



PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ
ESCOLA POLITÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO E
SISTEMAS (PPGEPS)

FRANCESCO SANCHES DOS SANTOS

MODELOS DE PREVISÃO DE SÉRIES TEMPORAIS APLICADOS A UM CASO
DE SISTEMA DE ABASTECIMENTO DE ÁGUA

CURITIBA

2023

FRANCESCO SANCHES DOS SANTOS

MODELOS DE PREVISÃO DE SÉRIES TEMPORAIS APLICADOS A UM CASO
DE SISTEMA DE ABASTECIMENTO DE ÁGUA

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia de Produção e Sistemas (PPGEPS). Área de concentração: Gerência de Produção e Logística, da Escola Politécnica, da Pontifícia Universidade Católica do Paraná, como requisito parcial à obtenção do título de Mestre em Engenharia de Produção e Sistemas.

Orientador: Dr. Leandro dos Santos Coelho
Coorientador: Dra. Viviana Cocco Mariani
(PPGEM-PUCPR)

CURITIBA

2023

FRANCESCO SANCHES DOS SANTOS

**MODELOS DE PREVISÃO DE SÉRIES TEMPORAIS APLICADOS A
UM CASO DE SISTEMA DE ABASTECIMENTO DE ÁGUA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção e Sistemas (PPGEPS). Área de concentração: Gerência de Produção e Logística, da Escola Politécnica, da Pontifícia Universidade Católica do Paraná, como requisito parcial à obtenção do título de Mestre em Engenharia de Produção e Sistemas.

COMISSÃO EXAMINADORA

Dr. Leandro dos Santos Coelho
Orientador

Pontifícia Universidade Católica do Paraná

Dra. Viviana Cocco Mariani (PPGEM-PUCPR)
Coorientadora

Pontifícia Universidade Católica do Paraná

Dr. Gilberto Reynoso Meza
Membro Interno

Pontifícia Universidade Católica do Paraná

Dr. Matheus Henrique Dal Molin Ribeiro
Membro Externo
Universidade Tecnológica Federal do Paraná

Curitiba, 4 de dezembro de 2023

*Dedico essa dissertação de mestrado à Deus, essa força maior, que me guia e ilumina meus
pensamentos para que eu desenvolva minha luz.*

Agradecimentos

Primeiramente, manifesto minha sincera gratidão a Deus por todas as bênçãos que recebi ao longo desta jornada. Foi Ele quem, com Sua infinita bondade, iluminou meu caminho, abriu portas e me conferiu a força necessária para superar os desafios, tornando possível a realização deste sonho.

Aos meus pais e irmãos, pilares fundamentais em minha vida, expresso meu profundo reconhecimento e eterno agradecimento. Com amor incondicional e apoio constante, vocês estiveram ao meu lado, incentivando-me a seguir em frente com determinação e a sempre almejar novos horizontes. Foi por meio do apoio e da dedicação de vocês que encontrei a motivação e a coragem necessárias para enfrentar os obstáculos e concretizar meus objetivos. Sou-lhes extremamente grato por tudo que fizeram e continuam fazendo por mim.

Expresso também minha sincera gratidão ao estimado professor doutor Leandro dos Santos Coelho pela valiosa oportunidade de colaborar em sua equipe durante o meu período de mestrado. Sua orientação cuidadosa e compartilhamento de conhecimentos foram de importância vital para o meu desenvolvimento pessoal e profissional, contribuindo significativamente para a concretização deste projeto.

À professora doutora Viviana Cocco Mariani, desejo expressar minha profunda gratidão pela disponibilidade e paciência demonstradas em me auxiliar nas adversidades, contribuindo de forma decisiva para o avanço de minha pesquisa.

Gostaria de estender meus agradecimentos à equipe da Pontifícia Universidade Católica do Paraná (PUCPR) e a todos os demais professores, com destaque para a dedicada secretária Denise da Mata Medeiros (PPGEPS), pela generosidade, carinho e suporte oferecidos ao longo desta jornada acadêmica.

Aos amigos, tanto os de longa data quanto os que conheci durante este desafiador percurso, manifesto meu mais profundo reconhecimento. Suas palavras de estímulo, apoio e os momentos de alegria compartilhados foram de valor inestimável e imprescindíveis para o meu progresso.

Expresso minha imensa gratidão ao investimento em bolsas de estudo concedidas pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), cujo apoio foi crucial para a realização deste sonho e para a conclusão desta etapa fundamental de minha trajetória acadêmica e profissional.

Por fim, a vocês, respeitáveis professores doutores Marcos Vinicius de Oliveira Peres e Rafael Mestrinheire Hungaro, bem como à estimada Thais Carreira Pfitzenreuter, manifesto minha profunda gratidão. Seus incentivos e presenças foram verdadeiros sustentáculos ao longo de minha jornada, oferecendo-me inspiração, apoio e orientação em cada etapa do percurso.

*Se vi mais longe, foi por estar de pé
sobre ombros de gigantes*

- Sir Isaac Newton

Resumo

Esta dissertação está inserida no contexto do abastecimento de água na cidade de Curitiba e concentra-se na análise de eficácia da previsão da demanda no Bairro Alto, por meio dos dados coletados pela Companhia de Saneamento do Paraná durante os anos de 2018 a 2020. A questão central investigada é a eficácia das previsões de abastecimento de água, visando assegurar que a infraestrutura existente seja capaz de atender às crescentes necessidades da população, prevenindo problemas de oferta inadequada. O propósito deste estudo é contribuir para o controle eficaz dos recursos hídricos, utilizando modelos de previsão, com ênfase na melhoria do abastecimento d'água. São explorados modelos de previsão, tais como *Auto-Regressive* (AR), *Auto-Regressive with Exogenous Inputs*, *Moving Average*, *Auto-Regressive Moving Average*, *Seasonal Auto-Regressive Integrated Moving Average*, ARIMA com *Exogenous Inputs*, e *Seasonal ARIMA* com *Exogenous Inputs*, *Decision Tree* (DT), *eXtreme Gradient Boosting* e *Recurrent Neural Network* para a previsão de séries temporais, com análise comparativa de eficácia dos modelos de previsão. O desempenho dos modelos de previsão é avaliado por meio de métricas que incluem *Symmetric Mean Absolute Percentage Error* (SMAPE), *Mean Absolute Error* (MAE) e *Root Relative Mean Square Error* (RRMSE). Nota-se que, o modelo DT se apresentou melhor nas previsões de curto prazo, destacando-se com resultados de SMAPE igual a 13,50, MAE de 0,58 e RRMSE de 0,16 para a previsão de 6 horas à frente. Com 1 hora à frente, o resultado foi de SMAPE igual a 7,83, MAE de 0,36 e RRMSE de 0,20, ficando atrás apenas dos valores de MAE e RRMSE do modelo AR, que apresentou os valores mais baixos nesse horizonte de previsão. No entanto, o modelo DT, no geral, até a previsão de 24 horas ou um dia à frente, se mostrou superior aos outros modelos. Na previsão de um dia à frente, o Prophet se destacou, apresentando resultados de SMAPE igual a 5,05, MAE de 0,17 e RRMSE com valor de 0,19. Conclui-se que a abordagem proposta contribui para a previsão da demanda d'água, proporcionando um planejamento eficiente e sustentável do abastecimento hídrico no Bairro Alto. Além disso, a previsão permite antecipar e prevenir possíveis escassezes d'água, prevendo a demanda futura e possibilitando a adoção de medidas proativas da Companhia de Saneamento do Paraná para evitar interrupções no fornecimento.

Palavras-chave: Previsão, Séries Temporais, Abastecimento de Água, Aprendizado de Máquina, Redes Neurais Artificiais.

Abstract

This dissertation is inserted in the context of water supply in the city of Curitiba and focuses on the analysis of the effectiveness of demand forecasting in the Bairro Alto, using data collected by the Sanitation Company of Paraná during the years 2018 to 2020. The central issue investigated is the effectiveness of water supply forecasts, aiming to ensure that the existing infrastructure is capable of meeting the growing needs of the population, preventing inadequate supply problems. The purpose of this study is to contribute to the effective control of water resources, using forecasting models, with emphasis on improving water supply. Forecasting models are explored, such as Auto-Regressive (AR), Auto-Regressive with Exogenous Inputs, Moving Average, Auto-Regressive Moving Average, Seasonal Auto-Regressive Integrated Moving Average, ARIMA with Exogenous Inputs, and Seasonal ARIMA with Exogenous Inputs, Decision Tree (DT), eXtreme Gradient Boosting, and Recurrent Neural Network for time series forecasting, with comparative analysis of forecast model effectiveness. The performance of the forecasting models is evaluated through metrics that include Symmetric Mean Absolute Percentage Error (SMAPE), Mean Absolute Error (MAE), and Root Relative Mean Square Error (RRMSE). It is noted that the DT model performed better in short-term forecasts, standing out with SMAPE results of 13.50, MAE of 0.58, and RRMSE of 0.16 for the forecast of 6 hours ahead. With 1 hour ahead, the result was SMAPE of 7.83, MAE of 0.36, and RRMSE of 0.20, ranking behind only the MAE and RRMSE values of the AR model, which presented the lowest values in this forecast horizon. However, the DT model, overall, up to the forecast of 24 hours or one day ahead, proved to be superior to the other models. In the forecast of one day ahead, Prophet stood out, presenting SMAPE results of 5.05, MAE of 0.17, and RRMSE with a value of 0.19. It is concluded that the proposed approach contributes to the prediction of water demand, providing efficient and sustainable planning of water supply in Bairro Alto. In addition, the forecast allows anticipating and preventing possible water shortages, predicting future demand and enabling the adoption of proactive measures by the Sanitation Company of Paraná to avoid supply interruptions.

Keywords: Forecasting, Time series, Water supply, Machine learning, Artificial neural networks.

Lista de Figuras

| | | |
|----|--|----|
| 1 | Etapas para análise dos dados. | 21 |
| 2 | Fluxograma da Revisão Sistemática da Literatura. | 24 |
| 3 | Modelos de previsão de series temporais na base de dados Scopus e WoS. | 26 |
| 4 | Elementos do modelo SARIMAX. | 39 |
| 5 | Fluxograma da árvore de decisão. | 41 |
| 6 | Fluxograma da floresta aleatória. | 43 |
| 7 | Fluxograma do XGBoost. | 44 |
| 8 | Comparação do crescimento em folha com o crescimento em nível. | 46 |
| 9 | Modelo de uma rede neural artificial MLP. | 48 |
| 10 | Fluxograma da RNN. | 49 |
| 11 | Diagrama do funcionamento de uma GRU. | 52 |
| 12 | Diferenças entre RNN, LSTM, e GRU. | 53 |
| 13 | Modelo de uma Rede Neural Convolucional. | 54 |
| 14 | Decomposição STL aditiva. | 66 |
| 15 | Correlação de Pearson. | 67 |
| 16 | Relação entre LT01 e PT01 cuja correlação de Pearson é 97%. | 67 |
| 17 | Autocorrelação. | 68 |
| 18 | Autocorrelação parcial. | 69 |
| 19 | Ruído branco. | 69 |
| 20 | Comparação dos modelos de previsão AR, ARX, MA e ARMA 24 passo à frente. | 73 |
| 21 | Comparação dos modelos de previsão ARIMA, ARIMAX, SARIMA e SARIMAX 24 passo à frente. | 73 |
| 22 | Comparação dos modelos DT, RF, XGBoost e LightGBM 1 passo à frente. | 74 |
| 23 | Modelo de previsão GRU para vários horizontes de previsão. | 74 |
| 24 | Modelo de previsão LSTM para vários horizontes de previsão. | 74 |
| 25 | Modelo de previsão RNN para vários horizontes de previsão. | 75 |
| 26 | Previsões do modelo Prophet para vários horizontes de previsão. | 75 |
| 27 | Comparação dos modelos AR, ARX, MA, ARIMA, SARIMA, ARIMAX, SARIMAX, DT, RF, XGBoost, LightGBM em vários horizontes na métrica SMAPE. | 86 |
| 28 | Comparação dos modelos AR, ARX, MA, ARIMA, SARIMA, ARIMAX, SARIMAX, DT, RF, XGBoost, LightGBM em vários horizontes na métrica MAE. | 86 |

| | | |
|----|--|----|
| 29 | Comparação dos modelos AR, ARX, MA, ARIMA, SARIMA, ARIMAX, SARIMAX, DT, RF, XGBoost, LightGBM em vários horizontes na métrica RRMSE. | 87 |
| 30 | Comparação dos modelos Prophet, MLP, CNN, RNN, LSTM, GRU em vários horizontes na métrica SMAPE. | 87 |
| 31 | Comparação dos modelos Prophet, MLP, CNN, RNN, LSTM, GRU em vários horizontes na métrica MAE. | 87 |
| 32 | Comparação dos modelos Prophet, MLP, CNN, RNN, LSTM, GRU em vários horizontes na métrica RRMSE. | 88 |
| 33 | Demanda média das variáveis de fluxo. | 90 |

Lista de Tabelas

| | | |
|----|---|----|
| 1 | Combinação de palavras-chave aplicando filtros. | 27 |
| 2 | Resumo dos artigos obtidos com a RSL nas bases Scopus e WoS. | 27 |
| 3 | Classificação dos principais periódicos obtidos na RSL. | 27 |
| 4 | Total de publicações dos principais autores obtidos na RSL. | 29 |
| 5 | Total de publicações dos principais países obtidos na RSL. | 30 |
| 6 | Principais modelos de previsão obtidos na RSL. | 32 |
| 7 | Descrição estatística dos dados do Bairro Alto em Curitiba de 2018 a 2019 disponibilizados pela SANEPAR. | 65 |
| 8 | Teste ADF. | 70 |
| 9 | Parâmetros para os modelos ARIMA utilizando a função autoARIMA. | 71 |
| 10 | Hiperparâmetros otimizados dos modelos. | 72 |
| 11 | Hiperparâmetros otimizados para RNA. | 72 |
| 12 | Comparação dos modelos de previsão através das métricas de desempenho para dados de treino. | 76 |
| 13 | Comparação dos modelos de previsão através das métricas de desempenho para dados de validação. | 77 |
| 14 | Comparação dos modelos de previsão através das métricas de desempenho para dados de teste. | 78 |
| 15 | Comparação dos modelos de previsão através das métricas de desempenho para todos dados. | 79 |
| 16 | Métricas de avaliação dos modelos com 24 passos à frente. | 80 |
| 17 | Teste de significância Nemenyi na métrica SMAPE. | 82 |
| 18 | Teste de significância Nemenyi na métrica MAE. | 83 |
| 19 | Teste de significância Nemenyi na métrica RRMSE. | 84 |
| 20 | Comparação dos modelos com o teste Ljung Box modelos ARIMA com defasagem de 10 para previsão de longo prazo na demanda d'água. | 89 |
| 21 | Demanda d'água. | 91 |

Lista de Abreviaturas e Siglas

| | |
|----------|--|
| ACF | <i>Autocorrelation Function</i> |
| ADY | <i>Average Documents per Year</i> |
| AGR | <i>Average Growth Rate</i> |
| ANN | <i>Artificial Neural Network</i> |
| AR | <i>Auto-Regressive</i> |
| ARIMA | <i>Auto-Regressive Integrated Moving Average</i> |
| ARIMAX | <i>Auto-Regressive Integrated Moving Average with eXogenous inputs</i> |
| ARMA | <i>Auto-Regressive Moving Average</i> |
| ARX | <i>Auto-Regressive with Exogenous Inputs</i> |
| BOA | <i>Bayesian Optimization Algorithm</i> |
| CART | <i>Classification And Regression Trees</i> |
| CMAES | <i>Covariance Matrix Adaptation Evolution Strategy</i> |
| CNN | <i>Convolutional Neural Networks</i> |
| DBN | <i>Dynamic Bayesian Network</i> |
| DT | <i>Decision Tree</i> |
| EFB | textitExclusive Feature Bundling |
| ERNN | <i>Elman Recurrent Neural Network</i> |
| ETS | <i>Exponential Smoothing</i> |
| FI | Fator de Impacto |
| FT | <i>Flow Transmitter</i> |
| GP | <i>Genetic Programming</i> |
| GARCH | <i>Generalized Autoregressive Conditional Heteroskedasticity</i> |
| GOSS | <i>Gradient-Based One-Side Sample</i> |
| GRU | <i>Gated Recurrent Unit</i> |
| HMM | <i>Hidden Markov Model</i> |
| INMET | Instituto Nacional de Meteorologia |
| LightGBM | <i>Light Gradient Boosting Machine</i> |
| LR | <i>Linear Regression</i> |
| LSTM | <i>Long Short-Term Memory</i> |
| MA | <i>Moving Average</i> |
| MAE | <i>Mean Absolute Error</i> |

| | |
|------------|--|
| MAPE | <i>Mean Absolute Percentage Error</i> |
| MIMO | <i>Multiple Inputs and Multiple Outputs</i> |
| MISO | <i>Multiple Inputs Single Output</i> |
| ML | <i>Machine Learning</i> |
| MSE | <i>Mean Squared Error</i> |
| PACF | <i>Partial Autocorrelation Function</i> |
| PDLY | <i>Percentage of Documents in Last Years</i> |
| PR | Estado do Paraná |
| RART | Reutilização de Águas Residuais Tratadas |
| RBAL | Recalque do Bairro Alto |
| RF | <i>Random Forest</i> |
| RMSE | <i>Root Mean Squared Error</i> |
| RNA | Rede Neural Artificial |
| RNN | <i>Recurrent Neural Network</i> |
| RRMSE | <i>Root of the Relative Mean Square Error</i> |
| RSL | Revisão Sistemática da Literatura |
| SANEPAR | Companhia de Saneamento do Paraná |
| SARIMA | <i>Seasonal Auto-Regressive Integrated Moving Averages</i> |
| SARIMAX | <i>Seasonal Auto-Regressive Integrated Moving Averages with Exogenous Inputs</i> |
| SJR | <i>Scimago Journal Rank</i> |
| SMAPE | <i>Symmetric Mean Absolute Percentage Error</i> |
| SVM | Support Vector Machines |
| SVM-VAR | Máquinas de Vetor de Suporte - Vetores Auto-Regressivos |
| TCN | <i>Temporal Convolutional Network</i> |
| TND | <i>Total Number of Documents</i> |
| TotalBoost | Impulso Total |
| WoS | <i>Web of Science</i> |
| XGBoost | <i>eXtreme Gradient Boosting</i> |

Sumário

| | | |
|----------|--|-----------|
| 1 | Introdução | 16 |
| 1.1 | Motivação | 19 |
| 1.2 | Objetivo Geral | 20 |
| 1.3 | Etapas para Análise dos Dados | 21 |
| 1.4 | Estrutura do Documento | 23 |
| 2 | Revisão da Literatura | 24 |
| 3 | Fundamentos dos Modelos de Previsão | 34 |
| 3.1 | Conceito de Séries Temporais | 34 |
| 3.2 | Estratégia Recursiva de Previsão Multipassos | 34 |
| 3.3 | Modelos Clássicos de Séries Temporais | 35 |
| 3.4 | Modelos de Aprendizado de Máquina | 39 |
| 3.4.1 | Prophet | 39 |
| 3.4.2 | Regressão Linear | 40 |
| 3.4.3 | Árvore de Decisão | 40 |
| 3.4.4 | Floresta Aleatória | 42 |
| 3.4.5 | <i>Gradient Boosting</i> | 44 |
| 3.4.6 | <i>LightGBM</i> | 45 |
| 3.5 | Redes Neurais Artificiais | 47 |
| 3.5.1 | MLP | 47 |
| 3.5.2 | Rede Neural Recorrente | 48 |
| 3.6 | Aprendizado Profundo | 49 |
| 3.6.1 | LSTM | 50 |
| 3.6.2 | GRU | 50 |
| 3.7 | Rede Neural Convolucional | 53 |
| 3.8 | Abordagem de Otimização | 54 |
| 3.9 | Decomposição STL | 55 |
| 3.10 | Abordagens de Correlações | 56 |
| 3.10.1 | Autocorrelação e Autocorrelação Parcial | 56 |
| 3.10.2 | Correlação de Pearson | 57 |
| 3.11 | Medidas de Desempenho | 58 |
| 3.12 | Testes de Hipóteses | 59 |
| 3.12.1 | Teste Run | 59 |
| 3.12.2 | Teste Dickey-Fuller | 60 |
| 3.12.3 | Teste de Ljung-Box | 61 |

| | | |
|----------|---|-----------|
| 3.12.4 | Testes de Friedman e Nemenyi | 62 |
| 4 | Resultados | 64 |
| 4.1 | Análise Exploratória dos Dados | 64 |
| 4.2 | Aplicando os Modelos de Previsão | 72 |
| 4.3 | Teste de Hipótese | 80 |
| 4.3.1 | Comparação dos Modelos | 85 |
| 4.4 | Aplicação | 89 |
| 4.4.1 | Estudo de Caso 1 - Análise de Padrões de Consumo | 90 |
| 4.4.2 | Estudo de Caso 2 – Estratégias de Economia de Energia | 90 |
| 5 | Conclusões | 92 |
| 5.1 | Propostas Futuras | 93 |
| | Referências | 94 |

1 Introdução

O acesso à água potável é vital para a saúde e bem-estar das pessoas, sendo um requisito fundamental para a sobrevivência e o desenvolvimento humano. A água potável é essencial para a higiene pessoal, a preparação de alimentos e a prevenção de doenças transmitidas pela água. Além disso, é um componente crucial para o funcionamento adequado de sistemas de saneamento básico. O acesso a água limpa não apenas reduz significativamente a incidência de doenças, mas também promove o crescimento econômico, a educação e a igualdade. Garantir o acesso universal à água potável não apenas salva vidas, mas também contribui para a construção de comunidades saudáveis e sustentáveis em todo o mundo. Os recursos d'água potável estão tornando-se escassos em algumas comunidades, em parte devido à crescente procura d'água potável em regiões urbanas, à pobreza de técnicas de gestão d'águas residuais e as secas ou enchentes induzidas pelo clima (KOEBELE et al., 2022).

Dada a crescente escassez d'água os órgãos que gerenciam tais recursos recomendam a implementação de novas iniciativas para expandir seus portfólios locais e regionais d'água, incluindo a reutilização d'águas residuais tratadas para fins potáveis para evitar a escassez d'água, questões que poderão piorar nas próximas décadas (BARNES; KRISHEN; HU, 2023). A reutilização d'águas residuais tratadas (RART) para fins potáveis ajuda a aliviar o uso das águas superficiais e subterrâneas locais que estão relacionadas com preocupações de escassez. Ao reutilizar águas residuais, as comunidades tornam-se menos dependentes dos recursos hídricos disponíveis localmente e operam de forma mais local, girando a economia circular da água centrada (TSATSOU; FRANTZESKAKI; MALAMIS, 2023). Algumas cidades conseguiram implementar este sistema RART para aumentar o abastecimento local d'água potável, porém no Brasil ainda não temos isto de forma efetiva, então uma possibilidade é fazer previsão da demanda de tal forma que o abastecimento para a população não seja descontinuado.

O presente estudo envolve dados de abastecimento d'água da cidade de Curitiba no Bairro Alto entre os anos de 2018 e 2020. No entanto, vale destacar, que durante o ano de 2022, os habitantes de Curitiba enfrentaram escassez d'água, sendo necessário implementar rodízios, alternando períodos com e sem fornecimento d'água potável. Os dados utilizados foram coletados pela Companhia de Saneamento do Paraná (SANEPAR). A previsão da demanda d'água ao longo do tempo, o que será abordado neste estudo, é essencial para um planejamento sustentável e eficiente do abastecimento hídrico, especialmente no contexto urbano, como é o caso da cidade de Curitiba, capital do estado do Paraná.

Existem vários modelos que podem ser utilizados para prever a demanda d'água, e a

escolha da abordagem dependerá da disponibilidade de dados, da complexidade do sistema e das necessidades específicas da aplicação. Entre os modelos estão: modelos estatísticos, tal como regressão linear, que podem ser usados quando há uma relação linear entre variáveis como temperatura, população, atividade econômica e consumo d'água, modelos tais como *Seasonal Auto-Regressive Integrated Moving Averages* (SARIMA) que podem ser eficazes para prever padrões sazonais e tendências ao longo do tempo (OLIVEIRA; STEFFEN; CHEUNG, 2017), as Redes Neurais Artificiais incluindo Redes Neurais Recorrentes (RNN) (ASEERI, 2023) e *Long Short-Term Memory* (LSTM) (SABZIPOUR et al., 2023) são eficazes para lidar com dados temporais e sequenciais que são mais comuns em aprendizado profundo, *deep learning* (DL), capturando dependências de longo prazo e os métodos de aprendizado de máquina, tais como Máquinas de Vetores de Suporte (SVM), que podem ser usados para processar dados não lineares e complexos (CANDELLIERI et al., 2019), *Random Forest* (RF) (ALI et al., 2023) *Gradient Boosting* (DONG et al., 2023), que são métodos de *ensemble learning* (métodos de aprendizado de comitês) que podem ser aplicados para prever padrões complexos, entre outros modelos.

Séries temporais são comumente tratadas no âmbito de aprendizado na área de Aprendizado de Máquina Supervisionado. Existem três tipos principais de Aprendizado de Máquina: aprendizado supervisionado, aprendizado não-supervisionado e aprendizado por reforço. No aprendizado supervisionado (LIU; FU, 2023), o modelo é treinado em um conjunto de dados rotulado, onde cada exemplo do conjunto de dados possui uma entrada e a saída desejada correspondente. O objetivo dessa forma de aprendizado é aprender uma função que mapeia as entradas para as saídas, permitindo ao modelo fazer previsões usando dados não processados na fase de treinamento. No aprendizado não-supervisionado (WANG et al., 2022), o modelo é treinado em um conjunto de dados não rotulado, e o sistema tenta aprender a estrutura e padrões presentes nos dados. O objetivo principal é explorar a estrutura intrínseca dos dados, identificando agrupamentos, associações ou padrões emergentes sem ter rótulos para orientar o processo. O aprendizado por reforço (CHEN et al., 2023) envolve um agente que interage com um ambiente dinâmico. O agente toma decisões sequenciais para alcançar um objetivo específico, recebendo realimentação na forma de recompensas ou penalidades. O objetivo é aprender uma política, ou seja, uma estratégia que maximize a recompensa cumulativa ao longo do tempo (SILVA; GOMES, 2021). Nesta dissertação, são avaliados modelos de aprendizado de máquina supervisionados, uma escolha apropriada para a concepção de modelos de previsão de séries temporais (UC-CASTILLO et al., 2023).

As séries temporais de abastecimento d'água potável geralmente exibem características específicas devido à natureza dinâmica e sazonal do consumo d'água, muitas vezes exibem padrões sazonais, com variações regulares ao longo do tempo. Isso pode ser in-

fluenciado por fatores como as estações do ano, dias da semana ou horas do dia. Podem haver tendências a longo prazo nas séries temporais, refletindo mudanças demográficas, crescimento urbano, desenvolvimento industrial ou outros fatores que afetam o consumo d'água ao longo do tempo (JI; AHN, 2023). A ocorrência de eventos anômalos, como vazamentos, interrupções no fornecimento devido a situações críticas tais como secas, enchentes, ou situações de emergência, pode ser evidenciada em picos ou quedas abruptas nas séries temporais. Também podem estar correlacionadas com fatores externos, tais como eventos climáticos (por exemplo, períodos de seca ou chuvas intensas) e feriados, impactando o comportamento do consumo (BERGLUND; SKARBEEK; KANTA, 2023).

O consumo d'água muitas vezes segue padrões diários e semanais previsíveis incluindo, picos de demanda durante o horário de pico diário e variações ao longo da semana (SIEGEL et al., 2020). Flutuações de curto prazo podem ocorrer devido a atividades específicas, eventos locais ou situações temporárias que afetam o consumo d'água em um período limitado. Mudanças na infraestrutura, como expansões urbanas, construção de novos empreendimentos ou implementação de políticas de conservação, podem ser refletidas nas séries temporais. Entender essas características é fundamental para o gerenciamento eficiente do abastecimento d'água, permitindo a implementação de estratégias proativas, otimização de recursos e resposta adequada a eventos inesperados. O uso de modelos de previsão e análise de séries temporais, isto é, modelagem preditiva, pode ser valioso para entender as dinâmicas complexas (UC-CASTILLO et al., 2023).

Por meio da utilização de métodos de séries temporais, neste estudo será realizada a previsão do nível do reservatório na estação de tratamento d'água no Bairro Alto em Curitiba, incorporando diversos modelos de previsão nesse processo. Dentre esses modelos, inclui-se os clássicos, tais como ARIMA e suas variantes, tais como *Auto-Regressive* (AR), *Auto-Regressive with Exogenous input* (ARX), *Moving Average* (MA), *Auto-Regressive Moving Average* (ARMA), *Seasonal Auto-Regressive Integrated Moving Average* (SARIMA), *Auto-Regressive Integrated Moving Average with Exogenous input* (ARIMAX), e *Seasonal Auto-Regressive Integrated Moving Average with Exogenous input* (SARIMAX). Também é utilizado o método de decomposição STL *Seasonal and Trend Decomposition Using locally estimated scatterplot smoothing* (Loess) para que se possa verificar a presença de tendência, sazonalidade e ruído nos modelos ARIMA. Se houver sazonalidade, podem ser utilizados os modelos SARIMA e SARIMAX. Sem o efeito sazonal, os modelos ficam melhor previstos com modelos mais simples, como AR, MA, ARMA e ARX. Além disso, são explorados modelos de aprendizado de máquina, como árvore de decisão, floresta aleatória, Prophet, *eXtreme Gradient Boosting* (XGBoost), *Light Gradient Boosting Machine* (LightGBM), e redes neurais artificiais, tais como LSTM, *Gated Recurrent Unit* (GRU) e *Convolutional Neural Network* (CNN). A diversidade de modelos

é utilizada buscando aprimorar a qualidade das previsões.

O estudo adota modelos consagrados na literatura, como GRU, LSTM, XGBoost, LightGBM, RNN e CNN, reconhecendo sua eficácia em diversas aplicações. A escolha destes modelos não se baseia em sua novidade, mas sim em sua comprovada robustez e desempenho, já demonstrados em diferentes contextos. Essa abordagem visa incorporar as melhores práticas já estabelecidas para aprimorar as previsões de demanda d'água, garantindo resultados confiáveis. Os modelos ARIMA e suas variantes foram aplicados nesta área, como demonstrado por (BUYUKSAHIN; ERTEKIN, 2019; BHANGU; SANDHU; SAPRA, 2022). Alguns outros modelos, apesar de suas vantagens, ainda não foram devidamente aplicados, como é o caso do modelo *Decision Tree* (DT) (FOUILLOY et al., 2018), que se mostrará em termos de métricas de desempenho aos demais avaliados nessa dissertação.

Torna-se evidente que a análise de séries temporais e previsões são ferramentas poderosas para apoiar o processo de tomada de decisão em curto, médio e longo prazo de previsão. Devido às não linearidades, sazonalidades e tendências que podem ocorrer, nos dados temporais de abastecimento d'água, o desenvolvimento de modelos de previsão eficientes torna-se uma tarefa desafiadora (RIBEIRO et al., 2021).

1.1 Motivação

A análise e previsão de séries temporais de abastecimento d'água é crucial por várias razões, pois isso permite uma gestão eficiente e sustentável dos recursos hídricos. Porém pode-se citar alguns motivos. A previsão de séries temporais ajuda a antecipar demandas futuras d'água, permitindo que os gestores planejem e desenvolvam infraestruturas adequadas para atender a essas demandas. Isso é vital para garantir que a água esteja disponível em quantidade suficiente para atender às necessidades da população. Fornece informação sobre os padrões sazonais e tendências de consumo d'água. Com essas informações, os gestores podem tomar decisões informadas sobre a alocação de recursos hídricos e implementar práticas de conservação. Permitem que as autoridades otimizem as operações de abastecimento d'água, ajustando a produção e a distribuição com base nas variações de demanda ao longo do tempo. Isso contribui para uma operação mais eficiente do sistema. Podem prever eventos climáticos extremos, tais como secas prolongadas ou inundações, para atuar em situações de emergência. A previsão de séries temporais ajudam a antecipar esses eventos, permitindo que medidas preventivas sejam tomadas para garantir a continuidade do abastecimento d'água.

Assim, uma gestão eficiente baseada em previsões precisas pode resultar em economia de recursos financeiros. Isso inclui evitar investimentos desnecessários em infra-

estrutura e garantir que os recursos sejam alocados de maneira eficaz. Ao compreender os padrões de consumo d'água, é possível implementar práticas que promovam a sustentabilidade ambiental, como a redução do desperdício d'água e a promoção de fontes alternativas e renováveis. Muitas áreas possuem regulamentações que exigem o monitoramento e relatório regular do abastecimento d'água. O estudo e previsão de séries temporais auxiliam as autoridades a cumprir essas normativas de maneira eficaz. A capacidade de prever variações nas condições de abastecimento d'água permite uma melhor gestão de riscos, tanto em termos de disponibilidade d'água quanto de eventos que possam impactar negativamente a infraestrutura.

