

Resumos Expandidos

Análise e Previsão de Séries Temporais na Gestão da Demanda de Água

Time Series Analysis and Forecasting in Water Demand Management

Franchesco Sanches dos Santos¹, Marcos Vinicius de Oliveira Peres², Eniuce Menezes³

¹ Graduado em Matemática – UNESPAR. Mestre em Engenharia de Produção e Sistemas – PUC/PR

² Doutor em Ciências – FMRP/USP, Professor Adjunto do Departamento de Estatística – UEM

³ Doutora em Ciências Cartográficas. Professora Associada do Departamento de Estatística – UEM

✉ franchesco.sanches@gmail.com, mvoperes@uem.br, emenezes@uem.br

Palavras-chave:

Previsão;
Séries Temporais;
Abastecimento de Água;
Aprendizado de Máquina.

Resumo

A previsão dos níveis dos reservatórios de água é fundamental para a gestão eficiente dos recursos hídricos, especialmente em áreas urbanas. Este estudo compara os modelos ARIMA, *Decision Tree* (DT) e Prophet aplicados a dados horários do Bairro Alto, Curitiba, coletados entre 2018 e 2020. As métricas *Symmetric Mean Absolute Percentage Error* (SMAPE), *Mean Absolute Error* (MAE) e *Root Relative Mean Square Error* (RRMSE) foram utilizadas para horizontes de 1, 6, 12 e 24 horas. O DT apresentou melhor desempenho em MAE e RRMSE, principalmente em previsões curtas, enquanto o Prophet obteve menores valores de SMAPE, destacando-se em previsões mais longas. Os resultados indicam que a abordagem adotada melhora a precisão na previsão da demanda, auxiliando a prevenção de escassez e o uso sustentável da água.

Keywords:

Forecasting;
Time Series;
Water Supply;
Machine Learning.

Abstract

The prediction of water reservoir levels is essential for the efficient management of water resources, especially in urban areas. This study compares the ARIMA, Decision Tree (DT), and Prophet models applied to hourly data from Bairro Alto, Curitiba, collected between 2018 and 2020. The metrics Symmetric Mean Absolute Percentage Error (SMAPE), Mean Absolute Error (MAE), and Root Relative Mean Square Error (RRMSE) were used for horizons of 1, 6, 12, and 24 hours. The DT model showed better performance in MAE and RRMSE, mainly in short-term forecasts, while the Prophet model achieved lower SMAPE values, standing out in longer forecasts. The results indicate that the proposed approach improves demand prediction accuracy, supporting scarcity prevention and the sustainable use of water.

1 INTRODUÇÃO

A previsão dos níveis dos reservatórios é essencial para o planejamento e a gestão eficiente da água, principalmente em cenários de crise hídrica. Flutuações resultantes de fenômenos hidrológicos e climáticos, com alta variabilidade e não linearidade, tornam a modelagem desse tipo de série temporal desafiadora. Modelos de aprendizado de máquina e estatísticos, como ARIMA, Prophet e *Decision Tree* (DT), permitem lidar com padrões complexos e são amplamente utilizados em dados hidrometeorológicos. O presente estudo aplica e compara esses modelos a dados horários de um reservatório no Bairro Alto, Curitiba, coletados entre 2018 e 2020, avaliando-os com as métricas SMAPE, MAE e RRMSE para horizontes de 1, 6, 12 e 24 horas.

O objetivo é identificar o modelo mais eficaz para melhorar a previsão da demanda de água, contribuindo para a tomada de decisão na operação e manutenção do sistema. O artigo está organizado em: Seção 2 – trabalhos relacionados; Seção 3 – métodos; Seção 4 – resultados e análise; Seção 5 – considerações finais.

2 TRABALHOS RELACIONADOS

Diversos estudos investigaram métodos de previsão de demanda de água, combinando técnicas estatísticas e de aprendizado de máquina. Balti et al. (2021) compararam ARIMA, Prophet e LSTM na previsão de secas, concluindo que o LSTM apresentou melhor desempenho na maioria dos casos. Ibrahim, Omar e Maghraby (2020) avaliaram ARIMA e *Support Vector Regression* (SVR) para dados do Kuwait, obtendo menor erro com SVR.

Liu, Savic e Fu (2023) demonstraram que o Prophet mantém desempenho estável mesmo com ruído nos dados. Landicho, Saengarunwong e Esichaikul (2019) destacaram o ARIMA como modelo mais preciso em dados domésticos. Em aplicações com IoT, estudos como Water (2020) mostraram superioridade do LSTM em comparação ao ARIMA. Liu et al. (2023) propuseram um modelo híbrido ARIMA-DNN, alcançando bom desempenho anual.

Modelos baseados em árvores, como XGBoost e Random Forest, também foram aplicados a previsões de curto prazo (Brédy et al., 2020), enquanto abordagens híbridas com SOM e RT mostraram ganhos em precisão (Bata, Carriveau e Ting, 2020). Revisões recentes (Alhendi et al., 2022; Niknam et al., 2022) reforçam o potencial da inteligência artificial para previsões no setor hídrico.

