



**UNIVERSIDADE ESTADUAL DA PARAÍBA  
CAMPUS I - CAMPINA GRANDE  
CENTRO DE CIÊNCIAS E TECNOLOGIA  
DEPARTAMENTO DE ESTATÍSTICA  
CURSO DE GRADUAÇÃO EM ESTATÍSTICA**

**LUCAS MANOEL BATISTA DE ALBUQUERQUE**

**ANÁLISE COMPARATIVA ENTRE OS MODELOS PROPHET E SARIMA NA  
PREVISÃO DE PRECIPITAÇÃO PLUVIOMÉTRICA EM CAMPINA GRANDE - PB:  
UM ESTUDO EM SÉRIES TEMPORAIS**

**CAMPINA GRANDE -PB**

**2023**

**LUCAS MANOEL BATISTA DE ALBUQUERQUE**

**ANÁLISE COMPARATIVA ENTRE OS MODELOS PROPHET E SARIMA NA  
PREVISÃO DE PRECIPITAÇÃO PLUVIOMÉTRICA EM CAMPINA GRANDE - PB:  
UM ESTUDO EM SÉRIES TEMPORAIS**

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de bacharel em Estatística.

**Área de concentração:** Estatística

**Orientador:** Prof.Dr. Tiago Almeida de Oliveira.

**CAMPINA GRANDE -PB**

**2023**

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

A345a Albuquerque, Lucas Manoel Batista de.  
Análise comparativa entre os modelos Prophet e SARIMA na previsão de precipitação pluviométrica em Campina Grande - PB: [manuscrito] : um estudo em séries temporais / Lucas Manoel Batista de Albuquerque. - 2023.

30 p. : il. colorido.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2023.

"Orientação : Prof. Dr. Tiago Almeida de Oliveira, Coordenação do Curso de Estatística - CCT. "

1. Prophet. 2. SARIMA. 3. Meteorologia. 4. Séries temporais. I. Título

21. ed. CDD 519.5

**LUCAS MANOEL BATISTA DE ALBUQUERQUE**

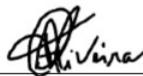
**ANÁLISE COMPARATIVA ENTRE OS MODELOS PROPHET E SARIMA NA  
PREVISÃO DE PRECIPITAÇÃO PLUVIOMÉTRICA EM CAMPINA GRANDE - PB:  
UM ESTUDO EM SÉRIES TEMPORAIS**

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de bacharel em Estatística.

**Área de concentração:** Estatística

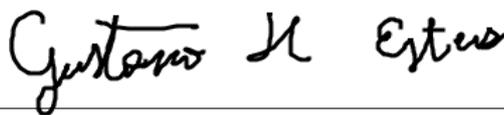
Aprovado em: 07/12/2023

**BANCA EXAMINADORA**



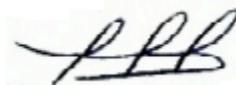
---

Prof. Dr. Tiago Almeida de Oliveira (Orientador)  
Universidade Estadual da Paraíba (UEPB)



---

Prof. Dr. Gustavo Henrique Esteves  
Universidade Estadual da Paraíba (UEPB)



---

Prof. Dr. Alvaro Ramon Paiva Sanz  
Universidade Estadual da Paraíba (UEPB)

Dedico este trabalho a minha família e amigos  
por confiarem em mim.

## **AGRADECIMENTOS**

Primeiramente, expresso minha gratidão a Deus por sempre guiar-me com Suas bênçãos e orientação por meio das minhas orações. Agradeço imensamente à minha mãe, Elizabeth Eva de Albuquerque, e à minha tia, Lucia Cavalcanti, por seu contínuo apoio e incentivo que foram fundamentais para que eu concluísse meu curso. Reconheço e sou grato a todos os professores que estiveram ao meu lado ao longo de toda a jornada acadêmica, contribuindo significativamente para o meu aprendizado. Também expresso minha gratidão a todos os amigos que fiz durante o curso, pois compartilhamos experiências e aprendizados valiosos ao longo desse período.

“Todos os modelos estão errados, mas alguns são úteis.”  
(George Box)

## RESUMO

A previsão precisa de padrões de precipitação pluviométrica é essencial para diversas áreas, especialmente em locais como Campina Grande – PB. Este estudo propõe ajustar um modelo de séries temporais utilizando o *Prophet* para prever a precipitação pluviométrica mensal na cidade, comparando-o com o modelo SARIMA. A principal meta é determinar qual entre os modelos, *Prophet* ou SARIMA, é mais preciso e confiável para prever a precipitação pluviométrica. Para isso, a série de dados fornecida pelo Instituto Nacional de Meteorologia (INMET) entre jan/2002 a maio/2023, compreendendo 257 observações, foi utilizada para calibração e avaliação dos modelos *Prophet* e SARIMA. A série foi dividida em dois subconjuntos: o primeiro (jan/2002 a maio/2021) para calibração e o segundo (jun/2021 a maio/2023) para avaliação das previsões. O SARIMA(2,0,1)×(3,1,1)<sub>12</sub> apresentou a menor raiz do erro quadrático médio (REQM), destacando-se pela precisão na previsão. No entanto, o modelo *Prophet* também demonstrou boa performance, superando o SARIMA em termos de REQM, apesar de ambos captarem efetivamente a sazonalidade anual da série. Os testes estatísticos realizados detectaram tendências e sazonalidades na série, contribuindo para uma melhor interpretação e aplicabilidade dos modelos. Para o período de previsão de jun/2023 a maio/2025, os modelos ajustados serão utilizados, visando fornecer projeções confiáveis da precipitação pluviométrica na região de Campina Grande – PB.

**Palavras-chaves:** *Prophet*; SARIMA; Séries Temporais; Meteorologia.

## ABSTRACT

The accurate forecast of rainfall patterns is essential for various sectors, especially in locations like Campina Grande - PB. This study aims to fine-tune a time series model using *Prophet* to predict monthly rainfall in the city, comparing it with the SARIMA model. The main objective is to determine which of the models, *Prophet* or SARIMA, is more precise and reliable in forecasting rainfall. To achieve this, the dataset provided by the National Institute of Meteorology (INMET) from Jan/2002 to May/2023, comprising 257 observations, was used for model calibration and evaluation of *Prophet* and SARIMA. The series was divided into two subsets: the first (Jan/2002 to May/2021) for calibration and the second (Jun/2021 to May/2023) for forecast evaluation. The SARIMA(2,0,1)×(3,1,1)<sub>12</sub> showed the lowest root mean square error (RMSE), standing out for its forecast precision. However, the *Prophet* model also exhibited good performance, surpassing SARIMA in terms of RMSE, although both effectively captured the annual seasonality of the series. The statistical tests conducted identified trends and seasonal patterns in the series, contributing to a better interpretation and applicability of the models. For the forecast period from Jun/2023 to May/2025, the adjusted models will be utilized to provide reliable projections of rainfall in the Campina Grande - PB region.

**Keywords:** Prophet; SARIMA; Time Series; Meteorology.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Gráfico de barras para a média de precipitação total mensal por ano . . . . .	21
Figura 2 – Série da precipitação total mensal para a cidade de Campina Grande no período de jan/2002 a maio/2021. . . . .	22
Figura 3 – Correlograma da função de autocorrelação (FAC) dos dados de treino. . . . .	22
Figura 4 – Correlograma das funções de autocorrelação (FAC) e autocorrelação parcial (FACP) dos dados de treino com uma diferença de ordem 12. . . . .	23
Figura 5 – Valores reais (azul), valores calibrados (vermelho) e intervalo com 80% de confiança em relação a previsão (azul claro) da precipitação total mensal (mm) em Campina Grande - PB de jan/2002 a maio/2025 para o modelo <i>Prophet</i> . . . . .	25
Figura 6 – Valores reais (azul), valores calibrados (vermelho) e intervalo com 80% de confiança em relação a previsão (azul claro) da precipitação total mensal (mm) em Campina Grande - PB de jan/2002 a maio/2025 para o modelo SARIMA(2,0,1)×(3,1,1) <sub>12</sub> . . . . .	26
Figura 7 – Função de autocorrelação FAC para os resíduos do modelo SARIMA e <i>Prophet</i> , respectivamente. . . . .	26

