TREINAMENTO E COMPARAÇÃO DE MODELOS DE PREVISÃO PARA ESTIMATIVA DA DEMANDA DE SERVIÇOS DE DESOBSTRUÇÃO DE REDES DE ESGOTO NA REGIÃO DO PLANO PILOTO, DISTRITO FEDERAL

Carolina Timo Pinheiro de Almeida[(1)]
Gerente de Manutenção na Caesb (Companhia de Saneamento Ambiental do Distrito Federal). Engenheira Civil pela UnB, MBA em Ciência de Dados pela USP, cursando mestrado no programa Profágua pela UnB.
Marco Lúcio do Nascimento[(2)]
Superintendente de Manutenção na Caesb (Companhia de Saneamento Ambiental do Distrito Federal). Engenheiro Civil pelo Centro Universitário do Distrito Federal - UDF, Pós-Graduação em Contabilidade e Controladoria Pública pela Universidade Tuiti do Paraná, cursando mestrado no programa Profágua pela UnB.

Endereço[(1)]**:** Sqsw 306, bloco B. Sudoeste. Brasília – Df – CEP: 70.673-432 – Brasil – Tel: 3312-2171 – *e-mail*: carolinatimo@gmail.com.

RESUMO

Este trabalho apresenta a comparação de modelos de previsão para estimativa da demanda de serviços de desobstrução de redes de esgoto na região do Plano Piloto, Distrito Federal, visando apoiar o planejamento operacional e a elaboração de contratos de manutenção. Foram utilizados dados mensais da Companhia de Saneamento Ambiental do Distrito Federal (CAESB) referentes ao período de janeiro de 2016 a dezembro de 2020. A série temporal foi dividida em conjunto de treinamento (2016 a 2019) e conjunto de validação (2020). Foram avaliados modelos de persistência, média simples, média móvel simples, média móvel exponencialmente ponderada (MMEP12), suavização exponencial de Holt, Holt-Winters, SARIMA e *XGBoost*. O desempenho foi avaliado por meio do Erro Médio Absoluto (MAE), da Raiz do Erro Quadrático Médio (RMSE), do Erro Percentual Médio Absoluto (MAPE) e do teste não paramétrico de Wilcoxon. A MMEP12 apresentou o melhor desempenho para previsão mensal, com MAE de 19,63, RMSE de 26,90 e MAPE de 16,13%. Na análise anual, também apresentou o menor erro percentual (7,19%), seguida pelo *XGBoost* (8,03%) e pela Média Simples (8,41%). O teste de Wilcoxon não identificou diferenças estatisticamente significativas entre a MMEP12 e os modelos Igual2019 e *XGBoost*. Os resultados indicam que métodos estatísticos simples podem apresentar desempenho comparável ao de técnicas mais complexas em séries temporais curtas e relativamente estáveis.

PALAVRAS-CHAVE: previsão de demanda; manutenção de ativos; séries temporais.

INTRODUÇÃO

O planejamento da manutenção constitui uma das atividades mais relevantes para a gestão dos sistemas de abastecimento de água e esgotamento sanitário. Empresas prestadoras desses serviços necessitam estimar antecipadamente a demanda por intervenções corretivas e preventivas para garantir a disponibilidade operacional dos ativos, a continuidade dos serviços prestados à população e a adequada aplicação dos recursos financeiros disponíveis.

No contexto das companhias de saneamento, parcela significativa das atividades de manutenção é executada por empresas terceirizadas contratadas por meio de processos licitatórios. A elaboração desses contratos depende da estimativa prévia dos quantitativos de serviços que deverão ser executados ao longo de sua vigência, uma vez que tais quantitativos compõem o orçamento-base utilizado para definição dos valores contratuais. Erros de previsão podem resultar tanto em insuficiência contratual quanto em superdimensionamento dos contratos, comprometendo a eficiência da gestão dos recursos públicos.

Historicamente, muitas organizações utilizam médias históricas ou estimativas baseadas na experiência dos gestores para definir quantitativos futuros. Embora simples, essas abordagens frequentemente desconsideram a estrutura temporal dos dados, incluindo possíveis padrões de tendência, sazonalidade e autocorrelação. Técnicas de previsão baseadas em séries temporais e aprendizado de máquina surgem como alternativas capazes de explorar essas informações e produzir estimativas mais precisas (MORETTIN; TOLOI, 2018; BOX et al., 2015).

Nos últimos anos, o avanço das ferramentas computacionais e o aumento da disponibilidade de dados operacionais possibilitaram a aplicação de métodos estatísticos e algoritmos de aprendizado de máquina em problemas relacionados à gestão de ativos, manutenção e planejamento operacional. Dentre essas técnicas destacam-se os modelos de suavização exponencial, os modelos da família ARIMA/SARIMA e os algoritmos de *Gradient Boosting*, amplamente utilizados em problemas de previsão devido à sua capacidade de identificar padrões temporais e relações não lineares (FRIEDMAN, 2001; CHEN; GUESTRIN, 2016).

OBJETIVO

Nesse contexto, este trabalho tem como objetivo comparar diferentes modelos de previsão aplicados à demanda de serviços de desobstrução de redes de esgoto da Companhia de Saneamento Ambiental do Distrito Federal (CAESB), buscando identificar técnicas capazes de apoiar o planejamento operacional e a elaboração de estimativas para processos licitatórios de manutenção.

METODOLOGIA

O estudo foi desenvolvido a partir de informações operacionais registradas pela Companhia de Saneamento Ambiental do Distrito Federal (CAESB). Os dados foram obtidos do Sistema de Gerenciamento Comercial (GCOM), utilizado pela companhia para registro e acompanhamento das ordens de serviço executadas nas redes de abastecimento de água e esgotamento sanitário. A extração da base foi realizada em dezembro de 2021.

Foram selecionadas exclusivamente as ordens de serviço relacionadas à desobstrução de redes de esgoto na região administrativa do Plano Piloto, Distrito Federal. Após as etapas de limpeza e tratamento dos dados, os registros foram agregados mensalmente, resultando em uma série temporal uni variada composta pela quantidade de serviços executados entre janeiro de 2016 e dezembro de 2020.

O processo de agregação foi adotado por representar a escala temporal utilizada no planejamento e acompanhamento dos contratos de manutenção da companhia, permitindo alinhar os resultados da pesquisa às necessidades práticas da gestão operacional.

