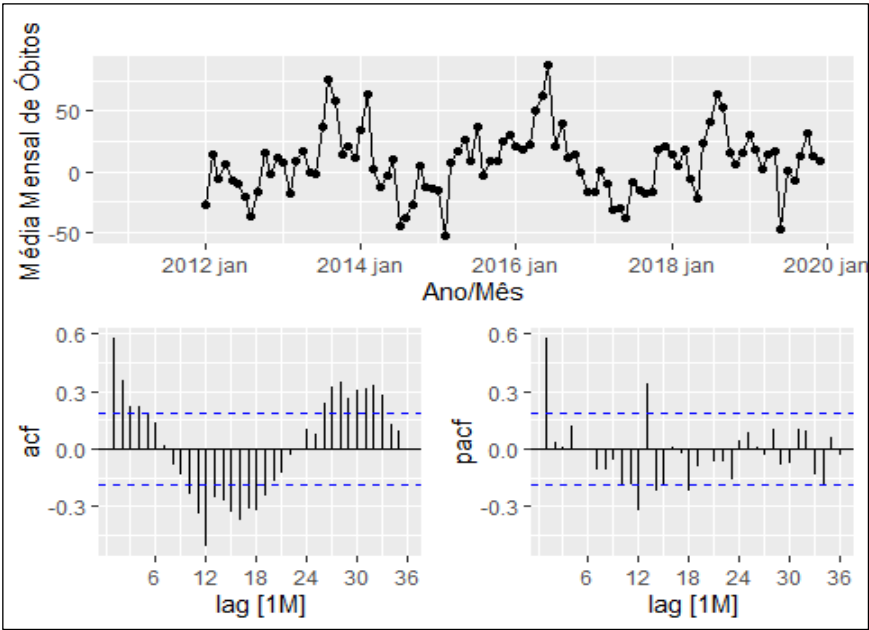
	x
kpss_stat	3.763882
kpss_pvalue	0.010000

Figura A.25
Decomposição STL do número de óbitos para a Região Sul — 2011-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Figura A.26
Análise dos resíduos da média diária de óbitos em 1.ª diferença sazonal (D = 1), Função Autocorrelação (ACF) e Função de Autocorrelação Parcial (PACF) para a Região Sul — 2011-19



Quadro A.47

Teste Dickey-Fuller Aumentado (ADF) em 1.^a diferença sazonal (D = 1)
para a Região Sul — 2011-19

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##   Min     1Q   Median     3Q    Max
## -57.975 -11.025  2.125 16.205 53.817
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## z.lag.1 -0.38868    0.08106  -4.795 6.08e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 21.77 on 94 degrees of freedom
## Multiple R-squared:  0.1965, Adjusted R-squared:  0.188
## F-statistic: 22.99 on 1 and 94 DF, p-value: 6.079e-06
##
##
## Value of test-statistic is: -4.7953
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.6 -1.95 -1.61
```

Quadro A.48

Teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
em 1.^a diferença sazonal (D=1)
para a Região Sul — 2011-19

	x
kpss_stat	0.4433850
kpss_pvalue	0.0584548

Teste de previsão e escolha do modelo

Quadro A.49

Descrição dos modelos de treinamento (2011-2017) e efetivo (2011-2019) para a Região Sul

