



**CIÊNCIAS  
EMPRESARIAIS**

ESCOLA SUPERIOR  
POLITÉCNICO SETÚBAL

AMÉRICO JOSÉ  
MENDONÇA  
PESTANA

**MODELOS DE PREVISÃO  
BASEADOS EM SÉRIES  
TEMPORAIS: UMA ABORDAGEM  
COMPARATIVA APLICADA À  
GESTÃO OPERACIONAL**

Relatório de Trabalho de Projeto do Mestrado em  
Ciência de Dados para Empresas

**ORIENTADORA**

Professora Doutora, Sandra Cristina Dias Nunes

**CO-ORIENTADORA**

Professora Doutora, Helena Alexandra Couceiro

Dezembro / 2025

AMÉRICO JOSÉ  
MENDONÇA  
PESTANA

**MODELOS DE PREVISÃO  
BASEADOS EM SÉRIES  
TEMPORAIS: UMA ABORDAGEM  
COMPARATIVA APLICADA À  
GESTÃO OPERACIONAL**

**JÚRI**

***Presidente:***

Professora Doutora, Ana de Jesus Pereira Barreira  
Mendes, Escola Superior de Ciências Empresariais

***Orientador:***

Professora Doutora, Sandra Cristina Dias Nunes,  
Escola Superior de Ciências

***Vogal Arguente:***

Professora Doutora, Sandra Maria Simões de Oliveira,  
Escola Superior de Ciências

## **Dedicatória**

Aos meus pais, cujas vidas foram um exemplo de amor, resiliência e dedicação.

Embora já não estejam ao meu lado, a vossa presença continua a guiar-me em cada passo que dou. Este trabalho é o reflexo dos valores que me ensinaram, da força que me transmitiram e do sonho que sempre partilharam comigo.

Dedico-vos este mestrado, não apenas como uma conquista académica, mas como uma homenagem ao vosso legado.

Obrigado por serem a minha inspiração, mesmo à distância.

Com todo o meu amor e gratidão,

## **Agradecimentos**

Aqui gostaria de deixar o meu agradecimento a todos os que tornaram a realização deste trabalho possível.

Em primeiro lugar, manifesto a minha imensa gratidão à minha orientadora, Professora Doutora Sandra Cristina Dias Nunes, e à minha co-orientadora, Professora Doutora Helena Alexandra Couceiro Feio de Almeida Penalva, por todo o apoio prestado ao longo destes meses, desde o primeiro momento, com a minha indecisão em relação à escolha do tema do meu projeto, até à sua finalização. Agradeço todas as horas que dedicaram para que eu conseguisse concretizar este trabalho de investigação, toda a troca de pensamentos e ideias estabelecidos.

Também quero prestar o meu agradecimento a todos os envolvidos neste projeto, pela disponibilidade que tiveram ao contribuir para este trabalho, sem eles não seria possível a sua realização. Quero destacar o CEO da CIBEN – IT and Business Solutions SA, Eng. João Manuel Garcia Luís, por me ter identificado os vários temas dentro das necessidades atuais ao nível da gestão das equipas na CIBEN, pelo que fica registado o meu sincero obrigado.

Quero fazer um agradecimento especial a minha companheira Maria João, por ouvir todos os meus desabaços, por me compreender nos dias de maior stress, por me acompanhar durante todo este percurso e por toda a motivação. Agradeço também aos meus filhos, Beatriz e Afonso, pela paciência e porque nunca duvidaram que conseguia alcançar as metas estipuladas. A todos eles agradeço sobretudo por fazerem parte da minha vida e por todo o apoio e compreensão demonstrada.

Quero agradecer aos meus familiares e amigos, por todos os momentos de descontração que me permitiram ter um refúgio da rotina, pelo apoio, força e carinho que me deram.

A todos um enorme obrigado!

## Resumo

Este projeto, desenvolvido no âmbito do Mestrado em Ciência de Dados para Empresas, teve como objetivo principal a realização de uma análise comparativa entre diferentes modelos de previsão de séries temporais, aplicados a dados históricos do sistema ATX (2021–2024). Foram avaliados cinco modelos: Seasonal Autoregressive Integrated Moving Average (SARIMA), Holt-Winters (método de suavização exponencial), Neural Basis Expansion Analysis for Time Series (N-BEATS), Deep Autoregressive Model (DeepAR) e Prophet.

A metodologia adotada seguiu o modelo Cross Industry Standard Process for Data Mining (CRISP-DM), abrangendo desde a compreensão do negócio até à avaliação e implementação dos modelos. O processo incluiu a preparação rigorosa dos dados, a aplicação de técnicas estatísticas e de aprendizagem profunda, e a validação através de métricas de erro como Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) e Mean Absolute Percentage Error (MAPE), complementadas pelo teste estatístico de Diebold-Mariano.

Os resultados indicaram que o modelo Prophet apresentou o melhor desempenho global, destacando-se pela sua robustez, consistência e adequação em contextos empresariais.

Este estudo evidencia a importância da previsão precisa para a eficiência operacional e o planeamento de recursos, propondo a integração dos modelos mais eficazes em sistemas internos, com recurso a dashboards interativos e Application Programming Interface (API). Recomenda-se, ainda, a continuação da investigação com modelos híbridos e a inclusão de variáveis exógenas, de forma a melhorar a precisão das previsões em cenários futuros.

## **Abstract**

This project, developed as part of the Master's in Data Science for Business, aimed to conduct a comparative analysis of different time series forecasting models applied to historical data from the ATX system (2021–2024). Five models were evaluated: SARIMA (Seasonal Autoregressive Integrated Moving Average), Holt-Winters (exponential smoothing method), N-BEATS (Neural Basis Expansion Analysis for Time Series), DeepAR (Deep Autoregressive Model), and Prophet.

The methodology adopted followed the Cross Industry Standard Process for Data Mining (CRISP-DM) framework, encompassing everything from business understanding to model evaluation and implementation. The process included rigorous data preparation, the application of statistical and deep learning techniques, and validation using error metrics such as MAE (Mean Absolute Error), MSE (Mean Squared Error), RMSE (Root Mean Squared Error), and MAPE (Mean Absolute Percentage Error), complemented by the Diebold-Mariano statistical test.

The results indicated that the Prophet model performed best overall, standing out for its robustness, consistency, and suitability in business contexts.

This study highlights the importance of accurate forecasting for operational efficiency and resource planning, proposing the integration of the most effective models into internal systems, using interactive dashboards and APIs. We also recommend further research with hybrid models and the inclusion of exogenous variables to improve forecast accuracy in future scenarios.

## **Palavras-chave**

SARIMA

Holt-Winters

N-Beats

DeepAR

Prophet

## Índice

|                                                  |      |
|--------------------------------------------------|------|
| Dedicatória .....                                | III  |
| Agradecimentos .....                             | IV   |
| Resumo .....                                     | V    |
| Abstract .....                                   | VI   |
| Palavras-chave .....                             | VII  |
| Índices de Tabelas .....                         | X    |
| Índices de Tabelas do Apêndice A .....           | XI   |
| Índices de Figuras .....                         | XIII |
| Índices de Figuras do Apêndice B .....           | XIV  |
| Lista de Acrónimos e Abreviaturas .....          | XVII |
| Lista de Símbolos .....                          | XX   |
| Introdução .....                                 | 1    |
| Motivação e Enquadramento .....                  | 1    |
| Motivação .....                                  | 1    |
| Enquadramento .....                              | 2    |
| Objetivo geral .....                             | 4    |
| Objetivos Específicos .....                      | 5    |
| Tipo de abordagem .....                          | 5    |
| Metodologia .....                                | 6    |
| Estrutura do Documento .....                     | 7    |
| 1. Revisão da Literatura .....                   | 9    |
| 1.1. Estado da Arte .....                        | 9    |
| 2. Objetivos e Metodologia .....                 | 17   |
| 2.1. Objetivos .....                             | 17   |
| 2.2. Metodologia .....                           | 17   |
| 3. Apresentação e discussão dos resultados ..... | 20   |
| 3.1. Implementação dos Modelos .....             | 20   |

|          |                                                                                                        |    |
|----------|--------------------------------------------------------------------------------------------------------|----|
| 3.1.1.   | Configuração do Ambiente .....                                                                         | 20 |
| 3.1.2.   | Implementação dos Modelos .....                                                                        | 21 |
| 3.2.     | Descrição dos Dados Utilizados .....                                                                   | 21 |
| 3.3.     | Análise Estatística e Resultados .....                                                                 | 23 |
| 3.3.1.   | Análise do resultado de cada modelo.....                                                               | 24 |
| 3.3.1.1. | SARIMA - Comparação entre os Modelos Geral e Simplificado .....                                        | 25 |
| 3.3.1.2. | Holt-Winters - Comparação entre os Modelos HWA e HWM.....                                              | 31 |
| 3.3.1.3. | N-BEATS - Comparação entre os Modelos Simples e Especializado.....                                     | 33 |
| 3.3.1.4. | DeepAR - Comparação entre os Modelos Padrão e Otimizado.....                                           | 37 |
| 3.3.1.5. | PROPHET - Comparação entre do Modelo Padrão e Otimizado.....                                           | 40 |
| 3.3.2.   | Análise comparativa dos resultados dos modelos selecionados .....                                      | 44 |
| 3.3.2.1. | Análise comparativa dos resultados dos modelos selecionados .....                                      | 44 |
| 3.3.3.   | Análise comparativa dos resultados dos modelos selecionados com base nas previsões out-of-sample ..... | 47 |
| 3.4.     | Discussão de Resultados.....                                                                           | 51 |
|          | Conclusão e Investigação Futura .....                                                                  | 54 |
|          | Referências Bibliográficas .....                                                                       | 56 |
|          | Apêndices A.....                                                                                       | 59 |
|          | Apêndices B .....                                                                                      | 72 |

## Índices de Tabelas

|                                                                                                                   |    |
|-------------------------------------------------------------------------------------------------------------------|----|
| Tabela 1 - Quadro comparativo dos modelos SARIMA.....                                                             | 26 |
| Tabela 2 - Quadro resumo com referência às Tabelas e Figuras .....                                                | 28 |
| Tabela 3 – Quadro comparativo dos modelos Holt-Winters.....                                                       | 31 |
| Tabela 4 – Quadro resumo com referência às Tabelas e Figuras .....                                                | 32 |
| Tabela 5 - Quadro comparativo dos modelos N-BEATS.....                                                            | 34 |
| Tabela 6 - Quadro resumo com referência às Tabelas e Figuras .....                                                | 35 |
| Tabela 7 - Quadro comparativo dos modelos DeepAR.....                                                             | 38 |
| Tabela 8 - Quadro resumo com referência às Tabelas e Figuras .....                                                | 38 |
| Tabela 9 - Quadro comparativo dos modelos Prophet.....                                                            | 41 |
| Tabela 10 - Quadro resumo com referência às Tabelas e Figuras .....                                               | 42 |
| Tabela 11 - Quadro comparativo dos modelos analisados .....                                                       | 44 |
| Tabela 12 – Quadro com os dados reais e as previsões para o ano 2024 de cada versão selecionada dos modelos ..... | 47 |
| Tabela 13 - Métricas de erro por modelo com base nas previsões do ano de 2024 .....                               | 48 |
| Tabela 14 - Teste Diebold-Mariano para comparações em pares com base no MSE.....                                  | 48 |
| Tabela 15 – Ranking dos modelos com base no MSE e no MAPE.....                                                    | 49 |

## Índices de Tabelas do Apêndice A

|                                                                                                                                  |    |
|----------------------------------------------------------------------------------------------------------------------------------|----|
| Tabela A. 1 – Dados 2012-2024 resumidos.....                                                                                     | 59 |
| Tabela A. 2 – Dados 2012-2024 agrupados por Ano-Mês.....                                                                         | 59 |
| Tabela A. 3 - Resumo dos resultados do teste ADF .....                                                                           | 60 |
| Tabela A. 4 - Resumo dos resultados do modelo SARIMA Geral .....                                                                 | 60 |
| Tabela A. 5 – Valores das métricas de validação Walk-Foward do modelo SARIMA Geral.                                              | 60 |
| Tabela A. 6 - Resumo dos resultados do modelo SARIMA Simplificado.....                                                           | 61 |
| Tabela A. 7 – Valores das métricas de validação Walk-Foward do modelo SARIMA Simplificado.....                                   | 61 |
| Tabela A. 8 - Comparação dos resultados do Teste de Ljung-Box entre as duas versões do modelo SARIMA.....                        | 61 |
| Tabela A. 9 – Resumo dos resultados do modelo HWA .....                                                                          | 62 |
| Tabela A. 10 – Valores das métricas de validação Walk-Forward do modelo HWA.....                                                 | 62 |
| Tabela A. 11 - Resumo dos resultados do modelo HWM .....                                                                         | 63 |
| Tabela A. 12 – Valores das métricas de validação Walk-Foward do modelo HWM .....                                                 | 63 |
| Tabela A. 13 - Comparação da informação geral entre o modelo HWA e HWM .....                                                     | 64 |
| Tabela A. 14 - Comparação dos resultados do Teste de Ljung-Box e do Teste Jarque-Bera entre os modelos HWA e HWM.....            | 64 |
| Tabela A. 15 - Resumo das métricas de informação do modelo N-BEATS Simples .....                                                 | 65 |
| Tabela A. 16 - Resumo das métricas de Erro do modelo N-BEATS Simples.....                                                        | 65 |
| Tabela A. 17 - Resumo das métricas de informação do modelo N-BEATS Especializado...                                              | 65 |
| Tabela A. 18 - Resumo das Métricas de Erro do modelo N-BEATS Especializado.....                                                  | 66 |
| Tabela A. 19 - Comparação dos resultados do Teste Ljung-Box e do Teste Jarque-Bera entre as duas versões do modelo N-BEATS ..... | 66 |
| Tabela A. 20 - Resumo das métricas de avaliação do modelo DeepAR Padrão.....                                                     | 67 |
| Tabela A. 21 - Resumo das métricas de Quantis do modelo DeepAR Padrão .....                                                      | 67 |
| Tabela A. 22 - Resumo das métricas de informação do modelo DeepAR Padrão .....                                                   | 67 |
| Tabela A. 23 - Resumo das métricas de avaliação do modelo DeepAR Otimizado .....                                                 | 68 |
| Tabela A. 24 - Resumo de métricas de Quantis do modelo DeepAR Otimizado.....                                                     | 68 |
| Tabela A. 25 - Resumo das métricas de informação do modelo DeepAR Otimizado.....                                                 | 68 |
| Tabela A. 26 – Comparação dos resultados do Teste Ljung-Box e do Teste Jarque-Bera entre as duas versões do modelo DeepAR .....  | 69 |
| Tabela A. 27 - Resumo das métricas de Erro do modelo Prophet Padrão .....                                                        | 69 |
| Tabela A. 28 - Resumo das métricas de informação do modelo Prophet Padrão .....                                                  | 70 |
| Tabela A. 29 - Resumo das métricas de Erro do modelo Prophet Otimizado .....                                                     | 70 |

|                                                                                                                                  |    |
|----------------------------------------------------------------------------------------------------------------------------------|----|
| Tabela A. 30 - Resumo das métricas de informação do modelo Prophet Otimizado.....                                                | 71 |
| Tabela A. 31 – Comparação dos resultados do Teste Ljung-Box e do Teste Jarque-Bera entre as duas versões do modelo Prophet ..... | 71 |

## Índices de Figuras

|                                                                                                      |    |
|------------------------------------------------------------------------------------------------------|----|
| Figura 1 – Fases da Metodologia CRISP-DM.....                                                        | 17 |
| Figura 2 – Cronograma dos dados reais dos dados histórico a 12 anos (2012 a 2024).....               | 23 |
| Figura 3 – Cronograma da série temporal com os dados originais dos últimos 3 anos (2022 a 2024)..... | 24 |
| Figura 4 – Gráficos comparativos dos modelos .....                                                   | 49 |

## Índices de Figuras do Apêndice B

|                                                                                                         |    |
|---------------------------------------------------------------------------------------------------------|----|
| Figura B. 1 – Cronograma dos dados reais dos dados histórico a 12 anos (2012 a 2024) ..                 | 72 |
| Figura B. 2 – Cronograma da série temporal com os dados originais dos últimos 3 anos (2021 a 2024)..... | 72 |
| Figura B. 3 – Gráfico de resíduos do modelo SARIMA Geral .....                                          | 72 |
| Figura B. 4 - Distribuição dos resíduos do modelo SARIMA Geral .....                                    | 73 |
| Figura B. 5 - Q-Q plot do modelo SARIMA Geral.....                                                      | 73 |
| Figura B. 6 - ACF dos resíduos do modelo SARIMA Geral.....                                              | 73 |
| Figura B. 7 – Cronograma dos dados reais e das previsões do modelo SARIMA Geral .....                   | 74 |
| Figura B. 8 – Gráfico de resíduos do modelo SARIMA Simplificado .....                                   | 74 |
| Figura B. 9 - Distribuição dos resíduos do modelo SARIMA Simplificado .....                             | 74 |
| Figura B. 10 - Q-Q plot do modelo SARIMA Simplificado .....                                             | 75 |
| Figura B. 11 - ACF dos resíduos do modelo SARIMA Simplificado .....                                     | 75 |
| Figura B. 12 – Cronograma dos dados reais e das previsões com o modelo SARIMA Simplificado.....         | 75 |
| Figura B. 13 – Gráfico de resíduos do modelo HWA .....                                                  | 76 |
| Figura B. 14 - Distribuição dos resíduos do modelo HWA .....                                            | 76 |
| Figura B. 15 - Probabilidade normal (Q-Q plot) do modelo HWA.....                                       | 76 |
| Figura B. 16 - Correlograma dos resíduos do modelo HWA .....                                            | 77 |
| Figura B. 17 – Validação Walk-Forward do modelo HWA.....                                                | 77 |
| Figura B. 18 – Cronograma da série temporal com os ajustes e a previsão do modelo HWA .....             | 77 |
| Figura B. 19 – Gráfico de resíduos do modelo HWM.....                                                   | 78 |
| Figura B. 20 - Distribuição dos resíduos do modelo HWM.....                                             | 78 |
| Figura B. 21 - Probabilidade normal (Q-Q plot) do modelo HWM .....                                      | 78 |
| Figura B. 22 - Correlograma dos resíduos do modelo HWM .....                                            | 79 |
| Figura B. 23 – Validação Walk-Forward do modelo HWM .....                                               | 79 |
| Figura B. 24 – Cronograma da série temporal com os ajustes e a previsão do modelo HWM .....             | 79 |
| Figura B. 25 – Modelo N-BEATS Simples .....                                                             | 80 |
| Figura B. 26 – Gráfico dos resíduos do modelo N-BEATS Simples.....                                      | 80 |
| Figura B. 27 - Distribuição dos resíduos do modelo N-BEATS Simples.....                                 | 80 |
| Figura B. 28 - Probabilidade normal (Q-Q plot) do modelo N-BEATS Simples .....                          | 81 |
| Figura B. 29 - Correlograma dos resíduos do modelo N-BEATS Simples .....                                | 81 |
| Figura B. 30 – Validação Walk-Forward do modelo N-BEATS Simples .....                                   | 81 |

|                                                                                                               |    |
|---------------------------------------------------------------------------------------------------------------|----|
| Figura B. 31 – Cronograma da série temporal com os ajustes e a previsão do modelo N-BEATS Simples.....        | 82 |
| Figura B. 32 – Modelo N-BEATS Especializado .....                                                             | 82 |
| Figura B. 33 – Gráfico de resíduos do modelo N-BEATS Especializado .....                                      | 82 |
| Figura B. 34 - Distribuição dos resíduos do modelo N-BEATS Especializado .....                                | 83 |
| Figura B. 35 - Probabilidade normal (Q-Q plot) do modelo N-BEATS Especializado .....                          | 83 |
| Figura B. 36 - Correlograma dos resíduos do modelo N-BEATS Especializado .....                                | 83 |
| Figura B. 37 – Validação Walk-Forward do modelo N-BEATS Especializado.....                                    | 84 |
| Figura B. 38 – Cronograma da série temporal com os ajustes e a previsão do modelo N-BEATS Especializado ..... | 84 |
| Figura B. 39 – Modelo DeepAR Padrão .....                                                                     | 84 |
| Figura B. 40 – Gráfico de resíduos do modelo DeepAR Padrão.....                                               | 85 |
| Figura B. 41 - Distribuição dos resíduos do modelo DeepAR Padrão.....                                         | 85 |
| Figura B. 42 - Probabilidade normal (Q-Q plot) do modelo DeepAR Padrão .....                                  | 85 |
| Figura B. 43 - Correlograma dos resíduos do modelo DeepAR Padrão.....                                         | 86 |
| Figura B. 44 - Modelo DeepAR Otimizado.....                                                                   | 86 |
| Figura B. 45 – Gráfico de resíduos do modelo DeepAR Otimizado .....                                           | 87 |
| Figura B. 46- Distribuição dos resíduos do modelo DeepAR Otimizado .....                                      | 87 |
| Figura B. 47 - Probabilidade normal (Q-Q plot) do modelo DeepAR Otimizado .....                               | 87 |
| Figura B. 48 - Correlograma dos resíduos do modelo DeepAR Otimizado .....                                     | 88 |
| Figura B. 49 - Cronograma da série temporal com as previsões dos modelos DeepAR.....                          | 88 |
| Figura B. 50 – Modelo Prophet Padrão .....                                                                    | 88 |
| Figura B. 51 – Gráfico de resíduos do modelo Prophet Padrão.....                                              | 89 |
| Figura B. 52 - Distribuição dos resíduos do modelo Prophet Padrão.....                                        | 89 |
| Figura B. 53 - Probabilidade normal (Q-Q plot) do modelo Prophet Padrão .....                                 | 89 |
| Figura B. 54 - Correlograma dos resíduos do modelo Prophet Padrão.....                                        | 90 |
| Figura B. 55 – Validação Walk-Forward do modelo Prophet Padrão .....                                          | 90 |
| Figura B. 56 – Cronograma da série temporal com os ajustes e a previsão do modelo Prophet Padrão .....        | 90 |
| Figura B. 57 – Previsão de séries temporais com intervalos de confiança do modelo Prophet Padrão .....        | 91 |
| Figura B. 58 – Modelo Prophet Otimizado.....                                                                  | 91 |
| Figura B. 59 – Gráfico de resíduos do modelo Prophet Otimizado .....                                          | 92 |
| Figura B. 60 - Distribuição dos resíduos do modelo Prophet Otimizado .....                                    | 92 |
| Figura B. 61 - Probabilidade normal (Q-Q plot) do modelo Prophet Otimizado.....                               | 92 |
| Figura B. 62 - Correlograma dos resíduos do modelo Prophet Otimizado .....                                    | 93 |

|                                                                                                                         |    |
|-------------------------------------------------------------------------------------------------------------------------|----|
| Figura B. 63 – Validação Walk-Forward do modelo Prophet Otimizado.....                                                  | 93 |
| Figura B. 64 – Cronograma da série temporal com os ajustes e a previsão do modelo Prophet Otimizado.....                | 93 |
| Figura B. 65 – Cronograma de previsão de séries temporais com intervalos de confiança do modelo Prophet Otimizado ..... | 94 |

## Lista de Acrónimos e Abreviaturas

|          |                                                                                         |
|----------|-----------------------------------------------------------------------------------------|
| ACF      | Função Autocorrelação (Autocorrelation Function)                                        |
| ADF      | Teste estatístico da estacionaridade da série temporal (Augmented Dickey-Fuller)        |
| AI       | Inteligência Artificial (Artificial Intelligence)                                       |
| AIC      | Critérios de Informação de Akaike (Akaike Information Criterion)                        |
| AICc     | Versão corrigida do AIC para amostras pequenas (Corrected Akaike Information Criterion) |
| API      | Application Programming Interface                                                       |
| ARIMA    | Autoregressive Integrated Moving Average                                                |
| ATX      | Solução de Gestão de Serviços Profissionais da CIBEN                                    |
| BD       | Base de Dados                                                                           |
| BCONS    | Departamento Business Consulting                                                        |
| BI       | Business Intelligence                                                                   |
| BIC      | Critérios de Informação Bayesiano (Bayesian Information Criteria)                       |
| BSERV    | Departamento Business Service                                                           |
| CMA      | Média Móvel Centrada (Centered Moving Average)                                          |
| CRM      | Customer Relationship Management                                                        |
| CRISP-DM | Cross-Industry Standard Process for Data Mining                                         |
| CS       | Departamento Administrativo                                                             |
| DEEPAR   | Deep Autoregressive Recurrent Network                                                   |
| EMA      | Média Móvel Exponencial (Exponential Moving Average)                                    |
| ERP      | Enterprise Resource Planning                                                            |
| ESCE     | Escola Superior de Ciências Empresárias                                                 |
| ESTS     | Escola Superior de Tecnologia de Setúbal                                                |
| F        | Estatística do teste F                                                                  |
| GARCH    | Generalized Autoregressive Conditional Heteroskedasticity                               |
| GRU      | Gated Recurrent Unit                                                                    |
| H        | Heterocedasticidade                                                                     |
| HQIC     | Critério de Informação de Hannan-Quinn (Hannan-Quinn Information Criterion)             |
| INFSO    | Departamento Infrastructure Solutions                                                   |
| IPS      | Instituto Politécnico de Setúbal                                                        |
| ISO      | International Organization for Standardization                                          |

|          |                                                                                           |
|----------|-------------------------------------------------------------------------------------------|
| JB       | Teste Jarque-Bera                                                                         |
| KDE      | Estimativa não paramétrica da densidade (Kernel Density Estimation)                       |
| LB       | Teste Ljung-Box                                                                           |
| LM       | Teste de multiplicador de Lagrange (Lagrange Multiplier Test)                             |
| LSTM     | Long Short-Term Memory                                                                    |
| MAE      | Erro Absoluto Médio (Mean Absolute Error)                                                 |
| MAPE     | Erro Percentual Absoluto Médio (Mean Absolute Percentage)                                 |
| MdAE     | Erro Absoluto Mediano (Median Absolute Error)                                             |
| MdAPE    | Erro Percentual Absoluto Mediano (Median Absolute Percentage Error)                       |
| MASE     | Erro Absoluto Escalado Médio (Mean Absolute Scaled Error)                                 |
| MCDE     | Mestrado em Ciências de Dados para Empresas                                               |
| MSIS     | Pontuação Média de Intervalo Escalado (Mean Scaled Interval Score)                        |
| ML       | Machine Learning                                                                          |
| MSE      | Erro Médio Quadrático (Mean Square Error)                                                 |
| N-BEATS  | Neural Basis Expansion Analysis for Time Series                                           |
| ND       | Normalized Deviation                                                                      |
| NRMSE    | Normalized RMSE                                                                           |
| OWA      | Overall Weighted Average                                                                  |
| PACF     | Função Autocorrelação Parcial                                                             |
| PMO      | Project Management Office                                                                 |
| PROPHET  | Modelo de previsão de séries temporais de código aberto desenvolvido pelo Facebook (Meta) |
| Q        | Estatística do Teste de Ljung-Box para Autocorrelação                                     |
| Q-Q Plot | Quantile-Quantile Plot                                                                    |
| RH       | Departamento de Recursos Humanos                                                          |
| RNN      | Recurrent Neural Networks                                                                 |
| RSME     | Raiz do Erro Quadrático Médio (Root Mean Squared Error)                                   |
| RSS      | Soma dos Quadrados dos Resíduos (Residual Sum of Squares)                                 |
| SALES    | Departamento Comercial                                                                    |
| SARIMA   | Seasonal Autoregressive Integrated Moving Average                                         |
| SMA      | Simple Moving Average                                                                     |
| SSE      | Soma dos quadrados dos resíduos (Sum of Squared Errors)                                   |

|       |                               |
|-------|-------------------------------|
| SVR   | Support Vector Regression     |
| SWFAC | Departamento Software Factory |
| WMA   | Weighted Moving Average       |

## Lista de Símbolos

|            |                            |
|------------|----------------------------|
| $h$        | Hora                       |
| $\sigma$   | Desvio Padrão Populacional |
| $\sigma^2$ | Variância Populacional     |
| $\alpha$   | Fator de Suavização        |

## Introdução

Neste capítulo apresenta-se uma breve introdução ao trabalho desenvolvido no âmbito do Trabalho Final do Mestrado em Ciências de Dados para Empresas (MCDE) da Escola Superior de Ciências Empresarias (ESCE) do Instituto Politécnico de Setúbal (IPS).

O tema central deste trabalho é “**Modelos de Previsão Baseados em Séries Temporais: Uma Abordagem Comparativa Aplicada à Gestão Operacional**”, o qual se revela particularmente relevante face aos desafios atuais enfrentados pelas organizações na otimização dos seus processos operacionais.

Este capítulo inicia-se com o enquadramento e a justificação do tema escolhido, sustentados em situações reais que evidenciam as necessidades emergentes às quais o projeto pretende dar resposta. São igualmente definidos os objetivos gerais e específicos, o tipo de abordagem adotada, a metodologia utilizada, bem como a estrutura geral do trabalho.

## Motivação e Enquadramento

### Motivação

Nas áreas de consultoria e serviços técnicos, a capacidade de estimar corretamente **os tempos de execução**, isto é, o intervalo de tempo necessário para completar uma tarefa ou conjunto de atividades desde o seu início até à sua conclusão, é crucial para o cumprimento dos prazos estabelecidos, a alocação eficiente de recursos e, conseqüentemente, para a prevenção de custos operacionais desnecessários.

Contudo, esta previsão revela-se frequentemente um desafio, devido à complexidade dos projetos, à heterogeneidade das equipas envolvidas e à influência de fatores externos imprevisíveis, como alterações nas prioridades dos clientes, constrangimentos logísticos ou limitações técnicas inesperadas.

A principal motivação para a escolha deste tema reside na necessidade de desenvolver **modelos de previsão precisos dos tempos de execução**, os quais são fundamentais para uma definição realista e eficiente dos objetivos operacionais, contribuindo diretamente para a satisfação dos clientes e para a melhoria contínua dos processos internos

Ao analisar dados históricos, pretende-se identificar padrões e tendências que possam ser utilizados para construir modelos robustos e fiáveis de previsão, permitindo uma definição mais rigorosa dos objetivos estratégicos e operacionais para os períodos seguintes.

Este trabalho representa também uma oportunidade para aplicar os conhecimentos adquiridos ao longo do Mestrado em Ciências de Dados para Empresas, nomeadamente nas áreas de **Análise de Dados** e **Machine Learning**, à resolução de um problema real e relevante. A utilização de modelos avançados de previsão, incluindo redes neurais, permitirá explorar o potencial dos dados históricos para antecipar cenários futuros e apoiar decisões estratégicas com maior confiança.

Além do contributo académico, este estudo proporciona uma preparação sólida para enfrentar desafios futuros num contexto profissional cada vez mais orientado pela análise de dados. A capacidade de interpretar e transformar dados em conhecimento útil assume hoje um papel determinante na competitividade das organizações.

Por fim, destaca-se o impacto prático que este trabalho poderá ter numa organização, ao disponibilizar metodologias que promovam previsões mais precisas e uma resposta mais ágil às necessidades dos clientes e às exigências do mercado, reforçando assim a sua capacidade de adaptação e inovação.

## **Enquadramento**

A CIBEN é uma empresa especializada no desenho, implementação e suporte de soluções de software de gestão empresarial (ERP, CRM, BI), sistemas e infraestruturas de comunicações e informática (CIBEN, nd).

A sua oferta de valor está segmentada e endereçada às pequenas, médias e grandes empresas.

Os seus fatores de diferenciação são:

- Abordagem em pré-venda orientada à solução;
- Equipa multidisciplinar orientada ao projeto de forma holística;
- Qualidade do serviço de pós-venda e de resolução de problemas;
- Relação de proximidade e confiança com os vários interlocutores no cliente.

Os valores que orientam a atividade da organização são:

- Atitude;
- Confiança;
- Eficácia;
- Eficiência;
- Ética;
- Inovação.

A empresa dispõe de uma equipa multidisciplinar, orientada para o projeto, que presta serviços de consultoria de negócio, suporte pós-venda, manutenção de soluções, desenvolvimento à medida e integração de soluções, ajustados às necessidades específicas de cada cliente.

Adicionalmente, está capacitada para disponibilizar soluções de infraestrutura, de forma a dar resposta aos requisitos técnicos e operacionais das organizações.

Alguns destes serviços são ainda disponibilizados em regime de outsourcing, permitindo aos clientes a externalização de determinadas funções, com ganhos de eficiência e otimização de recursos.

Internamente, a empresa utiliza a solução de Gestão de Serviços Profissionais OnTime e Online (ATX), desenvolvida pelas suas próprias equipas de consultoria e de desenvolvimento de software. O ATX é uma solução modular, em ambiente WEB desenhada para a gestão operacional de prestação de serviços, que incorpora o conhecimento de negócio acumulado pela empresa ao longo dos anos.

Como benefícios, esta solução permite:

- Planear eficazmente os recursos da equipa e garantir o cumprimento do nível de serviço acordado;
- Controlar OnTime as tarefas e os tempos de resposta e execução;
- Acelerar e simplificar os processos de imputação e faturação;
- Automatizar a renovação e faturação dos contratos;
- Analisar a rentabilidade dos serviços executados e dos contratos em vigor;
- Acompanhar eficientemente as oportunidades, as propostas e os ciclos de venda;
- Aumentar os níveis de satisfação e fidelização dos clientes.

O ATX regista toda a informação de negócio, mantendo a sequência do relacionamento com o cliente, numa lógica de pré-venda, venda, projeto e pós-venda.

Todas as tarefas atribuídas aos consultores, em qualquer fase do negócio, estão tipificadas e são registadas com os respetivos tempos de execução previstos e reais.

A informação gerada internamente por esta solução é atualmente utilizada em análises de dados, com o objetivo de construir uma visão 360° do cliente, abrangendo o planeamento de projetos e de recursos (consultores e técnicos), bem como o acompanhamento da execução dos projetos e dos serviços de manutenção em pós-venda.

A eficiência operacional é um dos pilares fundamentais para o sucesso de qualquer organização. Neste contexto, a previsão precisa do número total de horas executáveis por parte das equipas associadas aos vários departamentos torna-se crucial. Esta previsão permite um planeamento eficaz, uma distribuição equitativa de tarefas e uma gestão otimizada dos recursos.

O foco deste projeto centra-se na seleção de modelos de previsão adequados, com vista ao desenvolvimento de uma ferramenta de previsão, permitindo assim suportar de forma mais rigorosa o processo de planeamento e gestão de recursos.

## **Objetivo geral**

O objetivo geral deste projeto consiste em identificar o modelo de previsão mais adequado para estimar as horas de execução das tarefas dos consultores e técnicos dos departamentos Business Consulting (BCONS), Business Service (BSERV), Infraestrutura de Redes e Comunicações (INFSO) e Desenvolvimento (SWFAC), tendo por base as horas de execução registadas na plataforma ATX ao longo dos últimos anos.

Neste trabalho será realizada a comparação de cinco modelos de previsão de séries temporais com sazonalidade: SARIMA, Holt-Winters, N-Beats, DeepAR e Prophet, sendo o respetivo desempenho avaliado através da aplicação das seguintes métricas e técnicas:

- Erros de Previsão;
- Análise de Resíduos;
- Validação Cruzada Temporal;

- Análise de Sensibilidade;
- Avaliação de Robustez;
- Análise de Cenários;
- Utilidade Prática.

### **Objetivos Específicos**

Os objetivos específicos, identificados abaixo, são propostos para que o objetivo geral possa ser alcançado.

