

**UNIVERSIDADE FEDERAL DE MATO GROSSO
INSTITUTO FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA AMBIENTAL**

**MODELAGEM DE SÉRIES TEMPORAIS DA PROPAGAÇÃO DO COVID-19
UTILIZANDO SARIMA COM PARÂMETROS METEOROLÓGICOS: UM
ESTUDO DE CASO NOS 3 MAIORES CENTROS URBANOS DE MATO GROSSO**

RODRIGO ESTEVES ASCURRA

PROF. DR. CARLO RALPH DE MUSIS

**Cuiabá - MT
maio, 2023**

**UNIVERSIDADE FEDERAL DE MATO GROSSO
INSTITUTO FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA AMBIENTAL**

**MODELAGEM DE SÉRIES TEMPORAIS DA PROPAGAÇÃO
DO COVID-19 UTILIZANDO SARIMA COM PARÂMETROS
METEOROLÓGICOS: UM ESTUDO DE CASO NOS 3 MAIORES CENTROS
URBANOS DE MATO GROSSO**

RODRIGO ESTEVES ASCURRA

*Tese apresentada ao Programa de Pós-Graduação
em Física Ambiental da Universidade Federal de
Mato Grosso, como parte dos requisitos para
obtenção do título de Doutor em Física Ambiental.*

**PROF. DR. CARLO RALPH DE MUSIS
ORIENTADOR**

**Cuiabá - MT
maio, 2023**

Dados Internacionais de Catalogação na Fonte.

A815m Ascurra, Rodrigo Esteves Ascurra.

Modelagem de Séries Temporais da Propagação do COVID-19 Utilizando SARIMA com Parâmetros Meteorológicos: Um Estudo de Caso nos 3 Maiores Centros Urbanos de Mato Grosso [recurso eletrônico] / Rodrigo Esteves Ascurra Ascurra. -- Dados eletrônicos (1 arquivo : 92 f., il. color., pdf). -- 2023.

Orientador: Carlo Ralph de Musis Musis.

Tese (doutorado) - Universidade Federal de Mato Grosso, Instituto de Física, Programa de Pós-Graduação em Física Ambiental, Cuiabá, 2023.

Modo de acesso: World Wide Web: <https://ri.ufmt.br>.

Inclui bibliografia.

1. Série Temporal, modelos preditivos, Box-Jenkins, dados meteorológicos, ARIMA. I. Musis, Carlo Ralph de Musis, *orientador*. II. Título.

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

Permitida a reprodução parcial ou total, desde que citada a fonte.



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE MATO GROSSO
PRÓ-REITORIA DE ENSINO DE PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA AMBIENTAL

FOLHA DE APROVAÇÃO

TÍTULO: Modelagem de Séries Temporais da Propagação do Covid-19 Utilizando Sarima com Parâmetros Meteorológicos: Um Estudo de Caso nos 3 Maiores Centros Urbanos de Mato Grosso

AUTOR: DOUTORANDO RODRIGO ESTEVES ASCURRA

Tese defendida e aprovada em **28 de ABRIL de 2023**.

COMPOSIÇÃO DA BANCA EXAMINADORA

1. Prof. Dr. CARLO RALPH DE MUSIS (Presidente Banca / **ORIENTADOR**)

INSTITUIÇÃO: UNIVERSIDADE DE CUIABÁ

2. Profa. Dra. FLÁVIA MARIA DE MOURA SANTOS (Membro Interno)

INSTITUIÇÃO: UNIVERSIDADE FEDERAL DE MATO GROSSO

3. Prof. Dr. SÉRGIO ROBERTO DE PAULO (Membro Interno)

INSTITUIÇÃO: UNIVERSIDADE FEDERAL DE MATO GROSSO

4. Prof. Dr. ANTONIO RAMOS CORREIA (Membro Externo)

INSTITUIÇÃO: POLITEC-MT

5. Prof. Dr. JONATHAN WILLIAN ZANGESKI NOVAIS (Membro Externo)

INSTITUIÇÃO: UNIVERSIDADE DE CUIABÁ

CUIABÁ, 28/04/2023.



Documento assinado eletronicamente por **SERGIO ROBERTO DE PAULO, Coordenador(a) de Programas de Pós-Graduação em Física Ambiental - IF/UFMT**, em 28/04/2023, às 17:08, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **FLAVIA MARIA DE MOURA SANTOS, Docente da Universidade Federal de Mato Grosso**, em 28/04/2023, às 17:25, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **CARLO RALPH DE MUSIS, Usuário Externo**, em 28/04/2023, às 20:21, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Jonathan Willian Zangeski Novais, Usuário Externo**, em 02/05/2023, às 12:20, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Antonio Ramos Correia, Usuário Externo**, em 02/05/2023, às 22:18, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufmt.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5726705** e o código CRC **6FE11F2E**.

DEDICATÓRIA

À Deus pela vida e oportunidade, e aos queridos familiares, pelo amor, apoio e compreensão nas horas difíceis.

AGRADECIMENTOS

- Agradeço a Deus pelo privilégio de poder realizar este sonho e ter me fortalecido, para não desanimar no meio deste árduo caminho.
- Ao Prof. Dr. Carlo Ralph de Muis, pela orientação e principalmente pelo incentivo, apoio, pela confiança e grande amizade que possibilitaram a realização deste trabalho.
- À Secretaria de Estado de Saúde de Mato Grosso, pelo apoio para que a pesquisa se concretizasse.

SUMÁRIO

LISTA DE FIGURAS	IX
LISTA DE TABELAS	XII
RESUMO	XIV
ABSTRACT	XV
1. INTRODUÇÃO.....	1
1.1. PROBLEMÁTICA.....	3
1.2. JUSTIFICATIVA.....	3
1.3. OBJETIVOS.....	4
1.3.1OBJETIVO GERAL.....	4
1.3.2OBJETIVOS ESPECÍFICOS	4
2. REVISÃO BIBLIOGRÁFICA	5
2.1 O CLIMA	5
2.2 A COVID E O CLIMA	6
2.3 ESTUDO DE MODELOS PARA SERIES TEMPORAIS	11
3. MATERIAL E MÉTODOS	20
3.1 CARACTERIZAÇÃO DA ÁREA DE ESTUDO.....	20
3.2 DADOS METEOROLÓGICOS	26
3.3 BASE DE DADOS DA PROGRESSÃO DA PANDEMIA	26
3.4 METODOLOGIA BOX & JENKINS.....	28
4. RESULTADOS E DISCUSSÕES.....	33
4.1 CARACTERÍSTICAS GERAIS DAS VARIÁVEIS	33
4.1.1 INFECTADOS	33
4.1.2 TEMPERATURA DO AR.....	42
4.1.3 UMIDADE RELATIVA DO AR	46
4.1.4 PRECIPITAÇÃO PLUVIOMÉTRICA	49
4.2 MODELO AUTORREGRESSIVO.....	52
4.3 PREVISÃO DA SÉRIE TEMPORAL DOS INFECTADOS	63
5. CONSIDERAÇÕES FINAIS	68
6. REFERÊNCIAS BIBLIOGRÁFICAS	70

LISTA DE FIGURAS

Figura 1 - Mapa do Brasil, com climas tropicais e subtropicais.	8
Figura 2 - Ciclo interativo de ARIMA.	15
Figura 3 - Mato Grosso, localização no Brasil e na América do Sul.	20
Figura 4 - Densidade demográfica do Estado de Mato Grosso.	21
Figura 5 - Mapa de Mato Grosso com a localização dos municípios em destaque.	22
Figura 6 - Biomas do Estado de Mato Grosso.	23
Figura 7 - Dados meteorológicos, site do Instituto Nacional de Meteorologia.	26
Figura 8 - Tela do Site da Secretaria de Saúde do Estado de Mato Grosso que trazem as informações relativas à progressão da pandemia.	27
Figura 9 - Fluxograma do processo sequencial utilizado na análise das variáveis do modelo ARIMA.	29
Figura 10 - Série temporal dos dados brutos da variável Infectados do município de Cuiabá de 14/03/2020 a 27/09/2020.	33
Figura 11 - Série temporal da variável Infectados do município de Cuiabá de 14/03/2020 a 27/09/2020, com preenchimento de falhas.	34
Figura 12 - Série temporal dos dados brutos da variável Infectados, município de Várzea Grande, de 16/03/2020 a 27/09/2020.	34
Figura 13 - Série temporal da variável Infectados do município de Várzea Grande de 16/03/2020 a 27/09/2020.	35
Figura 14 - Série temporal dos dados brutos da variável Infectados para o município de Rondonópolis de 30/03/2020 a 27/09/2020.	35
Figura 15 - Série temporal com correção de falhas, da variável Infectados do município de Rondonópolis de 30/03/2020 a 27/09/2020.	36
Figura 16 - Histograma da série Infectados (Cuiabá-MT).	37
Figura 17 - Boxplot da série infectados.	37
Figura 18 - Histograma da série Infectados (Várzea Grande-MT).	38
Figura 19 - <i>Boxplot</i> da série infectados.	38
Figura 20 - Histograma da série Infectados (Rondonópolis-MT).	39
Figura 21 - <i>Boxplot</i> da série infectados.	39
Figura 22 - Decomposição da variável infectados utilizando transformada de <i>Fourier</i> para Cuiabá.	41

Figura 23 - Decomposição da variável infectados utilizando transformada de <i>Fourier</i> para Várzea Grande.....	41
Figura 24 - Decomposição da variável infectados	41
Figura 25 – Teste de estacionariedade série infectados (Cuiabá-MT).	41
Figura 26 - Teste de estacionariedade série infectados (Várzea Grande-MT).	41
Figura 27 - Teste de estacionariedade série.....	42
Figura 28 - Série temporal de temperatura média diária do ar do município de Cuiabá e Várzea Grande, período de 14/03/2020 a 27/09/2020.....	43
Figura 29 - Série temporal de temperatura média diária do ar do município de Rondonópolis, período de 30/03/2020 a 27/09/2020.....	43
Figura 30 - Histograma da temperatura do ar do município de Cuiabá/Várzea Grande.	44
Figura 31 - <i>Box-plot</i> da temperatura do ar do município de Cuiabá/Várzea Grande.	44
Figura 32 - Histograma da temperatura do ar do município de Rondonópolis.	44
Figura 33 - <i>Box-plot</i> da temperatura do ar do município de Rondonópolis.	44
Figura 34 - Decomposição da variável temperatura do ar do município de Cuiabá-MT.	45
Figura 35 - Decomposição da variável temperatura do ar do município de Rondonópolis-MT.	45
Figura 36 - Gráfico da série temporal da umidade relativa do ar, dos municípios de Cuiabá e Várzea Grande, durante o período de 14/03/2020 a 27/09/2020.	46
Figura 37 - Gráfico da série temporal da umidade relativa do ar, dos municípios de Rondonópolis, durante o período de 30/03/2020 a 27/09/2020.	46
Figura 38 - Histograma da umidade Cuiabá e Várzea Grande.	47
Figura 39 - <i>Box-plot</i> da umidade de Cuiabá e Várzea Grande.	47
Figura 40 - Histograma da umidade relativa do ar Rondonópolis.....	48
Figura 41 - <i>Boxplot</i> da umidade relativa do ar de Rondonópolis.....	48
Figura 42 - Decomposição da variável umidade média do município de Cuiabá-MT.....	49
Figura 43 - Decomposição da variável umidade média do município de Rondonópolis.....	49
Figura 44 - Série temporal da precipitação pluviométrica dos municípios de Cuiabá e Várzea Grande, durante o período de 14/03/2020 a 27/09/2020.....	50
Figura 45 - Histograma da precipitação pluviométrica de Cuiabá e Várzea Grande.	50
Figura 46 - <i>Box-plot</i> da precipitação pluviométrica de Cuiabá e Várzea Grande.....	50
Figura 47 - Decomposição da variável precipitação pluviométrica dos municípios de Cuiabá e Várzea Grande.....	51

- Figura 48** - Funções da autocorrelação (à esquerda) e autocorrelação parcial (à direita) da série temporal de infectados, registrado no período de 14/03/2020 a 27/09/2020 em Cuiabá-MT.....52
- Figura 49** - Funções da autocorrelação (à esquerda) e autocorrelação parcial (à direita) da série temporal de infectados, registrado no período de 16/03/2020 a 27/09/2020 em Várzea Grande-MT.....53
- Figura 50** - Funções da autocorrelação (à esquerda) e autocorrelação parcial (à direita) da série temporal de infectados, registrado no período de 30/03/2020 a 27/09/2020 em Rondonópolis-MT.....53
- Figura 51** - Funções de autocorrelação (ACF) e autocorrelação parcial (PACF) diff 1 da série temporal de infectados, registrado no período de 14/03/2020 a 27/09/2020 em Cuiabá-MT.....54
- Figura 52** - Funções de autocorrelação (ACF) e autocorrelação parcial (PACF) diff 1 da série temporal de infectados, registrado no período de 16/03/2020 a 27/09/2020 em Várzea Grande-MT.....54
- Figura 53** - Funções de autocorrelação (ACF) e autocorrelação parcial (PACF) diff 1 da série temporal de infectados, registrado no período de 30/03/2020 a 27/09/2020 em Rondonópolis-MT.....54
- Figura 54** - Gráfico de Previsão da série temporal Infectados (linha preta), do município de Cuiabá-MT, em covariância com a variável precipitação, com previsão (linha azul) de 5 dias, e intervalos de confiança de 80% (área azul claro) e 95% (área cinza claro).64
- Figura 55** - Gráfico de previsão da série temporal Infectados (linha preta), do município de Várzea Grande-MT, em covariância com a variável umidade relativa do ar, com previsão (linha azul) de 5 dias, e intervalos de confiança de 80% (área azul claro) e 95% (área cinza claro).65
- Figura 56** - Gráfico de previsão da série temporal Infectados (linha preta), do município de Rondonópolis-MT, com previsão de 5 dias (linha azul), e intervalos de confiança de 80% (área azul claro) e 95% (área cinza claro).67

LISTA DE TABELAS

Tabela 1 - População, área, Localização, Clima e Temperatura Média Anual das 4 cidades mais populosas de Mato Grosso.	23
Tabela 2 - Resumo estatístico da variável infectados dos municípios de Cuiabá, Várzea Grande e Rondonópolis.....	36
Tabela 3 - Resumo estatístico da variável temperatura do ar em Cuiabá, Várzea Grande e Rondonópolis.	43
Tabela 4 - Resumo estatístico da variável umidade relativa do ar em Cuiabá, Várzea Grande e Rondonópolis.	47
Tabela 5 - Resumo estatístico da variável precipitação pluviométrica, Cuiabá e Várzea Grande.	50
Tabela 6 - Quadro com as covariáveis utilizadas no modelo Auto-Arima agregadas a série temporal da variável Infectados da região de Cuiabá-MT.	55
Tabela 7 - Quadro com as covariáveis utilizadas no modelo Auto-Arima agregadas a série temporal da variável Infectados da região de Várzea Grande-MT.	59
Tabela 8 - Quadro com as covariáveis utilizadas no modelo Auto-Arima agregadas a série temporal da variável Infectados da região de Rondonópolis-MT.....	61
Tabela 9 - Tabela com a previsão da variável infectados de Cuiabá-MT, para 5 dias.	63
Tabela 10 – Tabela com a previsão da variável infectados para Várzea Grande-MT.....	64
Tabela 11 - Tabela com a previsão da variável infectados para Rondonópolis-MT.	66

LISTA DE ABREVIATURAS

ARIMA -Autoregressive integrated moving average

RSS -Root sum squared.

ACF -Auto-correlation function

PACF -Partial auto-correlation function

ADCF -Augmented Dickey-Fuller Test

SES – Secretária de Estado de Saúde

RESUMO

ASCURRA, R. E. **Modelagem de Séries Temporais da Propagação do Covid-19 Utilizando Sarima com Parâmetros Meteorológicos: Um Estudo de Caso nas 3 Cidades Mais Populosas de Mato Grosso.** 2023. 92f. Tese (Doutorado em Física Ambiental), Instituto Física, Programa De Pós-Graduação em Física Ambiental, Universidade Federal de Mato Grosso, Cuiabá.

O período correspondente da pandemia da Covid 19 teve início no final de 2019 e se estendeu pelos demais anos de 2020 a 2022. Chegou ao Brasil no final de fevereiro e se propagou rapidamente por todo o país, em Mato Grosso, o primeiro caso de transmissão comunitária ocorreu em 19/03/2020. Todavia, assim que as autoridades competentes tomaram conhecimento dos fatos, estas implementaram estratégias de mitigação e ações da vigilância epidemiológica, através de estratégias operacionais pelos serviços de saúde no desenvolvimento de suas ações preventivas e assistenciais. Surgindo assim, a necessidade de compreender a evolução da doença no estado de MT, como instrumento de auxílio das ações de combate e controle da pandemia. Desta forma, a presente pesquisa apresenta um modelo de previsão empregando técnicas de modelagem de série temporal a dados, por intermédio da utilização de base de dados epidemiológicos da Secretaria de Estado de Saúde de Mato Grosso (SES), por intercessão da utilização de base de dados de ocorrências de pessoas infectadas, tendo variáveis meteorológicas (temperatura do ar, umidade relativa do ar, e precipitação pluviométrica), como auxiliares, para a construção de modelos preditivos e verificar como estas variáveis explicam e refinam a previsão da trajetória do avanço do vírus, com o intuito de identificar o mecanismo estruturante da série e, por conseguinte, extrair periodicidades pertinentes, pormenorizar seu comportamento e desenvolver modelos preditivos. Os procedimentos foram desenvolvidos na linguagem R. A partir da avaliação dos pressupostos estatísticos, da sazonalidade e da validação, foi selecionado um modelo SARIMA para cada cidade: Cuiabá, ARIMA (4,1,0), a covariância com a precipitação pluviométrica obteve o melhor ajuste; Várzea grande, ARIMA (1,1,1), a previsão dos dados relacionados com a variável umidade relativa do ar é o modelo foi o mais adequado; Rondonópolis, ARIMA(1,1,0), a variável "Infectados" foi o melhor modelo a ser utilizado para descrever a dinâmica da série temporal na região no período analisado. Dessa forma, a aplicação do modelo deve ser interpretada como uma ferramenta de auxílio para as ações dos serviços de vigilância epidemiológica nas políticas públicas das gestões governamentais.

Palavras-Chave: Série Temporal, modelos preditivos, Box-Jenkins, dados meteorológicos, ARIMA.

ABSTRACT

ASCURRA, R. E. **Time Series Modeling of the Propagation of Covid-19 Using Sarima with Meteorological Parameters: A Case Study in the 3 Most Populous Cities of Mato Grosso.** 2023, 92f. Thesis (Doctorate in Environmental Physics), Institute of Physics, Graduate Program in Environmental Physics, Federal University of Mato Grosso, Cuiabá.

The corresponding period of the Covid 19 pandemic began at the end of 2019 and lasted for the other years from 2020 to 2022. It arrived in Brazil at the end of February and spread quickly throughout the country, in Mato Grosso, the first case of transmission community took place on 03/19/2020. However, as soon as the competent authorities became aware of the facts, they implemented mitigation strategies and epidemiological surveillance actions, through operational strategies by the health services in the development of their preventive and assistance actions. Thus, the need to understand the evolution of the disease in the state of MT arises, as an instrument to aid actions to combat and control the pandemic. In this way, the present research presents a forecast model employing techniques of modeling of time series data, through the use of epidemiological database of the State Department of Health of Mato Grosso (SES), through the intercession of the use of database data from occurrences of infected people, with meteorological variables (air temperature, relative humidity, and rainfall), as auxiliaries, for the construction of predictive models and to verify how these variables explain and refine the prediction of the trajectory of the virus's advance, with the aim of identifying the structuring mechanism of the series and, therefore, extracting relevant periodicities, detailing its behavior and developing predictive models. The procedures were developed in the R language, based on the evaluation procedures of statistical assumptions, seasonality and validation, a SARIMA model was selected for each city: Cuiabá, ARIMA (4,1,0), the covariance with rainfall obtained the best fit; Várzea Grande, ARIMA (1,1,1), the prediction of data related to the relative humidity variable is the most adequate model; Rondonópolis, ARIMA(1,1,0), the variable "Infected" was the best model to be used to describe the dynamics of the temporal series in the region in the analyzed period. Thus, the application of the model should be interpreted as an aid tool for the actions of epidemiological surveillance services in public policies of government administrations.

Keywords: Time series, predictive models, Box-Jenkins, meteorological data, ARIMA.

1. INTRODUÇÃO

A Organização Mundial da Saúde (OMS, 2020) através dos meios de comunicação declarou em março de 2020 a pandemia do novo coronavírus (Sars-Cov-2), este que se tornaria um dos maiores desafios do mundo em questão de saúde pública, devido ao grande crescimento mundial do número de infectados e dos óbitos motivados pela disseminação do vírus. Desta forma, não demorou para que o Estado de Mato Grosso (SES, 2020) tivesse confirmado o seu primeiro caso de COVID-19, numa noite de quinta-feira (19/03/2020), com o primeiro óbito registrado no início de abril do mesmo ano. Todavia, assim que as autoridades competentes tomaram conhecimento dos fatos, estas implementaram estratégias de mitigação e ações da vigilância epidemiológica, através de estratégias operacionais pelos serviços de saúde no desenvolvimento de suas ações preventivas e assistenciais. Surgindo assim, a necessidade de compreender a evolução da doença no estado de MT, como instrumento de auxílio para que estas autoridades públicas pudessem realizar o manejo das ações de combate e controle da pandemia. Pois bem, com relação a transmissão da COVID-19, o seu contágio se dá por diversas formas, sendo: contato direto, por saliva, aerossol, fezes e urina; que definem uma alta taxa de propagação, sendo também influenciadas pelo ambiente, em destaque, as condições meteorológicas, que são elementos que podem potencializar a transmissão de doenças infectocontagiosas virais (GONZAGA E FREITAS, 2020).

Sendo assim, a previsão acerca dos números da propagação da COVID-19 deve ser considerada no embasamento do estudo dos elementos estatísticos preditivos para a concepção das análises temporais dos dados relativos ao andamento da pandemia. Porém, o surto durante o andamento da pandemia representou um desafio para os modeladores, uma vez que existiam dados limitados sobre a trajetória inicial de crescimento e das características epidemiológicas do novo coronavírus.

Deste modo, o estudo das séries temporais representa etapa determinante para a elaboração de modelos preditivos, e estes estudos necessitam vislumbrar algumas variáveis que influenciem a qualidade da predição, em que o arranjo de metodologias e procedimentos de uma área em que seu uso é clássico (dados meteorológicos) para outra em que se leva em consideração os dados relativos à propagação da COVID-19. Pretende-se, assim, compreender técnicas de modelagem de séries temporais aos dados epidemiológicos da SES, por intercessão da utilização de base de dados de ocorrências de pessoas infectadas, dos recuperados e dos óbitos, tendo variáveis meteorológicas (temperatura do ar, umidade relativa do ar, e

precipitação pluviométrica), como auxiliares, para a construção de modelos preditivos e verificar como estas variáveis explicam e refinam a previsão da trajetória do avanço do vírus, com o intuito de identificar o mecanismo estruturante da série e, por conseguinte, extrair periodicidades pertinentes, pormenorizar seu comportamento e desenvolver modelos preditivos.

Esta tese se orientará a partir das seguintes etapas: revisão bibliográfica, composição das bases de dados, desenvolvimento dos modelos preditivos, validação da modelagem e a análise dos resultados.

A revisão bibliográfica será elaborada a partir da pesquisa em artigos científicos e publicações contendo contribuições de autores que tenham trabalhado com as condições meteorológicas, as relacionando com as infecções respiratórias virais, como a COVID-19.

A segunda etapa iniciou-se pelo levantamento das bases de dados com as informações relativas à progressão da pandemia para o estudo foi compilada a partir dos relatórios disponibilizados publicamente no site da Secretaria de Saúde do Estado de Mato Grosso (SES/MT) e variáveis meteorológicas mantidas e disponibilizadas pelo Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) do Instituto Nacional de Pesquisas Espaciais e do Instituto Nacional de Meteorologia (INMET).

A modelagem estatística computacional foi desenvolvida partindo-se da apreciação dos dados do passado com intuito de realizar a predição, a curto prazo (5 dias), dos números acerca da propagação da COVID-19. Também buscou a identificação dos dados que se diferenciam drasticamente de todos os outros (*outliers*), e análise dos recursos de autocorrelação e médias móveis, sendo todos os procedimentos estatísticos desenvolvidos com a utilização da linguagem R, no ambiente de processamento em nuvem mantido pelo *RStudio*.

Em seguida foi realizada a avaliação da estacionariedade (parâmetro d , transformação Box-Cox) e estabilização das variáveis. Com avaliação da modelagem utilizando o AIC (*Akaike Information Criterion*), via pacote *forecast*, afim de se, estabelecer este critério de informação para mensurar a qualidade do modelo estatístico. Assim, o mesmo permitirá a estimativa de previsões de 5 dias, das 3 cidades mais populosas de MT, assim como a avaliação da pertinência ao modelo das variáveis climáticas utilizadas a partir deste critério AIC.

1.1. PROBLEMÁTICA

Uma das problemáticas que surgiram com a propagação do *cluster* inicial dos casos de COVID-19, foi a falta de uma análise e um acompanhamento mais criterioso e aprofundado que levasse em consideração outras variáveis que se relacionassem com os dados dos números de infectados, recuperados e também os óbitos registrados em decorrência do vírus, e que pudessem apresentar uma maior confiabilidade e qualidade dos estudos sobre a compreensão da evolução da doença no estado de MT com base em dados estatísticos. No caso, as condições meteorológicas são elementos que podem potencializar a transmissão de doenças infectocontagiosas virais.

Pois, este surto na época em que se encontrava em andamento veio a representar um desafio para as modelagens estatísticas preditivas de séries temporais, uma vez que existiam dados limitados sobre a trajetória inicial de crescimento e as características epidemiológicas do novo coronavírus, e estes ainda não tinham sido totalmente elucidadas.

1.2. JUSTIFICATIVA

Desta forma, é de grande importância a necessidade de estudos que possam fomentar instrumentos de auxílio preditivo e automatizado de séries temporais para que as autoridades públicas possam realizar o manejo das ações de combate e controle da pandemia, e assim implementar estratégias de mitigação e ações dos serviços de vigilância epidemiológica. Contudo, o que se observa no presente tempo é que as pesquisas conduzidas não dispõem de metodologia favorável para evidenciar o potencial influência das variáveis meteorológicas na difusão do novo coronavírus.

Apesar de conhecermos muito pouco sobre como o vírus se comporta, as descobertas científicas até o momento, nos permitem perceber de forma evidente a relação direta da origem, disseminação, transmissibilidade, e óbitos causados pelo Sars-coV-2 com os parâmetros meteorológicos e também com a degradação e poluição ambiental (GONZAGA E FREITAS, 2020).

Neste sentido, a compreensão acerca dos mecanismos de transmissão do vírus Sars-coV-2 passou a ser largamente investigada por pesquisadores em todo o mundo. A pandemia do novo coronavírus, dentre outros efeitos sob o comportamento da humanidade, tem nos feitos refletir sobre a fragilidade da relação entre o homem e a natureza.

1.3. OBJETIVOS

1.3.1 Objetivo Geral

Analisar a influência dos parâmetros meteorológicos e escolher a melhor covariância na modelagem da previsão da série temporal da variável “infectados” da propagação da COVID-19 nas 3 cidades mais populosas de Mato Grosso, durante o período de 14 de março de 2020 a 27 de setembro de 2020, com emprego da modelagem estatística e computacional, com intuito de fornecer modelos estatísticos para previsão de curto prazo das séries temporais associadas a progressão da pandemia.

Espera-se que os resultados encontrados possam subsidiar ações de combate e controle da pandemia, e assim implementar estratégias de mitigação e ações dos serviços de vigilância epidemiológica, neste momento, e futuramente com novas infecções respiratórias causadas por novos vírus.

1.3.2 Objetivos Específicos

Antever a progressão da epidemia com modelos preditivos e compartimentais, com objetivo de colaborar na adoção de estratégias operacionais pelos serviços de saúde no desenvolvimento de suas ações preventivas e assistenciais.

Encontrar possíveis relações de dependências nos dados referenciados na escala temporal.

Identificar o mecanismo estruturante da base de uma série temporal, como os parâmetros meteorológicos, por conseguinte, extrair periodicidades pertinentes, pormenorizar seu comportamento e desenvolver modelos preditivos com melhor acurácia dentro de um intervalo de confiança.