## LISTA DE TABELAS

Tabela 1 – Hiperparâmetros e seus respectivos valores para um ajuste ótimo usando o <i>Prophet</i> . . . . .	15
Tabela 2 – Descrição dos parâmetros do modelo SARIMA . . . . .	16
Tabela 3 – Média e Desvio padrão para cada mês . . . . .	20
Tabela 4 – Métricas de avaliação dos modelos SARIMA e <i>Prophet</i> usados para validação através do AIC e REQM . . . . .	24
Tabela 5 – Valores dos hiperparâmetros do modelo do <i>Prophet</i> mais adequados para realização de previsão de precipitação pluviométrica na cidade de Campina Grande - PB . . . . .	24
Tabela 6 – Estimativas dos parâmetros do modelo SARIMA(2,0,1)×(3,1,1) <sub>12</sub> ajustado para descrever e prever a precipitação total mensal (mm) na cidade de Campina Grande – PB de jan/2002 a maio/2025. . . . .	25
Tabela 7 – previsão de precipitação pluviométrica em mm na cidade de Campina Grande - PB com o uso dos modelos <i>Prophet</i> e SARIMA(2,0,1)×(3,1,1) <sub>12</sub> . . . . .	27

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> . . . . .	<b>11</b>
<b>2</b>	<b>MATERIAL E MÉTODOS</b> . . . . .	<b>12</b>
<b>2.1</b>	<b>Banco de dados</b> . . . . .	<b>12</b>
<b>2.2</b>	<b>Série temporal</b> . . . . .	<b>12</b>
<b>2.3</b>	<b>Estacionariedade</b> . . . . .	<b>12</b>
<b>2.4</b>	<b>Metodologia Box-Jenkins</b> . . . . .	<b>12</b>
<b>2.5</b>	<b>Teste para identificar tendência</b> . . . . .	<b>13</b>
<b>2.6</b>	<b>Teste para identificar sazonalidade</b> . . . . .	<b>13</b>
<b>2.7</b>	<b>Modelo <i>Prophet</i> do Facebook</b> . . . . .	<b>14</b>
<b>2.8</b>	<b>Modelo SARIMA</b> . . . . .	<b>15</b>
<b>2.9</b>	<b>Função de autocorrelação FAC</b> . . . . .	<b>16</b>
<b>2.10</b>	<b>Função de Autocorrelação Parcial FACP</b> . . . . .	<b>17</b>
<b>2.11</b>	<b>Previsão usando SARIMA</b> . . . . .	<b>17</b>
<b>2.12</b>	<b>Critérios para escolha dos melhores modelos</b> . . . . .	<b>17</b>
<b>2.13</b>	<b>Análise dos resíduos</b> . . . . .	<b>18</b>
<b>3</b>	<b>RESULTADOS E DISCUSSÃO</b> . . . . .	<b>20</b>
<b>4</b>	<b>CONCLUSÃO</b> . . . . .	<b>28</b>
	<b>REFERÊNCIAS</b> . . . . .	<b>29</b>

## 1 INTRODUÇÃO

O município de Campina Grande, situada na região Nordeste do Brasil, é conhecida por suas particularidades climáticas, onde os padrões de precipitação têm um papel significativo na vida cotidiana, na agricultura e no meio ambiente local. Compreender as variações, tendências e ciclos sazonais dessas chuvas ao longo do tempo é essencial para diversos setores, desde o planejamento urbano até a gestão dos recursos hídricos e o desenvolvimento de estratégias de adaptação às mudanças climáticas.

Segundo Dantas, Santos e Olinda (2015), as chuvas intensas tendem a aumentar na cidade de Campina Grande, ocorrendo em períodos menores e específicos do ano. Um estudo feito por Santos e Brito (2007) observou uma tendência de aumento das condições de umidade para os Estados do Rio Grande do Norte e Paraíba.

A cidade de Campina Grande apresenta um clima semiárido e é afetada por mudanças sazonais de precipitação, o que a torna um local ideal para estudos de séries temporais. Possui uma área de 591,658km<sup>2</sup> e exerce forte influência econômica no estado da Paraíba. Com uma população estimada de 419.379 pessoas, possui densidade demográfica de 708,82 hab/km<sup>2</sup> (IBGE, 2023).

De acordo com Menezes et al. (2008), o clima no Estado da Paraíba sofre grande influência das anomalias de temperatura da superfície do mar, causando verões prolongados e fora de época, causando assim uma possível quebra na sazonalidade da temperatura e ocorrência de chuvas na região. Conforme Marcuzzo e Romero (2013), estas mudanças de clima abruptas causadas pelo aquecimento e resfriamento do mar são denominadas de El Niño e La Niña, respectivamente. Estes fenômenos desempenham um papel crucial na produção econômica, especialmente nas áreas ligadas à agricultura e setores relacionados.

Com o avanço da tecnologia e o crescente interesse em análises preditivas, a utilização de modelos de machine learning em séries temporais tem se mostrado uma técnica promissora para a previsão de valores futuros. Neste contexto, o objetivo deste trabalho é comparar o desempenho do modelo SARIMA com o modelo *Prophet* na previsão da precipitação pluviométrica em Campina Grande. Os resultados alcançados neste estudo oferecem ideias valiosas para embasar decisões em regiões impactadas pela precipitação pluviométrica, além de contribuir na seleção do modelo mais adequado para projeções futuras.

## 2 MATERIAL E MÉTODOS

### 2.1 Banco de dados

A variável utilizada para análise de série temporal é dada pela precipitação pluviométrica total mensal em milímetros (mm) na cidade de Campina grande, durante o período de janeiro de 2002 até maio de 2023, concedida pelo MAPA (2023). A série é constituída de dados da estação automática e da estação convencional, em que foi priorizado os da estação automática, ou seja, os dados faltantes da série da estação automática foram substituídos pelos da estação convencional. Para as análises foi utilizado o *Software R* (Versão 4.2.0) com os pacotes *seastests*, *forecast*, *Kendall* e *stats* e a linguagem de programação *Python* (Versão 3.9.7) com os pacotes *Prophet*, *matplotlib*, *numpy* e *pandas*.

### 2.2 Série temporal

Uma série temporal é um conjunto de observações organizados em sequência dado intervalos de tempo. De acordo com Morettin e Toloí (2006), uma série temporal é um processo estocástico, que por sua vez é descrito como uma família  $Z = \{Z(t), t \in T\}$ , tal que, para cada  $t \in T$ ,  $Z(t)$  é uma variável aleatória, sendo  $T$  um conjunto arbitrário.

### 2.3 Estacionariedade

Conforme afirmado por Morettin e Toloí (2006), uma série temporal demonstra estacionariedade quando se desenvolve de maneira aleatória ao longo do tempo, mantendo uma média constante, refletindo, assim, algum tipo de equilíbrio estável.

De acordo com Morettin e Toloí (2006) na análise de séries temporais, a tendência refere-se à direção geral de longo prazo nos dados ao longo do tempo. Ela descreve a variação sistemática dos dados que se estende por um período prolongado, indicando uma direção consistente de crescimento, declínio ou estabilidade. Já a sazonalidade em uma série temporal refere-se a padrões ou flutuações regulares e previsíveis que se repetem em intervalos fixos de tempo, geralmente em períodos curtos, como diários, semanais, mensais, trimestrais ou anuais. Esses padrões são influenciados por fatores sazonais, como estações do ano, feriados, eventos climáticos ou comportamentais que ocorrem regularmente ao longo do tempo.

### 2.4 Metodologia Box-Jenkins

Uma abordagem bastante usada para séries temporais univariadas é a de Box e Jenkins (1976). De acordo com Morettin e Toloí (2006), esta abordagem é subdividida em quatro etapas: especificação, identificação, estimação e verificação.

A primeira etapa, especificação, envolve a formulação de um modelo adequado para representar a série temporal. Isso inclui a escolha do tipo de modelo a ser utilizado (como

modelos Autorregressivos, modelos de Médias Móveis, modelos compostos de Autorregressão e Média Móvel, entre outros), levando em consideração a presença de tendências, sazonalidades e padrões na série.

Em seguida, vem a etapa de identificação, na qual os parâmetros do modelo são determinados. Aqui, são utilizadas técnicas como a análise de autocorrelação e autocorrelação parcial, além de métodos de seleção de modelos, como o critério de informação Akaike AIC, para identificar a estrutura do modelo mais apropriada.

A terceira etapa, estimação, consiste em estimar os parâmetros do modelo identificado utilizando métodos como a máxima verossimilhança. Esses parâmetros são essenciais para fazer previsões e inferências sobre a série temporal.

Por fim, a etapa de verificação é crucial para avaliar se o modelo escolhido é capaz de representar adequadamente a série temporal. Isso envolve a análise dos resíduos do modelo para verificar se eles são ruídos brancos (aleatórios), bem como a realização de testes estatísticos para confirmar a adequação do modelo aos dados.

É importante ressaltar que o sucesso dessa abordagem depende da correta aplicação de cada uma dessas etapas e da interpretação cuidadosa dos resultados obtidos em cada fase do processo de modelagem de séries temporais univariadas.