A Tabela 1 resume os trabalhos mais relevantes, apresentando modelos, horizontes e métricas utilizadas, servindo como base comparativa para este estudo.

Tabela 1: Trabalhos relacionados

Referência	Models	Escala temporal	Horizonte	Métricas
Landicho et al. (2019)	ARIMA, SVM, NN e LR	Diário	1- 3600	MAPE
Brédy et al. (2020)	RF, XGB, DT (basic)	Horas	1- 48	RMSE, NSE
Ibrahim et al. (2020)	SVR, ARIMA	Diário	1	RMSE, MAPE
Balti et al. (2021)	LSTM, Prophet, ARIMA	Diário	1 - 30	RMSE, MAE, R^2
Liu, Savic & Fu (2023)	RF	Diário	1	R^2
Liu et al. (2023)	ARIMA-DNN	Anual	1	RMSE, MAE, R^2

3 DESENVOLVIMENTO

Esta seção descreve o conjunto de dados e os métodos usados para desenvolver e executar os experimentos, incluindo o processo de seleção de recursos, a divisão de dados e os modelos comparados.

3.1 Conjunto de Dados

Entre 2018 e 2020, foram coletadas 26.306 amostras horárias em um reservatório da SANEPAR, em Curitiba. As variáveis incluem frequência das bombas B1, B2 e B3 (Hz), nível do reservatório LT01 (m), vazões FT01, FT02 e FT03 (m³/h) e pressões PT01SU e PT02RBAL (mca). Após análise de correlação, foram mantidas as variáveis com maior relação com LT01. Os dados foram normalizados e divididos em 70% para treino/validação e 30% para teste.

3.2 Modelos de Previsão

Foram utilizadas 26.306 amostras horárias coletadas entre 2018 e 2020 em um reservatório da SANEPAR, Curitiba. As variáveis mais correlacionadas ao nível do reservatório LT01 foram mantidas para modelagem. Os dados foram normalizados e divididos em 70% para treino/validação e 30% para teste.

Modelos avaliados:

- ARIMA: modelo estatístico que combina autorregressão, diferenciação e média móvel.
- Prophet: desenvolvido pelo Facebook, decompõe tendência, sazonalidade e feriados.
- Decision Tree (DT): baseado em regras hierárquicas, otimizado com algoritmo TPE.

Métricas: SMAPE (erro percentual médio simétrico), MAE (erro absoluto médio) e RRMSE (raiz do erro quadrático médio relativo). Testes KPSS, Ljung-Box, Friedman e Nemenyi foram aplicados para análise estatística.

4 RESULTADOS

Os modelos foram avaliados com dados horários de 2018–2019, utilizando LT01 como variável alvo. A Tabela 3 resume SMAPE, MAE e RRMSE para previsões de 1, 6, 12 e 24 horas.

- 1 hora: DT apresentou menor MAE e RRMSE; Prophet teve maior SMAPE.
- 6 horas: Prophet obteve menor SMAPE e MAE; DT manteve menor RRMSE.
- 12 horas: Prophet apresentou menor SMAPE; DT obteve menores MAE e RRMSE.
- 24 horas: Prophet se destacou em SMAPE e MAE; DT manteve melhor RRMSE.

O teste KPSS indicou não estacionariedade na série LT01, justificando transformações antes da modelagem. O teste Ljung-Box mostrou ausência de autocorrelação significativa para horizontes curtos, mas presença para 24 horas, especialmente no ARIMA. Nos testes de Friedman e Nemenyi, não houve diferença estatisticamente significativa para SMAPE e MAE, mas DT apresentou desempenho superior em RRMSE.

No geral, DT se mostrou mais consistente para erros absolutos e relativos, enquanto Prophet foi mais preciso em termos percentuais para horizontes longos.

A pesquisa aborda o abastecimento de água, focando em variáveis como Bombas de Sucção (B1, B2 e B3), Nível do Reservatório (LT01), Vazões (FT01, FT02, FT03) e Pressões (PT01SU, PT02RBAL). Os dados, de 2018 a 2020, foram divididos por hora, excluindo 2020 devido a irregularidades.

Tabela 3: Medidas de desempenho em treinamento, validação, teste e todos os dados.