2. REVISÃO BIBLIOGRÁFICA

2.1 O CLIMA

O clima é o conjunto de características e dinâmicas atmosféricas de uma determinada região, analisadas ao longo de um extenso período, que segundo Ayoade (2001) a síntese deste período dura aproximadamente 30-35 anos. Desta forma, o estudo científico do clima utiliza subdivisões, apesar dos vários fenômenos atmosféricos comporem um único espectro contínuo dos sistemas climáticos. As divisões utilizadas na climatologia possuem escalas relacionadas a sua abrangência nas escalas dos sistemas de circulação meteorológica, conforme são mostradas a seguir:

- a) macroclimatologia: aspectos dos climas de amplas áreas da Terra, com escala horizontal de abrangência na ordem de 5.000 a 20.000 km;
- b) mesoclimatologia: estudo do clima em áreas menores, entre 10 a 100 km de extensão, como por exemplo, o clima urbano;
- c) microclimatologia: estudo do clima próximo a superfície ou de áreas muito pequenas, com menos de 100 metros de extensão.

Pois bem, a atmosfera, camada de ar que envolve todo o planeta Terra composta por gases que não se dissipam, que se mantem por intermédio da gravidade, é representada por uma massa em movimento que vem a provocar variações nos parâmetros meteorológicos das regiões e o seu estado é apresentado por variáveis que caracterizam a condição energética (escala temporal), podendo ser instantânea ou condição média. Chama-se tempo à descrição instantânea, que é o estado da atmosfera em um local e instante determinado, sendo caracterizado pelas condições de temperatura, pressão, concentração de vapor, velocidade e direção do vento, precipitação. Enquanto que é denominado de clima a descrição média, que é valor mais provável, das condições atmosféricas nesse mesmo local, Pereira et al. (2002)

Segundo Romero (2001), o clima de um determinado local tem como denominação a associação dos elementos que se verificam em escalas diferentes, compreendendo desde a macro até a microescala. No caso da macroescala o clima regional cobre a maior parte de um continente e essas características regionais podem ser encontradas em latitudes análogas, vindo a estabelecer desta forma uma primeira ordem de características, como as temperaturas das estações e o regime das chuvas.

E com relação ao clima devido a presença da urbanização, as cidades se tornaram uma das manifestações mais óbvias da ação antrópica no meio ambiente. Ação esta que modifica o espaço rural, provocando alterações nas condições climáticas e ambientais, cuja as implicações são a degradação ambiental e a geração de um clima específico nos centros urbanos (LOMBARDO, 1985).

2.2 A COVID E O CLIMA

Embora o primeiro relato acerca do Coronavírus ter sido identificado em 1960, em função de epidemias que foram identificadas neste período, sabe-se que se trata de uma linhagem de vírus que ataca tanto o homem, como animais domésticos e selvagens. As notícias que se têm conhecimento e as informações que foram divulgadas até os dias de hoje, mostra que o início da pandemia ocorreu em um mercado de animais vivos, local este muito comum na China. Segundo Cohen (2020), ainda existem dúvidas sobre esta origem e a Organização Mundial da Saúde (OMS) almeja realizar novas investigações para determinar a verdadeira origem da pandemia.

De acordo com Prata et al. (2020), a epidemia respiratória COVID-19, que tem origem na família do coronavírus (2019-nCoV), causou comoção mundial em razão da possibilidade de rápida mortalidade de pessoas afetadas, que depende de suas condições e aparelhamentos gerais de saúde, além do aspecto sociodemográfico. Há questionamentos sobre a disseminação do vírus e se parâmetros meteorológicos ou características demográficas influenciariam em sua propagação, levando em consideração que uma variante semelhante, o vírus da gripe, foi afetado pelo clima.

Pois bem, desde março de 2020 o mundo convive em meio a pandemia da Covid 19, que acomete as pessoas de uma síndrome respiratória grave, esta qual é causada pelo SARS-CoV-2, um vírus derivado da família Coronavírus, seu início foi na cidade de Wuhan, na China, em novembro de 2019, inicialmente como uma epidemia com os primeiros casos confirmados em janeiro de 2020, (Franca et al., 2020). Sendo que, na data de 11 de março de 2020 a Organização Mundial da Saúde (OMS), diante da propagação mundial da doença, realiza a declaração que o planeta vivenciava uma Pandemia da Covid 19, fato qual veio a ser confirmado com o passar dos dias (Organização das Nações Unidas, [ONU], 2020).

Ademais, esta rápida disseminação do vírus em tão curto espaço de tempo tem desafiado estudiosos de diversas áreas do conhecimento, e na esfera do conhecimento da

climatologia, as publicações dos trabalhos publicados tanto no Brasil quanto no exterior têm apresentado variados resultados sobre a relação entre a disseminação da doença e variações na temperatura do ar.

Segundo Lipsitch, (2020), as pesquisas acerca da pandemia inicialmente apontavam para muitas dúvidas, até porque naquela época a pandemia estava acontecendo, basicamente no hemisfério norte na estação de inverno. Porém, com o passar dos dias a pandemia propagou-se para todo o planeta, inclusive em regiões nas quais estavam em término do verão. Desta forma, outras dúvidas surgiram, principalmente em função da sua disseminação em países e cidades de clima quente e úmido. No início a Covid 19 se disseminou por uma faixa de latitude entre 30° e 50° de latitude norte, com temperaturas médias entre 05°C e 11°C, (Sajadi et al., 2020).

Sagripani & Lytle (2020) comenta que, o tempo de sobrevivência em que o vírus da Covid 19 permanece ativo no ambiente, podendo infectar outras pessoas, depende das características da superfície e das condições observadas no ar, especialmente, com relação as condições meteorológicas, podendo citar três parâmetros climatológicos que podem ser destacados: a temperatura do ar, a umidade relativa do ar e a radiação solar.

Segundo Prata et al. (2020), análogo ao COVID-19, o coronavírus associado ao SARS (SARS-CoV) pode causar síndrome respiratória aguda grave (SARS), e também comenta que em outros trabalhos diversos, estudos entre 2007 e 2019, foram sinalizados que o SARS-CoV-2 e outras variantes de coronavírus relacionadas, como influenza e Ebola, tiveram relações significativas com os fatores climáticos. Também que alguns estudos (Chan et al., 2011; Van Doremalen et al., 2013; Hastie e Tibshirani, 1990) declararam que o tempo de sobrevivência dos coronavírus nas superfícies estão sujeitos ao aumento ou diminuição da temperatura, deste modo, a temperatura pode afetar o risco de transmissão do vírus. Portanto, é plausível e razoável conjecturar que as diferenças nas temperaturas médias anuais podem influenciar de forma significativa a transmissão do vírus. A prevalência de estudos disponíveis sobre a relação entre a transmissão do vírus e a temperatura tem sido realizada para países não tropicais (30° N e acima) em temperaturas que variam de -20 °C a um máximo de 20 °C. O Brasil é um país extenso e tropical, com a maior parte de seu território localizado entre o Trópico de Câncer, no Norte (cerca de 23° 26' N) até o Trópico de Capricórnio, no Sul (cerca de 23° 26' S). É interessante o alcance geográfico do Brasil, pois este permite comparar sub-regiões bem específicas, como as zonas climáticas temperadas, para examinar a disseminação do vírus em diferentes climas e territórios, Prata et al. (2020).

No Brasil, 93% da massa de terra está no Hemisfério Sul, e o restante (7%) está no Hemisfério Norte. Isso significa que o território está na zona tropical do planeta, com exceção da região sul, abaixo do Trópico de Capricórnio, correspondendo a 6,76% do território brasileiro, conforme mostra a Figura 01.

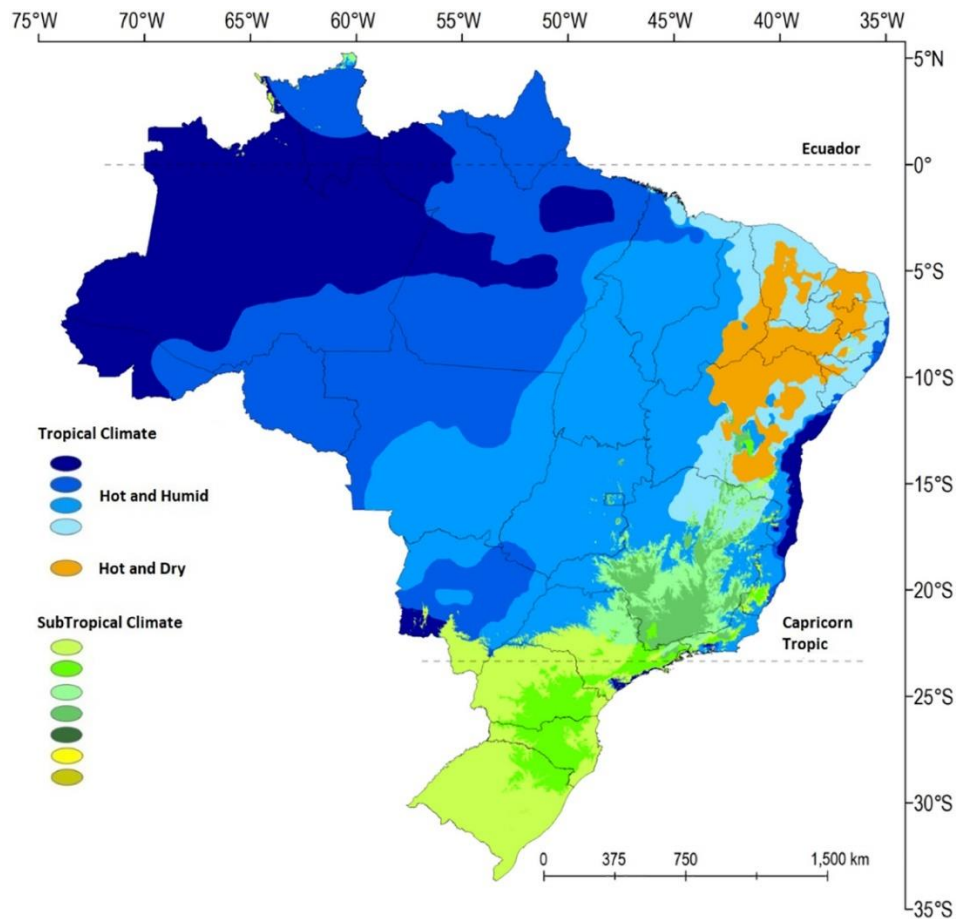


Figura 1 - Mapa do Brasil, com climas tropicais e subtropicais.
Fonte: Prata et al. (2020).

Segundo Sagripanti & Lytle, (2020), o coronavírus assim como outros vírus podem ser inativados pela radiação Ultra Violeta (UV), radiação esta que possui um comprimento de onda menor que a da luz visível e maior que a dos raios X, de 380 nm a 1 nm. Ou seja, através da análise dos valores de radiação que chegam à superfície em uma determinada região, podendo esta ser em maior ou menor quantidade em de, também pode ajudar a explicar uma maior ou menor propagação da Covid 19. Pois bem, a exposição ao sol do meio-dia por um período breve período de tempo, algo em torno de uma variação de 11 a 34 minutos pode causar

a inativação de 90% do vírus SARS-CoV-2, o que demonstra a sensibilidade do novo Coronavírus à radiação UV.

Outro ponto em questão, que deve ser analisado é a influência das nuvens na quantidade de radiação Ultravioleta que chega até a superfície terrestre, tema este que é abordada nos estudos feitos por Calbó, Pagès e González (2005), autores estes que mencionam e afirmam que as diferenciadas condições interferem na quantidade de radiação solar que atingem uma determinada área, condições estas como: ozônio, fatores astronômicos, aerossóis, albedo, nuvens e altitude. Portanto, em trabalhos que abordem este tema, deverão ser consideradas as nuvens, pois estas constituem um fator que provoca uma grande variabilidade na incidência da radiação UV, pois esta influência vai depender de sua densidade, espessura, altura e tempo de atuação no seu bloqueio. O novo Coronavírus pode ser inativado pela radiação UV, assim como outros vírus. Desta forma a análise dos índices de radiação que chegam à superfície em uma determinada região, em maior ou menor quantidade, também pode ajudar a explicar uma maior ou menor expansão da Covid 19 (Sagripanti & Lytle, 2020). Ainda segundo o autor, a sensibilidade do novo Coronavírus a radiação UV e a inativação de 90% do SARS-CoV-2 ocorre após exposição ao sol do meio-dia, por um período que pode variar entre 11 e 34 minutos.

Apesar de se pensar que a temperatura poderia ser um fator limitante na disseminação da COVID-19, estudos como o de Bukhari & Jameel (2020) mostram que a temperatura não impediu a disseminação do vírus em diferentes países, independente da estação do ano. No entanto, é importante destacar que a doença chegou a países tropicais e subtropicais durante o verão, como São Paulo, Manaus e Belém, que apresentavam temperaturas elevadas na época, de acordo com Prata, Rodrigues & Bermejo (2020) e Auler, Cássaro, Silva & Pires (2020).

Pois bem, na região ao sul do Brasil, de acordo com as suas características climáticas, perdurava a expectativa de que a propagação da pandemia seria visualizada de forma mais rápida, entretanto o que se percebeu foi uma maior demora de tempo para ao acometimento desta região em relação a disseminação do vírus, considerando que outras variáveis condicionantes que aparecem em paralelo devem ser levadas em conta, que neste caso provavelmente em virtude das medidas tomadas de isolamento e distanciamento social recomendadas pela OMS e pelo governo, assim como em função de um ano anômalo acerca das condições climáticas das quatro estações do ano, pois foram registradas menos chuvas e temperaturas acima da média histórica na estação verão, Cruz (2022). A estação de outono exibiu temperaturas abaixo da média, com o mês de maio entre os mais frios da série histórica,

registrando entradas de frentes frias. Também no inverno registraram valores de temperaturas e precipitações abaixo da média histórica.

Segundo Fedorova (2001), quando ocorre o aumento e avanço da área de contato entre as massas de ar frias e quentes sobre uma determinada região, acontece a frente fria, com aparecimento de chuvas frontais, com a presença de ventos variados, sendo que a temperatura tende a mostrar diminuição, com a sua pressão atmosférica vindo a se manter instável, e a formação dominante de nuvens baixas leva à instabilidade do tempo.

Os primeiros casos de Covid 19 no Brasil começaram a aparecer no começo de 2020, na pesquisa dirigida por Cruz (2022) sobre o Clima Urbano de Ponta Grossa passou a ser utilizada para ajudar a entender a ocorrência da doença na cidade, período esse que se estendeu, aproximadamente, do último terço do outono até a terceira semana da primavera, com o maior número de casos concentrados no inverno de 2020. Poderia se evidenciar a relação entre o clima e a Covid 19 à medida que as massas de ar frio foram chegando à região, com quedas na temperatura, e com a intensificação da nebulosidade, porém acontecia nesse momento uma liberação parcial do comércio, vindo a aumentar o contato entre as pessoas em virtude da diminuição do distanciamento e isolamento social. Desta forma, era então acreditado que os estados do Sul do país pudessem vivenciar o crescimento ritmo de propagação do vírus com a aproximação e chegada do inverno, fato este que se confirmou. Outro fato importante que corrobora na relação do clima com a disseminação vírus, foram que a pandemia se desenvolvia em outras regiões do país, mas foi no Sul que ela aconteceu de forma mais agressiva, provavelmente em virtude dos três primeiros meses do inverno e nas primeiras semanas da primavera.

Posteriormente, pesquisadores brasileiros como Prata et al. (2020) analisaram dados das 27 capitais brasileiras entre 27 de fevereiro e 1º de abril e concluíram que, para esse período, a cada 1º C de elevação da temperatura média anual do ar (até o limiar de 25.8°C) há redução em cerca de 4,8 % de casos da doença. Já Auler et al. (2020) investigaram a relação em 5 capitais brasileiras e encontraram resultados que sugerem que temperaturas e umidade relativa do ar médias mais altas favorecem a transmissão da COVID-19, diferente do que apontaram as demais pesquisas. O projeto de pesquisa elaborado pelo Laboratório de Climatologia Geográfica da Universidade de Brasília propõe a investigação do tema em diferentes unidades climáticas do país. O objetivo é analisar a variabilidade da curva epidêmica analogamente ao ritmo climático em cidades como Brasília, São Paulo e Manaus. O desenvolvimento de métodos para análise crítica das séries históricas é essencial. Dois dos maiores desafios nesse sentido

são a grande subnotificação e a defasagem entre a ocorrência da doença e o registro pelas secretarias de saúde. Além disso, mais além dos fatores climáticos, é necessário considerar que aspectos de ordem social apresentam importância-chave para a compreensão do avanço da doença no país. Espera-se, desse modo, contribuir para a elucidação do papel de atributos climáticos sobre a disseminação do novo vírus.

Desta forma, existe a importância do conhecimento sobre fenômenos climáticos e antrópicos para ações relacionadas a alterações no meio ambiente, pois essa compreensão pode ser relevante para entender a propagação da COVID-19, uma vez que há questionamentos sobre a influência de parâmetros meteorológicos na disseminação do vírus. Logo, o conhecimento dos fenômenos climáticos e antrópicos são importantes para as ações com relação as alterações no meio ambiente.

2.3 ESTUDO DE MODELOS PARA SERIES TEMPORAIS

Segundo Hamilton (1994), uma série temporal é uma coleção de observações feitas consecutivamente ao longo do tempo. Ou seja, dados que são coletados em sequência em um certo período, portanto, a observação em um determinado momento pode depender das observações anteriores. Em outras palavras, a ordem dos dados é importante e afeta a análise e interpretação da série temporal.

Desta forma, precisa-se de uma abordagem diferente dos modelos de regressão que consideram a independência das observações e não requerem ordenação entre as informações.

Ehlers (2009) mostra que existem muitos tipos diferentes de dados de séries temporais, desde a economia (produção industrial, taxas de juros, preços do mercado de ações) até a medicina (Eletrocardiograma, eletroencefalograma). Ele também menciona as principais características das séries temporais e sua complexidade:

- a) Séries temporais correlacionadas são mais difíceis de analisar e requerem técnicas específicas;
- b) A ordenação temporal das observações deve ser considerada;
- c) Fatores complicadores, como a presença de tendências e variações sazonais ou periódicas, podem ser complexas de estimar ou extrair;
- d) Os seguintes modelos de seleção podem ser muito complexos e as ferramentas podem ser difíceis de interpretar;

e) Dados discrepantes são mais difíceis de lidar devido à sua natureza sequencial.

O físico alemão A. Schuster (1906) é frequentemente citado como um dos trabalhos mais antigos sobre o uso de séries temporais no estudo de fenômenos naturais. Em seu trabalho, este propõe um modelo de circulação de 11 anos para explicar a atividade solar. Ferreira (2018) aponta que antes de iniciar os estudos de modelagem é necessário reconhecer diversos padrões comportamentais que muitas vezes são vistos em séries temporais e não necessariamente considerados de forma isolada:

- a) Tendência (T): componente observado quando uma série temporal segue uma direção específica (aumento ou diminuição), não necessariamente lineares;
- b) Sazonalidade (S): Um padrão que se repete em intervalos específicos dentro de um ano ou período de estudo;
- c) Ciclo (C): Não tem duração fixa e repete-se com alguma regularidade no período, como exemplo, tem-se os fenômenos meteorológicos.

A tendência contém todos os processos sistemáticos ou determinísticos de longo prazo, como um aumento linear, uma mudança de etapa ou um sinal sazonal, Mudelsee (2010).

As tendências contêm padrões como comportamento de longa duração, e podem possuir séries: com crescimento linear, constante para os dados, e com proporção linear; crescimento exponencial (um aumento percentual gradual ao longo do tempo que está diretamente relacionado a uma função exponencial); e crescimento suavizado (dados futuros menores que os dados atuais), Ehlers (2009).

A sazonalidade é um comportamento que tende a se repetir em diferentes momentos em uma série temporal. As variações sazonais são representadas como mudanças ao longo do padrão de comportamento da tendência, dependendo de suas características, em virtude das variáveis correlacionadas ao processo, Brocklebank & Dickey (2003).

Segundo Morettin e Tolo (2006), existe uma dificuldade em se definir quanto a sazonalidade, tanto conceitual quanto estatisticamente. Como exemplo, podemos citar acontecimentos que ocorrem regularmente a cada ano, como o aumento das vendas de passagens aéreas durante os meses de verão ou aumento das vendas no varejo durante as festas de final de ano, como o Natal, que são considerados sazonais.

A sazonalidade pode ser dividida em duas categorias, segundo Ehlers (2009):

- a) aditiva: série que apresenta variação sazonal estável sem considerar o nível geral da série);
- b) multiplicativa: variação sazonal depende do nível da série como um todo.

Os períodos em uma série temporal referem-se a movimentos cíclicos com valores de elevação e queda em torno do nível médio da tendência, portanto, também são valores relevantes que refletem a natureza de longo prazo das variáveis envolvidas (LAMOUNIER, 2007).

Deve -se enfatizar que as variáveis com elevação e queda dos valores em torno da tendência podem ser estritamente ou quase periódicos. Portanto, eles diferem em comprimento e tamanho, Souza (1989).

A análise de séries temporais pode ser elaborada sob vários focos objetivos que podem ser divididos em quatro grupos usando o exemplo de Chatfield (2013):

- a) Descrição: caracterização da tendência, sazonalidade, presença de outliers e mudança estrutural, com o objetivo de descrever o comportamento das séries;
- b) Explicação: Tem como objetivo identificar duas ou mais variáveis que contribuam para a covariância de uma série temporal, a fim de, explicar a variação relativa de uma variável com a outra;
- c) Previsão: tenta-se caracterizar o comportamento da série temporal em virtude dos valores descritos no passado, para sintetizar as propriedades da série e identifica e propor modelos que permitam a previsão de possíveis valores futuros;
- d) Controle: Nesta etapa de controle, busca determinar e mensurar a qualidade do processo de previsão a partir da análise do conjunto de valores representativos.

Segundo Lima (2015), a relevância dos eventos que ocorrem durante o estudo da série temporal tem como preocupação o domínio do tempo, e a relação entre as observações em momentos diferentes.

Uma metodologia amplamente utilizada para analisar padrões no domínio do tempo é a abordagem de Box e Jenkins (1970). Para busca modelar e prever o comportamento de uma série temporal, é comum utilizar modelos autorregressivos integrados de médias móveis (ARIMA) e modelos autorregressivos integrados de médias móveis sazonais (SARIMA), Lucas et al., 2009; Morettin e Troy, 2006).

O modelo ARIMA é um modelo estatístico que utiliza a regressão linear para estimar a relação entre uma série temporal e suas próprias defasagens. Segundo Brockwell e Davis (2016), o modelo ARIMA é capaz de modelar a autocorrelação em uma série temporal, bem como ajustar sazonalidade e tendências, possibilitando previsões precisas. O modelo é composto por três componentes principais: o termo autoregressivo (AR), o termo de média móvel (MA) e o termo de diferenciação (I). A combinação desses termos permite a modelagem de diferentes padrões de séries temporais.

Já o modelo SARIMA é uma extensão do modelo ARIMA que leva em consideração a sazonalidade presente em muitas séries temporais. Segundo Hyndman e Athanasopoulos (2018), o modelo SARIMA é capaz de modelar e prever séries temporais com padrões sazonais, como as vendas de varejo que têm picos sazonais durante as épocas de Natal e Páscoa. O modelo SARIMA tem como principal objetivo identificar e modelar padrões sazonais em dados históricos e, assim, prever valores futuros com base nesses padrões.

A aplicação desses modelos pode ser realizada por meio de softwares estatísticos, como o R e o Python, que possuem pacotes específicos para a implementação do modelo ARIMA e SARIMA.

Em suma, a utilização do modelo ARIMA e SARIMA é fundamental na análise e previsão de séries temporais, permitindo a identificação de padrões e tendências em dados sequenciais ao longo do tempo. A aplicação desses modelos são ferramentas úteis para modelar e prever o comportamento de séries temporais em diversas áreas do conhecimento, permitindo previsões precisas e valiosas para tomada de decisão.

Segundo Morettin e Toloí (2006), a construção de um modelo ARIMA ou SARIMA deve ser baseada no ciclo iterativo Box-Jenkins (BOX et al, 2015). As etapas seguem o diagrama abaixo:

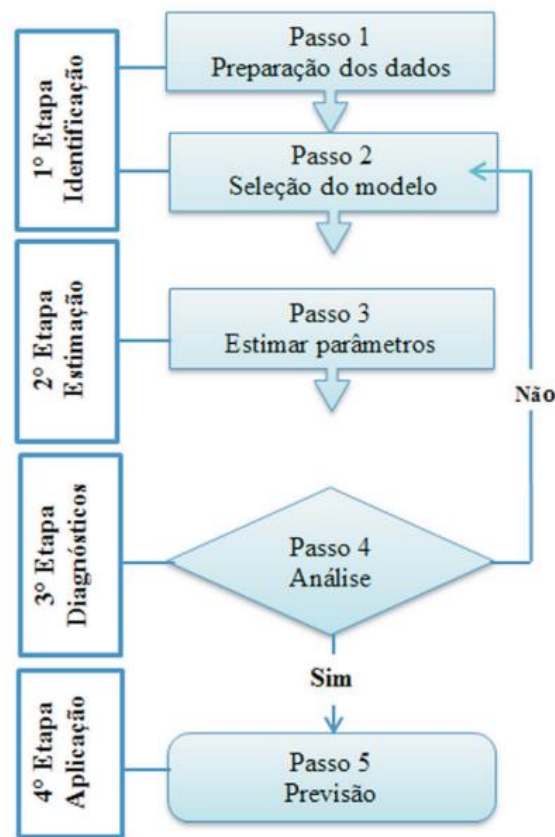


Figura 2 - Ciclo iterativo de ARIMA.

Fonte: Câmara *et al.* (2016).

Passo 1 (Preparação dos dados): nesta etapa é feita a preparação dos dados de acordo com as suas características;

Passo 2 (Seleção do modelo): A estrutura do modelo é escolhida de acordo com a característica dos dados. O ARIMA será o ideal no caso de a série contemplar o componente de tendência. Entretanto, se a série apresentar tanto tendência quanto sazonalidade, sugere-se a adoção de modelos mais complexos;

Passo 3 (Estimar parâmetros): Identificação das ordens do modelo, os valores de p , d e q , do ARIMA (p , d , q) são determinados pelos critérios de informação. Corresponde à determinação 1) do nível de diferenciação (d), a partir do qual a série se torna estacionária, ou seja, quando a sequência é diferenciada tantas vezes até que sua variância seja menor que a computada na versão original; 2) da ordem (máxima) dos termos autoregressivos (p); da ordem (máxima) dos termos médias-móveis (q) (GOMES, 1989). Também é possível pesquisar esses parâmetros utilizando como função objetivo, por exemplo, o Critério de Informação de Akaike (AIC);

Passo 4 (Análise): uma vez determinada a ordem do modelo (p, d, q) são estimados os seguintes parâmetros: 1) nível do processo; 2) parâmetros autoregressivos; 3) parâmetros médias-móveis; e 3) variância do ruído branco. Após a atribuição desses valores iniciais, os coeficientes são estimados otimizando-se a função de verossimilhança (MORETTIN e TOLOI, 2006);

Passo 5 (Previsão e diagnóstico): A verificação da adequação do modelo é efetuada em duas dimensões interrelacionadas, a saber 1) o exame do grau de ajustamento ou aderência do modelo - expresso em estatísticas como a variância do erro, o MSE (*Mean Square Error*) ou o MAPE (*Mean Absolute Percentage Error*), AIC e BIC; e 2) o exame da aleatoriedade dos resíduos (erros) do modelo - expresso nas suas funções de autocorrelação (GOMES, 1989).

Uma vez validado o modelo, pode ser criada uma função de previsão que não apenas forneça a previsão mais confiável dentro do horizonte de planejamento especificado, mas também forneça intervalos de confiança inferiores e superiores associados a níveis de significância predefinidos. É indicado a utilização de modelos com moderação, com o menor número possível de parâmetros devido ao teorema da dualidade, Box e Jenkins (1970). Segundo Sáfadi (2004), o bom ajuste dos modelos Box e Jenkins depende se a distribuição dos resíduos é ajustada ao ruído branco, ou seja, se os resíduos são variáveis aleatórias independentes e identicamente distribuídas.