A série histórica foi dividida em dois subconjuntos. Os registros compreendidos entre janeiro de 2016 e dezembro de 2019 foram utilizados para treinamento dos modelos, enquanto os dados referentes ao ano de 2020 foram reservados para validação.

O desempenho dos modelos foi avaliado por meio do Erro Médio Absoluto (*Mean Absolute Error* – MAE), do Erro Percentual Médio Absoluto (*Mean Absolute Percentage Error* – MAPE) e da Raiz do Erro Quadrático Médio (*Root Mean Squared Error* – RMSE).

O MAE mede a magnitude média dos erros de previsão na mesma unidade da variável analisada. O MAPE expressa os erros em termos percentuais, permitindo comparações relativas entre previsões e observações. Já o RMSE penaliza mais fortemente erros de grande magnitude devido à elevação dos resíduos ao quadrado. Para as três métricas, valores menores indicam melhor desempenho preditivo (HYNDMAN; KOEHLER, 2006).

Além das métricas de erro, foi empregado o teste não paramétrico de Wilcoxon para comparação pareada dos erros absolutos produzidos pelos modelos. O teste de postos sinalizados de Wilcoxon é amplamente utilizado para comparar amostras pareadas quando não se pode assumir a normalidade dos dados, sendo particularmente adequado para conjuntos de observações reduzidos, como o utilizado neste estudo (WILCOXON, 1945). O procedimento foi aplicado com o objetivo de verificar se as diferenças observadas entre os desempenhos dos modelos possuíam significância estatística ao nível de 5%, permitindo avaliar se os resultados obtidos poderiam ser atribuídos a diferenças reais de desempenho ou apenas à variabilidade amostral.

REVISÃO BIBLIOGRÁFICA

A previsão de demanda para serviços de manutenção em sistemas de saneamento envolve conhecimentos relacionados à gestão de ativos, análise de séries temporais e métodos estatísticos de previsão. A compreensão dos fundamentos teóricos dessas áreas é essencial para a seleção adequada dos modelos avaliados e para a interpretação dos resultados obtidos. Nesse contexto, esta revisão bibliográfica apresenta os principais conceitos

que sustentam o desenvolvimento do estudo, abordando aspectos da gestão da manutenção em sistemas de saneamento, os fundamentos das séries temporais, os métodos estatísticos tradicionalmente empregados em problemas de previsão e as abordagens de aprendizado de máquina utilizadas na comparação realizada.

Gestão da manutenção em sistemas de saneamento

Os sistemas de abastecimento de água e esgotamento sanitário constituem infraestruturas essenciais para a saúde pública, o desenvolvimento urbano e a preservação ambiental. Em razão de sua extensa distribuição geográfica e da elevada quantidade de componentes físicos envolvidos, esses sistemas demandam atividades permanentes de operação e manutenção. A gestão eficaz desses ativos requer planejamento adequado das intervenções, de forma a assegurar o desempenho operacional, a confiabilidade dos sistemas e a utilização eficiente dos recursos disponíveis (ABNT, 2014).

Segundo Silva Neto e Lima (2002), a manutenção deve ser compreendida como uma função estratégica voltada à maximização da disponibilidade dos ativos e à otimização dos custos operacionais. Nesse contexto, decisões associadas ao planejamento de recursos, contratação de serviços e programação das intervenções dependem diretamente da capacidade de estimar demandas futuras de manutenção.

Nas redes coletoras de esgoto, um dos problemas mais frequentes corresponde à ocorrência de obstruções provocadas pelo descarte inadequado de resíduos sólidos, acúmulo de gordura ou deterioração das condições hidráulicas das tubulações. Essas ocorrências resultam em aumento da demanda por serviços corretivos e representam importante parcela dos custos operacionais das companhias de saneamento, reforçando a necessidade de ferramentas que auxiliem o planejamento e a previsão dessas atividades.

Séries temporais aplicadas à previsão de demanda

Uma série temporal consiste em um conjunto de observações ordenadas cronologicamente cujo objetivo principal é descrever o comportamento de uma variável ao longo do tempo e permitir a realização de previsões futuras (MORETTIN; TOLOI, 2018).

Diversas séries temporais apresentam componentes estruturais como tendência, sazonalidade, ciclos e ruído aleatório (MORETTIN; TOLOI, 2018). A identificação adequada desses componentes constitui etapa fundamental para seleção do método de previsão mais apropriado. Em situações nas quais a série apresenta forte persistência temporal, técnicas relativamente simples podem produzir resultados competitivos. Em contrapartida, séries caracterizadas por estruturas mais complexas frequentemente demandam modelagens mais sofisticadas.

Métodos estatísticos de previsão

Os métodos de suavização exponencial estão entre as técnicas mais utilizadas para previsão de séries temporais (HYNDMAN; ATHANASOPOULOS, 2021). A suavização exponencial simples e a média móvel exponencialmente ponderada atribuem maior peso às observações recentes.

A partir dessas abordagens foram desenvolvidos os modelos de Holt e Holt-Winters. O método de Holt incorpora explicitamente a componente de tendência, enquanto o método de Holt-Winters adiciona também a modelagem da sazonalidade. Esses métodos permanecem amplamente utilizados em aplicações empresariais devido à simplicidade de implementação e bom desempenho em séries temporais de curto e médio prazo (MORETTIN; TOLOI, 2018).

Hyndman e Athanasopoulos (2021) também ressaltam que métodos de suavização exponencial, como os modelos de Holt e Holt-Winters, permanecem amplamente utilizados em aplicações de previsão devido à simplicidade computacional, facilidade de interpretação e bom desempenho em séries temporais de curto e médio prazo. Da mesma forma, os modelos ARIMA e sua extensão sazonal, denominada SARIMA, permitem representar simultaneamente dependências temporais de curto prazo e padrões sazonais recorrentes. A metodologia Box-Jenkins continua sendo uma das abordagens mais difundidas para construção desses modelos (BOX et al., 2015).

Aprendizado de máquina para previsão

O desenvolvimento recente do aprendizado de máquina ampliou o conjunto de ferramentas disponíveis para previsão de demanda. Entre essas técnicas destaca-se o *Gradient Boosting*, proposto por Friedman (2001), baseado na combinação sequencial de árvores de decisão ajustadas para reduzir os erros produzidos pelas iterações anteriores.