Horizontes	Métricas	Treinamento			Validação			Teste			Todos os dados		
		ARIMA	DT	Prophet	ARIMA	DT	Prophet	ARIMA	DT	Prophet	ARIMA	DT	Prophet
1 hora à frente	SMAPE	7,86	7,69	21,27	8,24	6,90	14,02	7,62	8,51	18,22	7,99	7,83	18,22
	MAE	0,25	0,35	0,77	0,25	0,31	0,46	0,25	0,39	0,50	0,26	0,36	0,50
	RRMSE	0,09	0,21	0,77	0,10	0,20	0,46	0,09	0,18	0,50	0,09	0,20	0,50
6 horas à frente	SMAPE	19,99	13,94	33,71	22,12	13,58	7,16	19,74	12,66	29,73	13,50	13,50	29,73
	MAE	0,64	0,59	1,12	0,69	0,57	0,25	0,65	0,56	0,93	0,58	0,58	0,93
	RRMSE	0,23	0,16	1,15	0,25	0,16	0,28	0,23	0,14	1,05	0,16	0,16	1,05
12 horas à frente	SMAPE	23,00	13,94	24,62	25,34	13,58	14,61	22,91	12,66	23,91	13,50	13,50	23,91
	MAE	0,74	0,59	0,83	0,80	0,57	0,54	0,77	0,56	0,80	0,58	0,58	0,80
	RRMSE	0,26	0,16	0,91	0,29	0,16	0,59	0,27	0,14	0,97	0,16	0,16	0,97
24 horas à frente	SMAPE	13,33	13,94	6,55	14,70	13,58	13,41	12,80	12,66	5,05	13,65	13,50	5,05
	MAE	0,43	0,59	0,23	0,46	0,57	0,43	0,42	0,56	0,17	0,44	0,58	0,17
	RRMSE	0,17	0,16	0,29	0,18	0,16	0,56	0,16	0,14	0,19	0,17	0,16	0,19

As medidas de desempenho estão em m^3/h

5 CONSIDERAÇÕES FINAIS

Este estudo comparou ARIMA, Prophet e *Decision Tree* (DT) na previsão da demanda de água no Bairro Alto, Curitiba, utilizando dados horários da SANEPAR. A análise mostrou que o DT foi mais consistente em MAE e RRMSE, especialmente em previsões curtas e médias, enquanto o Prophet se destacou em SMAPE para horizontes longos.

Os resultados indicam que a aplicação de modelos de séries temporais, aliada à otimização de hiperparâmetros, pode apoiar decisões operacionais e prevenir escassez. A metodologia proposta é de fácil implementação e pode ser adaptada a outros sistemas de abastecimento.

Como trabalhos futuros, sugere-se explorar modelos híbridos, incluir variáveis climáticas e aplicar métodos avançados de otimização para lidar com sazonalidades complexas e eventos extremos.

REFERÊNCIAS

ALHENDI, A. A. et al. Artificial intelligence for water-energy nexus demand forecasting: a review. *International Journal of Low-Carbon Technologies*, v. 17, p. 730-744, 2022. Disponível em: <https://doi.org/10.1093/ijlct/ctac043>.

BALTI, H. et al. Big data based architecture for drought forecasting using LSTM, ARIMA, and Prophet: case study of the Jiangsu Province, China. In: *2021 International Congress of Advanced Technology and Engineering (ICOTEN)*. [S.l.: s.n.], 2021. p. 1-8.

BATA, M.; CARRIVEAU, R.; TING, D. S.-K. Short-term water demand forecasting using hybrid supervised and unsupervised machine learning model. *Smart Water*, v. 5, n. 1, p. 2, maio 2020. Disponível em: <https://doi.org/10.1186/s40713-020-00020-y>.

BRÉDY, J. et al. Water table depth forecasting in cranberry fields using two decision-tree-modeling approaches. *Agricultural Water Management*, v. 233, p. 106090, 2020. Disponível em: <https://doi.org/10.1016/j.agwat.2020.106090>.

IBRAHIM, T.; OMAR, Y.; MAGHRABY, F. A. Water demand forecasting using machine learning and time series algorithms. In: *2020 International Conference on Emerging Smart Computing and Informatics (ESCI)*. [S.l.: s.n.], 2020. p. 325-329.

KWIATKOWSKI, D. et al. Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? *Journal of Econometrics*, v. 54, n. 1, p. 159-178, 1992. Disponível em: [https://doi.org/10.1016/0304-4076\(92\)90104-Y](https://doi.org/10.1016/0304-4076(92)90104-Y).

LANDICHO, J.; SAENGARUNWONG, A.; ESICHAIKUL, V. Modelling domestic water demand and management using multi-criteria decision making technique. *Water Supply*, v. 17, p. 56-70, jan. 2019. Disponível em: <https://doi.org/10.2166/ws.2016.113>.

LIU, G.; SAVIC, D.; FU, G. Short-term water demand forecasting using data-centric machine learning approaches. *Journal of Hydroinformatics*, v. 25, n. 3, p. 895-911, mar. 2023. Disponível em: <https://doi.org/10.2166/hydro.2023.163>.

LIU, G. et al. Water demand in watershed forecasting using a hybrid model based on autoregressive moving average and deep neural networks. *Environmental Science and Pollution Research*, v. 30, n. 5, p. 11946-11958, 2023. Disponível em: <https://doi.org/10.1007/s11356-022-22984-1>.

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

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

NIKNAM, A. et al. A critical review of short-term water demand forecasting tools – what method should I use? *Sustainability*, v. 14, n. 9, 2022. Disponível em: <https://doi.org/10.3390/su14095412>.