

PREVISÃO DE VAZÕES EM BACIAS DE CABECEIRA: DESAFIOS E OPORTUNIDADES NA APLICAÇÃO DE REDES NEURAIS

Gabriel Rodrigues Pereira¹; Rodrigo Perdigão Gomes Bezerra²; Bruno Melo Brentan³ & André Ferreira Rodrigues⁴

ABSTRACT

Streamflow forecasts are essential strategic tools for water resources management and planning, whether for the operation of hydroelectric power plants, extreme event prediction, early warning systems, or water availability estimation. The scenario becomes particularly challenging when applications are focused on headwater regions, near the sources of watercourses, where, due to their environmental characteristics, the hydrological response can exhibit greater spatial and temporal variability. In this context, the use of a recurrent neural network (RNN), specifically a Long Short-Term Memory (LSTM) model, is proposed to perform forecasts over horizons of one day, one week, one month, and six months. The database was obtained from the historical time series of the Brazilian National Water and Sanitation Agency (ANA), with daily frequency, covering the period from 1970 to 2023. To verify and validate the model, the following performance metrics are selected: Coefficient of Determination (R^2), Kling-Gupta Efficiency (KGE), Mean Squared Error (MSE), Nash-Sutcliffe Efficiency (NSE), and Percent Bias (PBIAS). As a result, better performance is observed for short and medium term horizons, while greater uncertainties were associated with seasonal forecasts. Despite the challenges, the predicted flow duration curve is close to the observed one, indicating the model's ability to simulate average flows, which is important for reservoir planning (e.g., water balance and energy generation). Thus, future applications of these techniques for streamflow forecasting for Small Hydropower Plants (SHPs) are highlighted, considering the energy potential available in these regions and the sector's demand for more accurate forecasts.

RESUMO

Previsões de vazões constituem instrumentos estratégicos fundamentais para a gestão e o planejamento de recursos hídricos, seja para operação de usinas hidroelétricas, predições de eventos

1) Afiliação: Departamento de Engenharia Hidráulica e de Recursos Hídricos (EHR), Universidade Federal de Minas Gerais (UFMG), gabriel_rodrigues07@hotmail.com

2) Afiliação: Departamento de Engenharia Hidráulica e de Recursos Hídricos (EHR), Universidade Federal de Minas Gerais (UFMG), rodrigo.pgb@gmail.com

3) Afiliação: Departamento de Engenharia Hidráulica e de Recursos Hídricos (EHR), Universidade Federal de Minas Gerais (UFMG), brentan@ehr.ufmg.br.

4) Afiliação: Departamento de Engenharia Hidráulica e de Recursos Hídricos (EHR), Universidade Federal de Minas Gerais (UFMG), afrodrigues@ehr.ufmg.br.

extremos, sistemas de alertas, ou ainda para estimativas de disponibilidade hídrica. O cenário torna-se desafiador quando o conjunto de aplicações se concentram em regiões de cabeceira, próximas às nascentes dos cursos d'água, onde, em razão de suas características ambientais, a resposta hidrológica pode apresentar maior variabilidade espacial e temporal. Nesse sentido, propõe-se o uso de uma rede neural recorrente (RNN), do tipo *Long Short-Term Memory* (LSTM), a fim de realizar previsões em horizontes de um dia, uma semana, um mês e seis meses. O banco de dados fora obtido da base de séries históricas da Agência Nacional de Águas e Saneamento (ANA), de frequência diária, entre 1970-2023. Para verificação e validação do modelo, as seguintes métricas foram selecionadas: *Coeficiente de Determinação* (R^2), *Kling-Gupta Efficiency* (KGE), *Mean Squared Error* (MSE), *Nash-Sutcliffe Efficiency* (NSE) e *Percent bias* (PBIAS). Como resultado, observou-se melhor desempenho para horizontes de curto e médio prazo, enquanto, para previsões sazonais, maiores foram as incertezas associadas ao modelo. Apesar das dificuldades, a curva de permanência prevista foi próxima à observada, indicando a capacidade do modelo em simular vazões médias, importante para o planejamento de reservatórios (e.g., balanço hídrico e geração de energia). Assim, destacam-se futuras aplicações dessas técnicas para previsão de vazões para Pequenas Centrais Hidrelétricas (PCHs), considerando o potencial energético disponível nessas regiões e a demanda do setor por previsões mais assertivas.