Em especial, a situação enfrentada por Curitiba e região metropolitana em 2020, conforme destacado por (VASCONCELOS, 2020) referente ao rodízio de abastecimento d'água, com períodos de 36 horas com abastecimento d'água, seguidos por 36 horas sem abastecimento mostra a necessidade de previsões mais acuradas do abastecimento d'água. Naquele ano a média geral dos reservatórios na região estava em torno de 27,96% de sua capacidade. A crise hídrica teve como principal gatilho a seca meteorológica e estava associada a como era realizado o planejamento e a gestão dos recursos hídricos. Deste modo este estudo visa contribuir para a área trazendo algumas percepções que poderão ser usadas na previsão.

1.2 **Objetivo Geral**

O objetivo geral dessa dissertação é avaliar modelos de previsão de séries temporais para a demanda d'água, integrando técnicas estatísticas e de aprendizado de máquina, visando proporcionar uma gestão eficiente e sustentável dos recursos hídricos, em específico na região do Bairro Alto em Curitiba (Paraná, Brasil), além de contribuir para a otimização do planejamento e operação de sistemas de abastecimento, promovendo a resiliência diante de variações sazonais e eventos imprevistos.

Os objetivos específicos do estudo são:

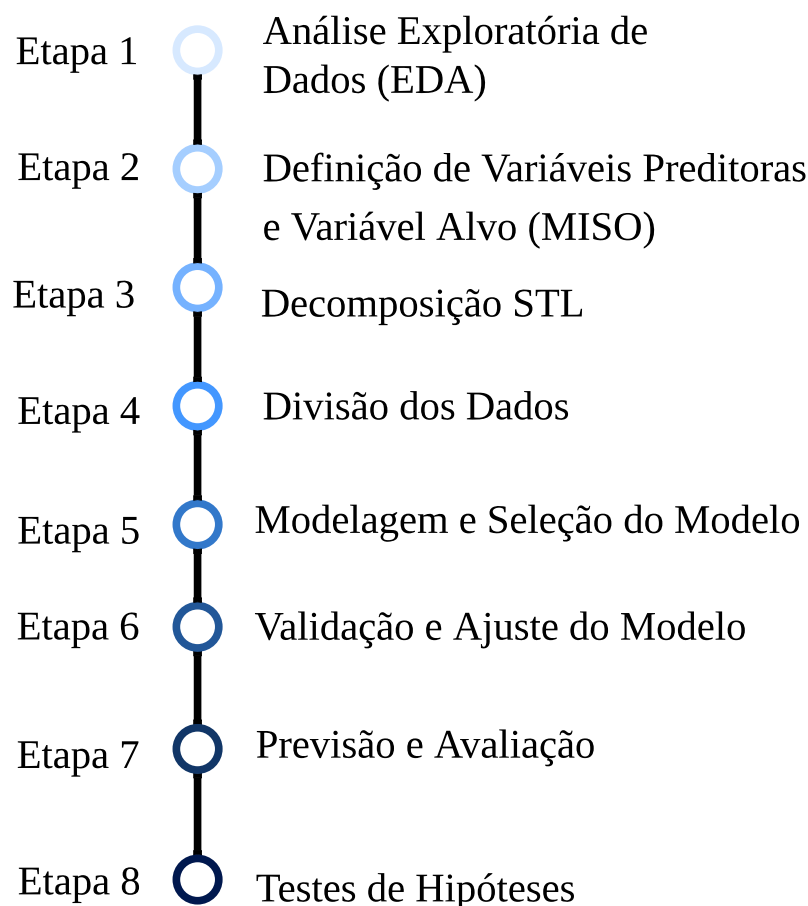
1. Aplicar diferentes modelos de previsão de séries temporais utilizando dados provenientes do Bairro Alto em Curitiba, fornecidos pela SANEPAR.
2. Avaliar a precisão, eficiência e capacidade de previsão desses modelos em conjuntos de dados específicos, utilizando métricas para análise de desempenho.
3. Explorar estratégias de otimização baseadas em otimização Bayesiana, empregando o algoritmo *Tree-structured Parzen Estimator* (TPE) para ajustar os hiperparâmetros dos modelos de previsão de séries temporais.

4. Identificar combinações de modelos de previsão de séries temporais em conjunto com a configuração otimizada.
5. Avaliar o impacto das variáveis exógenas na melhoria da precisão dos modelos de previsão de séries temporais.

1.3 Etapas para Análise dos Dados

Com o objetivo de realizar as previsões e fazer comparações entre os modelos preditores, a pesquisa adotará um processo bem definido, bem como a seleção dos modelos a serem utilizados na Análise Exploratória de Dados (EDA). A pesquisa foi conduzida seguindo as etapas delineadas, conforme apresentado na Figura 1. As etapas para análise dos dados incluem:

Figura 1: Etapas para análise dos dados.



1. EDA: Nesta etapa tem-se a identificação de valores ausentes, a observação de padrões temporais e a detecção de anomalias. Gráficos de linha são comuns para visualizar a convergência dos dados (ROSTAM et al., 2021).

2. Definição de Variáveis Preditoras e Variável Alvo (Modelo MISO): Na segunda etapa, as variáveis preditoras e a variável alvo para a previsão de Múltiplas Entradas e Uma Saída *Multiple Inputs Single Output* (MISO) são selecionadas. Diferentes modelos, podem incorporar variáveis exógenas na modelagem. Essas variáveis exógenas aprimoram as capacidade de previsão do modelo, especialmente quando o horizonte de previsão se estende além dos dados históricos (PAWŁOWSKI et al., 2022).
3. Decomposição STL: O método de decomposição STL separa uma série temporal em três componentes: sazonalidade, tendência e resíduo. Essa decomposição permite. Decompor séries temporais em sazonal captura variações periódicas e repetitivas. Decompor séries temporais em tendência reflete a evolução geral dos dados ao longo do tempo. Para a componente de resíduo engloba as variações não explicadas pelas anteriores (BANDARA; HYNDMAN; BERGMEIR, 2021).
4. Divisão dos Dados: É prática comum dividir o conjunto de dados em conjuntos de treinamento, validação e teste para avaliar o desempenho do modelo. Essa divisão permite uma análise da capacidade de generalização dos modelos, evitando problemas de ajuste excessivo ou insuficiente. A proporção de alocação pode variar, mas uma abordagem é alocar 70% para treinamento e validação, e os 30% restantes para o conjunto de testes. A porção de treinamento e validação pode ser subdividida em 80% para treinamento e 20% para validação (TAO et al., 2020).
5. Modelagem e Seleção do Modelo: Nesta etapa, diversos modelos são construídos e avaliados. Alguns modelos comumente usados para previsão de séries temporais incluem ARX, AR, MA, ARIMA, SARIMA, SARIMAX e modelos de aprendizado de máquina como RNN, LSTM, GRU, DT, *Linear Regression* (LR), XGBoost, LightGBM e Prophet. A escolha do modelo baseia-se em critérios como desempenho na validação, simplicidade do modelo e interpretabilidade dos resultados.
6. Validação e Ajuste do Modelo: Durante a construção do modelo, é importante avaliar seu desempenho usando dados de validação. Neste contexto, métricas de avaliação tais como Erro Médio Absoluto (MAE), Erro Médio Percentual Absoluto Simétrico (SMAPE) e Raiz do Erro Médio Quadrático Relativo (RRMSE) podem ser usadas para comparar e selecionar o melhor modelo. Além disso, técnicas de ajuste como otimização de hiperparâmetros dos modelos usando dados de treinamento e validação combinados podem melhorar o desempenho da previsão das séries temporais.
7. Previsão e Avaliação: Com o modelo final com os dados de treinamento e validação, é possível fazer previsões para o conjunto de testes, que representa dados futuros não

observados. Essas previsões são comparadas com os valores reais correspondentes para avaliar a qualidade e precisão do modelo.

8. Teste de Hipótese: Aplicar os modelos de previsão e fazer comparativo baseado em testes de significância estatística (*Friedman e Nemenjy*).

Cada uma dessas etapas desempenha um papel crucial na análise dos dados e no processo de modelagem das séries temporais, contribuindo para a compreensão dos dados, construção e validação dos modelos de previsão.

1.4 Estrutura do Documento

Esse documento de dissertação está organizado em 5 capítulos. O Capítulo 1 apresentou contextualização, justificativa, objetivos, contribuições, e a organização do documento. O Capítulo 2 menciona uma visão geral das principais pesquisas na área de previsão de séries temporais aplicadas a demanda d'água. No Capítulo 3 são descritos os modelos de previsão que serão utilizados nos dados de séries temporais de abastecimento d'água no Bairro Alto em Curitiba, dados estes fornecidos pela SANEPAR por meio da coleta horária durante 3 anos consecutivos. O Capítulo 4 apresenta os resultados obtidos ao longo da pesquisa com discussões. Os resultados de previsão obtidos são detalhados e a técnica de otimização e métricas estatísticas de desempenho são aplicadas e analisadas. O Capítulo 5 apresenta os principais resultados da pesquisa, as limitações e delineia possíveis estudos futuros na área.

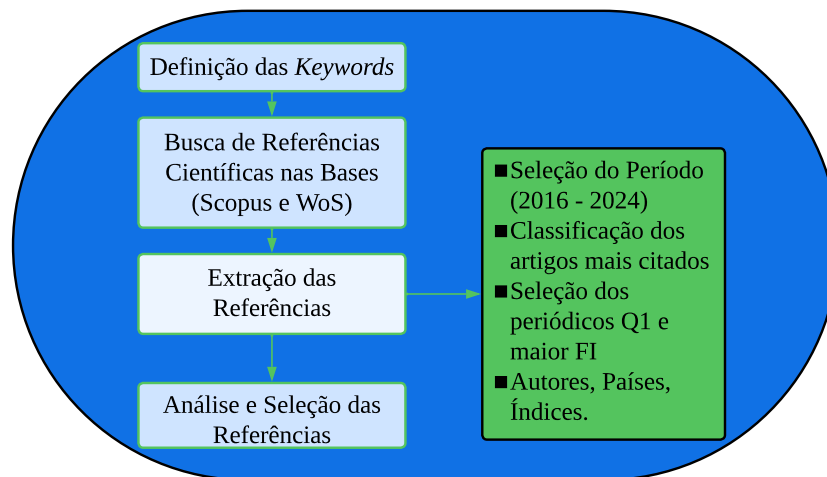
2 Revisão da Literatura

Este capítulo apresenta uma revisão sistemática da literatura (RSL) nos temas relacionados a previsão de séries temporais e aplicações em hidrologia e mais especificamente em abastecimento d'água. A revisão bibliográfica realizada consiste em uma análise abrangente e crítica das principais fontes de literatura. As informações extraídas da literatura são fundamentais para embasar a fundamentação teórica, a metodologia e a análise dos resultados deste estudo.

A seleção das referências foi baseada em critérios específicos tais como. Definição das bases de busca, escolha das *keywords*, seleção do período de busca (2016 a 2024), seleção do tipo de artigo, organização pela ordem de citações, verificação dos periódicos mais importantes (fator Q1 do Scimago e Fator de Impacto (FI)). Se um periódico pertencer ao quartil Q1 do *Scimago Journal Rank* (SJR) significa que tem um desempenho melhor do que pelo menos 75% das revistas dessa mesma categoria. Embora nem todas as referências obtidas tenham uma relação evidente com a área de aprendizado de máquina, elas contribuem como material de suporte a implementação de alguns modelos avaliados para previsão nesta dissertação e podem servir como base para outros estudos.

A Figura 2 apresenta um fluxograma de como a pesquisa foi realizada, destacando a importância da escolha dos periódicos Q1 e com maior fator de impacto, como base para esta RSL. A mesma figura apresenta uma adaptação da metodologia proposta por Martins e Gorschek (2016) para a realização da RSL, onde primeiramente foram realizadas buscas na base Scopus e *Web of Science* (WoS), selecionando referências relevantes para o tema da pesquisa. Para as duas bases de busca utilizadas foram usadas as palavras-chave “*time series forecasting*”, “*time series analysis*”, “*sanitation*” e “*water supply*” .

Figura 2: Fluxograma da Revisão Sistemática da Literatura.



Na etapa seguinte, foi realizada uma avaliação preliminar de cada artigo obtido.

Critérios de Inclusão:

1. Idioma: Os artigos considerados devem estar escritos em inglês.
2. Tipo de Publicação: Devem ser artigos (*papers*).
3. Base de Dados: Artigos provenientes de bases de dados como Scopus e WoS são considerados.

Critérios de Exclusão:

1. Filtro Anual: de 2016 a 2024.
2. Duplicatas: Artigos duplicados serão removidos durante o processo de revisão.
3. Número Elevado de Artigos: Diante da quantidade significativa de artigos encontrados, optou-se por realizar uma análise preliminar sem a aplicação de filtros anuais para otimizar o processo de busca. Na base de dados Scopus, por exemplo, existiam 831 artigos, enquanto na base de dados WoS, foram encontrados 98 artigos, totalizando 929 artigos.

Levando em consideração a diferença entre essa estimativa apresentada na Tabela 2 e a quantidade de artigos restantes após a remoção de duplicatas, tem-se menos de 929 artigos para análise. É válido lembrar que, ao remover as duplicatas, o número diminuiu ainda mais, chegando a 906 artigos.

Na etapa final, foi realizada uma análise dos conteúdos dos artigos selecionados, levando em consideração as áreas de especialização. Como esta revisão está inserida no contexto de um programa de mestrado em Engenharia de Produção e Sistemas, vale a pena analisar a relação dos artigos obtidos com áreas afins tais como Matemática. Assim, as áreas mais relevantes para a pesquisa foram Informática, Engenharia e Matemática, representando 50% das publicações.

São apresentados os resultados da RSL, utilizando um *software* VOSviewer de cada base de dados utilizada no trabalho. A Figura 3 mostra os modelos de previsão usados com frequência em conjunto com “*time series*” nos artigos obtidos nas bases de dados Scopus e WoS.

Nesse primeiro momento, foram obtidos 2.555 modelos, dos quais 83 modelos são dispostos na Figura 3. É importante destacar que as palavras-chave utilizadas foram *time series forecasting* ou *time series analysis* e *water supply* e *sanitation* em ambas as bases de dados. Os modelos apresentados na Figura 3 têm como base para as escolhas os modelos mais utilizados na literatura. Com essa visão mais geral, é possível ter um ponto de partida para a escolha dos modelos.

Tabela 1: Combinação de palavras-chave aplicando filtros.

| Bases | Palavras chaves | | | | | | | Resultados |
|--------|-----------------|-----|-------------|-----|--------|-----|------------|------------|
| Scopus | time series | AND | time series | | | | | 798 |
| | forecasting | | analysis | | | | | |
| WoS | time series | OR | time series | AND | water | AND | sanitation | 33 |
| | forecasting | | analysis | | supply | | | |
| WoS | time series | OR | time series | | | | | 79 |
| | forecasting | | analysis | | | | | |
| WoS | time series | OR | time series | AND | water | AND | sanitation | 19 |
| | forecasting | | analysis | | supply | | | |
| Total | | | | | | | | 929 |

Tabela 2: Resumo dos artigos obtidos com a RSL nas bases Scopus e WoS.

| | |
|--|-------|
| Quantidade de artigos obtidos | 929 |
| Quantidade de artigos da base WoS | 98 |
| Quantidade de artigos da base Scopus | 831 |
| Remoção de artigos duplicados | |
| Porcentagem de artigos duplicados | 87% |
| Quantidade de artigos duplicados | 23 |
| Quantidade de artigos sem duplicados | 906 |
| Porcentagem de artigos duplicados removidos da base WoS | 19,4% |
| Porcentagem de artigos duplicados removidos da base Scopus | 0,5% |
| Quantidade de artigos duplicados com diferentes citações | 3 |
| Porcentagem de artigos duplicados com diferentes citações | 13% |

tigos com as combinações utilizadas na RSL para o tema de estudo em questão. Todos os periódicos são listadas em ordem decrescente pela quantidade de publicações obtidas, incluindo a métrica do SJR que avalia a importância relativa de periódicos científicos com base em sua influência e prestígio na comunidade acadêmica. Ele é calculado usando algoritmos que levam em consideração a qualidade das citações recebidas por um periódico. Neste estudo a RSL procurou basear-se em periódicos classificados como Q1 e Q2, bem como o *h-index*.

Tabela 3: Classificação dos principais periódicos obtidos na RSL.

| Periódicos | No. de artigos | SJR | <i>h-index</i> |
|------------------------|----------------|-----|----------------|
| Neurocomputing | 27 | Q1 | 143 |
| IEEE Access | 18 | Q1 | 127 |
| Applied Soft Computing | 12 | Q1 | 143 |
| Energies | 11 | Q2 | 93 |
| Energy | 11 | Q1 | 343 |

O *software ScientoPy* obtém os principais tópicos de tendência com base na maior

taxa de crescimento médio *Average Growth Rate* (AGR). A AGR é a diferença média entre o número de documentos publicados em um ano e o número de documentos publicados no ano anterior (RUIZ-ROSETO; RAMIREZ-GONZALEZ; VIVEROS-DELGADO, 2019). Indicando como o número de documentos publicados para um tópico em específico cresceu (número positivo) ou diminuiu (número negativo) em média dentro de um período de tempo. Assim, o AGR é calculado por,

$$\text{AGR} = \frac{\sum_{i=Y_s}^{Y_e} P_i - P_{i-1}}{(Y_e - Y_s) + 1} \quad (2.1)$$

onde AGR é a taxa média de crescimento, Y_e é o ano final, Y_s é o ano inicial, P_i é o número de publicações no ano i . Para o ano final Y_e , o *ScientoPy* utiliza o ano final global por defeito configurado nas opções globais ou/em parâmetros do comando *ScientoPy*. O ano de início Y_s é calculado a partir do ano final Y_e , conforme calculado por,

$$Y_s = Y_e - (\text{WindowWidth} + 1) \quad (2.2)$$

onde a largura da janela (*Window Width*) é predefinida como 2 anos. Assim, se o ano final for 2018, o AGR é a taxa de crescimento média entre 2017 e 2018 (RUIZ-ROSETO; RAMIREZ-GONZALEZ; VIVEROS-DELGADO, 2019).

A média de documentos por ano *Average Documents per Year* (ADY) é um indicador absoluto que representa o número médio de documentos publicados num período de tempo para um tópico específico. O ADY é calculado por,

$$\text{ADY} = \frac{\sum_{i=Y_s(t)}^{Y_e(t)} P_i}{(Y_e(t) - Y_s(t)) + 1} \quad (2.3)$$

onde ADY é a média de documentos por ano, $Y_e(t)$ é o ano final, $Y_s(t)$ é o ano inicial, calculado como descrito na equação (2.3), P_i é o número de publicações no ano i .

A percentagem de documentos nos últimos anos *Percentage of Documents in Last Years* (PDLY) é um indicador relativo que representa a percentagem do ADY em relação ao número total de documentos para um tópico específico. Desta forma, o PDLY é calculado como,

$$\text{PDLY} = \frac{\sum_{i=Y_s(t)}^{Y_e(t)} P_i}{(Y_e(t) - Y_s(t) + 1) \cdot \text{TND}} \cdot 100\% \quad (2.4)$$

onde $PDLY$ é a percentagem de documentos nos últimos anos, $Y_e(t)$ é o ano final, Y_s é o ano inicial, calculado como descrito na equação (2.4), P_i é número de publicações no ano i , TND é o número total de documentos.

Tabela 4 descreve os principais autores obtidos na RSL descrita previamente, sobre o tema em análise. Essa abordagem visa evitar a inclusão de todos os autores e destacar aqueles que tiveram uma contribuição significativa na área. Dessa forma, é possível identificar o principal autor que se destacou, fornecendo uma visão geral da distribuição da produção científica entre os pesquisadores. Na Tabela 4 são descritos os valores da taxa de crescimento médio AGR, documentos médios por ano ADY a porcentagem de documentos nos últimos anos PDLY no período de 2021 a 2023.

Tabela 4: Total de publicações dos principais autores obtidos na RSL.

| Author | No. de artigos | AGR | ADY | PDLY | <i>h-index</i> |
|---------------------------|----------------|------|-----|------|----------------|
| Wang et al. (2016) | 11 | -0,5 | 2 | 36,4 | 8 |
| Shen e Wang (2022) | 11 | 0 | 3 | 54,5 | 5 |
| Xian et al. (2018) | 10 | 1 | 2,5 | 50 | 5 |
| Li et al. (2018) | 9 | -1,5 | 2 | 44,4 | 4 |
| Sang et al. (2016) | 7 | 1,5 | 2 | 57,1 | 3 |
| Sadaei et al. (2019a) | 7 | 1 | 2 | 57,1 | 3 |
| Hao et al. (2023) | 7 | 1 | 3 | 85,7 | 2 |
| Guo, Pedrycz e Liu (2018) | 7 | 1,5 | 3 | 85,7 | 3 |
| O'Donncha et al. (2022) | 6 | 0 | 1,5 | 50 | 4 |
| Xu et al. (2019) | 6 | 0 | 1,5 | 50 | 5 |

A Tabela 5 apresenta os países com maior número de artigos obtidos na RSL usando as palavras-chaves citadas previamente. Tais países estão ordenados de forma decrescente pelo número de publicações obtido. Os principais países que se destacam nessa análise são China, com 179 publicações, Estados Unidos da América com 74 publicações, Índia com 61 publicações, Brasil com 49 publicações, Espanha com 40 publicações, Reino Unido com 40 publicações, Austrália com 31 publicações, Itália com 26 publicações, Canadá com 25, Irã com 20 publicações.

O artigo de Ursu e Pereau (2016) destaca a importância do modelo ARIMA na previsão e demanda d'água, porém, é pertinente questionar até que ponto esse modelo pode ser aplicado eficazmente em cenários mais complexos e dinâmicos, especialmente diante dos desafios emergentes na gestão dos recursos hídricos.

O modelo ARIMA é reconhecido como valioso na previsão e demanda d'água, proporcionando uma base sólida para abordar desafios relacionados à oferta hídrica.

A limitação do ARIMA em cenários complexos e dinâmicos pode comprometer sua eficácia em situações que exigem considerações mais abrangentes.

A estratégia proposta por Graff et al. (2017), ao combinar o modelo ARIMA com

Tabela 5: Total de publicações dos principais países obtidos na RSL.

| País | No. de artigos | AGR | ADY | PDLY | <i>h-index</i> |
|---------------------------|----------------|------|------|------|----------------|
| China | 179 | 18,5 | 48 | 53,6 | 31 |
| Estados Unidos da América | 74 | 3 | 16 | 43,2 | 21 |
| Índia | 61 | 0 | 12 | 39,3 | 18 |
| Brasil | 49 | 3,5 | 12,5 | 51 | 17 |
| Espanha | 40 | 1,5 | 8,5 | 42,5 | 12 |
| Reino Unido | 40 | 3 | 10 | 50 | 15 |
| Austrália | 31 | 3,5 | 7,5 | 48,4 | 14 |
| Itália | 26 | 2 | 7 | 53,8 | 10 |
| Canadá | 25 | 1 | 5,5 | 44 | 11 |
| Irã | 20 | -1 | 3,5 | 35 | 11 |

programação genética para otimizar o ajuste do modelo GP junto à suavização exponencial, demanda uma análise crítica. Será que a introdução de técnicas mais avançadas realmente contribui para melhorias substanciais na precisão das previsões, ou há o risco de introduzir complexidade desnecessária?

A combinação de ARIMA com programação genética oferece uma abordagem inovadora, potencialmente melhorando a adaptação do modelo a padrões complexos nos dados.

A introdução de técnicas avançadas pode aumentar a complexidade do modelo, tornando-o menos acessível ou interpretável para usuários não especializados.

As observações de Tyralis e Papacharalampous (2017a) sobre o desempenho destacado dos modelos ARMA e RF em séries temporais de curto prazo levantam questões sobre a generalização desses resultados para diferentes contextos e horizontes de previsão. É fundamental avaliar criticamente a robustez desses modelos em face de condições variáveis e demandas específicas de diferentes sistemas de abastecimento d'água.

Os modelos ARMA e RF demonstraram desempenho destacado em séries temporais de curto prazo, indicando sua aplicabilidade em cenários de alta volatilidade.

A generalização desses resultados para diferentes contextos pode ser limitada, especialmente em situações que exigem previsões de longo prazo ou considerações mais complexas.

O método proposto, que modela ciclos não sazonais e sazonais separadamente, parece promissor. No entanto, uma análise crítica se faz necessária para avaliar se a simplificação dessa abordagem não compromete a capacidade de capturar nuances complexas nos dados de carga, especialmente em ambientes onde fatores externos podem desempenhar um papel significativo.

A abordagem de modelagem separada para ciclos não sazonais e sazonais oferece

uma estrutura clara, permitindo a consideração específica de diferentes padrões nos dados.

A simplificação pode resultar na perda de informações importantes, especialmente em ambientes onde fatores externos influenciam significativamente os padrões de carga.

Os modelos de aprendizado de máquina simples, como ANN, SVM e CART, mencionados por Chou e Tran (2018), oferecem uma alternativa interessante. Contudo, é preciso questionar até que ponto essas abordagens são suficientes para lidar com a complexidade inerente às séries temporais de demanda d'água. Será que a simplicidade desses modelos compromete a precisão em cenários mais desafiadores?

Modelos de aprendizado de máquina simples são acessíveis e eficazes para muitos casos, proporcionando uma solução prática.

Em cenários altamente complexos, a simplicidade desses modelos pode limitar sua capacidade de capturar padrões intrincados nos dados.

A diversidade de modelos baseados em dados destacada por Ahmad et al. (2018) levanta a questão da seleção adequada para diferentes contextos. A crítica aqui se concentra em entender se a escolha do modelo é guiada por uma compreensão profunda do problema específico em análise ou se há uma tendência a aplicar abordagens genéricas sem considerar as nuances do sistema.

A diversidade de modelos oferece flexibilidade na escolha da abordagem mais adequada para contextos específicos.

A seleção inadequada, sem uma compreensão aprofundada do problema, pode resultar em escolhas subótimas, comprometendo a precisão das previsões.

A combinação de RNNs com outras técnicas de aprendizagem profunda, como CNN, conforme mencionado por Chen et al. (2018), é uma estratégia intrigante. Contudo, a crítica necessária é em relação à necessidade de uma grande quantidade de dados de treinamento e parâmetros. Isso levanta a questão da viabilidade prática desses modelos em cenários onde dados abundantes podem não estar prontamente disponíveis.

A combinação de RNNs com outras técnicas de aprendizagem profunda oferece uma abordagem holística para capturar dependências temporais complexas.

A necessidade de grandes volumes de dados e parâmetros pode limitar a aplicabilidade prática desses modelos em situações com recursos limitados.

A transformação de séries temporais em imagens proposta por Sadaei et al. (2019b) oferece uma perspectiva única. No entanto, a crítica se volta para a aplicabilidade e interpretabilidade dessa abordagem. Como as imagens refletem efetivamente as complexidades das séries temporais e como os resultados são interpretados em termos práticos são questionamentos importantes.

A transformação de séries temporais em imagens pode proporcionar uma representação visual intuitiva dos padrões nos dados.

A interpretabilidade das imagens e a relação direta com os padrões nas séries temporais precisam ser cuidadosamente avaliadas para garantir a eficácia dessa abordagem.

A abordagem de Shih, Sun e Lee (2019a), ao combinar RNN, LSTM e CNN para modelar séries temporais, suscita dúvidas sobre sua eficácia. A fusão dessas técnicas complexas pode gerar uma estrutura computacionalmente intensiva, sem garantia clara de melhorias nas previsões. A introdução de pesos de atenção e o vetor de contexto adicionam incerteza ao processo, comprometendo a interpretabilidade dos resultados.

A combinação de RNN, LSTM e CNN oferece uma abordagem abrangente para modelar diferentes aspectos das séries temporais.

A complexidade computacional, juntamente com a incerteza introduzida pelos pesos de atenção, pode dificultar a interpretação clara dos resultados.

Em resumo, a análise crítica dos modelos de previsão de séries temporais destaca a necessidade de uma abordagem equilibrada. Cada modelo apresenta vantagens e desvantagens, e a escolha deve ser orientada pelos requisitos específicos de cada situação. É crucial considerar a aplicabilidade prática, a interpretabilidade dos resultados e a capacidade de lidar com a complexidade inerente aos dados de demanda d'água.

Na Tabela 6 é apresentada a quantidade de artigos foram obtidos pela RSL para cada modelo de previsão e também cita-se pelo menos um autor correspondente a cada modelo.

Tabela 6: Principais modelos de previsão obtidos na RSL.

| Modelos | Autores | No. | AGR | ADY | PDLY | <i>h-index</i> |
|--------------------|-------------------------------------|-----|-----|------|------|----------------|
| ARIMA | Buyuksahin e Ertekin (2019) | 84 | 1,7 | 16,7 | 59,5 | 27 |
| ARMA | Tyralis e Papacharalampous (2017b) | 7 | 0,3 | 0,7 | 28,6 | 6 |
| Árvores de Decisão | Fouilloy et al. (2018) | 12 | 0,7 | 3 | 75 | 7 |
| ARX | Gustin, McLeod e Lomas (2018) | 3 | 0 | 0,7 | 66,7 | 2 |
| CNN | Rostamian e O'Hara (2022) | 8 | 1,3 | 2,7 | 100 | 4 |
| GRU | Yang, Guo e Li (2022) | 5 | 0 | 1,3 | 80 | 4 |
| LR | Mohan et al. (2022) | 3 | 0 | 0,7 | 66,7 | 3 |
| LSTM | Sezer, Gudelek e Ozbayoglu (2020) | 35 | 3,3 | 10,7 | 91,4 | 16 |
| Prophet | Kulshreshtha e Vijayalakshmi (2020) | 3 | 0,3 | 1 | 100 | 3 |
| RF | Yang, Guo e Li (2022) | 9 | 1,7 | 2,7 | 88,9 | 5 |
| RNN | Shih, Sun e Lee (2019b) | 20 | 0 | 4,3 | 65 | 11 |
| SARIMA | Kushwah e Wadhvani (2022) | 5 | 1 | 1,7 | 100 | 4 |
| SARIMAX | Bhangu, Sandhu e Sapra (2022) | 2 | 0,3 | 0,7 | 100 | 2 |
| XGBoost | Liu et al. (2022) | 1 | 0,3 | 0,3 | 100 | 0 |

Além dos modelos prévios, também será utilizada a versão atualizada do ARIMA nesta dissertação, bem como os modelos SARIMA e SARIMAX serão comparados para determinar qual deles é o mais adequado. Além disso, serão empregados os modelos Light GBM e XGBoost. Os modelos de aprendizado profundo, como a RNN, ainda é considerada

um excelente modelo para previsão de séries temporais no tema de saneamento básico que está estudado.

Embora existam diversas variações do modelo ARIMA, o Prophet, desenvolvido pelo Facebook, destaca-se como uma opção superior em muitos aspectos. O Prophet, introduzido em 2017, simplifica consideravelmente muitas tarefas do processo de modelagem em comparação com o ARIMA, que tem origens que remontam à década de 1960. Essa disparidade temporal ressalta a evolução e a modernização contínuas no campo de modelagem de séries temporais ao longo das décadas (RAMOS, 2010).

No Prophet, as tarefas são mais automatizadas, tornando-o mais acessível, especialmente para usuários não especializados. Em contraste, os modelos ARIMA frequentemente requerem ajustes manuais dos parâmetros e uma compreensão mais profunda do processo de modelagem, exigindo uma intervenção mais manual. Essa diferença enfatiza a natureza mais automática do Prophet em comparação com o processo potencialmente mais manual dos modelos ARIMA.

Ademais, é relevante destacar que o Prophet demonstrou ser especialmente eficaz em previsões de longo prazo. Sua capacidade de lidar automaticamente com padrões sazonais, feriados e eventos específicos contribui para uma modelagem mais robusta em horizontes temporais estendidos. Essa característica, aliada à sua abordagem mais automatizada, confere ao Prophet uma vantagem adicional em cenários de previsão de longo prazo em comparação com modelos ARIMA, que podem exigir uma sintonização manual mais cuidadosa para estender suas previsões.

3 Fundamentos dos Modelos de Previsão

Neste capítulo são descritos os modelos de previsão de séries temporais usados neste estudo, bem como as medidas de desempenho e testes de hipóteses de performance.