Pois bem, o modelo ARIMA é uma mistura de duas partes: Auto regressão e média móvel, logo o modelo tem que conter as duas partes referenciadas, sendo assim a estrutura da equação matemática (1), é exibida logo abaixo:

$$Y_t = c + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} + \epsilon_t \quad (1)$$

Onde:

Y_t é o valor da série temporal no tempo t;

c é a constante do modelo (pode ser zero);

ϕ_1, \dots, ϕ_p são os coeficientes dos termos autorregressivos (AR) de ordem p;

Y_{t-1}, \dots, Y_{t-p} são os valores passados da série temporal, usados para modelar a correlação serial;

ϵ_t é o erro ou inovação no tempo t;

$\theta_1, \dots, \theta_q$ são os coeficientes dos termos de média móvel (MA) de ordem q;

$\epsilon_{t-1}, \dots, \epsilon_{t-q}$ são os erros passados

Segundo Shumway e Stoffer (2017), a equação (2) abaixo é usada para modelar séries temporais usando um modelo SARIMA, que é uma extensão do modelo ARIMA:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps})(1 - B)^d(1 - B^s)^D y_t \quad (2)$$

$$= (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)(1 + \theta_1 B^s + \theta_2 B^{2s} + \dots + \theta_q B^{qs}) \epsilon_t$$

Onde:

Y_t é a série temporal que estamos tentando modelar;

B é o operador de defasagem. Usa-se B para representar valores anteriores da série temporal;

p , d e q são os parâmetros do modelo ARIMA para as componentes autoregressiva, integrada e de média móvel, respectivamente;

s é o período da sazonalidade;

D é o número de diferenciações sazonais;

Φ e Θ são os coeficientes de autocorrelação e correlação cruzada, respectivamente.

A aplicação de modelos de séries temporais no domínio do tempo geralmente requer uma compreensão de algumas técnicas estatísticas exploratórias que enfatizam a autocorrelação (ACF) e a função de autocorrelação parcial (PACF). O ACF mede a dependência entre os valores de uma série temporal em diferentes períodos. Seu valor reflete a conhecimento de um procedimento estocástico, ou seja, quanto os valores observados são afetados pelos valores observados ao longo do tempo. O PACF pode ser definido como a autocorrelação parcial aplicada aos resíduos da série.

Os autores Asce e Asce (2003) publicaram um estudo sobre modelagem e previsão do fluxo de tráfego usando modelos ARIMA sazonais (SARIMA) utilizando dados de duas localizações de ruas nos EUA e no Reino Unido. Os dois agregados de dados não apenas representavam países diferentes, mas também continham diferentes tipos de rodovias e diferentes técnicas de detecção. Na análise de correlação, os autores encontraram séries de dados não estacionárias sobre as condições do tráfego. Isso envolve a suposição de que as autocorrelações são muito fortes em atrasos de 1 dia e 1 semana (96 e 672 com intervalos de 15 minutos). A autocorrelação atinge o pico em alguns pares de atrasos de 1 dia de 96 intervalos, e o pico de correlação aumenta em intervalos de 1 semana de 672 intervalos. Para os autores, a

justificativa teórica para modelar séries temporais univariadas de dados de fluxo de tráfego como processos sazonais ARIMA é baseada no teorema das séries temporais conhecido como decomposição de Wald, que se aplica a séries de dados discretos que são estacionárias sobre sua média e variância, bem como na afirmação de que uma primeira diferença sazonal de uma semana de *lag* aplicada a dados discretos de condição de tráfego de intervalo produzirá uma transformação fracamente estacionária.

Portanto, também é necessário apoiar uma afirmação de que uma diferença sazonal apropriada induzirá a estacionariedade. Além disso, os resultados empíricos usando dados reais do sistema de transporte inteligente são apresentados como consistentes com a hipótese teórica.

Sáfadi (2004) analisou o comportamento da série de vazão de água na barragem de Furnas/MG entre 1963 e 1994, estudando o efeito de sazonalidade, tendência e intervenção, considerando modelos de séries temporais com e sem a presença de intervenção. No trabalho foram citados os modelos de Box e Jenkins, sendo sua estimação efetuada de forma iterativa pelo Método da Máxima Verossimilhança. Embora a série em análise não tenha apresentado sazonalidade, houve correlação significativa nos *lags* múltiplos de 12, o que levou à utilização do modelo SARIMA (1,1,1) (1,0,0)₁₂.

Em concordância, para o ajuste do modelo com intervenção, notou-se que os acréscimos encontrados nas observações 197 (maio de 1980), 233 (maio de 1983) e 330 (junho de 1991) poderiam ser consideradas como possíveis intervenções, já que esses aumentos não desapareceram na série sem tendência. Os resultados apontaram que a série de vazão se apresenta bem ajustada utilizando modelos sazonais e a incorporação do parâmetro de intervenção forneceu informações complementares na análise. Lucas et al. (2009), para simular a vazão mensal das sub-bacias da região hidrográfica do Xingu no Estado do Pará, testou o modelo hidrológico determinístico mensal de dois parâmetros e o modelo estocástico (ARIMA). O modelo hidrológico determinístico mensal de dois parâmetros, aplicado às vazões, possui estrutura simples e pode ser facilmente utilizado no planejamento e ou gerenciamento dos recursos hídricos da região hidrográfica do Xingu. Tal modelo mostrou-se eficiente na simulação das vazões mensais, porém sensível a eventos extremos de precipitação, afetando fortemente o armazenamento de água no solo. O modelo estocástico ARIMA apresentou melhores resultados na simulação das vazões mensais para todas as estações, conseguindo, inclusive, capturar a dinâmica da série temporal na simulação das vazões mensais em todas as estações da região em estudo.

Lee et al. (2012) utilizaram o modelo SARIMA para previsão de índice de poluição atmosférica. O objetivo do referido estudo foi determinar as variações mensais e sazonais de um Índice de Poluição do Ar (API) em estações de monitoramento em Johor, Malásia, utilizando a abordagem Box-Jenkins. O API foi calculado pelo Departamento de Meio Ambiente com base em cinco conjuntos principais de concentração de poluentes atmosféricos, sendo Matérias Particuladas (PM10), Dióxido de enxofre (SO₂), Dióxido de Nitrogênio (NO₂), Monóxido de Carbono (CO) e Ozônio (O₃). A Região Sul da Malásia Peninsular, Johor, foi escolhida como o local de estudo, tendo em vista que lá existem quatro estações de monitoramento da qualidade do ar, indicando que a área mais poluída estava localizada em Pasir Gudang.

Taneja (2016) propôs uma análise de séries temporais para a profundidade óptica de aerossóis em Nova Deli, capital da Índia, usando a abordagem de Box-Jenkins, com enfoque na aplicação da técnica de modelagem estocástica na análise das tendências futuras das propriedades ópticas dos aerossóis. A média mensal de dados de AOD (500 nm) de 2004 a 2014 foi obtida do MODIS a bordo do satélite Terra. O conjunto de imagens reportou a um período de 2004 a 2014 Vários modelos ARIMA sazonais foram gerados e sua validação foi verificada pela avaliação de vários parâmetros de estimativa, usando o software IBM SPSS.

O ARIMA (1,0,0) (0,1,2) 12 foi identificado como o modelo de melhor ajuste a partir das sendo utilizadas as estatísticas: coeficiente de determinação estacionário (0,530), coeficiente de determinação (0,674), média quadrática dos erros (0,128); erro absoluto médio (0,095); erro médio de porcentagem absoluta (16,942); e critérios normalizados de informação bayesiana (-3,941). Os modelos selecionados foram ainda utilizados para prever os valores de AOD para o ano de 2014 com um nível de confiança de 95%. Um modelo com o menor valor de BIC normalizado (0,046) foi encontrado para um ARIMA (1,0,0) (0,1,2) 12, que pode ser considerado como o melhor modelo de ajuste e pode ser usado para gerar previsões de AOD obteve um BIC um pouco melhor (0,046).

Camelo et al. (2017) buscou em seu estudo realizar previsões de velocidade do vento para aplicação no setor de geração eólica na região Nordeste do país, nos municípios de Fortaleza, Parnaíba e São Luís, no período de janeiro de 2010 a dezembro de 2014. Foram utilizados modelos de previsão de séries temporais - ARIMA, Holt-Winters (HW) e Redes Neurais Artificiais (RNA), com a finalidade de identificar o de melhor performance após comparação. O modelo RNA apresentou os melhores valores das medidas de acurácia, principalmente o erro percentual entre as séries observadas e simuladas para Fortaleza e São

Luís (10,50%). Na localidade de Parnaíba os valores dos erros foram menores para o modelo ARIMA, com erro percentual de 9,70%. As séries temporais previstas pelos modelos conseguiram representar características importantes, particularmente, as menores intensidades da média mensal da velocidade do vento (período chuvoso), seguido das maiores intensidades da média mensal da velocidade do vento (período seco).

3. MATERIAL E MÉTODOS

3.1 CARACTERIZAÇÃO DA ÁREA DE ESTUDO

A área do estado de Mato Grosso é 903.207,047 km² (IBGE, 2021), sendo considerado o terceiro maior Estado em área da Federação brasileira, e encontra-se na região Centro-Oeste do país, centro do continente Sul-americano, conforme mostra a Figura 03. Possui uma privilegiada localização, pois a sua área territorial tem fronteira internacional e faz parte da Amazônia Brasileira, com a sua vegetação caracterizada por abranger três biomas distintos: Pantanal, Cerrado e o Amazônico, com 25% de área ao centro norte tipificada por transição entre cerrado e floresta amazônica. Esta diversidade associada a outros fatores influencia nas características climáticas locais da região mato-grossense (MORENO, 2005).

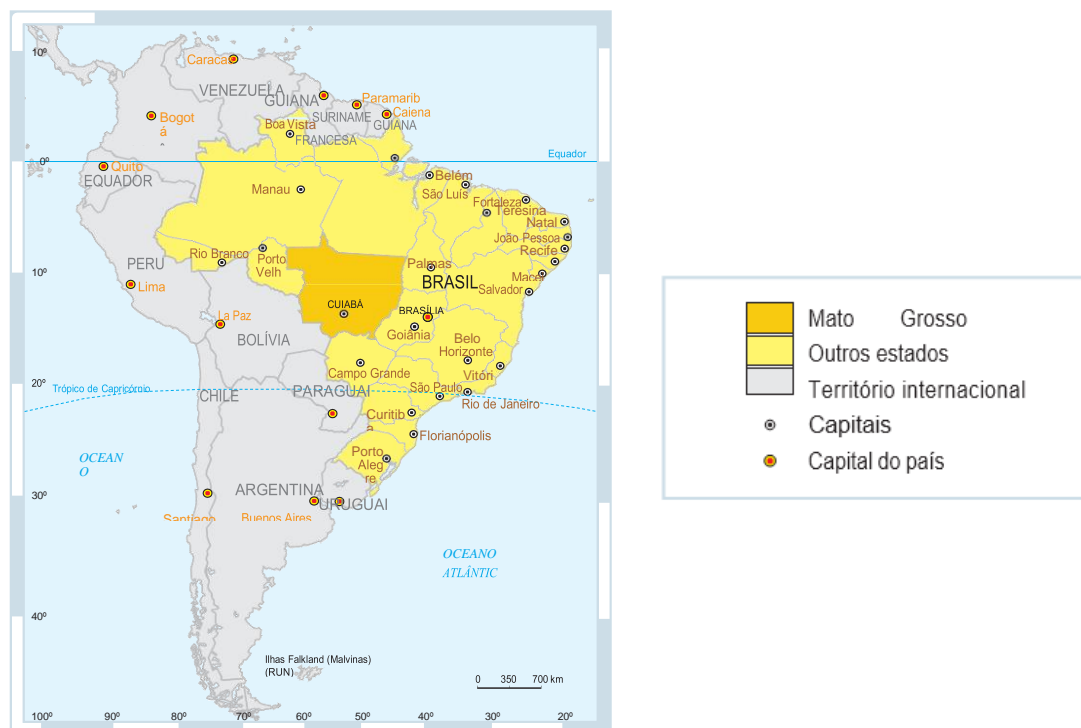


Figura 3 - Mato Grosso, localização no Brasil e na América do Sul.

Fonte: Moreno, 2005.

Pois bem, o estado de Mato Grosso está inserido na região ao centro da América do Sul, em uma zona intertropical com valores de latitude entre 15,5° a 8° Sul. Segundo DA ROCHA et al. (2023), o clima desta região é caracterizado pela sazonalidade. A estação chuvosa é de outubro a abril e o período seco, de maio a setembro. As oscilações de temperatura são amenas ao longo do ano, com temperatura média de 24,2 ° C e máxima média de 31,6 ° C, em concordância com o perfil do clima tropical.

Segundo Nimer (1989), a região Centro-Oeste possui em seu relevo alguns contrastes, devido as formações de suas chapadas compostas por rochas sedimentares (entre 700 a 900 metros de altitude), também apresenta superfícies baixas (altura menor que 200m), e as altas superfícies cristalinas (de 900 a mais de 1.200 m de altitude), inseridas a uma extensão latitudinal, esta qual lhe confere diversificação térmica ao todo o seu território. As diferenças térmicas da região acontecem devido a dois parâmetros de cunho geográfico, a sua latitude e o seu relevo. Com relação ao seu mecanismo atmosférico, este determina o comportamento semelhante da marcha estacional de precipitação pluviométrica (máximo no verão e mínimo no inverno), atuando no sentido de criar uma uniformidade regional. Portanto, na região Centro-Oeste, o mecanismo atmosférico (fator dinâmico) constitui o fator regional que assegura uma certa homogeneidade climática, enquanto o relevo, através da variação da altitude e latitudinal, levam à heterogeneidade.

A população de MT é aproximadamente 3.567.234 pessoas, conforme censo de 2021. A densidade demográfica é de 3,36 hab/km² (IBGE, 2021), e a sua evolução é mostrada na Figura 4.

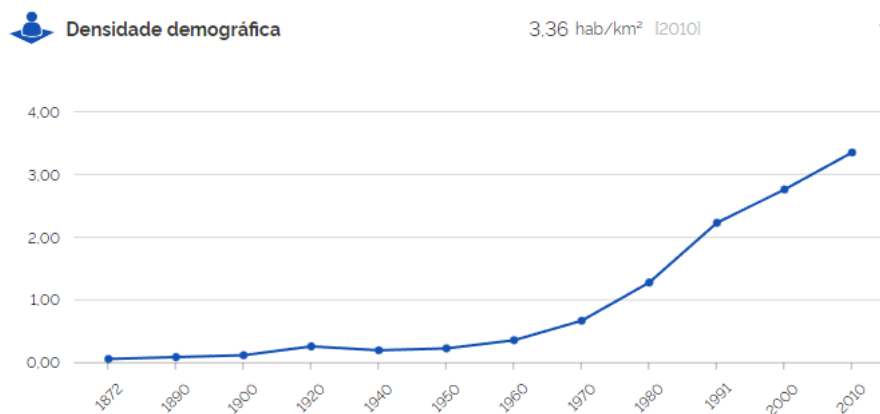


Figura 4 - Densidade demográfica do Estado de Mato Grosso.

Fonte: IBGE, 2021.

Pois bem, a área de estudo considerada para o presente trabalho refere-se às 3 cidades mais populosas de Mato Grosso: Cuiabá, Várzea Grande e Rondonópolis. Estado que possui como Capital o município de Cuiabá, este qual ocupa uma área total de 5.077,181 km² e apresenta uma população estimada de 623.614 habitantes (densidade demográfica 157,66 hab/km²), conforme mostra a Tabela 1 (IBGE, 2021). Na figura 05, é mostrado o mapa com traçado dos municípios de Mato Grosso, com destaque para as cidades que foram utilizadas na pesquisa.

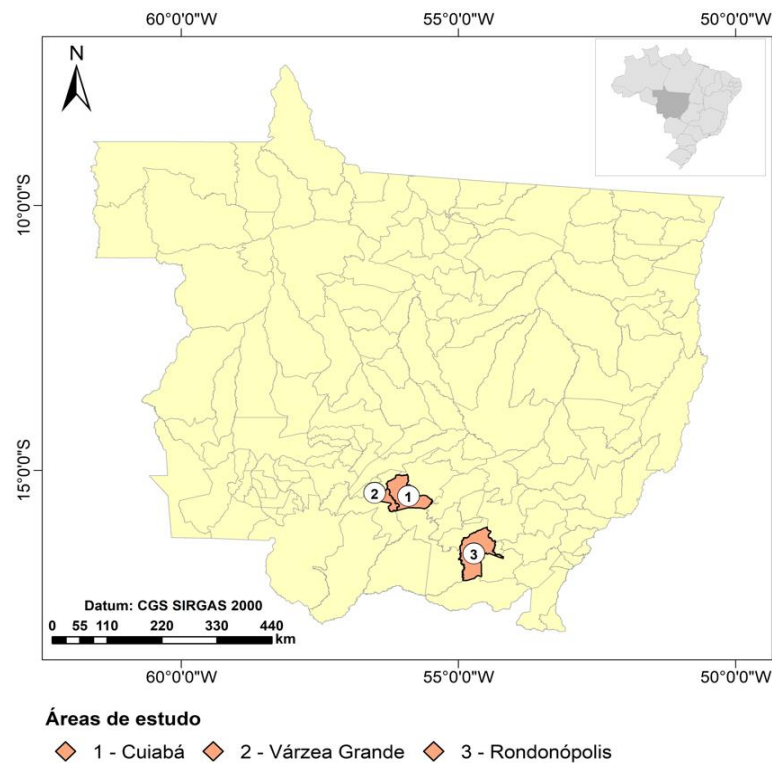


Figura 5 - Mapa de Mato Grosso com a localização dos municípios em destaque.

Fonte: Do autor, 2023

Portanto, a motivação da escolha destes municípios está relacionada com a sua maior quantidade amostral, em vista que, a propagação do vírus tem como variável o número de indivíduos existentes em uma determinada região, e que possua um potencial considerável de infectados devido a sua população, conforme mostra a estimativa calculada para ano de 2021 segundo o IBGE, mostrado na tabela 01, que traz além das informações sobre população também dados da sua área, localização, classificação do clima e temperatura média Anual.

Sendo que as informações acerca dos municípios de Cuiabá e Várzea Grande tiveram seus dados epidemiológicos integralizados, devido à proximidade entre ambos.

E de acordo com o Instituto Nacional de Meteorologia (INMET), a temperatura média anual destas cidades: em Cuiabá é de 26,4°C; em Várzea Grande é de 26,2°C; em Rondonópolis

é de 24,7°C (INMET, 2021). A análise da temperatura anual de Cuiabá, Várzea Grande e Rondonópolis permite entender melhor as diferenças climáticas entre as regiões do estado de Mato Grosso. Enquanto Cuiabá e Várzea Grande apresentam temperaturas médias anuais próximas de 26°C, Rondonópolis apresenta temperaturas médias anuais mais baixas, próximas de 24°C, respectivamente. Essas informações são importantes para a tomada de decisões em diferentes áreas, como agricultura, saúde pública e planejamento urbano.

Tabela 1 - População, área, Localização, Clima e Temperatura Média Anual das 4 cidades mais populosas de Mato Grosso.

Município	População	Área (km ²)	Localização	Clima	Temp. Anual
Cuiabá	623.614	5.077,181	-15.5989°S, -56.0949°O	Aw	26,4°C
Várzea Grande	290.383	938,057	-15.6469°S, -56.1322°O	Aw	26,2°C
Rondonópolis	239.613	4.824,020	-16.4713°S, -54.6358°O	Cwa	24,7°C

Fonte: IBGE, 2021; INMET, 2013; KÖPPEN, W, 1936.

Os três biomas distintos de Mato Grosso são mostrados na figura 6: Amazônia; Cerrado; e Pantanal. Cada um com suas particularidades e importância para o equilíbrio ecológico da região. A Amazônia, que ocupa a parte norte do estado, é o maior bioma brasileiro e abriga uma rica biodiversidade, com uma grande variedade de espécies de fauna e flora. Segundo o Ministério do Meio Ambiente (MMA, 2023), é considerada a maior floresta tropical do mundo, com áreas de floresta densa, várzea e igapó. Esse bioma é caracterizado por um clima quente e úmido, com altas temperaturas e chuvas frequentes ao longo do ano.

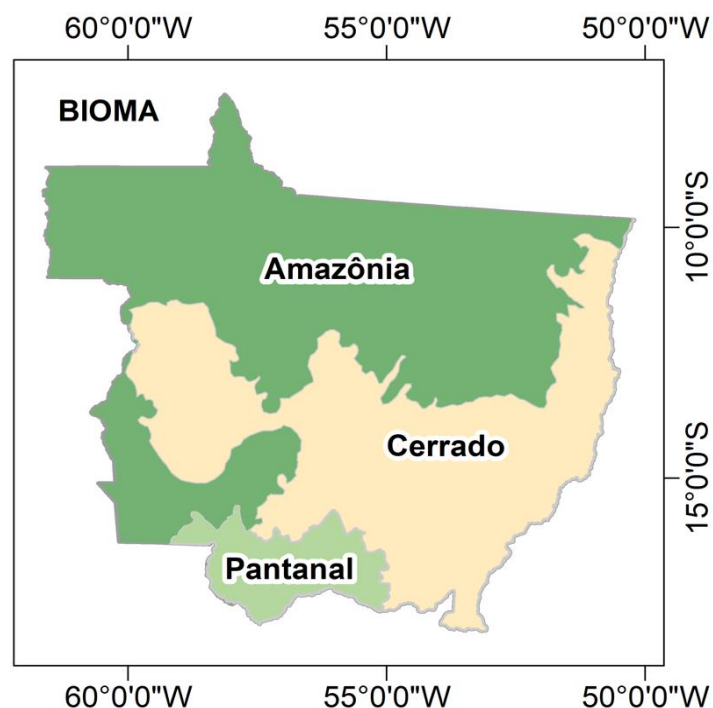


Figura 6 - Biomas do Estado de Mato Grosso.

Fonte: Do autor, 2023

O Cerrado, que ocupa a parte central e leste do estado, é o segundo maior bioma brasileiro e abrange diversas formações vegetais, como campos, cerrados, savanas e matas ciliares. Segundo o MMA (2023), é um bioma rico em espécies de fauna e flora, muitas delas endêmicas, e é considerado um importante centro de diversidade biológica. Esse bioma é caracterizado por um clima tropical sazonal, com duas estações bem definidas: uma estação chuvosa, que ocorre de outubro a março, e uma estação seca, que ocorre de abril a setembro.

O Pantanal, que ocupa a parte sudoeste do estado, é uma das maiores planícies inundáveis do mundo e abriga uma grande variedade de espécies de fauna e flora, muitas delas endêmicas. Segundo o Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA, 2023), é um ecossistema único e extremamente importante para a conservação da biodiversidade e para a manutenção dos recursos hídricos da região. Esse bioma é caracterizado por um clima tropical de altitude, com uma estação chuvosa, que ocorre de outubro a março, e uma estação seca, que ocorre de abril a setembro. Apesar de suas particularidades, esses três biomas estão interconectados e são fundamentais para a manutenção do equilíbrio ecológico do estado de Mato Grosso e de toda a região centro-oeste do Brasil.

Primeiramente foi analisado o município de Cuiabá, que se encontra na confluência de três ecossistemas brasileiros: o Cerrado no seu entorno, a floresta Amazônica ao norte, e o Pantanal ao sul. Nesta região a vegetação que predomina é o bioma Cerrado, sendo que as matas mais densas estão presentes nas proximidades dos cursos d'água (CALLEJAS, 2012). O seu bioma é caracterizado pela vegetação do cerrado, um dos principais biomas do Brasil, que ocupa cerca de 22% do território nacional (IBGE, 2004). Segundo Ab'Saber (1977), o cerrado é uma formação vegetal típica de regiões tropicais e subtropicais, caracterizada pela presença de árvores baixas e retorcidas além de gramíneas e arbustos, adaptados às condições de clima quente e seco. Em Cuiabá e Várzea Grande, a vegetação do cerrado é influenciada pelo clima tropical semiúmido, com duas estações bem definidas: uma estação chuvosa, que vai de outubro a abril, e uma estação seca, que vai de maio a setembro (FUNCEME, 2021). Segundo a Embrapa (2021), o cerrado do município é caracterizado por solos ácidos e pobres em nutrientes, o que exige adaptações das plantas para sobreviverem em tais condições.

O bioma de Rondonópolis é caracterizado como um Cerrado *stricto sensu*, e apresenta uma vegetação típica de savana com árvores baixas, arbustos e gramíneas, além de uma diversidade de fauna adaptada às condições do ambiente (Mendes et al., 2018). Ainda de acordo com Mendes et al. (2018), a região apresenta um clima tropical de altitude, com chuvas concentradas no verão e uma estação seca bem definida, o que influencia diretamente na

biodiversidade do bioma. Outra característica marcante do Cerrado de Rondonópolis é a ocorrência de solos pobres em nutrientes, mas com alta capacidade de armazenamento de água, o que favorece a adaptação das plantas às condições de seca e incêndios frequentes (Mendes et al., 2018). Além disso, o bioma de Rondonópolis é considerado uma das áreas mais ameaçadas do Cerrado devido à intensa pressão antrópica, com desmatamentos para a expansão da agropecuária e a construção de infraestruturas urbanas (Silva et al., 2019).

A classificação climática de *Köppen*, esta que é amplamente utilizada para descrever o clima de diferentes regiões do mundo, de acordo com Kottke et al. (2006), se divide em cinco tipos principais de clima, com base na temperatura e na precipitação: A (climas tropicais), B (climas secos), C (climas temperados), D (climas frios) e E (climas polares). Em relação às cidades de Cuiabá, Várzea Grande e Rondonópolis, localizadas no estado de Mato Grosso, a classificação climática de *Köppen* varia de acordo com as características climáticas de cada região. Segundo dados do Instituto Nacional de Meteorologia (INMET), a cidade de Cuiabá é classificada como Aw (clima tropical com estação seca no inverno), enquanto Rondonópolis é classificadas como Cwa (clima temperado com estação seca no inverno). Essa classificação climática indica que Cuiabá e Várzea Grande apresentam temperaturas elevadas ao longo de todo o ano, com média anual em torno de 25°C, e chuvas concentradas na estação chuvosa, que ocorre de outubro a abril. Já em Rondonópolis, as temperaturas são mais amenas, com médias anuais em torno de 23°C, e a estação seca ocorre de maio a setembro.

3.2 DADOS METEOROLÓGICOS

Os dados meteorológicos utilizados neste estudo apresentam as suas variáveis climáticas da base de dados do Instituto Nacional de Meteorologia (INMET), conforme mostra a Figura 7, estas quais são disponibilizadas através do sítio “<https://portal.inmet.gov.br/>”, que incluem, temperatura do ar, umidade relativa do ar e precipitação pluviométrica.

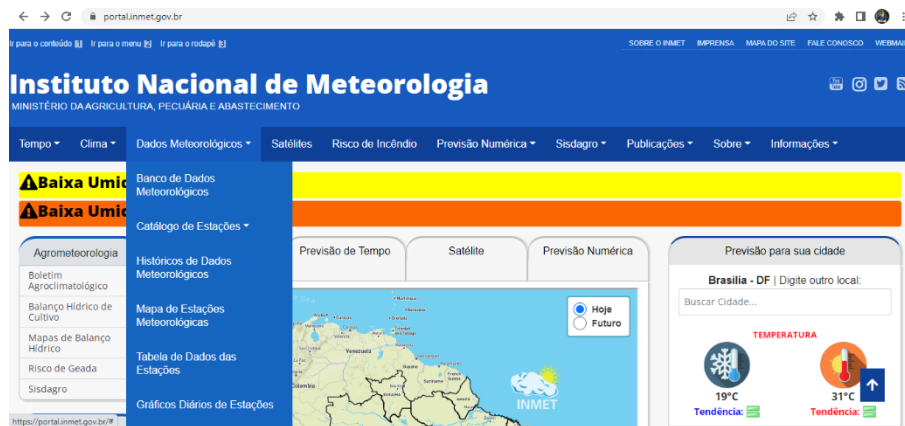


Figura 7 - Dados meteorológicos, site do Instituto Nacional de Meteorologia.

Fonte: INMET, 2022.

Para este trabalho, foram utilizados os dados de temperatura do ar, umidade relativa do ar, e precipitação pluviométrica a partir dos registros mensais da estação pertencente ao Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) do Instituto Nacional de Pesquisas Espaciais. Enquanto para valores de temperatura do ar e umidade relativa do ar, estes foram extraídos dos dados e parâmetros meteorológicos coletados através das bases do Instituto Nacional de Meteorologia (INMET), estação Cuiabá-A901, Código OMM 86705. O período considerado para todas as variáveis foi o compreendido entre 14/03/2020 e 27/09/2020, pois compreende o início dos primeiros casos de infectados pelo vírus, com pico e queda do seu gráfico antes do período de vacinação da população. Sendo que, assim como em muitos outros estados brasileiros, Mato Grosso enfrentou um aumento significativo no número de casos e óbitos por COVID-19 neste intervalo de tempo.