Palavras-Chave – Previsão de vazões, bacias de cabeceira, redes neurais.

1. INTRODUÇÃO

Previsões hidrológicas são essenciais para fundamentar tomadas de decisões, seja para planejamento de recursos hídricos, gestão de cheias, mitigação de secas, operações de sistemas de água, geração de energia elétrica, (HE; LEE, 2021) sistemas de alertas e horizontes de disponibilidade hídrica futura. A busca por previsões assertivas e confiáveis estão cada vez mais recorrentes, à medida que as demandas pela gestão de recursos hídricos e mitigação de desastres naturais aumentaram substancialmente (SABZIPOUR et al., 2023).

Nesse contexto, as bacias de cabeceira se inserem por possuírem um papel fundamental no planejamento de recursos hídricos, em termos quantitativos e qualitativos (ALEXANDER et al., 2007). Regiões de cabeceira tendem a apresentar uma maior variabilidade hidrológica espacial e temporal, quando comparada com canais maiores, em razão de menores área contribuintes, menores

distâncias de fluxos nas componentes superficiais e subterrâneas, bem como reduzidas áreas superficiais e subterrâneas de armazenamento de água (WOHL, 2017).

Entretanto, há de observar dificuldades, uma vez que os fenômenos hidrológicos em geral, em razão da complexa interação e dependência, apresentam uma aleatoriedade intrínseca, haja vista os inúmeros fatores influentes nas diversas fases do ciclo hidrológico (NAGHETTINI; PINTO, 2007), além do caráter não linear do processo de transformação chuva-vazão (MARACAJÁ, 2005). Além disso, bacias de cabeceira caracterizam-se por possuírem uma variação altimétrica acentuadamente maior, prevalecendo gradientes energéticos elevados. (LAVARINI; MAGALHÃES, 2013). Por essa razão, tais localidades configuram como ideais para o aproveitamento energético, sendo propícias para instalação de Usinas e Pequenas Centrais Hidrelétricas (PCHs). Nesta perspectiva, as previsões de vazão podem fornecer subsídios com informações de curto, médio e longo prazos, sobre os possíveis cenários de entradas e, consequentemente, afetar as decisões a serem tomadas sobre saídas ou armazenamento de água (CASSAGNOLE, 2021).

Atualmente, os métodos existentes de previsões de vazões são divididos em quatro categorias, a citar: conceituais, empíricos, físicos e orientado por dados (BESAW et al., 2010). Dentre eles, os modelos convencionais baseados fisicamente há muito tempo têm sido aplicados para simular regimes hidrológicos, mas que, em razão da complexidade e alta demanda de diferentes tipos de dados para calibração e validação, falham em atender às necessidades práticas para uma análise hidrológica dinâmica (DAMAVANDI, 2019).

Desta posta, emergem-se os modelos baseados em dados, mais especificamente as Redes Neurais Artificiais (RNAs). Tal assertiva se deve a sua capacidade de modelar relações complexas, que são uma característica dos processos hidrológicos (e.g, previsão de vazões, qualidade da água), mas que também exige uma série de dados de longo prazo e de boa qualidade (MAIER et al., 2010). Modelos baseados em dados, incluindo aprendizado de máquina e aprendizado profundo, possuem a capacidade de apreender padrões oculto de séries históricas, sem a necessidade explícita de modelos físicos subjacentes (YANG et al., 2020).