REGIAO	.model	fit_train_2011-2017	fit_2011-2019
Sul	auto_arima_STEP	<ARIMA(1,0,0)(1,1,1)[12] w/ drift>	<ARIMA(0,0,2)(2,1,0)[12] w/ drift>
Sul	auto_arima	<ARIMA(0,0,2)(1,1,1)[12] w/ drift>	<ARIMA(0,0,2)(2,1,0)[12] w/ drift>
Sul	ets	<ETS(M,N,A)>	<ETS(M,N,A)>
Sul	stl		
Sul	snaive		

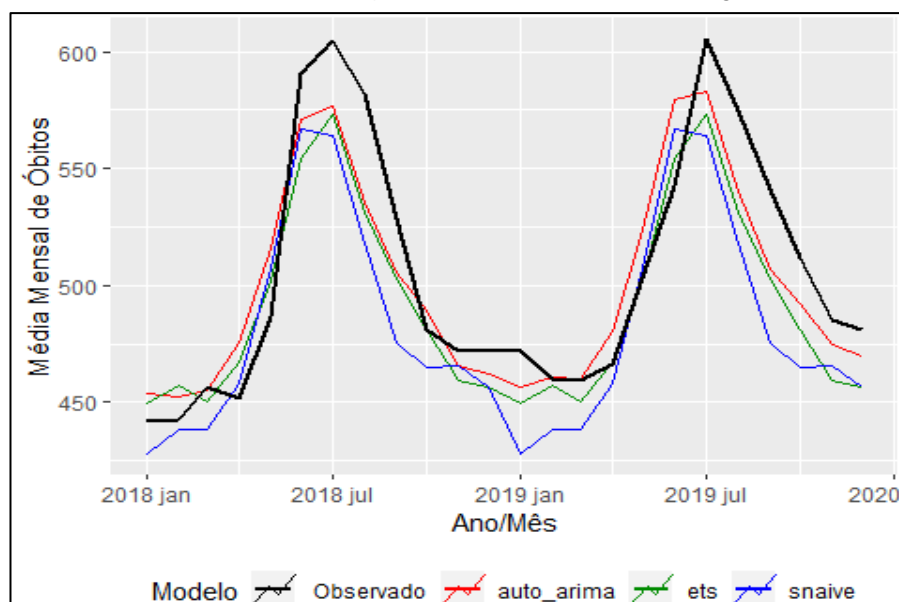
Quadro A.50

Análise dos resíduos de treinamento, ordenados por erro absoluto médio (MAE), para a Região Sul — 2011-17

REGIAO	.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
Sul	stl	Training	3.131	16.219	9.962	0.546	2.057	0.465	0.581	0.439
Sul	auto_arima	Training	-0.535	14.280	10.219	-0.201	2.162	0.477	0.511	0.032
Sul	auto_arima_STEP	Training	-0.315	14.747	10.847	-0.145	2.295	0.507	0.528	0.061
Sul	ets	Training	0.452	14.815	11.730	0.041	2.499	0.548	0.530	0.039
Sul	snaive	Training	5.513	27.928	21.407	1.026	4.452	1.000	1.000	0.605
Sul	auto_arima	Test	5.212	21.768	18.295	0.803	3.498	0.855	0.779	0.510
Sul	auto_arima_STEP	Test	5.773	22.419	18.788	0.916	3.593	0.878	0.803	0.516
Sul	ets	Test	14.252	24.127	19.770	2.583	3.756	0.924	0.864	0.511
Sul	snaive	Test	22.905	33.282	27.591	4.349	5.275	1.289	1.192	0.520

Figura A.27

Previsão de treinamento e observado para a média mensal de óbitos na Região Sul — 2018-19



Fonte dos dados brutos: Datasus (BRASIL, 2023).