## 2.5 Teste para identificar tendência

Um dos testes utilizados para detecção de tendência em uma série temporal é o teste de Mann-Kendall, desenvolvido por Mann (1945) e Sneyers (1975). A metodologia deste teste baseia-se na análise das tendências utilizando as posições das observações ao longo do tempo. Em linhas gerais, o teste compara as diferenças entre todos os pares de valores na série temporal, avaliando a direção dessas diferenças para identificar padrões de tendência. Ele verifica se há uma consistência predominante nas posições dos valores, procurando por uma tendência crescente ou decrescente. Em outras palavras, deseja-se testar duas hipóteses:

- $H_0$  : As observações da série não apresentam tendência
- $H_1$  : As observações da série apresentam uma tendência positiva ou negativa

Caso seja identificada a tendência na série de acordo com este teste, é necessário realizar uma diferenciação simples na série.

## 2.6 Teste para identificar sazonalidade

De acordo com Morettin e Tolo (2006) pode-se testar se uma série apresenta sazonalidade utilizando o teste não paramétrico de Kruskal e Wallis (1952). Para uma série temporal com dados mensais, cada mês é uma amostra de uma população que será adicionada ao teste. Desta forma, é de interesse testar duas hipóteses:

- $H_0$  : Não há sazonalidade na série
- $H_1$  : Há sazonalidade na série

É importante que este teste seja realizado após a retirada da tendência na série para que seja possível obter melhores resultados. Segundo Ollech e Webel (2020), para amostras grandes o teste de Kruskal-Wallis segue distribuição Qui-Quadrado com  $k - 1$  graus de liberdade, sendo  $k$  o número de amostras.

## 2.7 Modelo *Prophet* do Facebook

O modelo do *Prophet* foi projetado para a análise e previsão de dados de séries temporais introduzidos por Taylor e Letham (2017) e estão disponíveis nos Softwares Python e R. A formulação do modelo do *Prophet* é semelhante a um modelo aditivo generalizado (MAG), incluindo tendência, sazonalidade e aleatoriedade:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon(t), \quad (2.1)$$

em que  $g(t)$  é a função de tendência, que representa mudanças não periódicas nos valores da série temporal e  $s(t)$  representa mudanças periódicas (por exemplo, sazonalidade semanal ou mensal).  $h(t)$  representa os efeitos dos feriados, que não serão utilizados neste estudo, já que a precipitação pluviométrica não depende de feriados.  $\varepsilon(t)$  é o termo de erro e assume distribuição normal neste estudo. Em particular,  $s(t)$  usa séries de Fourier para modelar a sazonalidade, e a quantidade de ordens dessa série fica a escolha do pesquisador. Aqui usaremos uma série de Fourier igual a 30, mesma quantidade usada por Jana, Ghosh e Wallin (2022).

O *Prophet* realiza previsões usando modelos logísticos ou lineares para sua componente de tendência. Aqui usaremos apenas a componente de tendência linear, que é descrita como na equação (2.2):

$$g(t) = (k + \mathbf{a}(t)^T \boldsymbol{\delta})t + (m + \mathbf{a}(t)^T \boldsymbol{\gamma}), \quad (2.2)$$

em que  $k$  é a taxa de crescimento,  $\boldsymbol{\delta}$  é o ajuste dessa taxa,  $m$  como um parâmetro de correção geral, descrito por Taylor e Letham (2017) como *offset parameter* e  $\boldsymbol{\gamma}$  controla a magnitude das mudanças de tendência nos pontos de mudança de tendência. Usaremos o método de estimação de quantidade de pontos de mudança de tendência padrão do *Prophet*, que seleciona automaticamente a quantidade de pontos.

A sazonalidade é modelada usando uma série de Fourier como é mostrado na equação (2.3):

$$s(t) = \sum_{n=1}^N \left( a_n \cos \left( \frac{2\pi n t}{P} \right) + b_n \text{seno} \left( \frac{2\pi n t}{P} \right) \right), \quad (2.3)$$

aqui,  $P$  se refere a qual período a sazonalidade série temporal está ocorrendo, no nosso caso é anual, logo  $P = 365.25$  e  $N$  refere-se ao tamanho da série de Fourier, que já foi citada anteriormente com um valor igual à 30.

O modelo *Prophet* tem vários hiperparâmetros que podem ser considerados para o ajuste, mas apenas alguns são recomendados para um bom ponto de partida. Tais hiperparâmetros e seus respectivos valores recomendados para um ajuste ótimo são descritos a partir da Tabela 1:

Tabela 1 – Hiperparâmetros e seus respectivos valores para um ajuste ótimo usando o *Prophet*.

hiperparâmetro	Valor
<i>Change point prior scale</i>	0,001 ou 0,01 ou 0,1 ou 0,5
<i>Seasonality prior scale</i>	0,01 ou 0,1 ou 1 ou 10
<i>Seasonality mode</i>	aditivo ou multiplicativo
<i>Change point range</i>	0,8 ou 0,85 ou 0,9 ou 0,95

Fonte: Elaborada pelo autor, 2023.

Tem-se que *change point prior scale* determina a força da tendência e, em particular, o quanto a tendência muda caso haja alguma mudança de tendência na série; *Seasonality prior scale* controla a flexibilidade da sazonalidade; *Seasonality mode* determina se o modelo irá trabalhar com uma série do tipo multiplicativa ou aditiva; *Change point range* é a proporção dos dados que serão usados para estimar os possíveis pontos de mudança de tendência. Segundo Taylor e Letham (2019), com exceção do *seasonality mode*, quanto maior o valor atribuído ao hiperparâmetro, maior será a sua flexibilidade na modelagem. Para convergência dos resultados será usado o método otimização de Newton, que já está implementado na biblioteca do *Prophet* na linguagem de programação *Python*.

## 2.8 Modelo SARIMA

De acordo com Shumway e Stoffer (2000), o modelo SARIMA é uma técnica de modelagem de séries temporais que incorpora a sazonalidade estocástica e a autocorrelação dos dados e constrói um modelo composto de uma série temporal. De acordo com Box et al. (2015) o modelo é denotado como  $SARIMA(p,d,q) \times (P,D,Q)_{12}$ , onde cada um dos seus parâmetros são descritos de acordo com a Tabela 2:

Tabela 2 – Descrição dos parâmetros do modelo SARIMA

Parâmetro	Descrição
$p$	Número de termos autorregressivos
$d$	Número de diferenças não sazonais
$q$	Número de termos de médias móveis
$P$	Número de termos autorregressivos sazonais
$D$	Número de diferenças sazonais
$Q$	Número de termos de médias móveis sazonais
$S$	Período sazonal

Fonte: Elaborada pelo autor, 2023.

Tal modelo é descrito pela equação (2.4):

$$\phi(\beta) \Phi(\beta^S) (1 - \beta)^d (1 - \beta^S)^D Z_t = \theta(\beta) \Theta(\beta^S) \varepsilon_t. \quad (2.4)$$

Seus polinômios autorregressivos e de médias móveis são descritos pelas seguintes equações:

$$\phi(\beta) = 1 - \phi_1\beta - \dots - \phi_p\beta^p \quad (2.5)$$

$$\theta(\beta) = 1 - \theta_1\beta - \dots - \theta_q\beta^q \quad (2.6)$$

$$\Phi(\beta^S) = 1 - \Phi_1\beta^S - \dots - \Phi_P\beta^{Ps} \quad (2.7)$$

$$\Theta(\beta^S) = 1 - \Theta_1\beta^S - \dots - \Theta_Q\beta^{Qs} \quad (2.8)$$

a equação (2.5) é referente à ordem  $p$ , a (2.6) à ordem  $q$ , a (2.7) à ordem  $P$ , (2.8) à ordem  $Q$  e  $S$  representa o período sazonal da série. As diferenças são representadas por  $(1 - \beta)^d$  e  $(1 - \beta^S)^D$ .

Para a sua aplicação é necessário que a série apresente um comportamento estacionário, onde tal característica será verificada pelo teste de Mann-Kendall para tendência com 10% de significância estatística, ou seja, a série apresenta tendência caso seu valor  $p$  seja menor do que 0,10. Para identificar a sazonalidade característica será usado o gráfico de Função de Autocorrelação (FAC).

## 2.9 Função de autocorrelação FAC

Segundo Montgomery, Jennings e Kulahci (2015), para um processo estocástico estacionário  $\{Z_t\}$ , podemos escrever a covariância entre  $\{Z_t\}$  e  $\{Z_{t+k}\}$  como:

$$\gamma_k = Cov(Z_t, Z_{t+k}) \quad (2.9)$$

e a correlação entre  $\{Z_t\}$  e  $\{Z_{t+k}\}$  como:

$$\rho_k = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)}\sqrt{Var(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0} \quad (2.10)$$

$\rho_k$  é chamado de função de autocorrelação (FAC) pois representa a correlação entre  $Z_t$  e  $Z_{t+k}$  de um mesmo processo estocástico, separados apenas pelo lag de tempo  $k$ . O gráfico da FAC será usada para estimar os parâmetros  $q$  e  $Q$  do modelo SARIMA após todas as diferenciações necessárias serem aplicadas, assim como no estudo de Silva et al. (2022).