No escopo das arquiteturas de aprendizado profundo, destaca-se a *Long Short-Term Memory* (LSTM). Desenvolvida por Hochreiter e Schmidhuber (1997), a LSTM revolucionou os modelos de Redes Neurais Recorrentes (RNNs), uma vez que possibilitou o gerenciamento bem-sucedido de dependências de conexões de longo prazo, ideais para séries temporais (LUO et al., 2025). Em geral, RNNs funcionam a partir de um *loop-feedback*, que permite que a rede revise informações ao longo

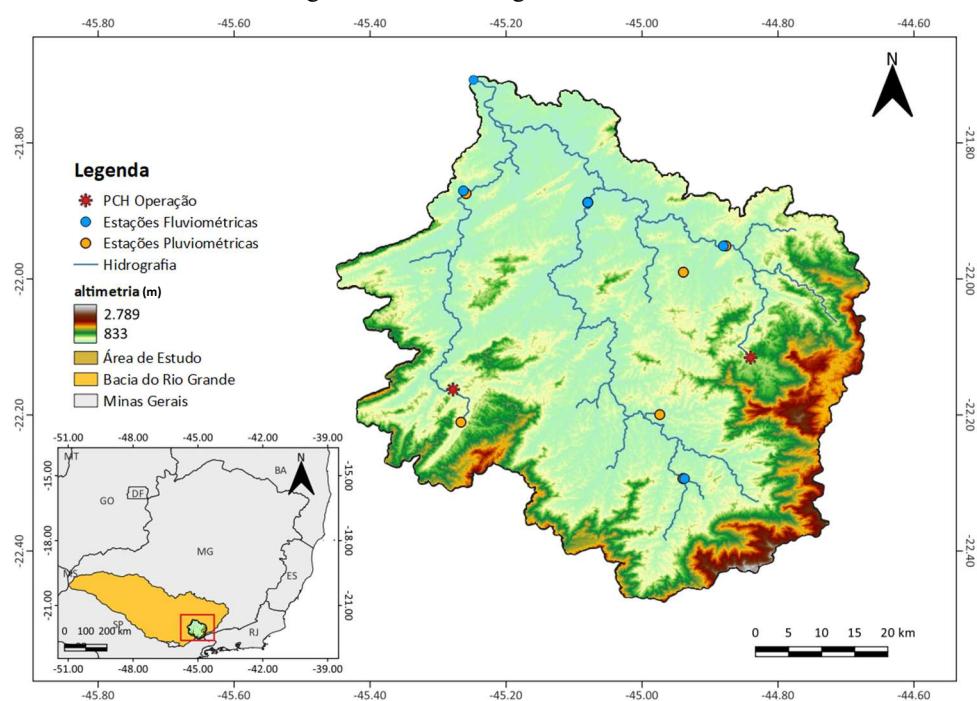
do tempo. No caso específico da LSTM, há a adição de uma unidade diferenciada, conhecida como célula de memória ou portão de esquecimento, que pode armazenar informações indefinidamente, o que permite que o modelo capte tendências de longo prazo (HUNT et al., 2022). Assim, observa-se um bom funcionamento do modelo ao lidar com dados caóticos resultantes da complexidade do ambiente real e da instabilidade de rios de médio ou pequeno porte (HU et al., 2020). Nesse sentido, propõe-se o uso da LSTM, no intuito de avaliar sua eficácia em previsões de vazões em uma bacia hidrográfica do Rio Grande, no estado de Minas Gerais, em horizontes temporais de um dia, uma semana, um mês e, por fim, seis meses.

2. MATERIAIS E MÉTODOS

2.1. Área de Estudo

A área de estudo pertence à região hidrográfica do Alto Rio Grande, localizada na porção sul do estado de Minas Gerais, próxima à divisa entre os estados do Rio de Janeiro e São Paulo. A bacia possui uma área de drenagem de 4.169 km², na qual flui um dos afluentes do Rio Grande, o Rio Verde. A localidade é conhecida por suas cadeias montanhosas (e.g., Serra da Mantiqueira, Parque Estadual da Serra do Papagaio, Parque Nacional do Itatiaia). Para delimitação da bacia, considerou-se como exutório a estação fluviométrica a ser prevista. Segundo a Agência Nacional de Águas, existem duas PCHs na área da bacia. A Figura 1 apresenta o mapa de localização da bacia de estudo.

Figura 1: Bacia hidrográfica de estudo.