3.3 BASE DE DADOS DA PROGRESSÃO DA PANDEMIA

Os dados inseridos acerca das informações relativas à progressão da pandemia neste estudo foram compilados a partir dos relatórios e levantamentos de dados disponibilizados publicamente no site da Secretaria de Saúde do Estado de Mato Grosso (SES/MT), e o seu

acesso é realizado através do endereço <http://www.saude.mt.gov.br/painelcovidmt2/>. Na Figura 8 os dados dos números de infectados, recuperados e também os óbitos registrados em decorrência do vírus

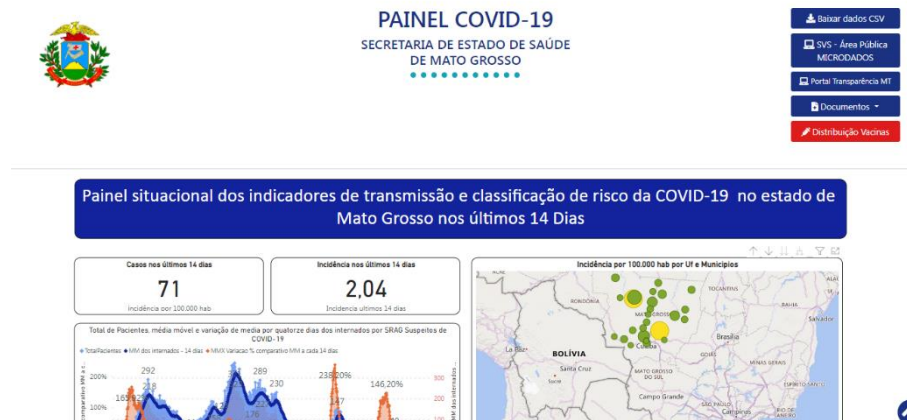


Figura 8 - Tela do Site da Secretaria de Saúde do Estado de Mato Grosso que trazem as informações relativas à progressão da pandemia.
Fonte: SES, 2022.

Os dados inseridos acerca das informações relativas à progressão da pandemia neste estudo foram compilados a partir dos relatórios e levantamentos de dados disponibilizados publicamente no site da Secretaria de Saúde do Estado de Mato Grosso (SES/MT), e o seu acesso e coleta é realizado através do endereço “<http://www.saude.mt.gov.br/painelcovidmt2/>”. Esta série histórica de dados epidemiológicos utilizados na calibração do sistema da modelagem proposta, contempla os registros dos números de infectados, recuperados e também os óbitos registrados em decorrência do vírus, dos casos ocorridos no estado de Mato Grosso no período de 14/03/2020 a 27/09/2020.

A análise dos dados das séries temporais da propagação do Covid-19 no período referenciado, em Cuiabá, Várzea Grande, através da sua variável infectados demonstrou tendência de aumento seguida de queda e novo aumento da sua curva ao longo da série, enquanto Rondonópolis manteve a tendência de aumento. Historicamente, pode-se explicar que o início do gráfico de infectados para Cuiabá e Várzea Grande foram confirmados no começo de março de 2020. A partir daí, o número de casos começou a aumentar gradualmente. Sendo que, no final de março, o governo estadual decretou medidas de distanciamento social para tentar conter a propagação do vírus, fechando escolas, comércios e serviços não essenciais. No entanto, essas medidas não foram bem-sucedidas em conter a disseminação do vírus, e o número de casos continuou a crescer. No mês de junho, o estado de Mato Grosso, incluindo Cuiabá,

começou a flexibilizar as medidas de distanciamento social, permitindo a reabertura de alguns estabelecimentos comerciais. No entanto, o número de casos continuou a aumentar e, em julho, o sistema de saúde de Cuiabá começou a ficar sobrecarregado, com a ocupação dos leitos de UTI chegando a 100%. Em agosto, o governo estadual implementou novas medidas restritivas para tentar conter a propagação do vírus, incluindo toque de recolher noturno e fechamento de bares e restaurantes. E a partir de setembro, o número de casos começou a diminuir gradualmente.

Já a cidade de Rondonópolis, teve seu primeiro caso confirmado e registrado no final de março de 2020. Sendo que, a falta de adesão às medidas de distanciamento social e a falta de equipamentos de proteção individual para profissionais de saúde foram alguns dos fatores que contribuíram para o avanço da pandemia na cidade. Porém em junho, o estado de Mato Grosso, incluindo Rondonópolis, começou a flexibilizar as medidas de distanciamento social, permitindo a reabertura de alguns estabelecimentos comerciais. No entanto, o número de casos continuou a aumentar, e em julho, a cidade registrou um aumento significativo no número de mortes relacionadas à COVID-19. Já em agosto, a cidade começou a implementar medidas mais restritivas, incluindo toque de recolher noturno e fechamento de alguns estabelecimentos comerciais. No entanto, o número de casos continuou a aumentar, e em setembro, o sistema de saúde de Rondonópolis começou a ficar sobrecarregado. A partir de outubro, a cidade começou a registrar uma queda gradual no número de casos e mortes relacionadas à COVID-19.

3.4 METODOLOGIA BOX & JENKINS

Para este estudo foi utilizada a metodologia Box & Jenkins para previsão, este qual se baseia no ajuste de modelo sequenciais denominados ARIMA. Segundo Gomes (1989), os modelos ARIMA são matematicamente muito robustos e apropriados no quesito de caracterização dos processos de geração de uma variedade de séries temporais, este procedimento metodológico abordado consiste no ajuste de modelos autorregressivos integrado e de médias móveis ao conjunto as séries temporais de dados dos parâmetros meteorológicos em conjunto com os registros da evolução da pandemia, de forma que a diferença entre os valores gerados pelos modelos e os valores observados resulte numa sequência de ruídos brancos, com base nas autocorrelações (ACF) e autocorrelações parciais (PACF) estimadas, para identificação dos pontos a serem utilizados no modelo ARIMA e para comparar com as quantidades teóricas e apontar um modelo possível para os dados, Hamilton (1994). Ao final

foi analisado os resíduos dos modelos gerados. Este passo a passo se enquadra nos passos do Box-Jenkins. Segundo Morettin e Tolloi (2006), com relação as séries utilizadas e a identificação da componente de tendência, existem três objetivos principais: “1) avaliar seu comportamento e integrá-lo a um modelo preditivo; 2) removê-lo da série temporal para promover a visibilidade dos demais componentes; ou 3) usá-lo para estimar o nível da série”. Melo (2019), descreve o processo sequencial da análise de dados conforme o fluxograma proposto na Figura 08:

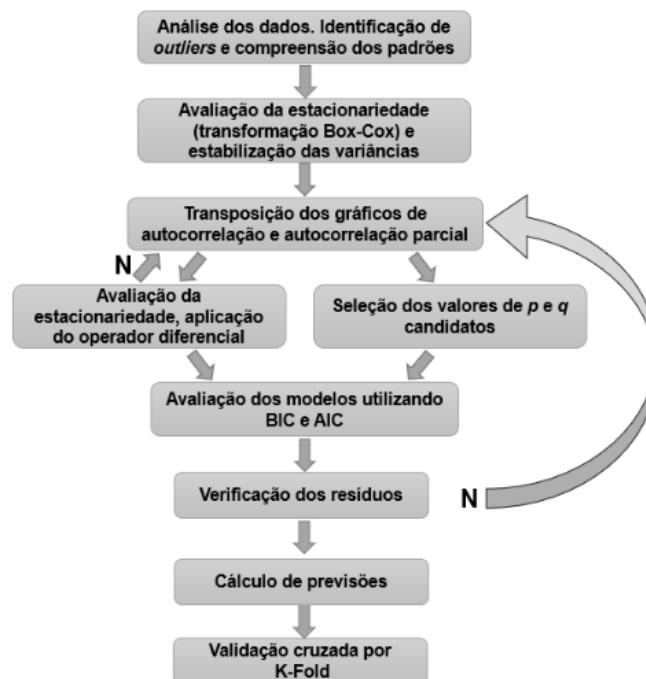


Figura 9 - Fluxograma do processo sequencial utilizado na análise das variáveis do modelo ARIMA.
Fonte: MELO, 2019.

Primeiramente, na etapa de análise dos dados, são realizadas a identificação de outliers, valores atípicos que apresentem um grande afastamento dos demais valores das séries, ou que são inconsistentes, pois a existência de outliers implica, tipicamente, em prejuízos à interpretação dos resultados dos testes estatísticos aplicados às amostras, com as falhas preenchidas através da utilização do filtro de *Kalman*, Modelo Auto Arima, este qual emprega a estimativa dos valores desconhecidos pela maximização da verossimilhança por intermédio da decomposição do erro de previsão, Tusell (2011).

Posteriormente, na Avaliação da estacionariedade (transformação Box-Cox) e estabilização das variáveis: é preciso analisar a estacionariedade da série temporal, sendo que, uma série estacionária oscila em torno de uma média ao longo do tempo, o que não ocorre na

série não estacionária, o qual apresenta elevações e diminuições abruptas. Essa análise pode ser realizada observando-se o desenvolvimento da série no transcorrer dos períodos analisados, ou também, pelo desempenho vislumbrado na função de autocorrelação (GUJARATI; PORTER, 2011). Para tornar a série estacionária e suprir essa necessidade, Box e Jenkins (1970), propõe uma metodologia capaz de estacionarizar processos não estacionários a partir da realização de sucessivas diferenças. Desta forma, verificam-se as diferenças até alcançar a estacionariedade. A avaliação da estacionariedade e aplicação do fator diferencial e seleção dos valores de p e q candidatos: apresenta a etapa de identificação de modelos de forma subdividida em duas partes. Na primeira parte, a identificação da ordem $I(d)$ do modelo a partir da verificação da estacionariedade da série, e caso necessário, a realização de sucessivas diferenciações até tornar o processo estacionário. Na segunda, a identificação das ordens p e q dos termos AR e MA do modelo ARIMA. Observa-se, portanto, que apenas se divide o processo iterativo tornando-o, de certa forma, mais detalhado e compreensível. E caso a estacionariedade não seja observada, retorna-se à etapa anterior, Maddala (2001). Ou seja, esse modelo será considerado satisfatório somente se, após a análise de diagnóstico, os resíduos do modelo satisfizerem algumas condições. Caso contrário, deve-se retroceder a etapa de identificação para que um novo modelo seja estimado e, novamente, realizar a análise.

A linguagem de programação utilizada na análise dos dados foi a linguagem R, através do pacote *forecast*, que é uma das ferramentas mais populares para gerar previsões em linguagem R, (HYNDMAN et al., 2019). Além de contar com os principais modelos embutidos, o pacote também traz diversos recursos complementares à tarefa de previsão. Por exemplo, funções para calcular erros de previsão por validação cruzada e gerar variações via *bootstrap* das séries temporais dinâmico, e voltado à manipulação, análise e visualização de dados, que utiliza uma técnica de reamostragem para se realizar inferências sobre acurácia de estimativas por amostra.

Este pacote também oferece diversas funções para ajustar modelos de previsão, incluindo os modelos SARIMA (Seasonal Autoregressive Integrated Moving Average). Quando se trata de correlação entre as variáveis, este modelo pode incluir termos autorregressivos (AR) e de média móvel (MA) para lidar com a dependência serial dos dados, enquanto o termo sazonal considera a dependência entre os dados sazonais. Segundo Hyndman e Khandakar (2008), "os modelos SARIMA são ajustados por máxima verossimilhança (*Maximum Likelihood* - ML), usando a função *arima()* no pacote *forecast* do R", que busca encontrar os parâmetros do modelo que maximizam a probabilidade dos dados observados.

Quando há correlação entre as variáveis de entrada, o modelo pode levar em conta essa correlação através da inclusão de outras variáveis exógenas no modelo. Segundo Hyndman e Khandakar, 2008, "Para incorporar variáveis exógenas em um modelo SARIMA, inclua as variáveis exógenas como um argumento adicional no modelo usando a opção `xreg`". Dessa forma, o pacote *forecast* através da função "auto.arima()" permite a inclusão de variáveis exógenas no modelo SARIMA para considerar a correlação entre as variáveis de entrada, utilizando um algoritmo que busca os melhores parâmetros de ordem e sazonalidade para o modelo. As variáveis exógenas devem ser fornecidas na forma de uma matriz ou *dataframe* com as mesmas dimensões do conjunto dos dados modelados. E para incluir múltiplas variáveis exógenas, basta adicionar as colunas correspondentes à matriz ou *dataframe* passada para o argumento `xreg`.

Avaliação dos modelos utilizando o BIC e AIC: o critério de informação de Akaike (AIC) e o critério de informação Bayesiano (BIC), desenvolvidos, respectivamente, por Akaike (1973) e Schwarz (1978), se baseiam nas estimativas da log verossimilhança e ainda penalizam os ajustes pelo número de parâmetros estimados. O AIC é o ideal para previsão, pois utiliza a validação cruzada, enquanto o BIC é mais bem utilizado para explicação, levando em consideração que permite uma estimativa coesa ao processo de composição de dados.

O AIC é calculado conforme a equação (3) abaixo:

$$\text{AIC} = -2\ln(L) + 2k \quad (3)$$

Onde L é a função de verossimilhança do modelo e k é o número de parâmetros no modelo. O AIC penaliza modelos mais complexos, adicionando um termo $2k$ à função de verossimilhança. Quanto menor o valor do AIC, melhor é o modelo.

O BIC, por outro lado, é uma versão do AIC que incorpora uma penalização mais forte para modelos mais complexos. O BIC foi proposto por Schwarz (1978) e é calculado conforme a equação (4):

$$\text{BIC} = -2\ln(L) + k\ln(n) \quad (4)$$

Onde L é a função de verossimilhança do modelo, k é o número de parâmetros no modelo e n é o tamanho da amostra. O BIC penaliza modelos mais complexos adicionando um

termo $k \ln(n)$ à função de verossimilhança. Como resultado, o BIC é mais rigoroso na escolha de modelos mais simples do que o AIC.

Os valores de referência para a análise do AIC e do BIC são relativos e dependem dos dados e modelos específicos que estão sendo usados. Em geral, o modelo com o menor valor de AIC ou BIC é considerado o modelo preferido, mas é importante considerar outros fatores, como a interpretabilidade dos modelos e a relevância teórica. Em resumo, tanto o AIC quanto o BIC são métodos estatísticos importantes para a seleção de modelos em ciências sociais e outras áreas de estudo, e seus valores podem ser interpretados de acordo com as referências disponíveis na literatura, bem como outros fatores relevantes ao contexto de estudo.

Verificação dos resíduos: nesta etapa o procedimento estimativo mais amplamente empregado é o método dos mínimos quadrados ordinários (MMQ ou MQO), que consiste em uma técnica de otimização do modelo matemático que procura a melhor ajustamento para o conjunto de dados, com intuito de tornar mínima a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Cálculo de previsões: para a metodologia de cálculo preditivo com um número elevado de registros, é plausível a utilização e ajustamento de modelos autorregressivos. Caso o número de dados for pequeno tal análise não será adequada, considerando que a hipótese básica de independência dos resíduos é quase sempre violada, concebendo estimadores inconsistentes. Todos os procedimentos estatísticos foram desenvolvidos utilizando a linguagem R no ambiente de processamento em nuvem.

4. RESULTADOS E DISCUSSÕES

Neste capítulo são mostrados os resultados, e serão apresentados os modelos ajustados à série temporal da taxa de propagação da Covid 19 nas 3 cidades mais populosas de Mato Grosso, levando-se em consideração a covariância com os parâmetros meteorológicos.

O comportamento da série epidemiológica traz informações sobre a série dos infectados, casos ativos, todos pacientes com confirmação de infecção por Covid-19, que não evoluíram a óbito e não se enquadram nos critérios de recuperados, entre as datas 14/03/2020 e 27/09/2020, totalizando 198 dias.

4.1 CARACTERÍSTICAS GERAIS DAS VARIÁVEIS

Com a finalidade de estudar a série temporal das dinâmicas estruturais dos dados através da covariância, este estudo estatístico traz possibilidade da comparação de dois grupos de dados e, dessa forma, entender como eles se relacionam entre si. Inicialmente foi realizada a análise dos dados de infectados pelo vírus e dos parâmetros meteorológicos (temperatura do ar, umidade relativa do ar, precipitação).

4.1.1 Infectados

Na figura 10, é exibida série de dados brutos referente aos casos ativos de Covid-19 do município de Cuiabá, que tiveram a confirmação da presença do vírus, e que não evoluíram a óbito e não foram contabilizados como pacientes considerados recuperados.

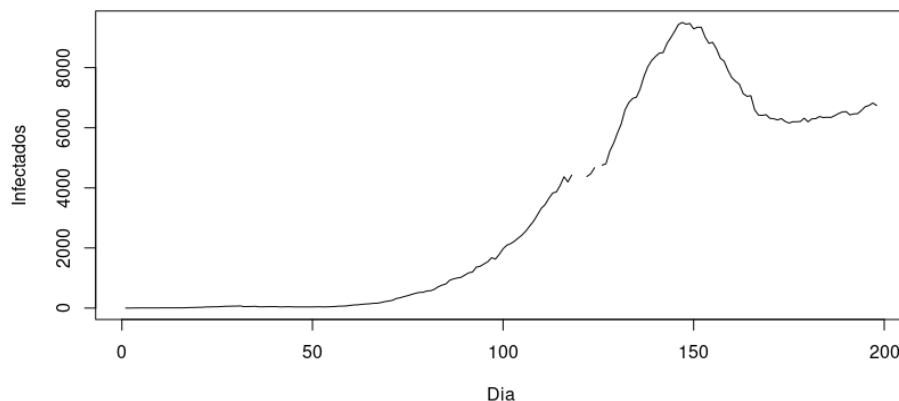


Figura 10 - Série temporal dos dados brutos da variável Infectados do município de Cuiabá de 14/03/2020 a 27/09/2020.

Fonte: Do Autor, 2023.

O gráfico contendo as correções e preenchimentos de falhas empregando o filtro de *Kalman*, através do modelo auto ARIMA, é mostrado na Figura 11.

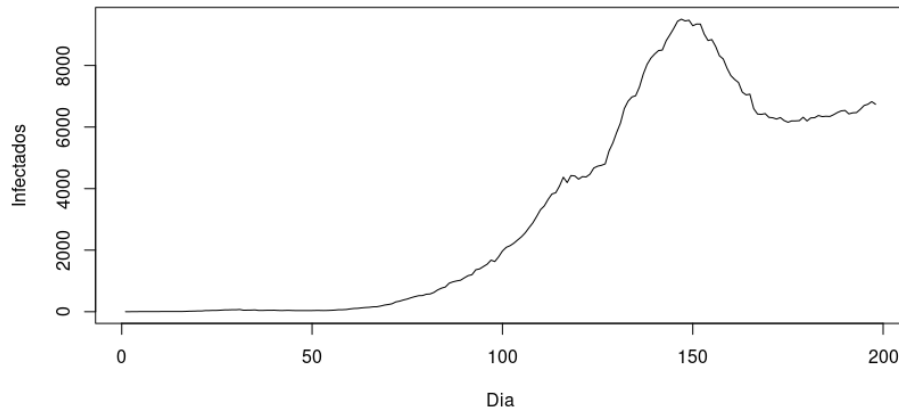


Figura 11 - Série temporal da variável Infectados do município de Cuiabá de 14/03/2020 a 27/09/2020, com preenchimento de falhas.

Fonte: Do Autor, 2023.

A figura 12 apresenta uma série de informações sobre os casos ativos de Covid-19 de Várzea Grande, que foram confirmados os casos de contaminação pelo vírus, mas que ainda não resultaram em morte ou recuperação dos pacientes, entre o período de 16/03/2020 a 27/09/2020.

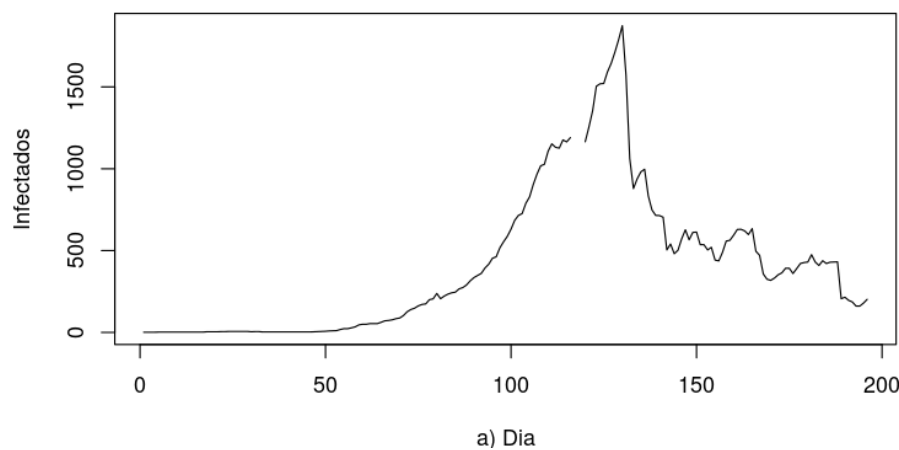


Figura 12 - Série temporal dos dados brutos da variável Infectados, município de Várzea Grande, de 16/03/2020 a 27/09/2020.

Fonte: Do Autor, 2023.

Na Figura 13, é apresentado o gráfico da série temporal da variável infectados para o município de Várzea Grande, com as suas falhas e lacunas preenchidas através da função automatizada do modelo auto ARIMA, filtro de *Kalman*.

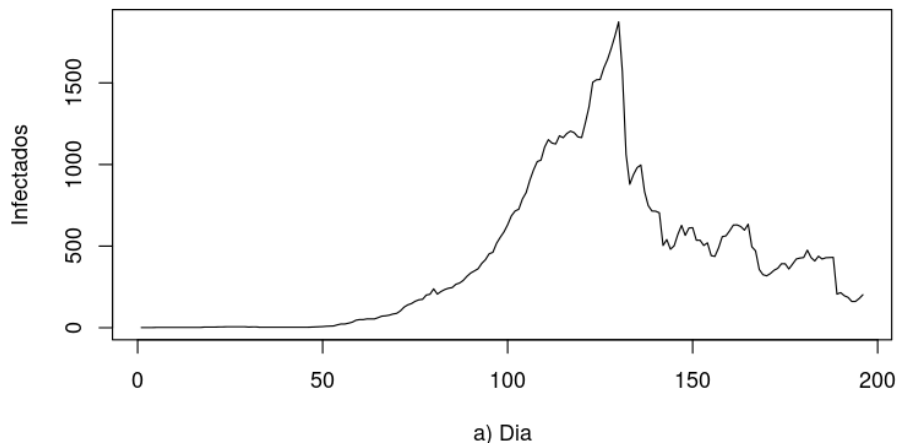


Figura 13 - Série temporal da variável Infectados do município de Várzea Grande de 16/03/2020 a 27/09/2020.
Fonte: Do Autor, 2023.

A figura 14 apresenta a série temporal dos dados brutos referentes aos casos ativos de Covid-19 de Rondonópolis-MT do período compreendido entre 30/03/2020 e 27/09/2020, a data inicial indica a ocasião que foram confirmados os primeiros casos de Covid nesta região.

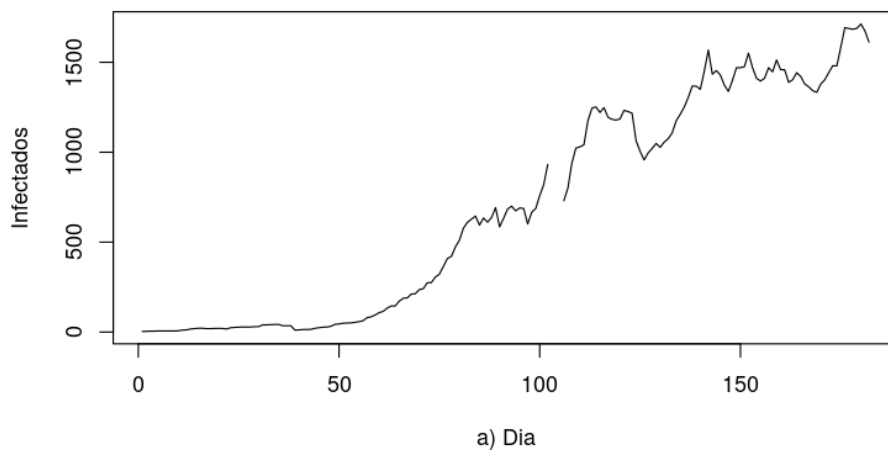


Figura 14 - Série temporal dos dados brutos da variável Infectados para o município de Rondonópolis de 30/03/2020 a 27/09/2020.
Fonte: Do Autor, 2023.

A figura 15 exibe o gráfico que representa a evolução temporal dos casos de infecção pelo coronavírus no município de Rondonópolis, no qual as informações inexistentes foram preenchidas utilizando o modelo auto ARIMA com o filtro de *Kalman* de forma automatizada.

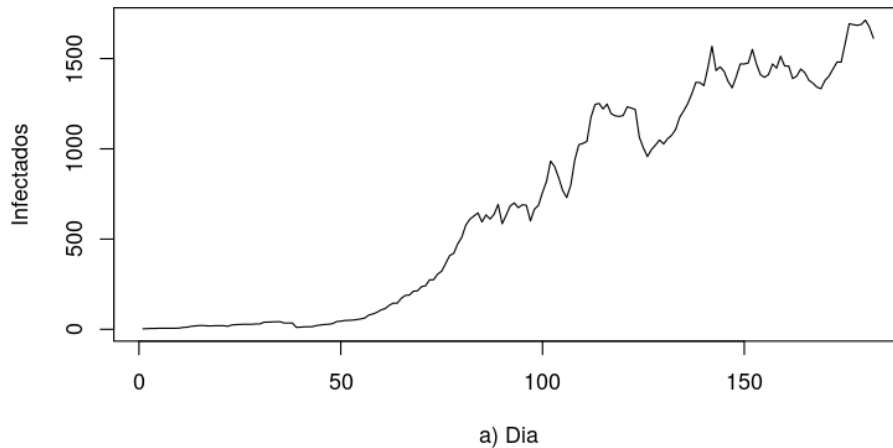


Figura 15 - Série temporal com correção de falhas, da variável Infectados do município de Rondonópolis de 30/03/2020 a 27/09/2020.

Fonte: Do Autor, 2023.

Primeiramente, a análise acerca dos gráficos de Cuiabá e Várzea Grande foram identificados dois modelos importantes a futuras observações: tendência de crescimento dos infectados, seguida de queda e flutuações sazonais não regulares ao longo da série, tal similaridade possivelmente é explicada devido à proximidade entre as duas cidades. A estatística descritiva da série com relação a sua variável quantitativa registrou os seguintes dados para os municípios analisados, conforme a tabela 2 abaixo:

Tabela 2 - Resumo estatístico da variável infectados dos municípios de Cuiabá, Várzea Grande e Rondonópolis.

Sumário	Cuiabá	Várzea Grande	Rondonópolis
Mínimo	1	1	3
1º Quartil	54	6	41
Mediana	1879	321	655
Média	3257	411	692
3º Quartil	6395	601	1291
Máximo	9504	1874	1713
Assimetria	0,4271869	1,205251	0,1565681
Curtose	-1,411373	0,8438047	-1,582865

Fonte: Do autor, 2023.

Cuiabá, Várzea Grande e Rondonópolis apresentam uma média registrada no valor de 3257, 411 e 692 infectados respectivamente, com 1 e 3 infectados no início da pandemia conforme mostra a tabela 2. Para município de Cuiabá o maior valor registrado foi de 9504, registro este computado em 07/08/2020.

A figura 17 e a figura 18, apresentam respectivamente, o histograma e o *box-plot*, do parâmetro Infectados para região de Cuiabá.

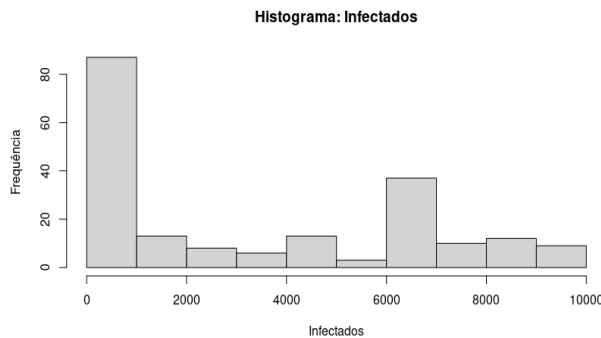


Figura 16 - Histograma da série Infectados (Cuiabá-MT).