Objetivos específicos:

1. Obter os dados históricos de registos de tempos de execução, procedendo ao seu tratamento e adequação para análise.
2. Realizar uma revisão teórica dos modelos de previsão supracitados, analisando as suas características e aplicações.
3. Implementar os modelos de previsão selecionados na linguagem Python, utilizando o ambiente de desenvolvimento Visual Studio Code Notebook.
4. Validar as previsões geradas por cada modelo, recorrendo à análise gráfica e à aplicação de diversas métricas de avaliação, nomeadamente: Erros de Previsão; Análise de Resíduos; Validação Cruzada Temporal; Análise de Sensibilidade; Avaliação de Robustez; Análise de Cenários e Utilidade Prática.
5. Comparar os vários modelos identificando o mais adequado, utilizando as métricas e técnicas referidas no ponto anterior.

### **Tipo de abordagem**

Neste projeto adotou-se uma abordagem de estudo de caso, com foco na análise exploratória e preditiva de séries temporais, recorrendo a técnicas de aprendizagem automática (*machine learning*). O estudo de caso permite realizar uma investigação aprofundada de um fenómeno específico no seu contexto real, através da combinação de métodos quantitativos e qualitativos com o objetivo de compreender padrões complexos e gerar insights aplicáveis (Yin, 2018).

As principais características desta abordagem incluem:

- o foco contextual, através da análise detalhada de um caso concreto (neste caso, a previsão das horas de execução realizadas);
- a utilização de dados multifacetados, com a integração de dados quantitativos (séries temporais) e informações qualitativas (variáveis como Grupo de Atividade, Atividade, Dep (Departamento) e Tec (Técnicos));
- a triangulação de métodos, através da combinação de técnicas estatísticas (modelos SARIMA e Holt-Winters) com algoritmos de machine learning (redes LSTM), permitindo comparar e validar a consistência dos resultados obtidos;
- e a profundidade analítica, não apenas na identificação de padrões numéricos, mas também na exploração dos fatores contextuais que influenciam os dados (Stake, 2006).

Neste estudo, o foco recai sobre dados reais ("Tempos de Execução Realizados" entre 2012-2024), cuja contextualização foi efetuada através da identificação de variáveis-chave. A análise quantitativa foi conduzida mediante a modelação da série temporal com métodos clássicos (SARIMA e Holt-Winters) comparando-os ainda com outras abordagens. A análise qualitativa apoiou-se em reuniões com os responsáveis da empresa, de modo a identificar e interpretar eventuais anomalias nos dados.

Esta abordagem proporciona:

- uma flexibilidade, permitindo a adaptação a casos específicos;
- uma riqueza informativa, possibilitando a captação de nuances que métodos puramente quantitativos poderiam desconsiderar;
- e uma aplicabilidade prática, com resultados orientados para a realidade concreta e necessidades específicas do contexto analisado.

"O estudo de caso é particularmente útil para investigar fenómenos contemporâneos dentro dos seus contextos da vida real, especialmente quando os limites entre o fenómeno e o contexto não são claros" (Yin, 2018, p. 16).

## **Metodologia**

Neste projeto foi adotada a metodologia CRISP-DM (Cross-Industry Standard Process for Data Mining) para identificar o modelo mais adequado para prever os tempos de execução

das tarefas de consultores e técnicos nos departamentos BCONS, BSERV, INFSO e SWFAC, com base nos dados históricos provenientes do sistema ATX. Esta abordagem sistemática, estruturada em seis fases interligadas, garante a robustez e a fiabilidade dos resultados obtidos, em conformidade com as melhores práticas de ciência de dados (Chapman et al., 2000).

O objetivo do estudo consiste na comparação de cinco modelos de previsão de séries temporais com sazonalidade: SARIMA, Holt-Winters, N-Beats, DeepAR e Prophet.

A metodologia baseia-se numa combinação de revisão bibliográfica, implementação prática e avaliação rigorosa dos modelos. Segundo Hyndman & Athanasopoulos (2018, Cap. 5.9), a escolha do modelo adequado é crucial para garantir previsões precisas e fiáveis, especialmente em séries temporais caracterizadas por padrões sazonais complexos.

## **Estrutura do Documento**

O presente documento encontra-se estruturado de forma a proporcionar uma análise rigorosa, sistemática e fundamentada para a previsão de horas executáveis. A organização do conteúdo segue uma lógica sequencial coerente com a metodologia CRISP-DM, que favorece uma leitura fluida e uma compreensão progressiva dos objetivos, procedimentos metodológicos e resultados obtidos. A estrutura compreende os seguintes capítulos principais:

- **Introdução:** esta secção apresenta o enquadramento do projeto, o objetivo geral e os objetivos específicos, bem como a relevância da previsão de horas de execução no contexto da eficiência operacional e da satisfação do cliente.
- **Revisão da Literatura:** apresentam-se os principais contributos teóricos e empíricos no âmbito dos modelos de previsão temporal, englobando tanto as abordagens clássicas - SARIMA, Holt-Winters - quanto as contemporâneas, nomeadamente N-BEATS, DeepAR e Prophet.
- **Metodologia:** descreve detalhadamente o percurso metodológico adotado, desde a compreensão do problema de negócio até à implementação dos modelos preditivos, com base na metodologia CRISP-DM. Inclui a preparação dos dados, a seleção e parametrização dos modelos, bem como os critérios de avaliação utilizados.
- **Apresentação e Discussão dos Resultados:** procede à análise comparativa do desempenho dos modelos aplicados, com base em métricas como Erro Absoluto

Médio (MAE), Raiz do Erro Quadrático Médio (RMSE), Erro Percentual Absoluto Médio (MAPE), Critérios de Informação de Akaike (AIC) e Critérios de Informação Bayesiano (BIC), complementada por testes estatísticos, nomeadamente o teste de Diebold-Mariano. São igualmente analisadas as propriedades dos resíduos e a robustez dos modelos.

- **Conclusões e Recomendações:** sintetiza os principais resultados alcançados, destacando-se os modelos com melhor desempenho, e propondo-se direções para investigação futura.
- **Referências Bibliográficas:** apresenta uma compilação abrangente de obras fundamentais e estudos recentes no domínio da análise de séries temporais, métodos de previsão e avaliação de modelos.

Esta estrutura visa assegurar a coerência entre os objetivos do estudo, a metodologia adotada e os resultados obtidos, promovendo a aplicabilidade prática das conclusões no contexto empresarial.

## 1. Revisão da Literatura

Neste capítulo, apresenta-se uma revisão abrangente da literatura existente, relacionada com o tema deste trabalho. O principal objetivo desta revisão é identificar os modelos de previsão em séries temporais que têm sido amplamente utilizados no contexto da aplicação a este tipo de dados, bem como posicionar este estudo no panorama da literatura atual.

Será também realizada uma introdução teórica aos modelos selecionados, destacando as suas principais características que sustentam a sua inclusão no presente trabalho. Por fim, será efetuada uma revisão das métricas adotadas para a avaliação do desempenho de cada modelo, evidenciando a sua relevância e adequação para a análise e interpretação dos resultados obtidos.

### 1.1. Estado da Arte

A previsão de séries temporais constitui uma área crítica em diversas áreas, tais como finanças, vendas, turismo, meteorologia, entre outras. A escolha do modelo de previsão adequado é crucial para obter resultados precisos e fiáveis.

Hyndman & Athanasopoulos (2018, Cap. 6), definem sazonalidade como um padrão que se repete em intervalos regulares, como dias, meses ou trimestres, podendo ser modelado quer por métodos de decomposição clássica, quer por modelos como o SARIMA. Por sua vez, Box et al. (2008, Cap. 9) destacam que o modelo SARIMA mostra-se particularmente eficaz em séries temporais com sazonalidade, ao conjugar componentes autoregressivos (AR), de média móvel (MA) e diferenciação sazonal.

Além disso, abordagens mais recentes, como redes neuronais recorrentes (RNNs) e Transformers, têm sido aplicadas para captar padrões sazonais complexos em grandes volumes de dados (Vaswani et al., 2017; Salinas et al., 2020).

A avaliação de desempenho dos modelos de previsão também é um aspeto crítico. Embora medidas de erro como o MAPE e o RMSE sejam amplamente utilizadas, o Erro Percentual Absoluto Mediano (MdAPE) revela-se preferível em cenários com *outliers* (Hyndman & Koehler, 2006; Chai & Draxler, 2014).

Segundo Chatfield & Xing (2019, Cap. 2), uma série temporal pode ser decomposta em diferentes componentes — tendência, sazonalidade, ciclos e irregularidades — que, em conjunto, descrevem o comportamento observado ao longo do tempo.

Na obra *Análise de Séries Temporais* (3ª ed., 2018), Morettin & Toloí apresentam a decomposição de uma série temporal em componentes fundamentais — tendência, sazonalidade, ciclos e ruído/irregularidade — como base para a compreensão e modelação do comportamento dos dados ao longo do tempo.

A Interpretação de Morettin & Toloí (2018) sobre os Componentes de Séries Temporais é a seguinte:

- **Tendência (T):** representa o movimento de longo prazo da série, refletindo uma direção geral — seja de crescimento, declínio ou estabilidade. É o componente que representa a evolução estrutural dos dados ao longo do tempo.
- **Sazonalidade (S):** refere-se a variações sistemáticas e periódicas que ocorrem em intervalos regulares, geralmente dentro de um ano. Estas flutuações são atribuídas a fatores como clima, calendário ou hábitos de consumo.
- **Ciclo (C):** embora semelhante à sazonalidade, o ciclo tem uma duração mais longa e não fixa. Está frequentemente associado a fenómenos económicos, como os ciclos de negócios, e pode ser mais difícil de identificar devido à sua irregularidade.
- **Ruído ( $\epsilon$ ):** representa as variações aleatórias e imprevisíveis que não são explicadas pelos componentes anteriores. É considerado o resíduo da série após a remoção da tendência, sazonalidade e ciclo, e assume papel importante na avaliação da qualidade dos modelos.

Um conceito basilar na análise de séries temporais é a **estacionariedade**, que implica que as propriedades estatísticas de uma série (como a média, variância e estrutura de autocorrelação) não mudam ao longo do tempo. Esta propriedade é crucial para a aplicação de muitos modelos estatísticos tradicionais, pois permite que os parâmetros do modelo sejam estimados de forma consistente. Existem dois tipos principais de estacionariedade: a **estacionariedade estrita**, onde a distribuição de probabilidade conjunta da série é invariante no tempo; e a **estacionariedade fraca** (ou de segunda ordem), na qual a média e a variância são constantes e a autocovariância depende apenas do *lag* temporal (Box et al., 2015). A não estacionariedade, comum em séries temporais reais, pode ser tratada através de métodos como a diferenciação (cálculo das diferenças entre observações consecutivas) ou

transformações (por exemplo, logaritmos), sendo estes passos críticos na modelação de séries temporais.

A modelação da sazonalidade é particularmente relevante na previsão, com diversas abordagens disponíveis. Modelos estatísticos clássicos, como o **SARIMA**, são amplamente utilizados e revelam-se eficazes para séries temporais que apresentem padrões sazonais

Como indicado anteriormente, existem diversos modelos de previsão que podem ser utilizados para séries temporais com sazonalidade. Este trabalho irá incidir sobre os seguintes modelos:

- Modelos Clássicos
  - SARIMA
  - Holt-Winters
- Modelos Baseados em Redes Neurais
  - N-Beats
  - DeepAR
- Modelos Híbridos
  - Prophet

Proceder-se-á à definição dos parâmetros iniciais de cada modelo, com base na literatura e nas características dos dados (Box et al., 2008, Cap. 7).

O modelo SARIMA é uma extensão do modelo Autoregressive Integrated Moving Average (ARIMA) que incorpora explicitamente a componente sazonal na modelação de séries temporais. Enquanto o ARIMA é eficaz para capturar tendências e padrões não sazonais, o SARIMA adiciona parâmetros específicos para modelar a sazonalidade, o que o torna adequado para séries com padrões sazonais estáveis e bem definidos. O modelo SARIMA é definido por sete parâmetros:  $(p, d, q) \times (P, D, Q, S)$ , onde  $p, d$  e  $q$  representam, respetivamente, a ordem dos componentes autoregressivo (AR), integrado (I) e de média móvel (MA) não sazonais, enquanto  $P, D$  e  $Q$  representam os mesmos componentes para a parte sazonal. O parâmetro  $S$  indica o período sazonal, ou seja, o número de observações que compõem um ciclo sazonal completo (Hyndman & Athanasopoulos, 2018, Cap. 6).

Segundo Hyndman & Athanasopoulos (2018, Cap. 8), o modelo SARIMA é particularmente eficaz para séries temporais com padrões sazonais estáveis e bem definidos. Por exemplo, é

amplamente utilizado na previsão de vendas para capturar flutuações sazonais, como aumentos nas vendas durante o Natal ou em épocas de férias. Também é aplicado no setor de energia para prever o consumo de energia elétrica, com uma variação significativa nas várias estações do ano, e no transporte para antecipar a procura por serviços de transporte público em horários de pico. Além disso, o SARIMA é utilizado na área financeira para a modelação de séries temporais com padrões sazonais, como as taxas de câmbio ou preços de *commodities*.

Uma das principais vantagens do SARIMA é a sua flexibilidade, pois permite modelar tanto a componente não sazonal como a sazonal de uma série temporal. No entanto, o modelo pode ter dificuldades em lidar com séries que apresentam sazonalidade múltipla (por exemplo, padrões diários e semanais simultaneamente) ou mudanças abruptas nos padrões sazonais. Além disso, a seleção dos parâmetros do SARIMA pode ser complexa e requer uma análise cuidadosa dos dados, incluindo a identificação da ordem de diferenciação e dos termos autoregressivos e de média móvel. Apesar destes desafios, o SARIMA, por ter uma grande flexibilidade, continua a ser uma escolha popular devido à sua capacidade de captar padrões complexos e à sua ampla aplicabilidade em diversas áreas (Hyndman & Athanasopoulos, 2018, Cap. 8).

O método de Holt-Winters é um modelo de alisamento exponencial que incorpora explicitamente a tendência e a sazonalidade. Este método foi desenvolvido a partir do modelo de *Holt* com o objetivo de resolver o problema da existência de sazonalidade nos dados. O método é composto por uma equação de previsão e três equações de alisamento, uma para o nível, outra para a tendência, e uma terceira para a sazonalidade (Hyndman & Athanasopoulos, 2018, Cap. 7). Este modelo é amplamente utilizado devido à sua simplicidade e eficácia em séries com padrões sazonais estáveis.

Este método apresenta duas variações principais de sazonalidade — aditiva e multiplicativa — cada uma disponível numa versão com tendência amortecida (*damped*).

A primeira variação, o Holt-Winters Aditivo, modela a sazonalidade como uma componente aditiva, ou seja, a magnitude do efeito sazonal é constante ao longo do tempo. Este modelo é adequado para séries temporais em que a amplitude da sazonalidade não varia com o nível da série. A segunda variação, o Holt-Winters Multiplicativo, faz a modelação da sazonalidade como uma componente multiplicativa, onde a magnitude do efeito sazonal é proporcional ao nível da série. Este modelo é particularmente adequado a séries em que a amplitude sazonal

varia com o nível da série. A variante Holt–Winters com tendência amortecida inclui um parâmetro que reduz gradualmente o impacto da componente de tendência ao longo do horizonte de previsão. Este modelo é indicado para séries em que a tendência tende a estabilizar-se ou eventualmente a decrescer ao longo do tempo.

De acordo com Makridakis et al. (1998, Cap. 5), o Holt-Winters é particularmente adequado para previsões de curto prazo, mas pode não ser eficaz nas séries com mudanças estruturais ou sazonalidade irregular ou com ruído excessivo. A sua implementação é relativamente simples, o que o torna numa escolha comum em aplicações práticas.

O N-Beats é uma arquitetura de rede neuronal concebida especificamente para a previsão de séries temporais. Desenvolvida para ultrapassar as limitações de modelos clássicos — como o SARIMA e os métodos de alisamento exponencial —, assenta em blocos neuronais modulares que captam padrões complexos, tais como tendências e sazonalidades, de forma eficaz. Uma das suas inovações é a capacidade de aprender representações hierárquicas diretamente dos dados, decompondo a série em componentes interpretáveis (por exemplo, tendência e sazonalidade) sem recurso a pré-processamento complexo. Oreshkin et al. (2019) demonstram que esta abordagem permite obter previsões precisas e interpretáveis, mesmo em cenários com padrões temporais variados.

Este modelo é altamente escalável e pode ser aplicado a múltiplas séries temporais simultaneamente, o que o torna particularmente útil nos cenários onde é necessário prever o comportamento de várias séries correlacionadas. Além disso, o modelo é eficaz nas tarefas de previsão de longo prazo, onde a captação de dependências temporais de longo alcance é essencial.

Uma das vantagens do N-Beats é a sua interpretabilidade. Ao contrário de muitas redes neuronais profundas, que funcionam como "caixas negras", o N-Beats permite decompor as previsões em componentes interpretáveis, como tendências e sazonalidades, o que facilita a análise e a tomada de decisões. No entanto, a sua implementação requer um volume significativo de dados para treino, o que pode ser uma limitação nos cenários com dados escassos. Além disso, a complexidade computacional do modelo pode ser um desafio nos ambientes com recursos limitados, embora a sua eficiência em termos de precisão e escalabilidade compense este custo em muitas aplicações práticas (Oreshkin et al., 2019).

Desenvolvido pela Amazon, o DeepAR é um modelo baseado em redes neurais recorrentes (RNNs) que utiliza uma abordagem probabilística para a previsão de séries temporais. Salinas et al. (2020) afirmam que o DeepAR é eficaz para captar dependências temporais e sazonalidade em séries com comportamento não linear. O modelo é particularmente útil para previsões de longo prazo e cenários com múltiplas séries temporais correlacionadas. Contudo, a sua implementação requer um volume significativo de dados para treino, o que pode ser uma limitação em alguns contextos.

Como principais características, o DeepAR é utilizado para previsão probabilística, ao contrário dos modelos tradicionais que fornecem previsões pontuais, e estima a distribuição de probabilidade dos valores futuros, permitindo uma melhor compreensão da incerteza. Este modelo utiliza RNNs para capturar dependências temporais nos dados, o que permite que aprenda com padrões complexos e não lineares. Também é projetado para lidar com grandes conjuntos de séries temporais relacionadas, o que permite que aprenda com padrões comuns entre as séries e melhore a precisão das previsões. Para além das características anteriores, este modelo pode incorporar variáveis exógenas (covariáveis) para melhorar a precisão das previsões.

Taylor & Letham (2018) destacam que o Prophet, um modelo de previsão de séries temporais desenvolvido pelo Facebook (atual Meta), é especialmente útil para séries temporais com padrões sazonais complexos e dados incompletos. O modelo adota uma abordagem aditiva para decompor a série temporal em tendência, sazonalidade e efeitos de feriados (correspondem às variações sazonais excecionais que ocorrem em dias festivos ou datas especiais e que não se repetem necessariamente com a mesma periodicidade anual padrão), conferindo-lhe elevada interpretabilidade. Contudo, o seu desempenho pode ser prejudicado em séries com ruído excessivo ou padrões não lineares acentuados.

Uma das principais vantagens do Prophet é a sua capacidade de lidar com dados incompletos e *outliers*, o que o torna robusto nos cenários do mundo real onde os dados podem apresentar algum ruído. No entanto, o Prophet pode ter limitações nas séries temporais com ruído excessivo ou padrões não lineares complexos, situações em arquiteturas baseadas em redes neurais, como o N-Beats ou variantes do Transformer, frequentemente obtêm um desempenho superior. Adicionalmente, o modelo pode não ser o mais adequado para previsões de longo prazo em séries sujeitas a mudanças estruturais abruptas. Apesar destas limitações, a simplicidade da implementação, a elevada interpretabilidade e a flexibilidade

para incorporar fatores exógenos tornam o Prophet uma ferramenta de referência em múltiplas aplicações práticas (Taylor & Letham, 2018, p. 37–45).

A **avaliação do desempenho** dos modelos de previsão é um aspecto crítico para determinar a sua eficácia. Para além das métricas amplamente utilizadas como o MAPE e o RMSE, que medem a magnitude dos erros, outras métricas oferecem perspectivas adicionais (Hyndman & Koehler, 2006). De seguida apresenta-se uma breve caracterização dessas métricas:

- **Erro Absoluto Médio (MAE):** média dos erros absolutos, sendo menos sensível a *outliers* do que o RMSE.
- **Erro Percentual Absoluto Médio (MAPE):** média dos erros percentuais absolutos, expressa a precisão em percentagem, facilitando a comparação entre séries com escalas distintas.
- **Raiz do Erro Quadrático Médio (RMSE):** raiz quadrada da média dos erros quadráticos; penaliza mais fortemente grandes desvios, sendo mais sensível a *outliers*.
- **Erro Absoluto Mediano (MdAE) e Erro Percentual Absoluto Mediano (MdAPE):** são métricas mais robustas à presença de *outliers* na série, uma vez que se baseiam na mediana dos erros.
- **sMAPE (Erro Percentual Absoluto Médio Simétrico):** uma alternativa ao MAPE que penaliza simetricamente os erros de sobrestimação e subestimação, evitando distorções quando os valores reais se aproximam de zero.
- **Erro Médio de Previsão (Estimativa do Viés):** média dos erros (previsão – valor real); indica a tendência sistemática de sobrestimação ou subestimação; valores próximos de zero são desejáveis.
- **Erro Absoluto Médio Escalado (MASE):** erro absoluto escalado em relação a uma previsão ingénua (o valor futuro será igual ao último valor observado); independente da escala dos dados e robusto a *outliers*; valores inferiores a 1 indicam desempenho superior ao da abordagem ingénua.
- **Pontuação Média de Intervalo Escalado (MSIS):** avalia a qualidade de intervalos de previsão probabilísticos, penalizando larguras excessivas e falhas na cobertura dos valores reais (Gneiting & Raftery, 2007).
- **Soma dos Quadrados dos Resíduos (RSS):** a soma dos quadrados dos resíduos, diretamente relacionada com o MSE, indicando a variância não explicada pelo modelo.
- **Critérios de Informação (AIC, AICC (Versão corrigida do AIC para amostras pequenas) e BIC):** métricas que penalizam a complexidade do modelo em relação ao

seu ajustamento. Valores mais baixos sugerem melhor equilíbrio entre parcimónia e qualidade do ajustamento (Akaike, 1974; Schwarz, 1978).

O tratamento de valores em séries temporais é uma etapa crucial no pré-processamento de dados, que permite lidar com os **valores ausentes (*missing values*)** e com os **valores atípicos (*outliers*)**. Ambos podem comprometer a robustez e a precisão das análises e previsões, exigindo a aplicação de técnicas específicas de imputação, interpolação e deteção/mitigação de outliers para assegurar a integridade e fiabilidade dos dados (Shumway & Stoffer, 2017; Chatfield, 2003).

O tratamento de **valores atípicos (*outliers*)**, observações que se afastam significativamente do padrão geral dos dados, potencialmente devido a erros de medição ou eventos excecionais (Aggarwal, 2017), recorre, neste trabalho, ao método de médias móveis. Esta abordagem desenvolve-se em duas fases:

- **Deteção:** considera-se um valor atípico sempre que o desvio entre o dado observado e a média móvel local excede um limiar pré-definido, sinalizando uma anomalia face ao comportamento médio da série.
- **Correção/Substituição:** o outlier identificado é tratado como valor em falta e substituído pela média móvel calculada a partir dos vizinhos imediatos. A escolha da variante de média móvel — simples, ponderada ou exponencial — dependerá das características da série e do grau de suavização pretendido.

## 2. Objetivos e Metodologia

### 2.1. Objetivos

O principal objetivo deste projeto é identificar o modelo mais adequado para prever os tempos de execução (em horas) das tarefas realizadas pelos consultores e técnicos dos departamentos de Business Consulting (BCONS), Business Services (BSERV), Infrastructure Solutions (INFSO) e Software Factory (SWFAC), utilizando os dados históricos extraídos do sistema ATX. Para tal, serão ajustados aos dados cinco modelos de previsão de séries temporais com sazonalidade: SARIMA, Holt-Winters, N-BEATS, DeepAR e Prophet. A abordagem adotada é um estudo de caso focado na análise exploratória e preditiva de séries temporais, utilizando técnicas de *Machine Learning*. A escolha criteriosa do modelo é fundamental para garantir previsões precisas e fiáveis, especialmente em séries com padrões sazonais complexos (Hyndman & Athanasopoulos, 2018, Cap. 5.9).

### 2.2. Metodologia

Neste projeto recorre-se à metodologia CRISP-DM para identificar o modelo de previsão mais adequado. Esta abordagem sistemática, estruturada em seis fases interligadas, garante a robustez e fiabilidade dos resultados, alinhando-se com as melhores práticas em ciência de dados (Chapman et al., 2000).

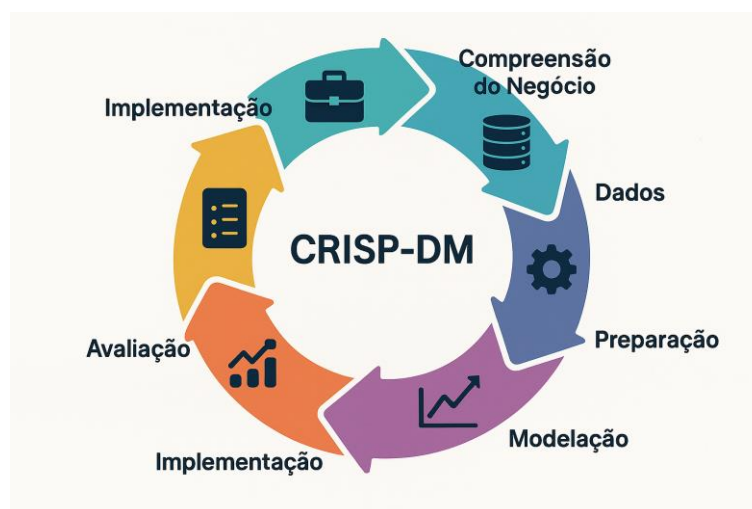


Figura 1 – Fases da Metodologia CRISP-DM

As fases da Metodologia CRISP-DM aplicadas neste projeto são:

- **Compreensão do Negócio (*Business Understanding*):** definição dos objetivos do projeto, dos requisitos do utilizador, das métricas de sucesso e da priorização de tarefas.
- **Compreensão dos Dados (*Data Understanding*):** recolha inicial de dados do sistema ATX; descrição das suas características; avaliação da qualidade (incluindo valores ausentes e anómalos); e realização de análises exploratórias para identificar tendências e padrões sazonais.
- **Preparação dos Dados (*Data Preparation*):** limpeza dos dados (tratamento de valores ausentes e anómalos), transformações (normalização, agregação) e engenharia de *features* (criação de variáveis *lag* e indicadores sazonais).
- **Modelação (*Modeling*):** treino com os dados pré processados, ajuste dos parâmetros para otimização do desempenho (ajuste de *hiperparâmetros*), e seleção dos modelos de previsão.
- **Avaliação (*Evaluation*):** análise rigorosa do desempenho dos modelos através de um conjunto diversificado de métricas e testes estatísticos. Este processo inclui:
  - **Métricas de Desempenho:**
    - Erro Absoluto Médio (MAE)
    - Raiz do Erro Quadrático Médio (RMSE)
    - Erro Percentual Absoluto Médio (MAPE)
    - Erro Absoluto Mediano (MdAE)
    - Erro Percentual Absoluto Mediano (MdAPE)
    - Viés
    - Erro Médio Absoluto Escalado (MASE)
    - Pontuação do Intervalo Escalado Médio (MSIS)
  - **Testes Estatísticos para Resíduos:**
    - **Teste de Ljung-Box (LB):** para avaliar a autocorrelação dos resíduos (Hyndman & Athanasopoulos, 2018).
    - **Testes de normalidade** (ex.: Shapiro-Wilk, Jarque-Bera): utilizados para verificar se os resíduos seguem uma distribuição aproximadamente normal (Jarque & Bera, 1987; Shapiro & Wilk, 1965).
  - **Validação Walk-Forward:** técnica de validação cruzada sequencial para séries temporais, que simula o processo de previsão em tempo real, testando o modelo em períodos sequenciais de dados futuros.
  - **Teste de Diebold-Mariano (DM):** comparação estatística da precisão das previsões entre diferentes modelos (Diebold, 1995; Harvey et al., 1997; Diebold, 2015).

- A seleção do modelo final é um processo complexo que exige uma análise rigorosa destas métricas, abrangendo tanto a precisão das previsões pontuais quanto a qualidade da quantificação da incerteza (Makridakis et al., 2020).
- **Implementação (*Deployment*):** integração do modelo selecionado no ambiente de produção, estabelecimento de mecanismos para monitorização contínua do seu desempenho e recolha de *feedback* dos utilizadores.

### **3. Apresentação e discussão dos resultados**

No presente capítulo procede-se à apresentação, análise e discussão dos resultados obtidos através da modelação de séries temporais. Aplicaram-se diversos modelos de previsão com o objetivo de gerar estimativas precisas e fiáveis, capazes de antecipar tendências e padrões futuros.

A estrutura deste capítulo foi cuidadosamente delineada para garantir uma apresentação clara e organizada dos resultados. Inicialmente, foram apresentados os resultados obtidos pela modelação, com recurso a gráficos e tabelas, que facilitam a visualização e interpretação comparativa do seu desempenho.

De seguida, realizou-se uma análise pormenorizada dos resultados, avaliando a sua relevância em função dos objetivos da investigação formulados no Capítulo I e confrontando-os com os contributos da literatura previamente analisada. Esta discussão tem em conta possíveis convergências e divergências, de modo a situar as descobertas no panorama científico atual.

A apresentação e análise detalhada dos resultados contribuem para validar a adequação dos modelos escolhidos e para extrair insights relevantes para a tomada de decisões empresariais, ao mesmo tempo que apontam direções para investigações futuras. Deste modo, o capítulo oferece uma interpretação integrada dos dados em estudo, assegurando que as conclusões alcançadas se apoiam numa base rigorosa e são passíveis de aplicação concreta em contexto profissional.

#### **3.1. Implementação dos Modelos**

##### **3.1.1. Configuração do Ambiente**

Foi utilizado o Visual Studio Code com Jupyter Notebook para implementação dos modelos em Python e foram instaladas as bibliotecas necessárias, como statsmodels para SARIMA e Holt-Winters, pytorch para N-Beats e DeepAR, fbprophet para Prophet, e scikit-learn para métricas de avaliação.

### 3.1.2. Implementação dos Modelos

Na fase de modelação do CRISP-DM procedeu-se à implementação prática dos modelos de previsão seleccionados para a série temporal em estudo, envolvendo a seleção, o treino e a otimização dos parâmetros das cinco arquiteturas distintas: SARIMA, Holt-Winters, N-BEATS, DeepAR e Prophet. Para cada modelo, a implementação incluiu a divisão dos dados em conjuntos de treino e de teste, o ajustamento aos dados de treino e a geração de previsões posteriormente avaliadas. A otimização dos parâmetros e hiperparâmetros constituiu uma etapa fundamental para maximizar o desempenho preditivo de cada modelo.

### 3.2. Descrição dos Dados Utilizados

Para a realização deste trabalho, foi utilizado um conjunto de dados com registos mensais de um período de 12 anos, de 2012 a 2024, contendo informação detalhada sobre as atividades realizadas pelos consultores e técnicos dos diferentes departamentos da organização.

Os dados (ver Tabela A.1 incluída no Apêndice A) são organizados em registos mensais detalhando o tempo dedicado a cada tarefa, o departamento responsável e o técnico envolvido. Além disso, o conjunto de dados inclui ainda diferentes tipos de atividades, como serviços contratados, projetos, investimentos em pesquisa e desenvolvimento (I&D) e atividades internas de gestão e suporte.

De forma mais detalhada, o conjunto de dados diz respeito às seguintes variáveis, para cada uma das tarefas realizadas:

- **GrupoAtividade:** categoria geral da atividade realizada, como "Cliente", "Comercial" e "Interno".
- **Atividade:** descrição mais específica da atividade, como "Serviço | Contrato", "Serviço | Faturado", "Serviços | Projeto", "Comercial", "Investimento | I&D", "Não Produtivas | Faltas e Comp.", "Operação | Gestão Interna", "Operação | Suporte Interno", "Operação | Projeto Interno", "Não Produtivas | Férias", "Investimento | Formação", "Não Produtivas | Baixa".
- **AnoMes:** data no formato de ano e mês (AAAA-MM) em que a atividade foi realizada.
- **Dep:** departamento responsável pela atividade, como "BCONS", "BSERV", "SWFAC", "INFSO", "PMO", "SALES", "CS", "FS", "HR"

- **Tec:** técnico, consultor ou profissional responsável pela execução da atividade, identificado por um código.
- **Tempo Realizado:** tempo associado a realização da atividade, medido em horas.

No contexto deste trabalho, estes dados serviram para analisar a eficiência operacional da organização, identificar padrões de alocação de recursos e avaliar o impacto de diferentes atividades no desempenho global. A análise temporal permitiu identificar períodos de maior e menor atividade, enquanto a segmentação por departamento e técnico forneceu insights sobre a distribuição de tarefas e a produtividade individual.

As atividades ao serem categorizadas em diferentes grupos (serviços, projetos, investimentos, entre outros), possibilitaram uma análise mais detalhada da distribuição dos recursos. Para cada grupo, a métrica Tempo Realizado poderá ser utilizada como variável de referência, permitindo a quantificação do esforço associado e servindo de base para o cálculo de indicadores de eficiência e produtividade.

De notar que foram eliminados os registos referentes aos departamentos de Recursos Humanos (RH), Comercial (SALES), Administrativo (CS) e Gestão de Projetos (PMO) por não serem departamentos associados à área operacional da organização e desta forma não fazer parte do âmbito da análise.

Para realizar a análise dos dados, estes foram agrupados pelo campo AnoMes, considerando o somatório dos valores do campo Tempo Realizado. Este procedimento permitiu a construção de uma série temporal representativa do total mensal dos tempos realizados pelos consultores e técnicos dos departamentos BCONS, BSERV, INFISO e SWFAC. Note-se que o campo AnoMes foi convertido para o tipo de dados datetime e formatado segundo a convenção YYYY-MM (ver Tabela A.2 Apêndice A).

O cronograma, Figura 2, "Resumo do Tempo Realizado por Ano Mês" revela um aumento notável nos tempos realizados a partir de janeiro de 2022. Este incremento significativo é justificado pela aquisição de outra empresa concorrente que operava na mesma área. A integração das operações e da força de trabalho da empresa adquirida resultou num aumento natural no volume total de trabalho e, consequentemente, nos tempos realizados.

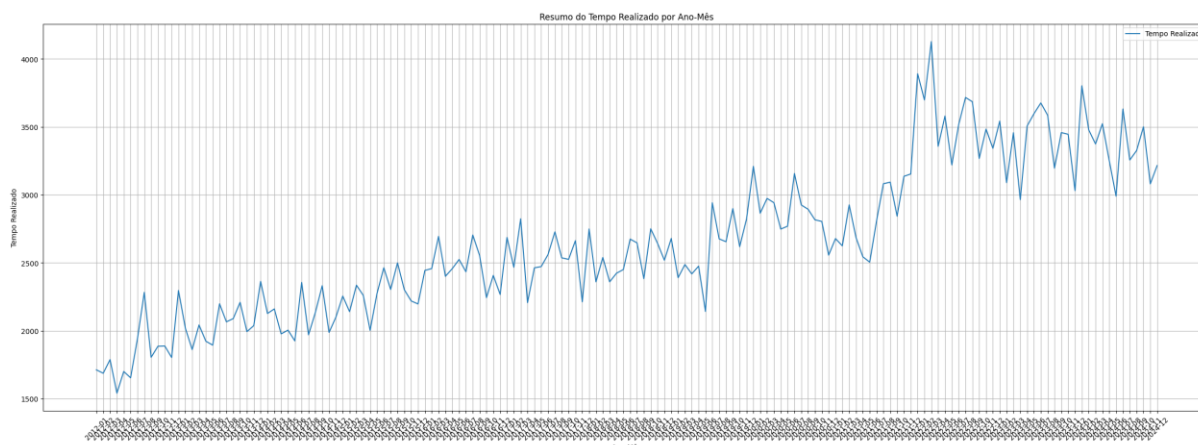


Figura 2 – Cronograma dos dados reais dos dados histórico a 12 anos (2012 a 2024)

### 3.3. Análise Estatística e Resultados

Nesta secção, os resultados são apresentados através de tabelas, gráficos e outras representações visuais que facilitam a interpretação dos dados. A análise estatística é conduzida com o objetivo de identificar padrões e tendências relevantes para o estudo, constituindo a base para a discussão subsequente.