2.2. Base de Dados

Os dados foram extraídos da série histórica hidrológica da Agência Nacional de Águas e Saneamento Básico (ANA), de frequência diária, entre os anos de 1970 e 2023. Como critério, selecionou-se apenas estações com no máximo 10% de ausência de dados. Nos dados faltantes, procedeu-se com o preenchimento utilizando o algoritmo *K-Nearest Neighbors* (KNN), o qual tem como entrada as estações mais semelhantes para preencher a estação de interesse. Abaixo apresenta-se a relação das estações selecionadas para o estudo.

Tabela 1: Estações fluviométricas e pluviométricas selecionadas para o estudo.

Tipo	Código	Nome	LAT	LONG
Pluviométrica	2144003	CAXAMBU	-21,9897	-44,9386
	2144004	BAEPENDI	-21,9511	-44,8761
	2145001	CONCEIÇÃO DO RIO VERDE	-21,8872	-45,0792
	2145008	FAZENDA JUCA CASIMIRO	-21,8742	-45,2583
	2244068	ITANHANDU	-22,2944	-44,9392
	2244071	POUSO ALTO	-22,1992	-44,9733
	2245065	CRISTINA - MONTANTE	-22,2103	-45,2658
Fluviométrica	61429000	ITANHANDU	-22,2944	-44,9369
	61460000	CONCEIÇÃO DO RIO VERDE	-21,8872	-45,0792
	61473000	BAEPENDI	-21,9511	-44,8797
	61500000	FAZENDA JUCA CASIMIRO	-21,8700	-45,2622
	61510000	TRÊS CORAÇÕES	-21,7031	-45,4750

2.3. LSTM e Critérios de Avaliação de desempenho

Um mecanismo de busca gradeada (*Grid Search*) foi para obtenção do cenário de ajuste ótimo dos hiperparâmetros da LSTM. Assim sendo, como resultado, o modelo é composto por 2 camadas, 64 neurônios, 64 lotes, 200 ciclos de treinamento e em uma estrutura 80% dos dados para treinamento e 20% dos dados para validação. Além disso, adotou-se o otimizador *Adam* a uma taxa de aprendizado de 0,001. O horizonte de dias passados a serem utilizados na previsão é de 14 dias, no intuito de se prever um dia, uma semana, um mês e, por fim, seis meses. A estação alvo a ser prevista é a estação Três Corações (cód. ANA 61510000), localizada no exutório da bacia hidrográfica de estudo. O desempenho das diferentes previsões realizadas pelo modelo é avaliado segundo o Coeficiente de Determinação (R^2), *Kling-Gupta Efficiency* (KGE), *Mean Squared Error* (MSE), *Nash-Sutcliffe Efficiency* (NSE) e Percent bias (PBIAS).

3. RESULTADOS E DISCUSSÕES

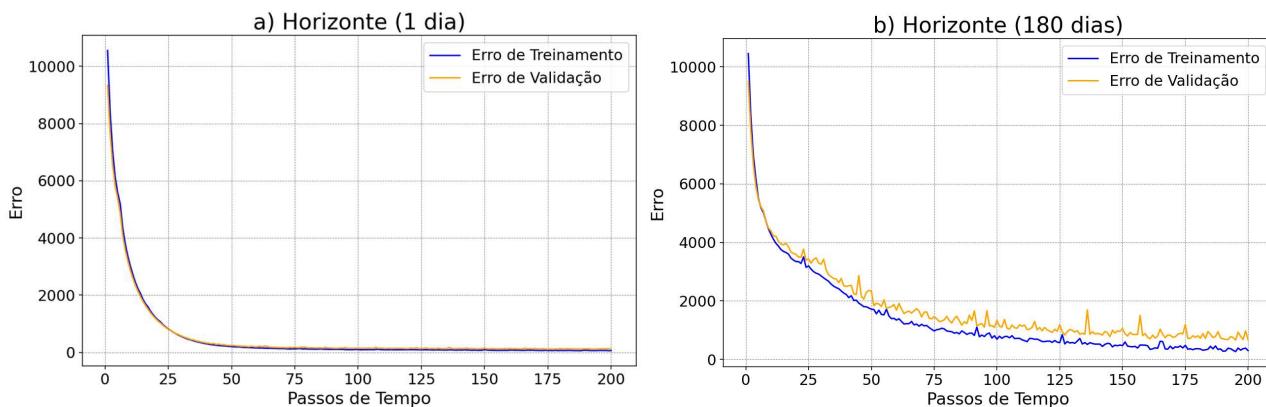
Abaixo é mostrada uma tabela síntese dos resultados obtidos ao longo dos horizontes de um dia, uma semana, um mês e, por fim, seis meses, das métricas escolhidas para o estudo.