Fonte: Do Autor, 2023.

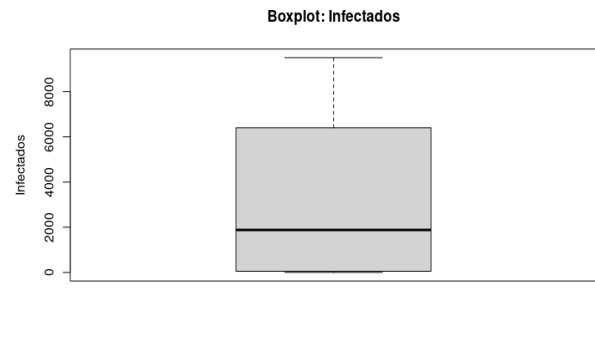


Figura 17 - Boxplot da série infectados. (Cuiabá-MT)

Fonte: Do Autor, 2023.

No gráfico de *box-plot* (figura 17), com relação a distribuição do seu conjunto de dados observados temos que:

- O valor mínimo é 1, o que significa que há pelo menos um valor no conjunto de dados igual a 1;
- O primeiro quartil (Q1) é 54, o que significa que 25% dos valores estão abaixo de 54;
- A mediana é 1879, o que significa que 50% dos valores estão abaixo de 1879.
- A média é 3257, que é a média aritmética de todos os valores no conjunto de dados.
- O terceiro quartil (Q3) é 6395, o que significa que 75% dos valores estão abaixo de 6395.
- O valor máximo é 9504, o que significa que há pelo menos um valor no conjunto de dados igual a 9504.

Na análise dos gráficos com relação aos coeficientes de assimetria (sendo que, um valor igual a 0 indica uma distribuição perfeitamente simétrica, enquanto valores acima de 0 indicam uma assimetria positiva e valores abaixo de 0 indicam uma assimetria negativa), para os dados apresentados, o gráfico de dados de Cuiabá apresenta um coeficiente de assimetria de 0,4271869, indicando uma assimetria positiva moderada. Isso significa que a distribuição dos dados está aumentando o seu valor conforme o passar do tempo, e a sua mediana estará próxima ao centro da distribuição, enquanto a média estará ligeiramente deslocada para a direita. Ou seja, com relação a progressão da doença, indica que existe um aumento dos casos de infectados.

Para o município de Várzea Grande, as figuras 18 e 19 mostram, respectivamente, a representação gráfica do parâmetro "Infectados" desta cidade por meio de histograma e box-plot.



Figura 18 - Histograma da série Infectados (Várzea Grande-MT).
Fonte: Do Autor, 2023.

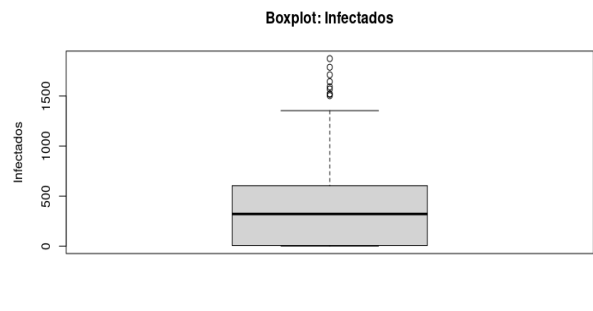


Figura 19 - Boxplot da série infectados. (Várzea Grande-MT)
Fonte: Do Autor, 2023.

Com base na configuração apresentada no gráfico de *box-plot* (figura 20), podemos interpretar o gráfico de boxplot da cidade de Várzea Grande da seguinte forma, a caixa do boxplot representa a distribuição da maioria dos dados, que estão concentrados entre 6 e 601. A mediana (321) está representada pela linha dentro da caixa, ou seja, metade dos dados estão abaixo desse valor e metade está acima, o valor mínimo (1) e o valor máximo (1874) estão representados pelas linhas que se estendem para fora da caixa, indicando que há valores extremos em ambos os lados da distribuição. Com base nessas informações, podemos concluir que os dados analisados possuem uma distribuição bastante assimétrica, com grande variabilidade entre os valores mínimo e máximo.

O gráfico de Várzea Grande apresenta um coeficiente de assimetria de 1,205251, indicando uma assimetria positiva forte. Nesse caso, a distribuição dos dados está altamente inclinada para a direita, com a cauda longa e pesada do lado direito do gráfico. Isso sugere que há uma alta concentração de valores abaixo da mediana e poucos valores muito altos puxando a média para a direita.

A figura 20 e a figura 21, apresentam respectivamente, o histograma e o *boxplot*, do parâmetro Infectados para região de Rondonópolis.

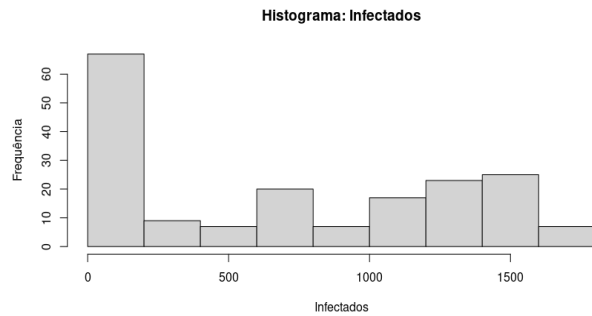


Figura 20 - Histograma da série Infectados (Rondonópolis-MT).
Fonte: Do Autor, 2023.

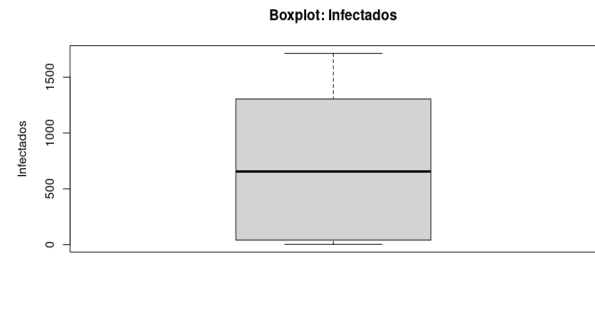


Figura 21 - *Boxplot* da série infectados. (Rondonópolis-MT)
Fonte: Do Autor, 2023.

No caso da série de propagação da Covid-19 para esta cidade, os valores estatísticos descritivos fornecem informações sobre a distribuição dos dados. O mínimo de 3 indica que há pelo menos uma observação tão baixa quanto esse valor. O primeiro quartil de 41 significa que 25% dos dados são menores que 41, enquanto o terceiro quartil de 1291 indica que 75% dos dados são menores que esse valor. A mediana de 655 é o ponto central dos dados e a média de 692 representa o valor médio da série.

Por fim, o gráfico de Rondonópolis apresenta um coeficiente de assimetria de 0,1565681, indicando uma assimetria positiva leve. Nesse caso, a distribuição dos dados está inclinada para a direita, mas de forma mais suave do que em Cuiabá. A distribuição é relativamente simétrica, mas com uma pequena cauda do lado direito do gráfico. Isso sugere que a maioria dos valores está próxima à mediana, com poucos valores muito altos puxando a média para a direita.

Em resumo, os coeficientes de assimetria indicam que as distribuições em Cuiabá, Várzea Grande e Rondonópolis têm assimetrias diferentes. Cuiabá e Rondonópolis apresentam uma assimetria positiva, mas de intensidades diferentes, enquanto Várzea Grande tem uma assimetria positiva forte.

Outra medida de dispersão analisada é a curtose, que é o grau de achatamento de uma distribuição, em relação a sua distribuição normal. Um valor de curtose igual a 0 indica que a distribuição dos dados tem a mesma forma que uma distribuição normal, enquanto valores acima de 0 indicam que a distribuição é mais "pontaguda" e estreita do que uma distribuição normal, e valores abaixo de 0 indicam que a distribuição é mais achatada e ampla do que uma distribuição normal. No caso apresentado, o gráfico de dados de Cuiabá apresenta um coeficiente de curtose de -1,411373, indicando que a distribuição dos dados é achatada e ampla em comparação com uma distribuição normal. Isso sugere que os dados têm uma dispersão maior do que o esperado em uma distribuição normal e que há poucos valores extremos. O

gráfico de Várzea Grande possui um pico dos casos de infectados e uma posterior queda destes valores, demonstrado por coeficiente de curtose igual a 0.8438047, indicando que a distribuição dos dados é mais "pontaguda" e estreita do que uma distribuição normal.

Isso sugere que os dados estão mais concentrados em torno da média e há menos valores extremos. Por fim, o gráfico de Rondonópolis apresenta um coeficiente de curtose de -1.582865, indicando que a distribuição dos dados é ainda mais achatada e ampla em comparação com uma distribuição normal do que a distribuição de Cuiabá. Isso sugere que os dados têm uma dispersão ainda maior do que o esperado em uma distribuição normal e que há ainda menos valores extremos. Em resumo, os coeficientes de curtose indicam que as distribuições em Cuiabá, Várzea Grande e Rondonópolis têm formas diferentes em relação à distribuição normal. Cuiabá e Rondonópolis apresentam uma distribuição mais achatada e ampla, enquanto Várzea Grande apresenta uma distribuição mais "pontaguda" e estreita.

A seguir, nas figuras 22, 23 e 24 a partir da transformada de *fourier* é apresentada a decomposição desta série temporal da variável infectados (Cuiabá, Várzea Grande e Rondonópolis), em valores observados, tendência, sazonalidade e parte aleatória.

Observados (*observed*): o primeiro gráfico mostra a representação da sua série de dados observados no período;

Tendência (*trend*): Segundo (FERREIRA et al., 2018) e (GONÇALVES, 2018), tendência é um elemento presente nas séries temporais quando a variável analisada demonstra uma propensão em seguir uma direção, seja ela crescente ou decrescente, e não obrigatoriamente há de ser linear;

Sazonalidade (*seasonal*): consiste em um padrão que se repete em um dado intervalo de tempo, nesta observação da parte sazonal, percebe-se um possível comportamento sazonal ao longo do gráfico;

Aleatória (*random*): chamado de ruído branco ou erro, gerada a partir de variáveis não correlacionadas, é usada como modelo para o ruído, são variações que não são explicadas. No gráfico analisado, o seu aspecto tende a aparentar comportamento aleatório.

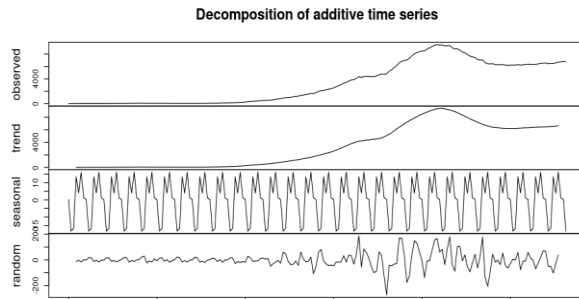


Figura 22 - Decomposição da variável infectados utilizando transformada de *Fourier* para Cuiabá.
Fonte: Do Autor, 2023.

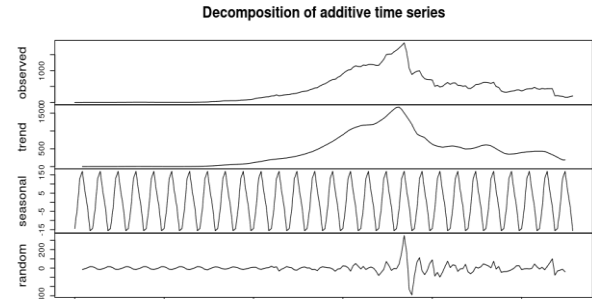


Figura 23 - Decomposição da variável infectados utilizando transformada de *Fourier* para Várzea Grande.
Fonte: Do Autor, 2023.

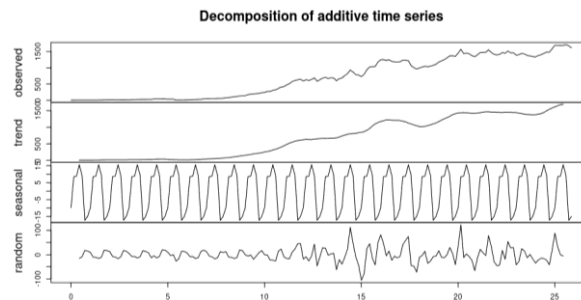


Figura 24 - Decomposição da variável infectados utilizando transformada de *Fourier* para Rondonópolis.
Fonte: Do Autor, 2023.

A figura 25, 26 e 27 apresentam os gráficos do teste de estacionariedade chamado KPSS, da série temporal da propagação do Covid-19 das cidades de Cuiabá, Várzea Grande e Rondonópolis e tem como objetivo a apresentação visual desta variável afim de saber se ela possui características de estacionariedade em uma série temporal. Portanto, através do teste de *Dickey-Fuller* (teste estatístico utilizado para verificar se uma série temporal é estacionária ou não - ADF), foi utilizada juntamente com a linguagem R o pacote "tseries", que fornece a função "adf.test()" para executar o teste ADF.

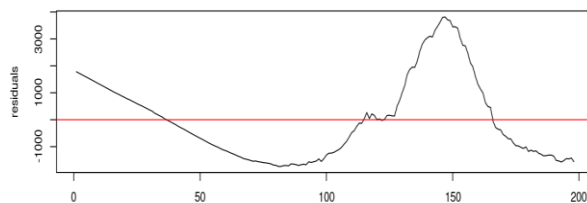


Figura 25 – Teste de estacionariedade série infectados (Cuiabá-MT).
Fonte: Do Autor, 2023.

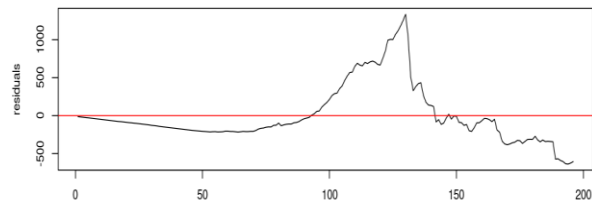


Figura 26 - Teste de estacionariedade série infectados (Várzea Grande-MT).
Fonte: Do Autor, 2023.

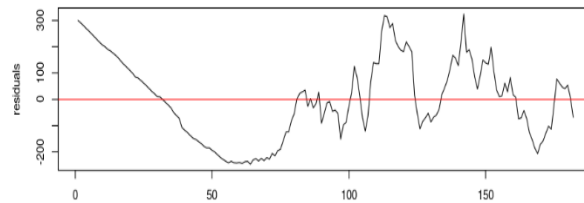


Figura 27 - Teste de estacionariedade série infectados (Rondonópolis-MT).

Fonte: Do Autor, 2023.

Pois bem, o KPSS é baseado em regressão linear, dividindo séries temporais em uma tendência determinística, valores aleatórios e erro. Através do método dos mínimos quadrados, para encontrar a equação de regressão. E se os dados são estacionários, a série é distribuída em torno de uma média constante, (STEPHANIE, 2016).

Os valores de “*p-value*” encontrados através do teste, indicam o nível de significância do teste de Dickey-Fuller aplicado às séries temporais de Cuiabá, Várzea Grande e Rondonópolis. O “*p-value*” é uma medida de quão improvável é a hipótese nula, que é a hipótese de que a série temporal não é estacionária. Um valor de *p* menor que o nível de significância escolhido (geralmente 0,05 ou 0,01) indica que há evidência suficiente para rejeitar a hipótese nula e concluir que a série temporal é estacionária. Por outro lado, um valor de *p* maior que o nível de significância escolhido sugere que não há evidência suficiente para rejeitar a hipótese nula e que a série temporal não é estacionária. No caso das cidades mencionadas, os valores de *p* são maiores do que o nível de significância de 0,05. Isso significa que não há evidência suficiente para rejeitar a hipótese nula de que as séries temporais não são estacionárias. Portanto, as séries temporais de Cuiabá, Várzea Grande e Rondonópolis podem ser consideradas não estacionárias, ou seja, podem apresentar uma tendência ou uma sazonalidade ao longo do tempo, que afeta a média, variância e autocorrelação dos dados.

4.1.2 Temperatura do Ar

A Figura 28 e 29 mostram respectivamente o comportamento do parâmetro temperatura do ar em Cuiabá, Várzea Grande e Rondonópolis, durante o período de 14/03/2020 a 27/09/2020 e 30/03/2020 a 27/09/2020. Os gráficos de temperatura do ar apresentam a temperatura média no período proposto, no caso, dias corridos. Os altos e baixos se repetem regularmente durante todos os períodos e não apresentam nenhuma tendência observável.

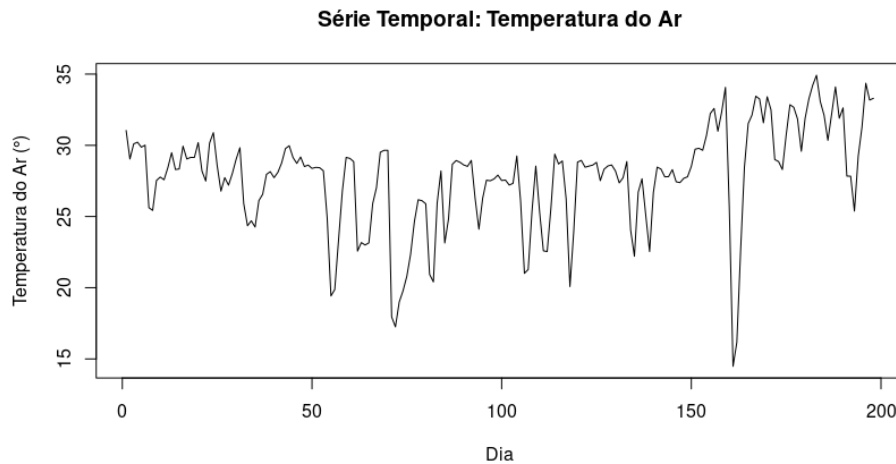


Figura 28 - Série temporal de temperatura média diária do ar do município de Cuiabá e Várzea Grande, período de 14/03/2020 a 27/09/2020.

Fonte: Do Autor, 2023.

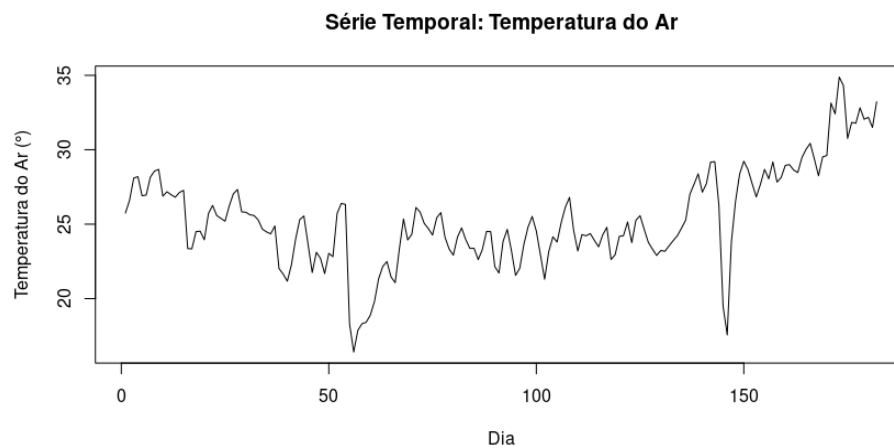


Figura 29 - Série temporal de temperatura média diária do ar do município de Rondonópolis, período de 30/03/2020 a 27/09/2020.

Fonte: Do Autor, 2023.

A estatística descritiva da série de temperatura do ar, com relação a sua variável quantitativa registrou os seguintes dados, conforme a tabela 3 abaixo:

Tabela 3 - Resumo estatístico da variável temperatura do ar em Cuiabá, Várzea Grande e Rondonópolis.

Sumário	Cuiabá/Várzea Grande	Rondonópolis
Mínimo	14,48	16,41
1º Quartil	26,14	23,36
Mediana	28,32	24,84
Média	27,72	25,37
3º Quartil	29,56	27,17
Máximo	34,92	34,88
Assimetria	-0,8597397	0,2787197
Curtose	1,250395	0,5555058

Fonte: Do autor, 2023.

A figura 30 e 31 trazem respectivamente, o histograma e o gráfico tipo *box-plot* da variável temperatura do ar para Cuiabá e Várzea Grande.

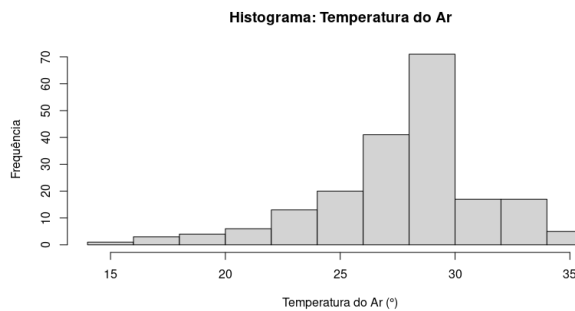


Figura 30 - Histograma da temperatura do ar do município de Cuiabá/Várzea Grande.

Fonte: Do Autor, 2023.

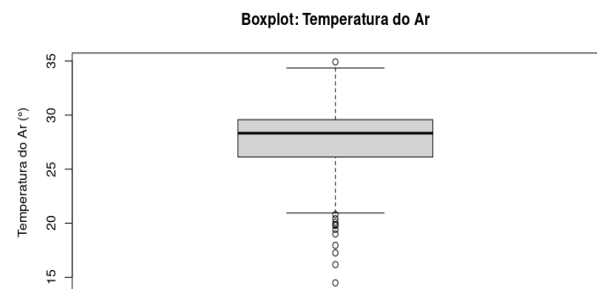


Figura 31 - *Box-plot* da temperatura do ar do município de Cuiabá/Várzea Grande.

Fonte: Do Autor, 2023.

No caso das cidades de Cuiabá e Várzea Grande, com um coeficiente de assimetria de $-0,8597397$, o gráfico de temperatura do ar pode apresentar uma distribuição assimétrica. É importante lembrar que o coeficiente de assimetria não fornece informações sobre a magnitude da assimetria, apenas sobre a direção. Com base nos valores estatísticos descritivos fornecidos, o gráfico de boxplot para a temperatura do ar em Cuiabá, a caixa começa no primeiro quartil (26,41) e termina no terceiro quartil (29,56), indicando que a maioria das observações de temperatura estão dentro desse intervalo. A mediana (28,32) é mostrada como uma linha dentro da caixa. O *Boxplot* se estende a partir da caixa até o mínimo (14,48) e o máximo (34,92), mostrando a variação completa dos dados. A partir deste gráfico, podemos inferir que a temperatura média na cidade de Cuiabá é de aproximadamente 27,72 graus Celsius, com uma variação que pode chegar a 20 graus Celsius entre o valor mínimo e o valor máximo. Além disso, a maioria das observações de temperatura (cerca de 50% a 75%) está concentrada entre 26,41 e 29,56 graus Celsius. A figura 32 e 33 trazem respectivamente, o histograma e o gráfico tipo *box-plot* da variável temperatura do ar para Rondonópolis.

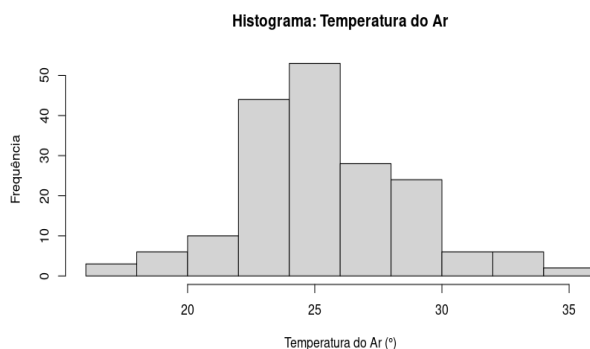


Figura 32 - Histograma da temperatura do ar do município de Rondonópolis.

Fonte: Do Autor, 2023.

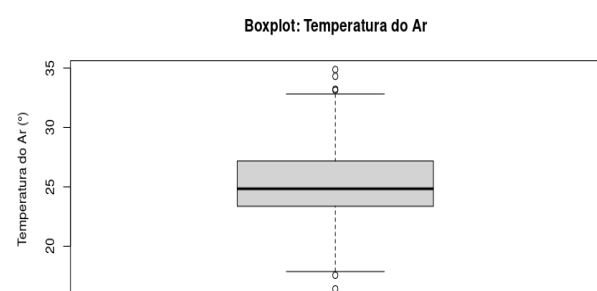


Figura 33 - *Box-plot* da temperatura do ar do município de Rondonópolis.

Fonte: Do Autor, 2023.

No caso de Rondonópolis, com um coeficiente de assimetria de 0,2787197, teoricamente o gráfico de temperatura do ar pode apresentar uma distribuição assimétrica para a direita. Isso significa que pode haver uma concentração maior de temperaturas mais altas, que estão puxando a média para cima.

O gráfico de *boxplot* para os dados de temperatura do ar da cidade de Rondonópolis mostraria que a maior parte dos valores se concentra entre 23,36 e 27,17 graus Celsius, com uma mediana de 24,84 graus Celsius. Além disso, os dados estão distribuídos em uma faixa relativamente ampla, que vai de 16,41 a 34,88 graus Celsius.

As figuras 34 e 35 apresentam a decomposição da série temporal de temperatura do ar dos municípios de Cuiabá/Várzea Grande e Rondonópolis, e exibe os valores observados (*observed*), tendência (*trend*), sazonalidade (*seasonal*) e valores aleatórios (*random*) no período de 14/03/2020 a 27/09/2020.

Observados (*observed*): o primeiro gráfico mostra a representação da sua série temporal observada no período;

Tendência (*trend*): verifica-se que a tendência apresenta o aumento dos seus valores;

Sazonalidade (*seasonal*): Há uma certa semelhança sazonal da temperatura média do ar na área de estudo e pouca variação;

Aleatória (*random*): A análise da variável apresentou comportamento aleatório, porém semelhante ao gráfico observado.

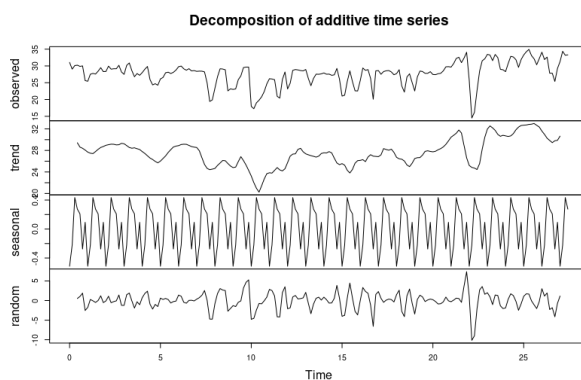


Figura 34 - Decomposição da variável temperatura do ar do município de Cuiabá-MT.

Fonte: Do Autor, 2023.

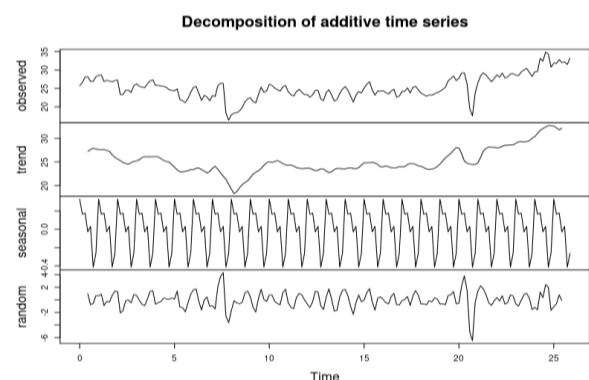


Figura 35 - Decomposição da variável temperatura do ar do município de Rondonópolis-MT.

Fonte: Do Autor, 2023.

4.1.3 Umidade Relativa do Ar

As Figuras 36 e 37 mostram respectivamente o comportamento do parâmetro umidade relativa do ar em Cuiabá, Várzea Grande e Rondonópolis e trazem as informações a respeito com os seus valores de porcentagem, durante o período de 14/03/2020 a 27/09/2020 e 30/03/2020 a 27/09/2020.