O gráfico da FAC também é usado para identificar se a série apresenta uma forte tendência caso as suas autocorrelações decaiam lentamente para cada defasagem. Para uma série com dados mensais, se as defasagens múltiplas de 12 apresentam um decaimento lento após ser removida a tendência, então a série apresenta uma sazonalidade característica, sendo necessária realizar uma diferenciação sazonal no lag 12.

## 2.10 Função de Autocorrelação Parcial FACP

Segundo Shumway e Stoffer (2000), a Função de Autocorrelação Parcial (FACP) para um processo estacionário  $Z_t$ , denotada por  $\phi_{tt}$ , para  $t = 1, 2, 3, \dots$ , é definida como:

$$\phi_{11} = \text{Corr}(Z_{t+1}, Z_t) \quad (2.11)$$

e

$$\phi_{tt} = \text{Corr}(Z_{t+k} - \hat{Z}_{t+k}, Z_t - \hat{Z}_t), h \geq 2 \quad (2.12)$$

De acordo com Shumway e Stoffer (2000) e Morettin e Toloí (2006), a FAC calcula as correlações entre  $Z_t$  e  $Z_{t+k}$  onde as suas dependências com as variáveis  $\{Z_{t+1}, \dots, Z_{t+k-1}\}$  são removidas.

O gráfico da FACP também será usada para estimar os parâmetros  $p$  e  $P$  do modelo SARIMA após todas as diferenciações necessárias serem aplicadas, assim como no estudo de Silva et al. (2022).

## 2.11 Previsão usando SARIMA

Segundo Silva et al. (2022), a previsão usando o SARIMA, descrita como  $\hat{Y}_t(h)$ , é a esperança matemática condicional de  $Y_{t+h}$ , dado  $Y_t, Y_{t-1}, \dots$ , ou seja:

$$\hat{Y}_t(h) = [Y_{t+h}] = E[Y_{t+h} | Y_t, Y_{t-1}, \dots] \quad (2.13)$$

## 2.12 Critérios para escolha dos melhores modelos

Após a coleta e limpeza dos dados, a próxima etapa é dividi-los em dados de treinamento e teste para estimar os parâmetros. Para métodos ingênuos, pode-se amostrar os dados aleatoriamente, pois os valores anteriores já foram levados em consideração para cada registro individual. No entanto, para métodos que levam em consideração o fator de série temporal, a divisão do treino e do teste é realizada considerando os valores mais antigos como parte do conjunto de

treinamento e os valores mais novos como parte do conjunto de teste. Portanto, os dados de jan/2002 a maio/2021 foram selecionados para o treino (modelagem), enquanto os dados de jun/2021 a maio/2023 para o teste.

O conjunto de dados de treino é usado para treinar os modelos em questão. Os algoritmos são ajustados aos dados históricos, permitindo que o modelo aprenda padrões, relações e características dos dados. Já o conjunto de dados de teste é usado para avaliar o desempenho do modelo após o treinamento. Após o modelo ter sido treinado com os dados históricos, ele é testado com esses dados que não foram usados durante o treinamento. Isso permite medir a capacidade do modelo de fazer previsões precisas ou classificações corretas em dados que não foram vistos durante o treinamento.

Uma forma de avaliar o modelo SARIMA é através do critério de informação de Akaike AIC, para os dados de treino, descrito como:

$$AIC = -2\log\hat{L} + 2k \quad (2.14)$$

em que  $\log\hat{L}$  é o logaritmo da função de máxima verossimilhança e  $k$  representa a quantidade de parâmetros do modelo, com a intenção de beneficiar modelos mais parcimoniosos, ou seja, com menos parâmetros mas que explicam bem a variabilidade dos dados. O modelo SARIMA cujo o valor do AIC seja menor será o mais adequado.

Outra forma para a avaliação dos modelos *Prophet* e SARIMA é uma medida chamada de raiz do erro quadrático médio (REQM), onde essa será aplicada entre os dados de teste que são de jun/2021 à maio/2023, que estão no mesmo período de tempo. Aquele modelo que obtiver a menor raiz do erro quadrático médio será o escolhido como o mais adequado. Sua equação é descrita como:

$$REQM = \sqrt{\frac{1}{n} \sum_{h=1}^n [Y_{t+h} - \hat{Y}(h)]^2}. \quad (2.15)$$

Já que a ideia é apresentar modelos que sejam bons em previsão, a métrica de avaliação REQM terá uma prioridade maior sobre o AIC para os modelos SARIMA estimados.

### 2.13 Análise dos resíduos

É necessário que todos os modelos tenham resíduos não autocorrelacionados, onde essa característica será averiguada pelo teste de LJUNG e BOX (1978). As hipóteses deste teste são descritas como:

- $H_0$  : Os resíduos são distribuídos de forma independente
- $H_1$  : Os resíduos apresentam autocorrelação

Caso o valor  $p$  deste teste seja maior do que 10%, quando aplicado para os resíduos do modelo estimado com a base de dados de treino, os resíduos daquele modelo não estão autocorrelacionados.

A estatística do teste é dada por:

$$W = n(n+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k}, \quad (2.16)$$

em que  $n$  é o tamanho da amostra usada,  $\hat{\rho}$  é a autocorrelação para a defasagem  $k$ , e  $h$  é o número de defasagens a serem testadas. Sob  $H_0$  a estatística  $W$  segue assintoticamente uma distribuição Qui-Quadrado com  $h$  graus de liberdade. Desta forma, com um nível de significância  $\alpha$  de 10%, a região crítica para rejeição da hipótese nula é:

$$W > \chi_{1-\alpha, h}^2, \quad (2.17)$$

em que  $\chi_{1-\alpha, h}^2$  é o quantil da distribuição Qui-Quadrado com  $h$  graus de liberdade.

### 3 RESULTADOS E DISCUSSÃO

Na Tabela 3 encontra-se as médias e desvios padrão para cada mês da série. É possível observar que os meses de março, abril, maio, junho e julho têm média de precipitação total em milímetros acima de 80, enquanto os meses de setembro, outubro, novembro e dezembro possuem uma média de precipitação total em milímetros abaixo de 26, mostrando que o período de fim de ano em Campina Grande é o que ocorre mais seca, enquanto o período de meio de ano é o que ocorre mais chuvas.

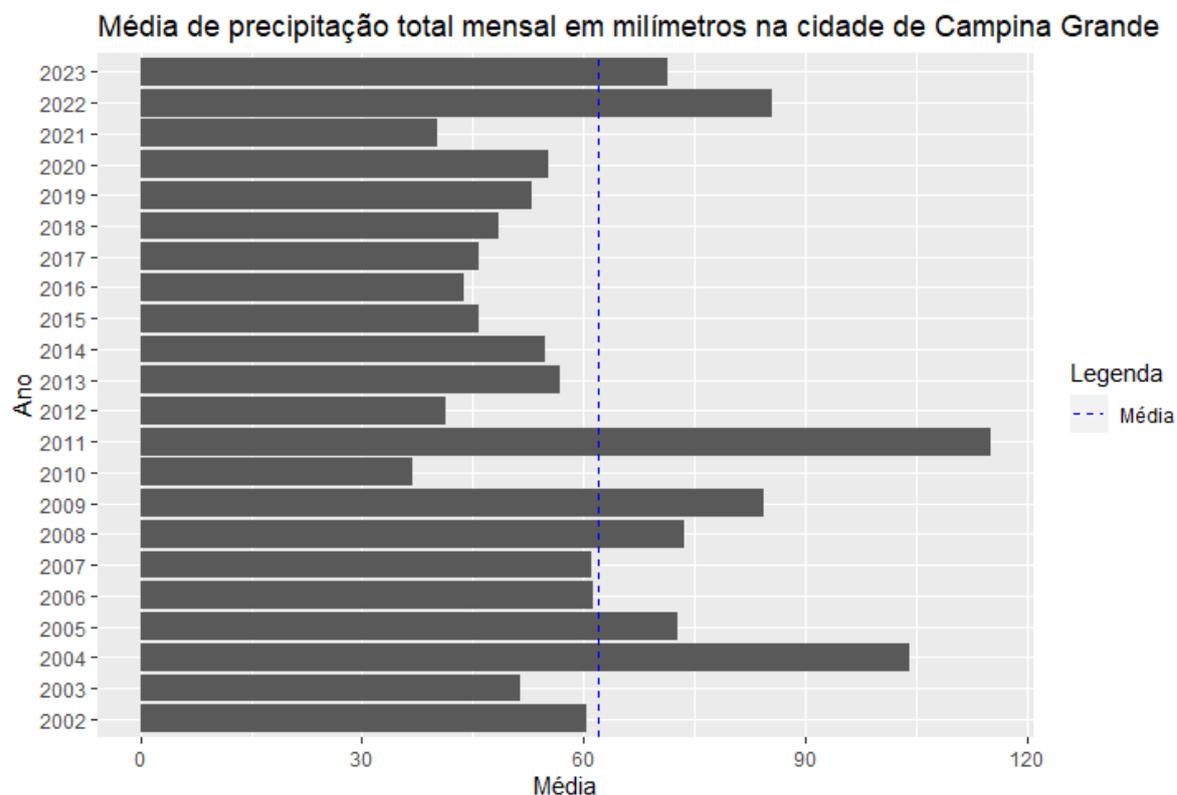
Tabela 3 – Média e Desvio padrão para cada mês