Adicionalmente, são consideradas as implicações dos resultados no contexto específico da investigação, bem como as suas limitações e o potencial de generalização a outros dados.

Pretende-se, assim, que a análise estatística e a apresentação dos resultados neste subcapítulo proporcionem uma visão abrangente e aprofundada dos dados recolhidos, contribuindo para o cumprimento dos objetivos definidos para o estudo.

Conforme referido no subcapítulo anterior, observa-se que, a partir de janeiro de 2022, os valores da variável Tempo Realizado registaram um aumento superior à tendência verificada nos anos anteriores, em consequência do processo de aquisição de outra empresa. Esta alteração introduziu uma mudança estrutural na série temporal, com potencial impacto na qualidade das previsões. Por este motivo, decidiu-se restringir a análise ao período correspondente aos últimos três anos, de forma a obter estimativas mais consistentes com o contexto atual.

Na análise restrita ao período 2022–2024 (ver Figura 3), identificou-se um valor atípico no mês de março de 2022, significativamente superior aos restantes registos da série.

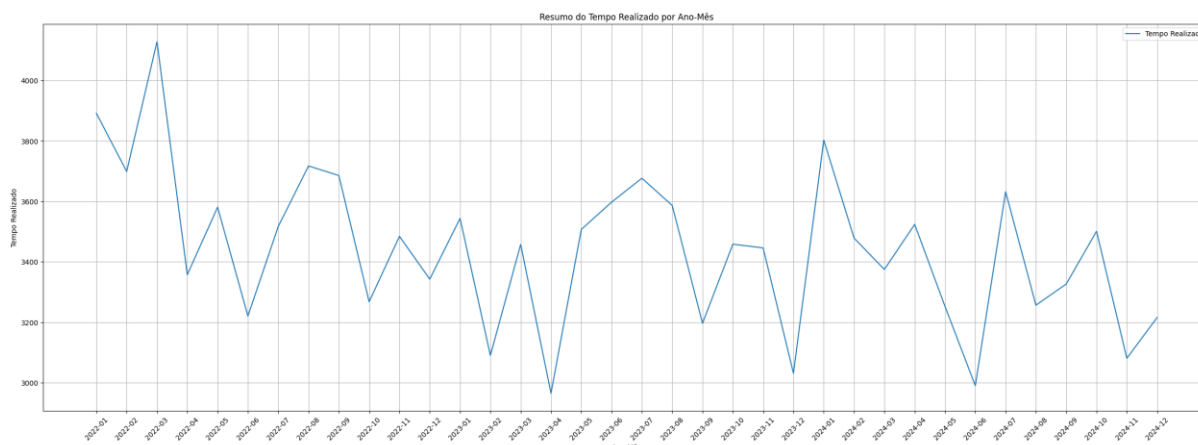


Figura 3 – Cronograma da série temporal com os dados originais dos últimos 3 anos (2022 a 2024)

Para o tratamento deste valor anômalo recorreu-se ao método das Médias Móveis, com o objetivo de suavizar as flutuações e reduzir o impacto do outlier na modelação subsequente. O valor estimado por este método foi de 3415.7073625 horas. Assim no novo *dataframe*, o valor original de março de 2022 (4127.515470 horas) foi substituído pelo valor suavizado (3415.7073625 horas).

### Estacionaridade da série temporal

O teste ADF (Augmented Dickey-Fuller) (Dickey & Fuller, 1979) é um método estatístico usado para verificar se uma série temporal é estacionária, isto é, uma série cujas propriedades estatísticas — como a média, a variância e a autocorrelação — se mantêm constantes ao longo do tempo. A estacionaridade é importante porque muitos modelos de previsão, como o SARIMA, assumem que os dados não apresentam variações de tendência ou de variância ao longo do tempo. Neste teste, a hipótese nula assume que a série não é estacionária, enquanto a hipótese alternativa considera que a série é estacionária.

Os resultados obtidos indicam que tanto a série com os dados originais como a série com os dados suavizados podem ser consideradas estacionárias, dado que em ambas as situações, se rejeitou a hipótese nula (ver Tabela A.3 Apêndice A).

#### 3.3.1. Análise do resultado de cada modelo

Nesta secção, cada modelo foi analisado através da comparação entre duas versões, uma versão base e outra que é simplificada ou otimizada. Os resultados das diversas métricas de

avaliação foram sintetizados em quadros comparativos das duas versões em cada um dos modelos (SARIMA, Holt-Winters, N-BEATS, DeepAR e Prophet). O objetivo é avaliar o desempenho de todos os modelos com base em métricas de qualidade de ajustamento, testes de diagnóstico dos resíduos e avaliação do desempenho preditivo fora da amostra.

A seleção entre modelos alternativos deve equilibrar múltiplos critérios, incluindo a capacidade preditiva, a interpretabilidade e a robustez. Os critérios de informação como AIC e BIC proporcionam orientação quantitativa, mas devem ser complementados por validação *out-of-sample* para avaliação da capacidade de generalização (Burnham & Anderson, 2002).

Salienta-se que, neste estudo, a avaliação *in-sample* foi realizada com base nas métricas AIC, BIC e AICc, enquanto a capacidade preditiva *out-of-sample* dos modelos foi validada através da técnica *walk-forward*, utilizando as métricas MAE, RMSE, MAPE, MdAE, MdAPE, Viés e MASE.

Os modelos foram treinados com uma janela de entrada de 36 observações, já descritas anteriormente, e utilizado um horizonte de previsão de 12 períodos.

#### **3.3.1.1. SARIMA - Comparação entre os Modelos Geral e Simplificado**

Nesta secção apresenta-se uma análise comparativa entre duas especificações do modelo SARIMA: o modelo geral, definido como SARIMAX(2,1,1)×(1,1,1,12), e o modelo simplificado, definido como SARIMAX(2,1,0)×(0,1,0,12). O modelo geral incorpora componentes autorregressivos, de integração e de média móvel na parte sazonal e não sazonal, enquanto o modelo simplificado remove todos os componentes de média móvel na parte sazonal e não sazonal. Na parte sazonal também se elimina a componente autorregressiva.

O modelo geral incorpora componentes autorregressivos (AR), de média móvel (MA) e de diferenciação (I) tanto na parte sazonal como na parte não sazonal. Esta configuração permite capturar dependências temporais complexas e padrões sazonais bem definidos, sendo particularmente recomendada para séries com estrutura rica e autocorrelação significativa (Makridakis et al., 1998).

Por outro lado, o modelo simplificado elimina os componentes de média móvel e autorregressivos na parte sazonal, bem como o termo MA na parte não sazonal. Esta simplificação visa reduzir a complexidade do modelo e evitar o sobreajuste, especialmente em contextos onde os padrões sazonais são menos pronunciados ou o ruído é elevado. A

abordagem segue o princípio da parcimônia, segundo o qual modelos mais simples devem ser preferidos quando oferecem desempenho comparável (Hyndman & Athanasopoulos, 2021, Cap. 5).

Ambos os modelos mantêm a diferenciação sazonal ( $D=1$ ), essencial para lidar com a sazonalidade anual ( $S=12$ ), conforme recomendado na literatura para séries com periodicidade mensal (Morettin & Toloi, 2018, Cap. 5).

Na Tabela 1 apresentam-se os resultados das métricas de qualidade do ajustamento AIC, BIC, MAE, RMSE, MAPE e viés, bem como os resultados dos testes de diagnóstico dos resíduos, nomeadamente Ljung-Box e Jarque-Bera.

Tabela 1 - Quadro comparativo dos modelos SARIMA

| Métrica       | Geral                         | Simplificado                  | Melhoria (%) | Observações                                                         |
|---------------|-------------------------------|-------------------------------|--------------|---------------------------------------------------------------------|
| Especificação | SARIMAX(2,1,1)<br>×(1,1,1,12) | SARIMAX(2,1,0)<br>×(0,1,0,12) | —            | Modelo geral mais completo; simplificado com menos parâmetros.      |
| MAE           | 187.93                        | 287.66                        | ↓ 34.7%      | Modelo geral com erro absoluto médio significativamente menor.      |
| MSE           | 47 566.91                     | 110 139.23                    | ↓ 56.8%      | Redução expressiva no erro quadrático médio no modelo geral.        |
| RMSE          | 218.09                        | 331.87                        | ↓ 34.3%      | Melhor desempenho na raiz do erro quadrático médio do modelo geral. |
| MAPE          | 5.56%                         | 8.59 %                        | ↓ 35.3%      | Modelo geral mais preciso em termos percentuais.                    |
| MdAE          | 175.55                        | 249.83                        | ↓ 29.8%      | Menor erro absoluto mediano no modelo geral.                        |
| MdAPE         | 5.33%                         | 7.80 %                        | ↓ 31.7%      | Redução significativa no erro percentual mediano no modelo geral.   |
| MASE          | NaN                           | NaN                           | —            | Métrica não calculada.                                              |

|                     |                |                |         |                                                                                |
|---------------------|----------------|----------------|---------|--------------------------------------------------------------------------------|
| Viés                | -19.57         | -103.12        | ↓ 81.0% | Modelo geral com menor tendência sistemática de erro.                          |
| sMAPE               | 5.60%          | 8.79%          | ↓ 36.3% | Modelo geral com menor erro percentual simétrico.                              |
| MSIS                | NaN            | NaN            | —       | Métrica não calculada.                                                         |
| RSS                 | 570 802.93     | 1 321 670.79   | ↓ 56.8% | Soma residual dos quadrados muito menor no modelo geral.                       |
| AIC                 | 331.99         | 311.04         | ↑ 6.3%  | Modelo simplificado favorecido pela penalização de complexidade.               |
| AICc                | NaN            | NaN            | —       | Não disponível.                                                                |
| BIC                 | 338.81         | 314.17         | ↑ 7.3%  | Modelo simplificado preferido por penalizar fortemente o número de parâmetros. |
| Log_Likelihood      | -160.00        | -152.52        | ↓ 4.7%  | Modelo geral com menor log-Likelihood (menos ajustado segundo este critério).  |
| JB                  | p_value > 0.05 | p_value > 0.05 | —       | Ambos os modelos com resíduos aproximadamente normais.                         |
| LB (Lags de 1 a 12) | p_value > 0.05 | p_value > 0.05 | —       | Ambos sem evidência de autocorrelação (Lag1 a 12)                              |

O modelo simplificado apresentou menores valores de AIC e BIC, evidenciando um melhor *trade-off* entre parcimônia e capacidade preditiva. O modelo geral demonstrou desempenho superior nas outras métricas onde se utilizou a validação *walk-forward*.

A análise dos resíduos confirma que ambos os modelos são estatisticamente adequados, mas o modelo geral apresenta, na generalidade, menor viés e erros mais baixos.

Para visualizar toda a análise que foi efetuada deve-se consultar os resultados incluídos nos apêndices A e B conforme indicado na Tabela 2.

Tabela 2 - Quadro resumo com referência às Tabelas e Figuras

| <b>Elemento</b>                                                                        | <b>Referência</b>  | <b>Descrição</b>                                                                                      |
|----------------------------------------------------------------------------------------|--------------------|-------------------------------------------------------------------------------------------------------|
| Parâmetros e resultados do modelo SARIMA Geral e do modelo SARIMA Simplificado         | Tabela A. 4 e A. 6 | Resumo dos resultados do modelo SARIMA Geral e do modelo SARIMA Simplificado                          |
| Validação Walk-Forward do modelo SARIMA Geral e do modelo SARIMA Simplificado          | Tabela A. 5 e A. 7 | Valores das métricas de validação Walk-Forward do modelo SARIMA Geral e do modelo SARIMA Simplificado |
| P_values do Teste LB dos modelos SARIMA                                                | Tabela A. 8        | Comparação dos resultados do Teste de Ljung-Box entre as duas versões do modelo SARIMA                |
| Resíduos do modelo SARIMA Geral e do modelo SARIMA Simplificado                        | Figura B.3 e B.8   | Gráfico de resíduos do modelo SARIMA Geral e do modelo SARIMA Simplificado                            |
| Histograma e KDE dos resíduos do modelo SARIMA Geral e do modelo SARIMA Simplificado   | Figura B.4 e B.9   | Distribuição dos resíduos do modelo SARIMA Geral e do modelo SARIMA Simplificado                      |
| Q-Q plot do modelo SARIMA Geral e do modelo SARIMA Simplificado                        | Figura B.5 e B.10  | Q-Q plot dos resíduos do modelo SARIMA Geral e do modelo SARIMA Simplificado                          |
| ACF dos resíduos do modelo SARIMA Geral e do modelo SARIMA Simplificado                | Figura B.6 e B.11  | Correlograma dos resíduos do modelo SARIMA Geral e do modelo SARIMA Simplificado                      |
| Cronograma com dados reais e previsões do modelo Geral e do modelo SARIMA Simplificado | Figura B.7 e B.12  | Cronograma dos dados reais e das previsões do modelo SARIMA Geral e do modelo SARIMA Simplificado     |

A seleção entre modelos alternativos deve equilibrar múltiplos critérios, incluindo a capacidade preditiva, a interpretabilidade e a robustez. Os critérios de informação como AIC

e BIC proporcionam orientação quantitativa, mas devem ser complementados por validação *out-of-sample* para avaliação da capacidade de generalização (Burnham & Anderson, 2002).

Com base na análise comparativa, o modelo SARIMA Geral ( $\text{SARIMAX}(2, 1, 1) \times (1, 1, 1, 12)$ ) demonstra um desempenho superior claro em relação ao modelo SARIMA Simplificado ( $\text{SARIMAX}(2, 1, 0) \times (0, 1, 0, 12)$ ) em termos de precisão das previsões. As métricas de erro (MAE, MSE, RMSE, MAPE, MdAE, MdAPE) do modelo geral são consistentemente mais baixas, o que indica que este modelo é mais adequado para a série temporal em análise, apesar da sua maior complexidade. A simplificação dos termos (particularmente os de média móvel e os sazonais de AR/MA) no modelo simplificado resultou num aumento significativo dos erros de previsão.

Ambos os modelos SARIMA produziram resíduos que, de acordo com os testes de Ljung-Box, não evidenciam autocorrelação significativa, o que constitui um bom indicador de que as dependências temporais lineares foram adequadamente captadas pelo modelo.

Estas dependências lineares permitem ao modelo:

- Capturar padrões recorrentes e correlações temporais;
- Estimar como valores anteriores influenciam os futuros;
- Incorporar sazonalidade e tendência de forma estruturada.

Segundo Hyndman & Athanasopoulos (2021), os modelos SARIMA são eficazes porque conseguem representar tanto o curto como o longo prazo de forma linear, desde que a série seja estacionária após diferenciação.

Os modelos SARIMA são adequados para séries temporais com padrões estruturados, pois conseguem capturar dependências lineares sazonais e não sazonais (Hyndman & Athanasopoulos, 2021, Cap. 10).

A normalidade dos resíduos é razoável para ambos, embora os Q-Q plots revelem algumas pequenas anomalias nas caudas.

Constatou-se uma subestimação do viés (valor negativo) em ambos os modelos SARIMA, indicando que as previsões tendem a situar-se sistematicamente abaixo dos valores observados, aspeto que deve ser considerado em contextos onde a precisão das estimativas

é fundamental. No entanto os valores que se obtiveram em ambos os casos não parecem ser significativos, indicando apenas um viés ligeiro.

Para investigações futuras, sugere-se a exploração de diferentes combinações de ordens SARIMA, recorrendo a processos de seleção de modelo mais abrangentes (por exemplo, a utilização da função *auto\_arima* do Phyton), ou a consideração de modelos SARIMAX, com a inclusão de variáveis exógenas sempre que existam fatores externos que influenciem a série (Hyndman & Athanasopoulos, 2018). Caso o viés sistemático negativo seja considerado problemático, uma possibilidade é a aplicação da transformação Box-Cox, utilizada principalmente para estabilizar a variância e aproximar a normalidade dos resíduos. Esta transformação pode, indiretamente, contribuir para reduzir o viés se este estiver associado a heterocedasticidade (variância não constante). Em alternativa, pode ser considerado um ajuste post-hoc, dado que, em alguns casos, um pequeno ajuste linear aplicado às previsões corrige desvios sistemáticos (Hyndman, 2017).

O teste de Diebold-Mariano (DM) é utilizado para avaliar se as diferenças no desempenho preditivo entre dois modelos são estatisticamente significativas (Diebold & Mariano, 1995). Neste estudo, o teste foi aplicado aos dois modelos em análise, considerando como função de perda o Erro Quadrático Médio (MSE), métrica frequentemente utilizada na implementação do teste DM. A hipótese nula assume que ambos os modelos apresentam a mesma capacidade preditiva, enquanto a hipótese alternativa considera que existe uma diferença estatisticamente significativa na precisão das previsões.

Foi também aplicado o teste Diebold-Mariano, para comparação dos dois modelos, obtendo-se um *p\_value* igual a 0.0363, inferior ao nível de significância comum de 0,05. Isso significa que a probabilidade de se observar uma diferença de precisão como a encontrada, se os modelos tivessem realmente a mesma precisão, é de apenas 3,63%. Como este valor é muito baixo, o resultado é considerado estatisticamente significativo.

A hipótese nula do teste Diebold-Mariano é que os dois modelos têm a mesma precisão de previsão. Como o *p\_value* é baixo, a hipótese nula é rejeitada. Isso fornece evidência estatística forte de que um modelo é, de facto, mais preciso que o outro.

O valor da Estatística DM é de -2.176, em que um valor negativo significa que o primeiro modelo (modelo geral) na comparação tem uma precisão de previsão superior à do segundo modelo (modelo simplificado). A magnitude do valor (-2.176) sugere uma diferença robusta.

O modelo SARIMA Geral oferece um melhor desempenho de previsão e será utilizado na análise comparativa entre os cinco modelos.

### 3.3.1.2. Holt-Winters - Comparação entre os Modelos HWA e HWM

Nesta secção apresenta-se uma análise comparativa entre duas variantes do modelo Holt-Winters aplicadas à série temporal mensal: uma versão com tendência e sazonalidade aditivas (modelo HWA) e outra com sazonalidade multiplicativa (modelo HWM).

A primeira variante considera componentes aditivas para a tendência e para a sazonalidade, enquanto a segunda aplica a transformação Box-Cox e modela a sazonalidade de forma multiplicativa.

Tabela 3 – Quadro comparativo dos modelos Holt-Winters

| Métrica | HWA       | HWM       | Melhoria (%) | Observações                    |
|---------|-----------|-----------|--------------|--------------------------------|
| MAE     | 229.65    | 229.33    | ↓ 0.14%      | Ligeira melhoria do modelo HWM |
| MSE     | 74 564.18 | 74 291.67 | ↓ 0.37%      | Ligeira melhoria do modelo HWM |
| RMSE    | 273.06    | 272.56    | ↓ 0.18%      | Ligeira melhoria do modelo HWM |
| MAPE    | 6.54%     | 6.53%     | ↓ 0.15%      | Ligeira melhoria do modelo HWM |
| MdAE    | 272.53    | 271.82    | ↓ 0.26%      | Ligeira melhoria do modelo HWM |
| MdAPE   | 7.61%     | 7.76%     | ↑ 1.97%      | Ligeira melhoria do modelo HWA |
| MASE    | 1.00      | 1.00      | —            | Desempenho idêntico            |
| Viés    | -216.81   | -216.99   | ↑ 0.08%      | Ligeira melhoria do modelo HWA |

|                     |                |                |          |                                                   |
|---------------------|----------------|----------------|----------|---------------------------------------------------|
| sMAPE               | NaN            | NaN            | —        | Métrica não calculada                             |
| MSIS                | NaN            | NaN            | —        | Métrica não calculada                             |
| RSS                 | NaN            | NaN            | —        | Métrica não calculada                             |
| AIC                 | 399.49         | 399.52         | ↑ 0.005% | Ligeira melhoria do modelo HWA                    |
| AICc                | 439.73         | 439.75         | ↑ 0.005% | Ligeira melhoria do modelo HWA                    |
| BIC                 | 424.83         | 424.85         | ↑ 0.005% | Ligeira melhoria do modelo HWA                    |
| Log-Likelihood      | -234.84        | -234.85        | ↑ 0.004% | Ligeira melhoria do modelo HWA                    |
| JB                  | p-value > 0.05 | p-value > 0.05 | —        | Ambos com resíduos aproximadamente normais        |
| LB (Lags de 1 a 12) | p-value > 0.05 | p-value > 0.05 | —        | Ambos os modelos sem autocorrelação significativa |

Para visualizar toda a análise que foi efetuada deve-se consultar os resultados incluídos nos apêndices A e B conforme indicado na Tabela 4.

Tabela 4 – Quadro resumo com referência às Tabelas e Figuras

| Elemento                                              | Identificador        | Descrição                                                                    |
|-------------------------------------------------------|----------------------|------------------------------------------------------------------------------|
| Parâmetros e resultados do modelo HWA e do modelo HWM | Tabela A. 9 e A. 11  | Resumo dos resultados do modelo HWA e do modelo HWM                          |
| Validação Walk-Forward do modelo HWA e do modelo HWM  | Tabela A. 10 e A. 12 | Valores das métricas de validação Walk-Forward do modelo HWA e do modelo HWM |
| Informação geral entre o modelo HWA e do modelo HWM   | Tabela A. 13         | Comparação da informação geral entre o modelo HWA e do modelo HWM            |
| P_values do Teste LB dos modelos Holt-Winters         | Tabela A. 14         | Comparação dos resultados do Teste de Ljung-Box entre os modelos HWA e HWM   |
| Resíduos do modelo HWA e do modelo HWM                | Figura B.13 e B. 19  | Gráfico de resíduos do modelo HWA e do modelo HWM                            |

|                                                                    |                     |                                                                                        |
|--------------------------------------------------------------------|---------------------|----------------------------------------------------------------------------------------|
| Histograma e KDE dos resíduos do modelo HWA e do modelo HWM        | Figura B.14 e B. 20 | Distribuição dos resíduos do modelo HWA e do modelo HWM                                |
| Probabilidade normal do modelo HWA e do modelo HWM                 | Figura B.15 e B. 21 | Probabilidade normal (Q-Q plot) do modelo HWA e do modelo HWM                          |
| ACF dos resíduos do modelo HWA e do modelo HWM                     | Figura B.16 e B. 22 | Correlograma dos Resíduos do Modelo HWA e do modelo HWM                                |
| Cronograma da validação Walk-Forward do modelo HWA e do modelo HWM | Figura B.17 e B. 23 | Validação Walk-Forward do modelo HWA e do modelo HWM                                   |
| Cronograma da série temporal do modelo HWA e do modelo HWM         | Figura B.18 e B. 24 | Cronograma da série temporal com os ajustes e a previsão do modelo HWA e do modelo HWM |

Os resultados indicam que o modelo HWM apresenta melhorias marginais nas métricas MAE, MSE, RMSE, MAPE e MdAE, mas não supera consistentemente o modelo HWA. Em termos de diagnóstico dos resíduos e penalização por complexidade, o modelo HWA mantém ligeira vantagem.

A comparação entre os modelos Holt-Winters revela um trade-off entre ajuste e desempenho preditivo. O modelo HWM apresenta melhor ajuste out-of-sample, enquanto o modelo HWA demonstra ligeira vantagem na previsão in-sample.

Com base nos resultados do teste Diebold-Mariano, constata-se que o  $p\_value$  é 0.936, superior ao nível de significância de 0,05. Isso significa que a diferença de precisão entre os dois modelos não é estatisticamente significativa.

Para a análise comparativa entre os cinco modelos será utilizado o modelo HWM.

### **3.3.1.3. N-BEATS - Comparação entre os Modelos Simples e Especializado**

Nesta secção apresenta-se uma análise comparativa entre duas variantes do modelo N-BEATS aplicadas à previsão de séries temporais: uma versão simples, utilizada como referência, e uma versão especializada, com maior capacidade de decomposição estrutural.

O objetivo é avaliar o desempenho preditivo e a adequação estatística de ambas as variantes, com base em métricas de erro, testes de diagnóstico dos resíduos e critérios de informação. A versão simples recorreu a uma arquitetura com um único *stack* do tipo *identity* (modelo simples), enquanto a versão especializada incorporou *stacks* dedicados à tendência e à sazonalidade, baseados em decomposição polinomial e harmónica (modelo especializado).

Tabela 5 - Quadro comparativo dos modelos N-BEATS

| Métrica | Simple       | Especializado | Melhorias (%) | Observações                                                           |
|---------|--------------|---------------|---------------|-----------------------------------------------------------------------|
| MAE     | 414.81       | 359.79        | ↓ 13.3%       | O modelo N-BEATS Especializado é mais preciso                         |
| MSE     | 250 632.51   | 232 321.78    | ↓ 7.3%        | O modelo N-BEATS Especializado é mais preciso                         |
| RMSE    | 500.63       | 481.99        | ↓ 3.7%        | O modelo N-BEATS Especializado é mais preciso                         |
| MAPE    | 12.05%       | 10.45%        | ↓ 13.3%       | O modelo N-BEATS Especializado é mais preciso                         |
| MdAE    | 270.06       | 271.28        | ↑ 0.5%        | O modelo N-BEATS Simple é mais preciso                                |
| MdAPE   | 8.01%        | 7.77%         | ↓ 3.0%        | O modelo N-BEATS Especializado é mais preciso                         |
| MASE    | 0.88         | 0.76          | ↓ 13.6%       | O modelo N-BEATS Especializado é mais preciso                         |
| Viés    | -381.77      | -334.44       | ↓ 12.4%       | O modelo N-BEATS Especializado é mais preciso                         |
| sMAPE   | 13.16%       | 11.35%        | ↓ 13.7%       | O modelo N-BEATS Especializado é mais preciso                         |
| MSIS    | NaN          | NaN           | —             | Métrica não calculada                                                 |
| RSS     | 3 007 590.00 | 2 787 861.00  | ↓ 7.3%        | O modelo N-BEATS Especializado é mais preciso                         |
| AIC     | 19 479.23    | 98 686.32     | ↑ 406.5%      | O modelo N-BEATS Simple é favorecido pela penalização de complexidade |
| AICc    | NaN          | NaN           | —             | Não disponível.                                                       |

|                |                |                       |          |                                                                                                  |
|----------------|----------------|-----------------------|----------|--------------------------------------------------------------------------------------------------|
| BIC            | 24 157.61      | 122 568.94            | ↑ 407.5% | O modelo N-BEATS Simples é preferido por penalizar fortemente o número de parâmetros             |
| Log-Likelihood | -94.72         | -91.41                | ↓ 3.5%   | O modelo N-BEATS Especializado com verossimilhança ligeiramente superior                         |
| JB             | p-value > 0.05 | p-value > 0.05        | —        | Ambos modelos com resíduos aproximadamente normais.                                              |
| LB (Lags 1–12) | p-value > 0.05 | Alguns p-value < 0.05 | —        | O modelo N-BEATS Simples sem autocorrelação e o modelo N-BEATS Especializado com autocorrelação. |

O modelo Especializado apresentou melhorias na maioria das métricas de erro, relativamente às previsões out-of-sample. Contudo, os critérios de informação (AIC e BIC) aumentaram significativamente, sugerindo um possível sobre-ajustamento. Além disso, o modelo Especializado violou o pressuposto da independência dos resíduos, conforme evidenciado pelo teste de Ljung-Box.

O modelo Simples é mais robusto estatisticamente, com resíduos normais e não autocorrelacionados, apesar de erros ligeiramente superiores. A escolha entre os modelos deve considerar o equilíbrio entre precisão preditiva e robustez estatística.

Tabela 6 - Quadro resumo com referência às Tabelas e Figuras

| Elemento                                                                            | Identificador        | Descrição                                                                                     |
|-------------------------------------------------------------------------------------|----------------------|-----------------------------------------------------------------------------------------------|
| Critérios de informação do modelo N-BEATS Simples e do modelo N-BEATS Especializado | Tabela A. 15 e A. 17 | Resumo das métricas de informação do modelo N-BEATS Simples e do modelo N-BEATS Especializado |
| Métricas de erro do modelo N-BEATS Simples e do modelo N-BEATS Especializado        | Tabela A. 16 e A. 18 | Resumo das métricas de Erro do modelo N-BEATS Simples e do modelo N-BEATS Especializado       |

|                                                                                                  |                      |                                                                                                                       |
|--------------------------------------------------------------------------------------------------|----------------------|-----------------------------------------------------------------------------------------------------------------------|
| P-values do Teste LB e do Teste JB dos modelos N-BEATS                                           | Tabela A. 19         | Comparação dos resultados do Teste de Ljung-Box e do Teste Jarque-Bera entre as duas versões do modelo <i>N-BEATS</i> |
| Modelo N-BEATS Simples e modelo N-BEATS Especializado                                            | Figura B.25 e B. 32  | Modelo N-BEATS Simples e modelo N-BEATS Especializado                                                                 |
| Resíduos do modelo N-BEATS Simples e do modelo N-BEATS Especializado                             | Figura B.26 e B. 33  | Gráfico de resíduos do modelo N-BEATS Simples e do modelo N-BEATS Especializado                                       |
| Histograma e KDE dos resíduos do modelo N-BEATS Simples e do modelo N-BEATS Especializado        | Figura B.27 e B. 34  | Distribuição dos resíduos do modelo N-BEATS Simples e do modelo N-BEATS Especializado                                 |
| Probabilidade normal do modelo N-BEATS Simples e do modelo N-BEATS Especializado                 | Figura B.28 e B. 35  | Probabilidade normal (Q-Q plot) do modelo N-BEATS Simples e do modelo N-BEATS Especializado                           |
| ACF dos resíduos do modelo N-BEATS Simples e do modelo N-BEATS Especializado                     | Figura B.29 e B. 36  | Correlograma dos resíduos do modelo N-BEATS Simples e do modelo N-BEATS Especializado                                 |
| Cronograma da Validação Walk-Forward do modelo N-BEATS Simples e do modelo N-BEATS Especializado | Figura B. 30 e B. 37 | Validação Walk-Forward do modelo N-BEATS Simples e do modelo N-BEATS Especializado                                    |
| Cronograma da série temporal do modelo N-BEATS Simples e do modelo N-BEATS Especializado         | Figura B.31 e B. 38  | Cronograma da série temporal com as previsões do modelo N-BEATS Simples e do modelo N-BEATS Especializado             |

O teste de Diebold–Mariano apresentou um p-value de 0,000145, inferior a 0,05, indicando uma diferença estatisticamente significativa entre os modelos. A hipótese nula de igual precisão preditiva é, assim, rejeitada. O valor negativo e de elevada magnitude da estatística DM (−4,261) evidencia que o modelo Especializado apresenta maior precisão de previsão do que o modelo Simples.

Embora o modelo N-BEATS Especializado tenha apresentado melhor desempenho preditivo, as violações estatísticas e o aumento excessivo da complexidade comprometem a sua robustez. Para aplicações que exigem previsões fiáveis e estatisticamente válidas, o modelo Simples é preferível. Sugere-se a aplicação de técnicas de regularização ou ajuste pós-modelo para mitigar o viés e melhorar a adequação estatística do modelo Especializado.

#### **3.3.1.4. DeepAR - Comparação entre os Modelos Padrão e Otimizado**

Nesta secção apresenta-se uma análise comparativa entre duas variantes do modelo DeepAR, designadas como DeepAR Padrão e DeepAR Otimizado, com foco na avaliação do desempenho preditivo, robustez estatística e adequação dos resíduos.

Na implementação do DeepAR, foi utilizado o GluonTS que integra a biblioteca do Python. De acordo com Alexandrov et al. (2020), o GluonTS fornece uma estrutura para modelação de séries temporais probabilísticas e neuronais.

A variante DeepAR padrão foi implementada com quatro camadas LSTM, *dropout* de 0.1 e 30 épocas de treino. A variante DeepAR otimizado utilizou três camadas, *dropout* de 0.2, *early stopping* e monitorização da taxa de aprendizagem. Ambos os modelos foram treinados com `context_length = 36` e `prediction_length = 12`, em dados de frequência mensal.

O modelo otimizado evidenciou melhorias significativas nas métricas de erro, registando uma redução de 43,7% no MSE e de 24,9% no RMSE. O valor do MASE manteve-se inalterado em 0,04 para ambos os modelos. As métricas percentuais MAPE e sMAPE apresentaram valores muito próximos entre si, variando entre 0,27% e 0,29%, o que indica um elevado nível de precisão das previsões.

Ambos os modelos não apresentaram autocorrelação significativa nos resíduos (teste de Ljung-Box). Ambos os modelos apresentaram normalidade dos resíduos (teste de Jarque-Bera).