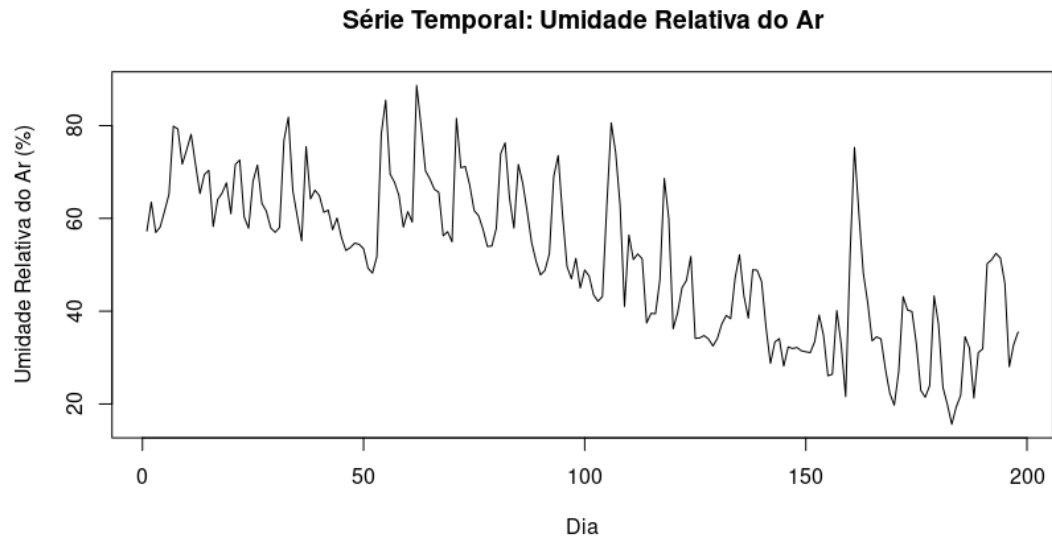


Figura 36 - Gráfico da série temporal da umidade relativa do ar, dos municípios de Cuiabá e Várzea Grande, durante o período de 14/03/2020 a 27/09/2020.

Fonte: Do Autor, 2023.

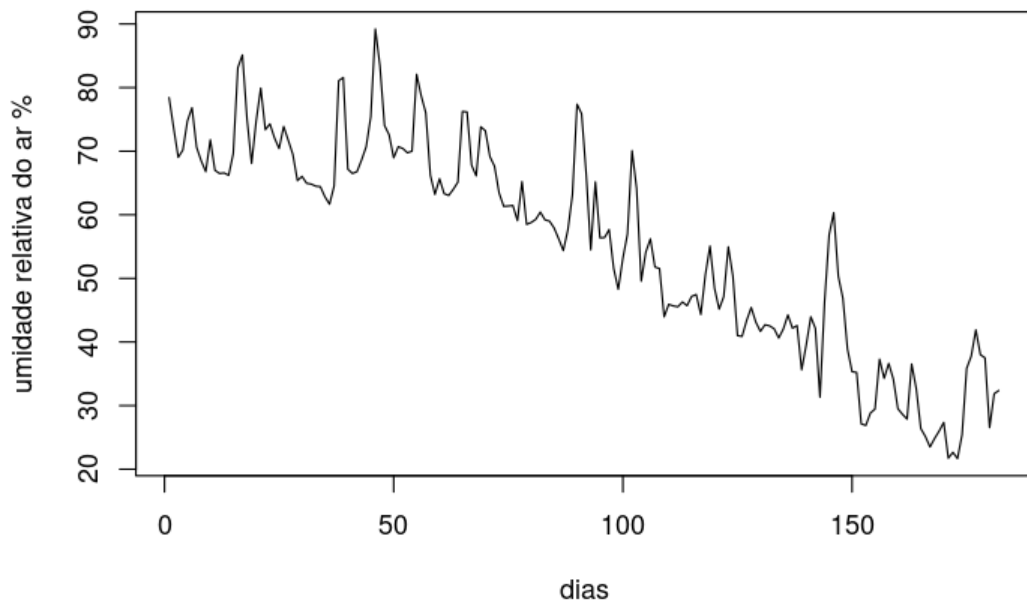


Figura 37 - Gráfico da série temporal da umidade relativa do ar, dos municípios de Rondonópolis, durante o período de 30/03/2020 a 27/09/2020.

Fonte: Do Autor, 2023.

Segundo Silva (2014), a umidade relativa do ar possui como influência a variável temperatura do ar, pois com a elevação da temperatura do ar a umidade tende a decrescer.

O resumo estatístico da série temporal umidade relativa do ar para os municípios de Cuiabá/Várzea Grande e Rondonópolis são mostrados na Tabela 4.

Tabela 4 - Resumo estatístico da variável umidade relativa do ar em Cuiabá, Várzea Grande e Rondonópolis.

Sumário	Cuiabá/Várzea Grande	Rondonópolis
Mínimo	15,63%	21,67%
1º Quartil	37,30%	42,55%
Mediana	52,33%	58,65%
Média	51,40%	55,51%
3º Quartil	63,91%	68,60%
Máximo	88,67%	89,21%
Assimetria	-0,08230566	-0,2967912
Curtose	-0,8610439	-0,9617847

Fonte: Do autor, 2023.

O histograma e *box-plot* da série temporal da variável umidade relativa do ar de Cuiabá e Várzea Grande são apresentados na figura 38 e 39, respectivamente.

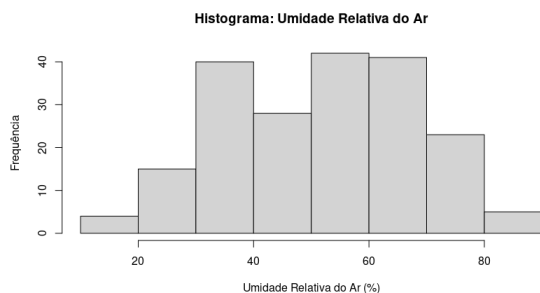


Figura 38 - Histograma da umidade de Cuiabá e Várzea Grande.

Fonte: Do Autor, 2023.

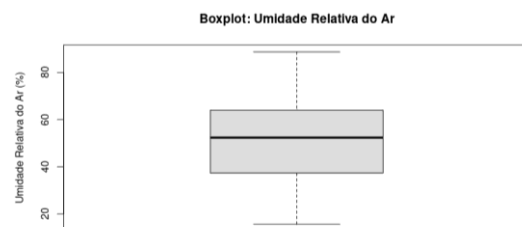


Figura 39 - *Box-plot* da umidade de Cuiabá e Várzea Grande.

Fonte: Do Autor, 2023.

Na análise do primeiro gráfico há assimetria negativa (assimetria é igual a -0,08230566), o que indica uma ligeira inclinação para a esquerda na distribuição.

Para o gráfico de *boxplot* da umidade relativa do ar da cidade de Cuiabá, temos as seguintes informações resumidas:

- O valor mínimo é de 15,63%;
- O primeiro quartil (25% dos dados) é de 37,30%;
- A mediana (50% dos dados) é de 52,33%;
- A média é de 51,40%;
- O terceiro quartil (75% dos dados) é de 63,91%;

f) O valor máximo é de 88,67%.

Com relação a análise do gráfico para as cidades de Cuiabá e Várzea Grande, pode-se perceber que a distribuição dos dados é assimétrica, com mais valores abaixo da mediana do que acima.

O histograma e o *boxplot* da série temporal da variável umidade relativa do ar de Rondonópolis são apresentados na figura 40 e 41, respectivamente.

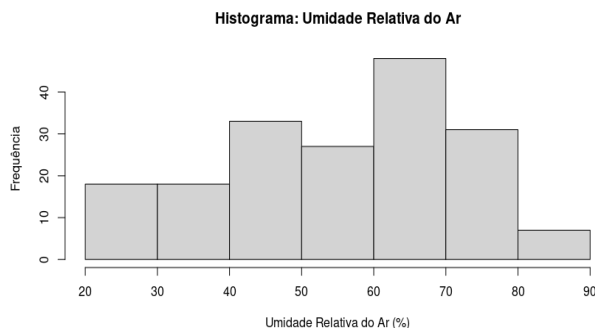


Figura 40 - Histograma da umidade relativa do ar Rondonópolis.

Fonte: Do Autor, 2023.

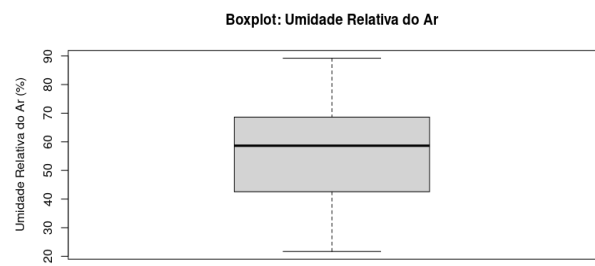


Figura 41 - *Boxplot* da umidade relativa do ar de Rondonópolis.

Fonte: Do Autor, 2023.

Ao analisar o coeficiente de assimetria igual a $-0,2967912$ para a umidade relativa do ar de Rondonópolis, isso sugere que a maioria dos valores desta umidade estão concentrados acima da média, mas há alguns valores extremos de baixa umidade relativa.

Em resumo, a distribuição da umidade relativa do ar em Rondonópolis é provavelmente concentrada em torno de valores mais altos de umidade, mas há alguns valores extremos de baixa umidade.

Para a análise do gráfico de *boxplot* da umidade relativa do ar de Rondonópolis, temos que:

- O valor mínimo de umidade relativa é de 21,67%, o que significa que pelo menos um ponto na amostra tem umidade relativa abaixo desse valor;
- O primeiro quartil é de 42,55%, o que significa que 25% dos dados estão abaixo desse valor;
- A mediana (ou segundo quartil) é de 58,65%, o que significa que metade dos dados estão acima desse valor e metade está abaixo;
- A média é de 55,51%, que é menor do que a mediana, indicando que a distribuição tem uma assimetria negativa (ou seja, há mais valores altos do que baixos);

- e) O terceiro quartil é de 68,60%, o que significa que 75% dos dados estão abaixo desse valor;
- f) O valor máximo é de 89,21%, o que significa que pelo menos um ponto na amostra tem umidade relativa acima desse valor.

Em resumo, o *boxplot* da umidade relativa do ar de Rondonópolis sugere que a maioria dos valores estão concentrados em torno da mediana e do terceiro quartil, indicando que a umidade relativa é relativamente alta na cidade. No entanto, existem alguns valores extremos abaixo do limite inferior que podem ser preocupantes, pois indicam condições climáticas extremamente secas.

As figuras 42 e 43 apresentam a decomposição da série temporal da umidade relativa do ar dos municípios de Cuiabá/Várzea Grande e Rondonópolis respectivamente, e exibe os valores observados (*observed*), tendência (*trend*), sazonalidade (*seasonal*) e valores aleatórios (*random*) no período de 14/03/2020 a 27/09/2020.

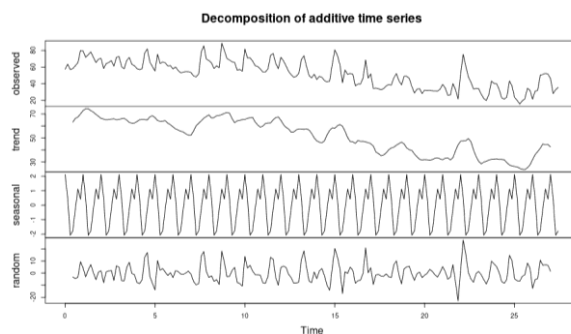


Figura 42 - Decomposição da variável umidade média do município de Cuiabá-MT.
Fonte: Do Autor, 2023.

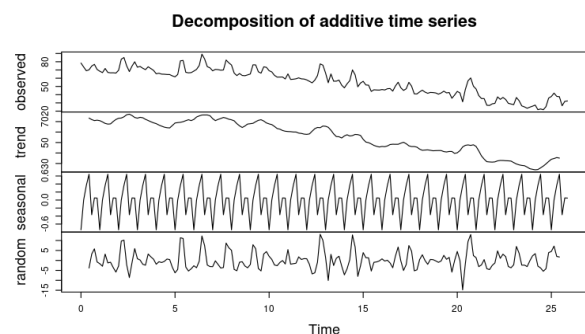


Figura 43 - Decomposição da variável umidade média do município de Rondonópolis.
Fonte: Do Autor, 2023.

4.1.4 Precipitação Pluviométrica

A Figura 44 mostra o comportamento da série de dados referentes a precipitação pluviométrica em Cuiabá e Várzea Grande, durante o período de 14/03/2020 a 27/09/2020. Não foi analisada a precipitação pluviométrica da cidade de Rondonópolis, pois o INMET não tem disponibilizado os dados deste parâmetro para o período pesquisado.

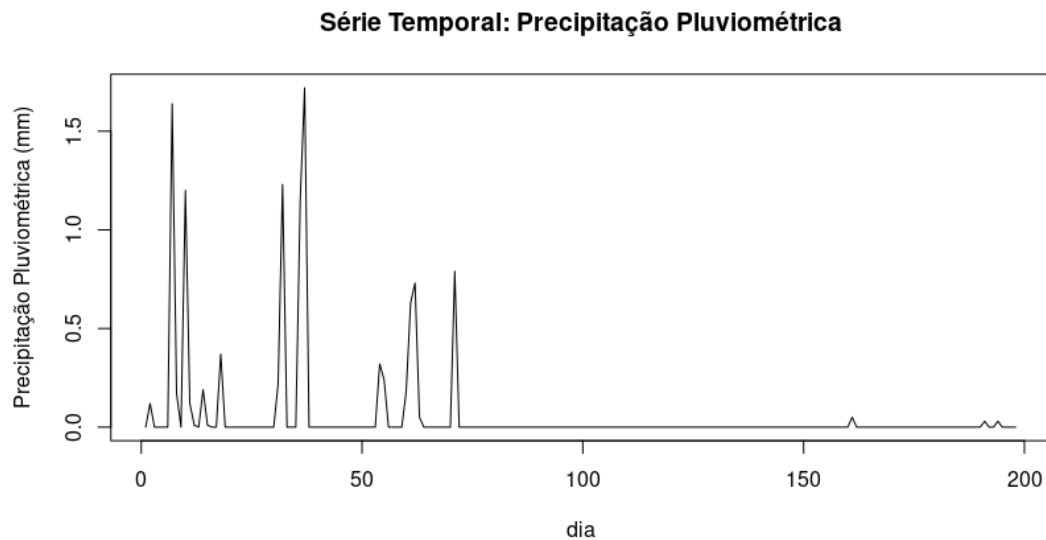


Figura 44 - Série temporal da precipitação pluviométrica dos municípios de Cuiabá e Várzea Grande, durante o período de 14/03/2020 a 27/09/2020.

Fonte: Do Autor, 2023.

O resumo estatístico da série temporal da precipitação pluviométrica para os municípios de Cuiabá e Várzea Grande são mostrados na Tabela 5.

Tabela 5 - Resumo estatístico da variável precipitação pluviométrica, Cuiabá e Várzea Grande.

Mínimo	1° Quartil	Mediana	Média	3° Quartil	Máximo	Assimetria	Curtose
0	0	0	0,05653	0	1,72	5,078983	26,70237

Fonte: Do autor, 2023.

O histograma e *boxplot* da série temporal da variável precipitação pluviométrica de Cuiabá/Várzea Grande são apresentados na figura 45 e 46, respectivamente.

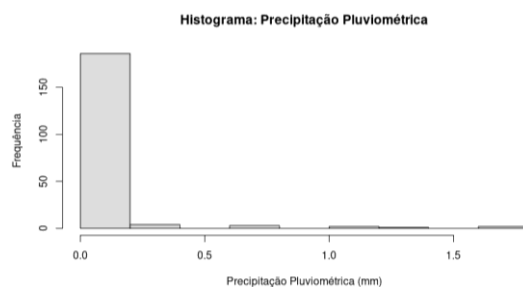


Figura 45 - Histograma da precipitação pluviométrica de Cuiabá e Várzea Grande.

Fonte: Do Autor, 2023.

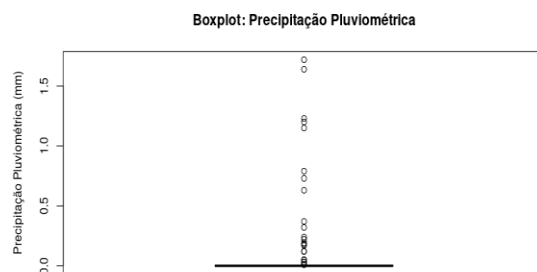


Figura 46 - Box-plot da precipitação pluviométrica de Cuiabá e Várzea Grande.

Fonte: Do Autor, 2023.

No caso desta variável, o gráfico de *boxplot* mostra que a distribuição dos dados é bastante assimétrica e com uma grande presença de valores extremos. A mediana, que divide a

distribuição em dois grupos iguais de valores, é igual a 0, o que significa que metade dos valores observados são menores ou iguais a 0. O primeiro quartil é próximo de 0, o que indica que a maioria dos valores observados é baixa. Já o terceiro quartil é de 0,05653, indicando que há alguns valores mais altos presentes nos dados. O mínimo é 0 e o máximo é 1,72, o que mostra que a maioria dos valores está concentrada em torno de 0, com alguns valores extremamente altos presentes nos dados.

O coeficiente de assimetria de 5,078983 indica que a distribuição é altamente assimétrica. Já o coeficiente de curtose de 26,70237 indica que a distribuição é extremamente leptocúrtica, ou seja, tem um pico muito alto e uma cauda muito longa, o que significa que há uma alta concentração de valores próximos à média e uma grande quantidade de valores extremos em relação à média.

Em resumo, o gráfico de *boxplot* da precipitação pluviométrica mostra que a maioria dos valores está concentrada em torno de 0, com alguns valores extremamente altos presentes nos dados, e a distribuição é altamente assimétrica e leptocúrtica.

A figura 47 apresenta a decomposição da série temporal da precipitação pluviométrica do município de Cuiabá, e exibe os valores observados (*observed*), tendência (*trend*), sazonalidade (*seasonal*) e valores aleatórios (*random*) no período de 14/03/2020 a 27/09/2020.

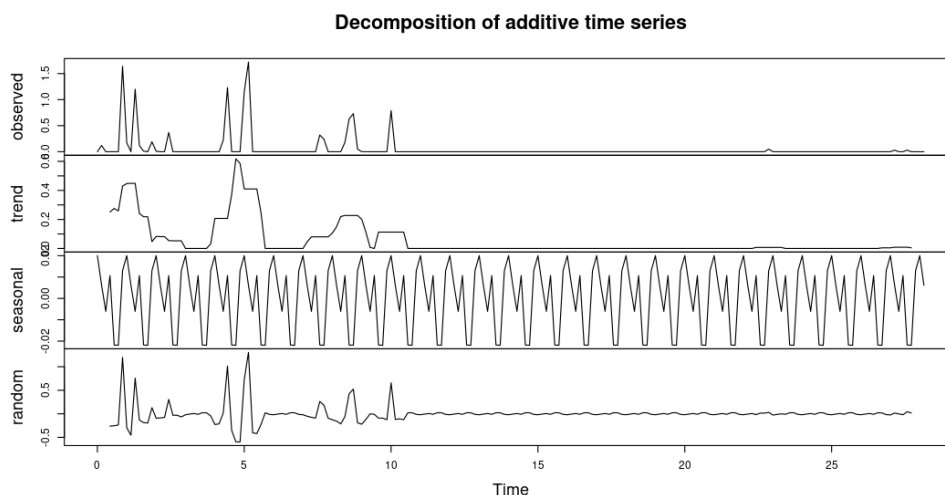


Figura 47 - Decomposição da variável precipitação pluviométrica dos municípios de Cuiabá e Várzea Grande.
Fonte: Do Autor, 2023.

4.2 MODELO AUTORREGRESSIVO

As funções de autocorrelação (ACF) e autocorrelação parcial (*Partial ACF*) contendo valores de autocorrelação em função do intervalo de tempo (*lag*) do município de Cuiabá são mostrados na figura 49, a partir da análise dos seus valores, e considerando um determinado nível de confiança, é possível afirmar se os valores são significativos ou não, apontando a intensidade da relação dos valores com os valores passados, (MINITAB, 2019).

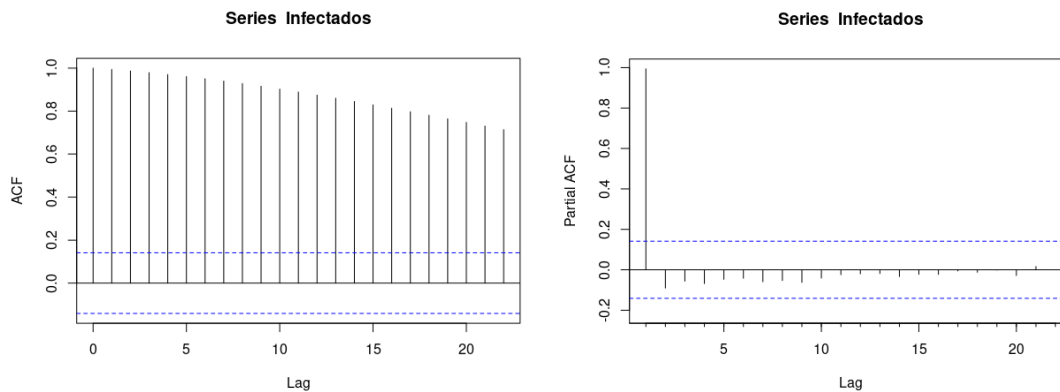


Figura 48 - Funções da autocorrelação (à esquerda) e autocorrelação parcial (à direita) da série temporal de infectados, registrado no período de 14/03/2020 a 27/09/2020 em Cuiabá-MT.

Fonte: Do autor, 2023.

Os termos autorregressivos e as médias móveis são identificadas pelo gráfico. O primeiro passo para identificação dos modelos ARIMA é a observação do seu ACF e do seu PACF utilizando como base o procedimento sugerido por Box & Jenkins (1970), que traz uma ideia inicial da configuração a ser testada, sendo está uma etapa fundamental para a aquisição de um modelo de série temporal.

No gráfico de ACF produzido, pode-se ver que a autocorrelação diminui gradualmente à medida que o *lag* aumenta, o que sugere que a série temporal pode ser estacionária e não apresenta tendência significativa. O *lag* 1 tem o maior valor de autocorrelação (1), o que significa que a série está altamente correlacionada consigo mesma no *lag* 1. Além disso, a autocorrelação permanece significativamente diferente de zero por vários *lags*, o que sugere a presença de algum padrão sazonal ou de alguma forma de autocorrelação.

A PACF (Função de Autocorrelação Parcial ou *Partial Autocorrelation Function*) é uma medida da correlação entre os valores de uma série temporal separados por unidades de tempo, após a remoção dos efeitos de correlações devidos a valores passados.

As funções de autocorrelação (ACF) e autocorrelação parcial (*Partial ACF*) contendo valores de autocorrelação em função do intervalo de tempo (*lag*) do município de Várzea Grande são mostrados na figura 49.

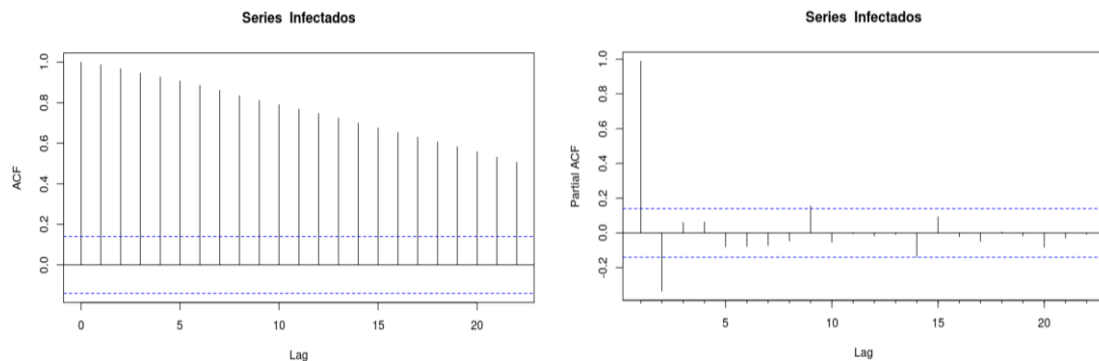


Figura 49 - Funções da autocorrelação (à esquerda) e autocorrelação parcial (à direita) da série temporal de infectados, registrado no período de 16/03/2020 a 27/09/2020 em Várzea Grande-MT.

Fonte: Do autor, 2023.

Este gráfico de ACF mostra os valores de autocorrelação para diferentes *lags* (atrasos) da série temporal. No eixo horizontal, temos os *lags* de 1 a 23. No eixo vertical, temos os valores de autocorrelação, que variam de -1 a 1. O valor de autocorrelação 1 indica uma correlação perfeita positiva, enquanto -1 indica uma correlação perfeita negativa.

As funções de autocorrelação (ACF) e autocorrelação parcial (*Partial ACF*) contendo valores de autocorrelação em função do intervalo de tempo (*lag*) do município de Rondonópolis são mostrados na figura 50.

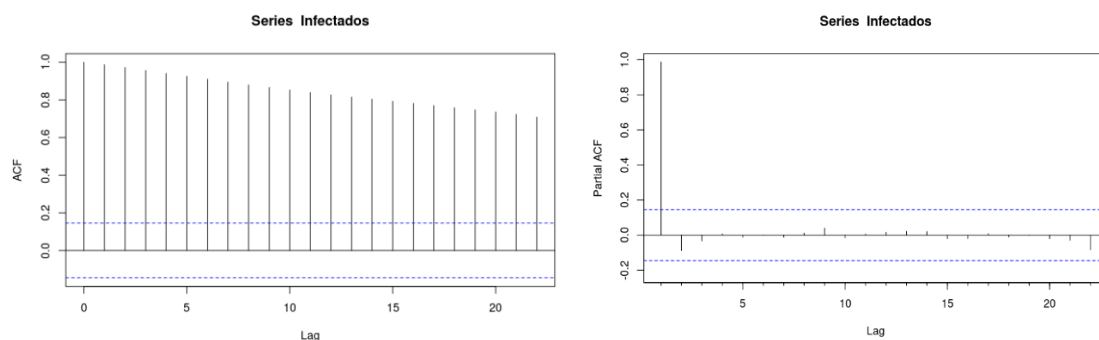


Figura 50 - Funções da autocorrelação (à esquerda) e autocorrelação parcial (à direita) da série temporal de infectados, registrado no período de 30/03/2020 a 27/09/2020 em Rondonópolis-MT.

Fonte: Do autor, 2023.

As funções de autocorrelação (ACF) e autocorrelação parcial (*Partial ACF*), que trata da ordem de diferenciação da sazonalidade, é obtido utilizando o comando R “*Diff*”, e são mostradas nos gráficos das figuras 51,52 e 53.

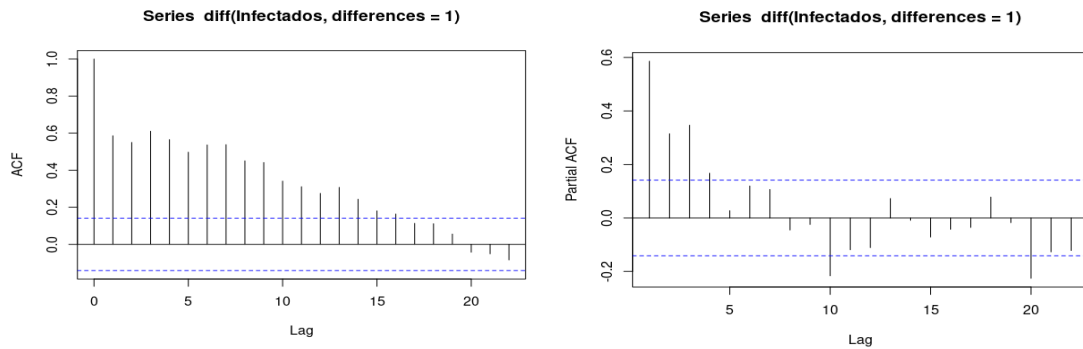


Figura 51 - Funções de autocorrelação (ACF) e autocorrelação parcial (PACF) diff 1 da série temporal de infectados, registrado no período de 14/03/2020 a 27/09/2020 em Cuiabá-MT.

Fonte: Do autor, 2023.

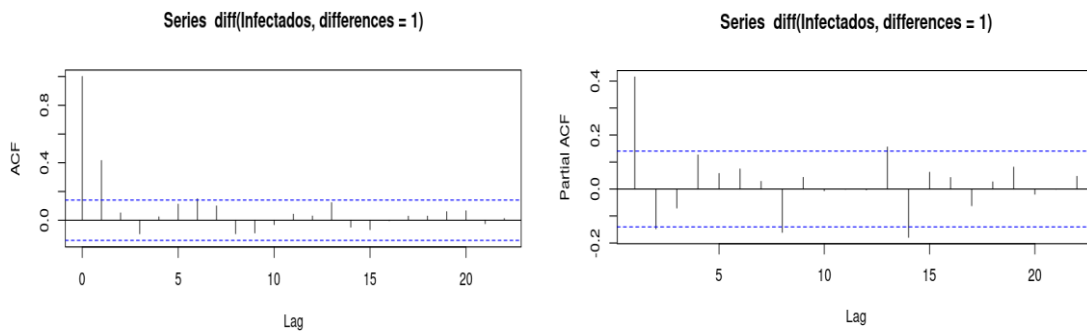


Figura 52 - Funções de autocorrelação (ACF) e autocorrelação parcial (PACF) diff 1 da série temporal de infectados, registrado no período de 16/03/2020 a 27/09/2020 em Várzea Grande-MT.