Uma das implementações mais difundidas dessa abordagem é o *XGBoost* (*Extreme Gradient Boosting*), desenvolvido por Chen e Guestrin (2016). O método foi concebido como uma evolução dos algoritmos tradicionais de *Gradient Boosting*, incorporando técnicas de regularização, otimização computacional e controle de sobreajuste, o que contribui para elevado desempenho em problemas de regressão e classificação.

Segundo Hyndman e Athanasopoulos (2021), a seleção de modelos de previsão deve considerar não apenas a complexidade da técnica empregada, mas principalmente sua capacidade de generalização em dados não utilizados durante o treinamento. Para esse propósito, recomenda-se a utilização de validação temporal, na qual os modelos são ajustados com observações históricas e avaliados em períodos subsequentes, simulando condições reais de previsão. Os autores destacam ainda a importância do uso de múltiplas métricas de desempenho, uma vez que diferentes indicadores capturam aspectos distintos da qualidade das previsões.

Avaliação e comparação de modelos de previsão

Entre as métricas mais utilizadas destacam-se o Erro Médio Absoluto (MAE), a Raiz do Erro Quadrático Médio (RMSE) e o Erro Percentual Absoluto Médio (MAPE). O MAE mede a magnitude média dos erros na unidade original da variável analisada, o RMSE atribui maior penalização a erros de grande magnitude e o MAPE permite interpretar os erros em termos percentuais, facilitando comparações entre modelos e conjuntos de dados distintos.

A comparação sistemática entre métodos simples e modelos mais sofisticados constitui prática recomendada na literatura de previsão. Em diversas aplicações, especialmente quando se dispõe de séries históricas curtas ou relativamente estáveis, modelos de menor complexidade podem apresentar desempenho semelhante ou até superior ao de técnicas mais avançadas, reforçando a necessidade de avaliações empíricas baseadas em métricas de erro e procedimentos formais de validação.

RESULTADO E ANÁLISE DE DADOS

A base de dados utilizada neste estudo foi extraída em 1º de dezembro de 2021 do sistema de gerenciamento de ordens de serviço (GCOM) da Companhia de Saneamento Ambiental do Distrito Federal (CAESB). Esse sistema registra todas as ordens de serviço executadas nas redes de abastecimento de água e esgotamento sanitário do Distrito Federal, contendo informações operacionais referentes à localização, datas de abertura e encerramento, situação da ordem de serviço e demais atributos administrativos.

Após a extração dos registros operacionais, os dados foram agregados mensalmente, resultando em uma série temporal composta pela quantidade de ordens de serviço executadas em cada mês.

A Figura 1 apresenta um resumo da estrutura do conjunto de dados obtido por meio do comando *info()* da biblioteca *Pandas*. Observa-se que a base é composta por 23 atributos, incluindo variáveis numéricas, datas e campos categóricos, totalizando o histórico das intervenções de desobstrução de redes de esgoto registradas pela companhia.

Figura 1 - Informações das colunas da planilha utilizada como banco de dados do trabalho.