Tabela 2: Comparação de desempenho do modelo, na validação, para os horizontes de um dia até seis meses.

Horizonte (dias)	MSE (m^3/s) ²	NSE	KGE	PBIAS (%)	R ²
1	91,70	0,98	0,97	0,10	0,98
7	577,54	0,88	0,90	-0,81	0,87
30	784,85	0,84	0,86	-1,30	0,82
180	1096,91	0,79	0,77	4,69	0,79

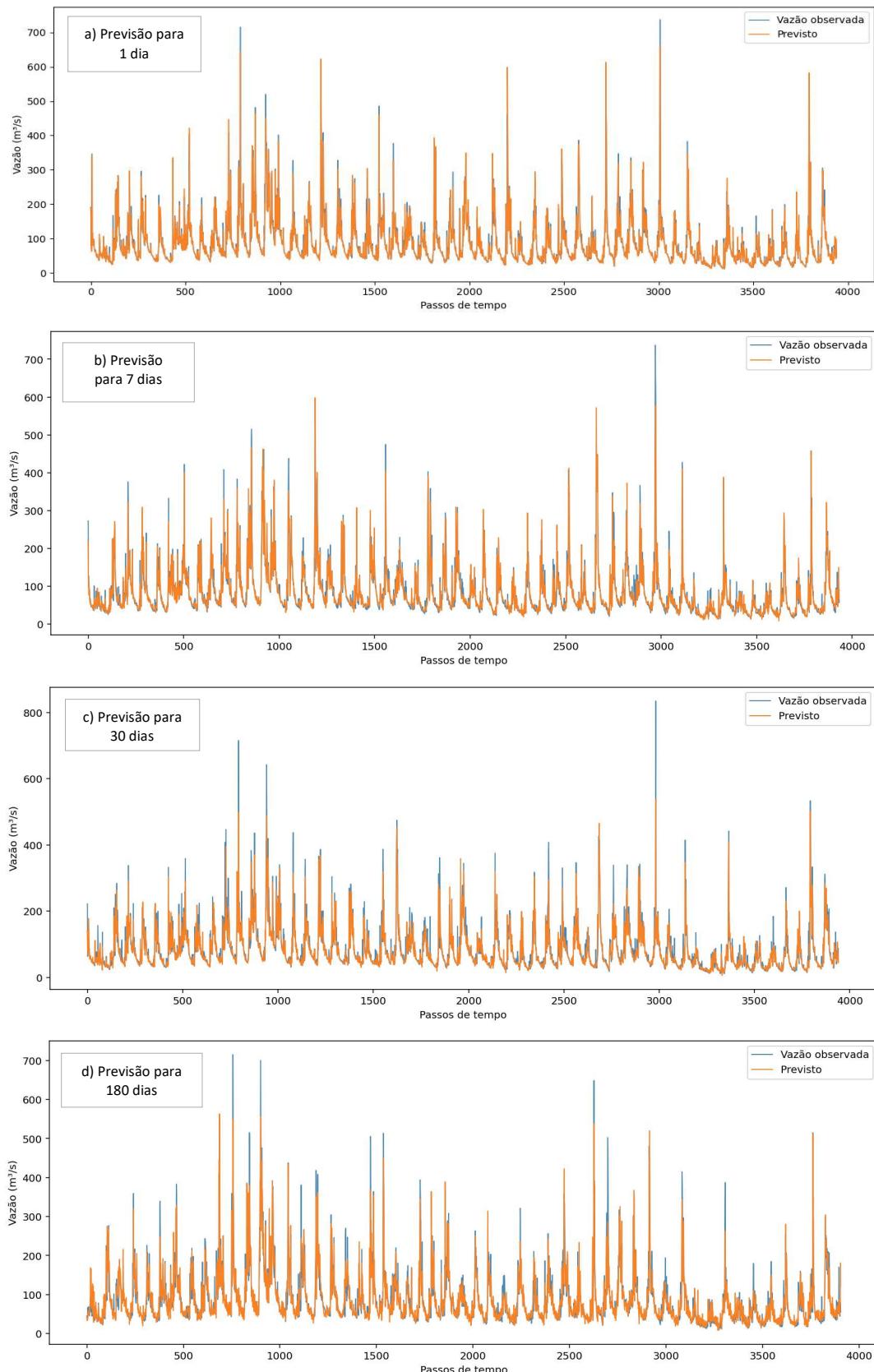
O modelo apresentou um bom ajuste para horizontes de curto prazo, tendo um decréscimo à medida que se aumenta a janela a ser prevista. Todavia, extrai-se que o modelo ainda se revela consistente, mesmo se tratando de seis meses de previsão à frente ($R^2 = 0,79$). Alguns fatores podem vir a explicar o bom desempenho do modelo como o bom monitoramento fluviométrico e pluviométrico da bacia hidrográfica, sob aspecto temporal e espacial, a organização dos dados e a taxa aprendizado do modelo. Abaixo é explanada a taxa de aprendizado do modelo.