Tabela 7 - Quadro comparativo dos modelos DeepAR

| Métrica        | Padrão         | Otimizado      | Melhoria (%) | Observações                                                       |
|----------------|----------------|----------------|--------------|-------------------------------------------------------------------|
| MAE            | NaN            | NaN            | —            | Métrica não calculada.                                            |
| MSE            | 158.66         | 89.42          | ↓ 43.7%      | Modelo otimizado com erro quadrático significativamente menor.    |
| RMSE           | 12.60          | 9.46           | ↓ 24.9%      | Redução expressiva na raiz do erro quadrático médio.              |
| MAPE           | 0.28%          | 0.29%          | ↑ 3.6%       | Modelo padrão ligeiramente mais preciso em termos percentuais.    |
| MdAE           | NaN            | NaN            | —            | Métrica não calculada.                                            |
| MdAPE          | NaN            | NaN            | —            | Métrica não calculada.                                            |
| MASE           | 0.04           | 0.04           | —            | Desempenho idêntico; sem melhoria.                                |
| Viés           | NaN            | NaN            | —            | Métrica não calculada.                                            |
| sMAPE          | 0.27%          | 0.29%          | ↑ 7.4%       | Modelo padrão com erro percentual simétrico ligeiramente menor.   |
| MSIS           | 0.39           | 0.43           | ↑ 10.3%      | Modelo padrão com intervalo de previsão mais estreito.            |
| RSS            | NaN            | NaN            | —            | Métrica não calculada.                                            |
| AIC            | NaN            | NaN            | —            | Métrica não calculada.                                            |
| AICc           | NaN            | NaN            | —            | Métrica não calculada.                                            |
| BIC            | NaN            | NaN            | —            | Métrica não calculada.                                            |
| Log-Likelihood | NaN            | NaN            | —            | Métrica não calculada.                                            |
| JB             | p-value > 0.05 | p-value > 0.05 | —            | Modelo otimizado com resíduos mais próximos da normalidade.       |
| LB (Lags 1-12) | p-value > 0,05 | p-value > 0,05 | —            | Ambos modelos com resíduos independentes; desempenho equivalente. |

Tabela 8 - Quadro resumo com referência às Tabelas e Figuras

| Elemento                                          | Referência           | Descrição                                                                             |
|---------------------------------------------------|----------------------|---------------------------------------------------------------------------------------|
| Tabela das métricas da avaliação do modelo DeepAR | Tabela A. 20 e A. 23 | Resumo das métricas de avaliação do modelo DeepAR Padrão e do modelo DeepAR Otimizado |

|                                                                                       |                      |                                                                                                               |
|---------------------------------------------------------------------------------------|----------------------|---------------------------------------------------------------------------------------------------------------|
| Padrão e do modelo DeepAR Otimizado                                                   |                      |                                                                                                               |
| Tabela de Quantis do modelo DeepAR Padrão e do modelo DeepAR Otimizado                | Tabela A. 21 e A. 24 | Resumo das métricas de Quantis do modelo DeepAR Padrão e do modelo DeepAR Otimizado                           |
| Tabela de métricas de informação do modelo DeepAR Padrão e do modelo DeepAR Otimizado | Tabela A. 22 e A. 25 | Resumo das métricas de informação do modelo DeepAR Padrão e do modelo DeepAR Otimizado                        |
| P-values do Teste LB e do Teste JB dos modelos DeepAR                                 | Tabela A. 26         | Comparação dos resultados do Teste de Ljung-Box e do Teste Jarque-Bera entre as duas versões do modelo DeepAR |
| Modelo DeepAR Padrão e modelo DeepAR Otimizado                                        | Figura B.39 e B. 44  | Modelo DeepAR Padrão e modelo DeepAR Otimizado                                                                |
| Resíduos do modelo DeepAR Padrão e do modelo DeepAR Otimizado                         | Figura B.40 e B.45   | Gráfico de resíduos do modelo DeepAR Padrão e do modelo DeepAR Otimizado                                      |
| Histograma e KDE dos resíduos do modelo DeepAR Padrão e do modelo DeepAR Otimizado    | Figura B.41 e B. 46  | Distribuição dos resíduos do modelo DeepAR Padrão e do modelo DeepAR Otimizado                                |
| Probabilidade normal do modelo DeepAR Padrão e do modelo DeepAR Otimizado             | Figura B.42 e B. 47  | Probabilidade normal (Q-Q plot) do modelo DeepAR Padrão e do modelo DeepAR Otimizado                          |
| ACF dos resíduos do modelo DeepAR Padrão e do modelo DeepAR Otimizado                 | Figura B.43 e B. 48  | Correlograma dos resíduos do modelo DeepAR Padrão e do modelo DeepAR Otimizado                                |
| Cronograma da série temporal dos modelos DeepAR                                       | Figura B.49          | Cronograma da série temporal com as previsões dos modelos DeepAR                                              |

A evolução do modelo DeepAR Padrão para a versão Otimizada demonstra melhorias significativas em precisão e qualidade estatística.

A implementação das melhorias resultou num modelo mais robusto e fiável para aplicações práticas de previsão, mantendo a flexibilidade e capacidade de modelação de dependências complexas que caracterizam a arquitetura DeepAR.

Com base nos resultados do teste de Diebold-Mariano, o valor de p-value obtido foi de 0,477, superior ao nível de significância de 0,05. Este resultado indica que não se rejeita a hipótese nula de igualdade de precisão entre os modelos. O valor da estatística DM, igual a 0,736, encontra-se próximo de zero, reforçando a ausência de diferenças significativas no desempenho. Deste modo, não se observa evidência estatística que suporte a superioridade de um modelo em relação ao outro em termos de precisão preditiva.

O modelo DeepAR Otimizado demonstrou melhorias substanciais na precisão das previsões e na normalidade dos resíduos. A escolha entre os modelos deve considerar o equilíbrio entre precisão preditiva e robustez estatística.

Para análise comparativa entre os cinco modelos será utilizado o modelo DeepAR Otimizado.

#### **3.3.1.5. PROPHET - Comparação entre do Modelo Padrão e Otimizado**

Nesta secção apresenta-se uma análise comparativa entre duas variantes do modelo Prophet, designadas como Prophet Padrão e Prophet Otimizado, com foco nas métricas de desempenho, testes estatísticos e comportamento dos resíduos. O objetivo é avaliar a eficácia preditiva de ambas as variantes e identificar as suas limitações e pontos fortes.

Ambos os modelos foram aplicados a uma série temporal mensal com 36 observações. O Prophet Padrão utilizou a configuração inicial fornecida pela biblioteca, enquanto o Prophet Otimizado incorporou ajustes de hiperparâmetros (nomeadamente `changepoint_prior_scale`), deteção de outliers pelo método IQR e desativação de sazonalidades irrelevantes. A avaliação foi realizada com base em métricas como MAPE, MASE e viés, complementada pelos testes de Ljung-Box e Shapiro-Wilk, bem como pela análise visual dos resíduos.

Ambos os modelos apresentaram desempenho semelhante em termos de precisão percentual ( $\text{MAPE} \approx 2.6\%$ ) e ausência de viés sistemático ( $\text{viés} \approx 0$ ). No entanto, o MASE de 1.00 em ambos os casos, indica que nenhum modelo superou uma previsão ingénua. O modelo

Otimizado não conseguiu eliminar a autocorrelação residual, conforme evidenciado pelos testes de Ljung-Box e pelos correlogramas. A normalidade dos resíduos foi confirmada pelo teste de Jarque-Bera e pelos gráficos Q-Q plot.

Tabela 9 - Quadro comparativo dos modelos Prophet

| Métrica | Padrão     | Otimizado  | Melhoria (%) | Observações                                                            |
|---------|------------|------------|--------------|------------------------------------------------------------------------|
| MAE     | 89.11      | 89.53      | ↑ 0.46%      | Modelo Otimizado com erro absoluto ligeiramente maior                  |
| MSE     | 13 374.29  | 13 398.52  | ↑ 0.18%      | Diferença mínima; desempenho equivalente                               |
| RMSE    | 115.65     | 115.75     | ↑ 0.09%      | Modelo Otimizado com erro quadrático marginalmente superior            |
| MAPE    | 2.60%      | 2.62%      | ↑ 0.77%      | Modelo Padrão ligeiramente mais preciso em termos percentuais          |
| MdAE    | 72.14      | 75.48      | ↑ 4.61%      | Modelo Padrão com erro mediano ligeiramente menor                      |
| MdAPE   | 2.07%      | 2.17%      | ↑ 4.83%      | Modelo Padrão com menor erro percentual mediano                        |
| MASE    | 1.00       | 1.00       | —            | Desempenho idêntico; sem melhoria                                      |
| Viés    | 0.04       | -0.02      | ↓ 150.0%     | Modelo Otimizado com viés mais próximo de zero                         |
| sMAPE   | 2.59%      | 2.61%      | ↑ 0.77%      | Diferença irrelevante; desempenho equivalente                          |
| MSIS    | NaN        | NaN        | —            | Métrica não calculada                                                  |
| RSS     | 481 474.48 | 482 346.65 | ↑ 0.18%      | Modelo Padrão com soma residual dos quadrados ligeiramente inferior.   |
| AIC     | 488.22     | 488.28     | ↓ 0.01%      | Modelo Padrão ligeiramente favorecido pela penalização de complexidade |
| AICc    | 566.06     | 566.13     | ↓ 0.01%      | Diferença irrelevante; desempenho equivalente                          |
| BIC     | 523.05     | 523.12     | ↑ 0.01%      | Diferença irrelevante; desempenho equivalente                          |

|                |                |                |         |                                                          |
|----------------|----------------|----------------|---------|----------------------------------------------------------|
| Log-Likelihood | -222.11        | -222,14        | ↑ 0.01% | Modelo Padrão com verossimilhança marginalmente superior |
| JB             | p_value > 0.05 | p_value > 0.05 | —       | Ambos modelos com resíduos aproximadamente normais.      |
| LB (Lags 1-12) | p_value < 0.05 | p_value < 0.05 | —       | Ambos modelos com autocorrelação                         |

Tabela 10 - Quadro resumo com referência às Tabelas e Figuras

| Elemento                                                                                 | Referência           | Descrição                                                                                                      |
|------------------------------------------------------------------------------------------|----------------------|----------------------------------------------------------------------------------------------------------------|
| Resumo das métricas de Erro do modelo Prophet Padrão e do modelo Prophet Otimizado       | Tabela A. 27 e A. 29 | Resumo das métricas de Erro do modelo Prophet Padrão e do modelo Prophet Otimizado                             |
| Resumo das métricas de informação do modelo Prophet Padrão e do modelo Prophet Otimizado | Tabela A. 28 e A.30  | Resumo das métricas de informação do modelo Prophet Padrão e do modelo Prophet Otimizado                       |
| P-values do Teste LB e Teste JB dos modelos Prophet                                      | Tabela A. 31         | Comparação dos resultados do Teste de Ljung-Box e do Teste Jarque-Bera entre as duas versões do modelo Prophet |
| Modelo Prophet Padrão e modelo Prophet Otimizado                                         | Figura B.50 e B. 58  | Modelo Prophet Padrão e modelo Prophet Otimizado                                                               |
| Resíduos do modelo Prophet Padrão e do modelo Prophet Otimizado                          | Figura B.51 e B. 59  | Gráfico de resíduos do modelo Prophet Padrão e do modelo Prophet Otimizado                                     |
| Histograma e KDE dos resíduos do modelo Prophet Padrão e do modelo Prophet Otimizado     | Figura B.52 e B. 60  | Distribuição dos resíduos do modelo Prophet Padrão e do modelo Prophet Otimizado                               |
| Probabilidade normal do modelo Prophet Padrão e do modelo Prophet Otimizado              | Figura B.53 e B. 61  | Probabilidade normal (Q-Q plot) do modelo Prophet Padrão e do modelo Prophet Otimizado                         |
| ACF dos resíduos do modelo Prophet <i>Padrão</i> e do modelo Prophet Otimizado           | Figura B.54 e B. 62  | Correlograma dos resíduos do modelo Prophet Padrão e do modelo Prophet Otimizado                               |

|                                                                                                   |                     |                                                                                                  |
|---------------------------------------------------------------------------------------------------|---------------------|--------------------------------------------------------------------------------------------------|
| Cronograma da validação Walk-Forward do modelo Prophet Padrão e do modelo Prophet Otimizado       | Figura B.55 e B. 63 | Validação Walk-Forward do modelo Prophet Padrão e do modelo Prophet Otimizado                    |
| Cronograma da série temporal do modelo Prophet Padrão e do modelo Prophet Otimizado               | Figura B.56 e B. 64 | Série temporal com os ajustes e previsões do modelo Prophet Padrão e do modelo Prophet Otimizado |
| Cronograma de previsão de séries temporais do modelo Prophet Padrão e do modelo Prophet Otimizado | Figura B.57 e B. 65 | Previsão do modelo Prophet Padrão e do modelo Prophet Otimizado com intervalos de confiança      |

Estas figuras e tabelas suportam a análise estatística e visual da adequação dos modelos.

O modelo Prophet Otimizado não apresentou melhorias significativas em relação ao modelo Padrão. Ambos os modelos demonstraram boa precisão percentual e ausência de viés, mas falharam em eliminar a autocorrelação dos resíduos. Futuras melhorias podem incluir abordagens híbridas ou a inclusão de variáveis exógenas para captar dependências não modeladas.

Com base nos resultados do teste de Diebold-Mariano, o valor de p-value obtido foi de 0,828, superior ao nível de significância de 0,05. Este resultado indica que não se rejeita a hipótese nula de igualdade de precisão entre os modelos. O valor da estatística DM, igual a -0,219, encontra-se próximo de zero, reforçando a ausência de diferenças significativas no desempenho. Deste modo, não se observa evidência estatística que suporte a superioridade de um modelo em relação ao outro em termos de precisão preditiva.

Dado que as métricas de erro são praticamente idênticas, o modelo Otimizado não melhora significativamente a autocorrelação dos resíduos, o teste de Diebold-Mariano não encontrou diferença estatisticamente significativa entre os dois modelos e o modelo Padrão é mais simples e computacionalmente mais eficiente será o modelo selecionada, com base no princípio da parcimônia (Princípio de Occam's Razor -quando dois modelos têm desempenho semelhante, deve-se preferir o mais simples).

### 3.3.2. Análise comparativa dos resultados dos modelos selecionados

A seleção do modelo de previsão mais adequado para uma série temporal é um processo complexo que exige uma análise rigorosa de um conjunto diversificado de métricas de desempenho, que abrangem tanto a precisão das previsões pontuais quanto a qualidade da quantificação da incerteza (Makridakis et al., 2020).

Nesta secção realizou-se a análise comparativa entre as cinco versões dos modelos de previsão selecionados. A avaliação baseou-se nas métricas de erro como o MAE (Erro Absoluto Médio), RMSE (Raiz do Erro Quadrático Médio) e MAPE (Erro Percentual Absoluto Médio), bem como nos critérios de informação e testes de diagnóstico estatístico. Esta abordagem permitiu não só quantificar a precisão dos modelos, mas também avaliar a adequação estatística dos seus resíduos, de forma a contribuir para uma escolha fundamentada do modelo mais robusto.

Na secção seguinte apresenta-se a tabela resumo com os principais indicadores de desempenho dos modelos analisados.

#### 3.3.2.1. Análise comparativa dos resultados dos modelos selecionados

Tabela 11 - Quadro comparativo dos modelos analisados

| Modelo | SARIMA     | Holt-Winters | N-BEATS      | DeepAR    | Prophet    |
|--------|------------|--------------|--------------|-----------|------------|
| Versão | Geral      | HWM          | Simples      | Otimizado | Otimizado  |
| MAE    | 187.93     | 229.33       | 414.81       | NaN       | 89.53      |
| MSE    | 47 566.91  | 74 291.67    | 250 632.51   | 89.42     | 13 398.52  |
| RMSE   | 218.09     | 272.56       | 500.63       | 9.46      | 115.75     |
| MAPE   | 5.56%      | 6.53%        | 12.05%       | 0.29%     | 2.62%      |
| MdAE   | 175.55     | 271.82       | 270.06       | NaN       | 75.48      |
| MdAPE  | 5.33%      | 7.76%        | 8.01%        | NaN       | 2.17%      |
| Viés   | -19.57     | -216.99      | -381.77      | NaN       | -0.02      |
| MASE   | NaN        | 1.00         | 0.88         | 0.04      | 1.00       |
| sMAPE  | 5.60%      | NaN          | 13.16%       | 0.29%     | 2.61%      |
| MSIS   | NaN        | NaN          | NaN          | 0.43      | NaN        |
| RSS    | 570 802.93 | NaN          | 3.007590e+06 | NaN       | 482 346.65 |

|                         |             |             |             |             |                 |
|-------------------------|-------------|-------------|-------------|-------------|-----------------|
| <b>AIC</b>              | 331.99      | 399.52      | 19 479.24   | NaN         | 488.28          |
| <b>AICC</b>             | NaN         | 439.75      | NaN         | NaN         | 566.13          |
| <b>BIC</b>              | 338.81      | 424.85      | 24 157.62   | NaN         | 523.12          |
| <b>Log_Likelihood</b>   | -160.00     | -234.85     | -94.72      | NaN         | -222.14         |
| <b>JB</b>               | >0.05       | >0.05       | >0.05       | >0.05       | >0.05           |
| <b>LB (Lag 1-12)</b>    | > 0.05      | > 0.05      | > 0.05      | >0.05       | < 0.05          |
| <b>Histograma</b>       | Normal.     | Normal.     | Normal.     | Normal.     | Normal.         |
| <b>KDE</b>              | Normal.     | Normal.     | Normal.     | Normal.     | Normal.         |
| <b>QQ-Plot</b>          | Normal.     | Normal.     | Normal.     | Normal.     | Normal.         |
| <b>Correlograma ACF</b> | S/ Autocor. | S/ Autocor. | S/ Autocor. | S/ Autocor. | Autocorrelação. |

### **Análise Comparativa de Modelos de Previsão: Interpretação e Seleção do Melhor Modelo**

A análise comparativa de modelos de previsão é essencial para identificar a abordagem mais adequada para um determinado conjunto de dados. Neste estudo, foram avaliados cinco modelos: SARIMA, Holt-Winters, N-BEATS, DeepAR e Prophet, com base em métricas de avaliação e análise de resíduos. A Tabela 11 resume os resultados obtidos, permitindo uma comparação direta entre os modelos.

O modelo Prophet destacou-se como o mais robusto em termos de desempenho preditivo. Apresentou os menores MAE (89.53), MdAE (75.48), MdAPE (2.17%) e viés (-0.02), o que indica uma elevada precisão nas previsões. Além disso, os testes de normalidade (histograma, KDE e QQ-Plot) sugerem que os resíduos seguem uma distribuição aproximadamente normal, e o correlograma ACF indica ausência de Autocorrelação, o que reforça a validade do modelo. Este resultado está em consonância com a literatura, que reconhece a capacidade do Prophet para modelar dados com sazonalidades múltiplas e tendências não lineares (Taylor & Letham, 2018).

O modelo DeepAR, embora tenha apresentado os menores RMSE (9,46) e MAPE (0,29%), possui valores ausentes em várias métricas fundamentais (como MAE, MASE e testes de diagnóstico), o que compromete a sua avaliação completa.

O modelo N-BEATS, apesar de ser uma abordagem baseada em redes neurais profundas, apresentou genericamente desempenho inferior ao Prophet. No entanto, os resíduos demonstraram normalidade e ausência de Autocorrelação.

Os modelos SARIMA e Holt-Winters apresentaram desempenhos mais modestos comparativamente ao Prophet, com valores de erro, na generalidade, mais elevados.

O AIC e o BIC são critérios importantes para avaliar a parcimónia dos modelos. O Prophet apresentou valores relativamente baixos de AIC (488,28) e BIC (523,12), o que, aliado ao seu bom desempenho preditivo, reforça a sua adequação.

A escolha do modelo ideal depende, em grande parte, das prioridades específicas do problema de previsão.

Se o principal objetivo for a minimização dos erros quadráticos e padronizados (RMSE, MSE, MAPE, MASE) e a qualidade das previsões probabilísticas, avaliada através do MSIS, for considerada crucial, o modelo DeepAR (Otimizado) revela-se o mais indicado. A sua performance nestas métricas é consistentemente superior, demonstrando elevada precisão preditiva. Além disso, a capacidade de fornecer intervalos de previsão fiáveis constitui um diferencial relevante para a gestão do risco.

Apesar da ausência de algumas métricas baseadas em erro absoluto (MAE, MdAE, Viés), que limita uma avaliação totalmente abrangente, o desempenho evidenciado nas métricas disponíveis é notável, assumindo que a omissão dessas medidas não reflete falhas de modelação.

Se a prioridade for a minimização de erros absolutos e medianos (MAE, MdAE, MdAPE) e a ausência de viés sistemático (Bias), o Prophet (Otimizado) é uma alternativa robusta e fiável. A sua consistência nestas métricas, aliada à sua estrutura interpretável, torna-o uma boa escolha, especialmente em cenários onde a robustez a outliers é importante. No entanto, o seu MASE de 1.00 sugere que não oferece uma melhoria significativa em relação a uma previsão ingénua em termos de erro padronizado.

O Prophet revela-se a escolha mais adequada, especialmente para séries que apresentam sazonalidade e tendência (Taylor & Letham, 2018). Para situações em que se privilegia a

interpretação do modelo, o SARIMA constitui uma alternativa apropriada, em conformidade com as recomendações de Hyndman & Athanasopoulos (2018) para séries estacionárias. Em trabalhos futuros, poderá ser interessante explorar técnicas de transformação de dados (por exemplo, a transformação Box-Cox) com o objetivo de melhorar a normalidade dos resíduos do modelo DeepAR. Poderá também ser analisada a combinação de alguns destes modelos ou a aplicação de técnicas de ensemble (por exemplo, SARIMA + Prophet) de forma a melhorar ainda mais a precisão preditiva (Hyndman & Athanasopoulos, 2018; Taylor & Letham, 2018).

Esta análise reforça a importância de selecionar modelos com base não apenas em métricas de erro, mas também em propriedades estatísticas dos resíduos e requisitos do problema em questão.

### 3.3.3. Análise comparativa dos resultados dos modelos selecionados com base nas previsões out-of-sample

Nesta secção, os modelos foram treinados com uma janela de entrada de 24 observações (2022 e 2023), e utilizado um horizonte de previsão de 12 períodos correspondente ao ano de 2024.

As previsões obtidas para 2024 encontram-se na Tabela 12.

Tabela 12 – Quadro com os dados reais e as previsões para o ano 2024 de cada versão selecionada dos modelos

| AnoMes  | Real    | SARIMA  | Holt-Winters | N-BEATS | DeepAR    | Prophet   |
|---------|---------|---------|--------------|---------|-----------|-----------|
|         |         | Geral   | HWM          | Padrão  | Otimizado | Otimizado |
| 2024-01 | 3802.37 | 3787.57 | 3723.16      | 3291.56 | 3427.45   | 3800.93   |
| 2024-02 | 3477.98 | 3301.92 | 3418.88      | 3089.84 | 3519.71   | 3458.79   |
| 2024-03 | 3374.32 | 3386.95 | 3419.14      | 3251.16 | 3584.59   | 3267.22   |
| 2024-04 | 3523.38 | 3489.38 | 3288.36      | 3379.25 | 3516.03   | 3478.88   |
| 2024-05 | 3252.28 | 3157.04 | 3474.77      | 3223.40 | 3134.37   | 3242.21   |
| 2024-06 | 2990.17 | 2958.45 | 3280.85      | 3262.06 | 3399.84   | 2941.50   |
| 2024-07 | 3631.67 | 3587.65 | 3587.93      | 3179.04 | 3149.31   | 3602.45   |
| 2024-08 | 3256.17 | 3189.27 | 3510.83      | 3008.82 | 3413.78   | 3327.94   |
| 2024-09 | 3326.12 | 3280.44 | 3375.04      | 3150.49 | 3532.29   | 3436.66   |
| 2024-10 | 3500.70 | 3452.10 | 3375.19      | 3122.10 | 3590.81   | 3453.13   |
| 2024-11 | 3080.60 | 3023.57 | 3324.21      | 3014.23 | 3529.34   | 3086.11   |
| 2024-12 | 3215.52 | 3165.52 | 3165.53      | 3220.93 | 3143.90   | 3200.63   |

Foi realizado o cálculo das métricas MAE, MSE, RMSE e MAPE e o teste Diebold-Mariano no conjunto dos vários modelos com base nas previsões do ano de 2024 obtendo-se os resultados que se encontram na Tabela 13.

Tabela 13 - Métricas de erro por modelo com base nas previsões do ano de 2024

| Modelo      | SARIMA  | Holt-Winters | N-BEATS  | DeepAR    | Prophet   |
|-------------|---------|--------------|----------|-----------|-----------|
| Versão      | Geral   | HWM          | Simples  | Otimizado | Otimizado |
| <b>MAE</b>  | 56.39   | 141.48       | 232.75   | 218.20    | 42.54     |
| <b>MSE</b>  | 4935.19 | 28980.51     | 80578.98 | 73578.28  | 3085.84   |
| <b>RMSE</b> | 70.25   | 170.24       | 283.86   | 271.25    | 55.55     |
| <b>MAPE</b> | 1.68%   | 4.32%        | 6.73%    | 6.54%     | 1.27%     |

O teste de Diebold-Mariano (DM) é utilizado para avaliar se as diferenças entre as previsões de dois modelos são estatisticamente significativas (Diebold & Mariano, 1995). Este teste foi aplicado às versões selecionadas dos modelos analisados (SARIMA, Holt-Winters, N-BEATS, DeepAR e Prophet), comparando-os em pares com base no MSE, uma métrica frequentemente utilizada no teste de DM.

Tabela 14 - Teste Diebold-Mariano para comparações em pares com base no MSE

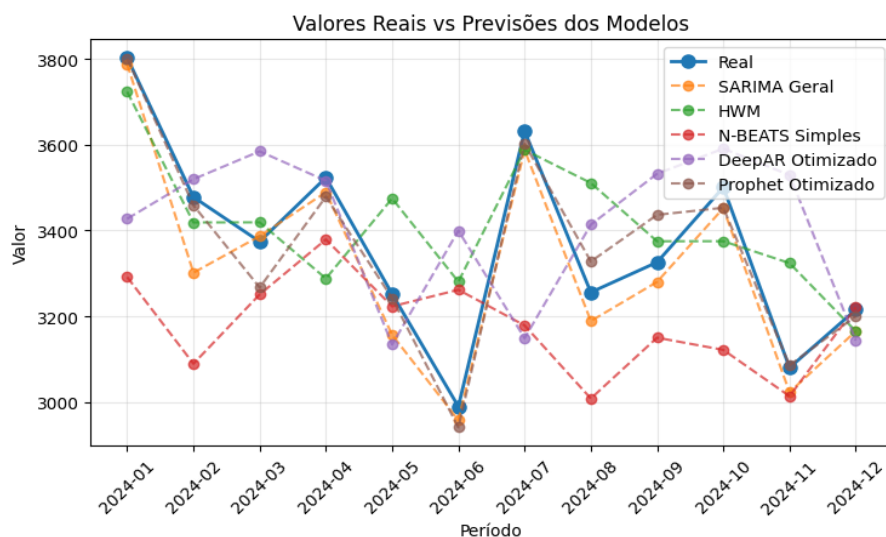
| Comparação                        | Estatística DM | p_value | Conclusão ( $\alpha=0.05$ )                   |
|-----------------------------------|----------------|---------|-----------------------------------------------|
| SARIMA Geral vs HWM               | -2.48          | 0.013   | Rejeita $H_0$ (com diferenças significativas) |
| SARIMA Geral vs N-BEATS Simples   | -2.96          | 0.003   | Rejeita $H_0$ (com diferenças significativas) |
| SARIMA Geral vs DeepAR Otimizado  | -2.65          | 0.008   | Rejeita $H_0$ (com diferenças significativas) |
| SARIMA Geral vs Prophet Otimizado | 0.61           | 0.544   | Não rejeita $H_0$ (sem diferença)             |
| HWM vs N-BEATS Simples            | -1.71          | 0.087   | Não rejeita $H_0$ (sem diferença)             |
| HWM vs DeepAR Otimizado           | -1.77          | 0.077   | Não rejeita $H_0$ (sem diferença)             |

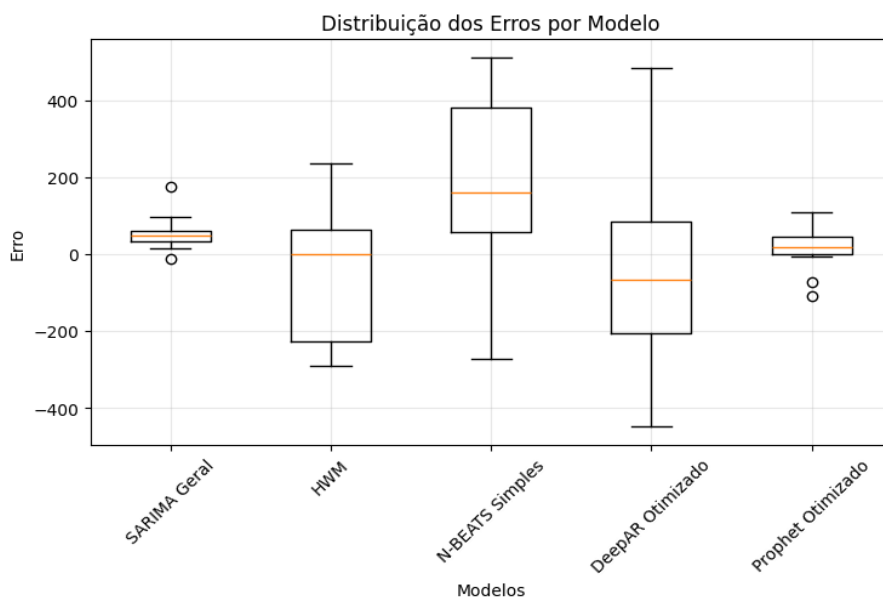
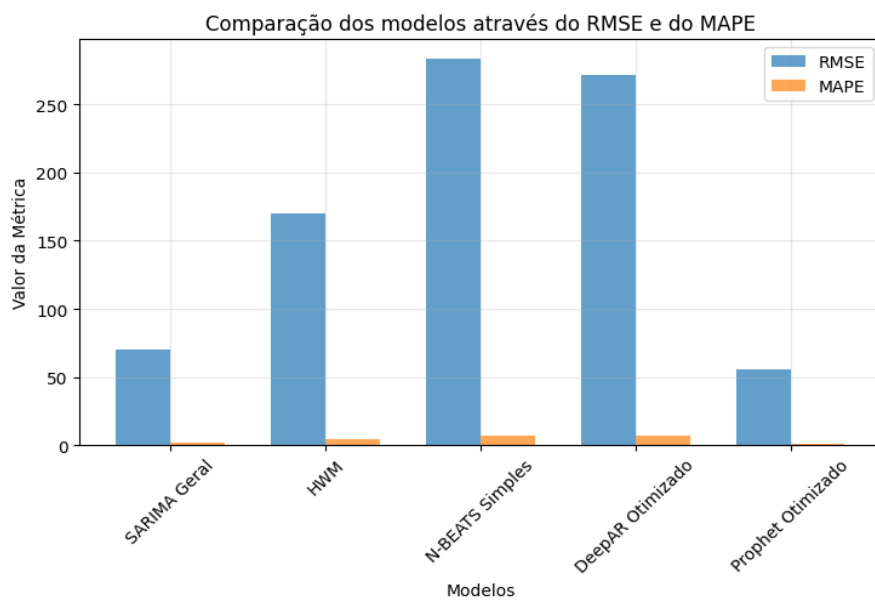
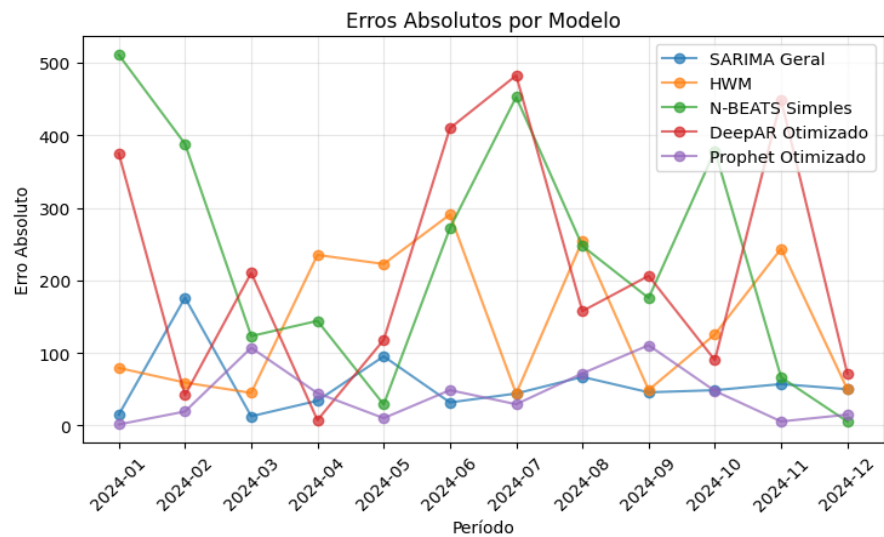
|                                       |      |       |                                               |
|---------------------------------------|------|-------|-----------------------------------------------|
| HWM vs Prophet Otimizado              | 2.78 | 0.006 | Rejeita $H_0$ (com diferenças significativas) |
| N-BEATS Simples vs DeepAR Otimizado   | 0.25 | 0.805 | Não rejeita $H_0$ (sem diferença)             |
| N-BEATS Simples vs Prophet Otimizado  | 2.98 | 0.003 | Rejeita $H_0$ (com diferenças significativas) |
| DeepAR Otimizado vs Prophet Otimizado | 2.79 | 0.005 | Rejeita $H_0$ (com diferenças significativas) |

Tabela 15 – Ranking dos modelos com base no MSE e no MAPE

| Ranking dos modelos (RMSE) |                                 | Ranking dos modelos (MAPE) |                                 |
|----------------------------|---------------------------------|----------------------------|---------------------------------|
| 1                          | Prophet Otimizado (RMSE: 55.55) | 1                          | Prophet Otimizado (MAPE: 1.27%) |
| 2                          | SARIMA Geral (RMSE: 70.25)      | 2                          | SARIMA Geral (MAPE: 1.68%)      |
| 3                          | HWM (RMSE: 170.24)              | 3                          | HWM (MAPE: 4.32%)               |
| 4                          | DeepAR Otimizado (RMSE: 271.25) | 4                          | DeepAR Otimizado (MAPE: 6.54%)  |
| 5                          | N-BEATS Simples (RMSE: 283.86)  | 5                          | N-BEATS Simples (MAPE: 6.73%)   |

Figura 4 – Gráficos comparativos dos modelos





Os resultados revelaram diferenças estatisticamente significativas entre o modelo **Prophet Otimizado** e os restantes modelos analisados. Em particular, as comparações entre Prophet e SARIMA Geral ( $p = 0,0038$ ), Prophet e Holt-Winters Multiplicativo ( $p = 0,0055$ ), Prophet e N-BEATS Simples ( $p = 0,0029$ ) e Prophet e DeepAR Otimizado ( $p = 0,0053$ ) confirmam a superioridade estatística do Prophet em termos de desempenho preditivo. Adicionalmente, a comparação entre SARIMA Geral e Holt-Winters ( $p = 0,0380$ ) revelou também diferenças significativas, favorecendo o primeiro.

Estes resultados reforçam a robustez do Prophet como modelo preditivo, corroborada igualmente pelas métricas de erro (MAE, RMSE, MAPE), onde apresentou os valores mais baixos no horizonte de previsão de 2024. Contudo, importa salientar que, dado o número reduzido de observações ( $n = 12$ ), o poder estatístico do teste pode ser limitado, pelo que as conclusões devem ser interpretadas com cautela (Makridakis et al., 2020).

### 3.4. Discussão de Resultados

A análise comparativa dos modelos de previsão aplicados neste estudo permitiu identificar diferenças no desempenho preditivo entre abordagens estatísticas tradicionais e modelos baseados em aprendizagem profunda. A Tabela 11 (Quadro comparativo dos modelos analisados, pág. 44) apresenta um conjunto abrangente de métricas de avaliação, permitindo uma análise multidimensional da eficácia de cada modelo.

O **modelo Prophet Otimizado** destacou-se como o mais eficaz, apresentando os melhores resultados na generalidade das métricas de erro, tais como: MAE = 89.53, RMSE = 115.75 e MAPE = 2.62%. Estes valores confirmam a robustez do Prophet na modelação de séries temporais com tendência e sazonalidade, conforme evidenciado por Taylor & Letham (2018). A capacidade do modelo para gerar previsões precisas, aliada a resíduos aproximadamente normais e ausência de autocorrelação (verificada através do correlograma - ACF), reforça a sua adequação em contextos empresariais.

O **modelo SARIMA Geral** apresentou o segundo melhor desempenho, com MAE = 187.93, RMSE = 218.09 e MAPE = 5.56%. Embora inferior ao Prophet, manteve resultados bastante competitivos, confirmando a relevância dos modelos estatísticos tradicionais em cenários de previsão com séries temporais relativamente curtas.

O **modelo Holt-Winters (HWM)** apresentou um desempenho intermédio, com MAE = 229.33, RMSE = 272.56 e MAPE = 6.53%. Apesar de apresentar erros mais elevados, manteve previsões consistentes e estáveis, sendo uma alternativa viável em contextos que privilegiam a simplicidade, a interpretabilidade e o baixo custo computacional (Hyndman & Athanasopoulos, 2018).

Os modelos de **aprendizagem profunda** apresentaram desempenhos inferiores aos métodos estatísticos. O **N-BEATS Simples** apresentou os piores resultados globais, com MAE = 414.81, RMSE = 500.63 e MAPE = 12,05%. Estes resultados sugerem que, neste caso, a complexidade adicional dos modelos de deep learning não se traduziu em ganhos de precisão, possivelmente devido ao reduzido número de observações disponíveis, em linha com as conclusões de Oreshkin et al. (2020). Enquanto o **DeepAR Otimizado**, embora tenha apresentado os menores RMSE (9,46) e MAPE (0,29%), possui valores ausentes em várias métricas fundamentais (como MAE, MASE e testes de diagnóstico), o que compromete a sua avaliação completa.