Quadro A.51

Descrição dos modelos para a Região Sul — amostra 2011-19

Model name	Orders
auto_arima	<ARIMA(0,0,2)(2,1,0)[12] w/ drift>
auto_arima_STEP	<ARIMA(0,0,2)(2,1,0)[12] w/ drift>
ets	<ETS(M,N,A)>
snaive	

Quadro A.52

Análise dos resíduos dos modelos efetivos, ordenados por erro absoluto médio (MAE), para a Região Sul — 2011-19

.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
stl	Training	2.704	16.495	10.408	0.448	2.103	0.493	0.601	0.441
auto_arima_STEP	Training	-0.268	15.992	11.676	-0.172	2.419	0.553	0.583	0.014
auto_arima	Training	-0.268	15.992	11.676	-0.172	2.419	0.553	0.583	0.014
ets	Training	0.622	15.372	11.783	0.059	2.446	0.558	0.560	0.036
snaive	Training	7.484	27.431	21.115	1.419	4.317	1.000	1.000	0.577

Quadro A.53

Análise dos critérios de informação dos modelos efetivos, ordenados por
Critério de Informação Akaike Corrigido (AICc),
para a Região Sul — 2011-2019

.model	sigma2	log_lik	AIC	AICc	BIC
auto_arima_STEP	303.518	-411.468	834.936	835.879	850.322
auto_arima	303.518	-411.468	834.936	835.879	850.322
ets	0.001	-542.800	1115.601	1120.818	1155.833

Quadro A.54

Especificação do modelo escolhido para a Região Sul — 2011-19

```
## Series: OBITOS
## Model: ARIMA(0,0,2)(2,1,0)[12] w/ drift
##
## Coefficients:
##      ma1  ma2  sar1  sar2 constant
## 0.6664 0.3507 -0.7232 -0.1893 15.2662
## s.e. 0.0959 0.1008 0.1076 0.1137 3.7469
##
## sigma^2 estimated as 303.5: log likelihood=-411.47
## AIC=834.94 AICc=835.88 BIC=850.32
```

Figura A.28

Resíduos do modelo escolhido para a Região Sul — 2011-19

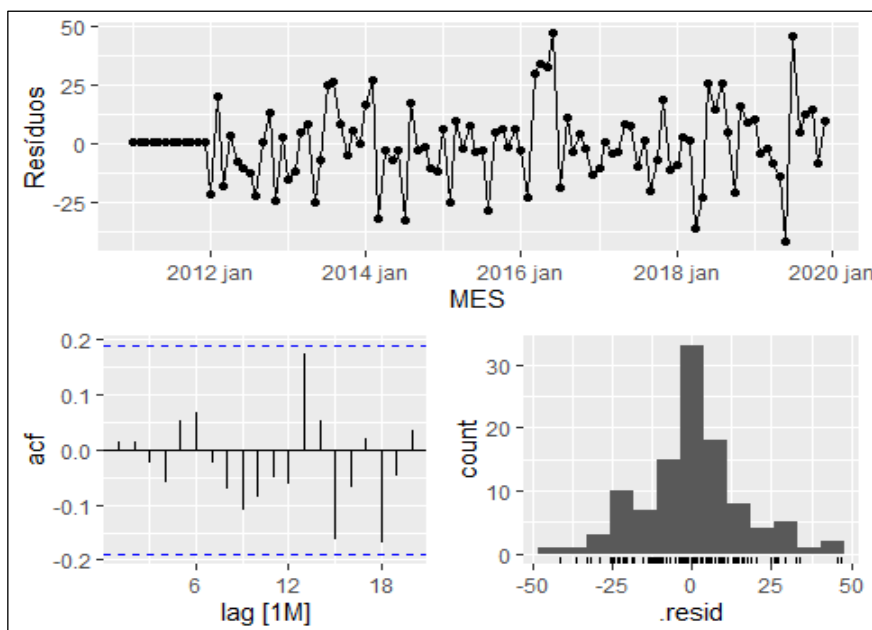
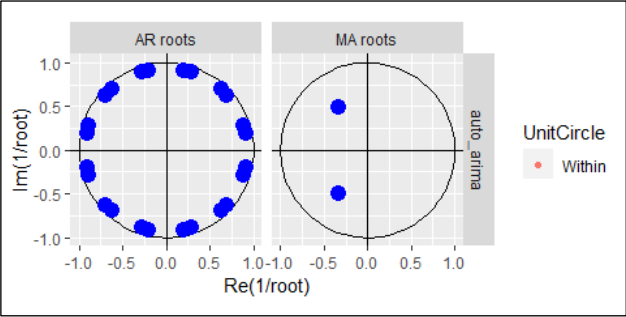


Figura A.29
Inverso das raízes unitárias do modelo escolhido para a
Região Sul — 2011-19



Quadro A.55
Teste Ljung-Box (lag = 36) com sete graus de
liberdade para a Região Sul — 2011-19

REGIAO	.model	lb_stat	lb_pvalue
Sul	auto_arima	47.47402	0.029536

Apêndice II

Informações sobre o R

Informações geradas com sessionInfo

```