3.1 Conceito de Séries Temporais

Uma série temporal é uma coleção de dados ordenados no tempo, utilizados para analisar o comportamento passado e prever o comportamento futuro de uma variável de interesse. Ela pode ser univariada, se envolve apenas uma variável, ou multivariada, se envolve mais de uma variável. A série temporal pode ser contínua, se os dados são observados continuamente ao longo do tempo, ou discreta, se os dados são observados em intervalos regulares ou irregulares. Ela pode apresentar diferentes componentes, como tendência, sazonalidade, ciclos e ruído, que podem ser modelados e decompostos por diversas técnicas estatísticas. Com aplicações em diversas áreas, como economia, finanças, meteorologia, epidemiologia, entre outras, a série temporal se revela como uma ferramenta versátil na compreensão de dados históricos e sistemas dinâmicos em geral.

3.2 Estratégia Recursiva de Previsão Multipassos

Segundo Ribeiro et al. (2021) diversas estratégias para previsões de vários passos à frente em séries temporais. Explora-se a eficácia das abordagens recursiva e direta, bem como a combinação destas com o uso de Múltiplas Entradas e Múltiplas Saídas (MIMO). Embora a estratégia recursiva ofereça flexibilidade no modelo de previsão, sua propensão a erros acumulativos pode ser um obstáculo significativo. Por outro lado, a estratégia direta, embora evite a acumulação de erros de previsão, pode exigir mais recursos computacionais devido à necessidade de ajustar um modelo para cada horizonte de previsão.

Comparativos entre as estratégias têm produzido resultados contraditórios quanto à sua precisão. Além disso, é fundamental considerar o custo computacional ao escolher a abordagem de previsão mais adequada. A escolha entre as estratégias pode depender do contexto específico, com a estratégia recursiva sendo preferível em situações onde os recursos computacionais são limitados, enquanto a estratégia direta pode ser mais apropriada para séries temporais mais longas ou quando se busca uma minimização de viés. A equação de previsão recursiva pode ser expressada como:

$$\hat{y}_{t+h} = \begin{cases} f[y_t, y_{t-1}, \dots, y_{t-n_y+1}] & \text{se } h = 1 \\ f[\hat{y}_{t+h-1}, \dots, \hat{y}_{t+1}, y_t, \dots, y_{t+h-n_y}] & \text{se } h \in [2, \dots, n_y] \\ f[\hat{y}_{t+h-1}, \dots, \hat{y}_{t+h-n_y}] & \text{se } h \in [n_y + 1, \dots, H] \end{cases} \quad (3.1)$$

onde \hat{y}_{t+h} é o valor previsto no tempo t e horizonte de previsão até h , y_{t+h-n_y} e \hat{y}_{t+h-n_y} são os casos observados e previstos anteriormente com defasagens em $n_y = 1, \dots, n$. Quando variáveis exógenas relacionadas à variável de entrada são adotadas, um vetor \mathbf{x} de variáveis exógenas deve ser considerado e incorporado na equação (3.1), com defasagens de acordo com o horizonte de previsão desejado.

Quanto à estratégia recursiva, a fraqueza está relacionada aos erros acumulativos conforme a expansão do horizonte de previsão. Além disso, é possível observar que para o h -ésimo horizonte de previsão, existem $n_y - (h - 1)$ valores observados, e $h - 1$ valores previstos são usados como entradas no caso de $h - 1 < n_y$. Caso contrário, o conjunto de entradas é composto apenas pelos valores previstos, reduzindo a precisão da previsão. A Figura 6 ilustra a estratégia recursiva quando quatro defasagens são usadas como entradas do modelo no horizonte de previsão de 3 passos à frente.

3.3 Modelos Clássicos de Séries Temporais

O modelo ARIMA é um modelo clássico utilizado para a previsão de séries temporais. Ele é composto por três componentes principais: o componente AR, o componente de MA e o componente de diferenciação (I), que é aplicado para tornar a série temporal estacionária (PRABHAKARAN, 2018).

O componente auto-regressivo do modelo ARIMA é representado por $AR(p)$, em que o parâmetro p determina o número de defasagens ou atrasos (*lags*) a serem usados. A equação do modelo $AR(p)$ é calculada da seguinte forma:

$$Y_t = c + \sum_{n=1}^p \alpha_n Y_{t-n} + \varepsilon_t \quad (3.2)$$

onde na equação (3.2), o termo ε_t representa o ruído branco que é caracterizado por um sinal com média zero e variância constante. Essa equação pode ser entendida como uma regressão múltipla, em que os valores defasados de Y_t são utilizados como preditores. Esse modelo é conhecido como modelo autorregressivo de ordem p , ou $AR(p)$.

O modelo ARX é uma extensão do modelo AR, que incorpora variáveis exógenas

nos dados para tentar melhorar as previsões. Esse modelo também é multivariado, e foi incluído nessa dissertação para fins de comparação com o modelo AR simples, considerando a presença de variáveis exógenas. Pode-se mencionar que de acordo com o valor de p tem-se alguns aspectos relevantes a citar: Se o parâmetro p for definido como zero AR(0), significa que não há termos autorregressivos no modelo. Nesse caso, a série temporal se comporta como um ruído branco. Com o parâmetro p definido como 1, o modelo AR leva em consideração o valor anterior da série temporal multiplicado por um coeficiente e, em seguida, adiciona ruído branco.

Se o parâmetro p estiver na faixa $0 < \alpha < 1 \in \mathbb{Z}$, ocorre o fenômeno de reversão média. Isso significa que os valores da série tendem a oscilar em torno de uma média central e a regressar em direção a ela após se afastarem. Esse padrão indica uma tendência de retorno à média ao longo do tempo. Aumentar ainda mais o parâmetro p no modelo AR significa considerar um número crescente de medições de tempo anteriores, cada uma multiplicada pelo seu próprio coeficiente. Isso permite levar em conta uma memória mais longa da série temporal e capturar padrões de dependência complexos ao longo do tempo. No entanto, é importante ter em mente que aumentar excessivamente o valor de p pode levar a problemas de sobre-treinamento (do inglês *overfitting*), onde o modelo se ajusta muito bem aos dados de treinamento, mas tem um desempenho ruim na previsão de novos dados. Portanto, é necessário encontrar um equilíbrio entre a complexidade do modelo e sua capacidade de generalização.

No modelo de MA, o componente não é uma média móvel simples, mas sim uma combinação de termos de erro de previsão defasados. O parâmetro q no modelo MA representa o número de termos de erro de previsão que são levados em consideração na previsão. Este componente não é uma média móvel, mas sim os atrasos no ruído branco (TRENBERTH, 1984). Em um modelo MA(1), por exemplo, a previsão é composta por um termo constante, o produto do termo de erro de previsão anterior por um multiplicador, e o termo de erro de previsão atual. Essa abordagem baseia-se em princípios estatísticos e de probabilidade, ajustando a previsão com base em termos anteriores de erro de previsão.

O modelo MA é uma alternativa ao modelo AR e é usado para capturar padrões de dependência na média móvel, ou seja, a influência de erros passados na previsão atual. Ao combinar o modelo AR e o modelo MA, como no modelo ARMA, é possível obter uma modelagem mais abrangente que considera tanto a dependência autorregressiva quanto a dependência na média móvel (VIDHYA, 2023), tal que

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (3.3)$$

onde ε_t representa o ruído branco, esse modelo é conhecido como um modelo de média móvel $MA(q)$, em que q é a ordem da média móvel. É importante ressaltar que não se observam diretamente os valores de ε_t , portanto, essa modelagem não se trata de uma regressão no sentido convencional.

Diferentemente de uma regressão comum em que se têm variáveis explicativas observadas, no modelo $MA(q)$, são usados os termos de ruído branco defasados para estimar e prever os valores da série temporal. O objetivo é capturar a dependência dos termos de erro passados na previsão atual (VIDHYA, 2023).

O modelo ARMA é uma combinação dos modelos AR e MA, onde o modelo AR é adicionado ao modelo MA. No modelo ARMA, é adicionada uma constante à soma dos termos autorregressivos multiplicados pelos seus coeficientes, juntamente com a soma dos termos de média móvel multiplicados pelos seus coeficientes, além do ruído branco. Essa estrutura é amplamente utilizada em diversos modelos de previsão em diferentes áreas. Esse modelo é bastante semelhante ao modelo ARIMA, pois calcula os termos, mas não inclui a diferenciação presente tanto no modelo ARMA quanto no modelo ARIMA, tal que

$$Y_t = \beta_2 + \omega_1 \varepsilon_{t-1} + \omega_2 \varepsilon_{t-2} + \dots + \omega_q \varepsilon_{t-q} + \varepsilon_t \quad (3.4)$$

onde Y_t representa a série temporal que foi diferenciada (possivelmente mais de uma vez). Os preditores no lado direito da equação incluem os valores defasados de Y_t e os erros defasados. Esse tipo de modelo é conhecido como ARIMA (p, d, q) .

O modelo ARIMA é uma extensão do modelo ARMA que incorpora uma etapa adicional de pré-processamento chamada de diferenciação. Essa etapa é representada pela notação $I(d)$, em que d denota a ordem de diferenciação, ou seja, o número de transformações necessárias para tornar a série temporal estacionária. Portanto, um modelo ARIMA é simplesmente um modelo ARMA aplicado à série temporal diferenciada. Isso permite lidar com séries temporais que possuem tendências ou padrões não estacionários.

Embora os modelos ARIMA sejam eficazes, incorporar variáveis sazonais e exógenas ao modelo pode potencializar sua capacidade de previsão. No entanto, é importante destacar que o modelo ARIMA pressupõe que a série temporal seja estacionária. Quando lidamos com séries temporais não estacionárias, é necessário recorrer a outros modelos para a análise e previsão adequadas (VIDHYA, 2023). Um exemplo é o do modelo SARIMA gerado por

$$Y_t = c + \sum_{n=1}^p \alpha_n y_{t-n} + \sum_{n=1}^q \theta_n \epsilon_{t-n} + \sum_{n=1}^P \phi_n y_{t-sn} + \sum_{n=1}^Q \eta_n \epsilon_{t-sn} + \epsilon_t \quad (3.5)$$

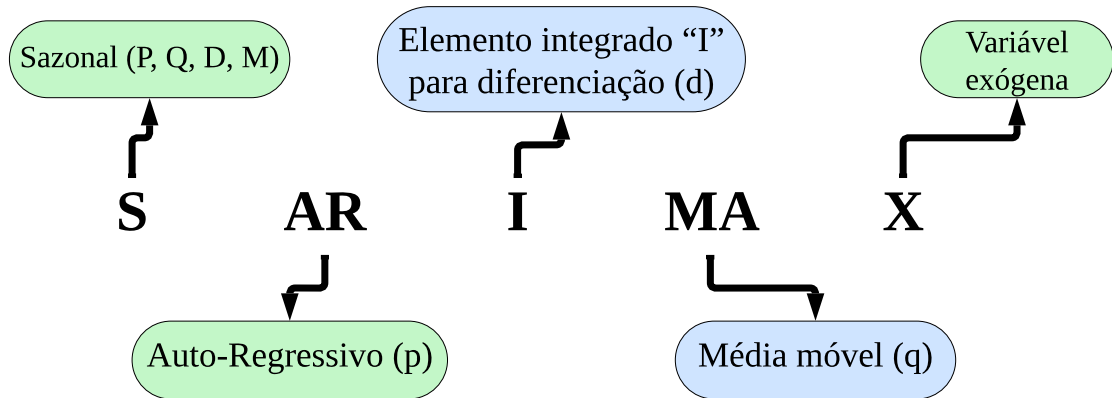
O modelo proposto é uma extensão do modelo ARIMA, com a adição de componentes autorregressivos e de média móvel sazonal. Esses componentes extras são ajustados levando em consideração os padrões sazonais presentes nos dados, utilizando atrasos correspondentes à frequência sazonal (por exemplo, 12 para dados mensais). Essa abordagem permite capturar e modelar de forma mais precisa as variações sazonais e melhorar a qualidade das previsões em séries temporais com esse comportamento cíclico (PREDUM, 2021).

Nos modelos ARIMAX e SARIMAX são consideradas variáveis exógenas, ou seja, são utilizados dados externos para a realização das previsões. É importante ressaltar que mesmo que essas variáveis exógenas sejam indiretamente modeladas no histórico de previsões do modelo, ao incluí-las diretamente, o modelo será capaz de responder de forma ágil aos efeitos dessas variáveis (PREDUM, 2021) tal que.

$$d_t = c + \sum_{n=1}^p \alpha_n d_{t-n} + \sum_{n=1}^q \theta_n \epsilon_{t-n} + \sum_{n=1}^r \beta_n x_{nt} + \sum_{n=1}^P \phi_n d_{t-sn} + \sum_{n=1}^Q \eta_n \epsilon_{t-sn} + \epsilon_t \quad (3.6)$$

onde p representa a ordem de autorregressão da tendência na *AutoCorrelation Function* (ACF), sendo o número de termos autorregressivos (parte AR) que incorporam o efeito de valores passados no modelo. d é a diferença de tendência, indicando o número de diferenças não sazonais necessárias para atingir estacionariedade, e q é a ordem da média móvel da tendência na *Partial AutoCorrelation Function* (PACF). Todas essas considerações são válidas somente se a série temporal for estacionária, d é o número de erros de previsão defasados na equação de previsão (parte MA). Enquanto os elementos sazonais em SARIMAX são, P é a ordem autorregressiva sazonal, D é a Ordem das diferenças sazonais, M é o número de etapas de tempo para um único período sazonal, M é igual à defasagem ACF com o valor mais alto (normalmente em uma defasagem alta). $D = 1$ se a série tiver um padrão sazonal estável ao longo do tempo, $D = 0$ se a série tiver um padrão sazonal instável ao longo do tempo, $P \geq 1$ se a PACF for positiva na defasagem M , senão $P = 0$, $Q \geq 1$ se a ACF for negativa na defasagem M , caso contrário $Q = 0$, e X é a variável exógena. Na Figura 4 é mostrado detalhes do SARIMAX.

Figura 4: Elementos do modelo SARIMAX.



3.4 Modelos de Aprendizado de Máquina

Os modelos de aprendizado de máquina para séries temporais têm sido amplamente reconhecidos e utilizados na literatura atual, especialmente por não serem baseados em métodos de gradiente. Esses modelos são relevantes por sua capacidade de capturar relações complexas e não lineares nos dados, permitindo previsões eficientes. Sua popularidade reflete o reconhecimento da eficácia desses modelos em abordar uma ampla gama de problemas de previsão de séries temporais em diferentes áreas de estudo (AL-SHABI, 2021; SEN et al., 2022; KHEIRI; KARIMI, 2023). A seguir são mencionados alguns dos modelos de aprendizado de máquina utilizados e analisados neste estudo.

3.4.1 Prophet

O Prophet é um modelo de previsão de séries temporais desenvolvido pelo Facebook. Foi projetado para simplificar a previsão de séries temporais que apresentam padrões sazonais, tendências e feriados. O Prophet é útil para usuários que desejam realizar previsões sem requerer um profundo conhecimento em estatística ou aprendizado de máquina (STEFENON et al., 2023). O modelo se baseia em uma abordagem aditiva que desagrega a série temporal em vários componentes individuais, como tendência de longo prazo, sazonalidade semanal e anual, e efeitos de feriados. Esses componentes são combinados para formar uma previsão geral. A equação básica do modelo Prophet pode ser representada da seguinte forma:

$$p(t) = g(t) + s(t) + h(t) + \varepsilon_t \quad (3.7)$$

onde $p(t)$ é o valor da série temporal no tempo t , que se deseja prever, $g(t)$ representa a tendência de longo prazo da série, $s(t)$ representa os componentes sazonais, que podem incluir padrões semanais e anuais, e $h(t)$ é a representação dos efeitos de feriados ou eventos especiais.

O modelo Prophet ajusta esses componentes aos dados históricos de séries temporais para criar uma previsão futura. Ele utiliza um procedimento de ajuste automático para estimar os parâmetros desses componentes com base nos dados fornecidos. A abordagem aditiva do Prophet permite que os padrões sazonais, tendências e feriados sejam capturados separadamente e, em seguida, somados para gerar uma previsão global (KULSHRESHTHA; VIJAYALAKSHMI, 2020).

3.4.2 Regressão Linear

A regressão linear é classificada como um modelo de aprendizado de máquina supervisionado. Essa classificação é baseada em sua definição, que pode ser descrita por:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p + \varepsilon \quad (3.8)$$

onde há p variáveis explicativas, denotadas por x . Existe uma variável alvo, denotada por y . O valor de y é calculado como uma constante β_0 , somada aos valores das variáveis x multiplicados por seus coeficientes β_1 a β_p .

Para utilizar a regressão linear (KORSTANJE, 2021), é necessário estimar os coeficientes beta com base em um conjunto de dados de treinamento. Esses coeficientes podem ser estimados por meio de,

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (3.9)$$

onde $\hat{\beta}$ é um vetor de coeficientes estimados que minimiza a soma dos quadrados dos resíduos no método de mínimos quadrados ordinários *Ordinary Least Squares method* (OLS). Cada $\hat{\beta}_i$ representa o coeficiente estimado para a variável independente X_i ; X é a matriz de dados independentes, onde cada coluna representa uma variável independente diferente e cada linha representa uma observação separada; resultando no vetor de coeficientes.

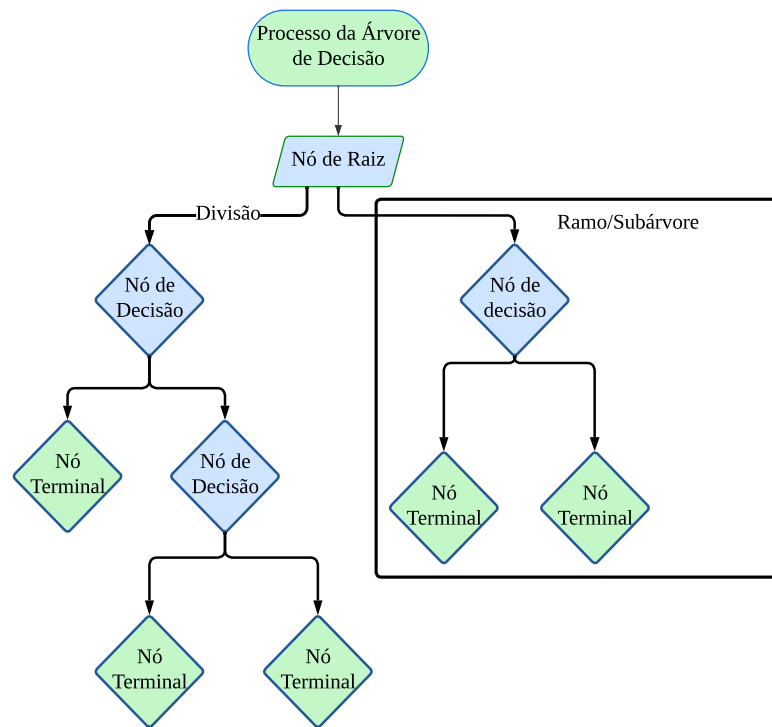
3.4.3 Árvore de Decisão

Uma árvore de decisão é um dos modelos de aprendizado de máquina frequentemente utilizados para resolver problemas de regressão e classificação. Como o nome sugere, o algoritmo utiliza um modelo de decisões semelhante a uma árvore para prever o

valor de destino (regressão) ou determinar a classe de destino (classificação) (SHI et al., 2023). É importante se familiarizar com as terminologias básicas associadas a uma árvore de decisão (Singh Kushwah et al., 2022).

Na Figura 5 tem-se o nó raiz que é o nó mais alto da árvore que representa todos os pontos de dados. A divisão refere-se à criação de um nó em dois ou mais sub-nós. O nó de decisão são os nós que são divididos em sub-nós, ou seja, esse nó que é dividido é chamado de nó de decisão. O nó folha/terminal são os nós que não se dividem. Esses nós estão geralmente no final da árvore. O ramo/subárvore é uma subseção de toda a árvore e é chamada de galho ou subárvore. O nó pai e filho é um nó, que é dividido em sub-nós e é chamado de um nó pai de sub-nós, enquanto sub-nós são os filhos do nó pai. Na Figura 5, o nó de decisão é o pai dos nós terminais (filhos). A poda é a remoção de sub-nós de um nó de decisão. A poda costuma ser feita em árvores de decisão para evitar o *overfitting* (READER, 2023).

Figura 5: Fluxograma da árvore de decisão.



A árvore de decisão pode ser mais robusta que modelo de regressão linear, tendo a capacidade de otimizar os parâmetros para trabalhar com horizontes de tempo mais longos. Vale destacar novamente, que a árvore de decisão representa um conjunto de regras de decisão hierárquicas, organizadas na forma de uma árvore. Cada nó na árvore representa uma decisão ou teste sobre um atributo, e cada ramo representa um possível resultado dessa decisão (GIFFORD; BAYRAK, 2023). Observa-se analisando a Figura

5 que repetir exatamente a mesma árvore de decisão várias vezes não adiciona valor significativo em comparação com o uso dessa árvore de decisão apenas uma vez. Em modelos de conjunto, é crucial que cada modelo individual apresente pequenas variações em relação aos demais.

Cada nó na árvore representa uma decisão ou teste sobre um atributo, e cada ramo representa um possível resultado dessa decisão. Considerando a Figura 5 como exemplo, a estrutura da árvore pode ser representada matematicamente da seguinte forma:

$$F_{DT}(x) = \begin{cases} f_1(x) & \text{se } x \text{ pertence à região do Nó 1} \\ f_2(x) & \text{se } x \text{ pertence à região do Nó 2} \\ \vdots & \\ f_k(x) & \text{se } x \text{ pertence à região do Nó } k \end{cases}$$

onde $F_{DT}(x)$ é a função de previsão do modelo de Árvore de Decisão para a instância x , k representa o número total de folhas na árvore, $f_i(x)$ é a previsão associada à i -ésima folha, determinada pela sequência de testes nos nós da árvore.

Cada teste nos nós da árvore compara um atributo específico de x com um valor de corte, decidindo qual ramo da árvore seguir com base no resultado. Este processo continua até que x chegue a uma folha, e a previsão associada a essa folha é atribuída a $F(x)$. Vale destacar que a robustez da árvore de decisão permite otimizar os parâmetros para lidar com horizontes de tempo mais longos, sendo uma alternativa mais flexível em comparação com modelos de regressão linear.

3.4.4 Floresta Aleatória

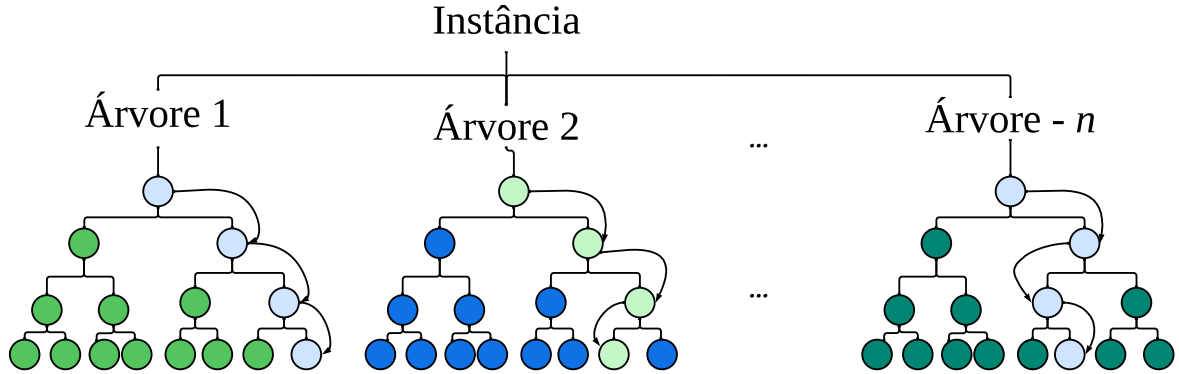
Dois modelos de previsão amplamente reconhecidos para criar conjuntos são *bagging* e *boosting*. A floresta aleatória RF utiliza o ensacamento para criar um conjunto de árvores de decisão, onde cada árvore é construída com uma amostra aleatória do conjunto de dados original. Isso assegura que as árvores sejam distintas e diversificadas, contribuindo para a robustez e eficácia do modelo (SEMAN et al., 2023).

Cada árvore da RF é construída por meio de um algoritmo de aprendizado individual que divide o conjunto de variáveis de entrada em subconjuntos, com base em um teste de valor de atributo, como o coeficiente de Gini. Ao contrário das árvores de decisão clássicas, as árvores do modelo RF são construídas sem poda e selecionam aleatoriamente um subconjunto de variáveis de entrada em cada nó. Atualmente, o número de variáveis utilizadas para dividir um nó em uma RF, denotado por m , corresponde à raiz quadrada do número total de variáveis de entrada.

Essa abordagem ajuda a aumentar a diversidade das árvores e aprimorar o desem-

penho do modelo (PELLETIER et al., 2016). Na Figura 6, um fluxograma do modelo RF ilustra como as árvores funcionam. Na construção da próxima árvore, os dois processos anteriores se repetirão, levando à criação de uma nova árvore. Provavelmente, essa árvore será diferente da primeira, pois tanto na seleção das amostras quanto na seleção das variáveis, o processo ocorre de maneira aleatória.

Figura 6: Fluxograma da floresta aleatória.



A opção pelo modelo RF em detrimento do modelo árvore de decisão é motivada por várias vantagens em termos de desempenho, robustez e generalização. Ele se destaca por apresentar uma redução significativa do *overfitting*, garantindo uma abordagem mais estável e resistente a variações nos dados. Além disso, a capacidade do RF de lidar com um grande número de características e determinar a importância relativa de cada uma contribui para uma melhor compreensão e interpretação do conjunto de dados. Essas características fazem do RF uma escolha valiosa em contextos onde é crucial alcançar resultados mais precisos e gerais.

Seja $F_{\text{RF}}(x)$ a função de previsão da Floresta Aleatória para a instância x , e $T_i(x)$ a previsão da i -ésima árvore na floresta. Então, a previsão final da Floresta Aleatória é dada por:

$$F_{\text{RF}}(x) = \frac{1}{N} \sum_{i=1}^N T_i(x) \quad (3.10)$$

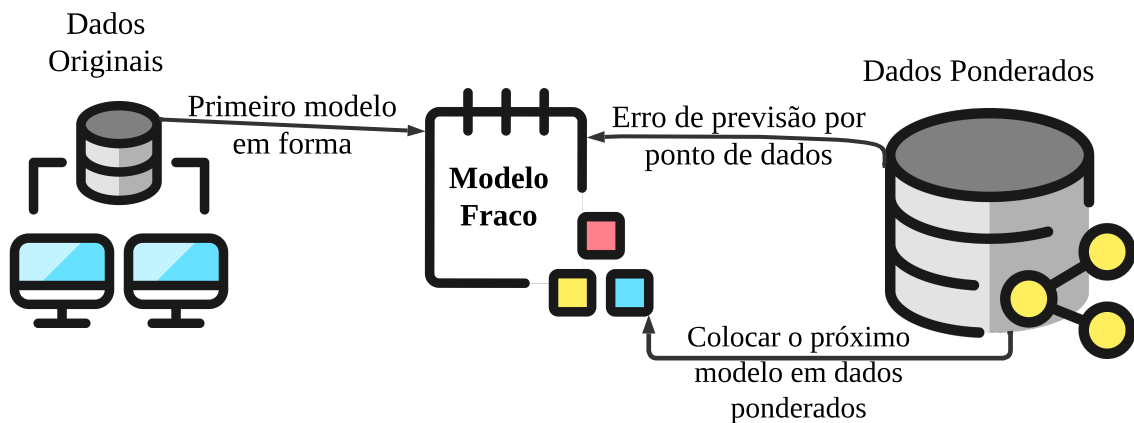
onde N é o número total de árvores na floresta. Cada árvore $T_i(x)$ é construída com uma amostra aleatória do conjunto de dados original, garantindo a diversidade e robustez do modelo. A previsão de cada árvore é combinada por meio da média, proporcionando uma abordagem de conjunto para melhorar a precisão e generalização do modelo.

3.4.5 Gradient Boosting

O *Gradient Boosting* é um modelo que combina vários modelos de árvore de decisão para realizar previsões. Cada uma dessas árvores de decisão é única, pois a diversidade é um elemento importante nesse processo. A diversidade é alcançada através de um processo chamado *boosting*, que é uma abordagem iterativa, que adiciona modelos fracos ao conjunto de forma inteligente, dando peso maior aos pontos de dados que ainda não foram previstos de forma adequada (BUEECHI et al., 2023).

A Figura 7 apresenta uma visão esquemática do modelo XGBoost. À medida que novos modelos fracos são adicionados, todos os modelos fracos intermediários são mantidos. O modelo final é uma combinação de todos esses modelos fracos, resultando em um ensemble que oferece uma melhor capacidade de previsão do que um único modelo. Neste contexto, tem-se que

Figura 7: Fluxograma do XGBoost.



$$G(x) = \sum_{i=1}^T f_t(x) \quad (3.11)$$

onde $G(x)$ é a função de previsão final do modelo, T é o número total de árvores de decisão no modelo, e $f_t(x)$ representa a saída da árvore de decisão t para a instância x .

O processo de treinamento do XGBoost envolve a minimização de uma função de perda regularizada, que incorpora termos para penalizar a complexidade do modelo e reduzir o *overfitting*. Isso é realizado por meio de um processo iterativo, onde novas árvores são adicionadas ao modelo, e a cada iteração, a função de perda é otimizada. A diversidade entre as árvores é garantida pela atribuição de pesos diferentes a cada uma, considerando seus erros residuais, tal que

$$\text{Objetivo} = \sum_{i=1}^n L(y_i, \hat{y}_i) + \sum_{k=1}^T \Omega(f_k) \quad (3.12)$$

onde $L(y_i, \hat{y}_i)$ é a função de perda que mede a discrepância entre a previsão \hat{y}_i e o rótulo verdadeiro y_i , e $\Omega(f_k)$ é um termo de regularização que penaliza a complexidade da k -ésima árvore, a função objetivo é uma combinação da função de perda e termos de regularização.

Esse processo de otimização é efetuado usando técnicas como *Gradient Boosting*, que ajusta os modelos fracos sequencialmente para melhorar a precisão global do modelo.

3.4.6 *LightGBM*

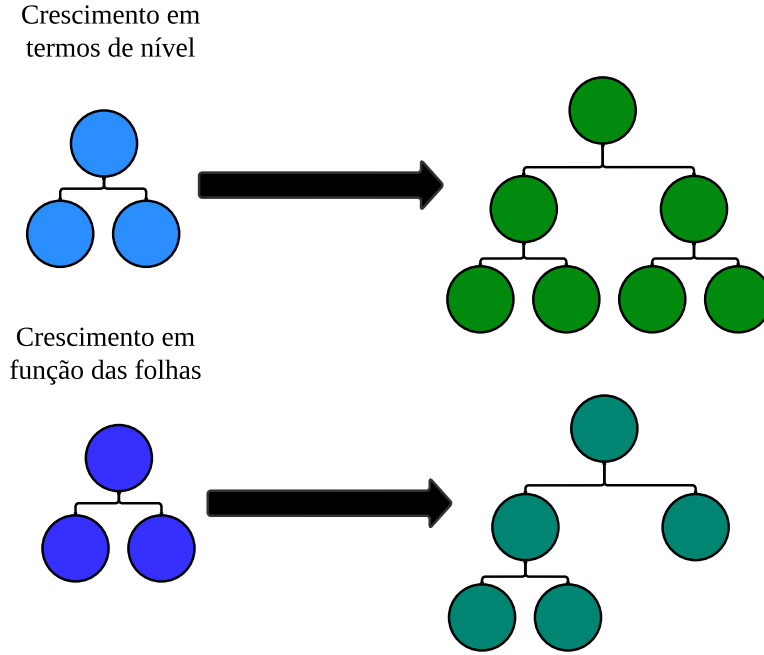
Uma alternativa proposta pelo XGBoost é a segmentação baseada em histograma. Nesse caso, em vez de iterar por todas as partições possíveis, o modelo constrói um histograma para cada variável e utiliza-os para encontrar a melhor divisão geral entre as variáveis. O LightGBM, desenvolvido pela Microsoft, adota uma abordagem eficiente para a definição das divisões. Essa abordagem é conhecida como amostragem *Gradient-Based One-Side Sample* (GOSS). O GOSS calcula o gradiente para cada ponto de dados e utiliza-o para filtrar os pontos de dados com gradientes baixos (SUN; LIU; SIMA, 2020).

O LightGBM também utiliza uma abordagem denominada *Exclusive Feature Bundling* (EFB), que acelera a seleção de variáveis correlacionadas. Outra diferença é que o modelo LightGBM é adequado para o crescimento de folhas, enquanto o XGBoost cultiva as árvores em níveis. Essa diferença pode ser observada na Figura 8 (YE; ZHAO; DENG, 2023). Essa diferença teoricamente favorece o LightGBM em termos de precisão, mas também apresenta um maior risco de sobre-ajuste quando há poucos dados disponíveis.