Fonte: Do autor, 2023.

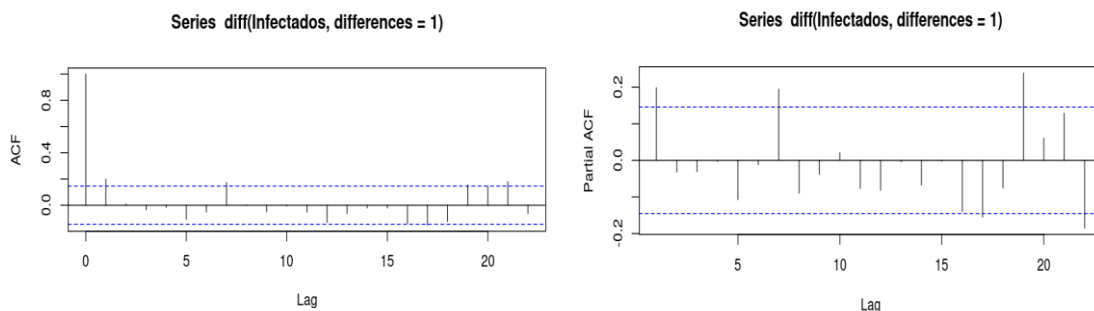


Figura 53 - Funções de autocorrelação (ACF) e autocorrelação parcial (PACF) diff 1 da série temporal de infectados, registrado no período de 30/03/2020 a 27/09/2020 em Rondonópolis-MT.

Fonte: Do autor, 2023.

Para prever uma série temporal usando o pacote "*forecast*", é importante entender a estrutura da série. A ACF e a PCF são úteis para identificar padrões na série, como sazonalidade, tendência e autocorrelação.

O algoritmo "auto.arima" começa com um modelo ARIMA simples, como ARIMA (0,1,0), que não inclui nenhum termo autorregressivo (AR) ou de média móvel (MA) e assume que a série tem uma tendência linear. Em seguida, a autocorrelação da série é avaliada por meio da ACF e PCF, e o modelo é ajustado incrementalmente para levar em consideração os padrões identificados. Ou seja, esta função utiliza a ACF e PCF para selecionar o número de termos AR

e MA no modelo. A ACF é usada para identificar a ordem do modelo MA, enquanto a PCF é usada para identificar a ordem do modelo AR. O algoritmo ajusta modelos com diferentes combinações de termos AR e MA, usando a abordagem de busca exaustiva.

Pois bem, escolher o melhor modelo ARIMA para uma determinada série temporal não é tão simples quanto se parece, partindo do princípio de que dificilmente haverá um modelo perfeito. Sendo assim, a predileção entre os modelos possíveis é, na verdade, uma proposta que traga a melhor configuração possível para previsão de uma série temporal, através de uma métrica que mesure a qualidade de um modelo estatístico, e a forma de se fazer isso é através da utilização de valores de referência, os chamados critério de informação (CI), (EMILIANO et al., 2010).

Estes critérios de informação são valores de ajuste do modelo baseado em probabilidade. Os mais comumente usados são o critério de informação de Akaike AIC (*Akaike Information Criterion*), critério de informação de Akaike modificado AICc (*Corrected Akaike Information Criterion*) e Critério de Informação Bayesiano BIC (*Bayesian Information Criterion*). Desta forma, são apresentados os resultados obtidos de AIC, AICc e BIC, após análise da programação feita por meio do script R, através do pacote *forecast*, estes que são utilizados para comparação dos modelos obtidos.

Sendo assim, a proposta da tabela 6, é analisar todas as combinações possíveis de variáveis, com o objetivo de buscar o modelo mais parcimonioso e identificar qual apresenta melhor ajuste aos dados e previsões, e que explique de forma adequada o comportamento da variável resposta.

Tabela 6 - Quadro com as covariáveis utilizadas no modelo Auto-Arima agregadas a série temporal da variável Infectados da região de Cuiabá-MT.

Variáveis	ARIMA	AIC	AICc	BIC
Infectados	(0,2,2)	281.12	281.25	290.96
Infectados + temperatura	(4,1,0)	281.24	281.83	304.22
Infectados + umidade	(4,1,0)	281.01	281.6	303.99
Infectados + precipitação	(4,1,0)	274.56	275.16	297.55
Infectados + temperatura + umidade	(4,1,0)	282.99	283.75	309.25
Infectados + temperatura + umidade + precipitação	(4,1,0)	278.53	279.49	308.08

Fonte: Do autor, 2023.

Essa tabela 6 apresenta as informações de modelos ARIMA ajustados para diferentes combinações de variáveis preditoras para prever o número de infectados por Covid-19. A primeira coluna apresenta as diferentes combinações de variáveis que foram incluídas nos

modelos, que incluem apenas a variável de resposta (Infectados) e diferentes combinações com variáveis climáticas (temperatura do ar, umidade relativa do ar e precipitação pluviométrica).

A segunda coluna apresenta o modelo ARIMA ajustado para cada combinação de variáveis, incluindo o valor da ordem de auto-regressão (AR), diferença (I) e média móvel (MA) das séries temporais.

As colunas seguintes apresentam diferentes critérios de informação para avaliar a qualidade de cada modelo ajustado: AIC (*Akaike Information Criterion*), AICc (corrigido para amostras pequenas) e BIC (*Bayesian Information Criterion*). Esses critérios avaliam a qualidade do ajuste do modelo e penalizam modelos mais complexos, ajudando a selecionar o melhor modelo dentre as diferentes combinações de variáveis. Em geral, o modelo com o menor valor de AIC, AICc ou BIC é considerado o melhor modelo.

Os modelos que incluem variáveis climáticas apresentam valores menores de AIC, AICc e BIC, indicando que são modelos mais adequados.

Considerando apenas a previsão da série temporal da variável “Infectados”, a previsão feita através do algoritmo Auto ARIMA com resultados ARIMA (0,2,2) significa que o modelo ARIMA ajustado não possui termos autorregressivos (AR), possui dois termos de média móvel (MA) e tem ordem de diferenciação de duas vezes ($d=2$). Isso significa que o modelo considera a diferença entre o valor atual e o valor de dois períodos anteriores.

O resultado da transformação de Box Cox com $\lambda = 0.358981$ indica que a série temporal foi transformada para uma distribuição mais normal, com a finalidade de tornar os dados mais adequados para modelagem e análise.

Os coeficientes estimados do modelo ARIMA são -0.8882 e 0.1433 para os termos MA1 e MA2, respectivamente. Esses coeficientes indicam o efeito dos erros de previsão passados na previsão atual. Um coeficiente MA1 negativo indica que os erros passados têm um efeito de correção negativo na previsão atual, enquanto um coeficiente MA2 positivo indica que os erros passados têm um efeito de correção positivo na previsão atual. O valor de $\sigma^2 = 0.2396$ representa a variância dos erros do modelo. Quanto menor for esse valor, melhor será o ajuste do modelo aos dados.

O log da probabilidade de ver os dados observados sob o modelo ajustado é -137.56, que é utilizado para comparar modelos alternativos. O AIC (critério de informação de Akaike) é 281.12, o AICc (AIC corrigido) é 281.25 e o BIC (critério de informação bayesiano) é 290.96.

A previsão feita com as variáveis correlacionadas entre Infectados e Temperatura do Ar, indica que o modelo ARIMA ajustado possui termos autorregressivos de ordem 4 (AR), não possui termos de média móvel (MA) e tem ordem de diferenciação de uma vez ($d=1$). O modelo também inclui um termo de regressão (drift) e uma variável exógena (xreg) para capturar a influência de uma ou mais variáveis externas na série temporal.

Os coeficientes estimados do modelo ARIMA incluem quatro termos autorregressivos (AR1-AR4), um termo de *drift* e um termo de regressão para a variável exógena. O valor do coeficiente indica o efeito dos valores passados da variável na previsão atual. O valor do coeficiente da variável exógena (xreg) indica o efeito que a variável exógena tem na previsão atual. O valor de $\sigma^2 = 0.2334$ representa a variância dos erros do modelo. Quanto menor for esse valor, melhor será o ajuste do modelo aos dados.

O log da probabilidade de ver os dados observados sob o modelo ajustado é -133.62, que é utilizado para comparar modelos alternativos. O AIC (critério de informação de Akaike) é 281.24, o AICc (AIC corrigido) é 281.83 e o BIC (critério de informação bayesiano) é 304.22.

Com relação a combinação Infectados e a Umidade Relativa do Ar, previsão feita através do algoritmo indica que o modelo ARIMA ajustado possui termos autorregressivos de ordem 4 (AR), não possui termos de média móvel (MA) e tem ordem de diferenciação de uma vez ($d=1$). O modelo também inclui um termo de regressão (drift) e uma variável exógena (xreg) para capturar a influência de uma ou mais variáveis externas na série temporal.

Os coeficientes estimados do modelo ARIMA incluem quatro termos autorregressivos (AR1-AR4), um termo de drift e um termo de regressão para a variável exógena. O valor do coeficiente indica o efeito dos valores passados da variável na previsão atual. O valor do coeficiente da variável exógena (xreg) indica o efeito que a variável exógena tem na previsão atual.

O valor de $\sigma^2 = 0.2331$ representa a variância dos erros do modelo. Quanto menor for esse valor, melhor será o ajuste do modelo aos dados.

O log da probabilidade de ver os dados observados sob o modelo ajustado é -133.5, que é utilizado para comparar modelos alternativos. O AIC (critério de informação de Akaike) é 281.01, o AICc (AIC corrigido) é 281.6 e o BIC (critério de informação bayesiano) é 303.99.

Infectados e Precipitação Pluviométrica: a especificação do modelo é dada pela ordem dos termos ARIMA, que no caso é (4,1,0), indicando que o modelo possui 4 termos autorregressivos e um termo de diferenciação sazonal. Além disso, o modelo também inclui

uma variável exógena, denotada como x_{reg} , que tem um coeficiente negativo significativo indicando que tem um impacto negativo na série temporal.

Os valores de AIC, AICc e BIC indicam o quão bem o modelo se ajusta aos dados, sendo que valores menores indicam um melhor ajuste. No caso, os valores são bastante próximos em todos os casos, mas o modelo que usa a variável exógena (data1) tem o menor valor de AIC e AICc, enquanto o modelo sem a variável exógena (data3) tem o menor valor de BIC. Por fim, a previsão de valores futuros é fornecida na tabela final, indicando o valor previsto para cada período juntamente com intervalos de confiança de 80% e 95%. Os valores dos intervalos de confiança são calculados a partir da variância do erro do modelo e podem ser usados para avaliar a incerteza da previsão.

Infectados, Temperatura do Ar e Umidade Relativa do Ar: O modelo escolhido é o modelo de regressão com erros ARIMA (4,1,0), também inclui duas variáveis exógenas: temperatura do ar e umidade relativa do ar. Os coeficientes para essas variáveis exógenas estão próximos de zero, indicando que elas não têm um grande impacto na previsão da série temporal.

O valor de σ^2 , que é a variância do ruído nos erros, é igual a 0,2343. O logaritmo da verossimilhança é -133,49, o que significa que o modelo ajustado descreve bem os dados de treinamento. Nesse caso, o valor de AIC é 282,99, o valor de AICc é 283,75 e o valor de BIC é 309,25, o que sugere um ajuste razoável do modelo.

Infectados, temperatura do ar, umidade relativa do ar e precipitação pluviométrica: o modelo selecionado pelo `auto.arima` foi um ARIMA(4,1,0), a precipitação pluviométrica, temperatura do ar e umidade relativa do ar são as variáveis exógenas adicionadas ao modelo.

No exemplo dado, o modelo tem um valor AIC de 278,53, AICc de 279,49 e BIC de 308,08. Por fim, a seção "Point Forecast" mostra as previsões de séries temporais para o período futuro, juntamente com os intervalos de confiança. Os valores previstos são dados para os próximos cinco períodos, junto com os limites inferior e superior dos intervalos de confiança de 80% e 95%.

Depois que o melhor modelo é selecionado, a função "*forecast*" é usada para prever a série. Esta função utiliza o modelo ARIMA selecionado e estima os valores futuros da série, juntamente com intervalos de confiança para as previsões.

Com base nos valores do AIC, o modelo que apresenta o menor valor é o melhor modelo. No quadro fornecido, o modelo que inclui a variável de precipitação pluviométrica juntamente com a série temporal de infectados apresenta o menor valor de AIC, sugerindo que

esse modelo tem um ajuste melhor aos dados do que os outros modelos para propagação do Covid-19 em Cuiabá-MT.

A tabela 7 apresenta as informações de modelos ARIMA ajustados para diferentes combinações de variáveis preditoras para prever o número de infectados para a região de Várzea Grande.

Tabela 7 - Quadro com as covariáveis utilizadas no modelo Auto-Arima agregadas a série temporal da variável Infectados da região de Várzea Grande-MT.

Variáveis	ARIMA	AIC	AICc	BIC
Infectados	(1,2,1)	53.48	53.6	63.28
Infectados + temperatura	(1,2,1)	52.51	52.72	65.58
Infectados + umidade	(1,1,1)	49.19	49.4	62.28
Infectados + precipitação	(1,1,1)	52.27	52.48	65.36
Infectados + temperatura + umidade	(1,1,1)	51.09	51.4	67.45
Infectados + temperatura + umidade + precipitação	(1,1,1)	52.26	52.7	71.9

Fonte: Do autor, 2023.

A primeira coluna apresenta as diferentes combinações de variáveis que foram incluídas nos modelos, que incluem apenas a variável de resposta (Infectados) e diferentes combinações com variáveis climáticas (temperatura do ar, umidade relativa do ar e precipitação pluviométrica). A segunda coluna apresenta o modelo ARIMA ajustado para cada combinação de variáveis, incluindo o valor da ordem de auto-regressão (AR), diferença (I) e média móvel (MA) das séries temporais. As colunas seguintes apresentam diferentes critérios de informação para avaliar a qualidade de cada modelo ajustado: AIC (Akaike Information Criterion), AICc (corrigido para amostras pequenas) e BIC (Bayesian Information Criterion).

Infectados: o modelo ARIMA (1,2,1) é um modelo que se ajusta a uma série temporal com uma diferença sazonal de ordem 2. Isso significa que é necessário aplicar a primeira diferença da série duas vezes antes de aplicar o modelo.

O coeficiente AR1 é 0.1178, que indica a dependência da série em relação aos seus valores passados imediatos, ou seja, o valor atual da série é uma combinação linear do valor anterior e de um valor aleatório.

O coeficiente MA1 é -0.9407. Ele indica a dependência da série em relação aos seus erros passados imediatos, ou seja, o erro atual da série é uma combinação linear do erro anterior e de um erro aleatório. O log da verossimilhança do modelo é -23.74. Quanto maior o valor de log da verossimilhança, melhor o modelo se ajusta aos dados.

Infectados e Temperatura do ar: o resultado da previsão feita através do algoritmo Auto ARIMA indica que foi realizado um modelo de regressão com erros ARIMA (1,2,1). Isso

significa que o modelo considerou um termo autorregressivo de ordem 1 (AR1) e um termo de média móvel de ordem 1 (MA1) nos erros, bem como uma transformação Box-Cox com um parâmetro lambda de 0.1597409.

Os coeficientes do modelo indicam que a variável explanatória (xreg) tem um coeficiente positivo de 0.0136, enquanto o termo AR1 tem um coeficiente positivo de 0.1347 e o termo MA1 tem um coeficiente negativo de -0.9417. O valor de σ^2 indica que a variância dos erros é de 0.07407. Além disso, o log-likelihood do modelo é -22.25, o que significa que o modelo ajustou bem os dados de treinamento. O valor de AIC é de 52.51, AICc (AIC corrigido) é de 52.72 e BIC é de 65.58.

Infectados e a umidade relativa do ar: O resultado da previsão feita através do algoritmo indica que foi realizado um modelo de regressão com erros ARIMA (1,1,1). Isso significa que o modelo considerou um termo autorregressivo de ordem 1 (AR1) e um termo de média móvel de ordem 1 (MA1) nos erros.

Os coeficientes do modelo indicam que a variável explanatória (xreg) tem um coeficiente negativo de -0.0037, enquanto o termo AR1 tem um coeficiente positivo de 0.9022 e o termo MA1 tem um coeficiente negativo de -0.7644.

O valor de σ^2 indica que a variância dos erros é de 0.07337. Além disso, o log-likelihood do modelo é -20.59, o que significa que o modelo ajustou bem os dados de treinamento. O valor de AIC é de 49.19, AICc (AIC corrigido) é de 49.4 e BIC é de 62.28.

Infectados e a variável Precipitação Pluviométrica: O resultado da previsão feita pelo modelo encontrou uma combinação ARIMA (1,1,1). Os coeficientes do modelo indicam que a variável explanatória "xreg" tem um coeficiente positivo de 0.0294. Além disso, o termo AR1 tem um coeficiente positivo de 0.9104 e o termo MA1 tem um coeficiente negativo de -0.7808.

O valor de σ^2 indica que a variância dos erros é de 0.07453. Além disso, o log-likelihood do modelo é -22.13, o que significa que o modelo se ajustou bem os dados de treinamento. O valor de AIC é de 52.27, AICc (AIC corrigido) é de 52.48 e BIC é de 65.36.

Infectados, temperatura do ar e a umidade relativa do ar: O resultado da previsão feita através do algoritmo Auto ARIMA indica que foi realizado um modelo de regressão com erros ARIMA (1,1,1). Isso significa que o modelo considerou um termo autorregressivo de ordem 1 (AR1) e um termo de média móvel de ordem 1 (MA1) nos erros.

Os coeficientes do modelo indicam que a variável explanatória "temperatura do ar" tem um coeficiente positivo de 0.0039, enquanto a variável "umidade" tem um coeficiente negativo de -0.0029. Além disso, o termo AR1 tem um coeficiente positivo de 0.9014 e o termo

MA1 tem um coeficiente negativo de -0.7625. O valor de σ^2 indica que a variância dos erros é de 0.07371. Além disso, o log-likelihood do modelo é -20.54, o que significa que o modelo ajustou bem os dados de treinamento. O valor de AIC é de 51.09, AICc (AIC corrigido) é de 51.4 e BIC é de 67.45.

Infectados, Temperatura do ar e Precipitação Pluviométrica: Mais especificamente, este modelo tem um termo ARIMA (1,1,1) e três variáveis exógenas: temperatura do ar, umidade relativa do ar e precipitação pluviométrica. A variável temperatura, o coeficiente listado é 0.0032, o que significa que um aumento de uma unidade na temperatura está associado a um aumento de 0.0032 unidades na resposta do modelo, todas as outras coisas iguais. O valor de $\sigma^2 = 0.07379$ é uma estimativa da variância do erro do modelo.

Para o município de Várzea Grande o modelo que inclui apenas a previsão dos dados da variável umidade relativa do ar apresentou o menor valor de AIC, AICc e BIC entre todos os modelos ajustados, indicando que é o modelo mais adequado para descrever a relação entre as variáveis.

A tabela 8 apresenta as informações de modelos ARIMA ajustados para diferentes combinações de variáveis preditoras para prever o número de infectados para a região de Várzea Grande.

Tabela 8 - Quadro com as covariáveis utilizadas no modelo Auto-Arima agregadas a série temporal da variável Infectados da região de Rondonópolis-MT

Variáveis	ARIMA	AIC	AICc	BIC
Infectados	(1,1,0)	63.57	63.71	73.17
Infectados + temperatura	(1,1,0)	65.54	65.77	78.33
Infectados + umidade	(1,1,0)	65.57	65.8	78.37
Infectados + temperatura + umidade	(1,1,0)	67.52	67.87	83.52

Fonte: Do autor, 2023.

Infectados: Esse resultado se refere a um modelo ARIMA com um termo autorregressivo (AR) de ordem 1, um termo de diferença de ordem 1 (I) e um termo de média móvel (MA) de ordem 0.

Os coeficientes do modelo indicam que o termo AR tem um valor de 0.1302, o que significa que o modelo usa a observação anterior multiplicada por esse coeficiente para prever a próxima observação. O termo drift é igual a 0.0911, que é um deslocamento constante que é adicionado a cada valor da série temporal, ajudando a explicar tendências lineares na série.

O valor de σ^2 é igual a 0.08137 e o logaritmo da verossimilhança é igual a -28.79. O valor de AIC é 63.57, o valor de AICc é 63.71 e o valor de BIC é 73.17.

Infectados e Temperatura: Esse resultado é referente a um modelo ARIMA com um termo autorregressivo de ordem 1 ARIMA (1,1,0). O modelo também inclui uma variável exógena (xreg) que é usada para explicar as flutuações na série temporal.

Os coeficientes do modelo indicam que a série temporal é influenciada tanto pelo seu próprio valor passado quanto pela variável exógena. O coeficiente AR1 é de 0,1318, o que significa que o valor atual da série temporal é uma combinação de 0,1318 vezes o valor imediatamente anterior mais uma constante de deriva. O coeficiente da variável exógena temperatura do ar é de -0,0026, o que significa que um aumento de um ponto na variável exógena está associado a uma diminuição de 0,0026 na série temporal.

O valor de σ^2 é a variância dos erros do modelo, que é de 0,08181. O log likelihood é uma medida da adequação do modelo aos dados observados e é de -28,77 neste caso. O modelo com AIC mais baixo é preferido, que é de 65,54. O AICc é uma versão corrigida do AIC para modelos com pequenas amostras. O BIC penaliza mais fortemente modelos com muitos parâmetros, e neste caso, o modelo com menor BIC é de 78,33.

Infectado e a umidade relativa do ar: Esta saída é resultado da modelagem de séries temporais utilizando o algoritmo ARIMA com uma regressão linear múltipla. O modelo inclui um termo AR(1) (autoregressivo de ordem 1), que representa a dependência serial da série temporal, e um termo xreg (regressão), que é um regressor externo que pode ajudar a explicar a variação na série temporal. Os coeficientes do modelo indicam que o valor AR(1) é 0,1302, o valor da deriva é 0,0911 e o valor do coeficiente de regressão é 0,0000. Os valores padrão dos erros padrão (s.e.) também são fornecidos. O modelo apresenta uma variância do erro σ^2 de 0,08182 e um log-verossimilhança negativo de -28,79. Neste caso, o AIC é de 65,57, AICc é de 65,8 e BIC é de 78,37.

Infectados, temperatura do ar e umidade relativa do ar: O modelo inclui um termo AR(1) (autoregressivo de ordem 1), que representa a dependência serial da série temporal, um termo drift (deriva), que representa a tendência da série temporal, e dois termos de regressão, temperatura e umidade, que são regressores externos que podem ajudar a explicar a variação na série temporal. Os coeficientes do modelo indicam que o valor AR(1) é 0,1319, o valor da deriva é 0,0911, o valor do coeficiente de temperatura do ar é -0,0038 e o valor do coeficiente de umidade relativa do ar é -0,0006. Os valores padrão dos erros padrão (s.e.) também são fornecidos. O modelo apresenta uma variância do erro σ^2 de 0,08226 e um log-verossimilhança negativo de -28,76. Neste caso, o AIC é de 67,52, AICc é de 67,87 e BIC é de 83,52.

Com base nos critérios de avaliação do modelo apresentados na tabela, o modelo ARIMA com apenas a variável "Infectados" é o melhor modelo a ser utilizado para descrever a dinâmica da série temporal na região de Rondonópolis-MT. Isso porque apresenta o menor valor de AIC, AICc e BIC, o que indica que é o modelo com o melhor ajuste aos dados observados. A adição das covariáveis "temperatura do ar" e "umidade relativa do ar" ao modelo aumenta o valor desses critérios, o que sugere que essas variáveis não são relevantes para explicar a variação na série temporal dos infectados. Portanto, o modelo mais simples com apenas a variável "Infectados" é o mais adequado para a análise da série temporal de infectados na região de Rondonópolis-MT.

4.3 PREVISÃO DA SÉRIE TEMPORAL DOS INFECTADOS

Após a modelagem da previsão da série temporal do registro dos infectados, realizada por meio do modelo ARIMA, em linguagem R através do pacote *forecast*. este veio a apresentar resultados satisfatórios até o momento, para as séries de umidade relativa do ar. Porém, também foi obtido um bom resultado com a variável temperatura média em conjunto com a série de dados dos infectados, ou seja, ambos casos são opções plausíveis para previsão de curto prazo destas séries temporais.

Conforme mostrado na tabela 9, após as análises dos valores de AIC, deve-se utilizar o melhor resultado para construir uma previsão aos registros de infectados do município de Cuiabá. Ao verificar os resultados obtidos, identificou-se o melhor modelo ARIMA (4,1,0), variáveis “infectados + precipitação pluviométrica”.

Tabela 9 - Tabela com a previsão da variável infectados de Cuiabá-MT, para 5 dias.

Variáveis	1	2	3	4	5
Infectados	6789	6816	6844	6872	6900
Infectados + temperatura do ar	6804	6845	6883	6907	6963
Infectados + umidade relativa do ar	6806	6847	6885	6911	6966
Infectados + precipitação pluviométrica	6805	6847	6882	6908	6964
Infectados + temperatura + umidade	6807	6849	6888	6914	6969
Infectados + temperatura + umidade + precipitação	6804	6845	6880	6907	6961

Fonte: Do autor, 2023.

A figura 54 mostra o gráfico da série temporal infectados, do município de Cuiabá-MT, combinado com a variável precipitação pluviométrica, para uma previsão de 5 dias.

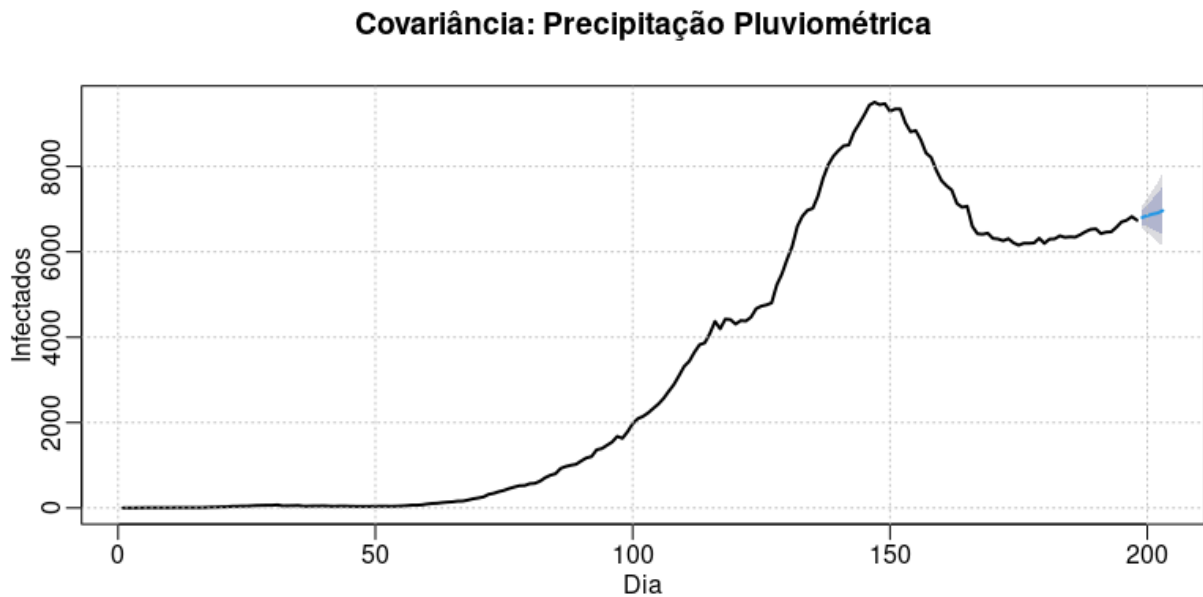


Figura 54 - Gráfico de Previsão da série temporal Infectados (linha preta), do município de Cuiabá-MT, em covariância com a variável precipitação, com previsão (linha azul) de 5 dias, e intervalos de confiança de 80% (área azul claro) e 95% (área cinza claro).

Fonte: Do Autor, 2023.