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 8837 entries, 10582 to 1746
Data columns (total 23 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   Código OS                       8837 non-null   int64
 1   (Executado) Cód. Tipo Serviço Mat.  8837 non-null   int64
 2   Tipo Status OS                  8837 non-null   object
 3   RA                              8837 non-null   object
 4   Serviço Material                8837 non-null   object
 5   Sigla Tipo de Serviço           8837 non-null   object
 6   Prazo Execução                  8837 non-null   int64
 7   Data Abertura                   8837 non-null   datetime64[ns]
 8   Data Execução Início            8836 non-null   datetime64[ns]
 9   Data Execução Fim               8836 non-null   datetime64[ns]
 10  Data Previsão Baixa             8837 non-null   datetime64[ns]
 11  Data Atendimento                8826 non-null   datetime64[ns]
 12  Desc. Serviço Solicitado        8837 non-null   object
 13  Data Ligação Água               8827 non-null   datetime64[ns]
 14  Data Ligação Esgoto             8656 non-null   datetime64[ns]
 15  Tipo Situação                   8837 non-null   object
 16  Tipo Vazamento                  8837 non-null   object
 17  Ind. Danificado                 7091 non-null   object
 18  Inscrição Imovel                8837 non-null   int64
 19  Latitude                        7403 non-null   float64
 20  Longitude                       7403 non-null   float64
 21  Motivo Não Exec.                8837 non-null   object
 22  Data Abertura_MES               8837 non-null   object
dtypes: datetime64[ns](7), float64(2), int64(4), object(10)
memory usage: 1.6+ MB
```

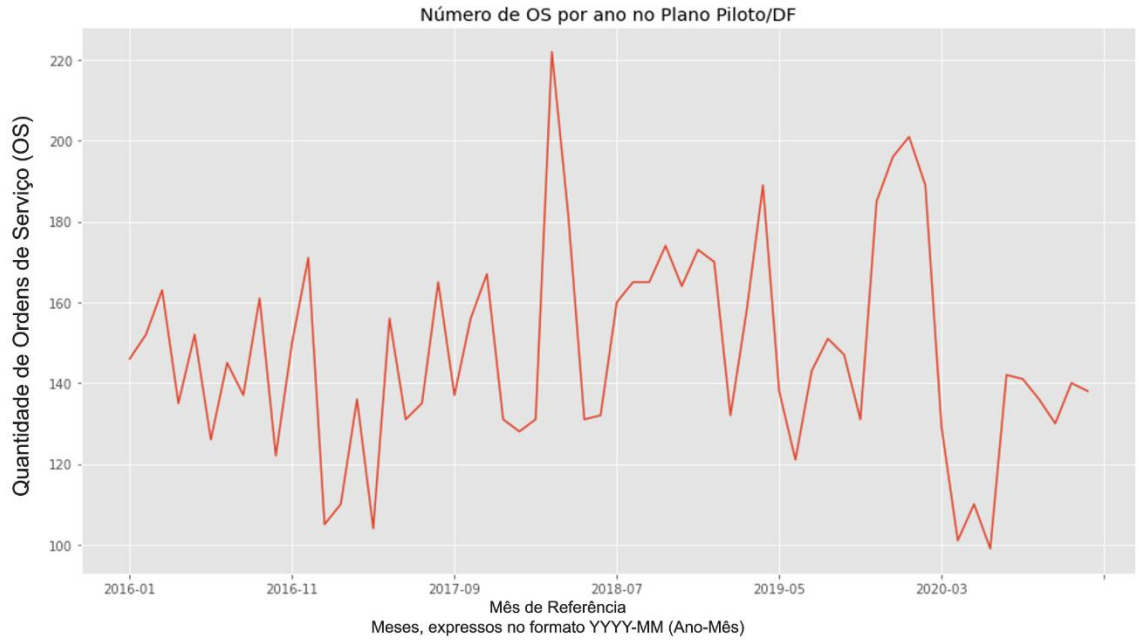
Fonte: Elaboração própria.

Com a etapa de pré-processamento, foram removidos os registros referentes ao ano de 2015. Durante a análise exploratória verificou-se que esse período apresentava quantidade reduzida de registros em comparação aos anos subsequentes, indicando provável incompletude da base ou diferenças no processo de registro das ordens de serviço. A permanência desses dados poderia introduzir vieses na modelagem da série temporal.

Após a limpeza da base, os registros foram agregados mensalmente utilizando a função *groupby(),* produzindo uma série temporal composta pelo número de serviços executados em cada mês entre janeiro de 2016 e dezembro de 2020.

A Figura 2 apresenta a evolução temporal das ordens de serviço de desobstrução de redes de esgoto na região do Plano Piloto.

Figura 2 - Gráfico com o número de ordens de serviço de desobstrução de esgoto na região do Plano Piloto no Distrito Federal nos anos de 2016 a 2020.



Fonte: Elaboração própria.

Visualmente observa-se que a série apresenta oscilações relativamente moderadas ao longo do período analisado, sem tendência crescente ou decrescente pronunciada. Apesar das flutuações mensais, os valores permanecem dentro de uma faixa relativamente estável, sugerindo um comportamento compatível com séries estacionárias.

A estatística descritiva da série é apresentada na Figura 3.

Figura 3 - Descrição dos dados de OS (quantidade, média, desvio padrão, mínimo, máximo e quartis)

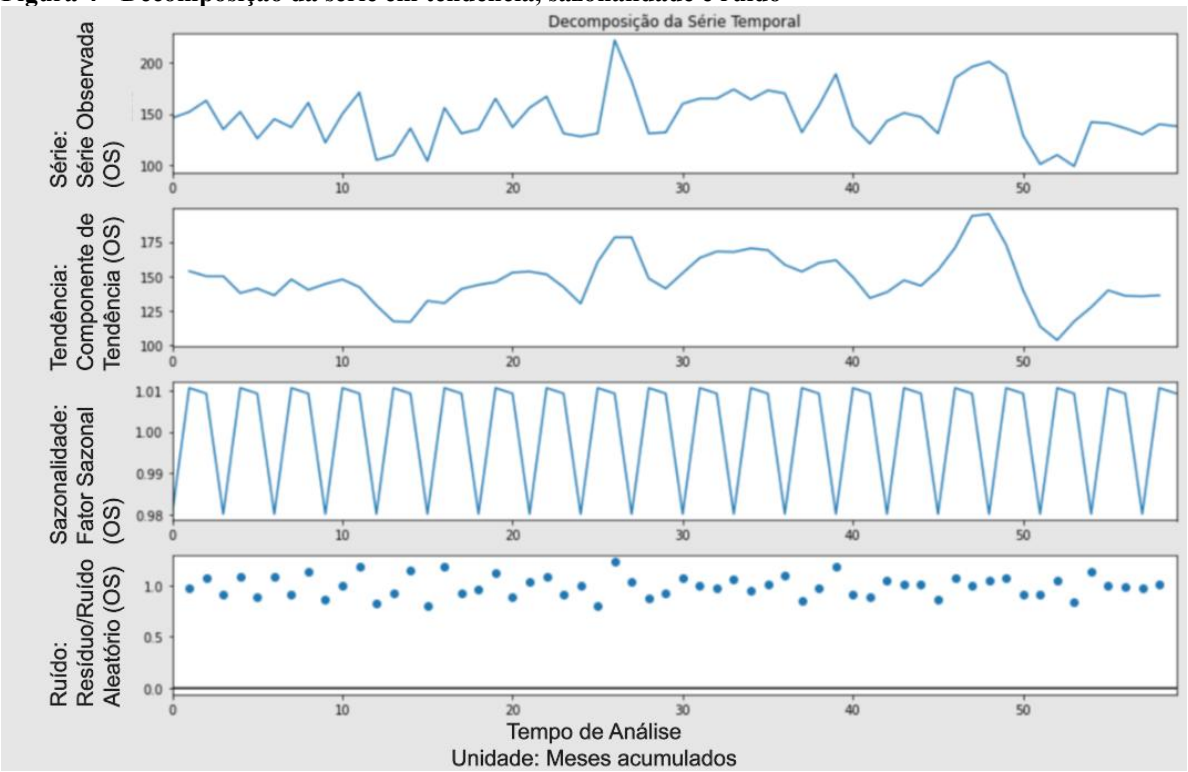
```
count     60.000000
mean     147.283333
std       25.552931
min       99.000000
25%      131.000000
50%      142.500000
75%      164.250000
max      222.000000
Name: Código OS, dtype: float64
```

Fonte: Elaboração própria.

Os resultados indicam média mensal de aproximadamente 147 ordens de serviço, mediana de 142, valor mínimo de 99 e máximo de 222 ocorrências. A proximidade entre média e mediana sugere ausência de forte assimetria na distribuição dos dados. O desvio-padrão relativamente reduzido em relação à média evidencia que a variabilidade da série permanece controlada, característica favorável para modelos de previsão baseados em dependência temporal.

A decomposição da série temporal foi realizada considerando período sazonal de doze meses, compatível com a frequência mensal dos dados e com a avaliação de possíveis padrões anuais na série. Os componentes obtidos são apresentados na Figura 4.

Figura 4 - Decomposição da série em tendência, sazonalidade e ruído



Fonte: Elaboração própria.

A decomposição evidencia que a tendência permanece praticamente constante durante todo o período analisado, reforçando a percepção obtida na inspeção visual da série original. O componente sazonal apresenta oscilações discretas, sugerindo a possível existência de um comportamento anual, porém de pequena magnitude quando

comparado à variabilidade total da série. Deve-se destacar que a identificação desse padrão sazonal é limitada pelo reduzido número de ciclos completos disponíveis (cinco anos), razão pela qual sua interpretação deve ser realizada com cautela.

A estacionariedade da série foi avaliada por meio do teste Dickey-Fuller Aumentado (DICKEY; FULLER, 1979). Os resultados encontram-se apresentados na Figura 5.

Figura 5 - Resultado da aplicação do teste de Dickey Fuller na série