A partir deste parágrafo, a análise e discussão dos resultados concentram-se nos dados referentes ao ano de 2024, com o objetivo de avaliar o desempenho preditivo dos modelos face aos valores efetivamente observados nesse período.

Com base nos resultados apresentados na Tabela 13 (Métricas de erro por modelo com base nas previsões do ano de 2024, pág. 48), é possível comparar o desempenho dos diferentes modelos de previsão de forma quantitativa e interpretativa, com base nas principais métricas de erro (MAE, MSE, RMSE e MAPE).

O **modelo Prophet Otimizado** destacou-se de forma expressiva como o mais preciso e consistente entre todas as abordagens avaliadas, apresentando MAE = 42.54, MSE = 3085.84, RMSE = 55.55 e MAPE = 1.27%. Estes valores evidenciam uma capacidade de previsão robusta, com erros médios e percentuais significativamente inferiores aos dos restantes modelos.

O **modelo SARIMA Geral** apresentou o segundo melhor desempenho, com MAE = 56.39, MSE = 4935.19, RMSE = 70.25 e MAPE = 1.68%, confirmando a relevância dos modelos estatísticos tradicionais em contextos de previsão com séries temporais relativamente curtas.

O **modelo Holt-Winters (HWM)** registou um desempenho intermédio, com MAE = 141.48, MSE = 28980.51, RMSE = 170.24 e MAPE = 4.32%. Apesar de apresentar erros mais elevados, manteve previsões consistentes e estáveis, sendo uma alternativa fiável em cenários que privilegiam a simplicidade, a interpretabilidade e o baixo custo computacional.

Os modelos de **aprendizagem profunda** apresentaram desempenhos inferiores aos métodos estatísticos. O **DeepAR Otimizado** obteve MAE = 218,20, MSE = 73578.28, RMSE = 271.25 e MAPE = 6.54%, enquanto o **N-BEATS Simples** apresentou os piores resultados, com MAE = 232.75, MSE = 80578.98, RMSE = 283.86 e MAPE = 6.73%. Estes resultados sugerem que, neste caso, a complexidade adicional dos modelos de deep learning não se traduziu em ganhos de precisão, possivelmente devido ao reduzido número de observações disponíveis.

O **teste de Diebold-Mariano** (ver Tabela 14 , pág. 48) foi fundamental para validar estatisticamente as diferenças de desempenho entre os modelos com base no MSE. Os resultados indicam que o **SARIMA Geral** apresenta desempenho significativamente superior ao **HWM** ( $p = 0,013$ ), ao **N-BEATS Simples** ( $p = 0,003$ ) e ao **DeepAR Otimizado** ( $p = 0,008$ ). Contudo, não foram encontradas diferenças estatisticamente significativas entre o **SARIMA Geral** e o **Prophet Otimizado** ( $p = 0,544$ ), sugerindo desempenhos equivalentes entre estas duas abordagens. O **HWM** não apresentou diferenças significativas face ao **N-BEATS Simples** ( $p = 0,087$ ) e ao **DeepAR Otimizado** ( $p = 0,077$ ), mas foi estatisticamente inferior ao **Prophet Otimizado** ( $p = 0,006$ ).

No grupo dos modelos de aprendizagem profunda, não se verificaram diferenças significativas entre o **N-BEATS Simples** e o **DeepAR Otimizado** ( $p = 0,805$ ). No entanto, ambos foram significativamente inferiores ao **Prophet Otimizado** ( $p = 0,003$  e  $p = 0,005$ , respetivamente).

Em síntese, a análise dos resultados de 2024 do teste DM reforçam a robustez do **Prophet Otimizado**, que se destacou face ao **HWM**, **N-BEATS Simples** e **DeepAR Otimizado**, mantendo desempenho equivalente ao **SARIMA Geral**. Por sua vez, o **SARIMA Geral** também demonstrou superioridade em relação ao **HWM** e aos modelos de deep learning, consolidando-se como uma alternativa estatística competitiva.

## Conclusão e Investigação Futura

O presente capítulo constitui o encerramento do trabalho desenvolvido no âmbito da previsão de séries temporais aplicadas a dados operacionais da CIBEN, sintetizando os principais resultados obtidos e refletindo sobre a aplicabilidade dos modelos de previsão de séries temporais no contexto organizacional. O capítulo encontra-se dividido em dois subcapítulos: o subcapítulo 4.1 – Conclusão, que destaca os principais resultados, as suas implicações e a forma como estes se relacionam com os objetivos definidos; e o subcapítulo 4.2 – Investigações Futuras, onde são apresentadas sugestões estudos complementares e possíveis desenvolvimentos com avista a potenciar os resultados alcançados.

### Conclusão

Este estudo teve como objetivo principal avaliar e comparar diferentes modelos de previsão de séries temporais, com vista à identificação da abordagem mais eficaz para antecipar padrões de atividade numa organização. Através da análise de dados operacionais recolhidos entre 2012 e 2024, foi possível aplicar e testar cinco modelos distintos: SARIMA, Holt-Winters, N-BEATS, DeepAR e Prophet.

A operacionalização do estudo incluiu a preparação dos dados, a aplicação de técnicas de validação cruzada temporal (walk-forward), e a avaliação dos modelos com base em diversas métricas de erro (MAE, RMSE, MAPE, entre outras), complementadas por testes estatísticos de diagnóstico (Ljung-Box, Jarque-Bera, entre outros). Esta abordagem metodológica permitiu uma análise rigorosa e comparativa do desempenho de cada modelo.

A investigação permitiu atingir os objetivos propostos, nomeadamente a comparação do desempenho entre diferentes modelos de previsão e a identificação da abordagem mais adequada aos dados em análise.

A metodologia CRISP-DM, revelou-se adequada para estruturar o processo analítico e garantir a coerência e a fiabilidade dos resultados.

O **modelo Prophet Otimizado** revelou-se o mais eficaz, sendo particularmente recomendado para aplicações empresariais que exijam previsões precisas, interpretáveis e de fácil implementação.

A análise estatística, complementada pelo teste de Diebold-Mariano, permitiu validar as diferenças de desempenho entre os modelos, reforçando a robustez das conclusões obtidas. Este procedimento confirma a importância de complementar as métricas de erro com testes formais de comparação, conforme recomendado por Diebold & Mariano (1995).

Em síntese, este trabalho contribuiu para uma compreensão aprofundada das potencialidades e limitações dos modelos de previsão de séries temporais, fornecendo uma base sólida para a sua aplicação prática em contextos organizacionais. Os resultados obtidos evidenciam a importância da modelação preditiva como ferramenta de apoio à decisão estratégica, permitindo antecipar padrões de atividade e otimizar a alocação de recursos de forma mais eficiente.

### **Investigações Futuras**

Apesar dos resultados promissores, este estudo apresenta algumas limitações que deverão ser consideradas em futuras investigações. Em primeiro lugar, o número de observações (36 meses) utilizadas para treino e validação dos modelos é relativamente reduzido, especialmente para modelos de aprendizagem profunda, que requerem volumes de dados maiores. A ampliação da série temporal ou a integração de múltiplas séries poderá melhorar a robustez dos resultados.

Em segundo lugar, a análise centrou-se exclusivamente em séries univariadas. A inclusão de variáveis exógenas (por exemplo: indicadores económicos, feriados, eventos internos, entre outras) poderá enriquecer os modelos e aumentar a sua capacidade explicativa, nomeadamente através da utilização de modelos SARIMAX ou Prophet com regressão.

Para investigações futuras, recomenda-se a ampliação da base de dados, a inclusão de variáveis externas relevantes, a exploração de modelos híbridos que combinem abordagens estatísticas e de aprendizagem automática, e a aplicação dos modelos em diferentes contextos organizacionais com vista à avaliação da sua generalização. Adicionalmente, a utilização de técnicas de ensemble e a análise de séries multivariadas poderão contribuir para melhorar a precisão e a robustez das previsões.

## Referências Bibliográficas

- Aggarwal, C. C. (2017). *Outlier analysis* (2nd ed.). Springer. <https://doi.org/10.1007/978-3-319-47578-3>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2), 211–252.
- Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. Holden-Day.
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time series analysis: Forecasting and control* (4th ed.). John Wiley & Sons.
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.
- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). Springer.
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2), 261–304. <https://doi.org/10.1177/0049124104268644>
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. SPSS. <https://the-modeling-agency.com/crisp-dm.pdf>
- Chatfield, C. (2003). *The analysis of time series: An introduction* (6th ed.). Chapman and Hall/CRC.
- Chatfield, C., & Xing, H. (2019). *The analysis of time series: An introduction* (7th ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9780429272045>
- CIBEN. (n.d.). A CIBEN. <https://www.ciben.pt/a-ciben/>
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a), 427–431.
- Diebold, F. X. (1995). Has forecasting performance improved? An evaluation of the Diebold-Mariano test. *Applied Econometrics*, 10(Suppl. S1), S3–S13.
- Diebold, F. X. (2015). *Forecasting in economics, business, finance, and just about everything else*. Princeton University Press.

- Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263. <https://doi.org/10.1080/07350015.1995.10524599>
- Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 20(1), 134–144. <https://doi.org/10.1198/073500102753410444>
- Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359–378.
- Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.
- Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction accuracy. *International Journal of Forecasting*, 13(2), 241–247.
- Holt, C. C. (1957). Forecasting seasonals and trends by exponentially weighted moving averages. *ONR Research Memorandum*, 52, 1–15.
- Hyndman, R. J. (2017). *Forecasting: Principles and practice*. OTexts. <https://otexts.com/fpp2/>
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts. <https://otexts.com/fpp2/>
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts. <https://otexts.com/fpp3/>
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, 55(2), 163–172. <https://doi.org/10.2307/1403192>
- Letham, B., Taylor, S. J., & Rubin, D. B. (2017). Prophet: Forecasting at scale. *Facebook Research*. <https://facebook.github.io/prophet/>
- Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303. <https://doi.org/10.1093/biomet/65.2.297>
- Makridakis, S., & Hibon, M. (2000). The M3-Competition: Results, conclusions and implications. *International Journal of Forecasting*, 16(4), 451–476.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *PLOS ONE*, 13(3), e0194889. <https://doi.org/10.1371/journal.pone.0194889>
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020a). The M4 Competition: 1000 monthly time series revisited. *International Journal of Forecasting*, 36(1), 1–24. <https://doi.org/10.1016/j.ijforecast.2019.04.014>
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020b). The M4 competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, 36(1), 54–74.

- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020c). The M4 Competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*, 36(1), 54–74.
- Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and applications* (3rd ed.). Wiley.
- Mariano, R. S., & Diebold, F. X. (1995). Comparing predictive accuracy in dynamic models. *Working Papers*, University of Pennsylvania.
- Oreshkin, B. N., Carпов, D., Chapados, N., & Bengio, Y. (2019). N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. *arXiv preprint* arXiv:1905.10437.
- Oreshkin, B. N., Carпов, D., Chapados, N., & Bengio, Y. (2020). N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. In *Proceedings of the International Conference on Learning Representations*. <https://arxiv.org/abs/1905.10437>
- Pandas Development Team. (2024). *pandas-dev/pandas: Pandas*. GitHub. <https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.ffill.html>
- Salinas, D., Flunkert, V., Gasthaus, J., & Januschowski, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *International Journal of Forecasting*, 36(3), 1181–1191. <https://doi.org/10.1016/j.ijforecast.2019.07.001>
- Salinas, D., Rangapuram, S., Hu, Y., & Gasthaus, J. (2020). DeepAR++: Probabilistic forecasting with deep ensembling. *International Journal of Forecasting*, 36(3), 1182–1191.
- Salinas, D., Flunkert, V., & Alexandrov, A. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. In G. E. P. Box, G. M. Jenkins, G. C. Reinsel, & G. M. Ljung, *Time series analysis: Forecasting and control* (5th ed., pp. 317–333). Wiley.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464. <https://doi.org/10.1214/aos/1176344136>
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591–611.
- Shumway, R. H., & Stoffer, D. S. (2017). *Time series analysis and its applications: With R examples* (4th ed.). Springer. <https://doi.org/10.1007/978-3-319-52452-8>
- Stake, R. E. (2006). *Multiple case study analysis*. Guilford Press.
- Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37–45. <https://doi.org/10.1080/00031305.2017.1380080>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems* (Vol. 30, pp. 5998–6008).
- Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3), 324–342. <https://doi.org/10.1287/mnsc.6.3.324>
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage.

## Apêndices A

### Tabelas com os dados obtidos

Tabela A. 1 – Dados 2012-2024 resumidos

|       | GrupoActividade | Actividade                 | AnoMes     | Dep   | Tec | Tempo Realizado |
|-------|-----------------|----------------------------|------------|-------|-----|-----------------|
| 0     | Cliente         | Serviço   Contrato         | 01/01/2012 | BCONS | BC1 | 2               |
| 1     | Cliente         | Serviço   Contrato         | 01/01/2012 | BCONS | BC3 | 0,383333        |
| 2     | Cliente         | Serviço   Contrato         | 01/01/2012 | BCONS | BC2 | 1,5             |
| 3     | Cliente         | Serviço   Contrato         | 01/01/2012 | BSERV | BS1 | 0,1             |
| 4     | Cliente         | Serviço   Contrato         | 01/01/2012 | SWFAC | SF1 | 2,083333        |
| 5     | Cliente         | Serviço   Contrato         | 01/01/2012 | SWFAC | SF2 | 10              |
| 6     | Cliente         | Serviço   Faturado         | 01/01/2012 | BCONS | BC1 | 84,366624       |
| 7     | Cliente         | Serviço   Faturado         | 01/01/2012 | BCONS | BC3 | 134,733299      |
| ...   | ...             | ...                        | ...        | ...   | ... | ...             |
| 20733 | Interno         | Operação   Suporte Interno | 01/12/2024 | INFSO | IS1 | 10,949996       |
| 20734 | Interno         | Operação   Suporte Interno | 01/12/2024 | INFSO | IS2 | 11,55           |
| 20735 | Interno         | Operação   Suporte Interno | 01/12/2024 | SWFAC | SF1 | 2,75            |
| 20736 | Interno         | Operação   Suporte Interno | 01/12/2024 | SWFAC | SF2 | 7,583333        |

Tabela A. 2 – Dados 2012-2024 agrupados por Ano-Mês

| Indice | AnoMes  | Tempo Realizado |
|--------|---------|-----------------|
| 0      | 2012-01 | 1712.549473     |
| 1      | 2012-02 | 1687.382844     |
| 2      | 2012-03 | 1787.799492     |
| 3      | 2012-04 | 1540.866240     |
| ...    | ...     | ...             |
| 151    | 2024-09 | 3326.115900     |
| 152    | 2024-10 | 3500.699203     |
| 153    | 2024-11 | 3080.599331     |
| 154    | 2024-12 | 3215.516012     |

Tabela A. 3 - Resumo dos resultados do teste ADF

| Séries      | ADF<br>Statistic | p_value                | Critical Values |         |         |
|-------------|------------------|------------------------|-----------------|---------|---------|
|             |                  |                        | '1%'            | '5%'    | '10%'   |
| Dados reais | -5.8452          | 3.6974543571563717e-07 | -3.6327         | -2.9485 | -2.6130 |
| Dados norm. | -6.3158          | 3.15182040693317e-08   | -3.6327         | -2.9485 | -2.9485 |

Tabela A. 4 - Resumo dos resultados do modelo SARIMA Geral

| SARIMAX Results         |                                |          |                   |          |           |         |
|-------------------------|--------------------------------|----------|-------------------|----------|-----------|---------|
| Dep. Variable:          | Tempo Realizado                |          | No. Observations: | 36       |           |         |
| Model:                  | SARIMAX(2, 1, 1)x(1, 1, 1, 12) |          | Log Likelihood    | -159.997 |           |         |
| Date:                   | Tue, 27 May 2025               |          | AIC               | 331.994  |           |         |
| Time:                   | 22:01:33                       |          | BIC               | 338.806  |           |         |
| Sample:                 | 0                              |          | HQIC              | 333.707  |           |         |
|                         | - 36                           |          |                   |          |           |         |
| Covariance Type:        | opg                            |          |                   |          |           |         |
|                         | coef                           | std err  | z                 | P> z     | [0.025    | 0.975]  |
| ar.L1                   | -0.3096                        | 0.303    | -1.021            | 0.307    | -0.904    | 0.285   |
| ar.L2                   | -0.3162                        | 0.325    | -0.973            | 0.330    | -0.953    | 0.320   |
| ma.L1                   | -0.9649                        | 1.455    | -0.663            | 0.507    | -3.816    | 1.886   |
| ar.S.L12                | -0.8020                        | 0.845    | -0.949            | 0.343    | -2.458    | 0.854   |
| ma.S.L12                | -0.1374                        | 3.210    | -0.043            | 0.966    | -6.430    | 6.155   |
| sigma2                  | 3.054e+04                      | 4.04e+04 | 0.756             | 0.450    | -4.86e+04 | 1.1e+05 |
| Ljung-Box (L1) (Q):     | 0.07                           |          | Jarque-Bera (JB): | 1.28     |           |         |
| Prob(Q):                | 0.80                           |          | Prob(JB):         | 0.53     |           |         |
| Heteroskedasticity (H): | 1.19                           |          | Skew:             | -0.28    |           |         |
| Prob(H) (two-sided):    | 0.81                           |          | Kurtosis:         | 1.99     |           |         |

Tabela A. 5 – Valores das métricas de validação Walk-Foward do modelo SARIMA Geral

| Métrica | Valor      |
|---------|------------|
| MAE     | 187.93     |
| MSE     | 47 566.91  |
| RMSE    | 218.09     |
| MAPE    | 5.56%      |
| MdAE    | 175.55     |
| MdAPE   | 5.33%      |
| MASE    | NaN        |
| Viés    | -19.57     |
| sMAPE   | 5.60       |
| MSIS    | NaN        |
| RSS     | 570 802.93 |

Tabela A. 6 - Resumo dos resultados do modelo SARIMA Simplificado

| SARIMAX Results                                                                       |                                |                   |                   |          |          |          |
|---------------------------------------------------------------------------------------|--------------------------------|-------------------|-------------------|----------|----------|----------|
| =====                                                                                 |                                |                   |                   |          |          |          |
| Dep. Variable:                                                                        | Tempo Realizado                |                   | No. Observations: | 36       |          |          |
| Model:                                                                                | SARIMAX(2, 1, 0)x(0, 1, 0, 12) |                   | Log Likelihood    | -152.520 |          |          |
| Date:                                                                                 | Tue, 27 May 2025               |                   | AIC               | 311.039  |          |          |
| Time:                                                                                 | 22:03:25                       |                   | BIC               | 314.173  |          |          |
| Sample:                                                                               | 0                              |                   | HQIC              | 311.719  |          |          |
|                                                                                       | - 36                           |                   |                   |          |          |          |
| Covariance Type:                                                                      | opg                            |                   |                   |          |          |          |
| =====                                                                                 |                                |                   |                   |          |          |          |
|                                                                                       | coef                           | std err           | z                 | P> z     | [0.025   | 0.975]   |
| -----                                                                                 |                                |                   |                   |          |          |          |
| ar.L1                                                                                 | -0.6650                        | 0.233             | -2.849            | 0.004    | -1.123   | -0.207   |
| ar.L2                                                                                 | -0.4889                        | 0.267             | -1.829            | 0.067    | -1.013   | 0.035    |
| sigma2                                                                                | 1.191e+05                      | 5.27e+04          | 2.261             | 0.024    | 1.59e+04 | 2.22e+05 |
| =====                                                                                 |                                |                   |                   |          |          |          |
| Ljung-Box (L1) (Q):                                                                   | 0.15                           | Jarque-Bera (JB): | 1.29              |          |          |          |
| Prob(Q):                                                                              | 0.70                           | Prob(JB):         | 0.53              |          |          |          |
| Heteroskedasticity (H):                                                               | 0.79                           | Skew:             | -0.51             |          |          |          |
| Prob(H) (two-sided):                                                                  | 0.77                           | Kurtosis:         | 2.34              |          |          |          |
| =====                                                                                 |                                |                   |                   |          |          |          |
| Warnings:                                                                             |                                |                   |                   |          |          |          |
| [1] Covariance matrix calculated using the outer product of gradients (complex-step). |                                |                   |                   |          |          |          |

Tabela A. 7 – Valores das métricas de validação Walk-Foward do modelo SARIMA Simplificado

| Métrica | Valor        |
|---------|--------------|
| MAE     | 287.66       |
| MSE     | 110 139.23   |
| RMSE    | 331.87       |
| MAPE    | 8.59 %       |
| MdAE    | 249.83       |
| MdAPE   | 7.80 %       |
| MASE    | NaN          |
| Viés    | -103.12      |
| sMAPE   | 8.79         |
| MSIS    | NaN          |
| RSS     | 1 321 670.79 |

Tabela A. 8 - Comparação dos resultados do Teste de Ljung-Box entre as duas versões do modelo SARIMA

| Lag | Geral | Simplificado |
|-----|-------|--------------|
| 1   | 0.55  | 0.32         |
| 2   | 0.55  | 0.56         |

|    |      |      |
|----|------|------|
| 3  | 0.83 | 0.76 |
| 4  | 0.94 | 0.88 |
| 5  | 0.98 | 0.88 |
| 6  | 0.94 | 0.94 |
| 7  | 0.97 | 0.96 |
| 8  | 0.99 | 0.97 |
| 9  | 0.99 | 0.99 |
| 10 | 1.00 | 0.99 |
| 11 | 1.00 | 0.99 |
| 12 | 0.75 | 0.27 |

Tabela A. 9 – Resumo dos resultados do modelo HWA

| ExponentialSmoothing Model Results |                      |                   |                  |
|------------------------------------|----------------------|-------------------|------------------|
| Dep. Variable:                     | Tempo Realizado      | No. Observations: | 36               |
| Model:                             | ExponentialSmoothing | SSE               | 976449.095       |
| Optimized:                         | True                 | AIC               | 399.494          |
| Trend:                             | Additive             | BIC               | 424.830          |
| Seasonal:                          | Additive             | AICC              | 439.729          |
| Seasonal Periods:                  | 12                   | Date:             | Sat, 20 Sep 2025 |
| Box-Cox:                           | False                | Time:             | 23:00:54         |
| Box-Cox Coeff.:                    | None                 |                   |                  |
|                                    | coeff                | code              | optimized        |
| smoothing_level                    | 1.4901e-08           | alpha             | True             |
| smoothing_trend                    | 1.4883e-08           | beta              | True             |
| smoothing_seasonal                 | 0.000000             | gamma             | True             |
| initial_level                      | 3528.1824            | l.0               | True             |
| initial_trend                      | -6.0502284           | b.0               | True             |
| initial_seasons.0                  | 295.93075            | s.0               | True             |
| initial_seasons.1                  | -21.386950           | s.1               | True             |
| initial_seasons.2                  | -21.720710           | s.2               | True             |
| initial_seasons.3                  | -149.62455           | s.3               | True             |
| initial_seasons.4                  | 21.081974            | s.4               | True             |
| initial_seasons.5                  | -150.11744           | s.5               | True             |
| initial_seasons.6                  | 194.76982            | s.6               | True             |
| ...                                |                      |                   |                  |
| initial_seasons.9                  | 13.642851            | s.9               | True             |
| initial_seasons.10                 | -52.368596           | s.10              | True             |
| initial_seasons.11                 | -186.71894           | s.11              | True             |

Tabela A. 10 – Valores das métricas de validação Walk-Forward do modelo HWA

| Métrica | Valor     |
|---------|-----------|
| MAE     | 229.65    |
| MSE     | 74 564.18 |
| RMSE    | 273.06    |
| MAPE    | 6.54%     |
| MdAE    | 272.53    |
| MpAPE   | 7.61%     |
| MASE    | 1.00      |

| Métrica | Valor   |
|---------|---------|
| Viés    | -216.81 |
| sMAPE   | NaN     |
| MSIS    | NaN     |
| RSS     | NaN     |

Tabela A. 11 - Resumo dos resultados do modelo HWM

| ExponentialSmoothing Model Results |                      |                   |                  |  |
|------------------------------------|----------------------|-------------------|------------------|--|
| =====                              |                      |                   |                  |  |
| Dep. Variable:                     | Tempo Realizado      | No. Observations: | 36               |  |
| Model:                             | ExponentialSmoothing | SSE               | 977037.801       |  |
| Optimized:                         | True                 | AIC               | 399.515          |  |
| Trend:                             | Additive             | BIC               | 424.852          |  |
| Seasonal:                          | Multiplicative       | AICC              | 439.751          |  |
| Seasonal Periods:                  | 12                   | Date:             | Sat, 17 May 2025 |  |
| Box-Cox:                           | True                 | Time:             | 21:48:17         |  |
| Box-Cox Coeff.:                    | 0.50000              |                   |                  |  |
| =====                              |                      |                   |                  |  |
|                                    | coeff                | code              | optimized        |  |
| -----                              |                      |                   |                  |  |
| smoothing_level                    | 1.4901e-08           | alpha             | True             |  |
| smoothing_trend                    | 6.8644e-09           | beta              | True             |  |
| smoothing_seasonal                 | 1.5747e-09           | gamma             | True             |  |
| initial_level                      | 109.62337            | l.0               | True             |  |
| initial_trend                      | -0.0973604           | b.0               | True             |  |
| initial_seasons.0                  | 1.1108777            | s.0               | True             |  |
| initial_seasons.1                  | 1.0615277            | s.1               | True             |  |
| initial_seasons.2                  | 1.0621050            | s.2               | True             |  |
| initial_seasons.3                  | 1.0408495            | s.3               | True             |  |
| initial_seasons.4                  | 1.0688238            | s.4               | True             |  |
| initial_seasons.5                  | 1.0408405            | s.5               | True             |  |
| initial_seasons.6                  | 1.0959252            | s.6               | True             |  |
| ...                                |                      |                   |                  |  |
| initial_seasons.9                  | 1.0674576            | s.9               | True             |  |
| initial_seasons.10                 | 1.0569596            | s.10              | True             |  |
| initial_seasons.11                 | 1.0350532            | s.11              | True             |  |
| -----                              |                      |                   |                  |  |

Tabela A. 12 – Valores das métricas de validação Walk-Foward do modelo HWM

| Métrica | Valor     |
|---------|-----------|
| MAE     | 229.33    |
| MSE     | 74 291.67 |
| RMSE    | 272.56    |
| MAPE    | 6.53%     |
| MdAE    | 271.82    |
| MdAPE   | 7.76%     |
| MASE    | 1.00      |
| Viés    | -216.99   |
| sMAPE   | NaN       |
| MSIS    | NaN       |
| RSS     | NaN       |

Tabela A. 13 - Comparação da informação geral entre o modelo HWA e HWM

| <b>Característica</b> | <b>HWA</b>           | <b>HWM</b>           | <b>Alteração Significativa?</b> |
|-----------------------|----------------------|----------------------|---------------------------------|
| Modelo                | ExponentialSmoothing | ExponentialSmoothing | Não                             |
| Otimizado             | True                 | True                 | Não                             |
| Tendência             | Aditiva              | Aditiva              | Não                             |
| Sazonal               | Aditiva              | Multiplicativa       | Sim                             |
| Períodos Sazonais     | 12                   | 12                   | Não                             |
| Nº Obs.               | 36                   | 36                   | Não                             |
| SSE                   | 977 038.08           | 977 037.80           | Não                             |
| AIC                   | 399.49               | 399.51               | Não                             |
| AICc                  | 439.73               | 439.75               | Não                             |
| BIC                   | 424.83               | 424.85               | Não                             |
| Log-Likelihood        | -234.84              | -234.85              | Não                             |

Tabela A. 14 - Comparação dos resultados do Teste de Ljung-Box e do Teste Jarque-Bera entre os modelos HWA e HWM

| <b>LB Lag</b> | <b>HWA</b> | <b>HWM</b> |
|---------------|------------|------------|
| 1             | 0.48       | 0.47       |
| 2             | 0.41       | 0.40       |
| 3             | 0.23       | 0.22       |
| 4             | 0.05       | 0.04       |
| 5             | 0.06       | 0.05       |
| 6             | 0.09       | 0.08       |
| 7             | 0.14       | 0.13       |
| 8             | 0.14       | 0.13       |
| 9             | 0.20       | 0.19       |
| 10            | 0.23       | 0.22       |
| 11            | 0.29       | 0.28       |
| 12            | 0.01       | 0.01       |
| <b>JB</b>     | <b>HWA</b> | <b>HWM</b> |
| P_value       | 0.62       | 0.63       |

Tabela A. 15 - Resumo das métricas de informação do modelo N-BEATS Simples

| Métrica        | Valor     |
|----------------|-----------|
| AIC            | 19 479.23 |
| BIC            | 24157.61  |
| AICc           | NaN       |
| Log_Likelihood | -94.72    |
| $\sigma$       | 323.86    |
| $\sigma^2$     | 104886.07 |

Tabela A. 16 - Resumo das métricas de Erro do modelo N-BEATS Simples

| Métrica     | Valor        |
|-------------|--------------|
| MAE         | 414.81       |
| MSE         | 250 632.51   |
| MASE        | 0.88         |
| MAPE        | 12.05%       |
| sMAPE       | 13.16%       |
| RMSE        | 500.63       |
| MdAE        | 270.06       |
| MdAPE       | 8.01%        |
| Bias        | -381.77      |
| Média y     | 3369.27      |
| Variância y | 49 026.93    |
| RSS         | 3.007590e+06 |

Tabela A. 17 - Resumo das métricas de informação do modelo N-BEATS Especializado

| Métrica        | Valor      |
|----------------|------------|
| AIC            | 98 686.33  |
| BIC            | 122 568.95 |
| AICc           | NaN        |
| Log_Likelihood | -91.413    |
| $\sigma$       | 419.906    |
| $\sigma^2$     | 176 321.53 |

Tabela A. 18 - Resumo das Métricas de Erro do modelo N-BEATS Especializado

| <b>Métrica</b> | <b>Valor</b> |
|----------------|--------------|
| MAE            | 398.03       |
| MSE            | 244 326.29   |
| MASE           | 0.85         |
| MAPE           | 11.58%       |
| sMAPE          | 12.66%       |
| RMSE           | 494.29       |
| MdAE           | 334.44       |
| MdAPE          | 10.26%       |
| Bias           | -334.438     |
| Média y        | 3 369.271    |
| Variância y    | 49 026.93    |
| RSS            | 2.931915e+06 |

Tabela A. 19 - Comparação dos resultados do Teste Ljung-Box e do Teste Jarque-Bera entre as duas versões do modelo N-BEATS

| <b>LB Lag</b> | <b>Simple</b> | <b>Especificado</b> |
|---------------|---------------|---------------------|
| 1             | 0.88          | 0.90                |
| 2             | 0.94          | 0.39                |
| 3             | 0.98          | 0.59                |
| 4             | 0.43          | 0.62                |
| 5             | 0.34          | 0.02                |
| 6             | 0.21          | 0.03                |
| 7             | 0.26          | 0.01                |
| 8             | 0.33          | 0.01                |
| 9             | 0.42          | 0.01                |
| 10            | 0.50          | 0.02                |
| 11            | 0.59          | 0.03                |
| 12            | 0.52          | 0.00                |
| <b>JB</b>     | <b>Simple</b> | <b>Especificado</b> |
| P_value       | 0.60          | 0.32                |

Tabela A. 20 - Resumo das métricas de avaliação do modelo DeepAR Padrão

| <b>Métrica</b>                    | <b>Valor</b> |
|-----------------------------------|--------------|
| MSE                               | 158.66       |
| abs_error (Erro Absoluto)         | 109.17       |
| abs_target_sum                    | 4 0431.25    |
| abs_target_mean                   | 3 369.27     |
| seasonal_error                    | 263.26       |
| MASE                              | 0.04         |
| MAPE                              | 0,28%        |
| sMAPE                             | 0,27%        |
| MSIS (Mean Scaled Interval Score) | 0.39         |

Tabela A. 21 - Resumo das métricas de Quantis do modelo DeepAR Padrão

| <b>Quantil</b> | <b>Quantile Loss</b> | <b>Coverage</b> |
|----------------|----------------------|-----------------|
| 0.1            | 78.89                | 0.08            |
| 0.5            | 109.17               | 0.58            |
| 0.9            | 105.82               | 1.0             |

Tabela A. 22 - Resumo das métricas de informação do modelo DeepAR Padrão

| <b>Métrica</b>                 | <b>Valor</b> |
|--------------------------------|--------------|
| RMSE                           | 12.60        |
| NRMSE (Normalized RMSE)        | 0.0037       |
| ND (Normalized Deviation)      | 0.0027       |
| wQuantileLoss[0.1]             | 0.0020       |
| wQuantileLoss[0.5]             | 0.0027       |
| wQuantileLoss[0.9]             | 0.0026       |
| mean_absolute_QuantileLoss     | 97.96        |
| mean_wQuantileLoss             | 0.0024       |
| MAE_Coverage                   | 0.5          |
| OWA (Overall Weighted Average) | NaN          |
| AIC                            | NaN          |
| BIC                            | NaN          |
| AICc                           | NaN          |
| Log_Likelihood                 | NaN          |

|                   |     |
|-------------------|-----|
| Nº de parâmetros  | NaN |
| Nº de observações | NaN |
| $\sigma$          | NaN |
| $\sigma^2$        | NaN |

Tabela A. 23 - Resumo das métricas de avaliação do modelo DeepAR Otimizado

| Métrica                  | Valor     |
|--------------------------|-----------|
| MSE                      | 89.42     |
| abs_error                | 117.02    |
| abs_target_sum           | 4 0431.25 |
| abs_target_mean          | 3 369.27  |
| seasonal_error           | 263.26    |
| MASE                     | 0.04      |
| MAPE                     | 0.29%     |
| sMAPE                    | 0.29%     |
| MSIS                     | 0.43      |
| num_masked_target_values | 0.0       |

Tabela A. 24 - Resumo de métricas de Quantis do modelo DeepAR Otimizado

| Quantil | Quantile Loss | Coverage |
|---------|---------------|----------|
| 0.1     | 75.46         | 0.0      |
| 0.5     | 114.77        | 0.42     |
| 0.9     | 90.03         | 1.0      |

Tabela A. 25 - Resumo das métricas de informação do modelo DeepAR Otimizado

| Métrica                    | Valor  |
|----------------------------|--------|
| RMSE                       | 9,46   |
| NRMSE                      | 0.0028 |
| ND                         | 0.0029 |
| wQuantileLoss[0.1]         | 0.0022 |
| wQuantileLoss[0.5]         | 0.0029 |
| wQuantileLoss[0.9]         | 0.0028 |
| mean_absolute_QuantileLoss | 106.22 |
| mean_wQuantileLoss         | 0.0026 |

|                   |     |
|-------------------|-----|
| MAE_Coverage      | 0.5 |
| OWA               | NaN |
| AIC               | NaN |
| BIC               | NaN |
| AICc              | NaN |
| Log_Likelihood    | NaN |
| Nº de parâmetros  | NaN |
| Nº de observações | NaN |
| $\sigma$          | NaN |
| $\sigma^2$        | NaN |