Com relação aos dados de previsão da Figura 54, como exemplo, para o dia 199 o valor esperado é de 6805 Infectados. Com um intervalo de confiança de 80%, espera-se que o valor da variável esteja entre 6632 e 6981. Com um intervalo de confiança de 95%, espera-se que o valor real esteja entre 6542 e 7075. Os valores de limite inferior e superior dos intervalos de confiança aumentam à medida que se projeta para períodos posteriores.

Na tabela 10, após as análises dos valores de AIC, são apresentados os melhores resultados para construção completa das previsões dos registros de infectados do município de Várzea Grande com todas as suas relações.

Tabela 10 – Tabela com a previsão da variável infectados para Várzea Grande-MT.

Variáveis	1	2	3	4	5
Infectados	199	194	188	182	177
Infectados + temperatura do ar	199	192	186	180	175
Infectados + umidade relativa do ar	202	202	202	202	201
Infectados + precipitação pluviométrica	201	200	199	198	198
Infectados + temperatura + umidade	202	201	201	201	200
Infectados + temperatura + umidade + precipitação	202	202	202	202	202

Fonte: Do autor, 2023.

Para a região Várzea-Grandense o modelo que inclui apenas a previsão dos dados da variável umidade relativa do ar apresentou o menor valor de AIC, AICc e BIC entre todos os modelos ajustados, indicando que é o modelo mais adequado para descrever a relação entre as variáveis.

A figura 55 retrata um gráfico que apresenta a evolução temporal dos casos de infecção em Várzea Grande-MT, acompanhado pela informação da variável umidade relativa do ar que ocorreu durante o mesmo período. Tal figura tem como finalidade ser utilizada para fazer uma previsão sobre o número de infectados para os próximos cinco dias.

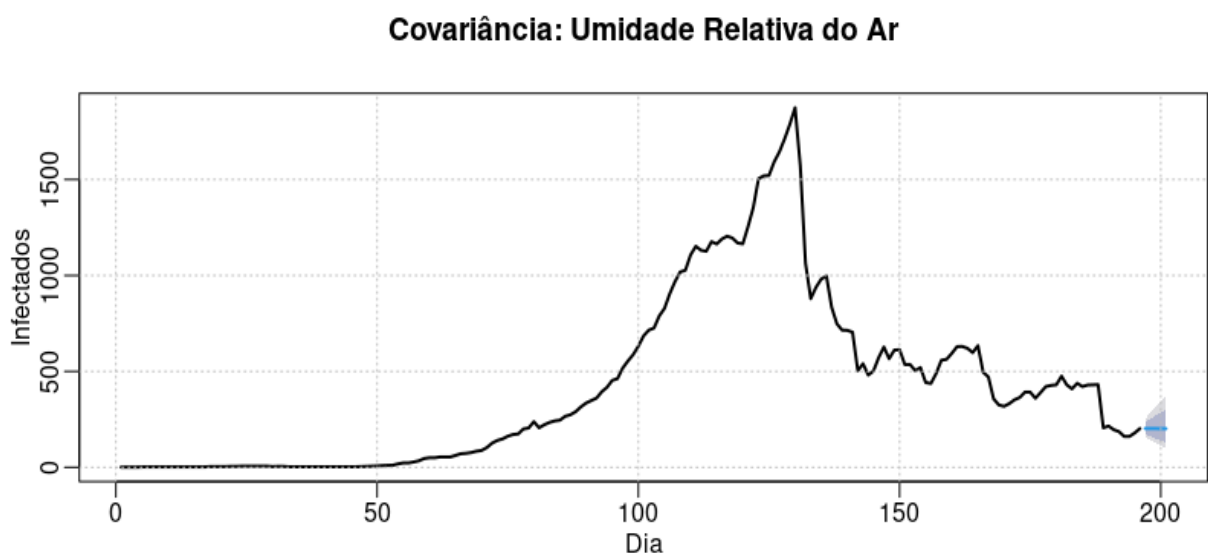


Figura 55 - Gráfico de previsão da série temporal Infectados (linha preta), do município de Várzea Grande-MT, em covariância com a variável umidade relativa do ar, com previsão (linha azul) de 5 dias, e intervalos de confiança de 80% (área azul claro) e 95% (área cinza claro).

Fonte: Do Autor, 2023.

A figura 55 apresenta uma série de previsões para a variável "infectados", considerando a variável exógena "umidade relativa do ar", para um horizonte de previsão de 5 dias. Com os limites do intervalo de confiança de 80%, representados na cor azul claro. Isso significa que há uma probabilidade de 80% de que o número real de infectados esteja dentro desse intervalo, considerando a umidade relativa do ar.

Os limites do intervalo de confiança de 95%, são representados pela área da cor cinza claro, que significa que há uma probabilidade de 95% de que o número real de infectados esteja dentro desse intervalo, considerando a umidade relativa do ar.

Por exemplo, a previsão para o primeiro dia (197) da previsão é de 202 infectados, com um intervalo de confiança de 80% de 174 a 234 infectados e um intervalo de confiança de

95% de 160 253 infectados. Os valores de limite dos intervalos de confiança aumentam à medida que se projeta para dias posteriores.

Conforme mostrado na tabela 11 são apresentadas todos os dados de previsão da série infectados e suas relações com os parâmetros meteorológicos combinados. Com objetivo da construção da série de previsão dos registros de infectados do município de Rondonópolis,

Tabela 11 - Tabela com a previsão da variável infectados para Rondonópolis-MT.

Variáveis	1	2	3	4	5
Infectados	1632	1662	1694	1727	1760
Infectados + temperatura do ar	1632	1662	1695	1727	1761
Infectados + umidade relativa do ar	1632	1662	1694	1727	1760
Infectados + precipitação pluviométrica	-	-	-	-	-
Infectados + temperatura + umidade	1632	1663	1696	1728	1762
Infectados + temperatura + umidade + precipitação	-	-	-	-	-

Fonte: Do autor, 2023.

Com base nos critérios de avaliação do modelo apresentados na tabela, o modelo ARIMA (1,0,0) com apenas a variável fundamental "Infectados" é o melhor modelo a ser utilizado para descrever a dinâmica da série temporal na região de Rondonópolis-MT.

Incluir as covariáveis "temperatura do ar" e "umidade relativa do ar" no modelo não melhora a sua performance, o que indica que essas variáveis não são importantes para explicar a variação dos casos de infectados na série temporal em questão. Dessa forma, o modelo mais simples que utiliza apenas a variável "Infectados" é o mais adequado para analisar a série temporal dos casos de infectados na região de Rondonópolis-MT.

A finalidade da figura 56 é mostrar o gráfico da evolução temporal da previsão dos casos de Covid-19 em Rondonópolis-MT e utilizá-la para fazer uma projeção do número de infectados para os próximos cinco dias.

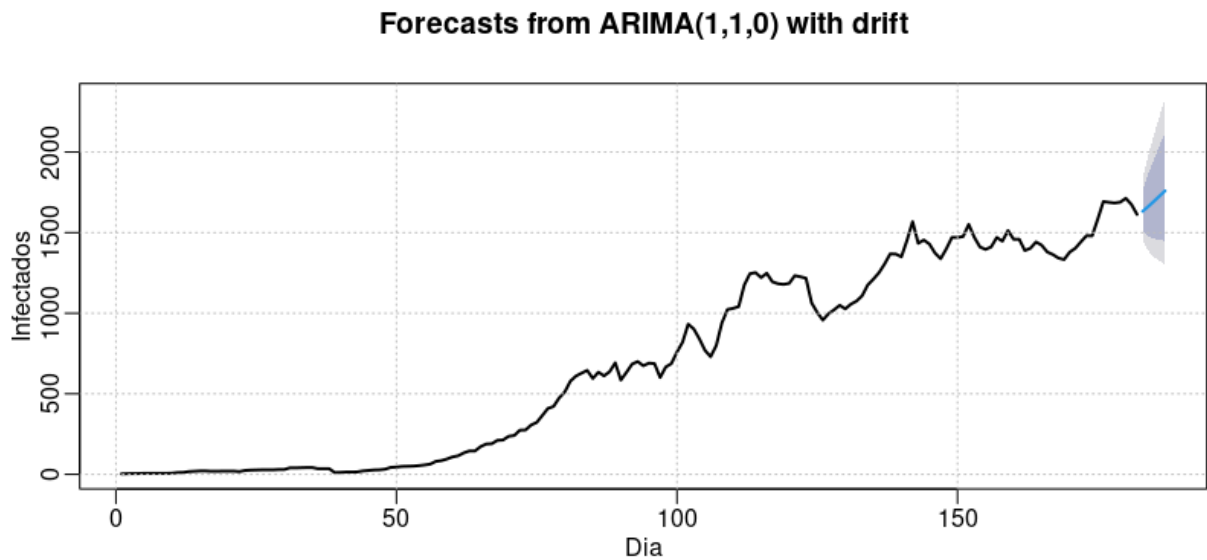


Figura 56 - Gráfico de previsão da série temporal Infectados (linha preta), do município de Rondonópolis-MT, com previsão de 5 dias (linha azul), e intervalos de confiança de 80% (área azul claro) e 95% (área cinza claro).

Fonte: Do Autor, 2023.

Como exemplo, com um nível de confiança de 80%, o gráfico mostrado na figura 56 prevê que o número de infectados varie entre 1509 e 1762 no dia 183, entre 1477 e 1864 no dia 184, entre 1461 e 1954 no dia 185, entre 1453 e 2038 no dia 186, e entre 1449 e 2119 no dia 187. Com um nível de confiança de 95%, espera-se que o número de infectados varie entre 1448 e 1834 no dia 183, entre 1387 e 1978 no dia 184, entre 1349 e 2105 no dia 185, entre 1323 e 2221 no dia 186, e entre 1304 e 2332 no dia 187.

5. CONSIDERAÇÕES FINAIS

Pois bem, após as análises dos históricos de casos de infectados, das variáveis e dos ajustes e modelagem ARIMA, para as 3 cidades mais populosas de Mato Grosso, o melhor resultado para as séries temporais de infectados e as suas respectivas covariações fora as seguintes:

- a) Cuiabá: o modelo que inclui a variável de precipitação pluviométrica observada na combinação para entre infectados e a precipitação pluviométrica obteve o melhor ajuste de dados através do modelo estatístico ARIMA (4,1,0);
- b) Várzea grande: a previsão dos dados relacionados com a variável umidade relativa do ar apresentou o menor valor de AIC, AICc e BIC entre todos os modelos ajustados, indicando que é o modelo mais adequado para descrever a relação entre as variáveis com a utilização do modelo ARIMA (1,1,1);
- c) Rondonópolis: com base nos critérios de avaliação dos modelos apresentados, o modelo ARIMA (1,1,0) com apenas a variável "Infectados" é o melhor modelo a ser utilizado para descrever a dinâmica da série temporal na região no período analisado.

De tal modo, essa modelagem tendo em vista históricos anteriores das séries estudadas, bem como a análise dos casos utilizando as covariáveis climáticas, é capaz de predizer dentro dos limites dos intervalos de confiança a previsão dos casos de infectados de Cuiabá, Várzea Grande e Rondonópolis no período compreendido entre 14/03/2020 e 27/09/2020, sendo instrumento prestante a orientação das políticas públicas na adoção de estratégias operacionais pelos serviços de saúde no desenvolvimento de suas ações preventivas e assistenciais. Nesse viés antever a progressão da epidemia com modelos preditivos se torna de grande importância.

Obviamente não pode se esperar que os parâmetros meteorológicos derrotem a Covid-19. Porém, esses parâmetros servem como subsídio base na previsão de séries temporais, uma vez que fazem parte do conjunto de variáveis envolvidas em estudos sobre a propagação de um vírus, mesmo sabendo-se que outras variáveis são mais importantes na adoção eficiente de políticas públicas, como foi o distanciamento social, estas quais podem ser alvo de novas pesquisas estatísticas. Ou seja, notavelmente, o distanciamento social pode ter tido um impacto direto nos resultados dessas pesquisas por causa da mudança do comportamento natural do vírus.

Logo, a modelagem estatística e computacional utilizando como variáveis auxiliares dados de séries climatológicas, pode fornecer modelos estatísticos para previsão de curto prazo das séries temporais associadas à progressão da epidemia da COVID-19 no Estado de Mato Grosso, estas quais estão inseridas como item básico na influência da variável infectados.

É importante destacar que a análise da propagação do COVID-19 é um problema complexo e multifacetado que envolve muitas variáveis, como densidade populacional, características socioeconômicas, estratégias de prevenção e controle e muitas outras. Portanto, é necessária uma abordagem multidisciplinar para entender completamente a epidemia e planejar intervenções efetivas para controlar sua disseminação.

6. REFERÊNCIAS BIBLIOGRÁFICAS

AB'SABER, A. N. **Os domínios de natureza no Brasil: potencialidades paisagísticas**. São Paulo: Ateliê Editorial, 1977.

ANGELO, H.; SILVA, G. F. da; MORAES E SILVA, V. S. **Análise econômica da indústria de madeiras tropicais: o caso do polo de Sinop- MT**. *Ciência Florestal*, Santa Maria, v. 14, n. 2, p. 91-101, 2004. ISSN 0103-9954.

AKAIKE, H. **Maximum likelihood identification of Gaussian autoregressive moving average models**. *Biometrika*, v.60, n.2, p.255{265, 1973.

AULER, A.C.; CÁSSARO, F.A.M.; SILVA V.O.; PIRES, L.F. (2020). **Evidence that high temperatures and intermediate relative humidity might favor the spread of COVID-19 in tropical climate: A case study for the most affected Brazilian cities** *Science of the Total Environment*, 729, 139090. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0048969720326073?dgcid=coauthor#bb0035>>. Acesso em: 01 dez. 2022.

AYOADE, J. O. **Introdução à climatologia dos trópicos**. Rio de Janeiro: Bertrand Brasil, 332 p., 2001.

BROCKWELL, P. J.; DAVIS, R. A. **Introduction to Time Series and Forecasting**. 3ª ed. Springer, 2016.

BUKHARI, Q.; JAMEEL, Y. (2020). **Will Coronavirus Pandemic Diminish by Summer?** *SSRN*, id3556998. Disponível em: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3556998>. Acesso em: 11 nov. 2022.

CALBÓ, J.; PAGÈS, D.; GONZÁLEZ, J.A. (2005). **Empirical studies of cloud effects on UV radiation: A review**. *Reviews of Geophysics*, 43(2), 1-28. Disponível em: <<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2004RG000155>>. Acesso em: 01 jan. 2023.

CALBÓ J.; PAGÈS, D.; GONZÁLEZ, J.A. (2005). **Empirical studies of cloud effects on UV radiation: A review**. *Reviews of Geophysics*, 43(2), 1-28. Recuperado de <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2004RG000155>

CAMARA, Renata Kelen Cardoso et al. **Modelagem Hidrológica Estocástica Aplicada ao Rio Tocantins para a Cidade de Marabá-PA**. *Revista Brasileira de Meteorologia*, São Paulo, v. 31, n. 1, p. 11-23, mar. 2016. Disponível em: <<http://www.scielo.br/pdf/rbmet/v31n1/0102-7786-rbmet-31-01-0011.pdf>>. Acesso em: 27 jan. 2022.

CHAN, KH.; PEIRIS, JS.; LAM, SY.; POON, LL.; YUEN, KY.; SETO, WH. **The Effects of Temperature and Relative Humidity on the Viability of the SARS Coronavirus**. *Adv Virol*. 2011; 2011:734690.

COHEN, J. (2020). **A WHO-led mission may investigate the pandemic's origin. Here are the key questions to ask**. *Science*. Disponível em: <<https://www.sciencemag.org/news/2020/07/who-led-mission-mayinvestigate-pandemic-s-origin-here-are-key-questions-ask>>. Acesso em: 24 nov. 2022.

CONSTANTINO, M.; MENDES, D. R. F.; MOREIRA, T. B. S. **Análise e previsão da taxa de inflação no Brasil: uma abordagem via modelo SARIMA**. *Revista Científica Hermes*, [S. l.], v. 24, p. 244–257, 2019. DOI: 10.21710/rch.v24i0.446. Disponível em: <<http://www.revistahermes.com.br/index.php/hermes1/article/view/446>>. Acesso em: 17 jan. 2023.

CRUZ, G. C. F. da. **Clima e COVID 19: as relações entre a temperatura do ar, radiação solar e os novos casos diários em Ponta Grossa, PR**. *Terr@ Plural*, [S. l.], v. 16, p. 1–15, 2022. Disponível em: <<https://revistas.uepg.br/index.php/tp/article/view/17446>>. Acesso em: 4 jan. 2023.

CRYER, J. D.; CHAN, K. S. **Time Series Analysis: With Applications in R**. 2^a ed. Springer, 2008.

DA ROCHA, Ângela F.; DE PAULA, D. C. J.; E SOUZA, N. S.; E SILVA, P. C. B. Da S.; DE MIRANDA, S. A.; ZAMADEI, T.; DE SOUZA, A. P.; MACHADO, N. G.; SANTOS, F. M. De M.; NOGUEIRA, J. De S.; NOGUEIRA, M. C. De J. A. **Variações Microclimáticas de Áreas Urbanas em Biomas no Estado de Mato Grosso: Cuiabá E Sinop.** Revista Gestão & Sustentabilidade Ambiental, [S. l.], v. 4, p. 246–257, 2015. DOI: 10.19177/rgsa.v4e02015246-257. Disponível em: https://portaldeperiodicos.animaeducacao.com.br/index.php/gestao_ambiental/article/view/3355. Acesso em: 13 mar. 2023.

DOLIN, R. **Common viral respiratory infections and severe acute respiratory syndrome (SARS).** In: Fauci, A. S. et al. 17 ed. Philadelphia: MacGraw-Hill, 2007

EHLERS, R. S. **Análises de séries temporais.** Universidade Federal do Paraná – Departamento de Estatística, v. 5, 2009.

EMBRAPA. **Cerrado.** Disponível em: <https://www.embrapa.br/cerrado>. Acesso em: 15 mar. 2023.

EMILIANO, P. C. et al. **Crítérios de informação de akaike versus bayesiano: análise comparativa.** 19º Simpósio Nacional de Probabilidade e Estatística, 2010.

FEDOROVA, N. **Meteorologia Sinótica.** Pelotas: Ed. UFPel FERREIRA, P. G. C. et al. **Análise de Séries Temporais em R: curso introdutório.** 1. ed. [S.l.]: Elsevier, 2018. ISBN 978-85-352-9087-5.

FRANCA, Rafael Rodrigues da et al. **Correlações entre variações climáticas e a disseminação do novo coronavírus no Brasil.** Participação, Brasília, ano 19, ed. esp., n. 34, p. 135-137, nov. 2020. Disponível em: https://drive.google.com/file/d/1_y95_7QMT_wC8vhwQUCJamcPgTvbjtBC/view. Acesso em: 02 out. 2022.

FUNCEME. **Clima de Cuiabá.** Disponível em: <http://www.funceme.br/apolo/apolo.cgi?&x=u&e=24&p=>>. Acesso em: 15 mar. 2023.

GARCÍA, A. C., & Tsai, H. T. (2007). **A new approach to forecasting automobile sales in Taiwan with Google trends data and the ARIMA model.** *Transportation Research Part A: Policy and Practice*, 41(10), 909-918.

GONÇALVES, L. **Características das séries temporais.** 2018. Disponível em: <<http://www.abgconsultoria.com.br/blog/caracteristicas-das-series-temporais/>>. Acesso em: 02 out. 2022.

GONZAGA A.C.; FREITAS, A. C. V. **Análise da relação entre condições meteorológicas, poluição atmosférica e infecções respiratórias virais: o caso COVID-19.** Itajubá: 2020. *Research, Society and Development*, v. 9, n. 8, e591985942, 2020 (CC BY 4.0) | ISSN 2525-3409 | DOI: Disponível em: <<http://dx.doi.org/10.33448/rsd-v9i8.5942>>. Acesso em: 07 out. 2022.

GOMES, F. C. **Os modelos ARIMA e a abordagem de Box-Jenkins: uma aplicação na previsão do IBOVESPA a curtíssimo prazo.** *RAE-Revista de Administração de Empresas*, [S. l.], v. 29, n. 2, p. 63–70, 1989. Disponível em: <<https://bibliotecadigital.fgv.br/ojs/index.php/rae/article/view/38778>>. Acesso em: 20 mar. 2023.

GUJARATI, D. M; PORTER, D. C. **Econometria básica.** 5. Ed. São Paulo: Makron books, 924 p., 2011.

HAMILTON, J. D. **Time Series Analysis.** Princeton, NJ: Princeton Univ. Press, 1994.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice.** 3ª ed. OTexts, 2018.

HYNDMAN, R.J; ATHANASOPOULOS, G; BERGMEIR, C; CACERES, G; CHHAY, L; O'HARA-WILD, M; PETROPOULOS, F; RAZBASH, S; WANG, E; YASMEEN, F. **Forecast: Forecasting functions for time series and linear models.** 2019. R package version 8.5.

HYNDMAN, R.J.; KHANDAKAR, Y. (2008). **Automatic time series forecasting: The forecast package for R**. Journal of Statistical Software, 26(3), 1-22.

HOTTA, Y.; LIU, H. (2020). **A seasonal autoregressive integrated moving average (SARIMA) model for forecasting Japanese tourism demand**. Tourism Management, 80, 104156.

IBAMA. **Biomass Brasileiros**. Disponível em: <<http://www.ibama.gov.br/biomass>>. Acesso em: 15 mar. 2023.

IBGE, 2022. **Instituto Brasileiro de Geografia e Estatística. Cidades e Estados**. Disponível em: <<https://www.ibge.gov.br/cidades-e-estados/mt.html>> Acesso em: 21/11/2022.

IBGE. **Manual técnico da vegetação brasileira**. Rio de Janeiro: IBGE, 2004.

INMET. **Normais Climatológicas do Brasil 1981-2010**. Disponível em: <<http://www.inmet.gov.br/portal/index.php?r=clima/normaisclimatologicas>>. Acesso em: 15 mar. 2023.

KOTTEK, M. et al. **World map of the Köppen-Geiger climate classification updated**. Meteorol. Z., Stuttgart, v. 15, n. 3, p. 259-263, 2006.

LIPSITCH, M. (2020). **Seasonality of SARS-CoV-2: Will COVID-19 go away on its own in warmer weather? Center for Communicable Disease Dynamics. Harvard Chan School of Public Health**. Disponível em: <<https://ccdd.hsph.harvard.edu/will-covid-19-go-away-on-its-own-in-warmer-weather/>>. Acesso em: 07 dez. 2022.

LOMBARDO, M. A. **Ilha de calor nas metrópoles: o exemplo de São Paulo**. Editora Hucitec com apoio de Lalekla SA Comércio e Indústria, 1985.

MELO, Fabiane de Sousa. **Modelagem de série temporal de homicídios dolosos registrados em Cuiabá-MT e Várzea Grande com Variáveis Mesoclimáticas pelo Método de Box-Jenkins**. Cuiabá-MT, 2019.

MENDES, P. F. et al. **Diversidade e similaridade florística em remanescentes de Cerrado no município de Rondonópolis, MT**. *Floresta e Ambiente*, v. 25, e20170208, 2018. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-09802018000200308>. Acesso em: 13 mar. 2023.

MINITAB. **Função de autocorrelação (ACF)**. 2019. Disponível em: <<https://support.minitab.com/pt-br/minitab/18/help-and-how-to/modeling-statistics/time-series/how-to/autocorrelation/interpret-the-results/autocorrelation-function-acf/>>. Acesso em: 23 dez. 2022.

MORETTIN, P. A.; TOLOI, C. M. **Análise de Séries Temporais**. 2. ed. São Paulo: Edgard Blucher, 535p., 2006.

MMA. **Biomás Brasileiros**. Disponível em: <<https://www.gov.br/mma/pt-br/assuntos/biodiversidade/biomass-brasileiros>>. Acesso em: 15 mar. 2023.

NIMER, E. (1989). **Climatologia do Brasil**. Série Recursos Naturais e Meio Ambiente, nº 4. Fundação IBGE. Brazil.

OMS, World Health Organization. **WHO Director- General's opening remarks at the media brief - ing on COVID-19 – 11 March 2020**. Disponível em: < <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>>. Acesso em: Out. 2022.

ONU. **Organização Mundial da Saúde declara novo coronavírus uma pandemia**. *ONU News*. (2020). Disponível em: <<https://news.un.org/pt/story/2020/03/1706881>>. Acesso em: 04 set. 2022.

PEREIRA, A. R.; ANGELOCCI, L. R. e SENTELHAS, P. C. **Agrometeorologia: fundamentos e aplicações práticas**. Guaíba: Agropecuária. 2002.

PRATA, D. et al. **Temperature significantly changes COVID-19 transmission in (sub) tropical cities of Brazil**. Science of the Total Environment. V. 729, 138862, 2020.

PRATA, D.N.; RODRIGUES, W.; BERMEJO, P.H. (2020). **Temperature significantly changes COVID-19 transmission in (sub)tropical cities of Brazil**. Science of the Total Environment, 729, 138862.

ROMERO, M. A. B. **Princípios bioclimáticos para o desenho urbano**. São Paulo: Projeto, 2001. 2ª edição, 128 p.

SAGRIPANTI, J.L.; LYTLE, C.D. (2020). **Estimated inactivation of coronaviruses by solar radiation with special reference to COVID-19**. Photochemistry and Photobiology, 96(4),731-737. Disponível em: <<https://onlinelibrary.wiley.com/doi/full/10.1111/php.13293>>. Acesso em: 02 out. 2022.

SAGRIPANTI, J.L.; Lytle, C.D. (2020). **Estimated inactivation of coronaviruses by solar radiation with special reference to COVID-19**. Photochemistry and Photobiology, 96(4),731-737. Disponível em: <<https://onlinelibrary.wiley.com/doi/full/10.1111/php.13293>>. Acesso em: 14 dez. 2022.

SAJADI, M.M.; HABIBZADEH, P.; VINTZILEOS, A.; SHOKOUHI, S.; MIRALLES-WILHELM, F.; AMOROSO, A. (2020). **Temperature, Humidity, and Latitude Analysis to Predict Potential Spread and Seasonality for COVID-19**. SSRN, id3550308. Disponível em: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3550308>. Acesso em: 25 ago. 2022.

SANTOS, F. M. de M. **Influência da ocupação do solo na variação termo-higrométrica na cidade de Cuiabá-MT**. Tese de doutorado. UFMT. 2012.

SCHWARZ, G. **Estimating the dimension of a model**. The annals of statistics. Hayward, v.6, n.2, p.461{464, Abr. 1978.

SES, Secretaria do Estado de Saúde do Mato Grosso. **SES confirma primeiro caso de coronavírus em MT.** 20/03/2020. Disponível em: <<http://www.saude.mt.gov.br/noticia/6199>>. Acesso em: 25 set. 2022.

SHUMWAY, R. H.; STOFFER, D. S. (2017). **Time series analysis and its applications: With R examples.** Springer.

SILVA, M. C. et al. **Degradação ambiental no Cerrado do estado de Mato Grosso: principais causas e desafios para a conservação.** Biota Amazônia, v. 9, n. 2, p. 42-52, 2019. Disponível em: <<https://periodicoscientificos.ufmt.br/ojs/index.php/biotaamazonia/article/view/8109>>. Acesso em: 13 mar. 2023.

SOUZA, R. C.; CAMARGO, M. E. **Análise e previsão de séries temporais: os modelos ARIMA.** Ijuí: Sedigraf, 1996.

STEPHANIE. **KPSS Test: Definition and Interpretation.** 2016. Acessado em 10/07/2019. Disponível em: <<https://www.statisticshowto.datasciencecentral.com/kpss-test/>>. Acesso em: 25 set. 2022.

TUSELL, F; **Kalman Filtering in R.** Journal of Statistical Software, [S. l.], v. 39, n. 2, p. 1–27, 2011. DOI: 10.18637/jss.v039.i02. Disponível em: <<https://www.jstatsoft.org/index.php/jss/article/view/v039i02>>. Acesso em: 21 mar. 2023.

YNOUE, R.Y.; REBOITA, M.S.; AMBRIZZI, T.; DA SILVA, G.A.M. (2017). **Meteorologia: noções básicas.** São Paulo: Oficina de Textos.

YUAN, J.; YUN, H.; LAN, W.; WANG, W.; SULLIVAN, S.G.; JIA, S.; et al. **A climatologic investigation of the SARS-CoV outbreak in Beijing, China.** American Journal of Infection Control, v.34, n.4, p.234-236, 2006.