```
ADF Statistic: -5.128772
p-value: 0.000012
Critical Values:
        1%: -3.546
        5%: -2.912
        10%: -2.594
```

Fonte: Elaboração própria.

O teste apresentou valor de *p-value* inferior a 0,05, permitindo rejeitar a hipótese nula de presença de raiz unitária. Dessa forma, conclui-se que a série pode ser considerada estacionária, não havendo evidências estatísticas de tendência estocástica que justifiquem transformações adicionais.

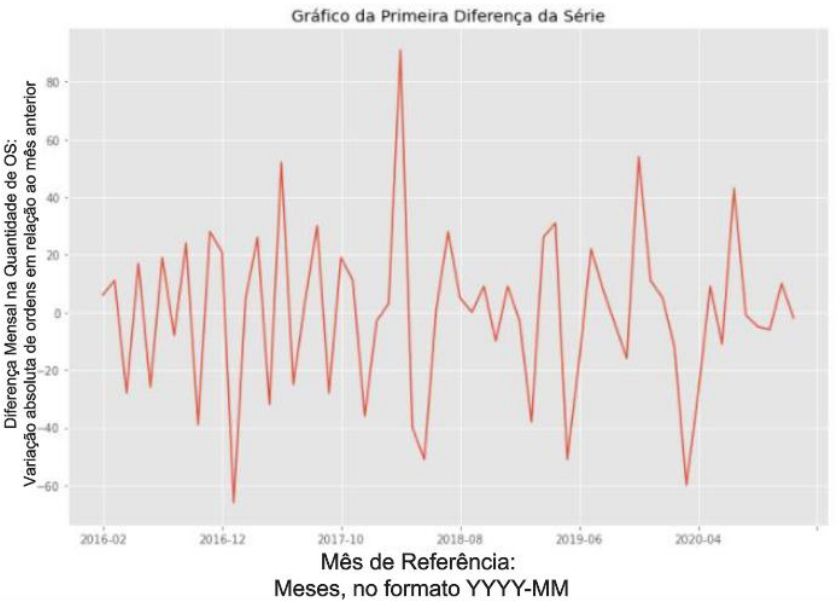
Como procedimento complementar, foi aplicada a primeira diferenciação da série, seguida da repetição do teste ADF. Os resultados são apresentados nas Figuras 6 e 7.

Figura 6 - Resultado da aplicação do teste de Dickey Fuller para a primeira diferença da série