Mês	Média	Desvio padrão
Janeiro	51,513	58,70
Fevereiro	63,295	68,10
Março	88,795	55,62
Abril	82,068	47,33
Maio	103,86	76,97
Junho	114,30	59,08
Julho	105,9	69,62
Agosto	56,69	37,39
Setembro	25,37	22,85
Outubro	11,10	10,54
Novembro	11,80	14,08
Dezembro	21,88	18,76

Fonte: Elaborada pelo autor, 2023.

A Figura 1 apresenta um gráfico de barras contendo a média de precipitação total mensal para cada ano, junto com uma linha vertical pontilhada em azul indicando a média destes anos. Vemos que os anos de 2004, 2005, 2008, 2009, 2011, 2022 e 2023 apresentaram valores maiores do que a média. Consequentemente, os demais anos apresentaram valores de precipitação total mensal menores do que a média. É interessante observar que, com os dados que possuímos, a média de precipitação total mensal em milímetros na cidade de Campina Grande para o mês de dezembro ainda está incompleta. Portanto, apesar de a média para o mês de dezembro estar relativamente alta, este valor apresentado sofrerá alterações até que o ano de 2023 encerre.

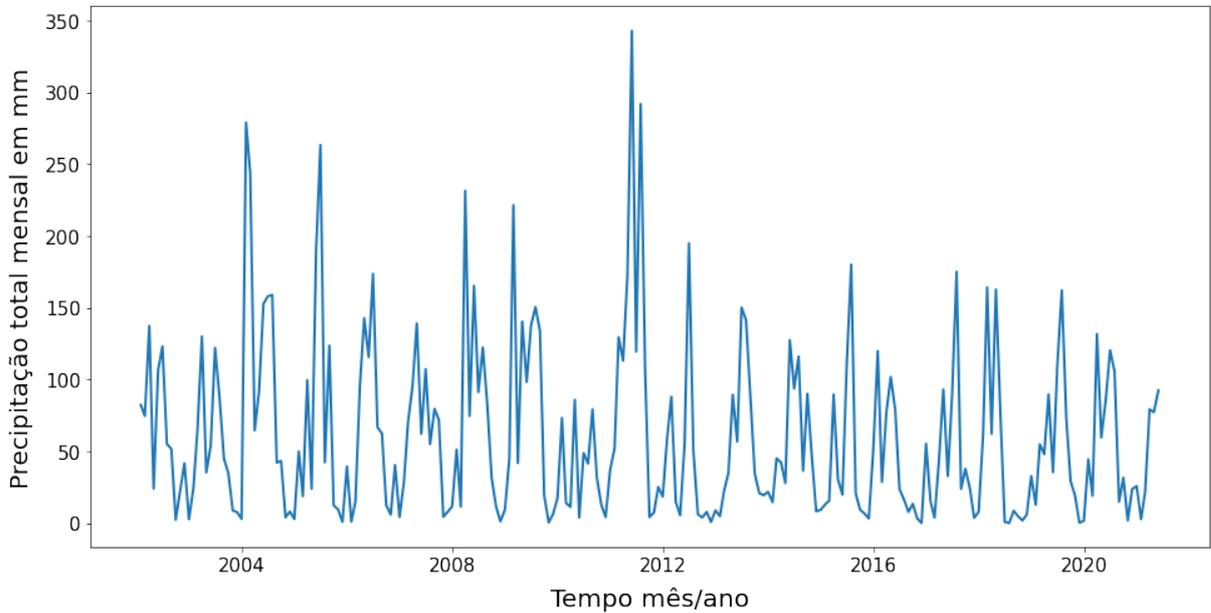
Figura 1 – Gráfico de barras para a média de precipitação total mensal por ano



Fonte: Elaborada pelo autor, 2023.

A Figura 2 apresenta o gráfico da série temporal de precipitação total mensal para a cidade de Campina Grande, em milímetros (mm), no período de jan/2002 a maio/2021 (dados de treino). É possível observar uma sazonalidade anual na série, evidenciada pelos picos e baixos igualmente espaçados ao longo dos meses, além de apresentar um leve declínio nos valores máximos da série para cada período anual a partir do ano de 2004. Esta característica dá indícios de que a sazonalidade na série exista, mas sendo menos forte ao passar dos anos. Além disso, a série temporal presente na Figura 2 não aparenta possuir uma tendência positiva ou negativa.

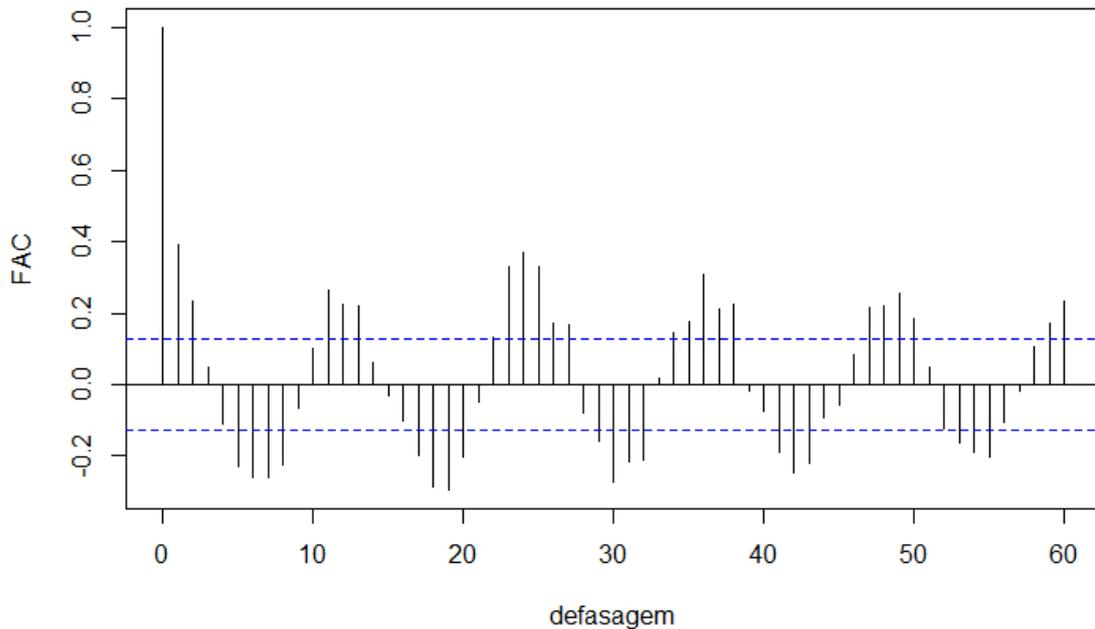
Figura 2 – Série da precipitação total mensal para a cidade de Campina Grande no período de jan/2002 a maio/2021.



Fonte: Elaborada pelo autor, 2023.

Na Figura 3 está exposto o gráfico da Função de Autocorrelação para os dados de treino sem nenhuma diferenciação. É possível observar que não há um decaimento lento nas defasagens de um a um, mas para as defasagens múltiplas de 6 há um decaimento lento e gradual, revelando evidências de que a série não possui uma tendência forte e sim uma sazonalidade.

Figura 3 – Correlograma da função de autocorrelação (FAC) dos dados de treino.