Na Figura 8, é possível visualizar como cada modelo é ajustado durante o processo de crescimento de árvore em folhas e em níveis. Essa representação gráfica oferece uma compreensão visual das diferenças entre os modelos. A Figura 8, apresenta um diagrama que ilustra o crescimento de uma árvore em termos de níveis, usando modelo XGBoost, e crescimento de uma árvore em termos de folhas, usando modelo LightGBM.

No crescimento de árvore em folhas, no modelo LightGBM, novas folhas são adicionadas à árvore de forma iterativa, visando maximizar a redução do erro de treinamento. Isso significa que as árvores são expandidas adicionando folhas, uma a uma, até que o critério de parada seja alcançado. No crescimento em níveis, no modelo XGBoost, as árvores são expandidas em profundidade de forma simultânea em todos os níveis. Ou seja, em cada nível, todas as folhas são expandidas ao mesmo tempo, resultando em um

Figura 8: Comparação do crescimento em folha com o crescimento em nível.



crescimento mais uniforme da árvore. Essa distinção no modo de crescimento das árvores pode afetar o comportamento e o desempenho do modelo.

XGBoost: A função de previsão do modelo XGBoost para a instância x é representada por $F_{\text{XGBoost}}(x)$. Esta função é obtida através do crescimento de árvores em níveis, onde cada árvore é construída simultaneamente em profundidade em todos os níveis. Assim, a previsão final é dada por:

$$F_{\text{XGBoost}}(x) = \sum_{i=1}^N T_i(x) \quad (3.13)$$

onde N é o número total de árvores na floresta, e $T_i(x)$ representa a previsão da i -ésima árvore.

LightGBM: A função de previsão do modelo LightGBM para a instância x é representada por $F_{\text{LightGBM}}(x)$. A abordagem do LightGBM inclui a amostragem *Gradient-Based One-Side Sample* (GOSS) para filtrar pontos de dados com gradientes baixos, e o *Exclusive Feature Bundling* (EFB) para acelerar a seleção de variáveis correlacionadas. A previsão final é dada por:

$$F_{\text{LightGBM}}(x) = \sum_{i=1}^N T_i(x) \quad (3.14)$$

onde N é o número total de árvores na floresta, e $T_i(x)$ representa a previsão da i -ésima árvore. No LightGBM, as árvores são cultivadas adicionando folhas iterativamente para maximizar a redução do erro de treinamento, resultando em um crescimento em folhas. Isso contrasta com o XGBoost, que expande as árvores em profundidade simultaneamente em todos os níveis, resultando em um crescimento mais uniforme da árvore.

3.5 Redes Neurais Artificiais

Uma rede neural é um modelo de processamento de informações inspirado pelo funcionamento do cérebro humano. Consiste em um conjunto interconectado de unidades de processamento, conhecidas como neurônios artificiais, que trabalham em conjunto para realizar tarefas de aprendizado a partir de dados (XIANG et al., 2018). Assim como os neurônios no cérebro estão interligados por sinapses, os neurônios artificiais são conectados por conexões ponderadas. Essas conexões permitem que a rede neural analise padrões complexos nos dados, reconhecendo relações e características importantes para executar tarefas como classificação, previsão, reconhecimento de padrões (BABU; REDDY, 2014). Conforme a rede é exposta a exemplos e informações, ela ajusta suas conexões para melhorar seu desempenho, tornando-a capaz de generalizar e lidar com novos dados (RAO; ZHAO; DENG, 2020).

3.5.1 MLP

Uma Rede Neural MLP é um tipo de arquitetura de rede neural artificial composta por várias camadas de neurônios (ou unidades) organizados em uma estrutura de múltiplas camadas. Cada neurônio em uma camada está conectado a todos os neurônios da camada seguinte, sem formar ciclos (rede *feedforward*). Essa arquitetura é projetada para realizar tarefas de aprendizado supervisionado, como classificação e regressão (QIN et al., 2023).

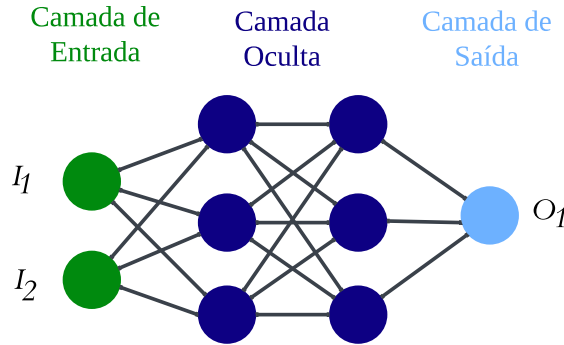
A topologia da MLP funciona como uma rede *feedforward*, rede progressiva, a saída de um neurônio se conecta com outro neurônio da próxima camada, no sentido esquerda/direita, formada por um conjunto de neurônios denominados nós, como na Figura 9. A rede possui uma camada de entrada, sem função de ativação, uma ou mais camadas ocultas e uma camada de saída. A complexidade da rede neural MLP se dá pela quantidade de camadas ocultas que houver e a quantidade de neurônios que essas camadas possuem (GRÜBLER, 2018).

$$I = [I_1, I_2] = \text{Vetor de Entrada}$$

$$O = [O_1] = \text{Vetor de Saída}$$

O modelo de rede neural artificial MLP é dado por,

Figura 9: Modelo de uma rede neural artificial MLP.



$$v_j = \sum_{i=0}^m w_i y_i + b \quad (3.15)$$

o funcionamento geral de uma rede MLP está representada na Figura 9. Cada neurônio recebe todos os valores das entradas, representadas pelo símbolo y , que são multiplicadas pelos pesos sinápticos simbolizados pelo w e somadas entre si junto com uma constante chamada de polarização ou bias, representada pelo símbolo b .

3.5.2 Rede Neural Recorrente

Uma Rede Neural Recorrente RNN é um tipo de arquitetura de rede neural que pode ser utilizada para usar dados sequenciais ou temporais (NASIRI; EBADZADEH, 2023). Ao contrário das redes neurais convencionais, onde as entradas e saídas são tratadas como dados independentes, as RNNs levam em consideração a ordem e a relação entre os elementos em uma sequência, tornando-as ideais para lidar com dados como séries temporais.

A característica principal das RNNs é que elas contêm laços em sua estrutura, permitindo que informações anteriores influenciem o processamento de informações subsequentes. Isso significa que a saída em um determinado passo de tempo não depende apenas da entrada atual, mas também das entradas anteriores na sequência.

$$h_t = f(W_{hh} \cdot h_{t-1} + W_{xh} \cdot x_t + b_h) \quad (3.16)$$

onde h_t é o estado oculto (ou saída) no tempo t , h_{t-1} é o estado oculto anterior no tempo $t - 1$, x_t é a entrada no tempo t , W_{hh} é a matriz de pesos que controla a influência do estado oculto anterior, W_{xh} é a matriz de pesos que controla a influência da entrada, b_h é

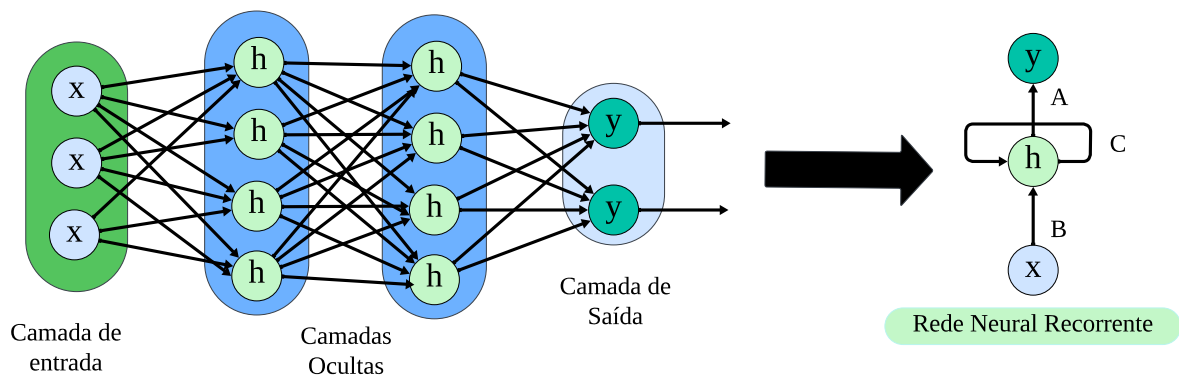
o vetor de viés, f é uma função de ativação, frequentemente a função tangente hiperbólica (tanh) ou a função sigmoide (TAM, 2023).

A equação representa a propagação do estado oculto ao longo do tempo em uma RNN. A cada novo passo de tempo, a RNN considera a entrada atual x_t e o estado oculto anterior h_{t-1} , calculando o novo estado oculto h_t usando as matrizes de pesos e a função de ativação.

No entanto, as RNNs tradicionais podem enfrentar dificuldades em capturar dependências de longo prazo, devido ao problema de dissipação do gradiente. Para entender isso, surgiram variações avançadas, como LSTM e GRU, que incorporam mecanismos de aprendizado de esquecimento e controle de informação, permitindo que informações relevantes sejam mantidas por períodos mais longos de tempo (WANG; YING, 2023), (ZHAO et al., 2023).

Como pode ser visto na Figura 10, a grande diferença da RNN é que há um laço de reações. Enquanto cada entrada de uma rede totalmente conectada é completamente independente, as entradas de uma RNN têm uma relação de realimentação entre si. Isso faz com que ela seja capaz de capturar padrões em dados sequenciais de uma maneira que redes neurais tradicionais talvez não consigam.

Figura 10: Fluxograma da RNN.



3.6 Aprendizado Profundo

Aprendizado Profundo ou *Deep Learning* (DL) refere-se a um subcampo do aprendizado de máquina que envolve a construção e o treinamento de modelos de rede neural profunda. O termo profundo se refere ao uso de arquiteturas de modelos que consistem em várias camadas, conhecidas como redes neurais profundas (KOTHONA et al., 2023).

3.6.1 LSTM

As redes neurais LSTMs são uma evolução das RNNs, projetadas para superar desafios na captura de dependências de longo prazo em sequências de dados. Diferentemente das RNNs convencionais, as LSTMs têm a capacidade de manter informações relevantes por longos períodos, tornando-as especialmente eficazes em tarefas que envolvem padrões complexos e dependências temporais distantes (ZHANG, 2021).

Uma das principais inovações das LSTMs é a introdução de unidades de memória chamadas células, que possuem três componentes principais: uma porta de entrada (*input gate*), uma porta de esquecimento (*forget gate*) e uma porta de saída (*output gate*). Essas portas permitem que as LSTMs controlem o fluxo de informações através da célula, decidindo quais informações devem ser mantidas, esquecidas ou passadas para a saída (ZHANG, 2021).

$$f_t = \sigma(W_{xf} \cdot x_t + W_{hf} \cdot h_{t-1} + b_f) \quad (3.17)$$

$$i_t = \sigma(W_{xi} \cdot x_t + W_{hi} \cdot h_{t-1} + b_i) \quad (3.18)$$

$$\tilde{C}_t = \tanh(W_{xc} \cdot x_t + W_{hc} \cdot h_{t-1} + b_c) \quad (3.19)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (3.20)$$

$$o_t = \sigma(W_{xo} \cdot x_t + W_{ho} \cdot h_{t-1} + b_o) \quad (3.21)$$

$$h_t = o_t \odot \tanh(C_t) \quad (3.22)$$

onde x_t é a entrada no tempo t , h_{t-1} é o estado oculto anterior no tempo $t - 1$, f_t é o valor da porta de esquecimento, i_t é o valor da porta de entrada, \tilde{C}_t é o candidato a novo estado de memória, C_t é o novo estado de memória, o_t é o valor da porta de saída, h_t é o novo estado oculto (saída) no tempo t , σ é a função de ativação sigmoide, \odot representa a multiplicação elemento a elemento.

Essa estrutura permite que as LSTMs controlem o fluxo de informações e aprendam a armazenar ou descartar informações relevantes para diferentes tarefas. As portas de entrada, esquecimento e saída funcionam como mecanismos de controle, permitindo que as LSTMs aprendam a manter informações importantes, esquecer informações desnecessárias e gerar saídas precisas ao longo de sequências temporais.

3.6.2 GRU

Uma rede neural GRU é um tipo de arquitetura de RNN que foi projetado para trabalhar com o problema de dissipação de gradiente e captura de dependências de longo prazo em sequências de dados. Essa variação das RNNs tradicionais introduz mecanismos

para controlar o fluxo de informação por meio das unidades de tempo. A GRU é uma alternativa vantajosa para a análise de séries temporais, devido à sua habilidade de lidar com sequências de dados de extensões variáveis e de capturar dependências de longo prazo presentes em informações sequenciais. Além disso, a GRU apresenta uma estrutura de simplicidade superior à LSTM, permitindo um processo de treinamento ágil (MIGLIATO; PONTI, 2021).

A estrutura da GRU inclui dois portões principais: o portão de atualização (*update gate*) e o portão de reinicialização (*reset gate*). Esses portões permitem que o GRU decida quais informações serão transmitidas para a próxima etapa de tempo e quais informações serão descartadas, tal que h_t representa o estado oculto na etapa de tempo t , h_{t-1} é o estado oculto na etapa de tempo anterior $t - 1$, x_t é a entrada na etapa de tempo t , r_t é o valor do portão de reinicialização na etapa t , z_t é o valor do portão de atualização na etapa t , \odot denota a multiplicação elemento a elemento, σ é a função sigmoid, que retorna valores entre 0 e 1, \tanh é a função tangente hiperbólica, que retorna valores entre -1 e 1 , W_r , W_z e W_h são matrizes de pesos que o modelo aprende durante o treinamento.

O portão de reinicialização (r_t) controla a quantidade de informação do passado a ser esquecida, dada por:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (3.23)$$

O portão de atualização (z_t) controla a quantidade de informação do passado a ser passada para o próximo estado, como:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (3.24)$$

A ativação do candidato (\tilde{h}_t), candidato a novo estado oculto calculado por:

$$\tilde{h}_t = \tanh(W_h \cdot [r_t \odot h_{t-1}, x_t]) \quad (3.25)$$

O novo estado oculto (h_t) é a combinação ponderada do estado anterior e do novo candidato, dado por:

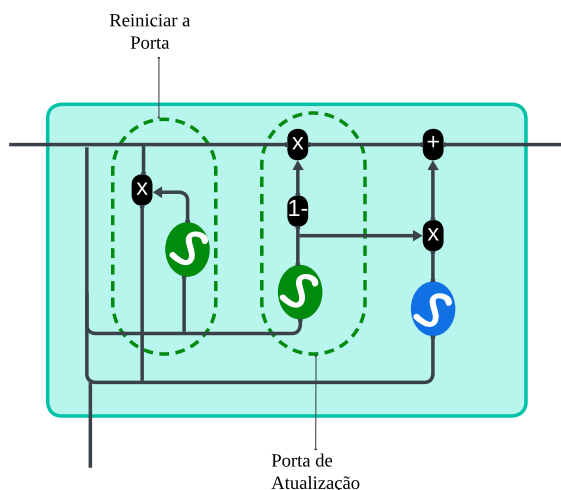
$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (3.26)$$

Na Figura 11 é representado um diagrama de um modelo de GRU para análise

de séries temporais. O modelo GRU é um tipo de RNN que possui dois portões: um portão de atualização e um portão de reinicialização. Esses portões controlam como a informação é armazenada e atualizada na memória oculta da rede. Um modelo GRU é capaz de aprender padrões temporais complexos e dependências de longo prazo nos dados sequenciais.

A Figura 11 apresenta uma representação simplificada do modelo com três portões: o portão de reinicialização, o portão de atualização e o portão de saída. Os portões são interconectados por linhas tracejadas, representando o fluxo de informação entre eles. O diagrama está rotulado em português, com Porta de Reinicialização, Porta de Atualização, e Porta de Saída (SARANYA; SIVAKUMAR, 2020; JORDAN; SOKÓŁ; PARK, 2021; KHAN et al., 2022).

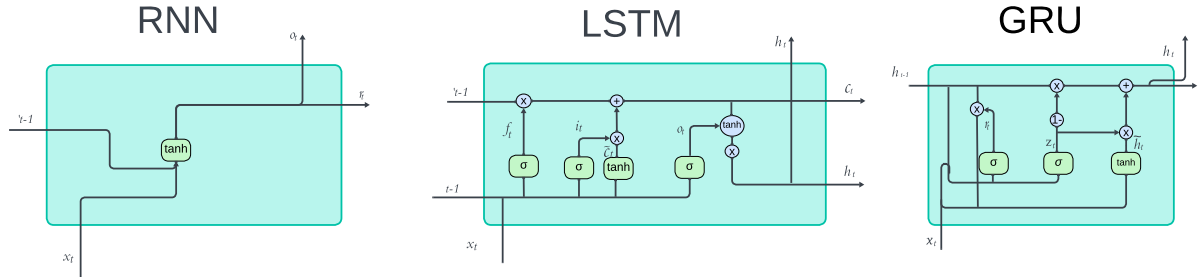
Figura 11: Diagrama do funcionamento de uma GRU.



As redes neurais GRUs, LSTMs, e RNNs apresentam variações em suas arquiteturas, todas projetadas para abordar a dificuldade de capturar dependências temporais em sequências de dados. Enquanto as RNNs tradicionais têm uma tendência a sofrer com o desvanecimento do gradiente ao longo do tempo, as LSTMs e GRUs foram desenvolvidas para superar essa limitação. As LSTMs introduzem células de memória e portas de controle que permitem armazenar e atualizar informações relevantes ao longo das etapas temporais, sendo especialmente adequadas para capturar relações de dependência de longo prazo. As GRUs, por sua vez, simplificam a arquitetura das LSTMs, utilizando portas de atualização para permitir o fluxo de informações e controle sobre o estado oculto. A Figura 12 ilustra as RNNs, LSTMs, e GRUs, permitindo uma visualização das diferenças entre essas redes neurais. Cada seção possui um diagrama da arquitetura da rede, com nós representando neurônios e arestas representando conexões entre neurônios. A RNN tem um único neurônio recorrente, a LSTM tem vários neurônios recorrentes com cone-

xões adicionais que formam portões e células de memória, e a GRU tem dois portões que controlam o fluxo de informação na memória oculta da rede.

Figura 12: Diferenças entre RNN, LSTM, e GRU.



As LSTMs e GRUs oferecem soluções mais sofisticadas em relação às RNNs tradicionais, apresentando mecanismos que permitem capturar dependências de longo prazo de maneira mais eficaz.

3.7 Rede Neural Convolucional

As Redes Neurais Convolucionais CNN são um tipo de rede neural que utiliza a operação de convolução em vez da multiplicação por matrizes em ao menos uma de suas camadas. Esse tipo de rede é efetiva em aplicações em que os dados são dispostos de forma que a relação de vizinhança entre os elementos é relevante, no caso de séries temporais, que são sequências unidimensionais de dados amostrados em intervalos de tempo regulares tem-se este tipo de característica (SILVA, 2021), (REICHMAN; MALOF; COLLINS, 2016) .

A camada convolucional tem como objetivo extrair as características mais importantes da entrada. Dessa forma, sua saída é um mapa de características obtido a partir da convolução da entrada com um *kernel* aprendido, seguido da aplicação de uma função de ativação não linear (LUCAS, 2019). Os mapas de características completos são obtidos por,

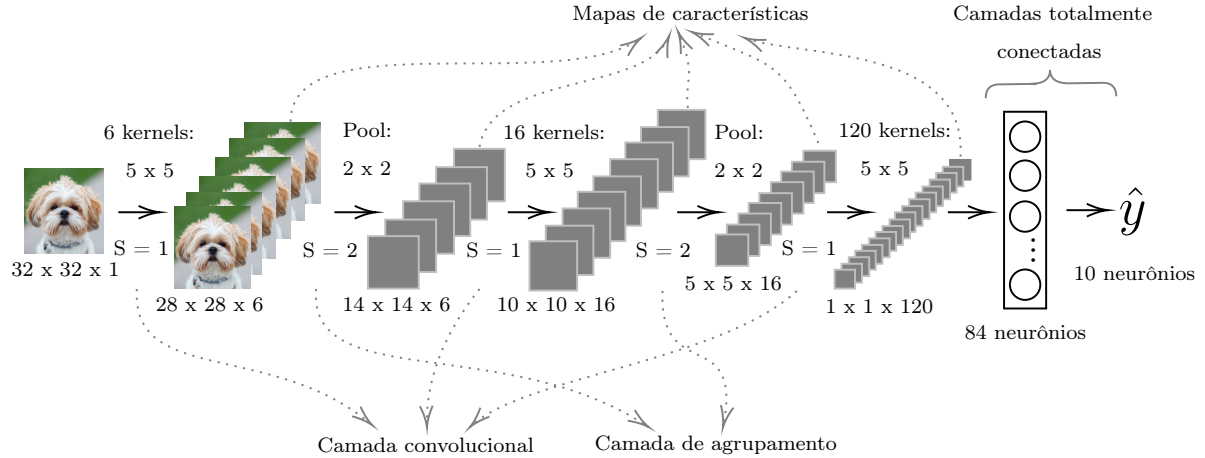
$$Z_{i,j,k}^L = W_k^L \cdot X_{i,j}^L + b_k^L \quad (3.27)$$

onde $Z_{i,j,k}^L$ é o mapa de características obtido pela convolução do k -ésimo filtro da L -ésima camada com a célula de entrada centrada na localização (i, j) , W_k^L vetor de pesos do k -ésimo filtro da L -ésima camada, b_k^L termo de polarização do k -ésimo filtro da L -ésima camada, $X_{i,j}^L$ é a célula de entrada centrada na localização (i, j) da L -ésima camada.

A profundidade dos mapas de características é dada pelo número de *kernels* (ou

filtros) de convolução. Observe na Figura 13 que a 1ª camada de convolução com 6 *kernels* gera uma saída de profundidade 6. Isso porque, cada *kernel* possui pesos diferentes para extrair diferentes características da entrada (LUCAS, 2019).

Figura 13: Modelo de uma Rede Neural Convolucional.



Uma vantagem das camadas de convolução é o compartilhamento do vetor de pesos para toda a circunvolução na construção de um mapa de características, pois reduz o número de parâmetros na rede, resultando em treinamento e previsões mais eficientes (LUCAS, 2019). A largura e a altura desses mapas são definidas pelo tamanho do *kernel* e do *stride* (passo da circunvolução) dado por,

$$T_{\text{map}} = \left(\frac{I - F}{S + 1} \right) \quad (3.28)$$

onde T_{map} é a altura ou largura do mapa de características, I é a altura ou largura da entrada, F é a altura ou largura do *kernel* de convolução, S é o tamanho do *stride*.

3.8 Abordagem de Otimização

A utilização do *Tree-structured Parzen Estimator* (TPE) na otimização de hiperparâmetros é crucial para melhorar o desempenho dos modelos. Integrado à biblioteca Optuna do Python, ele desempenha um papel vital ao guiar a busca eficiente por conjuntos ideais de hiperparâmetros durante a otimização automática (HVY, 2020).

Ao dividir o espaço de busca, o TPE destaca-se ao identificar áreas promissoras para exploração, direcionando a análise de maneira inteligente e eficiente. Essa capacidade de focar em regiões promissoras contribui significativamente para a melhoria do desempenho

do modelo.

No contexto da demanda d'água, a otimização de modelos é essencial para prever e gerenciar eficientemente o consumo d'água. A correlação aqui reside na importância de ajustar os hiperparâmetros do modelo para otimizar sua precisão na previsão da demanda d'água. O TPE, ao guiar a busca por configurações ideais, contribui para modelos mais precisos e adaptáveis (HVY, 2020).

A equação fundamental do TPE, ao maximizar a probabilidade de melhoria, reflete o compromisso em encontrar conjuntos ótimos de hiperparâmetros. Essa abordagem probabilística, aplicada à otimização, permite ajustes dinâmicos ao longo do tempo, garantindo que o modelo se adapte às mudanças nas demandas da previsão d'água.

$$P(x | y) = \begin{cases} p(x), & \text{se } y > y^* \\ q(x), & \text{se } y \leq y^* \end{cases} \quad (3.29)$$

onde x representa um conjunto de hiperparâmetros, y é o valor da função objetivo associado a esses hiperparâmetros, y^* é o valor de referência para a melhoria, $p(x)$ é a probabilidade de x ser um conjunto bom de hiperparâmetros, $q(x)$ é a probabilidade de x ser um conjunto ruim de hiperparâmetros.

Assim, ao considerar a otimização de hiperparâmetros com o TPE, não apenas aprimoramos a eficiência dos modelos, mas também fortalecemos sua capacidade de contribuir para soluções mais precisas e adaptáveis em domínios cruciais, como o gerenciamento da demanda d'água.

3.9 Decomposição STL

A decomposição STL é uma técnica amplamente utilizada para decompor séries temporais em seus componentes sazonais, de tendência e restantes (RIBEIRO et al., 2023). O método STL realiza a decomposição aditiva dos dados por meio de uma sequência de aplicações do Loess mais suave, onde regressões polinomiais ponderadas localmente são aplicadas em cada amostra do conjunto de dados, tendo como variáveis explicativas os valores próximos do amostra cuja resposta está sendo estimada (THEODOSIOU, 2011).

Ao aplicar a decomposição STL, a série temporal pode ser expressa como a soma dos componentes sazonais, de tendência e restantes. Essa técnica se revela valiosa para análise e modelagem de séries temporais, fornecendo uma compreensão clara dos padrões de variação presentes nos dados. Este método prepara o terreno para uma abordagem mais refinada na aplicação de modelos, permitindo uma interpretação mais precisa e uma melhor adaptação aos padrões intrínsecos dos dados ao utilizar o ARIMA ou SARIMA em sua análise (RIBEIRO et al., 2023).

A decomposição STL é formalmente definida como:

$$s_t = f(S_t, T_t, R_t) = \begin{cases} y_t = S_t + T_t + R_t & \text{modelo aditivo} \\ y_t = S_t T_t R_t & \text{modelo multiplicativo} \end{cases} \quad (3.30)$$

onde s_t é o valor da série temporal no tempo t , T_t é a componente de tendência no tempo t , S_t é a componente de sazonalidade no tempo t , R_t é a componente de resíduo no tempo t .

3.10 Abordagens de Correlações

A correlação é uma medida estatística que descreve a relação entre duas variáveis. Ela indica o grau e a direção da associação entre essas variáveis. Quando duas variáveis estão correlacionadas, significa que elas variam juntas de alguma maneira específica. A correlação pode ser positiva, quando as variáveis aumentam ou diminuem juntas, ou negativa, quando uma variável aumenta enquanto a outra diminui. Por exemplo, se houver uma correlação positiva entre o número de horas de estudo e as notas em um exame, isso significa que, em geral, quanto mais uma pessoa estuda, mais alta é sua nota. A correlação pode ser quantificada através de coeficientes de correlação.

3.10.1 Autocorrelação e Autocorrelação Parcial

A representação gráfica da Função de Autocorrelação (ACF) e da Função de Autocorrelação Parcial (PACF) é uma ferramenta essencial na análise de séries temporais, especialmente ao trabalhar com modelos ARIMA. Ela é expressa da seguinte forma:

$$R_k = \frac{\gamma_k}{\gamma_0} \quad (3.31)$$

$$\phi_{kk} = \frac{\gamma_{kk} - \sum_{j=1}^{k-1} \phi_{k-1,j} \gamma_{k-j,k}}{\gamma_{0,k}} \quad (3.32)$$

onde R_k é a autocorrelação de ordem k , γ_k é a função de autocovariância de ordem k , ϕ_{kk} é a autocorrelação parcial de ordem k , γ_{kk} é a autocovariância de ordem k , $\gamma_{0,k}$ é a variância amostral no tempo k , e $\phi_{k-1,j}$ é a autocorrelação parcial de ordem j na equação da PACF.

O ACF é uma medida estatística utilizada para identificar a presença de correlação serial em uma série temporal. Ele calcula a autocorrelação entre os valores da série em diferentes defasagens, ou seja, a correlação entre os valores atuais e os valores passados

da série. O ACF é útil para analisar a dependência temporal dos dados e identificar padrões de sazonalidade, tendência ou outros efeitos temporais. Por meio do ACF, é possível avaliar se a série exibe autocorrelação significativa em defasagens específicas, o que pode indicar a presença de não estacionariedade ou estrutura temporal que precisa ser considerada na análise ou modelagem da série temporal.

Enquanto o ACF mede a correlação total em um determinado lag, o PACF mede a correlação apenas entre a observação atual e uma observação em um lag específico, controlando os efeitos das observações intermediárias. O gráfico PACF é especialmente útil para identificar a ordem do componente autorregressivo (AR) no modelo ARIMA. Ele ajuda a identificar os lags que têm uma influência direta na observação atual, sem a interferência de outros lags. Ao usar esses gráficos em conjunto, os analistas de séries temporais podem identificar a ordem adequada do modelo ARIMA.

A ordem do ARIMA é denotada como (p, d, q) , onde p é a ordem do componente AR, d é a ordem de diferenciação, que representa o número de vezes que a série temporal é diferenciada para torná-la estacionária, e q é a ordem do componente de MA. Os pontos nos gráficos ACF e PACF que ultrapassam as bandas de confiança indicam lags significativos. Analisando esses gráficos, pode fazer escolhas sobre os valores de p , d , e q ao ajustar um modelo ARIMA à série temporal em questão.

3.10.2 Correlação de Pearson

A correlação de Pearson é uma medida estatística que avalia a relação linear entre duas variáveis. No contexto de séries temporais, a correlação de Pearson pode ser útil para entender se existe uma relação linear entre duas séries temporais, ou entre diferentes variáveis de uma mesma série temporal (Cesar de Lima Nogueira et al., 2023). No entanto, é importante ter em mente algumas considerações ao aplicar a correlação de Pearson a séries temporais:

(i) a correlação de Pearson é sensível a tendências e sazonalidades nas séries temporais. Se houver uma tendência ou sazonalidade em ambas as séries, a correlação pode indicar uma relação mesmo que a relação real seja mais complexa.

(ii) a correlação de Pearson pressupõe uma relação linear entre as variáveis. Se a relação entre as séries temporais for não linear, a correlação de Pearson pode não capturar essa relação de maneira adequada, e

(iii) a correlação de Pearson pode ser sensível a valores extremos (*outliers*) (COELHO; AYALA; MARIANI, 2024). Valores extremos podem distorcer a medida de correlação, tornando-a menos representativa da relação geral entre as séries.

Ao usar a correlação de Pearson para análise de séries temporais, é importante

interpretar os resultados com cautela e considerar outros métodos estatísticos e gráficos para uma compreensão mais completa da relação entre as séries. A equação do coeficiente de correlação de Pearson é dada por:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{(\sum (x_i - \bar{x})^2)(\sum (y_i - \bar{y})^2)}} \quad (3.33)$$

onde x_i e y_i representam os valores das variáveis X e Y , respectivamente. \bar{x} e \bar{y} são as médias dos valores x_i e y_i .

O coeficiente de correlação de Pearson mede a força e a direção da relação linear entre as variáveis X e Y . Valores próximos a 1 indicam uma correlação positiva forte, valores próximos a -1 indicam uma correlação negativa forte, e valores próximos a 0 indicam ausência de correlação entre as variáveis.

As métricas estatísticas são utilizadas na análise de previsão de séries temporais para avaliar se o modelo preditor possui um desempenho adequado e/ou superior. Quanto menor for o erro da maioria das métricas, melhor será o desempenho do modelo aplicado.

3.11 Medidas de Desempenho

Medidas de desempenho são métricas quantitativas utilizadas para avaliar a qualidade ou eficácia de um modelo, algoritmo, sistema ou processo em relação a um determinado objetivo ou critério. Elas fornecem uma maneira de quantificar o quão bem uma determinada entidade realiza uma tarefa específica, seja ela prever valores futuros, classificar dados, otimizar processos ou tomar decisões.

A Raiz do Erro Médio Quadrático Relativo (RRMSE) é uma variante do erro quadrático médio (RMSE) sendo calculado por,

$$RRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\sum_{i=1}^n (\hat{y}_i)^2} \quad (3.34)$$

onde n número total de observações ou amostras no conjunto de dados, y_i valor real da observação i , \hat{y}_i valor previsto ou estimado da observação i pelo modelo, $\sum_{i=1}^n$ soma sobre todas as observações no conjunto de dados.

O Erro Absoluto Médio (MAE) é utilizado como uma métrica para avaliar o desempenho de modelos de previsão. Em vez de calcular a média das diferenças entre os valores reais e previstos, o MAE calcula a média dos valores absolutos dessas diferenças, garantindo que os erros positivos e negativos não se anulem, calculada por,

$$MAE = \frac{1}{n} \sum |y_i - \hat{y}_i| \quad (3.35)$$

sua interpretação é similar ao RRMSE, em que o erro é expresso na mesma escala ou ordem de grandeza da variável estudada.

