

PREVISÃO DE VAZÕES MENSIS POR SÉRIES TEMPORAIS LONGAS E CURTAS

RICARDO ARAGÃO E SILVA¹, ÂNGELO MÁRCIO OLIVEIRA SANT'ANNA², MAURICIO SANTANA LORDELO³ e LIDIANE MENDES KRUSCHEWSKY LORDELO⁴

¹Pesquisador SGB/CPRM, Mestrando PEI UFBA, Salvador-BA, ricardo.aragao@ufba.br;

²Dr. Prof. Adjunto, UFBA, Salvador-BA, angelo.santanna@ufba.br;

³Dr. Prof. Adjunto, UEFS, Feira de Santana-BA, mslordelo@uefs.br;

⁴Dr. Prof. Adjunto, UFRB, Cruz das Almas-BA, lidiane@ufrb.edu.br

RESUMO: O presente trabalho procurou aferir a acurácia na modelagem por aprendizado de máquina para previsão de vazões com duas estações de monitoramento hidrológico na Bahia. A metodologia utilizou séries temporais com o software R. O banco de dados utilizado da estação de Santa Maria da Vitória é de longo período (1977-2022) e da estação de Batalha é curto (2015-2023). Os resultados indicaram melhor ajuste no modelo de séries temporais ARIMA (3,0,3)(3,1,1) que incorpora a sazonalidade temporal. O modelo apresentou nível de estacionariedade satisfatória no teste Dickey-Fuller ($D-F = -5.3257$; $p\text{-value}=0.01$) e ($D-F = -5.2138$; $p\text{-value}=0.01$) e os erros de estimação ($MAPE=16,20$; $RMSE=51,22$; $MAE=35,53$; $ACF= -0,0037$) e ($MAPE=0,4532$; $RMSE=0,0955$; $MAE=0,0730$; $ACF1= -0,1106$) para as estações de Santa Maria da Vitória e Batalha. Os valores previstos acompanharam o comportamento da série e obtiveram medidas de qualidade excelentes, com um intervalo de confiança de (80% e 95%).

PALAVRAS-CHAVE: Vazões, modelagem, séries temporais, estatística aplicada.

FORECASTING MONTHLY FLOWS USING LONG AND SHORT TIME SERIES

ABSTRACT: This study aimed to assess the accuracy of machine learning modeling for flow prediction using two hydrological monitoring stations in Bahia, Brazil. The methodology employed time series analysis with the R software. The dataset from the Santa Maria da Vitória station spans a long period (1977-2022), while the Batalha station dataset covers a shorter period (2015-2023). The results indicated a better fit for the ARIMA time series model (3,0,3)(3,1,1) that incorporates temporal seasonality. The model demonstrated satisfactory stationarity levels in the Dickey-Fuller test ($D-F = -5.3257$; $p\text{-value} = 0.01$) and ($D-F = -5.2138$; $p\text{-value} = 0.01$) and estimation errors ($MAPE = 16.20$; $RMSE = 51.22$; $MAE = 35.53$; $ACF = -0.0037$) and ($MAPE = 0.4532$; $RMSE = 0.0955$; $MAE = 0.0730$; $ACF1 = -0.1106$) for the Santa Maria da Vitória and Batalha stations, respectively. The predicted values closely followed the behavior of the series and achieved excellent quality measures, with confidence intervals of 80% and 95%.

KEYWORDS: Flows, modeling, time series, applied statistics.

INTRODUÇÃO

A importância dos estudos para um melhor gerenciamento dos recursos hídricos permanece relevante para governos e entidades que utilizam a água, um recurso natural indispensável para a vida humana. O estudo e descrição dos rios e cursos d'água é chamado de potamografia, termo atribuído ao geógrafo e hidrólogo francês Alfred de Martonne. Segundo Heath (1983), o ciclo hidrológico refere-se ao constante movimento da água acima da superfície terrestre, sobre a superfície e abaixo dessa camada da terra. No contexto de aquíferos sedimentares porosos, há uma interação significativa entre as águas superficiais e subterrâneas, assim, realizar monitoramento hidrológico torna-se uma ferramenta importantíssima no gerenciamento da bacia hidrográfica. Esta interação é notável na região do aquífero Urucuia, que é um reservatório de grande potencial subterrâneo e contribui significativamente para a manutenção das vazões do rio São Francisco, especialmente entre a divisa do estado de Minas Gerais e a montante da UHE Sobradinho (Serviço Geológico do Brasil/SGB-CPRM, 2019). Esta região é também marcada pela expansão agrícola, com uso intensivo dos recursos hídricos superficiais e subterrâneos para irrigação.