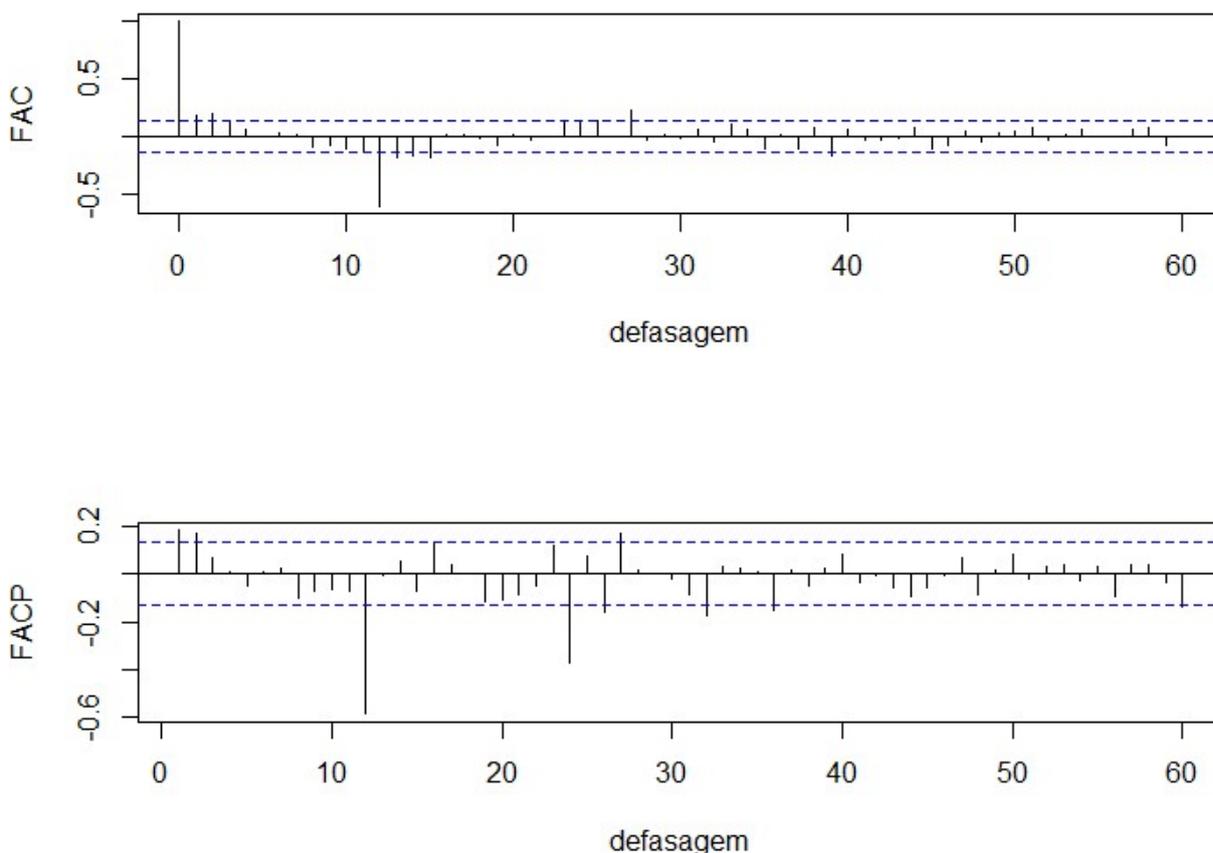
Fonte: Elaborada pelo autor, 2023.

Aplicando o teste de MannKendall para tendência na série, obteve-se um valor  $\tau$  de

Kendall de -0,0866, junto com um valor  $p$  significativo de 0,0494. Como o valor tau de Kendall foi muito baixo, a série não apresenta tendência, apesar do valor  $p$  significativo. Por conseguinte, foi aplicado o teste de Kruskal-Wallis na série, que retornou um valor  $p$  menor do que 0,01, rejeitando, ao nível de 1% de significância, a hipótese nula de ausência de sazonalidade, sendo assim, foi preciso uma diferenciação de ordem 12 na série.

Na Figura 4 estão representados os correlogramas da FAC e FACP para a série temporal com uma diferença de ordem 12. O correlograma da FACP nos indica, a partir das defasagens 1 e 2, as possíveis ordens do modelo para o parâmetro autorregressivo  $p$ . Existem duas defasagens significativas múltiplas de 12 na FACP, indicando que a ordem sazonal de  $P$  é 1 ou 2. No correlograma da FAC, as defasagens significativas são 1, 2, 3 e 12, sendo a décima segunda para a ordem sazonal de  $Q$  e as demais para o parâmetro  $q$ . Assim foi possível propor vários modelos SARIMA  $(p,d,q) \times (P,D,Q)_s$ .

Figura 4 – Correlograma das funções de autocorrelação (FAC) e autocorrelação parcial (FACP) dos dados de treino com uma diferença de ordem 12.



Fonte: Elaborada pelo autor, 2023.

Conforme mencionado previamente, foi conduzida a estimação de um modelo *Prophet* para fins de comparação com o modelo SARIMA identificado. Observou-se que o modelo SARIMA(1,0,1)×(1,1,1)<sub>12</sub>, previamente identificado por Dantas, Santos e Olinda (2016) em seu estudo utilizando simulações de Monte Carlo para preenchimento de dados ausentes, apresentou o menor valor de Critério de Informação de Akaike. Em contrapartida, o modelo

SARIMA(2,0,1)×(3,1,1)<sub>12</sub> demonstrou possuir o menor valor de Raiz do Erro Quadrático Médio. O modelo *Prophet* considerado mais apropriado revelou uma REQM de 46,01.

Na Tabela 4 estão as estatísticas de AIC e REQM dos modelos SARIMA e *Prophet* estimados:

Tabela 4 – Métricas de avaliação dos modelos SARIMA e *Prophet* usados para validação através do AIC e REQM

Modelo	AIC	REQM
SARIMA(1,0,1)×(1,1,1) <sub>12</sub>	2394,102	48,8
SARIMA(2,0,1)×(3,1,1) <sub>12</sub>	2398,825	46,52
<i>Prophet</i>	–	46,01

Fonte: Elaborada pelo autor, 2023.

Na tabela 5 apresenta-se os valores dos hiperparâmetros mais adequados encontrados para o uso do modelo do *Prophet*:

Tabela 5 – Valores dos hiperparâmetros do modelo do *Prophet* mais adequados para realização de previsão de precipitação pluviométrica na cidade de Campina Grande - PB

Hiperparâmetro	Valores
change point prior scale	0,5
change point range	0,9
seasonality prior scale	1
seasonality mode	multiplicativo

Fonte: Elaborada pelo autor, 2023.

Devemos nos atentar ao detalhe de que o hiperparâmetro *seasonality mode* como multiplicativo foi o mais eficiente, evidenciando que a série tem um comportamento majoritariamente multiplicativo, em outras palavras, uma série que a variância de sua sazonalidade não é constante, corroborando com o resultado do gráfico da FAC para a série sem a diferenciação nas defasagens múltiplas de 6 e com o teste de Kruskal-Wallis.

Prosseguiu-se então para a aplicação dos modelos, mas agora usando todas as observações da série, desde jan/2002 até maio/2023, com o intuito de fazer as previsões da precipitação total mensal em milímetros na cidade de Campina Grande – PB no horizonte de maio/2023 a maio/2025. A quantidade de parâmetros e hiper parâmetros usados para os novos modelos SARIMA e *Prophet* foram os mesmos para os obtidos com as metodologias usadas anteriormente. Teremos então um modelo para previsão usando o SARIMA e um modelo para previsão usando o *Prophet*.

A tabela 6 apresenta as estimativas dos parâmetros do modelo SARIMA ajustado para descrever a série temporal, obtidos pelo método de otimização de máxima verossimilhança e os seus respectivos erros padrão.

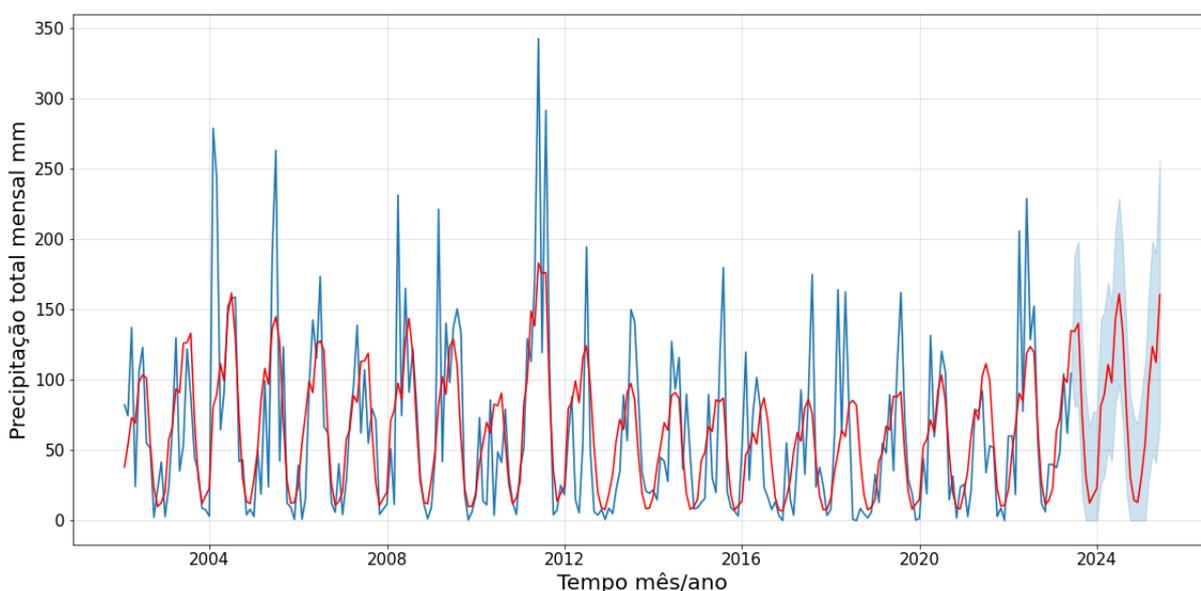
Tabela 6 – Estimativas dos parâmetros do modelo SARIMA(2,0,1)×(3,1,1)<sub>12</sub> ajustado para descrever e prever a precipitação total mensal (mm) na cidade de Campina Grande – PB de jan/2002 a maio/2025.