O Erro Percentual Absoluto Médio Simétrico (SMAPE) é outra métrica comumente utilizada para avaliar a precisão de modelos de previsão. O SMAPE é expresso como uma porcentagem, facilitando a compreensão da precisão relativa do modelo. O SMAPE é adequado para trabalhar com valores nulos nos dados, pois a divisão por zero é evitada no cálculo da métrica. O SMAPE é sensível a valores extremos nos dados. Se houver valores discrepantes que não representem a tendência geral, eles podem influenciar significativamente a métrica. Seu cálculo é dado por,

$$SMAPE = \frac{1}{n} \sum_{i=1}^n \frac{2|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)} \times 100 \quad (3.36)$$

3.12 Testes de Hipóteses

Um teste de hipótese é uma técnica estatística utilizada para avaliar a validade de uma afirmação sobre uma população com base em amostras de dados. Em geral, o teste de hipótese envolve a formulação de duas hipóteses: a hipótese nula (H_0), que geralmente afirma que não há efeito ou diferença entre as populações, e a hipótese alternativa (H_1), que contradiz a hipótese nula e geralmente afirma que existe um efeito ou diferença.

3.12.1 Teste Run

O teste de Run, também conhecido como teste de sequências de *Wald-Wolfowitz*, verifica se uma série de dados foi gerada aleatoriamente (PAIVA; SÁFADI, 2021).

As hipóteses são:

$$\begin{cases} H_0 : \text{a sequência foi gerada aleatoriamente (sem tendência)} \\ H_1 : \text{a sequência não foi gerada aleatoriamente (possui uma tendência)}. \end{cases} \quad (3.37)$$

A estatística T_1 utiliza os valores acima e abaixo da mediana. Primeiramente, considera-se N como o número de observações em uma série. Em seguida, esses valores são ordenados e o símbolo A é atribuído quando o valor da observação é maior ou igual à mediana, $A \geq m$, e o símbolo B se o valor for menor que a mediana, $B < m$.

O número total de observações é $N = (n_1 \text{ pontos } A) + (n_2 \text{ pontos } B)$, com a estatística T_1 igual ao número de sequências ou número de execuções.

O valor de T_1 é obtido contando o número de sequências iguais, ou seja, o número de oscilações no conjunto de dados entre os valores A e B . Por exemplo, suponha que o número de observações seja igual a 20. Entre essas observações, os primeiros 5 valores eram iguais a A , os próximos 6 eram iguais a B e os outros valores eram iguais a A novamente. Então, 3 sequências (ABA) foram obtidas, ou seja, o número de sequências (ou execuções) é igual a 3.

Conforme PAIVA e SÁFADI (2021), no caso de poucas sequências em relação à série, a hipótese nula H_0 é rejeitada. Para valores de n_1 ou $n_2 > 20$, a distribuição normal é utilizada sob H_0 , ou seja, $T_1 \sim N(\mu, \sigma^2)$, onde

$$\mu = \frac{2n_1n_2}{N} + 1, \quad (3.38)$$

$$\sigma = \sqrt{\frac{2n_1n_2(2n_1n_2 - N)}{N^2(N - 1)}} \quad (3.39)$$

Por meio dos cálculos de μ e σ , pode-se encontrar o valor da estatística Z dada por:

$$Z = \frac{T_1 - \mu}{\sigma} \quad (3.40)$$

Analisando o valor de Z , H_0 é rejeitada se $|Z| > Z_{\frac{\alpha}{2}}$, sendo α o nível de significância adotado.

3.12.2 Teste Dickey-Fuller

O teste Dickey-Fuller (DF) segue uma regressão linear AR de primeira ordem e testa a hipótese nula de que uma raiz unitária está presente em uma série temporal, indicando a não estacionariedade. A forma mais comum do teste Dickey-Fuller é conhecida como *Augmented Dickey-Fuller* (ADF), que inclui termos adicionais na regressão para trabalhar com possíveis problemas de autocorrelação (AGIAKLOGLOU; NEWBOLD, 1992). O procedimento geral do teste DF envolve as seguintes etapas:

Formulação da Hipótese Nula (H_0): A hipótese nula assume a presença de raízes unitárias, o que indica não estacionariedade na série temporal. Formulação da Hipótese Alternativa (H_1): A hipótese alternativa busca rejeitar a hipótese nula, sugerindo a esta-

cionariedade da série temporal. Realização do Teste: O teste DF é realizado calculando uma estatística de teste. Se o valor de p associado à estatística de teste for menor que um determinado nível de significância pré-definido, a hipótese nula é rejeitada. Interpretação dos Resultados: Se a hipótese nula for rejeitada, isso sugere que a série temporal é estacionária. Caso contrário, a não estacionariedade da série temporal não pode ser descartada.

De acordo com o Reisen et al. (2017), o teste DF é representado por,

$$\begin{aligned}\Delta y(t) = & \alpha + \beta \cdot t + \gamma \cdot y(t-1) + \delta_1 \cdot \Delta y(t-1) + \delta_2 \cdot \Delta y(t-2) + \dots + \delta_p \\ & \Delta y(t-p) + \varepsilon(t)\end{aligned}\tag{3.41}$$

onde $\Delta y(t)$ é a diferença entre os valores consecutivos da série temporal no tempo t , $y(t-1)$ é o valor da série temporal no tempo anterior, t é uma variável de tendência temporal, α, β, γ são parâmetros a serem estimados, $\delta_1, \delta_2, \dots, \delta_p$ são os coeficientes associados às diferenças defasadas, $\varepsilon(t)$ é o termo de erro.

3.12.3 Teste de Ljung-Box

A ideia do teste de Ljung-Box é verificar se as autocorrelações dos resíduos em diferentes defasagens são estatisticamente diferentes de zero. Se a estatística de teste do Ljung-Box indicar significância estatística, isso sugere que há autocorrelações remanescentes nos resíduos, indicando que o modelo não está capturando completamente a estrutura da série temporal (BOX; PIERCE, 1970), (LJUNG; BOX, 1978), (DAVIES; NEWBOLD, 1979). As etapas básicas do teste de Ljung-Box:

A formulação da Hipótese Nula (H_0): A hipótese nula assume que não há autocorrelação nos resíduos. Formulação da Hipótese Alternativa (H_1): A hipótese alternativa sugere que há autocorrelação nos resíduos. Cálculo da Estatística de Teste: A estatística de teste de Ljung-Box é calculada usando as autocorrelações dos resíduos em várias defasagens. A fórmula é baseada na soma dos quadrados dessas autocorrelações. Distribuição da Estatística de Teste: A estatística de teste é comparada a uma distribuição qui-quadrado com um número apropriado de graus de liberdade. Isso depende do número de defasagens considerado no teste. Decisão Estatística: Se o valor de p associado à estatística de teste for menor que um nível de significância escolhido (por exemplo, 0,05), a hipótese nula é rejeitada, indicando a presença de autocorrelação nos resíduos.

A estatística de teste Box-Pierce é uma versão simplificada da estatística de Ljung-Box para a qual estudos de simulação subsequentes mostraram baixo desempenho (DAVIES; NEWBOLD, 1979). A estatística de teste é (LJUNG; BOX, 1978):

$$Q = n(n+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k} \quad (3.42)$$

onde n é o tamanho da amostra, $\hat{\rho}_k$ é a autocorrelação da amostra no lag k e h é o número de defasagens que estão sendo testadas. Debaixo H_0 a estatística Q segue assintoticamente um $\chi_{(h)}^2$. Para o nível de significância α , a região crítica para rejeição da hipótese de aleatoriedade é:

$$Q > \chi_{1-\alpha, h}^2 \quad (3.43)$$

onde $\chi_{1-\alpha, h}^2$ é o $(1 - \alpha)$ -quantil (BROCKWELL; DAVIS, 2002) da distribuição qui-quadrada com graus h de liberdade.

O teste de Ljung-Box é comumente usado na modelagem de média móvel integrada ARIMA. Note que ele é aplicado aos resíduos de um modelo ARIMA ajustado, não à série original, e em tais aplicações a hipótese que está sendo testada é que os resíduos do modelo ARIMA não têm autocorrelação. Ao testar os resíduos de um modelo ARIMA estimado, os graus de liberdade precisam ser ajustados para refletir a estimação do parâmetro. Por exemplo, para um modelo ARIMA $(p, 0, q)$, os graus de liberdade devem ser definidos como $h - p - q$ (DAVIDSON, 2000). O teste Box-Pierce utiliza a estatística do teste, na notação descrita previamente, dada por (BOX; PIERCE, 1970)

$$Q_{BP} = n \sum_{k=1}^h \hat{\rho}_k^2 \quad (3.44)$$

e usa a mesma região crítica definida previamente. Estudos de simulação mostraram que a distribuição para a estatística Ljung-Box é mais próxima de um $\chi_{(h)}^2$ 6 distribuição do que é a distribuição para a estatística Box-Pierce para todos os tamanhos de amostra, incluindo os pequenos.

3.12.4 Testes de Friedman e Nemenyi

O teste de Friedman classifica os modelos K em cada conjunto de dados em relação ao valor absoluto dos resultados dados por esses algoritmos. A classificação do algoritmo com maior desempenho é 1, e o com menor desempenho é classificado como K . Em seguida, o valor da estatística com base em todas as classificações é calculado como

mostrado em equações (3.45) e (3.46) com r_{eu}^j sendo a classificação do desempenho do j -ésimo algoritmo no i -ésimo conjunto de dados. Essa estatística obedece à distribuição do qui-quadrado com $K - 1$ graus de liberdade (LIU; XU, 2022).

$$\chi_F^2 = \frac{12N}{K(K+1)} \left[\sum_{j=1}^K R_j^2 - \frac{K(K+1)^2}{4} \right] \quad (3.45)$$

$$R_j = \frac{1}{N} \sum_{i=1}^N r_{ei}^j \quad (3.46)$$

$$F_F = \frac{(N-1)\chi_F^2}{N(K-1)\chi_F^2} \quad (3.47)$$

As estatísticas F_F mostrados na equação (3.47) obedecem à distribuição F com graus de liberdade $K - 1$ e $(K - 1)(N - 1)$. Pode-se obter o valor crítico abaixo do nível de significância especificado (geralmente $\alpha = 0,05$ ou $0,01$). Ao comparar esse valor crítico com o valor calculado com a equação (3.47), a hipótese nula é rejeitada se o valor estatístico F_F é maior que o valor crítico, indicando que há diferenças significativas entre os algoritmos K . Em seguida, pode-se realizar um procedimento *post hoc* para analisar se o desempenho do algoritmo de controle é significativamente superior ao de cada algoritmo de referência nos experimentos. Ao contrário, se o valor for menor ou igual ao valor crítico, a hipótese nula é aceita, indicando que não há diferenças significativas entre os algoritmos K .

Adicionalmente, foi empregada a *Critical Difference* (CD) para avaliar se dois regressores eram significativamente distintos entre si. O CD foi calculado de acordo com a fórmula mencionada anteriormente:

$$CD = q_\alpha \sqrt{\frac{K(K+1)}{6N}} \quad (3.48)$$

onde CD, q_α representa o valor crítico, o teste de Nemenyi, k é o número de regressores, e N é o número total de amostras (LIU; XU, 2022).

4 Resultados

Neste capítulo, é fornecida uma síntese dos resultados obtidos neste estudo aplicando os diferentes modelos de previsão, citados no capítulo prévio, bem como a aplicação de testes e métricas estatísticas.

4.1 Análise Exploratória dos Dados

A descrição do problema, centrada no abastecimento d'água, é crucial. Apresenta variáveis-chave, como Bombas de Sucção (B1, B2 e B3), cuja frequência máxima é de 60 Hz, Nível do Reservatório (Câmara 1) representado por LT01 (m^3), Vazão de entrada (FT01) em (m^3/h), Vazão de gravidade (FT02) em (m^3/h), Vazão de recalque (FT03) em (m^3/h), Pressão de Sucção (PT01SU) medida em metros de coluna d'água (mca) e Pressão de Recalque (PT02RBAL) também em metros de coluna d'água (mca).

A pesquisa fará uso da variável LT01, que representa o nível do reservatório e desempenha um papel de extrema importância. A separação dos dados foi feita por hora a hora, mesmo que os dados obtidos da SANEPAR sejam de 2018 a 2020, sendo que o ano de 2020 causou muitas irregularidades. É possível remover esse ano para melhor trabalhar com os dados.

Mesmo havendo 9 variáveis nesse conjunto de dados, poderia-se trabalhar com 1 para previsão e as outras 8 como variáveis exógenas. No entanto, todas as variáveis podem ter correlação com o tanque, mas nem todas são necessárias, causando ruído na série temporal. Com isso em mente, foram retiradas as variáveis B3 e FT02 restando assim as variáveis de previsão com as variáveis que tiveram correlação significativa.

Neste caso, avalia-se 6 variáveis no formato de variáveis exógenas e uma variável para previsão. Com essa abordagem, restam 17.522 observações, com o intervalo temporal compreendido entre 2018 e 2019. Essa decisão foi tomada para evitar que o modelo sofra excessivamente com variações temporais.

É crucial destacar a importância de manter as outras variáveis como exógenas durante o processo de previsão. Embora haja a possibilidade de trabalhar com apenas uma variável para previsão, ao incorporar as demais como variáveis exógenas, o modelo é enriquecido, proporcionando uma visão mais abrangente do sistema. Neste contexto, as variáveis como Bombas de Sucção, Vazões e Pressões possuem potencial impacto no nível do reservatório, e ao mantê-las como exógenas, permite-se que o modelo considere suas influências durante a previsão. Essa abordagem contribui para uma análise mais completa e realista, levando em conta as interações complexas entre as variáveis e fortalecendo a capacidade preditiva do modelo em cenários de abastecimento d'água.

Os dados foram processados usando a EDA resumindo suas principais características, e formulando hipóteses que possam direcionar a coleta adicional de dados, se necessária. Existem dados anômalos *Not a Number* (NaN) que representam a ausência de dados coletados, logo tais dados foram interpolados usando os valores existentes e vizinhos a ele.

Assim como em qualquer empresa de saneamento básico e tratamento d'água, é utilizado um mecanismo de acionamento automático denominado trava de segurança para evitar que o nível do tanque se aproxime de zero e haja falta d'água nos locais abastecidos por esse tanque. O nível máximo que o tanque pode alcançar é de 5,26 m³ (equivalente a 5264.56 litros). As bombas são ativadas em sua potência máxima para evitar que sejam acionadas quando o nível do tanque estiver dentro dessa faixa. No entanto, a bomba 1 ainda estaria operando para completar o nível do tanque caso esteja dentro dessa faixa.

Na Tabela 7, o desvio padrão é representado por *STandard Deviation* (STD). A quantidade de dados medidos hora a hora pela SANEPAR de 2018 a 2019 são de 17.522 dados.

Tabela 7: Descrição estatística dos dados do Bairro Alto em Curitiba de 2018 a 2019 disponibilizados pela SANEPAR.

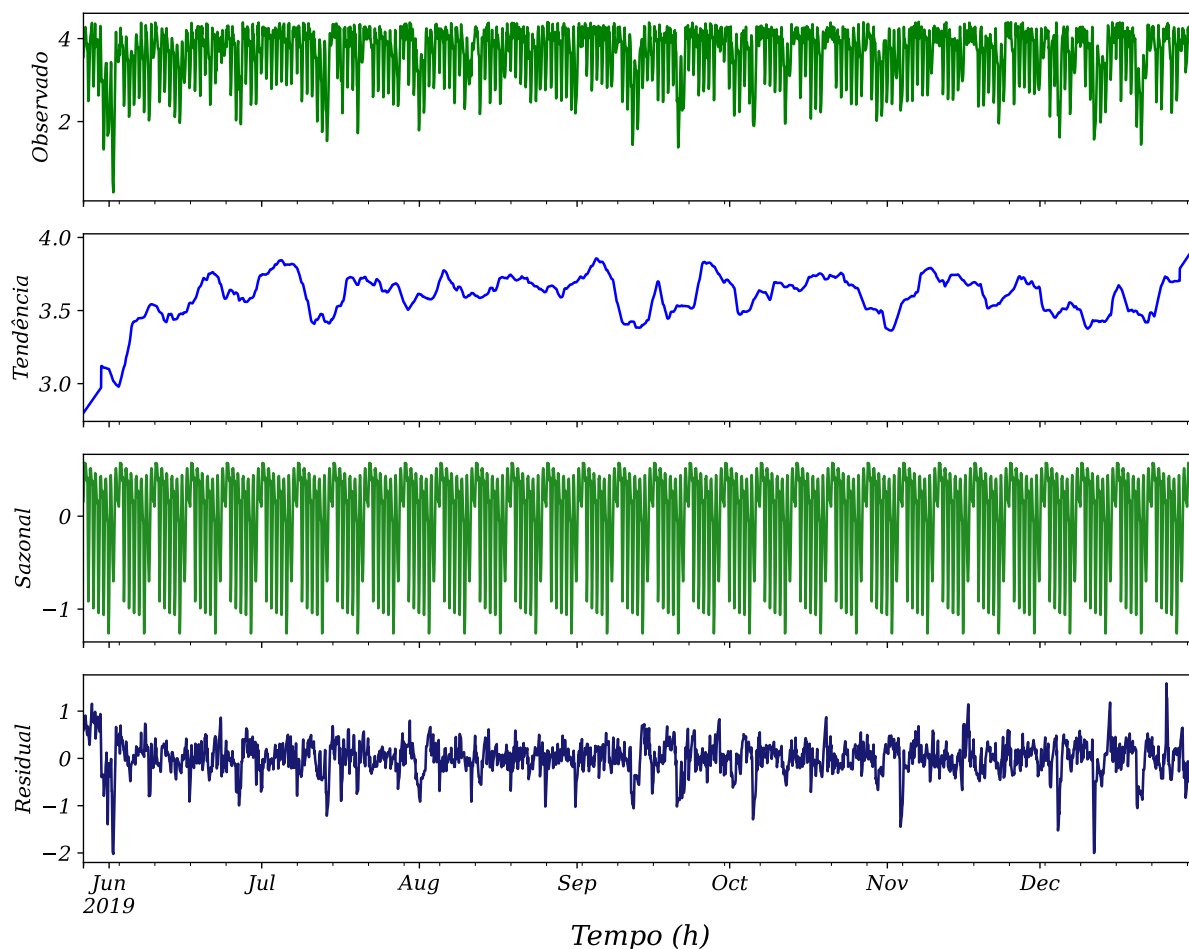
| Métricas | B1 | B2 | B3 | LT01 | FT01 | FT02 | FT03 | PT01 | PT02 |
|--------------|--------|--------|--------|-------|---------|---------|---------|-------|--------|
| Média | 52,289 | 18,421 | 3,338 | 3,513 | 215,699 | 114,832 | 104,195 | 4,448 | 20,724 |
| STD | 11,421 | 19,742 | 12,624 | 0,670 | 110,223 | 43,604 | 25,636 | 0,700 | 3,610 |
| Min. | 0 | 0 | 0 | 0,294 | 0 | 0 | 0 | 0,842 | 0 |
| 25% | 49,519 | 0 | 0 | 3,077 | 255,454 | 74,912 | 81,430 | 4,015 | 18,072 |
| 50% | 57,925 | 0,050 | 0 | 3,715 | 265,357 | 122,149 | 109,911 | 4,602 | 21,791 |
| 75% | 57,989 | 36,796 | 0 | 4,047 | 272,609 | 145,865 | 123,189 | 4,990 | 23,051 |
| Max. | 59,988 | 59,992 | 59,988 | 4,445 | 390,683 | 400,415 | 183,900 | 5,639 | 29,008 |

Com a decomposição STL é possível analisar se a série apresenta tendência, sazonalidade e ruídos. Ao observar a Figura 14, é evidente que os dados exibem ambos os padrões.

Essa decomposição divide os dados em sazonalidade, para descobrir no modelo ARIMA e seus antecessores que os dados podem ser coletados por hora, dia, mês ou ano. Saber que uma série tem sazonalidade torna o modelo ARIMA mais robusto, permitindo o uso do modelo ARIMA com sazonalidade, conhecido como SARIMA. Esses dados, por representarem uma sazonalidade difícil de ser notada, resultam em modelos com essa sazonalidade que não apresentam muita redução nos valores dos erros.

A tendência dos dados reflete como a série se comporta ao longo prazo, se ela pode subir ou reduzir o padrão. Tanto uma série pode ser estacionária ou não estacionária. Pela Figura 14, a tendência parece ser linear, caracterizando a série como estacionária.

Figura 14: Decomposição STL aditiva.



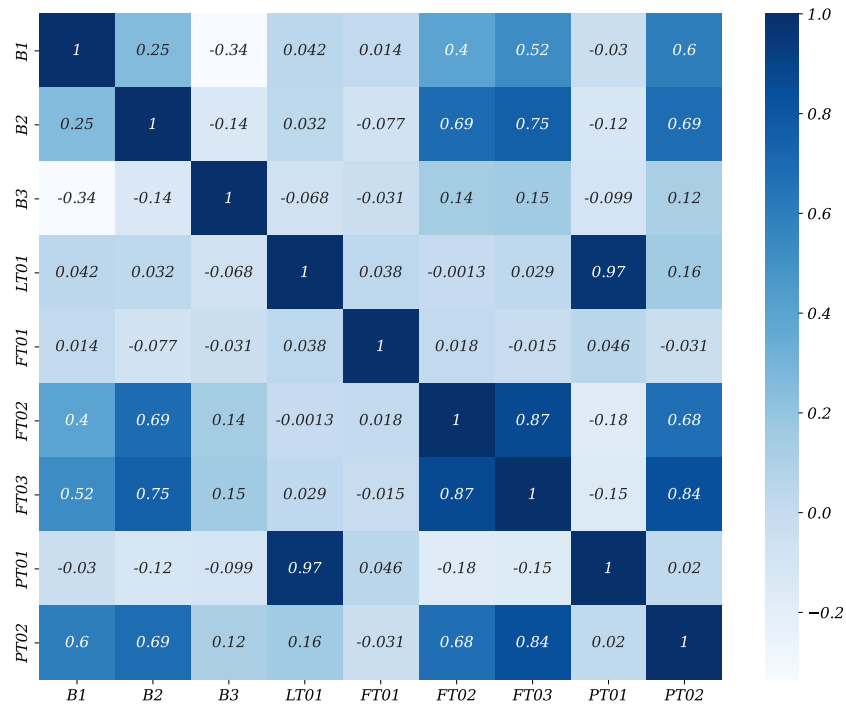
No entanto, essa série não possui um padrão sazonal, o que faz com que os modelos de ARIMA com sazonalidade apresentem erros maiores em comparação aos modelos clássicos.

O resíduo nos dados da série temporal é uma forma de avaliar que a série trabalhada possui várias irregularidades, levando aos resíduos que ajudam a verificar a verdadeira essência dos dados. Como esses dados foram coletados por hora, eles têm muita oscilação, trazendo muitos ruídos ao longo da análise.

A Figura 15 mostra a correlação de Pearson entre as variáveis do conjunto de dados deste estudo. Essa figura representa a matriz (simétrica) da dependência/correlação entre as variáveis. Analisamos as correlações das variáveis de entrada com a variável de saída. Serão descartadas as variáveis com correlação menor que 5% e maior que 90%.

No contexto das análises de dados, várias técnicas de EDA têm sido adotadas. Neste estudo, são realizadas diversas análises, incluindo a correlação de Pearson, para identificar quais variáveis podem ser excluídas devido a ruídos (correlação baixa) ou por estarem altamente correlacionadas com a variável de saída do estudo, que é a variável LT01, representando o nível do tanque de armazenamento d'água pela SANEPAR no

Figura 15: Correlação de Pearson.

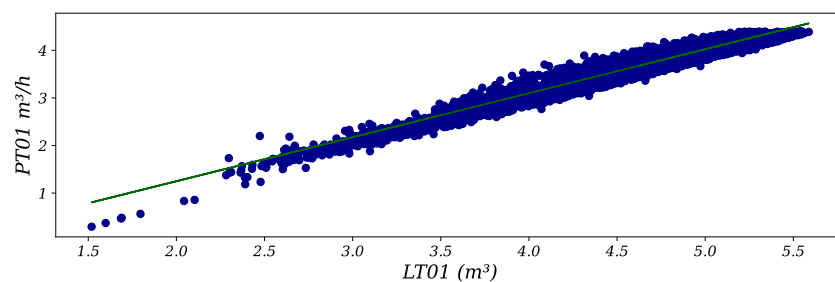


Bairro Alto. As variáveis removidas têm pouca correlação com a LT01; por exemplo, as variáveis B3 e FT02 apresentam correlação baixa e serão excluídas.

O termo “correlação baixa” indica que há pouca relação linear entre as variáveis, sugerindo uma associação fraca entre elas. Um teste de hipótese é comumente usado para determinar se essa correlação é estatisticamente significativa. Nos testes de correlação de Pearson, se o valor-p for menor que um determinado nível de significância (como 0,05), então a correlação é considerada estatisticamente significativa.

Nesse conjunto de dados que está sendo trabalhado, há uma forte correlação da variável PT01 com o LT01 conforme visto na Figura 16 que fornece uma representação dos coeficientes β_0 e β_1 , que são os coeficientes da correlação linear entre as variáveis. Um aumento de 1 na variável x está associado a um aumento proporcional de β_1 na variável y . O valor de β_0 representa o valor de y quando x é igual a 0.

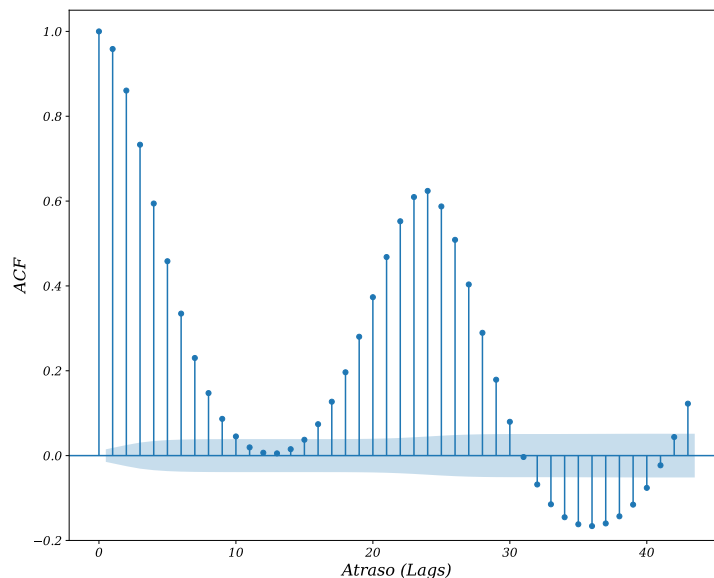
Figura 16: Relação entre LT01 e PT01 cuja correlação de Pearson é 97%.



Na Figura 17, observa-se a distinção entre ACF, e PACF, mostrada na Figura 18. A autocorrelação é uma métrica que indica a relação entre os valores de uma série temporal em diferentes defasagens, capturando tanto correlações diretas quanto indiretas. Por sua vez, a autocorrelação parcial mensura apenas a correlação direta entre os valores, ignorando a influência de defasagens intermediárias. Essas análises são cruciais para detectar padrões e dependências entre os valores da série temporal, o que é essencial para a modelagem e previsão desses dados.

O intervalo de confiança padrão de 95%, destacado pela faixa azul nas Figuras 17 e 18, serve como referência crítica. Observações fora desse intervalo indicam correlações estatisticamente significativas, sugerindo a presença de padrões ou estrutura na série temporal.

Figura 17: Autocorrelação.



A correlação visualizada na Figura 17 é fundamental para a interpretação do teste ADF. Em uma série de ruído branco, os valores são completamente aleatórios e não apresentam correlação significativa. Portanto, quando há correlação presente na série, isso indica a existência de padrões ou dependências entre os valores, o que pode ser explorado para a modelagem e previsão da série temporal.

Demonstrar que uma série temporal tem ou pode ter um ruído branco também é conveniente para a análise da EDA. Na Figura 19, é possível observar uma série temporal que pode ser caracterizada como ruído branco, se suas variáveis forem independentes e distribuídas de forma idêntica, com média zero. Isso implica que todas as variáveis possuem a mesma variância (σ^2) e que cada valor não possui correlação com os demais valores da série.

Um valor de ADF de $-12,515$ indica a evidência de estacionariedade na série

Figura 18: Autocorrelação parcial.

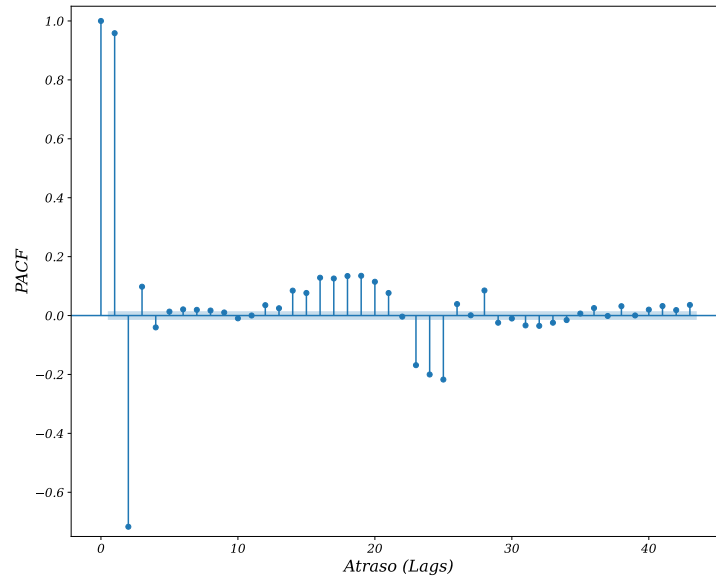
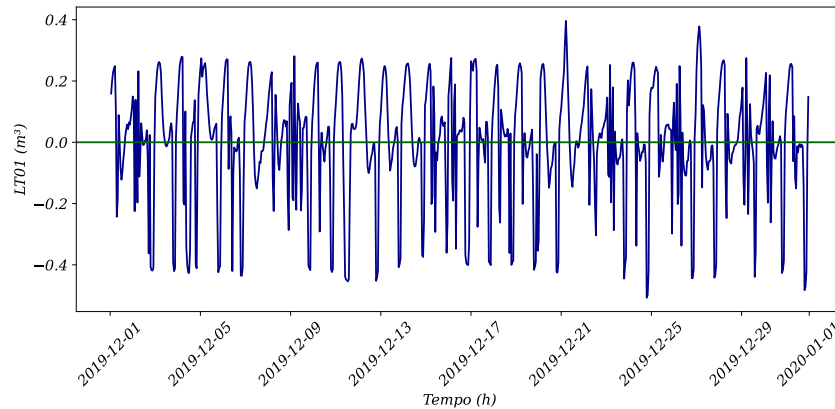


Figura 19: Ruído branco.



temporal do nível do tanque LT01. O valor de p expresso como $2,62 \times 10^{-23}$ está associado ao teste ADF. No contexto do teste ADF, a hipótese nula é a presença de raiz unitária na série temporal, indicando não estacionariedade. Dado o valor de p de $2,62 \times 10^{-23}$, evidencia-se uma probabilidade muito baixa, indicando forte suporte contra a hipótese nula e sugerindo que a série temporal é estacionária. Na Tabela 8, são apresentados todos os dados do teste para estacionariedade. Os resultados indicam fortes evidências contra a hipótese nula. Com um teste ADF de $-12,515$ e um valor de p extremamente baixo de $2,62 \times 10^{-23}$, rejeita-se a hipótese nula de presença de raiz unitária. Os 44 atrasos utilizados e as 17.477 observações corroboram a análise estatística. Ao comparar a estatística de teste ADF com os valores críticos, observa-se que está significativamente abaixo deles em todos os níveis de significância (1%, 5%, 10%). Portanto, a conclusão é de que os dados não possuem raiz unitária, indicando que são estacionários.

Para a previsão os dados foram divididos em conjuntos de treinamento, validação,

Tabela 8: Teste ADF.

| | |
|-----------------------|------------------------|
| Valor de p | $2,62 \times 10^{-23}$ |
| Atrasos utilizados | 44 |
| Número de observações | 17.477 |
| Valor crítico (1%) | -3,431 |
| Valor crítico (5%) | -2,862 |
| Valor crítico (10%) | -2,567 |

e teste (RASCHKA, 2015; GÉRON, 2017). Quanto à divisão dos dados, foi adotada uma estratégia básica em que 70% dos dados foram destinados ao conjunto de treinamento e 30% restantes foram reservados para o conjunto de teste. Dentro dos 70% de treinamento, foi realizada uma subdivisão em que 80% desses dados foram usados novamente para treinamento e os 20% restantes foram utilizados para validação.