Tabela A. 26 – Comparação dos resultados do Teste Ljung-Box e do Teste Jarque-Bera entre as duas versões do modelo DeepAR

| LB Lag  | Padrão | Otimizado |
|---------|--------|-----------|
| 1       | 0.36   | 0.36      |
| 2       | 0.66   | 0.65      |
| 3       | 0.79   | 0.78      |
| 4       | 0.87   | 0.87      |
| 5       | 0.92   | 0.92      |
| 6       | 0.96   | 0.96      |
| 7       | 0.98   | 0.98      |
| 8       | 0.98   | 0.98      |
| 9       | 0.99   | 0.99      |
| 10      | 1.00   | 1.00      |
| 11      | 1.00   | 1.00      |
| 12      | 1.00   | 1.00      |
| JB      | Padrão | Otimizado |
| P_value | 0.23   | 0.91      |

Tabela A. 27 - Resumo das métricas de Erro do modelo Prophet Padrão

| Métrica | Valor     |
|---------|-----------|
| MAE     | 89.11     |
| MSE     | 13 374.29 |

|             |            |
|-------------|------------|
| MASE        | 1.00       |
| MAPE        | 2.60%      |
| sMAPE       | 2.59       |
| RMSE        | 115.65     |
| MdAE        | 72.14      |
| MdAPE       | 2.07%      |
| Bias        | 0.04       |
| Média y     | 3 421.01   |
| Variância y | 51 569.23  |
| RSS         | 481 474.48 |

Tabela A. 28 - Resumo das métricas de informação do modelo Prophet Padrão

| <b>Métrica</b>    | <b>Valor</b> |
|-------------------|--------------|
| AIC               | 488.22       |
| BIC               | 523.05       |
| AICc              | 566.06       |
| Log_Likelihood    | -222.11      |
| Nº de parâmetros  | 22           |
| Nº de observações | 36           |
| $\sigma$          | 117.29       |
| $\sigma^2$        | 13781.41     |

Tabela A. 29 - Resumo das métricas de Erro do modelo Prophet Otimizado

| <b>Métrica</b> | <b>Valor</b> |
|----------------|--------------|
| MAE            | 89.53        |
| MSE            | 13 398.52    |
| MASE           | 115.75       |
| MAPE           | 2.62%        |
| sMAPE          | 2.61%        |
| RMSE           | 2.17%        |
| MdAE           | 1.00         |
| MdAPE          | 75.48        |
| Bias           | -0.017       |
| Média y        | 3 421.01     |

|             |            |
|-------------|------------|
| Variância y | 51 569.23  |
| RSS         | 482 346.65 |

Tabela A. 30 - Resumo das métricas de informação do modelo Prophet Otimizado

| <b>Métrica</b>    | <b>Valor</b> |
|-------------------|--------------|
| AIC               | 488.28       |
| BIC               | 523.12       |
| AICc              | 566.13       |
| Log_Likelihood    | -222.14      |
| Nº de parâmetros  | 22           |
| Nº de observações | 36           |
| $\sigma$          | 117.39       |
| $\sigma^2$        | 13 781.33    |

Tabela A. 31 – Comparação dos resultados do Teste Ljung-Box e do Teste Jarque-Bera entre as duas versões do modelo Prophet

| <b>Lag</b> | <b>Padrão</b> | <b>Otimizado</b> |
|------------|---------------|------------------|
| 1          | 0.57          | 0.58             |
| 2          | 0.10          | 0.10             |
| 3          | 0.10          | 0.10             |
| 4          | 0.16          | 0.17             |
| 5          | 0.01          | 0.01             |
| 6          | 0.01          | 0.01             |
| 7          | 0.00          | 0.00             |
| 8          | 0.00          | 0.00             |
| 9          | 0.00          | 0.00             |
| 10         | 0.00          | 0.00             |
| 11         | 0.00          | 0.00             |
| 12         | 0.57          | 0.58             |
| <b>JB</b>  | <b>Padrão</b> | <b>Otimizado</b> |
| P_value    | 0.78          | 0.76             |

## Apêndices B

### Representação gráfica

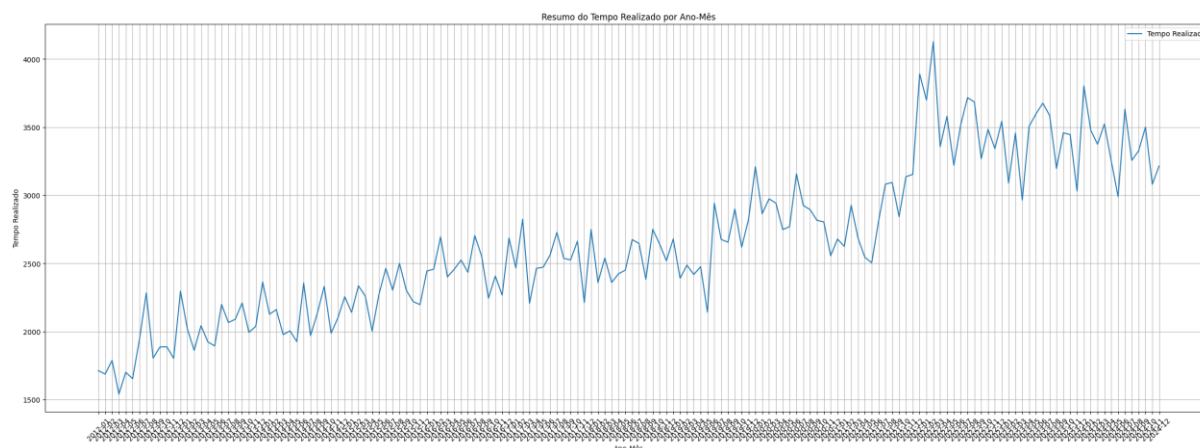


Figura B. 1 – Cronograma dos dados reais dos dados histórico a 12 anos (2012 a 2024)

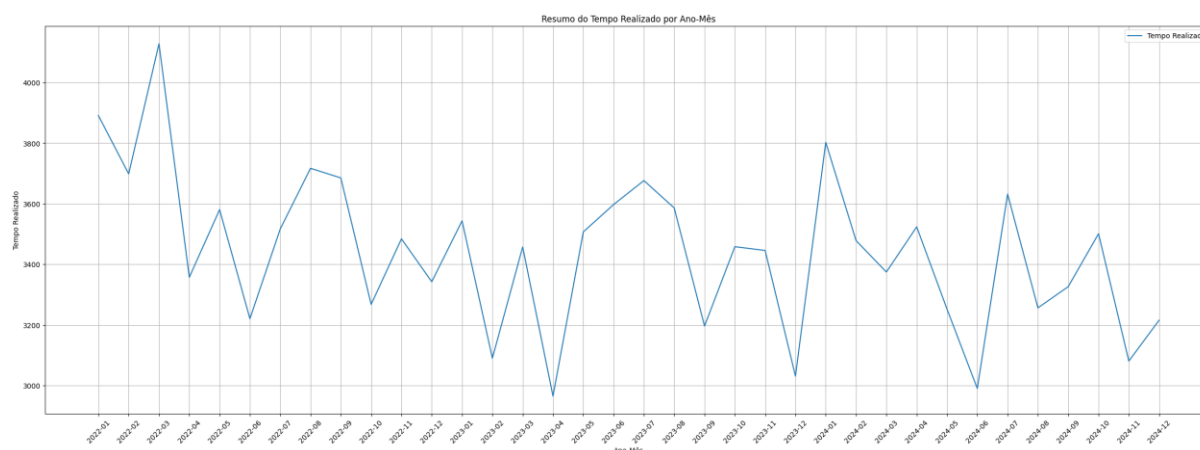


Figura B. 2 – Cronograma da série temporal com os dados originais dos últimos 3 anos (2021 a 2024)

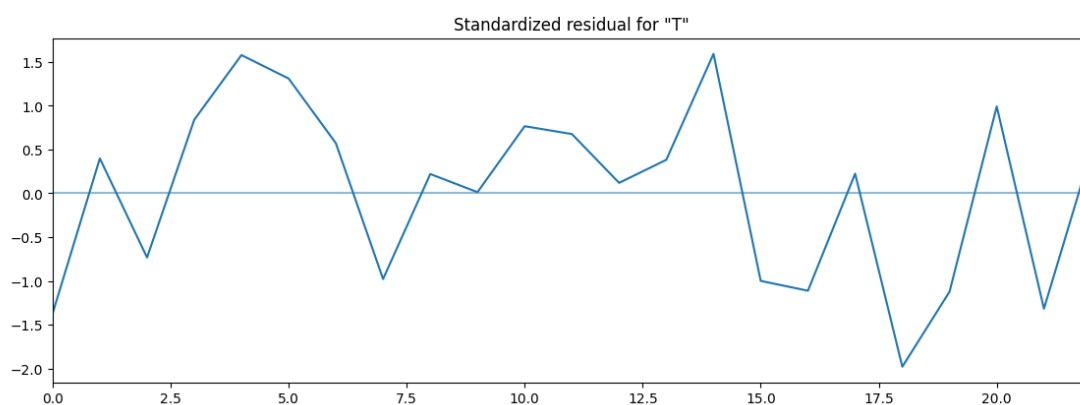


Figura B. 3 – Gráfico de resíduos do modelo SARIMA Geral

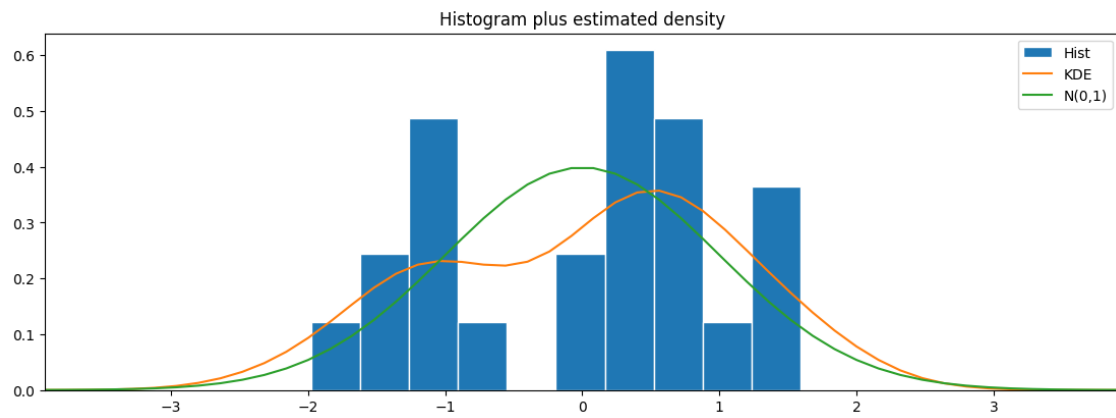


Figura B. 4 - Distribuição dos resíduos do modelo SARIMA Geral

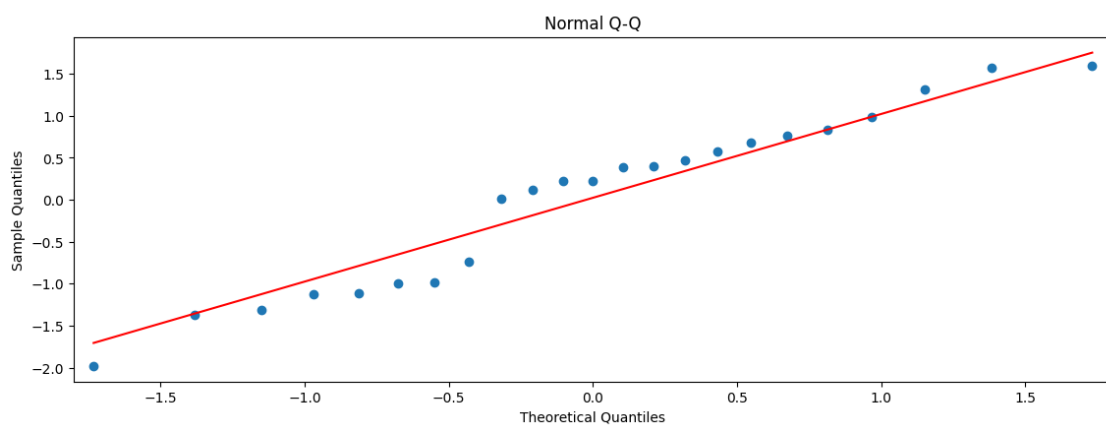


Figura B. 5 - Q-Q plot do modelo SARIMA Geral

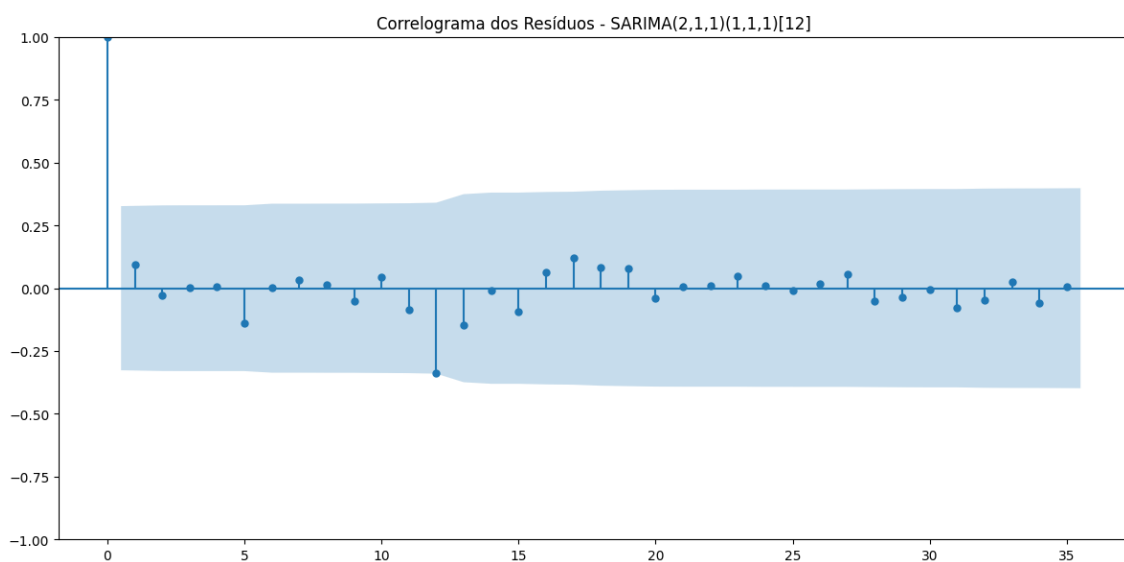


Figura B. 6 - ACF dos resíduos do modelo SARIMA Geral

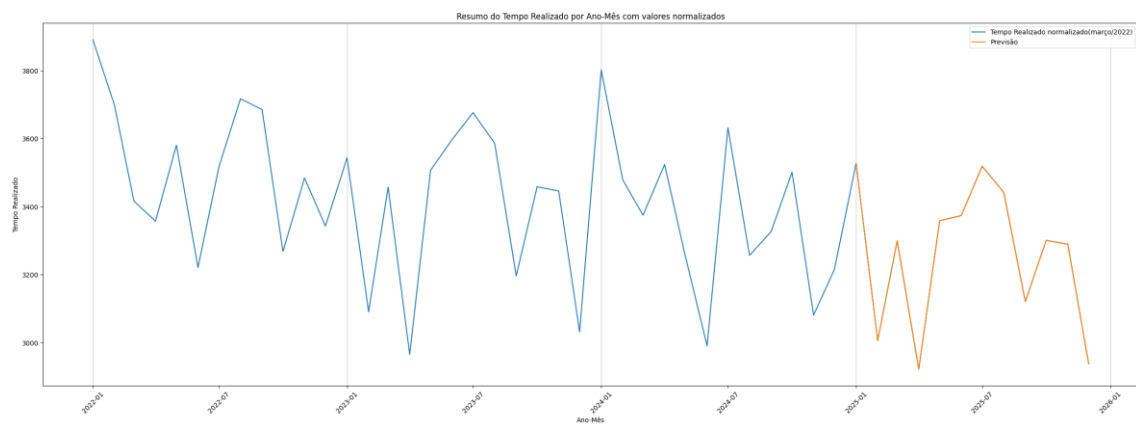


Figura B. 7 – Cronograma dos dados reais e das previsões do modelo SARIMA Geral

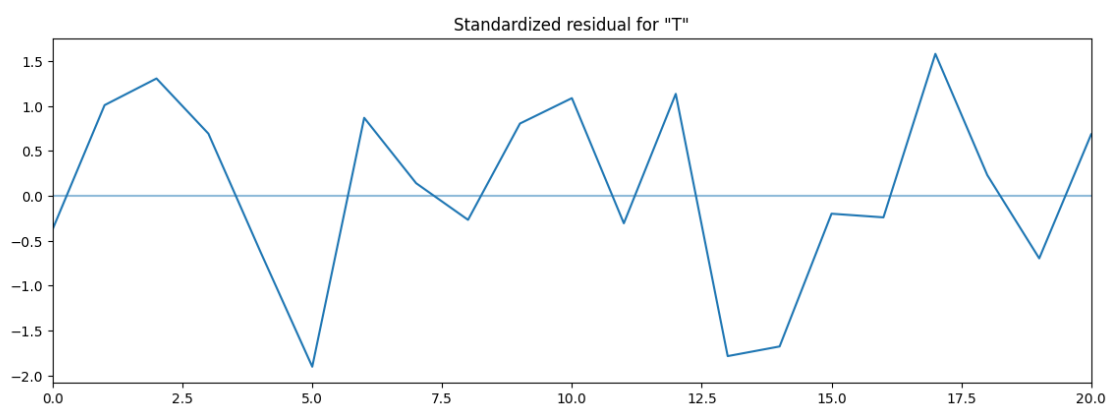


Figura B. 8 – Gráfico de resíduos do modelo SARIMA Simplificado

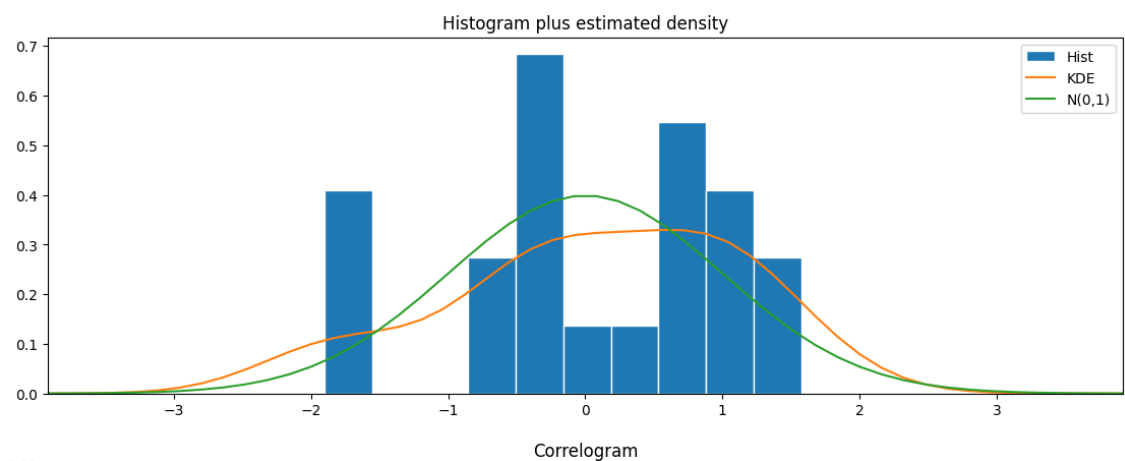


Figura B. 9 - Distribuição dos resíduos do modelo SARIMA Simplificado

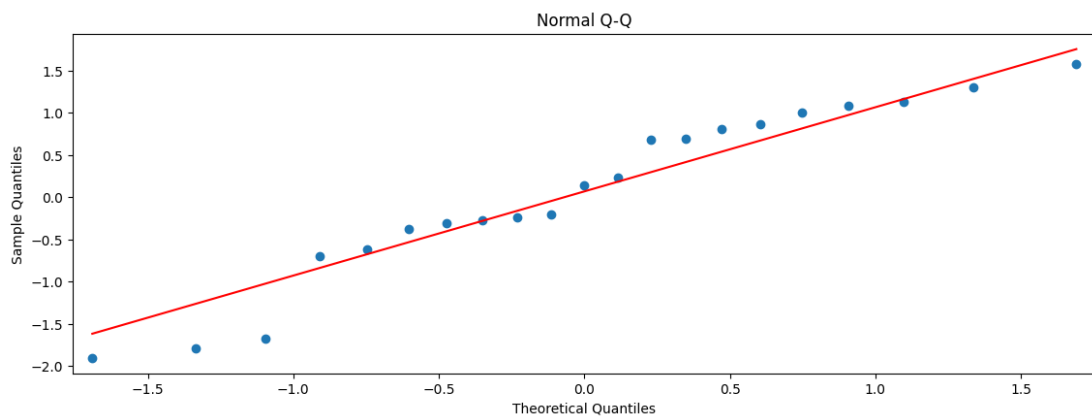


Figura B. 10 - Q-Q plot do modelo SARIMA Simplificado

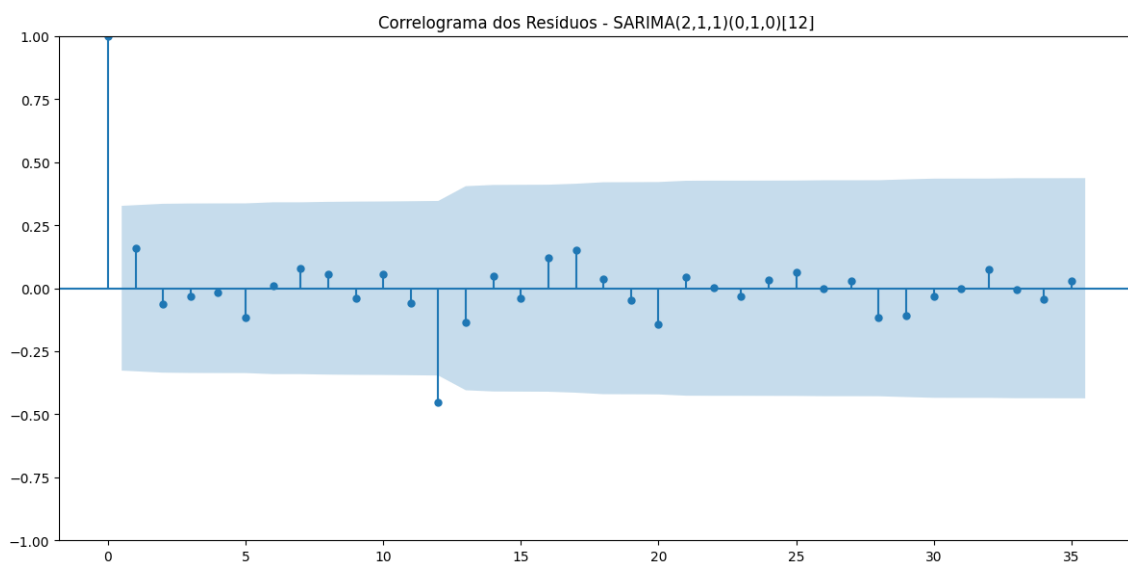


Figura B. 11 - ACF dos resíduos do modelo SARIMA Simplificado

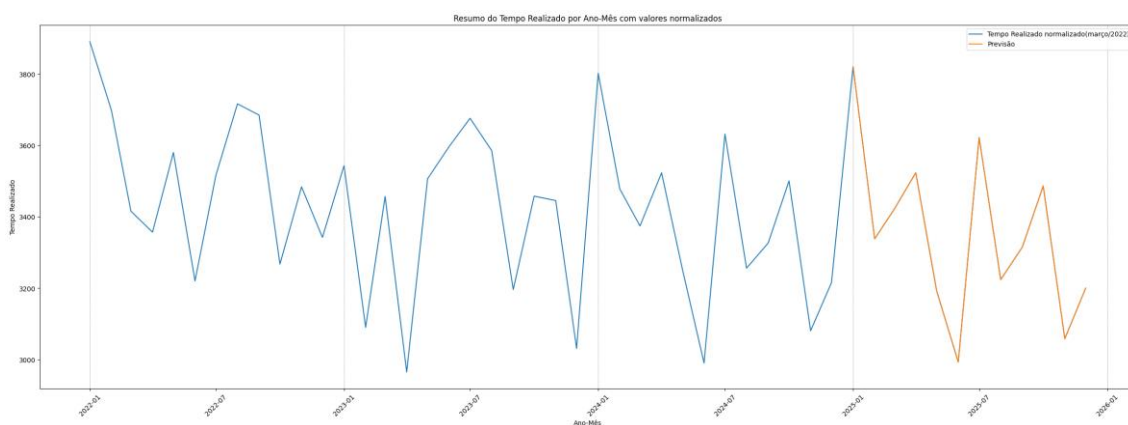


Figura B. 12 – Cronograma dos dados reais e das previsões com o modelo SARIMA Simplificado

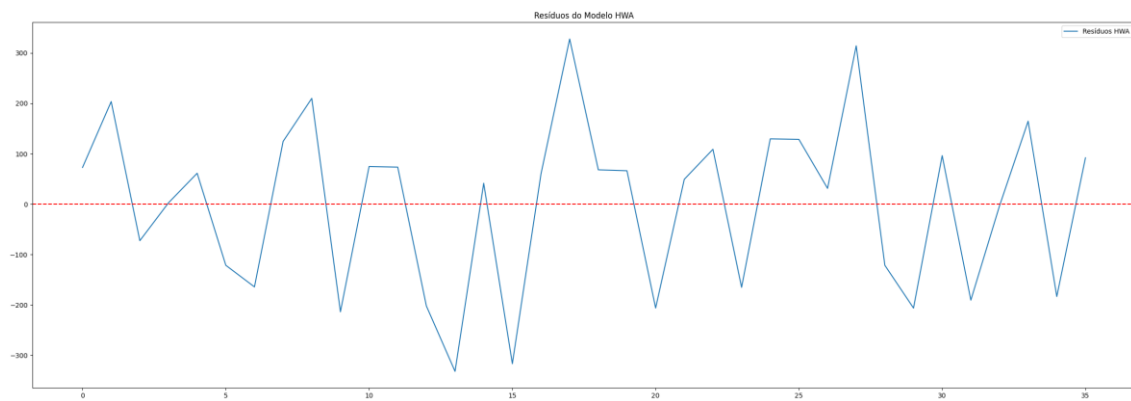


Figura B. 13 – Gráfico de resíduos do modelo HWA

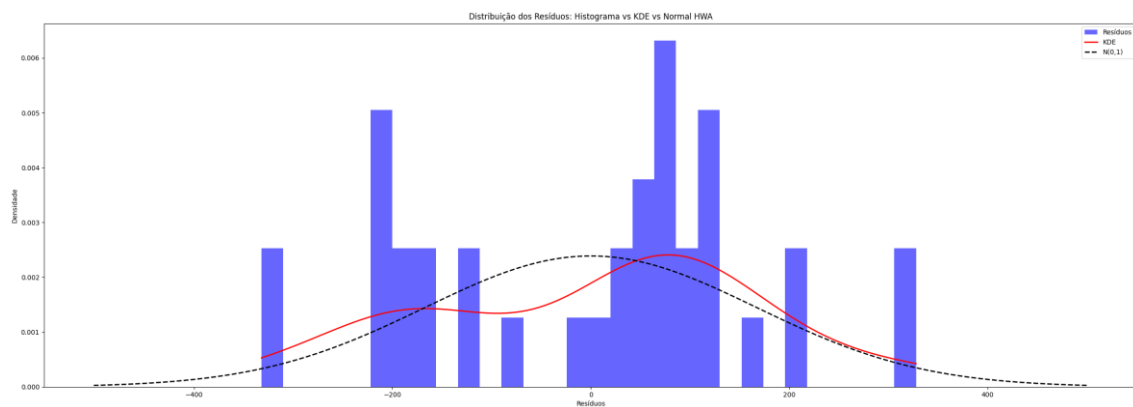


Figura B. 14 - Distribuição dos resíduos do modelo HWA

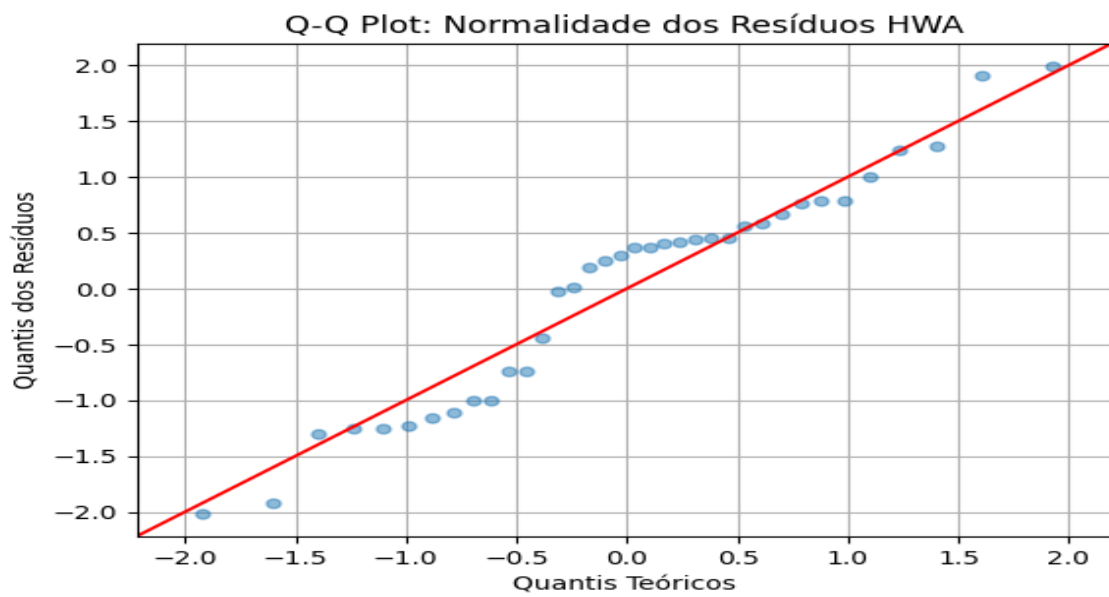


Figura B. 15 - Probabilidade normal (Q-Q plot) do modelo HWA

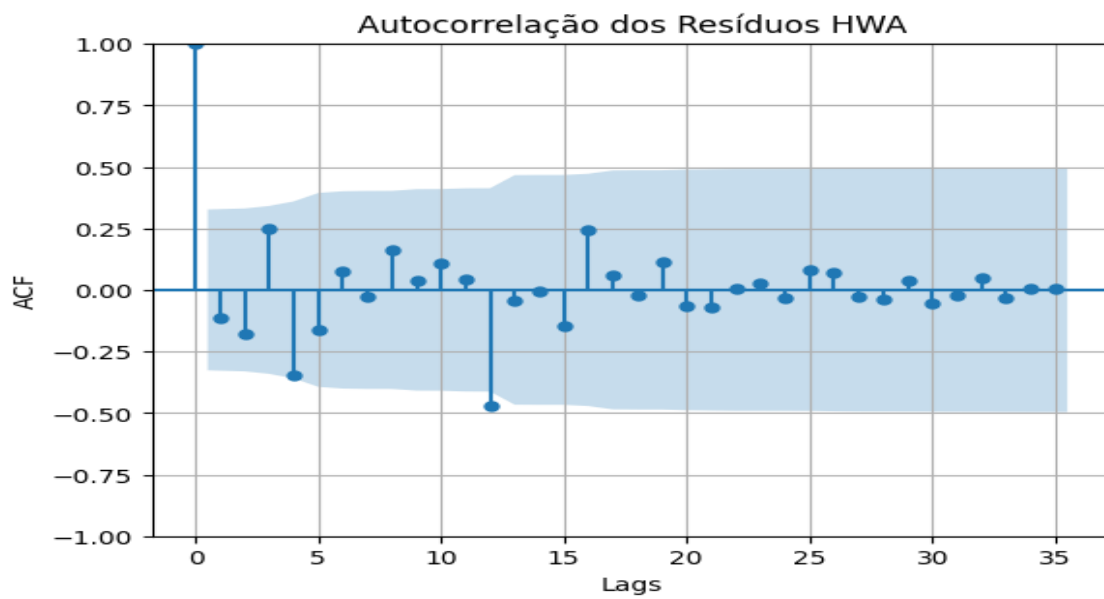


Figura B. 16 - Correlograma dos resíduos do modelo HWA

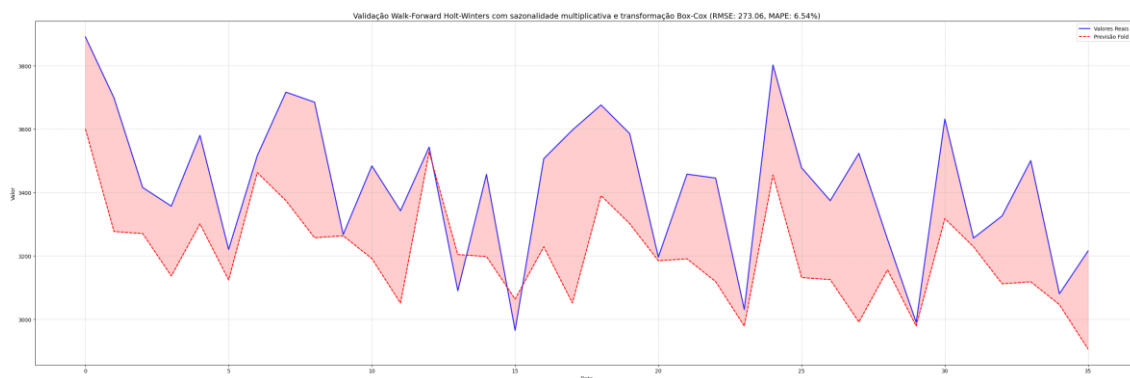


Figura B. 17 – Validação Walk-Forward do modelo HWA

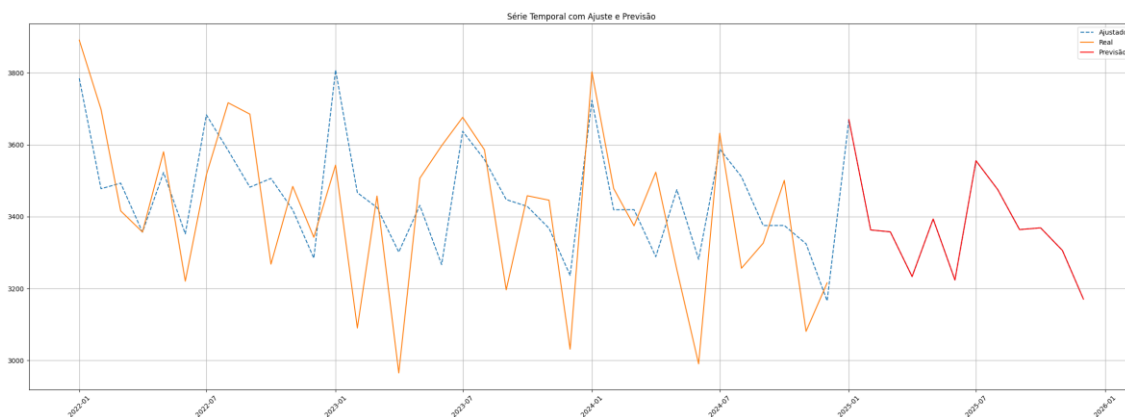


Figura B. 18 – Cronograma da série temporal com os ajustes e a previsão do modelo HWA