Parâmetro	Coefficiente	Erro Padrão
$\phi_1$	0.3761	0.0129
$\phi_2$	0.0959	0.0303
$\theta_1$	-0.2519	0.1217
$\Phi_1$	-0.0632	0.0293
$\Phi_2$	0.1253	0.0712
$\Phi_3$	0.0981	0.0368
$\Theta_1$	-0.9658	0.0879

Fonte: Elaborada pelo autor, 2023.

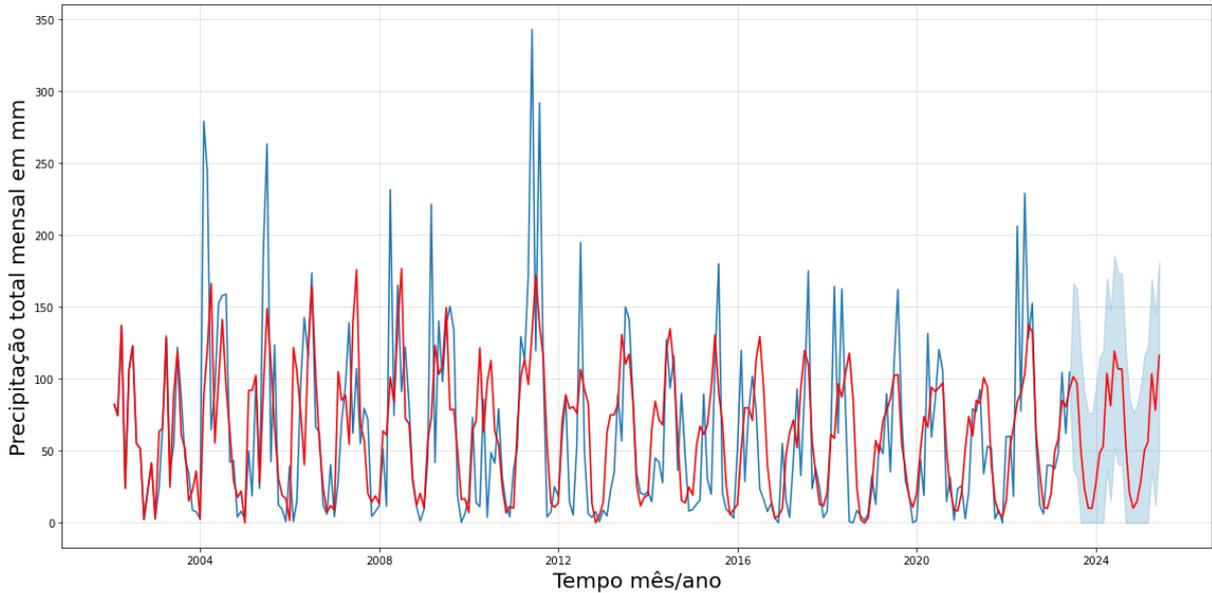
As Figuras 5 e 6 ilustram os gráficos dos modelos *Prophet* e SARIMA ajustados aos dados, respectivamente. Neles estão contidos a série de precipitação total mensal (mm) em Campina Grande - PB (em azul), o ajuste do modelo aos dados (em vermelho) e os intervalos com 80% de confiança para os meses de previsão de jun/2023 a maio/2025 (em azul claro). Observa-se que ambos captaram de forma satisfatória a sazonalidade presente na série.

Figura 5 – Valores reais (azul), valores calibrados (vermelho) e intervalo com 80% de confiança em relação a previsão (azul claro) da precipitação total mensal (mm) em Campina Grande - PB de jan/2002 a maio/2025 para o modelo *Prophet*.



Fonte: Elaborada pelo autor, 2023.

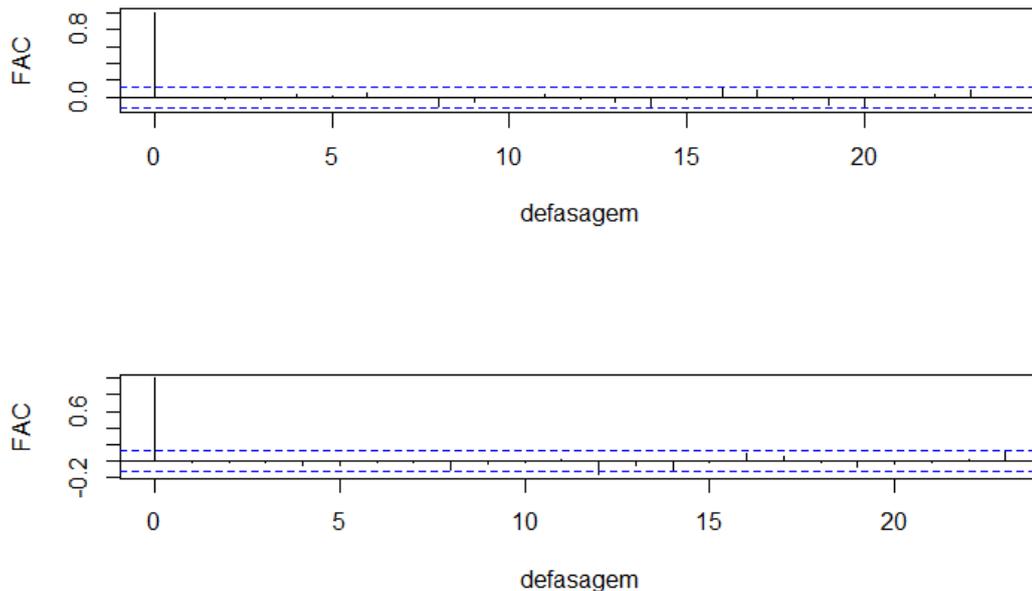
Figura 6 – Valores reais (azul), valores calibrados (vermelho) e intervalo com 80% de confiança em relação a previsão (azul claro) da precipitação total mensal (mm) em Campina Grande - PB de jan/2002 a maio/2025 para o modelo SARIMA(2,0,1)×(3,1,1)<sub>12</sub>.



Fonte: Elaborada pelo autor, 2023.

Na Figura 7 podemos ver que os resíduos não são autocorrelacionados de acordo com gráfico da função de autocorrelação FAC para ambos os modelos, corroborando com os resultados do teste de Ljung-Box que nos apontou que os resíduos não são autocorrelacionados.

Figura 7 – Função de autocorrelação FAC para os resíduos do modelo SARIMA e *Prophet*, respectivamente.



Fonte: Elaborada pelo autor, 2023.

A tabela 7 apresenta os valores previstos pelo modelo do *Prophet* e SARIMA e seus

intervalos de 80% de confiança para a precipitação total mensal (mm) na cidade de Campina Grande - PB. É interessante notar todas as previsões realizadas pelo *Prophet* foram maiores do que as realizadas pelo SARIMA, exceto para o mês de novembro de 2024. Isto se deve ao fato do último ponto de mudança de tendência encontrado pelo modelo ser positivo, e também pelo uso de séries de Fourier para modelagem da sazonalidade junto com o modo multiplicativo, fazendo com que as previsões apresentem picos sazonais. Wang et al. (2021) encontraram problemas semelhantes usando o *Prophet* em suas previsões para consumo diário de energia elétrica em nível urbano de três áreas metropolitanas nos Estados Unidos: Sacramento, Los Angeles e Nova York.

Tabela 7 – previsão de precipitação pluviométrica em mm na cidade de Campina Grande - PB com o uso dos modelos *Prophet* e SARIMA(2,0,1)×(3,1,1)<sub>12</sub>.