A Rede Hidrometeorológica do Aquífero Urucuia (RHU) foi instalada pelo Serviço Geológico do Brasil-CPRM em 2013, com o objetivo de intensificar o monitoramento na região e promover um maior conhecimento do comportamento hidrológico. Nesta rede foram instaladas 12 estações climatológicas, 17 estações pluviométricas e 16 estações fluviométricas (SGB, 2019). Algumas estações fluviométricas dessa rede foram descontinuadas, mas a estação de Batalha permanece monitorada. Com o banco de dados atual da RHU o objetivo deste estudo foi realizar a previsão de vazões dessa estação (Batalha, 2015-2022) e compará-la com outra estação mais antiga da Rede Hidrometeorológica Nacional de Referência (RHNR) (Santa Maria da Vitória, 1977-2022). Utilizou-se o banco de dados dos resumos de descargas dessas duas estações, aplicando aprendizado de máquina por séries temporais para previsão de vazões. O software utilizado foi o R, de uso livre.

MATERIAL E MÉTODOS

O presente trabalho usa os dados do acordo de cooperação técnica entre o Serviço Geológico do Brasil/SGB-CPRM e a Agência Nacional de Águas/ANA para monitorar redes hidrológicas. A estação de Santa Maria da Vitória dispõe de um banco de dados iniciado no ano de 1977, enquanto a estação de Batalha, localizada no aquífero Urucuia, teve sua série de dados iniciada em 2015, apesar de a RHU ter sido implantada em 2013. O rio São Francisco é medido na estação de Santa Maria da Vitória e o rio Grande, afluente do mesmo, é medido na estação de Batalha. O rio Grande corre na direção sudoeste-nordeste, recebendo seus principais afluentes na margem esquerda, como os rios das Fêmeas, de Ondas, Branco e Preto (SGB, 2021). O rio Grande é perene, tem conformação retilínea e fundo regular.

Estudo do Serviço Geológico do Brasil de 2021 chamado “Sistema Aquífero Urucuia: caracterização hidrológica com dados da rede hidrometeorológica do aquífero Urucuia” (<https://rigeo.cprm.gov.br/handle/doc/22606>) motivou a descontinuidade do monitoramento de estações fluviométricas; a estação de Batalha, no entanto, permaneceu ativa. A metodologia utilizada no presente trabalho foi resultado de dados do monitoramento hidrológico da Rede Hidrometeorológica Nacional (RHN), operada pelo Serviço Geológico do Brasil/SGB em parceria com a Agência Nacional de Águas/ANA. A validade das informações é ratificada com a publicação das séries históricas das estações no sistema *hidroweb*¹ da ANA. A estação de Batalha, por ser um projeto institucional do SGB, não tem os dados publicados nesse sistema. As duas estações dispõem de curvas-chave, sem do que estas curvas precisam estar aprovadas a fim de que suas vazões sejam geradas e validadas, bem como definidos prazos de validade das mesmas. As medições de vazão foram coletadas por equipamentos acústicos conhecidos como *Acoustic Doppler Current Profile* (ADCPs).

A série histórica de Santa Maria da Vitória dispõe de 386 dados de descarga, e Batalha possui 92 medições, sendo que em ambas as estações as variáveis medidas em campo são cota, vazão, área molhada, largura, velocidade média e profundidade. Após a avaliação das séries históricas das estações foi realizada a modelagem por séries temporais estabelecendo o modelo mais adequado e estatisticamente significativo.

RESULTADOS E DISCUSSÃO

As vazões geradas pela modelagem com séries temporais para a estação da RHNR Santa Maria da Vitória (1977-2022) e para a estação da RHU Batalha, com dados brutos e consistidos, tiveram boa acurácia em relação às vazões medidas nas seções de medição das estações no campo. O banco de dados da estação de Santa Maria da Vitória dispõe de maior histórico de dados medidos, com equipamentos analógicos e digitais, e, pelo maior período de monitoramento, esses dados apresentam maior erro percentual médio, seja por estarem suscetíveis à ocorrência de eventos de cheias e estiagens ao longo do tempo, bem como à natureza dos dados estocásticos e à histerese. As estações recebem técnicos e engenheiros de 4 a 6 vezes por ano, sendo que eventos de cheias devem ser monitorados diariamente a fim de obtenção de mais dados de campo em cotas mais altas. Essas informações