A estratégia recursiva é mencionada por Petropoulos et al. (2022) como uma abordagem eficaz na previsão de séries temporais de múltiplos passos. De acordo com o autor, essa estratégia envolve o uso de previsões anteriores como entradas para prever os próximos passos da série temporal. A abordagem recursiva tem demonstrado potencial para melhorar a acurácia das previsões de séries temporais de longo prazo.

Modelos de Séries Temporais (AR, ARIMA, ARIMAX, ARMA, ARX, MA, SARIMA, SARIMAX, Prophet):

Para todos os modelos dessa categoria, as características de entrada e variáveis exógenas são semelhantes. Ele utiliza apenas o histórico de dados do LT01 como entrada e considera outras variáveis, como FT01, FT03, PT01, PT02, B1 e B2, como variáveis exógenas. A configuração consiste em organizar os dados históricos do nível do reservatório em séries temporais e ajustar cada modelo de acordo com sua técnica específica para prever valores futuros com base nos padrões identificados nos dados históricos do LT01.

Redes Neurais Recorrentes (RNN, LSTM, GRU):

Assim como nos modelos de séries temporais, as entradas e variáveis exógenas dessas redes neurais são as mesmas mencionadas anteriormente. Ele também utiliza apenas o histórico de dados do LT01 como entrada e considera outras variáveis como exógenas. No entanto, a configuração difere, pois os dados históricos do nível do reservatório são organizados em sequências temporais, e a rede é treinada com camadas recorrentes para aprender padrões temporais e fazer previsões futuras.

Outros Modelos (CNN, MLP, DT, RF, XGBoost, LightGBM):

Similarmente aos modelos anteriores, esses modelos também compartilham as mesmas entradas e variáveis exógenas. Ele utiliza apenas o histórico de dados do LT01 como entrada e considera outras variáveis como exógenas. A configuração, no entanto, varia dependendo do modelo. Os dados históricos do nível do reservatório são organizados em

conjuntos de treinamento para cada modelo, e cada um é treinado para aprender padrões nos dados e fazer previsões futuras. Uma exceção é o modelo DT, que utiliza o LT01 para previsão e o PT01 como variável alvo no eixo y .

O MISO foi o modelo usado neste estudo. O modelo ARIMA, juntamente com suas variantes e extensões, foi amplamente estudado durante a pesquisa, assim como modelos regressivos que envolvem múltiplas variáveis de entrada e uma variável de saída, neste caso, a LT01. As demais variáveis, como FT01, FT03, PT01, PT02, B1 e B2, foram utilizadas como suporte para melhorar o modelo do tipo ARIMAX, modelos com variáveis exógenas.

Quando aplicado sem o uso de variáveis exógenas, o modelo ARIMA apresenta apenas uma entrada, semelhante ao modelo de regressão linear. No entanto, ao incluir variáveis exógenas, o modelo se torna MISO, permitindo uma modelagem abrangente e considerando a correlação das várias para prever a variável de interesse, LT01.

A previsão dos dados foi feita com diferentes horizontes de previsão como 1 hora, 6 horas, 12 horas, e 24 horas. Essas estratégias de previsão permitem a comparação entre os modelos de regressão e modelos ARIMA em diferentes horizontes temporais.

Além desses modelos de previsão, vários outros modelos foram utilizados no estudo, tais como DT, RF, XGBoost, LightGBM, MLP, LSTM, GRU, Prophet, RNN, e CNN, a fim de obter o melhor resultado para a previsão de séries temporais de abastecimento d'água.

Na Tabela 9 são exibidos todos os valores obtidos pela biblioteca do *Python* `autoARIMA`, resultantes para p , que representa a ordem do componente AR; d , que indica o número de diferenciações não sazonais; q , que denota a ordem do componente MA; P , que corresponde à ordem do componente AR sazonal; D , que representa o número de diferenciações sazonais; Q , que indica a ordem do componente MA sazonal; e M , que representa o período sazonal (o número de observações em um ciclo sazonal).

Tabela 9: Parâmetros para os modelos ARIMA utilizando a função `autoARIMA`.

| Modelos | Parâmetros |
|--------------------------------------|---|
| AR(p) | $p = 3$ |
| ARX(p) | $p = 3$ |
| MA(q) | $q = 3$ |
| ARMA(p, q) | $p = 3, q = 3$ |
| ARIMA(p, d, q) | $p = 3, d = 1, q = 3$ |
| ARIMAX(p, d, q) | $p = 3, d = 1, q = 3$ |
| SARIMA(p, d, q)(P, D, Q) | $p = 3, d = 1, q = 3, P = 0, D = 1, Q = 2$ |
| SARIMAX(p, d, q)(P, D, Q, M) | $p = 3, d = 1, q = 3, P = 0, D = 1, Q = 2, M = 4$ |

Os hiperparâmetros dos modelos foram otimizados usando a biblioteca `Optuna`

do *Python*. Nesse contexto, foram empregadas técnicas bayesianas, especificamente o algoritmo TPE visando uma otimização mais eficiente.

Na Tabela 10 são descritos os hiperparâmetros dos modelos XGBoost, LightGBM, RF, e DT, onde NE é o número de estimadores, PM é a profundidade máxima, MAD é o mínimo de amostras por divisão, MAF é o mínimo de amostras por folha, e TA é a taxa de aprendizado.

Tabela 10: Hiperparâmetros otimizados dos modelos.

| Modelo | Estimadores | PM | MAD | MAF | TA |
|----------|-------------|-----|-----|-----|-------|
| DT | N/A | 976 | 14 | 14 | N/A |
| LightGBM | 419 | 18 | 8 | 13 | 0,021 |
| RF | 477 | 18 | 4 | 11 | N/A |
| XGBoost | 1752 | 4 | 23 | 24 | 0,014 |

Os hiperparâmetros dos modelos de rede neural artificial (RNA), como RNN, MLP, CNN, GRU e LSTM, obtidos através da otimização TPE da biblioteca Optuna em Python, são exibidos na Tabela 11. CNN utilizou um kernel de tamanho 5 e uma camada densa, enquanto MLP utilizou uma única camada densa.

Tabela 11: Hiperparâmetros otimizados para RNA.

| Modelo | Layers | Tamanho do Batch | No. Épocas | Dropout/ Learning Rate |
|--------|--------|---------------------|------------|---------------------------|
| CNN | 4 | 13 | 100 | 0,24; 0,00067 |
| GRU | 2 | 16 | 69 | – ; 0,01 |
| LSTM | 4 | 32 | 122 | 0,2 ; 0,01 |
| MLP | 3 | 32 | 65 | 0,218, 0,00030 |
| RNN | 4 | 16 | 50 | 0,2 ; 0,01 |

4.2 Aplicando os Modelos de Previsão

Na análise dos modelos desenvolvidos, observou-se que o modelo DT obteve o melhor desempenho, tanto para previsões de curto prazo, durante as horas de pico entre 18h e 21h, quanto para outros períodos. Além disso, o Prophet foi o segundo modelo que mais apresentou resultados satisfatório em previsão de longo prazo.

Uma observação recorrente foi a superioridade dos modelos que incorporam variáveis exógenas em termos de capacidade de previsão, evidenciada nas Figuras 20 a 26 e nas Tabelas 12 a 15, onde os valores menores foram destacados em **negrito**. O modelo DT destacou-se como o melhor em termos de eficácia nas previsões realizadas, seguido pelo modelo Prophet. Desde a 20 a 26 são ilustrados cenários distintos de previsão e

comparação entre os modelos DT e Prophet, respectivamente, sendo o primeiro a escolha mais adequada. As Figuras são exibidas com um corte temporal da data 2019/12/01, 2020/01/02.

Figura 20: Comparação dos modelos de previsão AR, ARX, MA e ARMA 24 passo à frente.

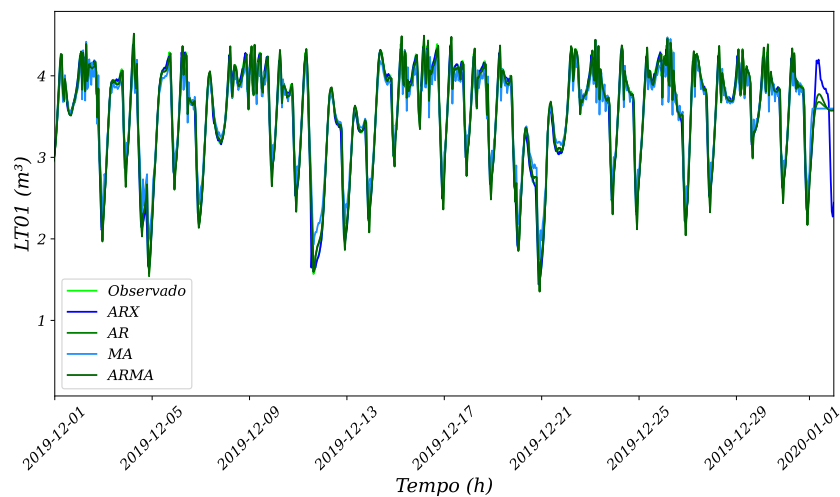


Figura 21: Comparação dos modelos de previsão ARIMA, ARIMAX, SARIMA e SARIMAX 24 passo à frente.

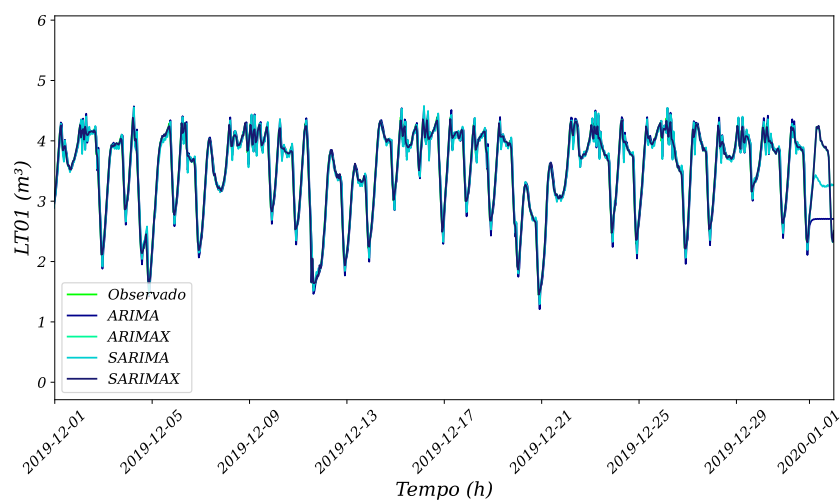


Figura 22: Comparação dos modelos DT, RF, XGBoost e LightGBM 1 passo à frente.

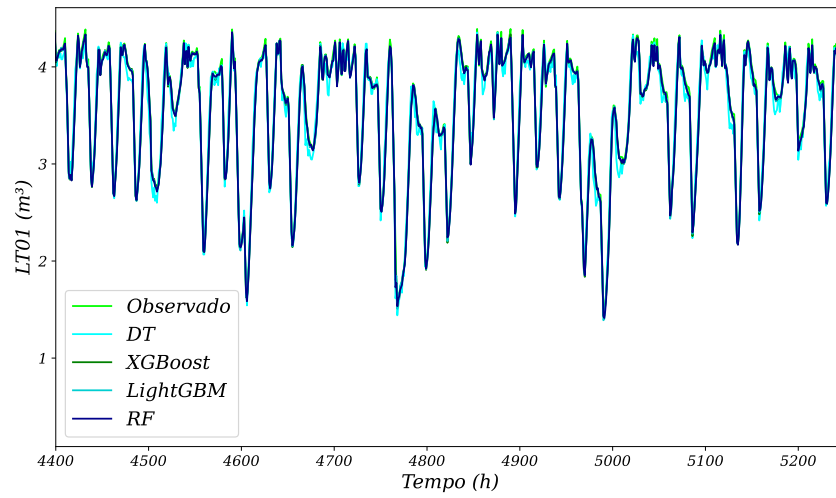


Figura 23: Modelo de previsão GRU para vários horizontes de previsão.

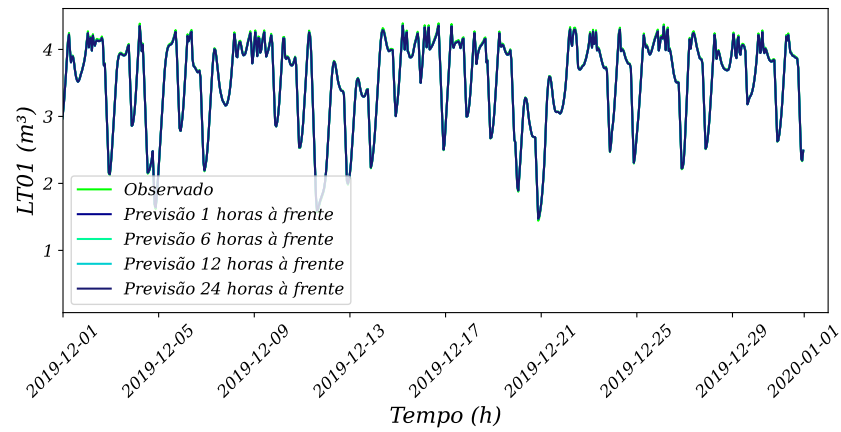


Figura 24: Modelo de previsão LSTM para vários horizontes de previsão.

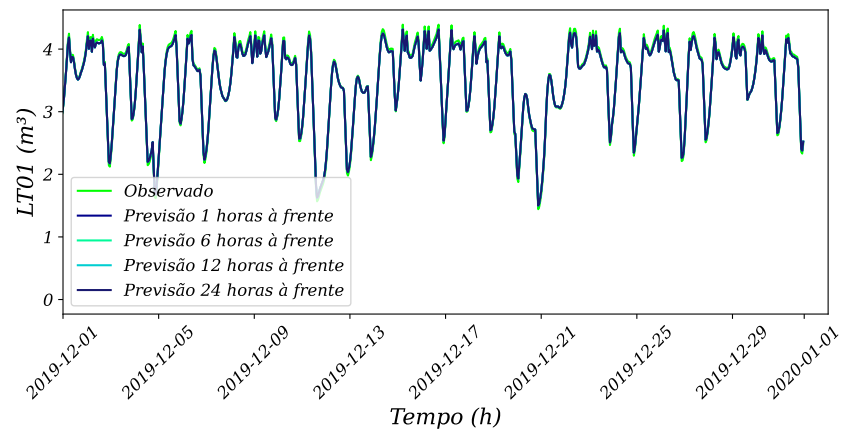


Figura 25: Modelo de previsão RNN para vários horizontes de previsão.

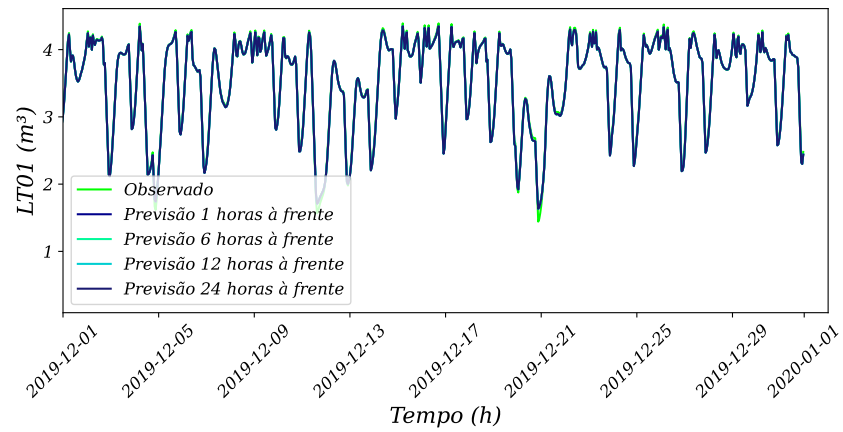


Figura 26: Previsões do modelo Prophet para vários horizontes de previsão.

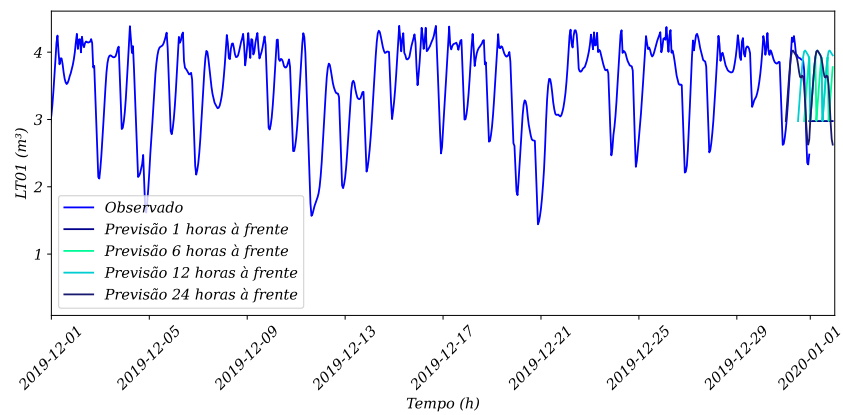


Tabela 12: Comparação dos modelos de previsão através das métricas de desempenho para dados de treino.

| Horizontes | Métricas | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|---|-----------|-------|-------|-------|-------|-------|-------|--------------|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 hora à frente | SMAPE (%) | 7,81 | 7,85 | 7,83 | 7,79 | 8,01 | 18,77 | 7,69 | 35,89 | 24,60 | 18,55 | 8,37 | 18,77 | 21,27 | 18,38 | 23,36 | 8,03 | 7,79 | 19,56 |
| | MAE | 0,25 | 0,25 | 0,25 | 0,25 | 0,25 | 0,10 | 0,35 | 1,44 | 0,93 | 0,65 | 0,27 | 0,10 | 0,77 | 0,64 | 0,87 | 0,26 | 0,25 | 0,69 |
| | RRMSE (%) | 0,09 | 0,09 | 0,10 | 0,09 | 0,10 | 0,42 | 0,21 | 0,65 | 1,36 | 0,21 | 0,10 | 0,42 | 0,77 | 0,21 | 1,28 | 0,10 | 0,10 | 0,22 |
| 6 horas à frente | SMAPE (%) | 19,99 | 19,93 | 20,10 | 20,01 | 20,08 | 20,37 | 13,94 | 83,75 | 58,21 | 18,55 | 18,34 | 20,37 | 33,71 | 24,25 | 50,20 | 20,12 | 20,13 | 26,23 |
| | MAE | 0,64 | 0,64 | 0,64 | 0,64 | 0,64 | 0,20 | 0,59 | 4,94 | 2,77 | 0,65 | 0,60 | 0,20 | 1,12 | 0,88 | 2,25 | 0,64 | 0,64 | 0,97 |
| | RRMSE (%) | 0,23 | 0,23 | 0,23 | 0,23 | 0,23 | 0,44 | 0,16 | 1,70 | 4,15 | 0,21 | 0,21 | 0,44 | 1,15 | 0,32 | 3,42 | 0,23 | 0,23 | 0,34 |
| 12 horas à frente | SMAPE (%) | 23,23 | 23,03 | 23,31 | 23,23 | 23,14 | 19,04 | 13,94 | 98,06 | 60,06 | 18,55 | 21,31 | 19,04 | 24,62 | 24,22 | 51,28 | 23,37 | 23,38 | 26,26 |
| | MAE | 0,75 | 0,75 | 0,75 | 0,75 | 0,75 | 0,12 | 0,59 | 6,62 | 2,91 | 0,65 | 0,70 | 0,12 | 0,83 | 0,88 | 2,32 | 0,76 | 0,76 | 0,97 |
| | RRMSE (%) | 0,27 | 0,26 | 0,27 | 0,27 | 0,26 | 0,42 | 0,16 | 2,23 | 4,36 | 0,21 | 0,25 | 0,42 | 0,91 | 0,32 | 3,53 | 0,27 | 0,27 | 0,34 |
| 24 horas à frente | SMAPE (%) | 13,53 | 13,37 | 13,69 | 13,53 | 13,70 | 19,63 | 13,94 | 104,82 | 60,10 | 18,55 | 12,90 | 19,63 | 6,55 | 24,25 | 51,13 | 13,77 | 13,70 | 26,75 |
| | MAE | 0,43 | 0,43 | 0,44 | 0,43 | 0,44 | 0,16 | 0,59 | 7,59 | 2,91 | 0,65 | 0,42 | 0,16 | 0,23 | 0,88 | 2,31 | 0,44 | 0,44 | 1,00 |
| | RRMSE (%) | 0,17 | 0,17 | 0,17 | 0,17 | 0,17 | 0,43 | 0,16 | 2,53 | 4,37 | 0,21 | 0,16 | 0,43 | 0,29 | 0,32 | 3,52 | 0,17 | 0,17 | 0,35 |
| A – AR, B – ARIMA, C – ARIMAX, D – ARMA, E – ARX, F – CNN, G – DT , H – GRU, I – LSTM, J – LigthGBM, K – MA, L – MLP, M – Prophet, N – RF, O – RNN, P – SARIMA, Q – SARIMAX, R – XGBoost | | | | | | | | | | | | | | | | | | | |

Tabela 13: Comparação dos modelos de previsão através das métricas de desempenho para dados de validação.

| Horizontes | Métricas | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|-------------------|-----------|-------|-------------|-------|-------|-------|-------|--------------|--------|-------|-------|-------|-------|--------------|-------|-------|-------|-------|-------|
| 1 hora à frente | SMAPE (%) | 7,88 | 7,66 | 8,31 | 7,87 | 8,29 | 20,52 | 8,51 | 15,84 | 23,63 | 16,76 | 8,20 | 20,52 | 18,22 | 16,64 | 22,26 | 7,88 | 8,31 | 17,81 |
| | MAE | 0,26 | 0,25 | 0,27 | 0,26 | 0,27 | 1,88 | 0,39 | 0,54 | 0,92 | 0,60 | 0,27 | 1,88 | 0,50 | 0,60 | 0,86 | 0,26 | 0,27 | 0,65 |
| | RRMSE (%) | 0,09 | 0,09 | 0,10 | 0,09 | 0,10 | 0,50 | 0,18 | 0,33 | 1,50 | 0,19 | 0,09 | 0,50 | 0,50 | 0,19 | 1,40 | 0,09 | 0,10 | 0,20 |
| 6 horas à frente | SMAPE (%) | 19,39 | 19,92 | 19,91 | 19,47 | 19,89 | 18,90 | 12,66 | 72,77 | 55,41 | 16,76 | 17,81 | 18,90 | 29,73 | 21,03 | 47,27 | 19,32 | 19,89 | 22,81 |
| | MAE | 0,64 | 0,66 | 0,66 | 0,65 | 0,66 | 1,80 | 0,56 | 4,04 | 2,68 | 0,60 | 0,60 | 1,80 | 0,93 | 0,78 | 2,15 | 0,64 | 0,66 | 0,86 |
| | RRMSE (%) | 0,23 | 0,23 | 0,23 | 0,23 | 0,23 | 0,48 | 0,14 | 1,42 | 4,43 | 0,19 | 0,21 | 0,48 | 1,05 | 0,28 | 3,61 | 0,23 | 0,23 | 0,30 |
| 12 horas à frente | SMAPE (%) | 22,16 | 23,05 | 22,46 | 22,18 | 22,47 | 19,62 | 12,66 | 93,58 | 57,15 | 16,76 | 20,40 | 19,62 | 23,91 | 20,99 | 48,28 | 22,29 | 22,45 | 22,82 |
| | MAE | 0,74 | 0,77 | 0,75 | 0,74 | 0,75 | 1,83 | 0,56 | 6,25 | 2,80 | 0,60 | 0,69 | 1,83 | 0,80 | 0,78 | 2,22 | 0,75 | 0,75 | 0,86 |
| | RRMSE (%) | 0,26 | 0,27 | 0,26 | 0,26 | 0,26 | 0,49 | 0,14 | 2,09 | 4,64 | 0,19 | 0,24 | 0,49 | 0,97 | 0,28 | 3,72 | 0,26 | 0,26 | 0,29 |
| 24 horas à frente | SMAPE (%) | 12,32 | 12,85 | 12,70 | 12,29 | 12,70 | 17,69 | 12,66 | 102,11 | 57,20 | 16,76 | 11,65 | 17,69 | 5,05 | 21,02 | 48,12 | 12,54 | 12,70 | 22,81 |
| | MAE | 0,41 | 0,42 | 0,42 | 0,40 | 0,41 | 1,74 | 0,56 | 7,42 | 2,81 | 0,60 | 0,39 | 1,74 | 0,17 | 0,78 | 2,21 | 0,41 | 0,41 | 0,86 |
| | RRMSE (%) | 0,16 | 0,17 | 0,16 | 0,16 | 0,16 | 0,46 | 0,14 | 2,44 | 4,65 | 0,19 | 0,15 | 0,46 | 0,19 | 0,28 | 3,71 | 0,16 | 0,16 | 0,30 |

A – AR, B – ARIMA, C – ARIMAX, D – ARMA, E – ARX, F – CNN, **G – DT**, H – GRU, I – LSTM, J – LigthGBM, K – MA, L – MLP, **M – Prophet**, N – RF, O – RNN, P – SARIMA, Q – SARIMAX, R – XGBoost

Tabela 14: Comparação dos modelos de previsão através das métricas de desempenho para dados de teste.

| Horizontes | Métricas | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|-------------------|-----------|-------|-------|-------|-------|-------|-------|--------------|--------|-------|-------|-------|-------|--------------|-------|-------|-------|-------|-------|
| 1 hora à frente | SMAPE (%) | 8,16 | 8,12 | 8,53 | 8,16 | 8,56 | 19,57 | 6,90 | 29,44 | 25,55 | 20,80 | 8,53 | 19,57 | 14,02 | 20,70 | 24,51 | 7,99 | 8,51 | 21,81 |
| | MAE | 0,25 | 0,25 | 0,26 | 0,25 | 0,26 | 0,45 | 0,31 | 1,08 | 0,94 | 0,72 | 0,26 | 0,45 | 0,46 | 0,71 | 0,89 | 0,25 | 0,26 | 0,76 |
| | RRMSE (%) | 0,10 | 0,10 | 0,10 | 0,10 | 0,10 | 0,42 | 0,20 | 0,56 | 1,36 | 0,23 | 0,10 | 0,42 | 0,46 | 0,23 | 1,29 | 0,10 | 0,10 | 0,24 |
| 6 horas à frente | SMAPE (%) | 21,80 | 21,75 | 22,19 | 21,82 | 22,23 | 22,08 | 13,58 | 83,96 | 60,90 | 20,80 | 19,96 | 22,08 | 7,16 | 27,31 | 53,00 | 22,09 | 22,17 | 29,34 |
| | MAE | 0,68 | 0,68 | 0,69 | 0,68 | 0,69 | 0,62 | 0,57 | 4,80 | 2,87 | 0,72 | 0,63 | 0,62 | 0,25 | 0,98 | 2,35 | 0,69 | 0,69 | 1,08 |
| | RRMSE (%) | 0,25 | 0,25 | 0,25 | 0,25 | 0,25 | 0,47 | 0,16 | 1,72 | 4,22 | 0,23 | 0,23 | 0,47 | 0,28 | 0,35 | 3,51 | 0,25 | 0,25 | 0,38 |
| 12 horas à frente | SMAPE (%) | 25,41 | 25,38 | 25,53 | 25,42 | 25,54 | 24,14 | 13,58 | 99,75 | 62,83 | 20,80 | 23,31 | 24,14 | 14,61 | 27,28 | 54,14 | 26,07 | 25,54 | 29,38 |
| | MAE | 0,80 | 0,80 | 0,81 | 0,80 | 0,81 | 0,73 | 0,57 | 6,64 | 3,01 | 0,72 | 0,74 | 0,73 | 0,54 | 0,98 | 2,42 | 0,82 | 0,81 | 1,08 |
| | RRMSE (%) | 0,29 | 0,29 | 0,29 | 0,29 | 0,29 | 0,49 | 0,16 | 2,31 | 4,44 | 0,23 | 0,27 | 0,49 | 0,59 | 0,35 | 3,62 | 0,30 | 0,29 | 0,38 |
| 24 horas à frente | SMAPE (%) | 14,59 | 14,56 | 14,86 | 14,59 | 14,89 | 23,90 | 13,58 | 106,86 | 62,88 | 20,80 | 13,71 | 23,90 | 13,41 | 27,30 | 54,00 | 14,99 | 14,84 | 29,91 |
| | MAE | 0,46 | 0,45 | 0,46 | 0,46 | 0,46 | 0,72 | 0,57 | 7,67 | 3,01 | 0,72 | 0,43 | 0,72 | 0,43 | 0,98 | 2,41 | 0,47 | 0,46 | 1,10 |
| | RRMSE (%) | 0,18 | 0,18 | 0,18 | 0,18 | 0,18 | 0,49 | 0,16 | 2,64 | 4,45 | 0,23 | 0,17 | 0,49 | 0,56 | 0,35 | 3,61 | 0,18 | 0,18 | 0,38 |

A – AR, B – ARIMA, C – ARIMAX, D – ARMA, E – ARX, F – CNN, **G – DT**, H – GRU, I – LSTM, J – LigthGBM, K – MA, L – MLP, **M – Prophet**, N – RF, O – RNN, P – SARIMA, Q – SARIMAX, R – XGBoost

Tabela 15: Comparação dos modelos de previsão através das métricas de desempenho para todos dados.

| Horizontes | Métricas | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|---|-----------|-------------|-------|-------|-------|-------|-------|--------------|--------|-------|-------|-------------|-------|--------------|-------|-------|-------|-------|-------|
| 1 hora à frente | SMAPE (%) | 7,87 | 7,90 | 8,03 | 7,88 | 8,29 | 18,20 | 7,83 | 17,37 | 24,44 | 18,33 | 8,36 | 18,20 | 18,22 | 18,18 | 23,19 | 8,05 | 7,99 | 19,35 |
| | MAE | 0,25 | 0,25 | 0,26 | 0,25 | 0,27 | 1,61 | 0,36 | 0,58 | 0,93 | 0,64 | 0,27 | 1,61 | 0,50 | 0,64 | 0,87 | 0,26 | 0,26 | 0,69 |
| | RRMSE (%) | 0,09 | 0,09 | 0,10 | 0,09 | 0,10 | 0,43 | 0,20 | 0,31 | 1,39 | 0,21 | 0,10 | 0,43 | 0,50 | 0,20 | 1,31 | 0,10 | 0,10 | 0,21 |
| 6 horas à frente | SMAPE (%) | 20,08 | 20,03 | 20,29 | 20,12 | 20,32 | 19,82 | 13,50 | 74,59 | 57,75 | 23,86 | 18,43 | 19,82 | 29,73 | 23,71 | 49,71 | 19,92 | 20,31 | 25,64 |
| | MAE | 0,65 | 0,65 | 0,65 | 0,65 | 0,66 | 1,71 | 0,58 | 4,08 | 2,76 | 0,87 | 0,60 | 1,71 | 0,93 | 0,87 | 2,23 | 0,64 | 0,65 | 0,95 |
| | RRMSE (%) | 0,23 | 0,23 | 0,23 | 0,23 | 0,23 | 0,45 | 0,16 | 1,45 | 4,22 | 0,31 | 0,21 | 0,45 | 1,05 | 0,31 | 3,47 | 0,23 | 0,23 | 0,33 |
| 12 horas à frente | SMAPE (%) | 23,24 | 23,13 | 23,36 | 23,23 | 23,27 | 18,47 | 13,50 | 95,48 | 59,57 | 23,93 | 21,34 | 18,47 | 23,91 | 23,68 | 50,78 | 23,15 | 23,42 | 25,66 |
| | MAE | 0,76 | 0,75 | 0,76 | 0,76 | 0,76 | 1,63 | 0,58 | 6,32 | 2,89 | 0,88 | 0,70 | 1,63 | 0,80 | 0,86 | 2,30 | 0,75 | 0,76 | 0,95 |
| | RRMSE (%) | 0,27 | 0,27 | 0,27 | 0,27 | 0,27 | 0,43 | 0,16 | 2,14 | 4,43 | 0,31 | 0,25 | 0,43 | 0,97 | 0,31 | 3,58 | 0,27 | 0,27 | 0,33 |
| 24 horas à frente | SMAPE (%) | 13,32 | 13,22 | 13,57 | 13,29 | 13,63 | 19,08 | 13,50 | 103,98 | 59,62 | 24,07 | 12,65 | 19,08 | 5,05 | 23,71 | 50,63 | 13,51 | 13,57 | 26,17 |
| | MAE | 0,43 | 0,42 | 0,43 | 0,43 | 0,44 | 1,67 | 0,58 | 7,51 | 2,89 | 0,88 | 0,41 | 1,67 | 0,17 | 0,87 | 2,29 | 0,43 | 0,43 | 0,98 |
| | RRMSE (%) | 0,17 | 0,17 | 0,17 | 0,17 | 0,17 | 0,44 | 0,16 | 2,50 | 4,43 | 0,32 | 0,16 | 0,44 | 0,19 | 0,31 | 3,57 | 0,17 | 0,17 | 0,34 |
| A – AR, B – ARIMA, C – ARIMAX, D – ARMA, E – ARX, F – CNN, G – DT , H – GRU, I – LSTM, J – LigthGBM, K – MA, L – MLP, M – Prophet , N – RF, O – RNN, P – SARIMA, Q – SARIMAX, R – XGBoost | | | | | | | | | | | | | | | | | | | |

Na Tabela 16 são mostrados os valores das métricas estatísticas para todos os modelos de previsão para 24 passos à frente (1 dia) usando dados de teste.