Figura 2: Taxa de treino e validação (MSE) para diferentes horizontes de previsão



Como não poderia deixar de ser diferente, a previsão para um dia apresentou uma curva com menor ruído associado, quando comparada à previsão de seis meses. Mesmo por se tratar de uma região de cabeceira, área caracterizada por uma variabilidade temporal e espacial dos processos de transformação chuva-vazão (e.g., solos pouco profundos e rápida formação do escoamento), o modelo conseguiu compreender os padrões hidrológicos e manter o erro de treinamento próximo ao erro de validação. Outra questão foi a representatividade do modelo para valores máximos e mínimos. A seguir são evidenciados os resultados obtidos pelo modelo no período de validação.

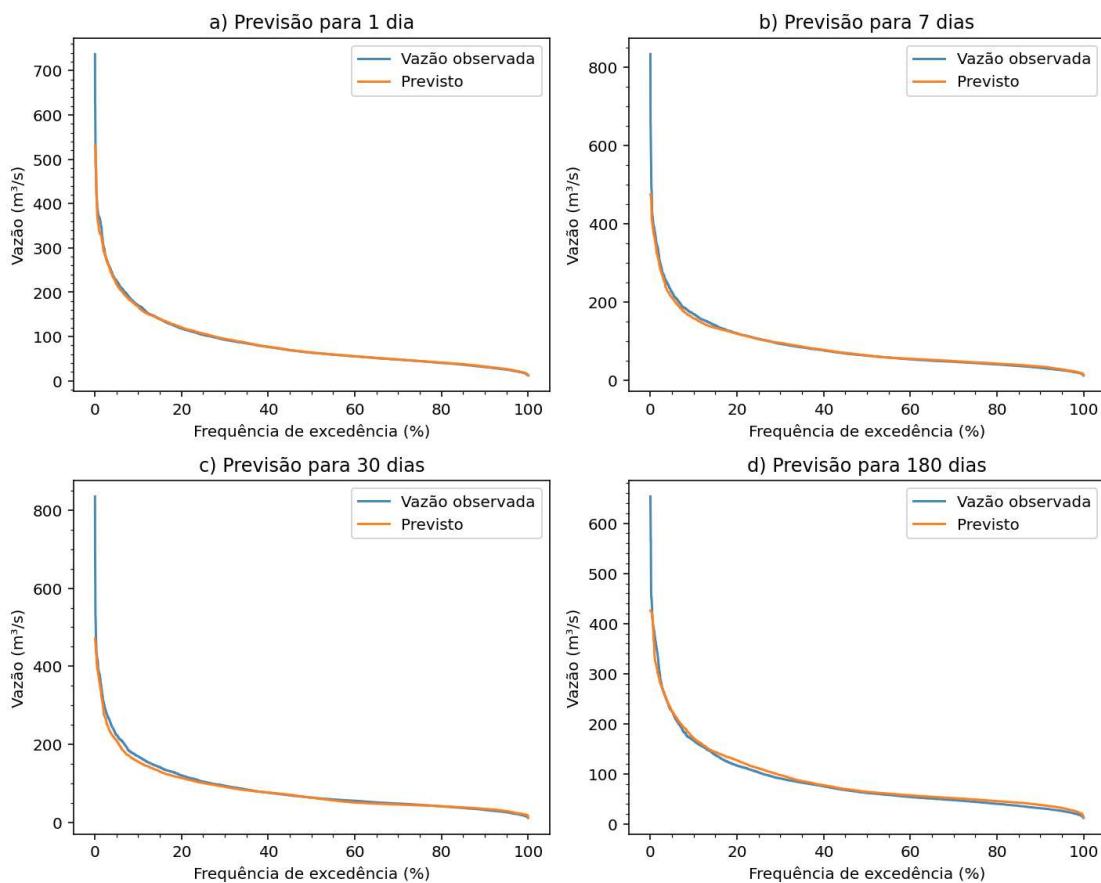
Figura 3: Comparação entre a vazão observada e a vazão prevista na validação