¹ *HidroWeb* é uma ferramenta integrante do Sistema Nacional de Informações sobre Recursos Hídricos (SNIRH). <https://www.snirh.gov.br/hidroweb/apresentacao/>

melhoram a eficiência e eficácia dos trabalhos em escritório, minimizando os erros, otimizando custos e gerando dados mais confiáveis para a sociedade. A relação entre cota-vazão, chamada de curva chave, apresenta menores desvios em cotas mais altas quando são realizadas medições diárias nesses eventos extremos. Esse fato também melhora o conhecimento dos elementos de tendência e sazonalidade da série temporal, podendo reduzir os ruídos ou um melhor entendimento dos mesmos. A métrica dos erros foram MAPE=15,04; RMSE=51,22; MAE=35,53; ACF= -0,0037, sendo que o MAPE foi 15,04 (1977-2022/dados brutos e consistidos) e 15,91 (1977-2014/dados consistidos). O mês de referência foi Fevereiro, estando os valores de previsão do modelo dentro da relação cota-vazão (curva chave) para este mês. Os resultados da estação de Batalha foram muito bons, com MAPE de MAPE = 0.4532; RMSE = 0.0955; MAE = 0.0730; ACF1 = -0.1116. A estação necessita de mais dados de campo, principalmente em eventos extremos. Com os valores do *forecast* foram obtidos com a previsão por julgamento, sendo necessário um bom conhecimento da área e informações atualizadas (Laurence et al., 2006). Ainda assim, para obtenção de melhor confiabilidade, importante destacar as análises descritas por Hyndman & Athanasopoulos (2021), nas quais “os dados estão disponíveis, as previsões estatísticas e com julgamento são geradas independentemente e depois combinadas”. Gerenciar recursos hídricos é uma atividade complexa, que depende de dados confiáveis, modelagens e análises estatísticas com boa acurácia, conhecimento local e da bacia hidrográfica, bem como da dinâmica social e econômica.

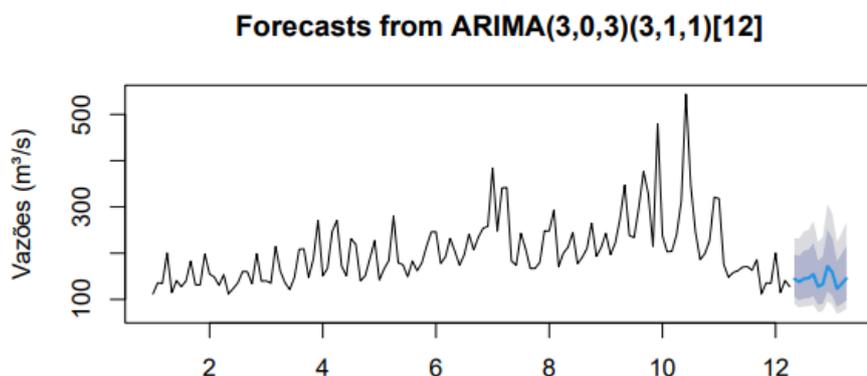


Figura 1. Valores reais ajustados da série e previsão das vazões médias mensais para Santa Maria da Vitória um período de 12 meses à frente. Fonte: Elaborado pelo autor (2023).

Forecasts from ARIMA(3,0,3)(3,1,1)[12]

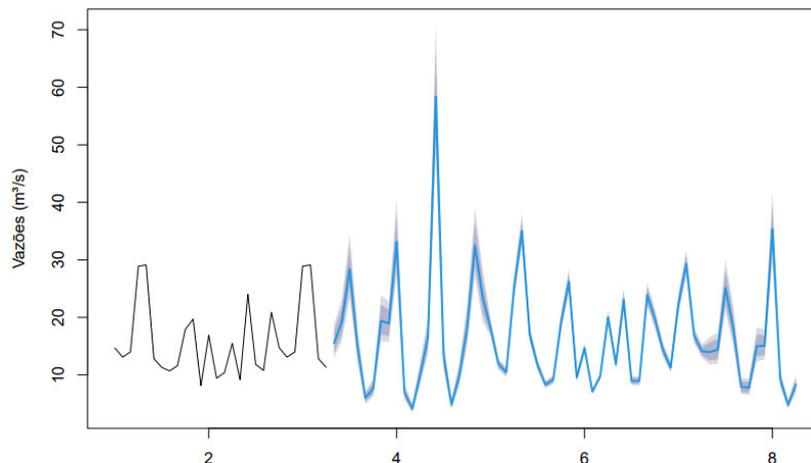


Figura 2. Valores reais ajustados da série e previsão das vazões médias mensais para Batalha um período de 12 meses à frente. Fonte: Elaborado pelo autor (2024).