```
ADF Statistic: -3.648581
p-value: 0.004897
Critical Values:
        1%: -3.575
        5%: -2.924
        10%: -2.600
```

Fonte: Elaboração própria.

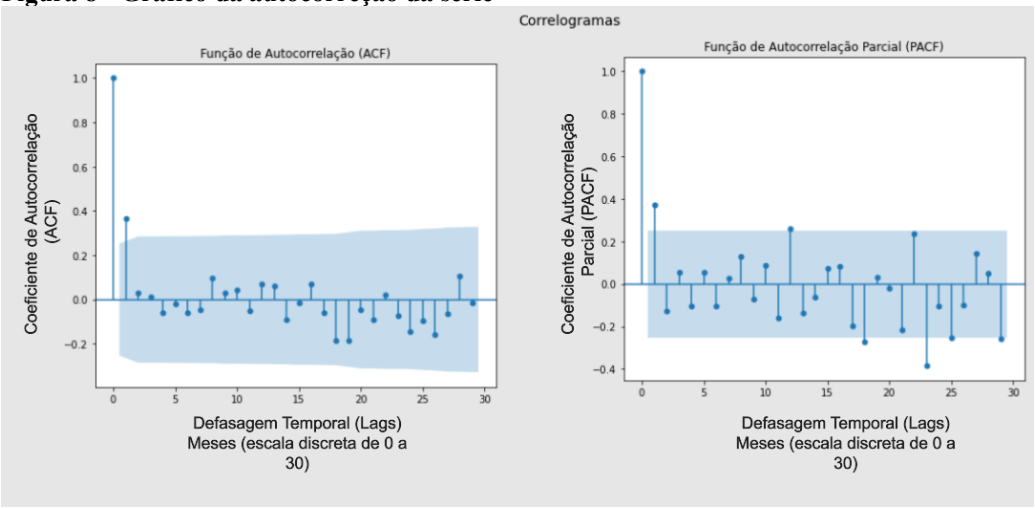
Figura 7 - Gráfico da primeira diferença da série



Os resultados permaneceram praticamente inalterados após a diferenciação, indicando que a série já apresentava comportamento estacionário em sua forma original. Dessa forma, conclui-se que procedimentos adicionais de diferenciação não proporcionam ganhos relevantes para a modelagem dos dados.

A Figura 8 apresenta o correlograma da Função de Autocorrelação (ACF), utilizado para investigar a dependência temporal entre as observações.

Figura 8 - Gráfico da autocorreção da série



Fonte: Elaboração própria.

Observa-se que alguns coeficientes ultrapassam os limites do intervalo de confiança de 95%, indicando que a série não possui comportamento puramente aleatório. A presença dessas autocorrelações confirma a existência de dependência temporal entre observações consecutivas, característica fundamental para a aplicação de modelos de previsão baseados em séries temporais, como o SARIMA e os métodos de suavização exponencial.

Concluída a etapa exploratória, foram ajustados os modelos de previsão utilizando os dados compreendidos entre janeiro de 2016 e dezembro de 2019 como conjunto de treinamento, sendo o ano de 2020 reservado exclusivamente para validação. Essa estratégia permitiu avaliar a capacidade de generalização dos modelos em observações não utilizadas durante o processo de ajuste, simulando uma situação real de previsão futura. Foram selecionados métodos com diferentes níveis de complexidade, desde abordagens estatísticas tradicionais até algoritmos de aprendizado de máquina, possibilitando uma comparação abrangente entre técnicas amplamente utilizadas em problemas de previsão de demanda.

Inicialmente foi avaliado o modelo de persistência, denominado neste estudo como *igual2019*, no qual os valores observados em 2019 foram utilizados diretamente como previsão para os meses correspondentes de 2020.

A Média Móvel Simples de 12 períodos (MMS-12) foi calculada por meio da média aritmética móvel das doze observações mensais anteriores. Já a Média Móvel Exponencialmente Ponderada de 12 períodos (MMEP12) foi obtida utilizando suavização exponencial com parâmetro *span* igual a 12, atribuindo maior peso às observações mais recentes da série. Dessa forma, ambos os métodos incorporam informações do comportamento histórico recente da demanda, porém com diferentes mecanismos de ponderação das observações passadas.

Em seguida foram ajustados os modelos de suavização exponencial de Holt e Holt-Winters. O método de Holt foi empregado para capturar possíveis componentes de tendência presentes na série, enquanto o modelo de Holt-Winters incorporou simultaneamente componentes de tendência e sazonalidade anual, considerando periodicidade de 12 meses. Ambos os métodos são amplamente utilizados em aplicações de previsão devido à sua simplicidade computacional e boa capacidade de adaptação a séries temporais de curto e médio prazo.

Para representar a classe dos modelos estatísticos baseados em autocorrelação temporal foi utilizado um modelo da família SARIMA (*Seasonal AutoRegressive Integrated Moving Average*). Diferentes combinações de parâmetros foram avaliadas durante a etapa de modelagem, sendo selecionada uma estrutura autorregressiva e de média móvel com termo constante, compatível com um modelo SARIMA (1,0,1), para geração das previsões do período de validação. O ajuste foi realizado utilizando exclusivamente os dados do período de treinamento.

Como representante das técnicas de aprendizado de máquina foi utilizado o algoritmo *XGBoost (Extreme Gradient Boosting)*. O modelo foi implementado para regressão utilizando a função objetivo *reg:squarederror*, 1000 estimadores (*n_estimators* = 1000), profundidade máxima das árvores (*max_depth* = 3), taxa de aprendizado (*learning_rate* = 0,05), fração de amostragem dos registros (*subsample* = 0,8), fração de amostragem das variáveis (*colsample_bytree* = 0,8) e semente aleatória (*random_state* = 42). Essa configuração foi adotada para os experimentos realizados no estudo.

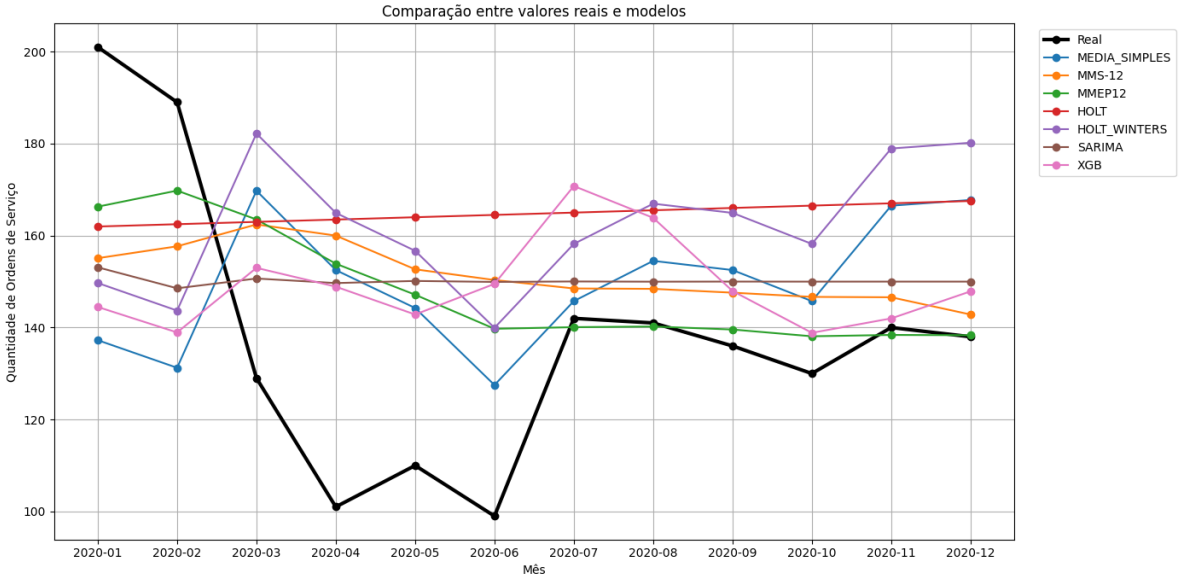
A Figura 9 e Figura 10 apresentam os valores previstos por cada um dos modelos para os meses do ano de 2020, possibilitando uma análise visual do comportamento das diferentes técnicas frente aos valores observados.

Figura 9 - Resultado das previsões para o ano de 2020

data	2020	igual2019	MEDIA_ SIMPLES	MMS-12	MMEP12	HOLT	HOLT_ WINTER	SARIMA	XGB
2020-01	201	170	137,25	155,0833	166,2858	161,9699	149,688	153,098	144,4771
2020-02	189	132	131,25	157,6667	169,7803	162,4746	143,687	148,538	138,9382
2020-03	129	158	169,75	162,4167	163,5064	162,9793	182,188	150,6624	152,979
2020-04	101	189	152,5	160	153,89	163,484	164,938	149,6727	148,8909
2020-05	110	138	144,25	152,6667	147,1377	163,9887	156,687	150,1338	142,8363
2020-06	99	121	127,5	150,3333	139,7319	164,4935	139,937	149,919	149,4941
2020-07	142	143	145,75	148,5	140,0809	164,9982	158,187	150,0191	170,7752