Tabela 16: Métricas de avaliação dos modelos com 24 passos à frente.

| Modelo | SMAPE (%) | MAE | RRMSE (%) |
|----------------|--------------|--------------|--------------|
| Prophet | 5,050 | 0,169 | 0,191 |
| MLP | 19,076 | 1,668 | 0,444 |
| CNN | 19,076 | 1,668 | 0,444 |
| RNN | 50,630 | 2,294 | 3,567 |
| LSTM | 59,622 | 2,894 | 4,433 |
| GRU | 103,984 | 7,506 | 2,505 |
| AR | 13,318 | 0,428 | 0,169 |
| ARIMA | 13,217 | 0,424 | 0,167 |
| ARIMAX | 13,566 | 0,434 | 0,173 |
| ARMA | 13,292 | 0,427 | 0,168 |
| ARX | 13,634 | 0,436 | 0,173 |
| DT | 13,505 | 0,576 | 0,158 |
| LigthGBM | 24,067 | 0,881 | 0,315 |
| MA | 12,650 | 0,411 | 0,159 |
| RF | 23,708 | 0,865 | 0,311 |
| SARIMA | 13,508 | 0,434 | 0,170 |
| SARIMAX | 13,572 | 0,434 | 0,173 |
| XGBoost | 26,169 | 0,978 | 0,338 |

Na tabela anterior, os diversos modelos de previsão de séries temporais foram avaliados para um horizonte de previsão de 24 horas. Para cada métrica SMAPE, MAE e RRMSE, identificou-se o modelo que apresentou o menor valor. A métrica SMAPE indicou que o modelo Prophet obteve o menor valor. Quanto à métrica MAE, novamente o modelo Prophet demonstrou o menor valor, e com a métrica RRMSE, o modelo MA demonstrou o menor valor da métrica.

Para validar estatisticamente as diferenças entre os modelos, foi realizado um teste estatístico denominado Teste de Friedman. Esse teste avalia o desempenho dos modelos em todas as métricas simultaneamente. O resultado do teste de Friedman revelou evidências estatísticas que pelo menos um dos modelos apresenta superioridade estatística em relação aos demais, considerando um nível de significância de 0.05.

4.3 Teste de Hipótese

O teste de Friedman e o teste de Nemenyi são usados para comparar os modelos de previsão. O teste de Nemenyi é uma ferramenta de comparação múltipla frequentemente empregada após a aplicação de testes não paramétricos com três ou mais fatores.

Usando os resultados obtidos na Tabela 17 para calcular o teste de Nemenyi, em comparação aos modelos que foram previstos, no teste Nemenyi. Nesse teste, em comparação com os modelos, calcula-se entre as métricas estatísticas qual desses modelos contém o menor valor de p . Dentro de todos os modelos de previsão, o RNN foi o modelo que obteve o menor valor registrado nas métricas estatísticas.

O teste de Friedman para o erro SMAPE, com um valor de estatística de teste e um valor de p de $7,29 \times 10^{-28}$, e para o erro MAE um valor de p de $4,78 \times 10^{-26}$, e para o erro RRMSE um valor de p de $1,23 \times 10^{-40}$ indica que há evidências estatísticas sugerindo que, pelo menos, um dos modelos (ou todos) apresenta um desempenho significativamente diferente dos demais. O valor extremamente baixo de p , próximo de zero, sugere que as diferenças observadas não são simplesmente devidas ao acaso. Portanto, há uma diferença estatisticamente significativa entre os grupos testados.

Na análise das tabelas, são identificadas diferenças significativas entre os modelos. Por exemplo, na Tabela 17, há valores baixos na linha DT (G) nas colunas GRU (H), LSTM (I) e RNN (O), indicando diferenças críticas entre esses modelos. Na Tabela 18, os valores na interseção da linha DT (G) com as colunas GRU (H) e RNN (O) são estatisticamente significativos. Na Tabela 19, os valores entre a linha CNN (F) e a coluna DT (G) são baixos, ressaltando outra diferença relevante. Esses resultados evidenciam a necessidade de uma análise minuciosa para determinar o modelo mais adequado.

Ao analisar as tabelas, fica evidente que o modelo DT se destaca como tendo desempenho superior em comparação com os outros modelos. Isso é corroborado pelos valores baixos nas interseções da linha DT (G) com as colunas GRU (H), RNN (O) e LSTM (I) na Tabela 17, bem como pelos valores significativos na Tabela 18. Portanto, o modelo DT parece ser a escolha mais promissora para os propósitos da análise.

Tabela 17: Teste de significância Nemenyi na métrica SMAPE.

| Modelo | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|----------|-------|-------|-------|-------|-------|-------|-------|----------|----------|-------|-------|-------|-------|-------|----------|-------|-------|-------|
| A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,092 | 0,070 | 1 | 1 | 1 | 1 | 0,976 | 0,185 | 1 | 1 | 0,846 |
| B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,093 | 0,071 | 1 | 1 | 1 | 1 | 0,977 | 0,186 | 1 | 1 | 0,847 |
| C | 1 | 1 | 1 | 1 | 1 | 1 | 0,999 | 0,158 | 0,124 | 1 | 1 | 1 | 1 | 0,991 | 0,285 | 1 | 1 | 0,918 |
| D | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,094 | 0,071 | 1 | 1 | 1 | 1 | 0,977 | 0,187 | 1 | 1 | 0,848 |
| E | 1 | 1 | 1 | 1 | 1 | 1 | 0,999 | 0,164 | 0,129 | 1 | 1 | 1 | 1 | 0,992 | 0,294 | 1 | 1 | 0,922 |
| F | 1 | 1 | 1 | 1 | 1 | 1 | 0,971 | 0,456 | 0,394 | 1 | 1 | 1 | 1 | 1 | 0,636 | 1 | 1 | 0,993 |
| G | 1 | 1 | 0,999 | 1,000 | 0,999 | 0,971 | 1 | O | O | 0,940 | 1 | 0,971 | 0,989 | 0,204 | O | 0,999 | 0,999 | 0,051 |
| H | 0,092 | 0,093 | 0,158 | 0,094 | 0,164 | 0,456 | 0 | 1 | 1 | 0,575 | 0,030 | 0,456 | 0,326 | 0,997 | 1 | 0,122 | 0,158 | 1 |
| I | 0,070 | 0,071 | 0,124 | 0,071 | 0,129 | 0,394 | 0 | 1 | 1 | 0,511 | 0,021 | 0,394 | 0,272 | 0,995 | 1 | 0,095 | 0,125 | 1 |
| J | 1 | 1 | 1 | 1 | 1 | 1 | 0,940 | 0,575 | 0,511 | 1 | 1 | 1 | 1 | 1 | 0,744 | 1 | 1 | 0,998 |
| K | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,030 | 0,021 | 1 | 1 | 1 | 1 | 0,903 | 0,071 | 1 | 1 | 0,651 |
| L | 1 | 1 | 1 | 1 | 1 | 1 | 0,971 | 0,456 | 0,394 | 1 | 1 | 1 | 1 | 1 | 0,636 | 1 | 1 | 0,993 |
| M | 1 | 1 | 1 | 1 | 1 | 1 | 0,989 | 0,326 | 0,272 | 1 | 1 | 1 | 1 | 0,999 | 0,499 | 1 | 1 | 0,979 |
| N | 0,976 | 0,977 | 0,991 | 0,977 | 0,992 | 1 | 0,204 | 0,997 | 0,995 | 1 | 0,903 | 1 | 0,999 | 1 | 1 | 0,986 | 0,992 | 1 |
| O | 0,185 | 0,186 | 0,285 | 0,187 | 0,294 | 0,636 | 0 | 1 | 1 | 0,744 | 0,071 | 0,636 | 0,499 | 1 | 1 | 0,232 | 0,286 | 1 |
| P | 1 | 1 | 1 | 1 | 1 | 1 | 0,999 | 0,122 | 0,095 | 1 | 1 | 1 | 1 | 0,986 | 0,232 | 1 | 1 | 0,886 |
| Q | 1 | 1 | 1 | 1 | 1 | 1 | 0,999 | 0,158 | 0,125 | 1 | 1 | 1 | 1 | 0,992 | 0,286 | 1 | 1 | 0,918 |
| R | 0,846 | 0,847 | 0,918 | 0,848 | 0,922 | 0,993 | 0,051 | 1 | 1 | 0,998 | 0,651 | 0,993 | 0,979 | 1 | 1 | 0,886 | 0,918 | 1 |

A – AR, B – ARIMA, C – ARIMAX, D – ARMA, E – ARX, F – CNN, **G** – **DT**, **H** – **GRU**, **I** – **LSTM**, J – LigthGBM, K – MA, L – MLP, M – Prophet, N – RF, **O** – **RNN**, P – SARIMA, Q – SARIMAX, R – XGBoost.

Tabela 18: Teste de significância Nemenyi na métrica MAE.

| Modelo | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|----------|-------|-------|-------|-------|-------|-------|-------|--------------|-------|-------|-------|-------|-------|-------|--------------|-------|-------|-------|
| A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,054 | 0,036 | 1 | 1 | 1 | 1 | 0,947 | 0,102 | 1 | 1 | 0,736 |
| B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,059 | 0,039 | 1 | 1 | 1 | 1 | 0,952 | 0,109 | 1 | 1 | 0,750 |
| C | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,093 | 0,065 | 1 | 1 | 1 | 1 | 0,975 | 0,162 | 1 | 1 | 0,829 |
| D | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,058 | 0,039 | 1 | 1 | 1 | 1 | 0,951 | 0,108 | 1 | 1 | 0,748 |
| E | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,101 | 0,071 | 1 | 1 | 1 | 1 | 0,978 | 0,175 | 1 | 1 | 0,843 |
| F | 1 | 1 | 1 | 1 | 1 | 1 | 0,999 | 0,752 | 0,679 | 1 | 1 | 1 | 1 | 1,000 | 0,855 | 1 | 1 | 1 |
| G | 1 | 1 | 1 | 1 | 1 | 0,999 | 1 | 0,012 | 0,008 | 0,999 | 1 | 0,999 | 1 | 0,803 | 0,027 | 1 | 1 | 0,467 |
| H | 0,054 | 0,059 | 0,093 | 0,058 | 0,101 | 0,752 | 0,012 | 1 | 1 | 0,732 | 0,022 | 0,752 | 0,217 | 0,997 | 1,000 | 0,062 | 0,093 | 1 |
| I | 0,036 | 0,039 | 0,065 | 0,039 | 0,071 | 0,679 | 0,008 | 1 | 1 | 0,657 | 0,014 | 0,679 | 0,163 | 0,995 | 1,000 | 0,042 | 0,064 | 1 |
| J | 1 | 1 | 1 | 1 | 1 | 1 | 0,999 | 0,732 | 0,657 | 1 | 1 | 1 | 1 | 1 | 0,840 | 1 | 1 | 0,999 |
| K | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,022 | 0,014 | 1 | 1 | 1 | 1 | 0,870 | 0,046 | 1 | 1 | 0,570 |
| L | 1 | 1 | 1 | 1 | 1 | 1 | 0,999 | 0,752 | 0,679 | 1 | 1 | 1 | 1 | 1 | 0,855 | 1 | 1 | 1 |
| M | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,217 | 0,163 | 1 | 1 | 1 | 1 | 0,996 | 0,331 | 1 | 1 | 0,942 |
| N | 0,947 | 0,952 | 0,975 | 0,951 | 0,978 | 1 | 0,803 | 0,997 | 0,995 | 1 | 0,870 | 1,000 | 0,996 | 1 | 0,999 | 0,955 | 0,975 | 1 |
| O | 0,102 | 0,109 | 0,162 | 0,108 | 0,175 | 0,855 | 0,027 | 1,000 | 1,000 | 0,840 | 0,046 | 0,855 | 0,331 | 0,999 | 1 | 0,114 | 0,162 | 1 |
| P | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,062 | 0,042 | 1 | 1 | 1 | 1 | 0,955 | 0,114 | 1 | 1 | 0,759 |
| Q | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,093 | 0,064 | 1 | 1 | 1 | 1 | 0,975 | 0,162 | 1 | 1 | 0,828 |
| R | 0,736 | 0,750 | 0,829 | 0,748 | 0,843 | 1 | 0,467 | 1,000 | 1,000 | 0,999 | 0,570 | 1 | 0,942 | 1 | 1 | 0,759 | 0,828 | 1 |

A – AR, B – ARIMA, C – ARIMAX, D – ARMA, E – ARX, F – CNN, **G** – **DT**, **H** – **GRU**, I – LSTM, J – LigthGBM, K – MA, L – MLP, M – Prophet, N – RF, **O** – **RNN**, P – SARIMA, Q – SARIMAX, R – XGBoost.

Tabela 19: Teste de significância Nemenyi na métrica RRMSE.

| Modelo | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|----------|-------|-------|-------|-------|-------|-------|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| A | 1 | 1 | 1 | 1 | 1 | 0,306 | 1 | 0,028 | 0,001 | 1 | 1 | 0,306 | 0,414 | 0,993 | 0,003 | 1 | 1 | 0,975 |
| B | 1 | 1 | 1 | 1 | 1 | 0,314 | 1 | 0,029 | 0,001 | 1 | 1 | 0,314 | 0,423 | 0,994 | 0,003 | 1 | 1 | 0,976 |
| C | 1 | 1 | 1 | 1 | 1 | 0,413 | 1 | 0,050 | 0,002 | 1 | 1 | 0,413 | 0,529 | 0,998 | 0,007 | 1 | 1 | 0,989 |
| D | 1 | 1 | 1 | 1 | 1 | 0,314 | 1 | 0,029 | 0,001 | 1 | 1 | 0,314 | 0,423 | 0,994 | 0,003 | 1 | 1 | 0,976 |
| E | 1 | 1 | 1 | 1 | 1 | 0,405 | 1 | 0,048 | 0,002 | 1 | 1 | 0,405 | 0,521 | 0,998 | 0,007 | 1 | 1 | 0,989 |
| F | 0,306 | 0,314 | 0,413 | 0,314 | 0,405 | 1 | 0,023 | 1 | 0,999 | 0,778 | 0,133 | 1 | 1 | 1 | 1 | 0,342 | 0,412 | 1 |
| G | 1 | 1 | 1 | 1 | 1 | 0,023 | 1 | 0 | 0 | 0,999 | 1 | 0,023 | 0,043 | 0,759 | 0 | 1 | 1 | 0,600 |
| H | 0,028 | 0,029 | 0,050 | 0,029 | 0,048 | 1 | 0 | 1 | 1 | 0,231 | 0,006 | 1 | 1 | 0,938 | 1 | 0,034 | 0,050 | 0,979 |
| I | 0,001 | 0,001 | 0,002 | 0,001 | 0,002 | 0,999 | 0 | 1 | 1 | 0,022 | 0 | 0,999 | 0,998 | 0,535 | 1 | 0,001 | 0,002 | 0,703 |
| J | 1 | 1 | 1 | 1 | 1 | 0,778 | 0,999 | 0,231 | 0,022 | 1 | 1 | 0,778 | 0,860 | 1,000 | 0,057 | 1 | 1 | 1 |
| K | 1 | 1 | 1 | 1 | 1 | 0,133 | 1 | 0,006 | 0 | 1 | 1 | 0,133 | 0,203 | 0,958 | 0,001 | 1 | 1 | 0,894 |
| L | 0,306 | 0,314 | 0,413 | 0,314 | 0,405 | 1 | 0,023 | 1 | 0,999 | 0,778 | 0,133 | 1 | 1 | 1 | 1 | 0,342 | 0,412 | 1 |
| M | 0,414 | 0,423 | 0,529 | 0,423 | 0,521 | 1 | 0,043 | 1 | 0,998 | 0,860 | 0,203 | 1 | 1 | 1 | 1 | 0,454 | 0,528 | 1 |
| N | 0,993 | 0,994 | 0,998 | 0,994 | 0,998 | 1 | 0,759 | 0,938 | 0,535 | 1 | 0,958 | 1 | 1 | 1 | 0,715 | 0,995 | 0,998 | 1 |
| O | 0,003 | 0,003 | 0,007 | 0,003 | 0,007 | 1 | 0 | 1 | 1 | 0,057 | 0,001 | 1 | 1 | 0,715 | 1 | 0,004 | 0,007 | 0,848 |
| P | 1 | 1 | 1 | 1 | 1 | 0,342 | 1 | 0,034 | 0,001 | 1 | 1 | 0,342 | 0,454 | 0,995 | 0,004 | 1 | 1 | 0,981 |
| Q | 1 | 1 | 1 | 1 | 1 | 0,412 | 1 | 0,050 | 0,002 | 1 | 1 | 0,412 | 0,528 | 0,998 | 0,007 | 1 | 1 | 0,989 |
| R | 0,975 | 0,976 | 0,989 | 0,976 | 0,989 | 1 | 0,600 | 0,979 | 0,703 | 1 | 0,894 | 1 | 1 | 1 | 0,848 | 0,981 | 0,989 | 1 |

A – AR, B – ARIMA, C – ARIMAX, D – ARMA, E – ARX, **F** – **CNN**, **G** – **DT**, H – GRU, I – LSTM, J – LigthGBM, K – MA, L – MLP, M – Prophet, N – RF, O – RNN, P – SARIMA, Q – SARIMAX, R – XGBoost.

O valor crítico CD foi utilizado para determinar se dois modelos eram significativamente diferentes entre si. Esse valor é calculado com base no valor crítico obtido das Tabelas 17 a 19 de teste de Nemenyi, o número de modelos e o número total de amostras. O valor CD é uma métrica que auxilia na interpretação das diferenças entre os modelos, ajudando a identificar quais pares de modelos apresentam diferenças estatisticamente significativas.

Os resultados da pesquisa indicam a existência de evidências estatísticas que sugerem a superioridade de pelo menos um modelo em relação aos demais. A análise de comparação significativa entre os modelos revelou pares de modelos que apresentam diferenças estatisticamente significativas em seus desempenhos, conforme exibido nas Tabelas 17 a 19, onde cada valor com três casas decimais após a vírgula representa modelos que têm significância estatística entre si.

4.3.1 Comparação dos Modelos

Com o objetivo de obter uma análise mais aprofundada do desempenho de cada modelo, foi realizada uma comparação por meio de um gráfico de barra. Dessa forma, pôde-se observar qual dos modelos apresentava o melhor desempenho.

Ao examinar os modelos representados nas Figuras 27 a 32 identificou os modelos que se destacam em relação à natureza dos dados. Nessas Figuras compara os modelos ARIMA e XGBoost com outros, torna-se evidente que os modelos ARIMA como AR, ARX, MA, ARMA, ARIMAX e SARIMAX demonstram um desempenho sólido. Os modelos baseados em gradientes e regressão, como o XGBoost, exibem resultados comparáveis, as redes neurais com o modelo Prophet, é importante destacar que os modelos de redes neurais, incluindo RNN, LSTM, GRU, MLP e CNN, foram avaliados em conjunto com o modelo Prophet. A análise das métricas demonstrou que o modelo DT se sobressai como o vencedor entre as métricas avaliadas. Essa conclusão é respaldada pelas evidências de que pelo menos um modelo é superior aos demais. Os modelos com valores de valor de p abaixo de 0,05 foram realçados em **negrito** para enfatizar sua significância. Beneficiando-se da otimização por meio do Optuna, uma abordagem de bayesiana usando o método TPE.

Figura 27: Comparação dos modelos AR, ARX, MA, ARIMA, SARIMA, ARIMAX, SARIMAX, DT, RF, XGBoost, LightGBM em vários horizontes na métrica SMAPE.

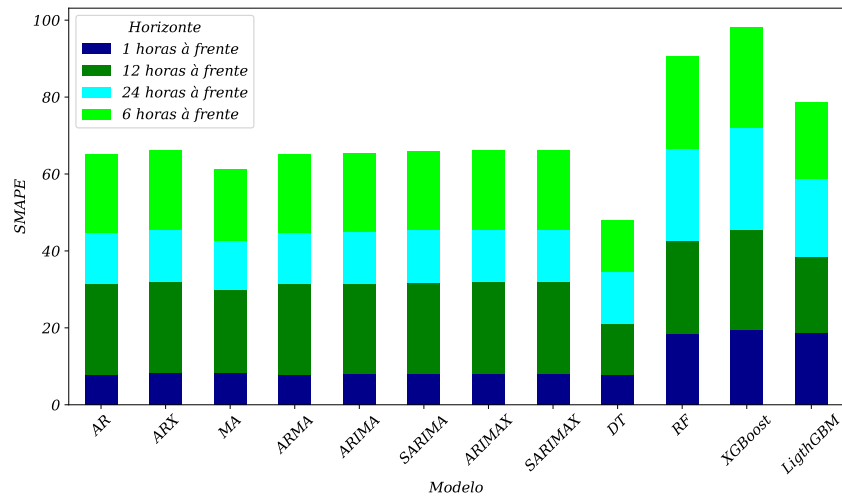


Figura 28: Comparação dos modelos AR, ARX, MA, ARIMA, SARIMA, ARIMAX, SARIMAX, DT, RF, XGBoost, LightGBM em vários horizontes na métrica MAE.

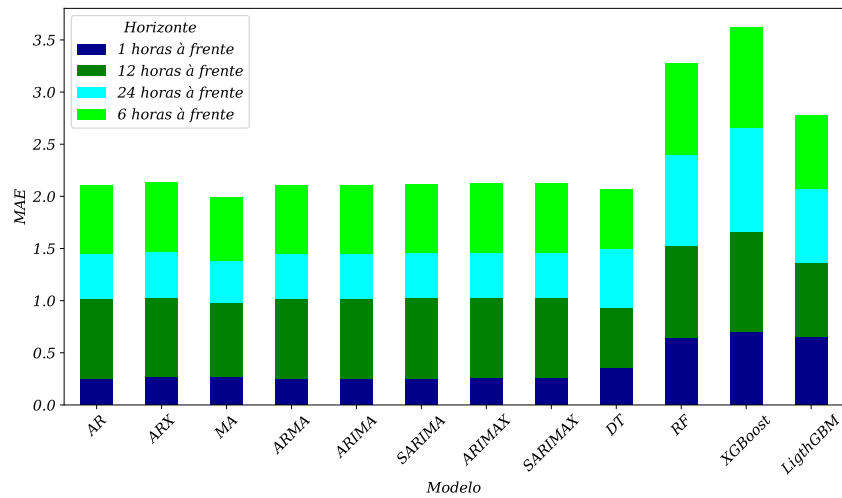


Figura 29: Comparação dos modelos AR, ARX, MA, ARIMA, SARIMA, ARIMAX, SARIMAX, DT, RF, XGBoost, LightGBM em vários horizontes na métrica RRMSE.

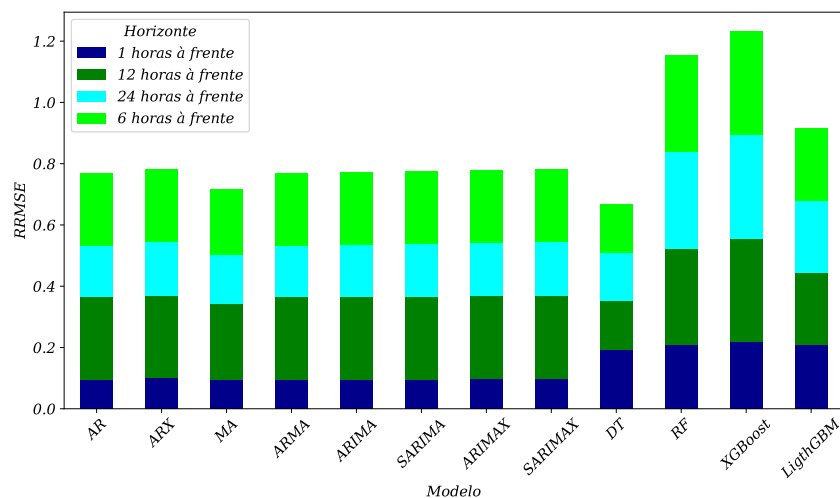


Figura 30: Comparação dos modelos Prophet, MLP, CNN, RNN, LSTM, GRU em vários horizontes na métrica SMAPE.

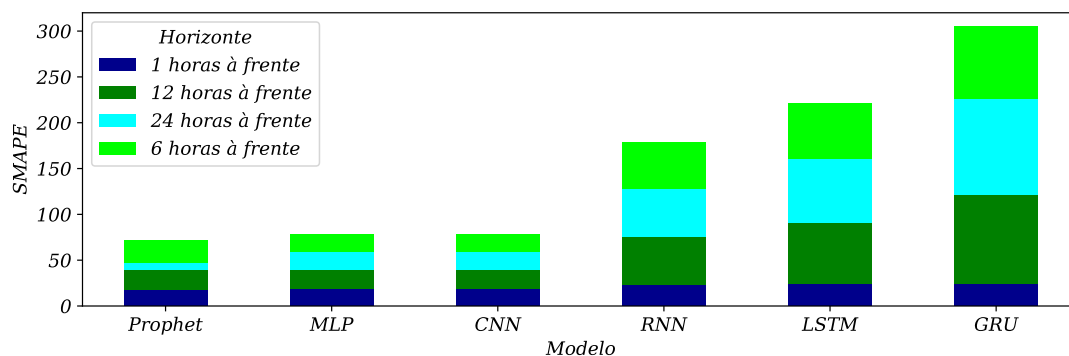


Figura 31: Comparação dos modelos Prophet, MLP, CNN, RNN, LSTM, GRU em vários horizontes na métrica MAE.

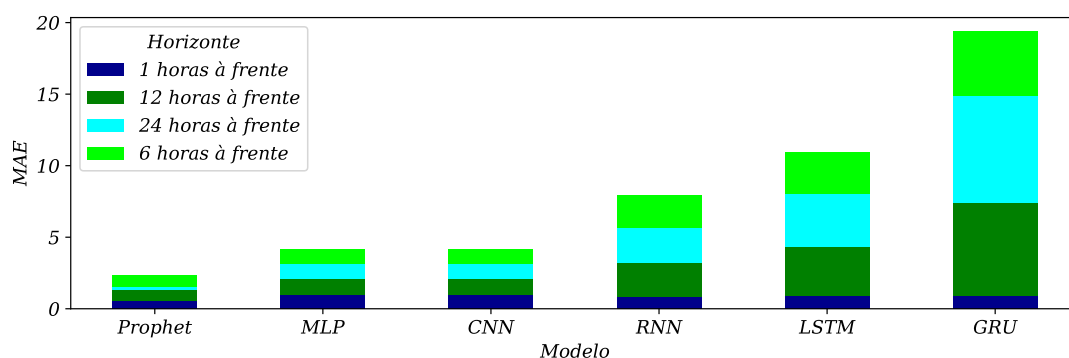
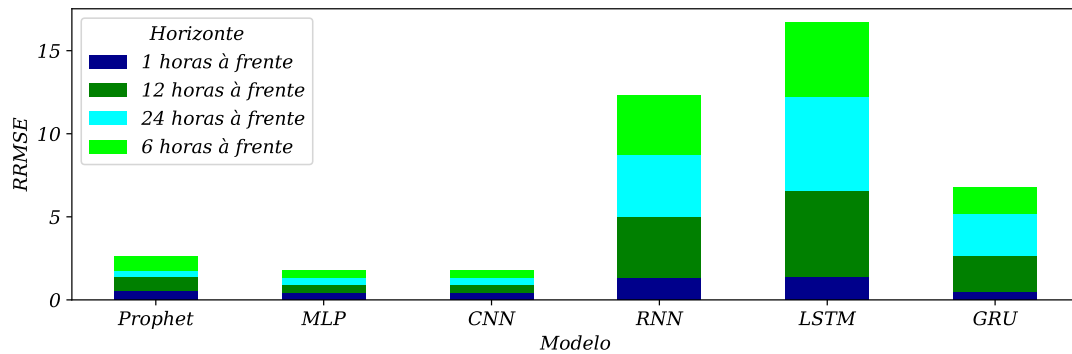


Figura 32: Comparação dos modelos Prophet, MLP, CNN, RNN, LSTM, GRU em vários horizontes na métrica RRMSE.



A avaliação da eficácia dos modelos ARIMA em previsões de longo prazo emprega o teste de Ljung-Box, conforme detalhado nas Tabelas 20a a 20d ilustram a acurácia dos modelos ARIMA ao longo do tempo, com valores menores sendo destacados em **negrito**. Modelos tais como ARX, ARIMAX e SARIMAX, que incorporam variáveis exógenas, demonstram um desempenho superior nesse contexto. Esses modelos não lineares apresentam uma capacidade de previsão robusta em horizontes temporais mais longos, diferenciando-se positivamente dos outros modelos ARIMA. Nas Figuras 27 a 29, são selecionados os modelos ARIMA e seus antecessores. Esses modelos têm suas limitações, tanto para horizontes de previsão de curto prazo quanto para horizontes de longo prazo. Esse gráfico pode fornecer várias informações, mas o objetivo aqui é identificar apenas o melhor modelo entre os modelos ARIMA.

Como essa série não apresentou uma estacionariedade bem definida e os dados não a tornaram estacionária, os modelos que não têm sazonalidade mostraram-se superiores, tais como AR, MA, ARX, ARMA, ARIMA e ARIMAX. O modelo ARIMAX demonstrou ser robusto para este caso, mas mesmo assim, modelos mais básicos como AR, ARMA e ARIMA ainda apresentaram resultados melhores.

Tabela 20: Comparação dos modelos com o teste Ljung Box modelos ARIMA com defasagem de 10 para previsão de longo prazo na demanda d'água.

| (a) Treinamento. | | | (b) Validação. | | |
|-------------------------|-------------------------|---------------|-----------------------|-------------------------|---------------|
| Ljung Box | Estatística de Teste | Valor de p | Ljung Box | Estatística de Teste | Valor de p |
| ARX | 86,332 | 0 | ARX | 6,514 | 0,770 |
| AR | 84,118 | 0 | AR | 0,764 | 0,376 |
| MA | 87,940 | 0 | MA | 76,952 | 0 |
| ARMA | 26,204 | 0,003 | ARMA | 4,900 | 0,898 |
| ARIMA | 21,4 | 0,018 | ARIMA | 11,025 | 0,356 |
| SARIMA | 97,735 | 0 | SARIMA | 68,136 | 0 |
| ARIMAX | 42,151 | 0 | ARIMAX | 10,809 | 0,373 |
| SARIMAX | 47,310 | 0 | SARIMAX | 13,287 | 0 |

| (c) Teste. | | | (d) Inteiro. | | |
|-------------------|-------------------------|---------------|---------------------|-------------------------|---------------|
| Ljung Box | Estatística de Teste | Valor de p | Ljung Box | Estatística de Teste | Valor de p |
| ARX | 36,884 | 0 | ARX | 22,944 | 0 |
| AR | 36,799 | 0 | AR | 22,715 | 0 |
| MA | 71,945 | 0 | MA | 5725,300 | 0 |
| ARMA | 9,706 | 0,467 | ARMA | 72,446 | 0 |
| ARIMA | 60,289 | 0 | ARIMA | 72,446 | 0 |
| SARIMA | 70,992 | 0 | SARIMA | 72,446 | 0 |
| ARIMAX | 48,435 | 0 | ARIMAX | 84,362 | 0 |
| SARIMAX | 45,749 | 0 | SARIMAX | 23,537 | 0 |

4.4 Aplicação

A previsão da demanda d'água é uma preocupação fundamental para muitas organizações e autoridades responsáveis pelo abastecimento d'água. Neste estudo de caso, explorou-se como a análise de séries temporais pode ser aplicada para prever a demanda d'água ao longo do tempo.