## R version 4.2.3 (2023-03-15 ucrt)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 22621)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=Portuguese_Brazil.utf8 LC_CTYPE=Portuguese_Brazil.utf8
## [3] LC_MONETARY=Portuguese_Brazil.utf8 LC_NUMERIC=C
## [5] LC_TIME=Portuguese_Brazil.utf8
##
## attached base packages:
## [1] stats graphics grDevices utils datasets methods base
##
## other attached packages:
## [1] knitr_1.42 beepr_1.3 rvcheck_0.2.1 ggrepel_0.9.3
## [5] highcharter_0.9.4 remotes_2.4.2 scales_1.2.1 reshape_0.8.9
## [9] plotly_4.10.1 forcats_1.0.0 purrr_1.0.1 readr_2.1.4
## [13] tidyr_1.3.0 tibble_3.2.1 ggplot2_3.4.1 tidyverse_2.0.0
## [17] lubridate_1.9.2 tsibble_1.1.3 feasts_0.3.1 fable_0.3.3
## [21] fabletools_0.3.2 stringr_1.5.0 readxl_1.4.2 tidyselect_1.2.0
## [25] vars_1.5-9 strucchange_1.5-3 sandwich_3.0-2 MASS_7.3-58.2
## [29] lmtest_0.9-40 forecast_8.21 quantmod_0.4.20 dplyr_1.1.1
## [33] TTR_0.24.3 xts_0.13.0 zoo_1.8-11 urca_1.3-3
## [37] microdatasus_2.1.0
##
## loaded via a namespace (and not attached):
## [1] nlme_3.1-162 httr_1.4.5 tools_4.2.3
## [4] backports_1.4.1 utf8_1.2.3 R6_2.5.1
## [7] lazyeval_0.2.2 colorspace_2.1-0 nnet_7.3-18
## [10] withr_2.5.0 curl_5.0.0 compiler_4.2.3
## [13] progressr_0.13.0 cli_3.6.0 tseries_0.10-53
## [16] fracdiff_1.5-2 quadprog_1.5-8 yulab.utils_0.0.6
## [19] digest_0.6.31 rmarkdown_2.21 pkgconfig_2.0.3
## [22] htmltools_0.5.5 fastmap_1.1.1 htmlwidgets_1.6.2
## [25] rlang_1.1.0 rstudioapi_0.14 farver_2.1.1
## [28] generics_0.1.3 jsonlite_1.8.4 distributional_0.3.2
## [31] magrittr_2.0.3 rlist_0.4.6 Rcpp_1.0.10
## [34] munsell_0.5.0 fansi_1.0.4 lifecycle_1.0.3
## [37] stringi_1.7.12 yaml_2.3.7 plyr_1.8.8
## [40] grid_4.2.3 parallel_4.2.3 lattice_0.20-45
## [43] hms_1.1.3 anytime_0.3.9 pillar_1.9.0
## [46] igraph_1.4.1 glue_1.6.2 evaluate_0.20
## [49] BiocManager_1.30.20 data.table_1.14.8 vctrs_0.6.1
## [52] tzdb_0.3.0 cellranger_1.1.0 gtable_0.3.3
## [55] assertthat_0.2.1 xfun_0.38 broom_1.0.4
## [58] viridisLite_0.4.1 audio_0.1-10 timeDate_4022.108
## [61] timechange_0.2.0 ellipsis_0.3.2
```