Data	Previsto <i>Prophet</i>	Previsto SARIMA
Junho/2023	134,45 [76,88; 189,44]	101,57 [36,5; 166,65]
Julho/2023	140,27 [85,75; 197,86]	97,1 [31,53; 162,68]
Agosto/2023	74,48 [22,13; 131,14]	50,98 [0; 117,21]
Setembro/2023	31,50 [0; 89,70]	23,63 [0; 90]
Outubro/2023	12,30 [0; 65,28]	10,22 [0; 76,63]
Novembro/2023	17,82 [0; 73,64]	10,07 [0; 76,49]
Dezembro/2023	23,10 [0; 82,29]	26,79 [0; 93,22]
Janeiro/2024	81,36 [23,05; 139,92]	48,31 [0; 114,69]
Fevereiro/2024	89,43 [32,87; 145,20]	53,14 [0; 119,52]
Março/2024	110,92 [52,25; 168,73]	103,89 [37,5; 170,27]
Abril/2024	98,03 [37,85; 156,39]	81,29 [14,9; 147,67]
Mai/2024	143,26 [81,81; 201,70]	119,27 [52,89; 185,65]
Junho/2024	161,15 [96,29; 226,73]	106,89 [40,5; 173,27]
Julho/2024	132,46 [67,61; 195,65]	106,96 [40,58; 173,35]
Agosto/2024	76,19 [16,25; 134,51]	55 [0; 121,38]
Setembro/2024	31,01 [0; 86,62]	21,04 [0; 87,43]
Outubro/2024	14,42 [0; 73,36]	10,21 [0; 76,6]
Novembro/2024	12,89 [0; 69,14]	14,59 [0; 80,98]
Dezembro/2024	31,49 [0; 88,40]	28,4 [0; 94,78]
Janeiro/2025	55,16 [0; 119,54]	50,41 [0; 116,75]
Fevereiro/2025	95,65 [32,90; 167,20]	56,28 [0; 122,62]
Março/2025	123,79 [51,21; 207,09]	103,62 [37,27; 169,96]
Abril/2025	112,39 [43,27; 190,69]	78,15 [11,81; 144,49]
Mai/2025	160,52 [74,84; 255,13]	116,15 [49,81; 182,49]

Fonte: Elaborada pelo autor, 2023.

## 4 CONCLUSÃO

Foi evidenciado que a série temporal da precipitação pluviométrica em Campina Grande - PB apresenta uma marcante componente de sazonalidade, identificada por meio do teste de Kruskal-Wallis para sazonalidade. Além disso, foi observado que o modelo *Prophet* teve um desempenho superior quando o hiperparâmetro *seasonality mode* foi configurado como "multiplicativo". Essa constatação ganha relevância ao considerar a importância de aproveitar da melhor forma possível os períodos do ano com maior ou menor incidência de chuva, especialmente em suas respectivas áreas de estudo.

Ambos os modelos, *Prophet* e SARIMA, demonstraram uma excelente adaptação aos dados, evidenciando uma boa capacidade de capturar a sazonalidade presente na série temporal. Ao compará-los pela Raiz do Erro Quadrático Médio (REQM) usando os dados de treino, notou-se que o *Prophet* teve um desempenho ligeiramente melhor, alcançando um valor de 46,01, enquanto o SARIMA(2,0,1)×(3,1,1)<sub>12</sub> registrou um valor de 46,52, ou seja, em um cenário em que não tivéssemos os dados de treino o *Prophet* erraria menos. Adicionalmente, observou-se que os meses de setembro, outubro, novembro e dezembro apresentaram as menores previsões para os anos de 2023 a 2025, em ambos os modelos. Por outro lado, os meses de maio, junho e julho foram identificados com previsões mais elevadas, também para ambos os modelos. Estas conclusões específicas estão detalhadas na Tabela 7.

## REFERÊNCIAS

- BOX, G. E.; JENKINS, G. M. Time series analysis: Forecasting and control san francisco. *Calif: Holden-Day*, 1976. Citado na página 12.
- BOX, G. E. et al. *Time series analysis: forecasting and control*. [S.l.]: John Wiley & Sons, 2015. Citado na página 15.
- DANTAS, L. G.; SANTOS, C. A. C.; OLINDA, R. A. *Aplicação de modelos para séries temporais e pluviométricas no Estado da Paraíba*. Universidade Federal de Campina Grande, 2016. Disponível em: <<http://dspace.sti.ufcg.edu.br:8080/jspui/handle/riufcg/1390>>. Acesso em: 28 dez. 2022. Citado na página 23.
- DANTAS, L. G.; SANTOS, C. A. C. d.; OLINDA, R. A. d. Tendências anuais e sazonais nos extremos de temperatura do ar e precipitação em campina grande-pb. *Revista Brasileira de Meteorologia*, SciELO Brasil, v. 30, 2015. Citado na página 11.
- IBGE, I. B. de Geografia e E. *Cidades e Estados*. 2023. Disponível em: <<https://www.ibge.gov.br/cidades-e-estados/pb/campina-grande.html>>. Acesso em: 15 nov. 2023. Citado na página 11.
- JANA, R. K.; GHOSH, I.; WALLIN, M. W. Taming energy and electronic waste generation in bitcoin mining: Insights from facebook prophet and deep neural network. *Technological Forecasting and Social Change*, v. 178, p. 4, 2022. Citado na página 14.
- KRUSKAL, W. H.; WALLIS, W. A. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, [American Statistical Association, Taylor Francis, Ltd.], v. 47, n. 260, p. 583–621, 1952. ISSN 01621459. Disponível em: <<http://www.jstor.org/stable/2280779>>. Citado na página 13.
- LJUNG, G. M.; BOX, G. E. On a measure of lack of fit in time series models. *Biometrika*, v. 65, n. 2, 1978. Citado na página 18.
- MANN, H. B. Nonparametric tests against trend. *Econometrica*, [Wiley, Econometric Society], v. 13, n. 3, p. 245–259, 1945. ISSN 00129682, 14680262. Disponível em: <<http://www.jstor.org/stable/1907187>>. Citado na página 13.
- MAPA, M. da Agricultura e P. *Instituto Nacional de Meteorologia*. 2023. Disponível em: <<https://portal.inmet.gov.br/>>. Acesso em: 12 jun. 2023. Citado na página 12.
- MARCUZZO, F. F. N.; ROMERO, V. Influência do el niño e la niña na precipitação máxima diária do estado de goiás. *Revista Brasileira de Meteorologia*, SciELO Brasil, v. 28, 2013. Citado na página 11.
- MENEZES, H. E. A. et al. A relação entre a temperatura da superfície dos oceanos tropicais e a duração dos veranicos no estado da paraíba. *Revista Brasileira de Meteorologia*, SciELO Brasil, v. 23, 2008. Citado na página 11.
- MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. *Introduction to time series analysis and forecasting*. [S.l.]: John Wiley & Sons, 2015. Citado na página 16.
- MORETTIN, P.; TOLOI, C. Análise de séries temporais—2ª edição revista e ampliada. *ABE–Projeto Fisher*, Editora Edgar Blücher, 2006. Citado 3 vezes nas páginas 12, 13 e 17.

OLLECH, D.; WEBEL, K. A random forest-based approach to identifying the most informative seasonality tests. *Deutsche Bundesbank Discussion Paper*, 2020. Citado na página 14.

SANTOS, C. A. C. d.; BRITO, J. I. B. d. Análise dos índices de extremos para o semi-árido do Brasil e suas relações com tsm e ivdn. *Revista Brasileira de Meteorologia*, SciELO Brasil, v. 22, 2007. Citado na página 11.

SHUMWAY, R. H.; STOFFER, D. S. *Time series analysis and its applications*. [S.l.]: Springer, 2000. v. 3. Citado 2 vezes nas páginas 15 e 17.

SILVA, M. dos S. et al. Análise da energia eólica no Brasil usando séries temporais. *Research, Society and Development*, v. 11, n. 1, 2022. Citado na página 17.

SNEYERS, R. *Sur l'analyse statistique des séries d'observations*. [S.l.]: Secrétariat de l'Organisation météorologique mondiale, 1975. Citado na página 13.

TAYLOR, S. J.; LETHAM, B. Forecasting at scale. *PeerJ Preprints*, 2017. Citado na página 14.

TAYLOR, S. J.; LETHAM, B. *Forecasting at scale*. 2019. Disponível em: <<https://facebook.github.io/prophet/>>. Acesso em: 20 nov. 2022. Citado na página 15.

WANG, Z. et al. Predicting city-scale daily electricity consumption using data-driven models. *Advances in Applied Energy*, Elsevier, v. 2, 2021. Citado na página 27.