2020-08	141	151	154,5	148,4167	140,2223	165,5029	166,937	149,9725	163,8609
2020-09	136	147	152,5	147,5833	139,5727	166,0076	164,937	149,9942	147,9493
2020-10	130	131	145,75	146,6667	138,1	166,5124	158,188	149,9841	138,8542
2020-11	140	185	166,5	146,5833	138,3923	167,0171	178,938	149,9888	141,9852
2020-12	138	196	167,75	142,8333	138,3319	167,5218	180,188	149,9866	147,8676

Fonte: Elaboração própria.

Figura 10 – Comparação entre a série real e os modelos.



Fonte: Elaboração própria.

Complementarmente, a Figura 11 apresenta os indicadores de desempenho utilizados na avaliação quantitativa dos modelos, incluindo o Erro Médio Absoluto (MAE), a Raiz do Erro Quadrático Médio (RMSE) e o Erro Percentual Médio Absoluto (MAPE). A utilização conjunta dessas métricas permite avaliar simultaneamente a magnitude média dos erros, a ocorrência de desvios extremos e a precisão relativa das previsões em termos percentuais, proporcionando uma comparação mais abrangente do desempenho dos modelos avaliados

Figura 11 – MAE, RMSE e MAPE

Modelo	MAE	RMSE	MAPE
igual2019	31,75	40,49177	24,47445
MEDIA_SIMPLES	31,85417	36,51619	23,47709
MMS-12	26,4375	32,55669	21,14688
MMEP12	19,62576	26,90443	16,13417
HOLT	37,67176	40,26408	30,10912
HOLT_WINTERS	40,14578	42,15182	30,74755
SARIMA	26,89142	31,56606	20,96966
XGB	28,83979	33,99824	21,84269

Fonte: Elaboração própria.

Observa-se que a Média Móvel Exponencialmente Ponderada de 12 períodos (MMEP12) apresentou o melhor desempenho entre os modelos avaliados, alcançando os menores valores de MAE (19,63), RMSE (26,90) e MAPE (16,13%). Esses resultados indicam maior capacidade de aproximação dos valores observados ao longo do período de validação.

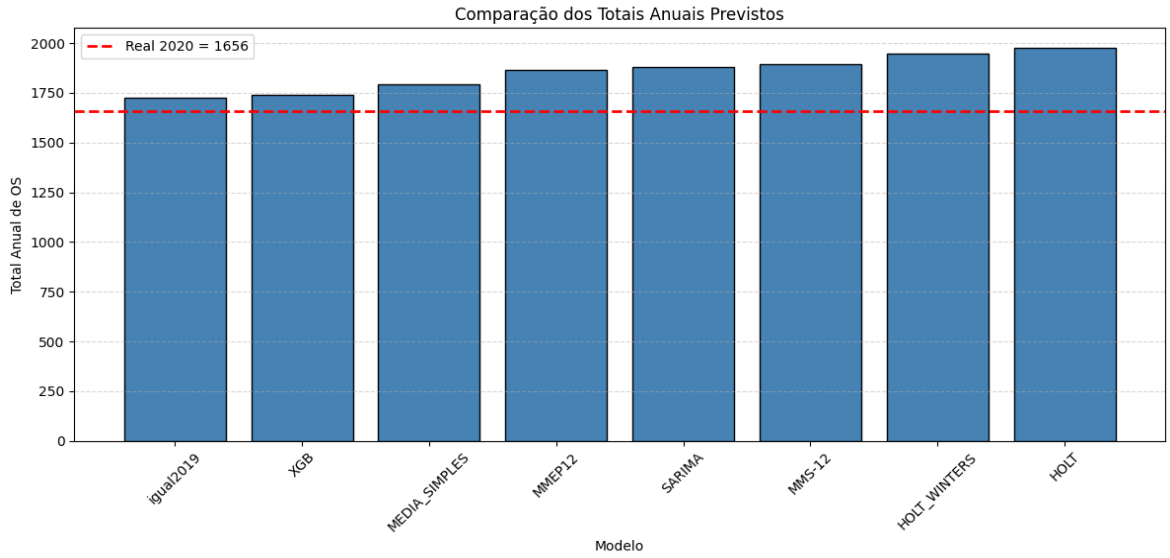
O modelo MMS-12 apresentou o segundo melhor resultado em termos de MAE (26,44), seguido pelo modelo SARIMA (26,89). A pequena diferença entre esses modelos sugere desempenho semelhante para a série analisada. O modelo *XGBoost* apresentou erro ligeiramente superior aos modelos de suavização, com MAE de 28,84 e MAPE de 21,84%, evidenciando capacidade moderada de capturar as oscilações da demanda.

Os modelos Holt e Holt-Winters apresentaram os maiores valores de erro. O método Holt-Winters registrou MAE de 40,15 e MAPE de 30,75%, enquanto o modelo Holt apresentou MAE de 37,67 e MAPE de 30,11%. Esses resultados indicam que os componentes de tendência e sazonalidade incorporados por esses métodos não contribuíram para melhorar a qualidade das previsões na série analisada.

Além da comparação mensal, foi realizada a análise dos totais anuais previstos para o ano de 2020. O valor real observado foi de 1.656 ordens de serviço. Os resultados são apresentados na Figura 12 e na Figura 13.

Novamente, o modelo MMEP12 apresentou o melhor desempenho, produzindo erro absoluto anual de apenas 119 ordens de serviço, correspondente a um erro percentual de 7,19%. O modelo *XGBoost* apresentou resultado semelhante, com erro percentual de 8,03%, seguido pela Média Simples (8,41%) e pelo SARIMA (8,81%). Os modelos Holt e Holt-Winters apresentaram as maiores divergências em relação ao valor real observado, superestimando a demanda anual em aproximadamente 19,38% e 17,42%, respectivamente.

Figura 12 – Gráfico com a comparação entre o valor total anual real do ano de 2020 e o valor encontrado para cada modelo.



Fonte: Elaboração própria.

Figura 13 - Erro percentual do quantitativo anual previsto por modelo

Modelo	Erro percentual (%)
MMEP12	7,19
XGB	8,03
MEDIA SIMPLES	8,41
SARIMA	8,81
MMS-12	9,82
IGUAL2019	12,38
HOLT_WINTERS	17,42
HOLT	19,38

Fonte: Elaboração própria.

Com o objetivo de verificar se as diferenças observadas entre os modelos eram estatisticamente significativas, foi aplicado o teste não paramétrico de Wilcoxon aos erros absolutos das previsões, utilizando o modelo MMEP12 como referência por ter apresentado o melhor desempenho geral. Os resultados estão apresentados na Figura 14.

Figura 14 – Resultado do teste Wilcoxon comparando os erros absolutos do MMEP12 com os erros dos demais modelos.

Modelo	p-value
igual2019	0,266113
MEDIA_SIMPLES	0,020996
MMS-12	0,000977
HOLT	0,000977
HOLT_WINTERS	0,000488
SARIMA	0,03418
XGB	0,063965

Fonte: Elaboração própria.

Considerando um nível de significância de 5% ($\alpha = 0,05$), observa-se que os modelos Média Simples, MMS-12, Holt, Holt-Winters e SARIMA apresentaram diferenças estatisticamente significativas em relação ao MMEP12, uma vez que os respectivos valores de p foram inferiores a 0,05.

Por outro lado, os modelos Igual2019 ($p = 0,2661$) e *XGBoost* ($p = 0,0640$) não apresentaram diferenças estatisticamente significativas quando comparados ao MMEP12. Embora o MMEP12 tenha apresentado os menores valores de MAE, RMSE e MAPE, os resultados do teste indicam que, para a amostra analisada, não há evidências estatísticas suficientes para afirmar que seu desempenho seja superior ao dos modelos Igual2019 e *XGBoost* ao nível de significância adotado.

DISCUSSÃO