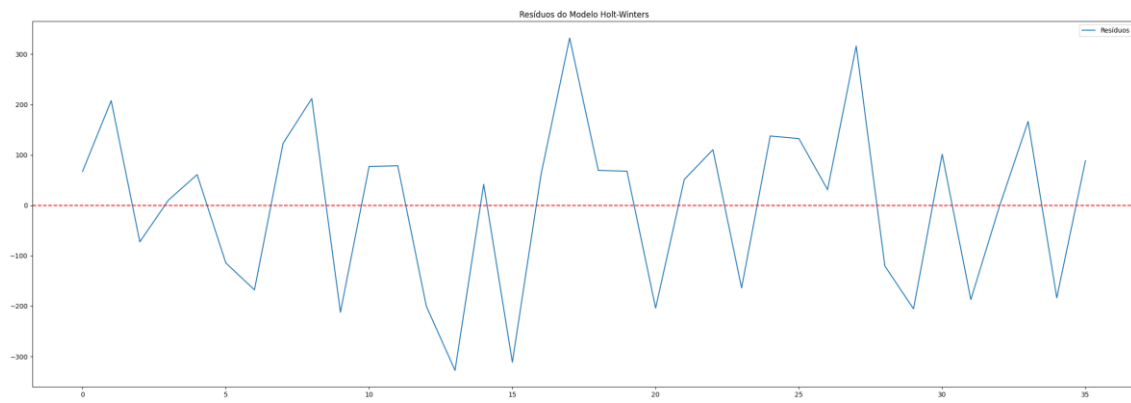


Figura B. 19 – Gráfico de resíduos do modelo HWM

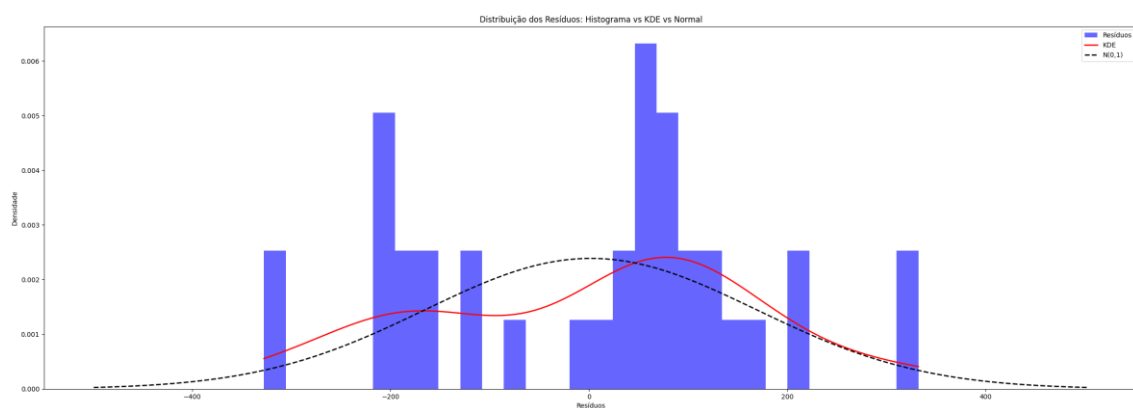


Figura B. 20 - Distribuição dos resíduos do modelo HWM

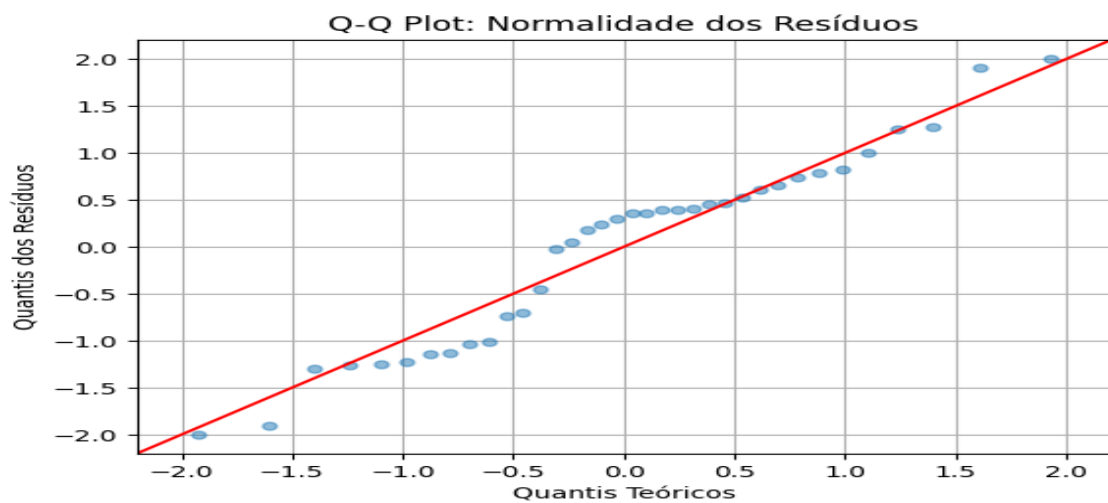


Figura B. 21 - Probabilidade normal (Q-Q plot) do modelo HWM

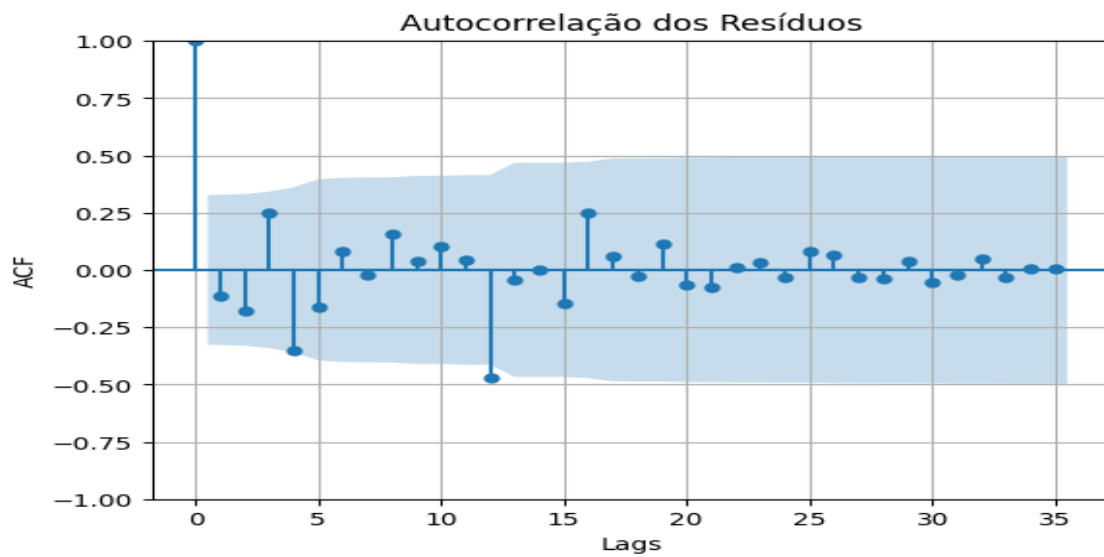


Figura B. 22 - Correlograma dos resíduos do modelo HWM

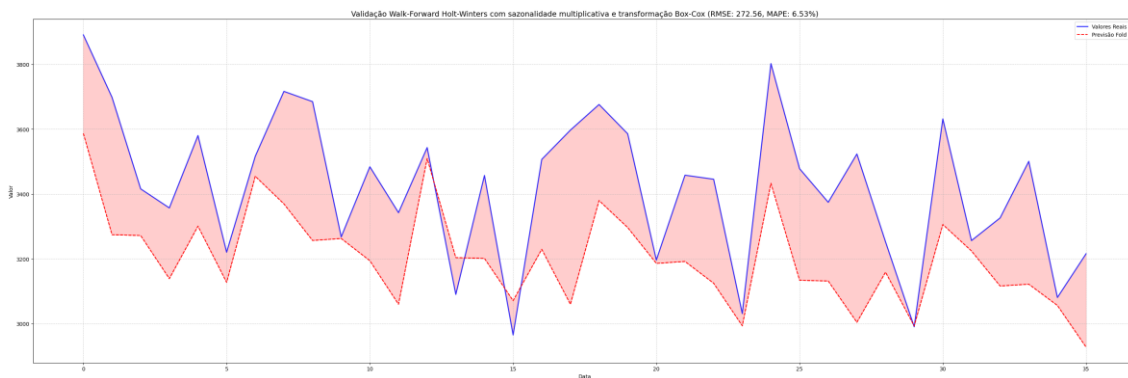


Figura B. 23 – Validação Walk-Forward do modelo HWM

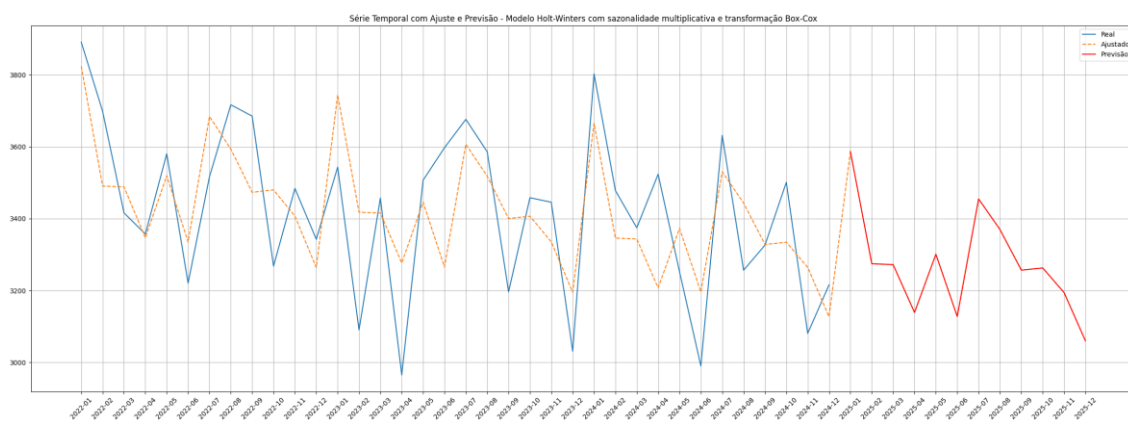


Figura B. 24 – Cronograma da série temporal com os ajustes e a previsão do modelo HWM

```

model_n = NBEATS (
    h=12,
    input_size=36,
    max_steps=200,
    early_stop_patience_steps=3,
    val_check_steps=12,
    learning_rate=1e-3,
    stack_types=['identity'],
    n_blocks=[1],
    mlp_units=[[64, 64]],
    activation='ReLU',
    start_padding_enabled=True
)

```

Figura B. 25 – Modelo N-BEATS Simples

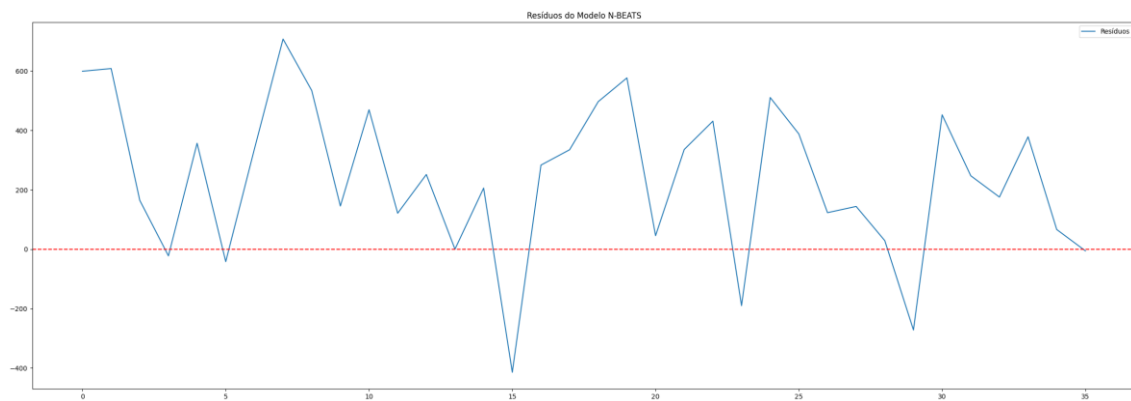


Figura B. 26 – Gráfico dos resíduos do modelo N-BEATS Simples

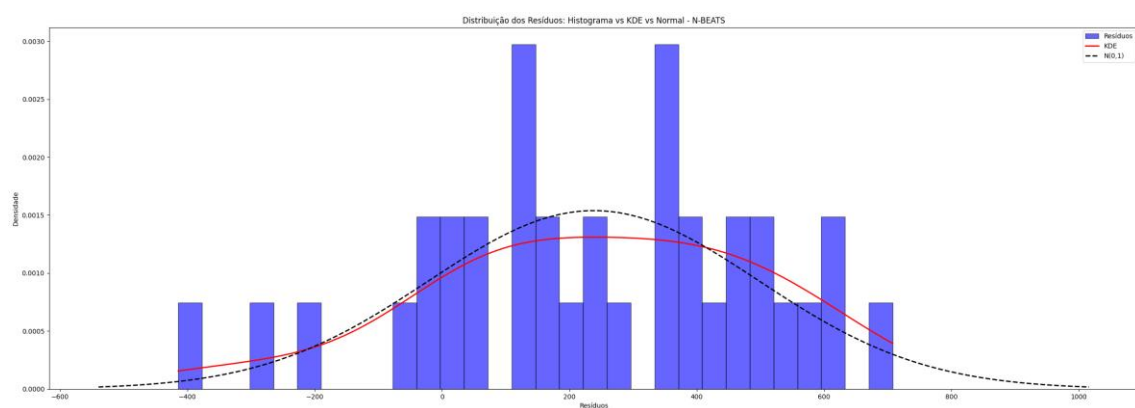


Figura B. 27 - Distribuição dos resíduos do modelo N-BEATS Simples

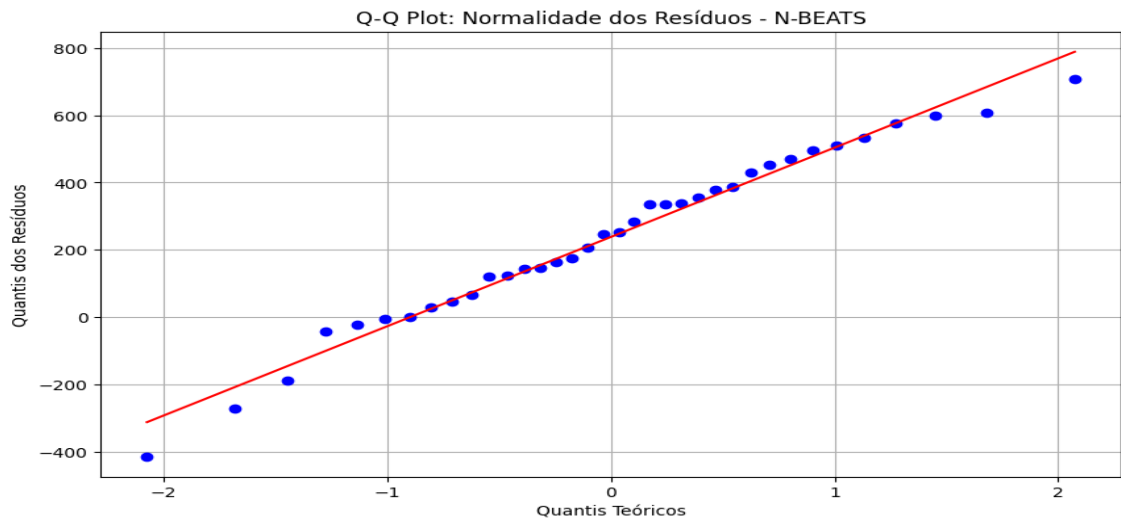


Figura B. 28 - Probabilidade normal (Q-Q plot) do modelo N-BEATS Simples

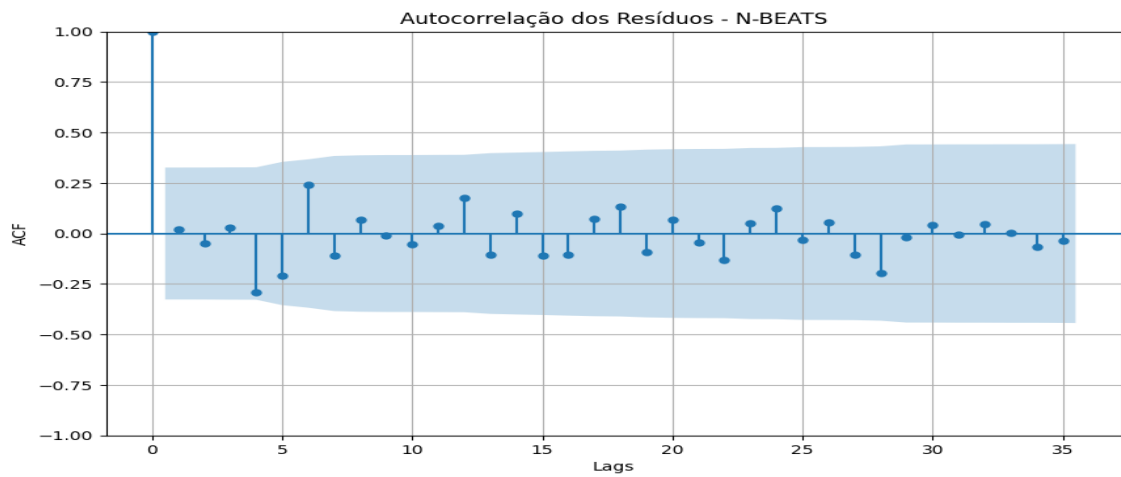


Figura B. 29 - Correlograma dos resíduos do modelo N-BEATS Simples

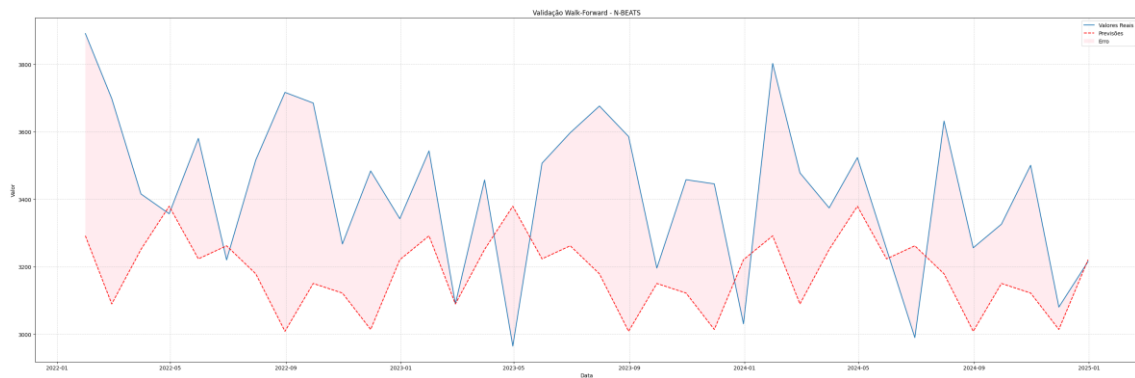


Figura B. 30 – Validação Walk-Forward do modelo N-BEATS Simples

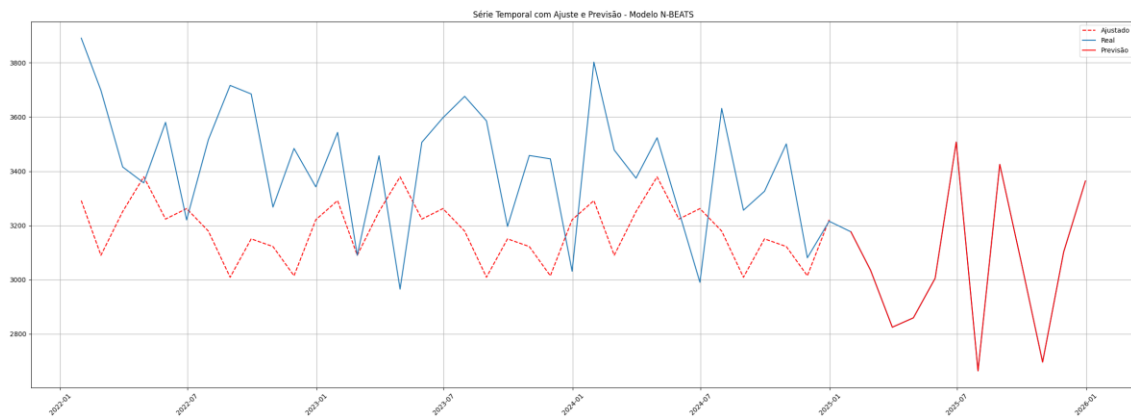


Figura B. 31 – Cronograma da série temporal com os ajustes e a previsão do modelo N-BEATS  
Simples

```
model_f = NBEATS(
    h=12,
    input_size=36,
    max_steps=200,
    early_stop_patience_steps=3,
    val_check_steps=12,
    learning_rate=1e-3,
    stack_types=['trend', 'seasonality'],
    n_blocks=[2, 2],
    n_polynomials=2,
    n_harmonics=2,
    mlp_units=[[64, 64], [64, 64]],
    activation='ReLU',
    start_padding_enabled=True,
    optimizer_kwargs={'weight_decay': 1e-5}, # Parâmetros do otimizador
)
```

Figura B. 32 – Modelo N-BEATS Especializado

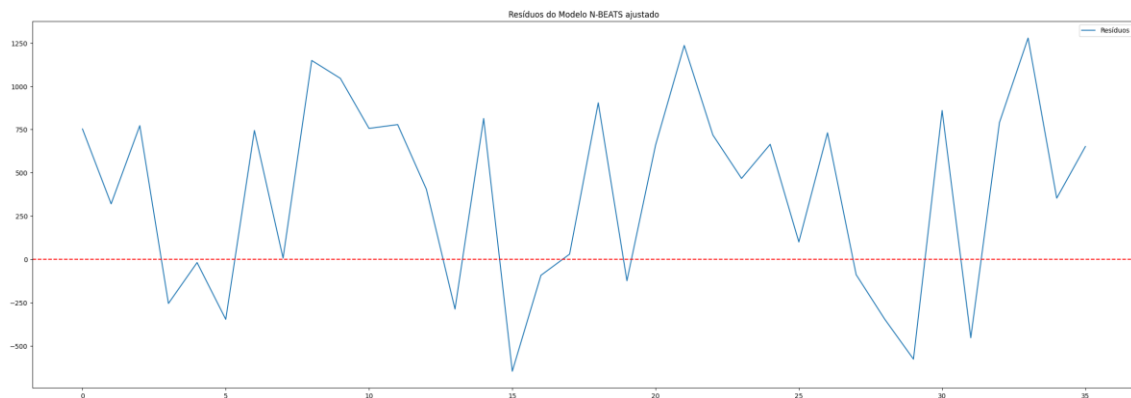


Figura B. 33 – Gráfico de resíduos do modelo N-BEATS Especializado

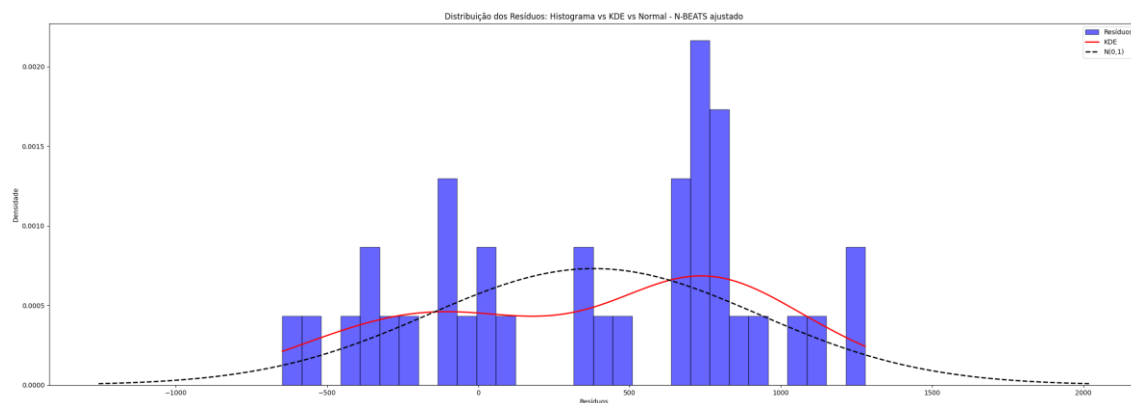


Figura B. 34 - Distribuição dos resíduos do modelo N-BEATS Especializado

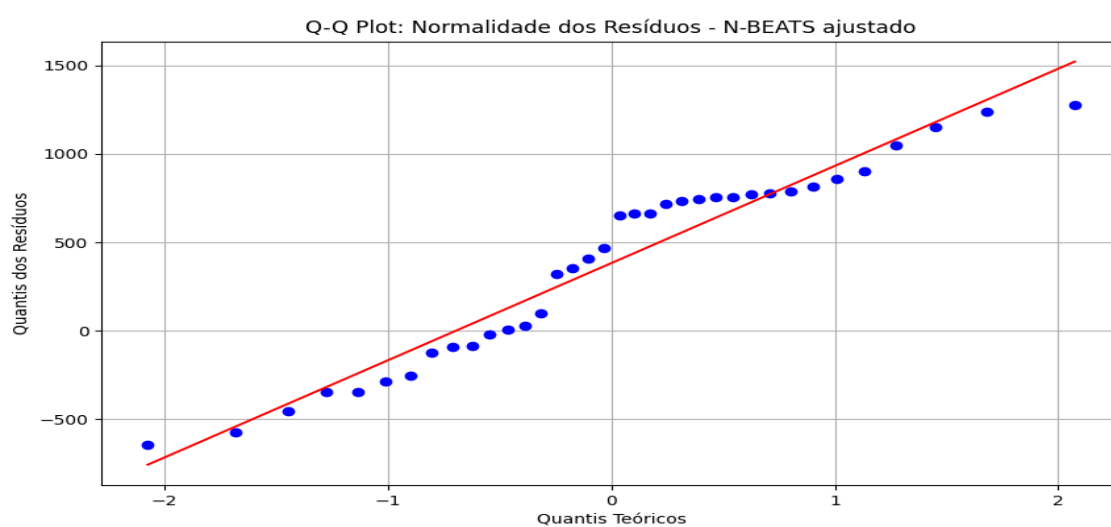


Figura B. 35 - Probabilidade normal (Q-Q plot) do modelo N-BEATS Especializado

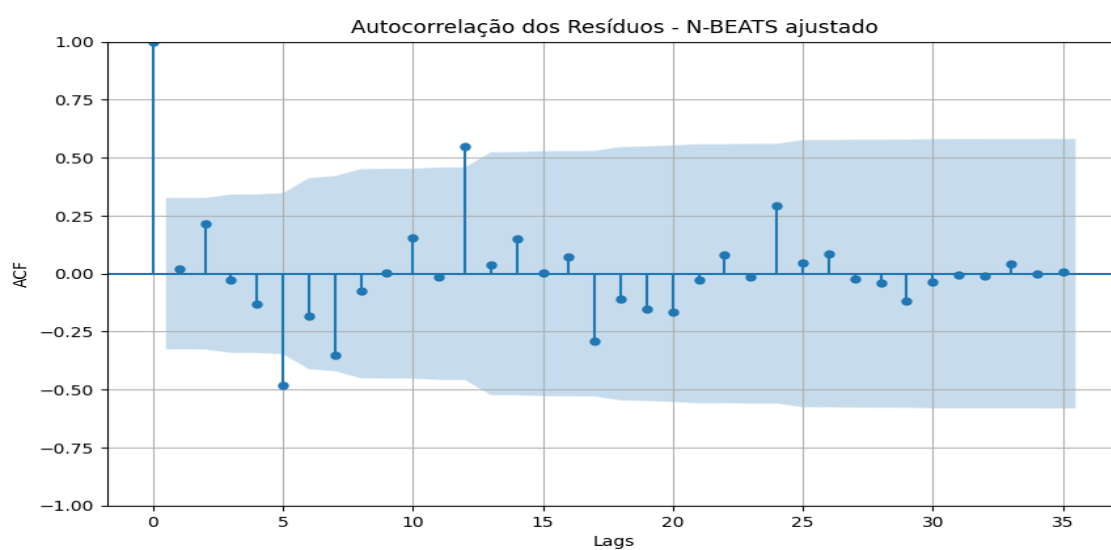


Figura B. 36 - Correlograma dos resíduos do modelo N-BEATS Especializado



Figura B. 37 – Validação Walk-Forward do modelo N-BEATS Especializado

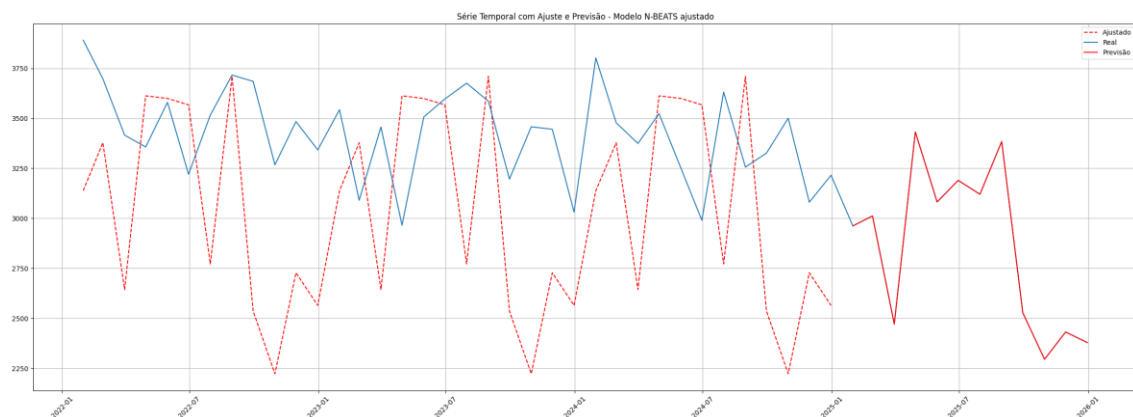


Figura B. 38 – Cronograma da série temporal com os ajustes e a previsão do modelo N-BEATS Especializado

```
estimator = DeepAREstimator(
    freq="ME",          # Frequência dos dados
    prediction_length=12, # Horizonte de previsão
    context_length=36,   # Janela de contexto
    num_layers=4,        # Profundidade da rede
    hidden_size=64,      # Tamanho das camadas ocultas
    batch_size=64,       # Tamanho do batch
    dropout_rate=0.1,
    trainer_kwargs={"max_epochs": 30} # Número de épocas de treino
)
```

Figura B. 39 – Modelo DeepAR Padrão

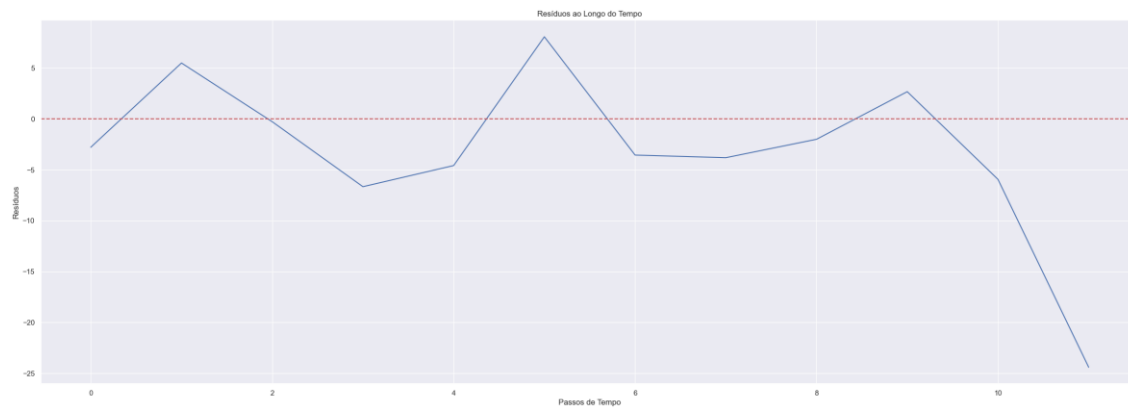


Figura B. 40 – Gráfico de resíduos do modelo DeepAR Padrão

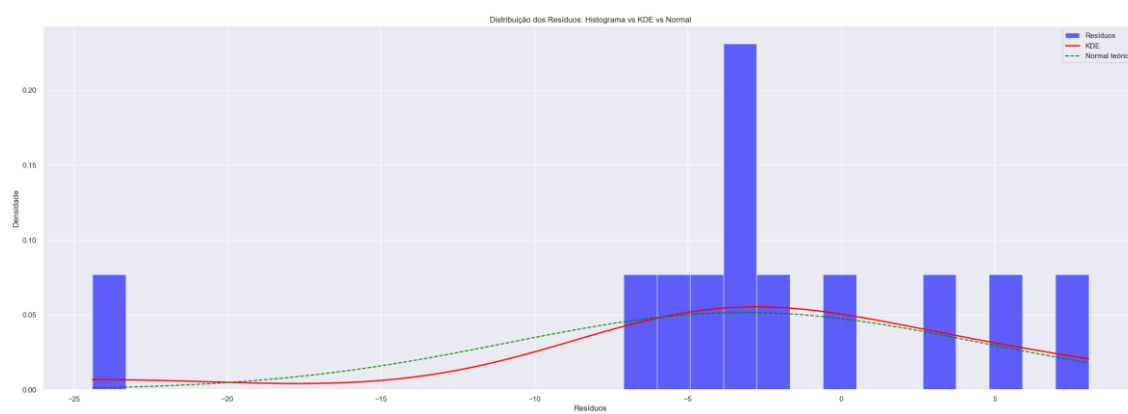


Figura B. 41 - Distribuição dos resíduos do modelo DeepAR Padrão

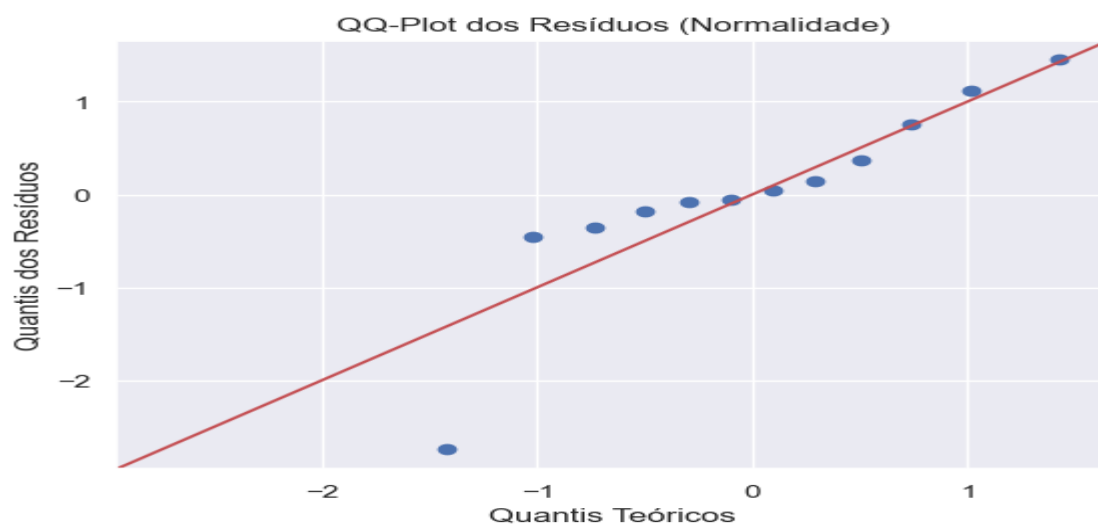


Figura B. 42 - Probabilidade normal (Q-Q plot) do modelo DeepAR Padrão

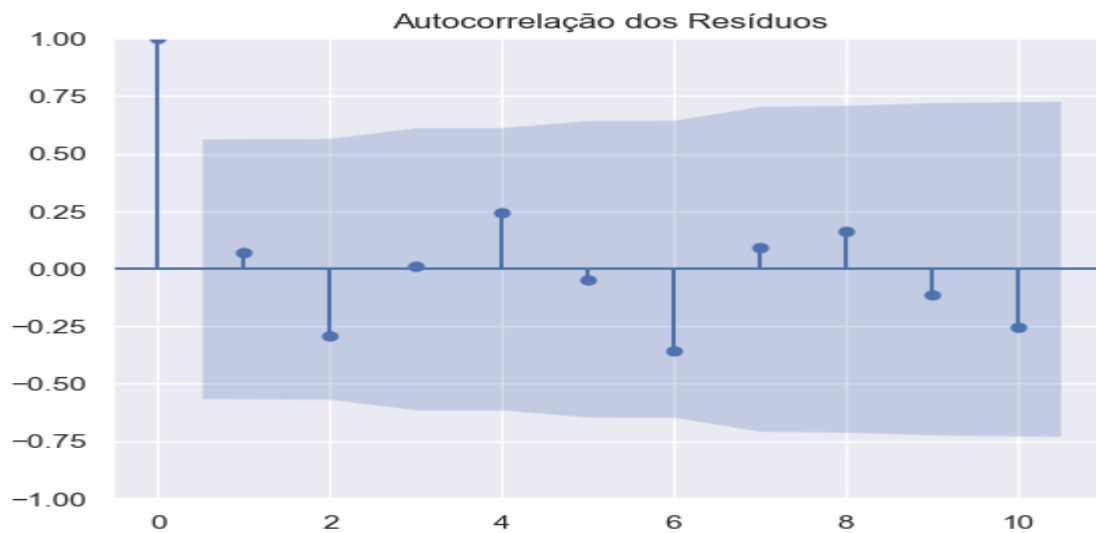


Figura B. 43 - Correlograma dos resíduos do modelo DeepAR Padrão

```
callbacks = [
    EarlyStopping(
        monitor="train_loss",
        patience=10,
        mode="min",
    ),
    LearningRateMonitor()
]

estimator_mm_2 = DeepAREstimator(
    freq='ME',
    prediction_length=12,
    context_length=36,
    num_layers=3,
    hidden_size=64,
    dropout_rate=0.2,
    num_parallel_samples=100,
    lr=1e-3,
    trainer_kwargs={
        "max_epochs": 50,
        #"callbacks": '[<lightning.pytorch.callbacks.early_stopping.EarlyStopping object at 0x000002338BFAF410>',
        '<lightning.pytorch.callbacks.lr_monitor.LearningRateMonitor object at 0x000002338BEA6D20>]',
        "callbacks": callbacks,
        "enable_progress_bar": True,
        "enable_model_summary": True,
        # REMOVIDO: "enable_checkpointing": False
        # Deixar o GluonTS gerir seus próprios checkpoints
    }
)
```