A partir da análise da Figura 3, observa-se a seguinte questão: para a previsão a curto prazo, o modelo conseguiu captar os valores de picos e mínimos, o que não se traduziu nas vazões de longo prazo, em que maiores foram as incertezas. Tal alegação se deve a proporção entre a quantidade de entradas, janela de dias utilizada para entrada, e a quantidade de saídas, dias a serem previstos. Ademais, vale salientar a possibilidade de que o modelo durante o período de treinamento não tenha tido contato com a frequência necessária a valores extremos, o que, de certa maneira, reduz a acurácia na previsão de valores máximos e mínimos de vazão. Outra possibilidade seria mudar a função objetivo para focar a minimização do erro nas vazões mínimas ou máximas (MAN et al., 2023; LI et al., 2024). Isso poderia melhorar os resultados para vazões específicas, a depender do problema a ser modelado (e.g., cheias e secas).

De maneira geral, o modelo indicou representatividade para vazões médias segundo a curva de permanência e, na maioria dos cenários, as vazões mínimas, sendo as maiores diferenças encontradas para o horizonte de previsão de 180 dias. A maior dificuldade encontrada na previsão foi para as vazões máximas (Figura 4). Tal constatação é comum e ocorre devido à pequena representatividade (observações) desses eventos na série histórica.

Figura 4: Comparação entre a curva de permanência observada e a prevista.



4. CONSIDERAÇÕES FINAIS

A LSTM apresentou uma boa representatividade nos horizontes temporais de curto prazo, de acordo com as métricas propostas pelo estudo, mesmo se tratando de uma bacia hidrográfica de cabeceira. No entanto, à medida que se aumentou o horizonte de passos à frente, o modelo apresentou dificuldades na previsão de valores máximos e, em alguns casos, para valores mínimos (principalmente para 180 dias de horizonte de previsão). Isso se deve em razão da taxa de aprendizado, haja vista a quantidade de entradas e saídas do modelo, bem como a variabilidade espacial e temporal da área de estudo e a quantidade de eventos extremos máximos observados. Já para valores médios, segundo a curva de permanência, resultados satisfatórios foram obtidos. Assim, a pesquisa revela o potencial uso da LSTM na previsão de vazões em bacias de cabeceira, como ferramenta de planejamento e gestão de recursos hídricos. Como possível aplicação, há de se citar o setor hidrelétrico, especialmente, em PCHs, haja vista o potencial energético disponível nessas regiões e a demanda do setor por previsões mais assertivas.