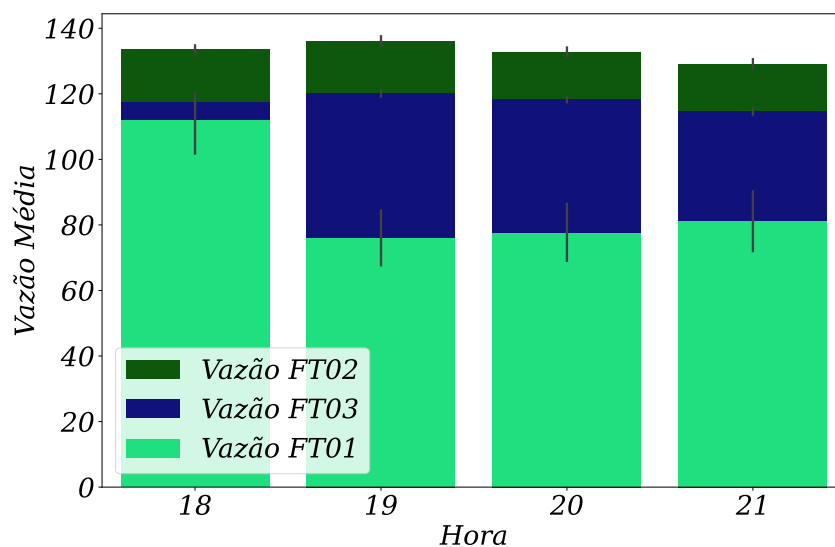
4.4.1 Estudo de Caso 1 - Análise de Padrões de Consumo

Questão de Pesquisa: Existe tendência, padrão, sazonalidade para os dados destes três anos do Bairro Alto?

Na análise dos dados dos últimos 2 anos do Bairro Alto utilizando o método de decomposição STL mostrado na Figura 14, constatou-se que identificar tendências, sazonalidades e padrões de consumo d'água não é uma tarefa fácil. Essa dificuldade em compreender os dados torna-se um desafio na hora de planejar o abastecimento de forma eficiente.

O gráfico de barras apresentado na Figura 33 mostra a demanda média das variáveis de fluxo (Vazão de Entrada – FT01, Vazão de Gravidade – FT02 e Vazão de Recalque – FT03) durante o intervalo das 18h às 21h. Cada barra representa a média da demanda para cada variável em um horário específico dentro desse intervalo. A altura de cada barra indica a magnitude da demanda média para a respectiva variável. Essa visualização permite que seja identificado o horário em que as variáveis de fluxo apresentaram maior demanda, o que é útil para o planejamento e gerenciamento adequado do sistema.

Figura 33: Demanda média das variáveis de fluxo.



4.4.2 Estudo de Caso 2 – Estratégias de Economia de Energia

Questão de Pesquisa: Qual é o impacto do acionamento das bombas durante o horário de pico?

Confirma-se que a ativação das bombas de sucção durante o período de 18h às 21h resulta em um maior custo energético para a SANEPAR. Portanto, é recomendado evitar o acionamento das bombas durante esse período, utilizando estratégias de armazenamento e gerenciamento eficientes.

Questão de Pesquisa: Qual é o nível ideal no reservatório para evitar a ativação das bombas da SANEPAR durante o período de maior demanda, das 18h às 21h, sem comprometer o abastecimento d'água para a população?

Verifica-se que, para evitar o acionamento das bombas durante o horário de pico (18h às 21h) sem comprometer o abastecimento d'água para a população, é necessário manter o nível do reservatório acima de 4.445 litros.

A Tabela 21 apresenta os resultados para as três variáveis estudadas: vazão de entrada – FT01, vazão de gravidade – FT02 e vazão de recalque – FT03. Os resultados destacam os horários específicos em que cada variável apresentou maior demanda dentro do intervalo das 18h às 21h, fornecendo importantes informações para o planejamento e gerenciamento adequado do sistema.

Tabela 21: Demanda d'água.

| Vazões | Horário de Maior Demanda | Demanda |
|------------------|--------------------------|----------------|
| Entrada – FT01 | 2020/10/08 21:00:00 | 383,87 m^3/h |
| Gravidade – FT02 | 2020/10/20 18:00:00 | 326,17 m^3/h |
| Recalque – FT03 | 2020/11/26 19:00:00 | 194,35 m^3/h |

Durante as horas de pico, é necessário que o nível do reservatório esteja mantido dentro na média de 4.445 litros para evitar o acionamento das bombas. Manter o nível do reservatório dentro dessa faixa permitirá que o sistema opere de forma eficiente, atendendo à demanda d'água sem a necessidade de acionar as bombas.

É importante destacar que a vazão de recalque exerce um impacto significativo no nível do reservatório em comparação com as outras vazões. Essa diferença se deve ao fato de que a vazão de recalque está diretamente relacionada à injeção d'água no reservatório por meio da bomba localizada próxima à sua base. Em contraste, as demais vazões possuem alguns valores ausentes, o que limita sua influência na análise geral do sistema.

5 Conclusões

Nessa dissertação teve como objetivo, conduziu-se um estudo abrangente sobre a previsão da demanda d'água por meio da análise de séries temporais. Através da análise exploratória dos dados e da aplicação da decomposição STL, identificaram-se padrões sazonais e tendências na demanda d'água. Durante o estudo, os modelos DT e Prophet foram empregados para validar o estudo de caso da SANEPAR.

Os objetivos específicos do estudo foram alcançados, contribuindo significativamente para a compreensão e aplicação de modelos de previsão de séries temporais na gestão da demanda d'água.

Aplicar diferentes modelos de previsão de séries temporais utilizando dados provenientes do Bairro Alto em Curitiba, fornecidos pela SANEPAR. A aplicação de diferentes modelos de previsão, utilizando dados específicos do Bairro Alto em Curitiba fornecidos pela SANEPAR, permitiu avaliar sua precisão, eficiência e capacidade de previsão em conjuntos de dados específicos, utilizando medidas de desempenho para análise.

Avaliar a precisão, eficiência e capacidade de previsão desses modelos em conjuntos de dados específicos, utilizando métricas para análise de desempenho.

Explorar estratégias de otimização baseadas em otimização Bayesiana, empregando o algoritmo TPE para ajustar os hiperparâmetros dos modelos de previsão de séries temporais. A exploração de estratégias de otimização baseadas em otimização Bayesiana, empregando o algoritmo TPE para ajustar os hiperparâmetros dos modelos, resultou na identificação de combinações eficazes de modelos de previsão de séries temporais em conjunto com a configuração otimizada.

Identificar combinações eficazes de modelos de previsão de séries temporais em conjunto com a configuração otimizada. Avaliar o impacto das variáveis exógenas na melhoria da precisão dos modelos de previsão de séries temporais. Os resultados obtidos demonstram que a abordagem de séries temporais é uma ferramenta eficaz para prever a demanda futura d'água, destacando a importância de considerar as flutuações sazonais e as diferentes partes do dia ao determinar a vazão e o volume mínimo de reserva no reservatório.

Esses objetivos específicos não apenas cumpriram sua função original, mas também forneceram informações para a otimização do planejamento e operação de sistemas de abastecimento d'água, promovendo a resiliência diante de variações sazonais e eventos imprevistos. Essa pesquisa pode ser aplicado na prática, auxiliando gestores e tomadores de decisão na implementação de estratégias mais eficientes para garantir o abastecimento d'água de forma sustentável e resiliente.

O estudo de casos separados em dois tipos de análise de padrões e consumo e o

outro estratégia de economia de energia tem como foco melhorar o consumo d'água em horários de pico e otimizar a demanda d'água para que não seja afetada novamente no futuro e falte água novamente. Usando os modelos, consegue-se responder a três questões de pesquisa feitas nessa dissertação. A primeira foi entender os dados se existe tendência, padrão e sazonalidade, para o tratamento que foi realizado, o qual foi muito bom para entender os dados.

A segunda questão já teve a ver com o impacto no acionamento das bombas em horário de pico, confirmando assim que no horário de pico o gasto de energia e a demanda d'água são maiores, mas com uma pequena estratégia pode-se contornar isso facilmente. A terceira questão é quanto deve ser armazenado previamente no reservatório LT01 para que não seja necessário ter esse gasto de energia, 4.445 litros é o suficiente para que o reservatório passe o horário de pico com o menor gasto energético.

5.1 Propostas Futuras

Apesar dos resultados promissores evidenciados por esta pesquisa, é essencial que se reconheçam suas limitações e que se instigue a exploração de novos horizontes em pesquisas subsequentes.

Uma análise mais profunda e abrangente pode ser realizada, investigando modelos de redes neurais mais avançados. A implementação de técnicas de otimização matemática mais refinadas, como o uso do método *Covariance Matrix Adaptation Evolution Strategy* (CMAES), pode ser considerada.

Seria prudente incluir cuidadosamente variáveis exógenas em todos os modelos pertinentes, como o uso de variáveis climáticas e dados de precipitação do tempo. Implementa modelos que utilizam sistemas *fuzzy* para aprimorar a previsão do tanque. Usa essa previsão juntamente com modelos existentes na literatura, como a otimização *Bayesian Optimization Algorithm* (BOA), que não foi abordada neste contexto.

Referências

- AGIAKLOGLOU, C.; NEWBOLD, P. Empirical evidence on dickey-fuller-type tests. **Journal of Time Series Analysis**, v. 13, n. 6, p. 471–483, 1992.
- AHMAD, T. et al. A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review. **ENERGY AND BUILDINGS**, v. 165, p. 301–320, 2018. ISSN 0378-7788.
- AL-SHABI, M. Q. Machine learning: Algorithms, real-world applications and research directions. **SN Computer Science**, Springer, v. 2, n. 3, p. 1–12, 2021.
- ALI, M. et al. Ensemble robust local mean decomposition integrated with random forest for short-term significant wave height forecasting. **Renewable Energy**, v. 205, p. 731–746, 2023.
- ASEERI, A. O. Effective rnn-based forecasting methodology design for improving short-term power load forecasts: Application to large-scale power-grid time series. **Journal of Computational Science**, v. 68, p. 101984, 2023.
- BABU, C. N.; REDDY, B. E. A moving-average filter based hybrid arima-ann model for forecasting time series data. **Applied Soft Computing**, v. 23, p. 27–38, 2014.
- BANDARA, K.; HYNDMAN, R. J.; BERGMEIR, C. Mstl: A seasonal-trend decomposition algorithm for time series with multiple seasonal patterns. **arXiv preprint arXiv:2107.13462**, 2021. Disponível em: <<https://arxiv.org/abs/2107.13462>>.
- BARNES, J. L.; KRISHEN, A. S.; HU, H. fen. Public tap water perceptions and potable reuse acceptance: A cognitive dissonance theoretical understanding. **Journal of Cleaner Production**, v. 429, p. 139587, 2023.
- BERGLUND, E. Z.; SKARBEEK, M.; KANTA, L. A sociotechnical framework to characterize tipping points in water supply systems. **Sustainable Cities and Society**, v. 97, p. 104739, 2023.
- BHANGU, K.; SANDHU, J.; SAPRA, L. Time series analysis of covid-19 cases. **World Journal of Engineering**, v. 19, p. 40–48, 2022. ISSN 17085284.
- BOX, G. E. P.; PIERCE, D. A. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. **Journal of the American Statistical Association**, v. 65, n. 332, p. 1509–1526, 1970. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/01621459.1970.10481180>>.
- BROCKWELL, P. J.; DAVIS, R. A. **Introduction to Time Series and Forecasting**. [S.l.]: Springer New York, 2002. ISBN 978-0-387-95351-9.
- BUEECHI, E. et al. Crop yield anomaly forecasting in the pannonian basin using gradient boosting and its performance in years of severe drought. **Agricultural and Forest Meteorology**, v. 340, p. 109596, 2023.
- BUYUKSAHIN, U.; ERTEKIN. Improving forecasting accuracy of time series data using a new arima-ann hybrid method and empirical mode decomposition. **Neurocomputing**, v. 361, p. 151–163, 2019. ISSN 09252312.

CANDELIERI, A. et al. Tuning hyperparameters of a svm-based water demand forecasting system through parallel global optimization. **Computers & Operations Research**, v. 106, p. 202–209, 2019.

Cesar de Lima Nogueira, S. et al. Prediction of the nox and co2 emissions from an experimental dual fuel engine using optimized random forest combined with feature engineering. **Energy**, v. 280, p. 128066, 2023. ISSN 0360-5442. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0360544223014603>>.

CHEN, M. et al. The predictive management in campus heating system based on deep reinforcement learning and probabilistic heat demands forecasting. **Applied Energy**, v. 350, p. 121710, 2023.

CHEN, Y. Y. et al. Applications of Recurrent Neural Networks in Environmental Factor Forecasting: A Review. **NEURAL COMPUTATION**, v. 30, n. 11, p. 2855–2881, 2018. ISSN 0899-7667.

CHOU, J.-S.; TRAN, D.-S. Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders. **Energy**, v. 165, p. 709–726, 2018.

COELHO, L. dos S.; AYALA, H. V. H.; MARIANI, V. C. Co and nox emissions prediction in gas turbine using a novel modeling pipeline based on the combination of deep forest regressor and feature engineering. **Fuel**, v. 355, p. 129366, 2024. ISSN 0016-2361. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0016236123019804>>.

DAVIDSON, J. **Econometric Theory**. Wiley-Blackwell, 2000. ISBN 978-0-631-21584-4. Disponível em: <<https://www.wiley.com/en-cn/Econometric+Theory-p-9780631215844>>.

DAVIES, N.; NEWBOLD, P. Some power studies of a portmanteau test of time series model specification. **Biometrika**, v. 66, n. 1, p. 153–155, 04 1979. ISSN 0006-3444. Disponível em: <<https://doi.org/10.1093/biomet/66.1.153>>.

DONG, J. et al. Enhancing short-term forecasting of daily precipitation using numerical weather prediction bias correcting with xgboost in different regions of china. **Engineering Applications of Artificial Intelligence**, v. 117, p. 105579, 2023.

FOUILLOY, A. et al. Solar irradiation prediction with machine learning: Forecasting models selection method depending on weather variability. **Energy**, v. 165, p. 620–629, 2018. ISSN 03605442.

GIFFORD, M.; BAYRAK, T. A predictive analytics model for forecasting outcomes in the national football league games using decision tree and logistic regression. **Decision Analytics Journal**, v. 8, p. 100296, 2023.

GRAFF, M. et al. Time series forecasting with genetic programming. **Natural Computing**, v. 16, n. 1, p. 165–174, 2017.

GRÜBLER, M. Entendendo o funcionamento de uma rede neural artificial. **Medium**, Jun 2018. Disponível em: <<https://medium.com/brasil-ai/entendendo-o-funcionamento-de-uma-rede-neural-artificial-4463fcf44dd0>>.

GUO, H.; PEDRYCZ, W.; LIU, X. Hidden markov models based approaches to long-term prediction for granular time series. **IEEE Transactions on Fuzzy Systems**, v. 26, p. 2807–2817, 2018. ISSN 10636706.

GUSTIN, M.; MCLEOD, R.; LOMAS, K. Forecasting indoor temperatures during heatwaves using time series models. **Building and Environment**, v. 143, p. 727–739, 2018. ISSN 03601323.

GÉRON, A. **Hands-On Machine Learning with Scikit-Learn and TensorFlow**. 2nd. ed. O'Reilly Media, 2017. Disponível em: <<https://www.oreilly.com/library/view/hands-on-machine-learning/9781491962282/>>.

HAO, J. et al. A bi-level ensemble learning approach to complex time series forecasting: Taking exchange rates as an example. **Journal of Forecasting**, v. 42, p. 1385–1406, 2023. ISSN 02776693.

HVY. “multivariate” tpe makes optuna even more powerful. **Optuna**, 2020. Disponível em: <<https://medium.com/optuna/multivariate-tpe-makes-optuna-even-more-powerful-63c4bfbae2>>.

JI, S.; AHN, K.-H. Temperature change-informed future multisite streamflow generation to support water supply vulnerability assessments under climate change. **Journal of Hydrology**, v. 624, p. 129928, 2023.

JORDAN, I. D.; SOKÓŁ, P. A.; PARK, I. M. Gated recurrent units viewed through the lens of continuous time dynamical systems. **Frontiers in Computational Neuroscience**, v. 15, p. 678158, 2021. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fncom.2021.678158/full>>.

KHAN, M. et al. Cyclic gate recurrent neural networks for time series classification with missing values. **Neural Processing Letters**, v. 55, n. 1, p. 1–32, 2022. Disponível em: <<https://link.springer.com/article/10.1007/s11063-022-10950-2>>.

KHEIRI, K.; KARIMI, H. Sentimentgpt: Exploiting gpt for advanced sentiment analysis and its departure from current machine learning. **arXiv preprint arXiv:2307.10234**, 2023.

KOEBELE, E. A. et al. A role for water markets in enhancing water security in the western United States?: Lessons from the Walker River Basin. **Water Policy**, v. 24, n. 11, p. 1757–1771, 2022.

KORSTANJE, J. **Advanced Forecasting with Python**. [S.l.]: Springer, 2021.

KOTHONA, D. et al. Deep learning forecasting tool facilitating the participation of photovoltaic systems into day-ahead and intra-day electricity markets. **Sustainable Energy, Grids and Networks**, v. 36, p. 101149, 2023.

KULSHRESHTHA, S.; VIJAYALAKSHMI, A. An arima-lstm hybrid model for stock market prediction using live data. **Journal of Engineering Science and Technology Review**, v. 13, p. 117–123, 2020. ISSN 17919320.

KUSHWAH, A.; WADHVANI, R. Trend triplet based data clustering for eliminating nonlinear trend components of wind time series to improve the performance of statistical forecasting models. **Multimedia Tools and Applications**, v. 81, p. 33927–33953, 2022. ISSN 13807501.

LI, P. et al. Dynamic similar sub-series selection method for time series forecasting. **IEEE Access**, v. 6, p. 32532–32542, 2018. ISSN 21693536.

LIU, H. et al. Dual-stage time series analysis on multifeature adaptive frequency domain modeling. **International Journal of Intelligent Systems**, v. 37, p. 7837–7856, 2022. ISSN 08848173.

LIU, J.; FU, Y. Renewable energy forecasting: A self-supervised learning-based transformer variant. **Energy**, v. 284, p. 128730, 2023.

LIU, J.; XU, Y. T-friedman test: A new statistical test for multiple comparison with an adjustable conservativeness measure. **International Journal of Computational Intelligence Systems**, v. 15, p. 29–43, 2022. Disponível em: <<https://doi.org/10.1007/s44196-022-00083-8>>.

LJUNG, G. M.; BOX, G. E. P. On a measure of lack of fit in time series models. **Biometrika**, v. 65, n. 2, p. 297–303, 08 1978. ISSN 0006-3444. Disponível em: <<https://doi.org/10.1093/biomet/65.2.297>>.

LUCAS, P. d. O. e. **Previsão de Séries Temporais de Evapotranspiração de Referência com Redes Neurais Convolucionais**. Tese (Doutorado) — Universidade Federal de Minas Gerais, 2019. Disponível em: <<https://www.ppgee.ufmg.br/defesas/1748M.PDF>>.

MARTINS, L. E. G.; GORSCHKE, T. Requirements engineering for safety-critical systems: A systematic literature review. **Information and Software Technology**, v. 75, p. 71–89, 2016. ISSN 0950-5849. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0950584916300568>>.

MIGLIATO, A. L. T.; PONTI, M. A. **Detecção de Outliers em Dados não Vistos de Séries Temporais por meio de Erros de Predição com SARIMA e Redes Neurais Recorrentes LSTM e GRU**. Dissertação (Mestrado) — Universidade de São Paulo, 2021.

MOHAN, S. et al. Predicting the impact of the third wave of covid-19 in india using hybrid statistical machine learning models: A time series forecasting and sentiment analysis approach. **Computers in Biology and Medicine**, v. 144, 2022. ISSN 00104825.

NASIRI, H.; EBADZADEH, M. M. Multi-step-ahead stock price prediction using recurrent fuzzy neural network and variational mode decomposition. **Applied Soft Computing**, v. 148, p. 110867, 2023.

O'DONNCHA, F. et al. A spatio-temporal lstm model to forecast across multiple temporal and spatial scales. **Ecological Informatics**, v. 69, 2022. ISSN 15749541.

- OLIVEIRA, P. J.; STEFFEN, J. L.; CHEUNG, P. Parameter estimation of seasonal arima models for water demand forecasting using the harmony search algorithm. **Procedia Engineering**, v. 186, p. 177–185, 2017. XVIII International Conference on Water Distribution Systems, WDSA2016.
- PAIVA, D. d. A.; SÁFADI, T. Study of tests for trend in time series. **Brazilian Journal of Biometrics**, v. 39, n. 2, p. 311–333, Jun. 2021. Disponível em: <<https://biometria.ufra.br/index.php/BBJ/article/view/471>>.
- PAWŁOWSKI, A. et al. Model predictive control using miso approach for drug co-administration in anesthesia. **Journal of Process Control**, v. 117, p. 98–111, 2022. ISSN 0959-1524. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0959152422001287>>.
- PELLETIER, C. et al. Assessing the robustness of random forests to map land cover with high resolution satellite image time series over large areas. **Remote Sensing of Environment**, v. 187, p. 156–168, 2016. Cited By 296. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84992151859&doi=10.1016%2fj.rse.2016.10.010&partnerID=40&md5=09efc79bab8e893b97fd21cb4844b98d>>.
- PETROPOULOS, F. et al. Forecasting: theory and practice. **International Journal of Forecasting**, v. 38, n. 3, p. 705–871, 2022. ISSN 0169-2070. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169207021001758>>.
- PRABHAKARAN, S. Arima model – complete guide to time series forecasting in python. **Machine Learning Plus**, 2018.
- PREDUM, R. **Time Series Forecasting with ARIMA, SARIMA, and SARIMAX**. 2021. Disponível em: <<https://towardsdatascience.com/time-series-forecasting-with-arima-sarima-and-sarimax-ee61099e78f6>>.
- QIN, Y. et al. Spatio-temporal hierarchical mlp network for traffic forecasting. **Information Sciences**, v. 632, p. 543–554, 2023.
- RAMOS, A. S. **Previsões de Séries Temporais combinando modelos ARMA e Redes Neurais Artificiais**. Tese (Doutorado) — Universidade Federal de Pernambuco, 2010.
- RAO, X.; ZHAO, H.; DENG, Q. Artificial-neural-network (ann) based proxy model for performances forecast and inverse project design of water huff-n-puff technology. **Journal of Petroleum Science and Engineering**, v. 195, p. 107851, 2020.
- RASCHKA, S. A practical guide to machine learning in python. **Machine Learning with Python**, 2015. Disponível em: <https://sebastianraschka.com/pdf/books/machine_learning_with_python/mlwp.pdf>.
- READER, T. C. Decision tree regression explained with implementation in python. **Medium**, 2023. Disponível em: <<https://medium.com/@theclickreader/decision-tree-regression-explained-with-implementation-in-python-1e6e48aa7a47>>.
- REICHMAN, D.; MALOF, J. M.; COLLINS, L. M. Leveraging seed dictionaries to improve dictionary learning. In: **2016 IEEE International Conference on Image Processing (ICIP)**. [S.l.: s.n.], 2016. p. 3723–3727.

REISEN, V. et al. Robust dickey–fuller tests based on ranks for time series with additive outliers. **Metrika**, v. 80, n. 1, p. 115–131, 2017. Cited By 1. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84986317325&doi=10.1007%2fs00184-016-0594-8&partnerID=40&md5=c83f82d0c372e22d5970aff448f05411>>.

RIBEIRO, M. H. D. M. et al. Cooperative ensemble learning model improves electric short-term load forecasting. **Chaos, Solitons & Fractals**, v. 166, p. 112982, 2023.

RIBEIRO, M. H. D. M. et al. Time series forecasting based on ensemble learning methods applied to agribusiness, epidemiology, energy demand, and renewable energy. Pontificia Universidade Católica do Paraná, 2021.

ROSTAM, N. A. P. et al. A complete proposed framework for coastal water quality monitoring system with algae predictive model. **IEEE Access**, Institute of Electrical and Electronics Engineers Inc., v. 9, p. 108249 – 108265, 2021. ISSN 21693536. Cited by: 12; All Open Access, Gold Open Access. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85112659996&doi=10.1109%2fACCESS.2021.3102044&partnerID=40&md5=a078d7fe0d04f37177flae6f798ff71b>>.

ROSTAMIAN, A.; O'HARA, J. G. Event prediction within directional change framework using a cnn-lstm model. **NEURAL COMPUTING & APPLICATIONS**, v. 34, p. 17193–17205, 2022. ISSN 0941-0643.

RUIZ-ROSETO, J.; RAMIREZ-GONZALEZ, G.; VIVEROS-DELGADO, J. Software survey: Scientopy, a scientometric tool for topics trend analysis in scientific publications. **Scientometrics**, v. 121, n. 2, p. 1165–1188, Nov 2019. ISSN 1588-2861. Disponível em: <<https://doi.org/10.1007/s11192-019-03213-w>>.

SABZIPOUR, B. et al. Comparing a long short-term memory (lstm) neural network with a physically-based hydrological model for streamflow forecasting over a canadian catchment. **Journal of Hydrology**, v. 627, p. 130380, 2023.

SADAEI, H. et al. Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series. **Energy**, v. 175, p. 365–377, 2019. ISSN 03605442.

SADAEI, H. et al. Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series. **Energy**, v. 175, p. 365–377, 2019.

SANG, Y.-F. et al. Wavelet-based hydrological time series forecasting. **Journal of Hydrologic Engineering**, v. 21, 2016. ISSN 10840699.

SARANYA, S.; SIVAKUMAR, R. Gated recurrent units (gru) for time series forecasting in higher education. **International Journal of Engineering Research and Technology**, v. 13, n. 7, p. 1809–1813, 2020. Disponível em: <<https://www.ijert.org/gated-recurrent-units-gru-for-time-series-forecasting-in-higher-education>>.

SEMAN, L. O. et al. Ensemble learning methods using the hodrick–prescott filter for fault forecasting in insulators of the electrical power grids. **International Journal of Electrical Power & Energy Systems**, v. 152, p. 109269, 2023.

SEN, J. et al. Machine learning: Algorithms, models, and applications. **arXiv preprint arXiv:2201.01943**, 2022.

- SEZER, O. B.; GUDELEK, M. U.; OZBAYOGLU, A. M. Financial time series forecasting with deep learning : A systematic literature review: 2005-2019. **APPLIED SOFT COMPUTING**, v. 90, 2020. ISSN 1568-4946.
- SHEN, L.; WANG, Y. Tcct: Tightly-coupled convolutional transformer on time series forecasting. **Neurocomputing**, v. 480, p. 131–145, 2022. ISSN 09252312.
- SHI, M. et al. Ensemble regression based on polynomial regression-based decision tree and its application in the in-situ data of tunnel boring machine. **Mechanical Systems and Signal Processing**, v. 188, p. 110022, 2023.
- SHIH, S.-Y.; SUN, F.-K.; LEE, H.-Y. Temporal pattern attention for multivariate time series forecasting. **Machine Learning**, v. 108, n. 8-9, p. 1421–1441, 2019.
- SHIH, S.-Y.; SUN, F.-K.; LEE, H.-Y. Temporal pattern attention for multivariate time series forecasting. **Machine Learning**, v. 108, p. 1421–1441, 2019. ISSN 08856125.
- SIEGEL, J. E. et al. Safe energy savings through context-aware hot water demand prediction. **Engineering Applications of Artificial Intelligence**, v. 90, p. 103481, 2020.
- SILVA, A. C.; GOMES, L. F. A. M. Inteligência artificial: estado atual, desafios e oportunidades de pesquisa. **Estudos Avançados**, v. 35, n. 101, p. 7–26, 2021. Disponível em: <<https://www.scielo.br/j/ea/a/wXBdv8yHBV9xHz8qG5RCgZd/>>.
- SILVA, J. P. **Como funcionam as Redes Neurais Convolucionais (CNNs)**. 2021. <<https://medium.com/data-hackers/como-funcionam-as-redes-neurais-convolucionais-cnns-71978185c1>>. Acessado em 04/05/2023.
- Singh Kushwah, J. et al. Comparative study of regressor and classifier with decision tree using modern tools. **Materials Today: Proceedings**, v. 56, p. 3571–3576, 2022. First International Conference on Design and Materials.
- STEFENON, S. F. et al. Aggregating prophet and seasonal trend decomposition for time series forecasting of italian electricity spot prices. **Energies**, v. 16, n. 3, 2023.
- SUN, X.; LIU, M.; SIMA, Z. A novel cryptocurrency price trend forecasting model based on lightgbm. **Finance Research Letters**, v. 32, p. 101084, 2020.
- TAM, A. **LSTM for Time Series Prediction in PyTorch**. Machine Learning Mastery, 2023. Disponível em: <<https://machinelearningmastery.com/lstm-for-time-series-prediction-in-pytorch/>>.
- TAO, H. et al. Training and testing data division influence on hybrid machine learning model process: Application of river flow forecasting. **Complexity**, Hindawi, Oct 2020.
- THEODOSIOU, M. Forecasting monthly and quarterly time series using stl decomposition. **International Journal of Forecasting**, v. 27, n. 4, p. 1178–1195, 2011. Cited By 86. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-80052160927&doi=10.1016%2fj.ijforecast.2010.11.002&partnerID=40&md5=e8242471ba1ec14ada46ab567f3a364d>>.

- TRENBERTH, K. E. Signal versus noise in the southern oscillation. **Monthly Weather Review**, v. 112, n. 2, p. 326–332, 1984.
- TSATSOU, A.; FRANTZESKAKI, N.; MALAMIS, S. Nature-based solutions for circular urban water systems: A scoping literature review and a proposal for urban design and planning. **Journal of Cleaner Production**, v. 394, p. 136325, 2023.
- TYRALIS, H.; PAPACHARALAMPOUS, G. Variable selection in time series forecasting using random forests. **Algorithms**, v. 10, n. 4, 2017.
- TYRALIS, H.; PAPACHARALAMPOUS, G. Variable selection in time series forecasting using random forests. **Algorithms**, v. 10, 2017. ISSN 19994893.
- UC-CASTILLO, J. L. et al. A systematic review and meta-analysis of groundwater level forecasting with machine learning techniques: Current status and future directions. **Environmental Modelling & Software**, v. 168, p. 105788, 2023.
- URSU, E.; PEREAU, J. C. Application of periodic autoregressive process to the modeling of the Garonne river flows. **STOCHASTIC ENVIRONMENTAL RESEARCH AND RISK ASSESSMENT**, v. 30, n. 7, p. 1785–1795, 2016. ISSN 1436-3240.
- VASCONCELOS, F. **Falta d'água em Curitiba e região metropolitana não É culpa só da estiagem**. 2020. Disponível em: <<https://www.brasildefato.com.br/2020/11/03/falta-d-agua-em-curitiba-e-regiao-metropolitana-nao-e-culpa-so-da-estiagem>>.
- VIDHYA, A. **Time Series Forecasting and Analysis — ARIMA and Seasonal ARIMA**. Medium, 2023. Disponível em: <<https://medium.com/analytics-vidhya/time-series-forecasting-and-analysis-arima-and-seasonal-arima-cacaf61ae863>>.
- WANG, J. et al. Financial time series prediction using elman recurrent random neural networks. **Computational Intelligence and Neuroscience**, v. 2016, 2016. ISSN 16875265.
- WANG, L. et al. Hybrid application of unsupervised and supervised learning in forecasting absolute open flow potential for shale gas reservoirs. **Energy**, v. 243, p. 122747, 2022.
- WANG, M.; YING, F. Point and interval prediction for significant wave height based on lstm-gru and kde. **Ocean Engineering**, v. 289, p. 116247, 2023.
- XIAN, S. et al. A novel fuzzy time series forecasting method based on the improved artificial fish swarm optimization algorithm. **Soft Computing**, v. 22, p. 3907–3917, 2018. ISSN 14327643.
- XIANG, Y. et al. A svr–ann combined model based on ensemble emd for rainfall prediction. **Applied Soft Computing**, v. 73, p. 874–883, 2018.
- XU, W. et al. A hybrid modelling method for time series forecasting based on a linear regression model and deep learning. **Applied Intelligence**, v. 49, p. 3002–3015, 2019. ISSN 0924669X.
- YANG, S.; GUO, H.; LI, J. Cnn-grua-fc stock price forecast model based on multi-factor analysis. **Journal of Advanced Computational Intelligence and Intelligent Informatics**, v. 26, p. 600–608, 2022. ISSN 13430130.

YE, J.; ZHAO, B.; DENG, H. Photovoltaic power prediction model using pre-train and fine-tune paradigm based on lightgbm and xgboost. **Procedia Computer Science**, v. 224, p. 407–412, 2023.

ZHANG, E. **Recurrent Neural Network is All You Need**. 2021. <<https://medium.com/mcgill-mma-intro-to-ai/recurrent-neural-network-is-all-you-need-f576782c5d2>>. Acessado em: 22 de Março de 2023.

ZHAO, L. et al. A hybrid vmd-lstm/gru model to predict non-stationary and irregular waves on the east coast of china. **Ocean Engineering**, v. 276, p. 114136, 2023.