Figura B. 44 - Modelo DeepAR Otimizado

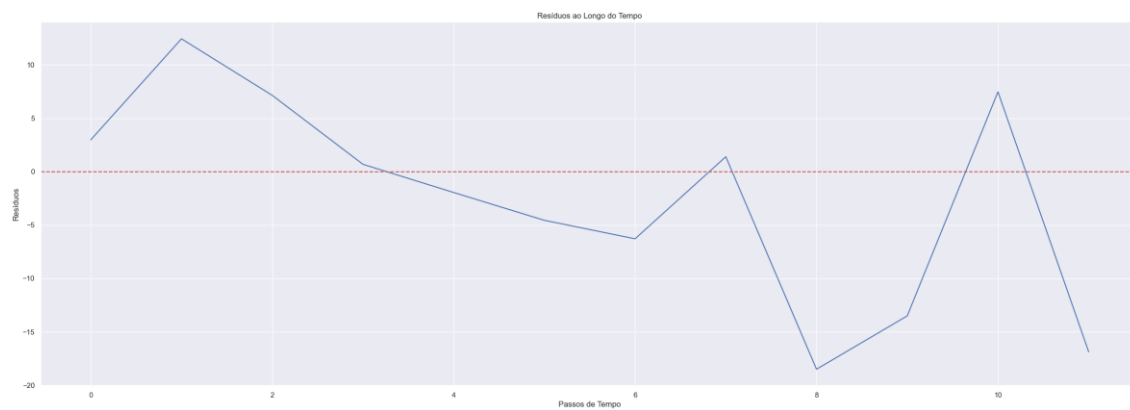


Figura B. 45 – Gráfico de resíduos do modelo DeepAR Otimizado

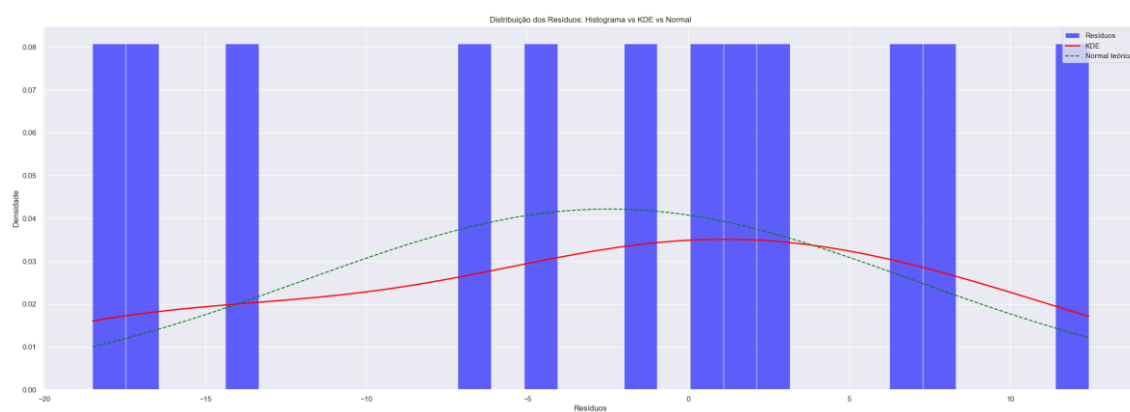


Figura B. 46- Distribuição dos resíduos do modelo DeepAR Otimizado

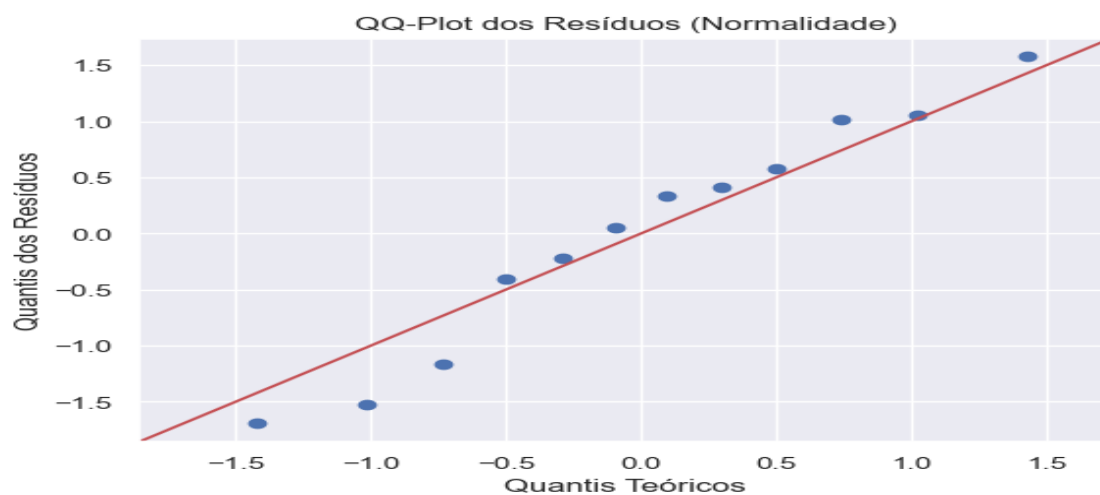


Figura B. 47 - Probabilidade normal (Q-Q plot) do modelo DeepAR Otimizado

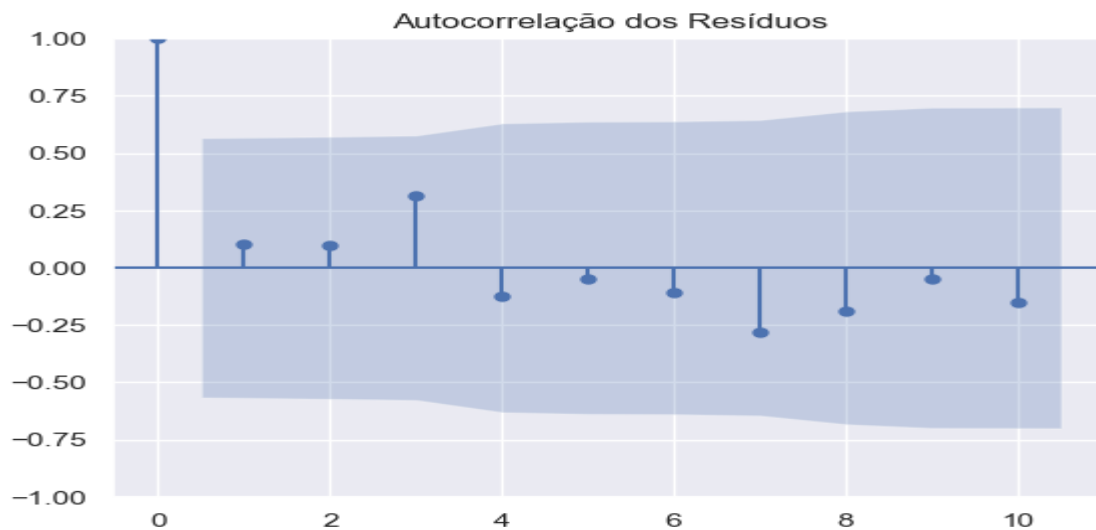


Figura B. 48 - Correlograma dos resíduos do modelo DeepAR Otimizado

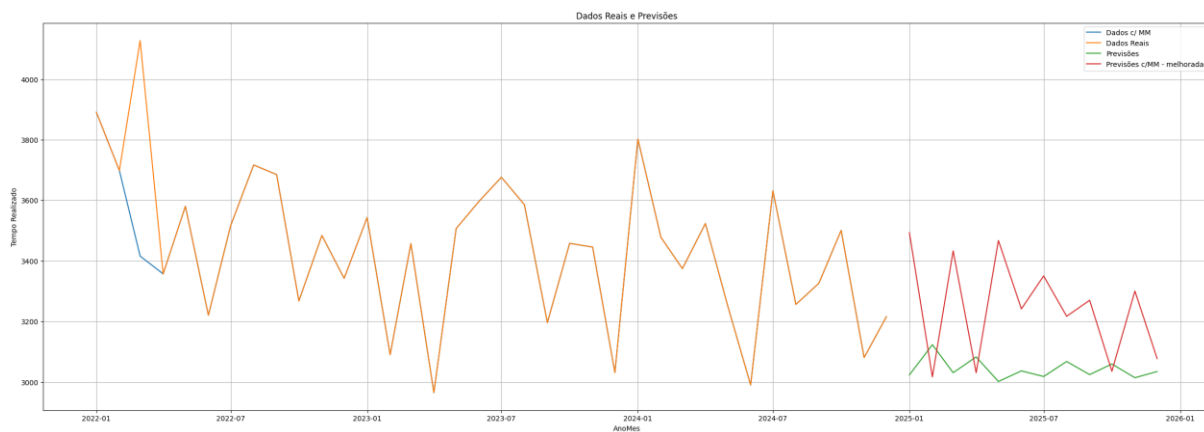


Figura B. 49 - Cronograma da série temporal com as previsões dos modelos DeepAR

```
# Criar o modelo Prophet
m = Prophet()

# Treinar o modelo
m.fit(resumo_mm[['ds', 'y']])

# Criar um DataFrame para as futuras datas
future = m.make_future_dataframe(periods=numero_meses, freq='ME')

# Fazer a previsão
forecast = m.predict(future)
```

Figura B. 50 – Modelo Prophet Padrão

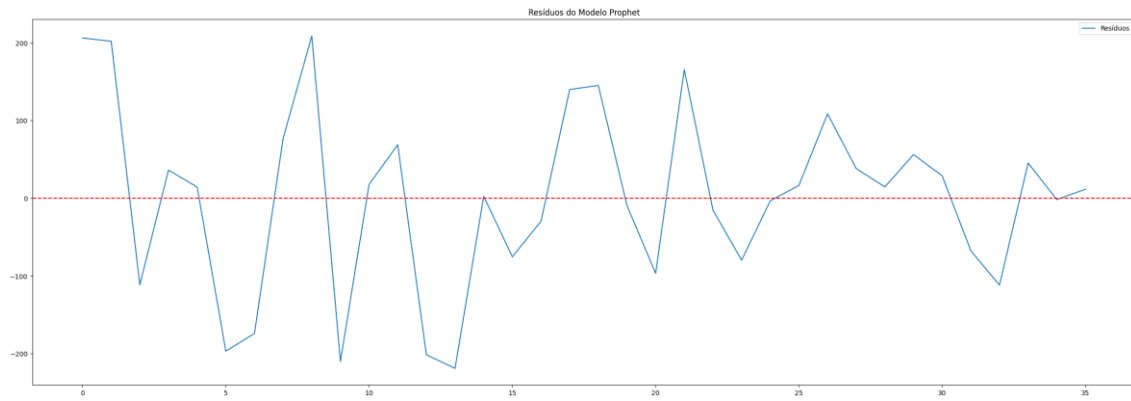


Figura B. 51 – Gráfico de resíduos do modelo Prophet Padrão

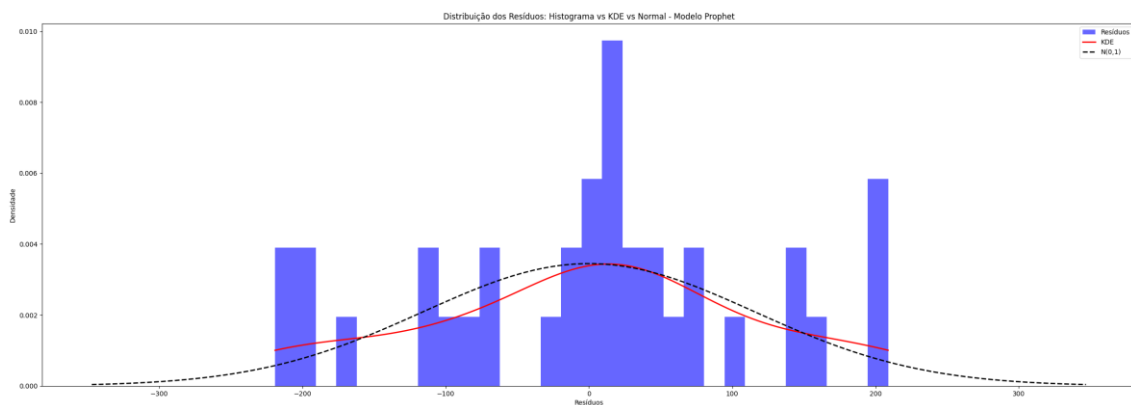


Figura B. 52 - Distribuição dos resíduos do modelo Prophet Padrão

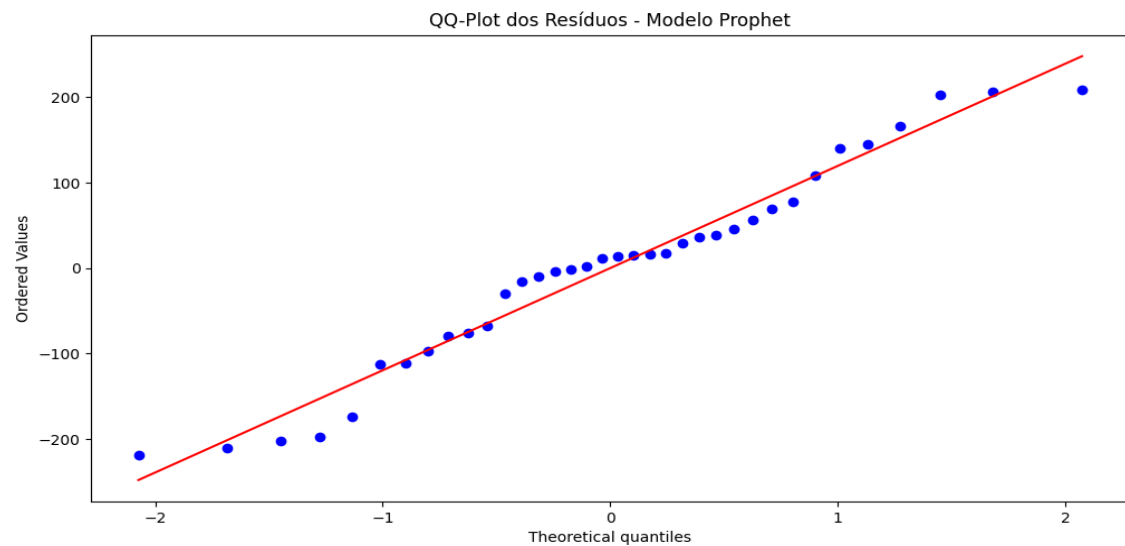


Figura B. 53 - Probabilidade normal (Q-Q plot) do modelo Prophet Padrão

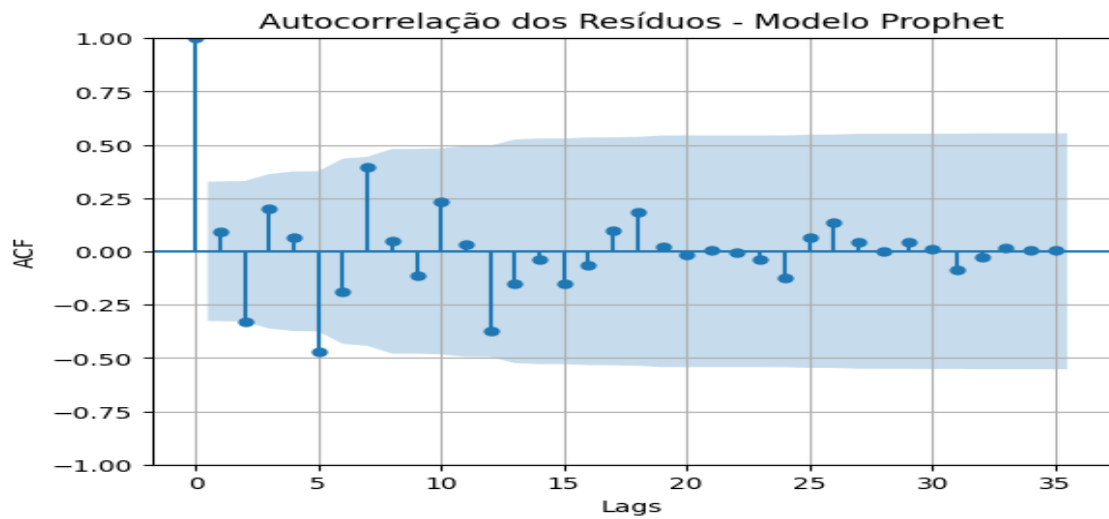


Figura B. 54 - Correlograma dos resíduos do modelo Prophet Padrão

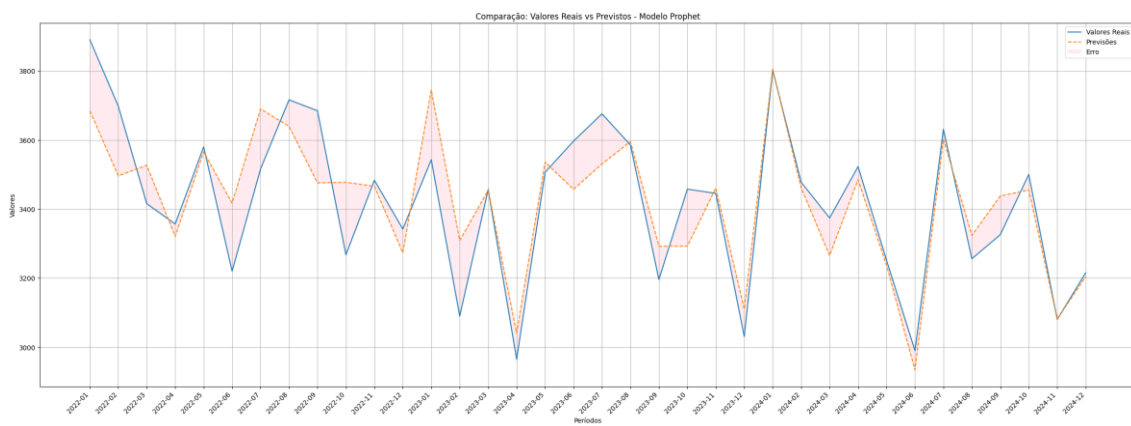


Figura B. 55 – Validação Walk-Forward do modelo Prophet Padrão

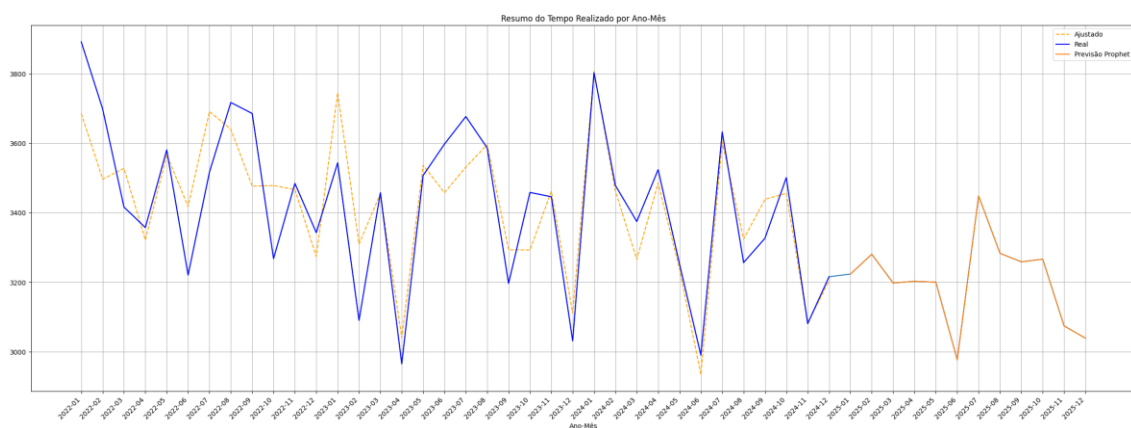


Figura B. 56 – Cronograma da série temporal com os ajustes e a previsão do modelo Prophet Padrão

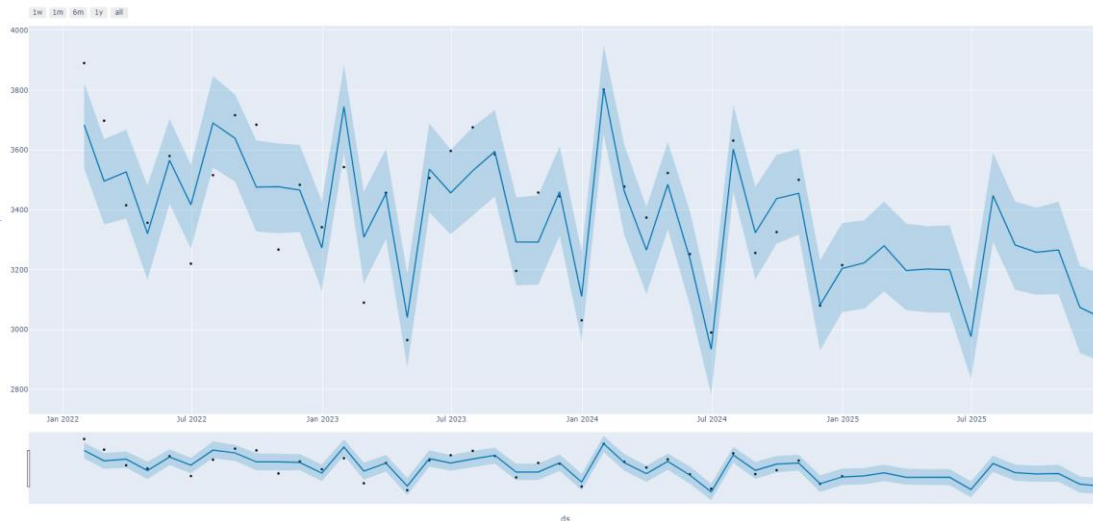


Figura B. 57 – Previsão de séries temporais com intervalos de confiança do modelo Prophet Padrão

```
# Agora a função pode ser usada:
best_scale = validate_changepoint_scale([0.001, 0.01, 0.05, 0.1, 0.5], cv_data)

# changepoint_scales = [0.001, 0.01, 0.05, 0.1, 0.5],
# best_scale = validate_changepoint_scale(changepoint_scales, cv_data),

# Configuração corrigida e mais conservadora
m_corrected = Prophet(
    # Manter crescimento linear (mais estável)
    growth='linear',

    # Sazonalidade aditiva (mais robusta para esta série)
    seasonality_mode='additive',

    # Flexibilidade moderada da tendência
    # changepoint_prior_scale=0.01, # Reduzido para maior estabilidade
    # Teste sistemático de changepoint prior scale
    changepoint_prior_scale=best_scale, # Corrigido

    # Força sazonal moderada
    seasonality_prior_scale=1.0, # Valor padrão mais conservador

    # Parâmetros de incerteza standard
    uncertainty_samples=1000,

    # Configuração temporal básica
    yearly_seasonality=True,
    weekly_seasonality=False,
    daily_seasonality=False,

    # Intervalos de confiança padrão
    interval_width=0.80
)
```

Figura B. 58 – Modelo Prophet Otimizado

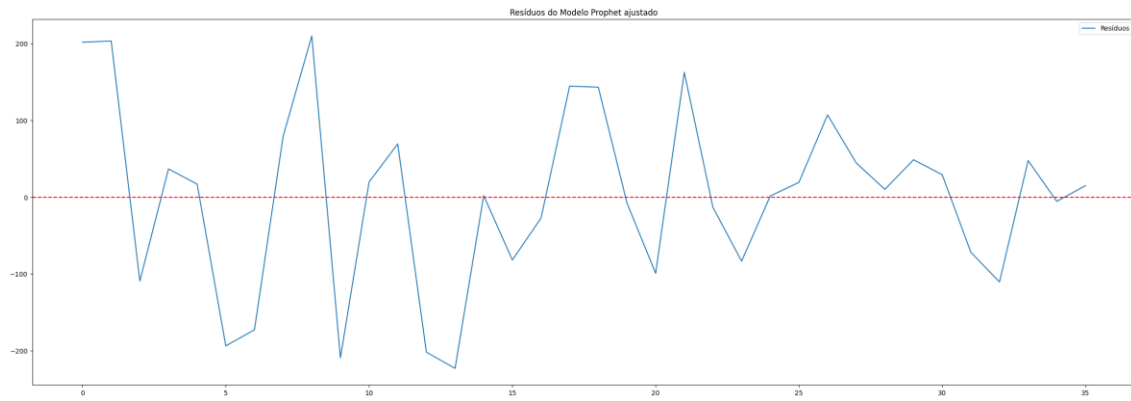


Figura B. 59 – Gráfico de resíduos do modelo Prophet Otimizado

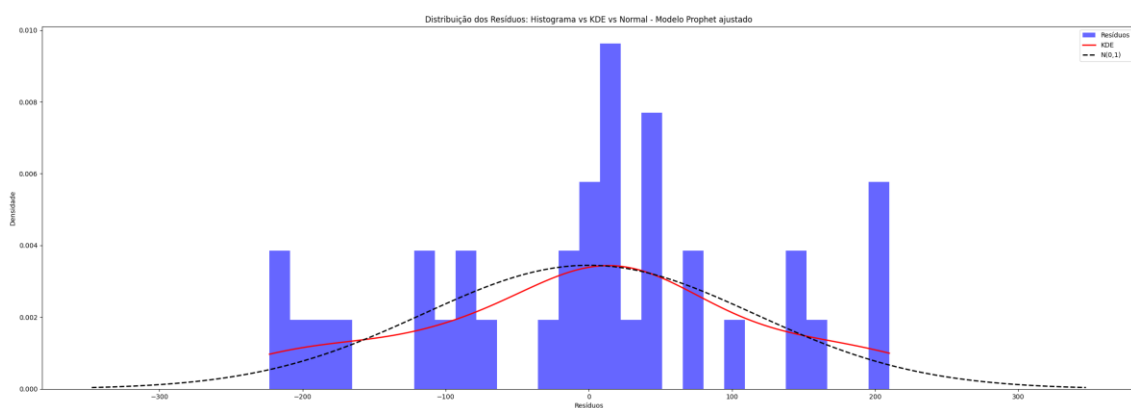


Figura B. 60 - Distribuição dos resíduos do modelo Prophet Otimizado

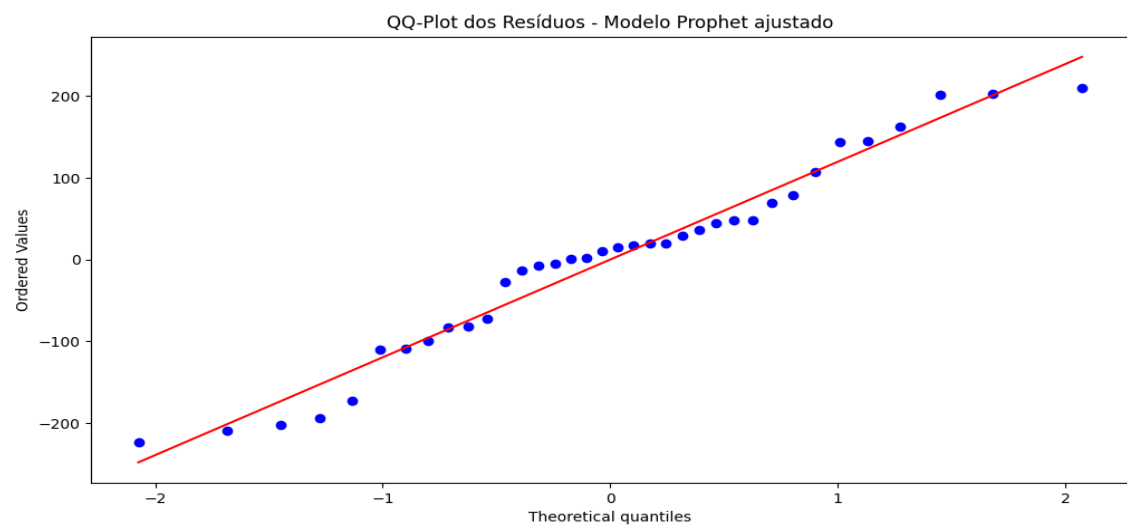


Figura B. 61 - Probabilidade normal (Q-Q plot) do modelo Prophet Otimizado

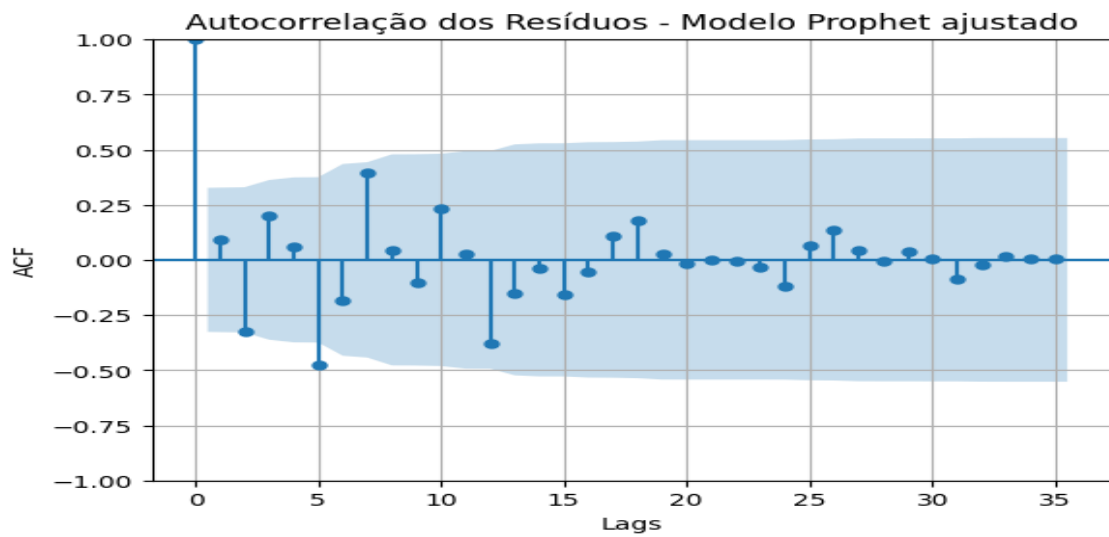


Figura B. 62 - Correlograma dos resíduos do modelo Prophet Otimizado

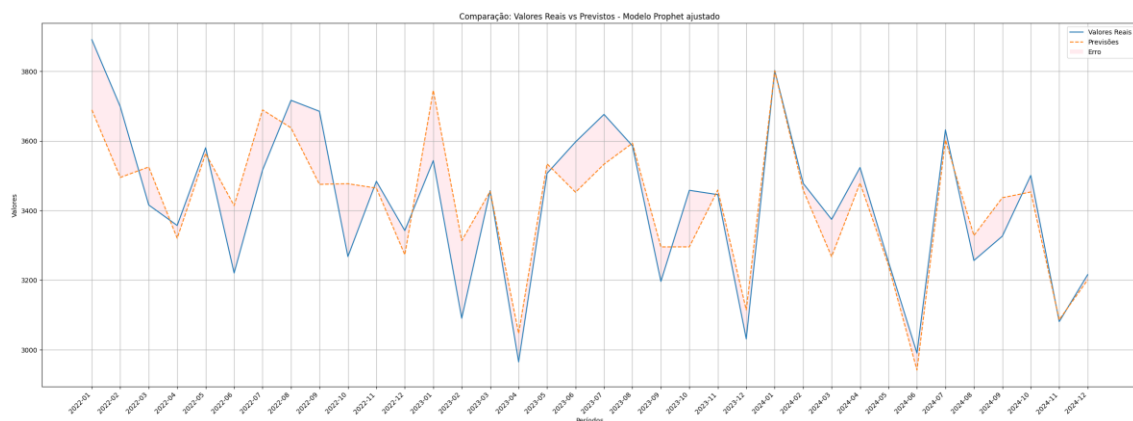


Figura B. 63 – Validação Walk-Forward do modelo Prophet Otimizado

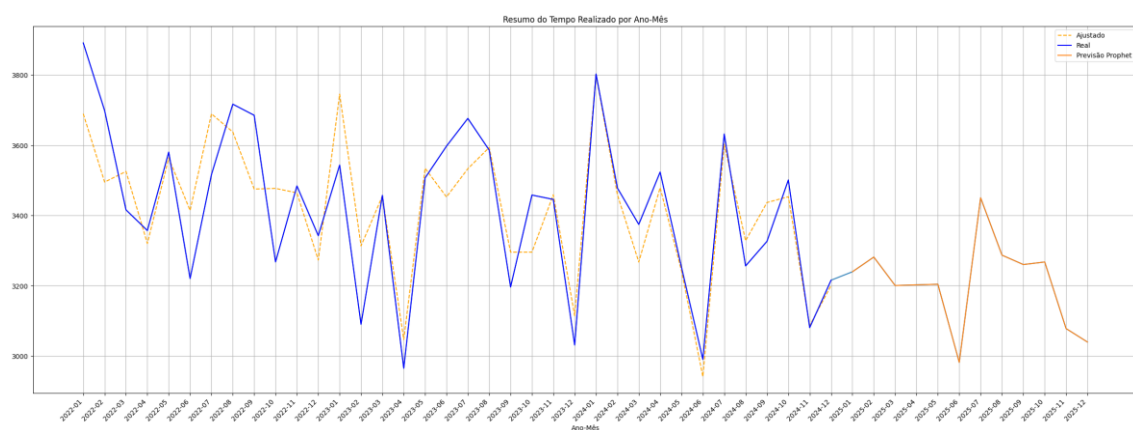


Figura B. 64 – Cronograma da série temporal com os ajustes e a previsão do modelo Prophet Otimizado



Figura B. 65 – Cronograma de previsão de séries temporais com intervalos de confiança do modelo Prophet Otimizado