CONCLUSÃO

Desenvolver e validar modelos hidrológicos para previsão de eventos extremos são mais uma ferramenta de gestão de recursos hídricos. Avaliações conjuntas com reservatórios subterrâneos melhoram o entendimento da disponibilidade hídrica, como recargas com memória curta (resposta rápida) com aplicação conjunta de *Long Short Temporal Network/LSTN* aumentam o conhecimento de séries temporais estocásticas.

Utilizar e aprimorar metodologias de aprendizado de máquina para auxiliar na metodologia adotada pela Agência Nacional de Águas (e executada pelo Serviço Geológico do Brasil) para o monitoramento das curvas chave e melhoria dos sistemas de alerta hidrológicos são um benefício que reduz a probabilidade de perdas humanas, além dos significativos danos materiais, financeiros e ambientais.

AGRADECIMENTOS

Ao Serviço Geológico do Brasil e ao Programa de Engenharia Industrial da UFBA pela orientação e treinamento concedidos ao primeiro autor.

REFERÊNCIAS

1. Acosta, SM., Oliveira, RMA., Sant'Anna, AMO. Machine learning algorithms applied to intelligent tire manufacturing. *International Journal of Computer Integrated Manufacturing*, ahead-of-print, 2023. <https://doi.org/10.1080/0951192X.2023.2177734>
2. Acosta, SM., Sant'Anna, AMO. Machine learning-based control charts for monitoring fraction nonconforming product in smart manufacturing. *International Journal of Quality and Reliability Management*, ahead-of-print, 1-35, 2022. <https://doi.org/10.1108/IJQRM-07-2021-0210>
3. Bayer, D. M.; Castro, N. M. R.; Bayer, F.M. Modelagem e Previsão de Vazões Médias Mensais do Rio Potiribu Utilizando Modelos de Séries Temporais. In: *Revista Brasileira de Recursos Hídricos*, v. 17 n.2, pp. 229-239, 2012.
4. Huang, W.; Jones, W. K.; Yang, Q. Field experiment study of transient stratified flow in a estuary. In: *Journal of Engineering Mechanics*, 129(10), pp. 1220-12236, 2003.
5. Hyndman, R. J. (2015). *Forecasting Functions for Time Series and Linear Models*. R package version 6.

6. Morettin, P. A., Toloi, C. M. Análise de Séries Temporais. São Paulo: Edgard Blucher. 2004.
7. Muchiri, P., Pintelon, L. Performance measurement using overall equipment effectiveness (OEE): literature review and practical application discussion, International Journal of Production Research, 46(13), 3517-3535, 2008. <https://doi.org/10.1080/00207540601142645>
8. Patil, A., Soni, G., Prakash, A., Karwasra, K. Maintenance strategy selection: a comprehensive review of current paradigms and solution approaches, International Journal of Quality & Reliability Management, 39(3),675-703. 2022. <https://doi.org/10.1108/IJQRM-04-2021-0105>
9. Pinto, W. P.; Lima, G. B.; Zanetti, J.B. Análise Comparativa de Modelos de Séries temporais para modelagem e previsão de regimes de vazões médias mensais do Rio Doce, Colatina – Espírito Santo. In: Ciência e Natura, v. 37 n.4, 2015. Revista do Centro de Ciências Naturais e Exatas da Universidade Federal de Santa Maria/RS.
10. R. R language and environment for statistical computing. R Foundation for Statistical Computing. ISBN 3-900051-07-0. 2022. Available at <http://www.r-project.org>.
11. RHNR. (2023). Rede Hidrometeorológica Nacional de Referência. Disponível em: <<https://www.sgb.gov.br/publique/Hidrologia/Monitoramento-Hidrologico-e-Hidrogeologico/Rede-Hidrometeorologica-Nacional-de-Referencia---RHNR-6589.html>>. Acessado em Janeiro 5, 2023.
12. WMO. World Meteorological Organization. Manual on Stream Gauging. Vol. II: Computation of Discharge, 2010.