Os resultados obtidos indicam que a série analisada, caracterizada por tamanho reduzido da amostra histórica e comportamento relativamente estável, pode ser adequadamente representada por métodos estatísticos de menor complexidade. Embora tenham sido avaliadas técnicas com diferentes níveis de sofisticação, os ganhos obtidos com modelos mais complexos foram limitados para o conjunto de dados estudado.

A aplicação do teste de Wilcoxon mostrou que, apesar do melhor desempenho observado para a MMEP12, não foram identificadas diferenças estatisticamente significativas em relação aos modelos *XGBoost* e Igual2019 ao nível de significância de 5%. Esse resultado sugere que a superioridade observada nas métricas de erro deve ser interpretada com cautela, considerando a variabilidade inerente à amostra disponível.

Do ponto de vista operacional, os resultados possuem relevância prática para empresas de saneamento, uma vez que métodos mais simples apresentam menor custo computacional, facilidade de implementação e maior transparência na interpretação dos resultados. Dessa forma, podem constituir ferramentas adequadas para apoio ao planejamento, dimensionamento de contratos e alocação de recursos em atividades de manutenção.

Como análise complementar, foi realizada uma validação retrospectiva utilizando os anos de 2016 a 2018 para treinamento e o ano de 2019 para teste. Os resultados mantiveram o padrão observado para a validação de 2020, com destaque novamente para a MMEP12. Embora os erros tenham sido, em geral, inferiores aos observados em 2020, a hierarquia de desempenho dos modelos permaneceu semelhante. Esse resultado sugere que as conclusões obtidas são consistentes e não dependem exclusivamente das particularidades do período marcado pelos impactos da pandemia da COVID-19.

CONCLUSÃO

Este trabalho comparou diferentes modelos de previsão aplicados à estimativa da demanda de serviços de desobstrução de redes de esgoto na região do Plano Piloto, utilizando dados operacionais da CAESB e técnicas estatísticas e de aprendizado de máquina amplamente utilizadas em problemas de previsão.

Os resultados indicaram que a Média Móvel Exponencialmente Ponderada de 12 períodos (MMEP12) apresentou o melhor desempenho geral entre os modelos avaliados. Entretanto, a análise estatística mostrou que

não houve evidências suficientes para afirmar superioridade significativa desse modelo em relação ao *XGBoost* e ao modelo de persistência Igual2019, sugerindo desempenho comparável para a amostra analisada.

De forma geral, os resultados sugerem que séries temporais curtas, relativamente estáveis e com baixa intensidade de tendência e sazonalidade podem ser adequadamente modeladas por métodos estatísticos de menor complexidade. Para o conjunto de dados estudado, o aumento da complexidade dos modelos não resultou em ganhos substanciais de desempenho preditivo.

A validação complementar utilizando o ano de 2019 como período de teste produziu resultados consistentes com aqueles observados para 2020, reforçando a robustez das conclusões obtidas e reduzindo a possibilidade de que os resultados sejam consequência exclusiva das condições atípicas associadas à pandemia da COVID-19.

Sob a perspectiva gerencial, os resultados indicam que métodos simples de suavização temporal podem ser utilizados como ferramentas eficazes de apoio ao planejamento operacional e orçamentário de serviços de manutenção em sistemas de esgotamento sanitário.

Como trabalhos futuros, recomenda-se ampliar a base histórica analisada, incorporar variáveis externas relacionadas a fatores climáticos, operacionais e socioeconômicos, bem como avaliar técnicas mais avançadas de aprendizado de máquina e validação temporal. Tais iniciativas poderão contribuir para o desenvolvimento de modelos mais robustos e aderentes às necessidades de planejamento de longo prazo e à gestão eficiente dos serviços de saneamento.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. ABNT NBR ISO 55000:2014: Gestão de ativos: visão geral, princípios e terminologia. Rio de Janeiro: ABNT, 2014.

2. BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. Time Series Analysis: Forecasting and Control. 5. ed. Hoboken: Wiley, 2015

3. CHEN, T.; GUESTRIN, C. XGBoost: A Scalable Tree Boosting System. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, 2016. p. 785-794.

4. DICKEY, D. A.; FULLER, W. A. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, v. 74, n. 366, p. 427-431, 1979.

5. FRIEDMAN, J. H. Greedy Function Approximation: A Gradient Boosting Machine. The Annals of Statistics, v. 29, n. 5, p. 1189–1232, 2001.

6. HYNDMAN, R. J.; KOEHLER, A. B. Another look at measures of forecast accuracy. *International Journal of Forecasting*, Amsterdam, v. 22, n. 4, p. 679-688, 2006.

7. HYNDMAN, R. J.; ATHANASOPOULOS, G. *Forecasting: Principles and Practice*. 3. ed. Melbourne: OTexts, 2021. Disponível em: https://otexts.com/fpp3/. Disponível em: <https://otexts.com/fpp3/>. Acesso em: 29 jul. 2026.

8. MORETTIN, P. A.; TOLOI, C. M. Análise de séries temporais. 3. ed. São Paulo: Blucher, 2018.

9. SILVA NETO, J.C.; LIMA, A.M.G. Implantação do Controle de Manutenção. *Revista Club de Mantenimiento*, n. 10, 2002.

10. WILCOXON, F. Individual comparisons by ranking methods. Biometrics Bulletin, v. 1, n. 6, p. 80-83, 1945