AGRADECIMENTOS

À Fapemig pelo financiamento à participação no congresso sob processo número PCE-00429-25.

REFERÊNCIAS

- ALEXANDER, R. B. et al. The Role of Headwater Streams in Downstream Water Quality1. JAWRA Journal of the American Water Resources Association, v. 43, n. 1, p. 41–59, 26 jan. 2007.
- BEHMARD SABZIPOUR et al. Comparing a long short-term memory (LSTM) neural network with a physically-based hydrological model for streamflow forecasting over a Canadian catchment. Journal of hydrology, v. 627, p. 130380–130380, 1 dez. 2023.
- BESAW, L. E. et al. Advances in ungauged streamflow prediction using artificial neural networks. Journal of Hydrology, v. 386, n. 1-4, p. 27–37, maio 2010.
- CASSAGNOLE, M. et al. Impact of the quality of hydrological forecasts on the management and revenue of hydroelectric reservoirs – a conceptual approach, Hydrol. Earth Syst. Sci., 25, 1033–1052, <https://doi.org/10.5194/hess-25-1033-2021>, 2021.
- DAMAVANDI, H. G. et al. Accurate Prediction of Streamflow Using Long Short-Term Memory Network: A Case Study in the Brazos River Basin in Texas. International Journal of Environmental Science and Development, v. 10, n. 10, p. 294–300, 2019.
- HE, M.; LEE, H. Advances in Hydrological Forecasting. Forecasting 2021, 3, 517–519. [https://doi.org/10.3390/ forecast3030032](https://doi.org/10.3390/forecast3030032)

HOCHREITER, S.; SCHMIDHUBER, J. Long Short-Term Memory. *Neural Computation*, v. 9, n. 8, p. 1735–1780, nov. 1997.

HU, Y. et al. Stream-Flow Forecasting of Small Rivers Based on LSTM. *arXiv* (Cornell University), 1 jan. 2020.

KIERAN et al. Using a long short-term memory (LSTM) neural network to boost river streamflow forecasts over the western United States. v. 26, n. 21, p. 5449–5472, 1 nov. 2022.

LAVARINI, C., & MAGALHÃES JR., A. P. (2013). ANÁLISE MORFOMÉTRICA DE BACIAS DE CABECEIRA COMO FERRAMENTA DE INVESTIGAÇÃO GEOMORFOLÓGICA EM MÉDIA E LARGA-ESCALA ESPACIAL. *Revista Brasileira De Geomorfologia*, 14(1). <https://doi.org/10.20502/rbg.v14i1.331>

LI, X. et al. Enhancing hydrological extremes prediction accuracy: Integrating diverse loss functions in Transformer models. *Environmental Modelling & Software*, v. 177, p. 106042, 15 abr. 2024.

LUO, J.; ZHU, D.; LI, D. Classification-enhanced LSTM model for predicting river water levels. *Journal of Hydrology*, v. 650, p. 132535, 14 dez. 2024.

MAIER, H. R. et al. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environmental Modelling & Software*, v. 25, n. 8, p. 891–909, ago. 2010.

MAN, Y. et al. Enhanced LSTM Model for Daily Runoff Prediction in the Upper Huai River Basin, China. *Engineering*, v. 24, p. 229–238, 28 abr. 2023.

MARACAJÁ, J. R.A. Previsão regionalizada de vazão sazonal utilizando redes neurais artificiais, 2005, 122f. Dissertação (Mestrado em Engenharia Civil) - Universidade Federal de Campina Grande, Campina Grande, 2005.

NAGHETTINI, Mauro; PINTO, Éber José de Andrade. *Hidrologia estatística*. Belo Horizonte: CPRM, 2007.

WOHL, E. The significance of small streams. *Frontiers of Earth Science*, v. 11, n. 3, p. 447–456, 6 abr. 2017.

YANG, S. et al. A physical process and machine learning combined hydrological model for daily streamflow simulations of large watersheds with limited observation data. *Journal of Hydrology*, v. 590, p. 125206, nov. 2020.