Antecipação Probabilística de Cheias no Rio Jaguarão (RS) com Redes LSTM e Precipitação por Satélite

Brayan Soares [1] ; Bruno Brentan [2] & André Ferreira [3]

Palavras-Chave – Previsão de Cheias, Redes Neurais LSTM, Precipitação por Satélite, Quantificação de Incerteza, Sistemas de Alerta.

INTRODUÇÃO

A previsão de cheias em bacias hidrográficas brasileiras é dificultada pela distribuição espacial limitada da Rede Hidrometeorológica Nacional, que apesar de reunir milhares de estações pluviométricas e fluviométricas (ANA, 2020), concentra-se majoritariamente em grandes bacias de interesse hidrelétrico, deixando bacias de pequeno e médio porte como a do Rio Jaguarão (RS) com cobertura esparsa de monitoramento em tempo real. Modelos físicos amplamente utilizados na hidrologia operacional demandam parametrização extensiva e séries de dados consistentes de múltiplas variáveis, requisitos frequentemente indisponíveis nesse contexto. As redes neurais do tipo *Long Short-Term Memory* (LSTM) têm demonstrado capacidade de aprender relações não lineares entre precipitação e resposta hidrológica diretamente a partir de dados observados, sem parametrização física explícita, com resultados competitivos ou superiores aos modelos tradicionais (Kratzert *et al.*, 2018). A análise da série histórica de cota do Rio Jaguarão (2015-2024) evidencia aumento expressivo na frequência e intensidade de eventos extremos nos anos mais recentes, reforçando a urgência de ferramentas de previsão mais robustas. Este trabalho apresenta um sistema de previsão probabilística da cota do Rio Jaguarão baseado em redes LSTM, cuja arquitetura incorpora padrões físicos da bacia e mecanismos específicos para corrigir o problema de inércia observado em modelos recorrentes, utilizando precipitação por satélite como forçante principal.

METODOLOGIA

O projeto foi estruturado em duas etapas sequenciais: tratamento e caracterização dos dados, e modelagem preditiva.

Os dados de precipitação foram obtidos do produto MERGE (Rozante *et al.*, 2010), versão V07B, que combina estimativas do satélite GPM-IMERG (NASA) com observações de superfície do INMET, ANA e CEMADEN, cobrindo 10 anos em frequência horária (2015-2024, 87.648 registros) a partir de 30 pixels de aproximadamente 11 km × 11 km distribuídos sobre a bacia. Falhas no sensor físico de nível foram preenchidas por meio da técnica *k-Nearest Neighbors* (*k*-NN) *Imputer*, garantindo série temporal contínua para o treinamento. A partir da série tratada, foram extraídos padrões físicos da bacia como tempo de concentração, taxa de recessão hidrológica e percentis históricos de cota, utilizados para orientar as janelas temporais de entrada, as variáveis derivadas e os limiares da função de perda.

A modelagem foi realizada por meio de uma rede LSTM Residual com dois ramos: o primeiro processa o histórico recente de cota via *Long Short-Term Memory* (LSTM), capturando a dinâmica temporal do rio; o segundo processa a previsão de chuva futura via perceptron multicamadas (MLP). A combinação dos dois ramos gera uma estimativa da variação esperada do nível do rio, convertida

1) Afiliação: [1] Universidade Federal de Minas Gerais (UFMG), Escola de Engenharia, Departamento de Engenharia Ambiental.
E-mail: brayan.mendonca@ufmg.br

em previsão de cota com intervalo de confiança. A arquitetura incorpora três mecanismos de correção do problema de inércia, tendência de modelos recorrentes a replicar o estado atual em vez de detectar mudanças: o alvo de treinamento corresponde à variação do nível e não ao valor absoluto; aplica-se *dropout* assimétrico com maior intensidade no ramo de histórico, incentivando dependência da informação de chuva futura; e a função de perda incorpora penalidade assimétrica para subestimações acima do limiar de atenção, refletindo o custo desproporcional de não detectar uma cheia. São treinados modelos para os horizontes de 1 e 3 dias, com arquitetura *multi-escala* para o horizonte estendido.

A quantificação de incerteza combina *Monte Carlo Dropout* que gera uma distribuição de previsões por meio de múltiplas execuções com *dropout* ativo e *Conformal Prediction* estratificado por regime de cota (normal, atenção e extremo), calibrando empiricamente os intervalos de confiança para cobertura próxima ao nível nominal de 95%.

RESULTADOS

O sistema foi avaliado para os horizontes de 1 e 3 dias, utilizando janeiro de 2023 a dezembro de 2024 como conjunto de validação (20% da série histórica). Para o horizonte de 1 dia, o modelo demonstrou elevada acurácia global (NSE=0,93; KGE=0,925), sem tendência sistemática de superestimar ou subestimar o nível do rio (PBIAS=2,5%). Na detecção de eventos de inundação acima de 4,30 m, o sistema identificou 93,8% das cheias reais antes de ocorrerem, com antecedência média de 23,1 horas e precisão no *timing* do pico de -0,6 hora, ou seja, o momento exato da cheia foi antecipado com menos de 1 hora de erro. O intervalo de confiança cobriu 95,6% das observações reais (PICP), com apenas 15,0% de alarmes falsos (FAR). Para o horizonte de 3 dias, manteve-se desempenho operacionalmente relevante (NSE=0,85; POD=84,3%; FAR=18,1%; PICP=94,7%), com antecedência máxima de 72 horas para mobilização de recursos. A redução de desempenho em relação ao horizonte de 1 dia é esperada e consistente com a literatura (Kratzert *et al.*, 2018). A Tabela 1 resume as métricas obtidas.

Tabela 1 – **Métricas de desempenho do sistema para os horizontes de 1 e 3 dias.**

Métrica	1 dia	3 dias
NSE	0,93	0,85
KGE	0,925	0,913
POD	93,8%	84,3%
FAR	15,0%	18,1%
LEAD TIME MÉDIO	23,1h	72 h
PICP GLOBAL	95,6%	94,7%

Os dois horizontes configuram um sistema de alerta em camadas complementares: o modelo de 1 dia viabiliza o alerta operacional imediato, enquanto o de 3 dias amplia a antecipação para decisões logísticas de maior escopo. No maior evento da série em setembro de 2023 com cota de 9,39 m, o modelo de 1 dia emitiu alerta com 5 horas de antecedência ao cruzamento do limiar, com o intervalo de confiança cobrindo o valor observado em todos os instantes. Esse resultado ilustra a capacidade do sistema de antecipar eventos extremos que extrapolam os padrões históricos médios, ainda que com subestimação da magnitude do pico, limitação esperada em eventos de cauda da distribuição.

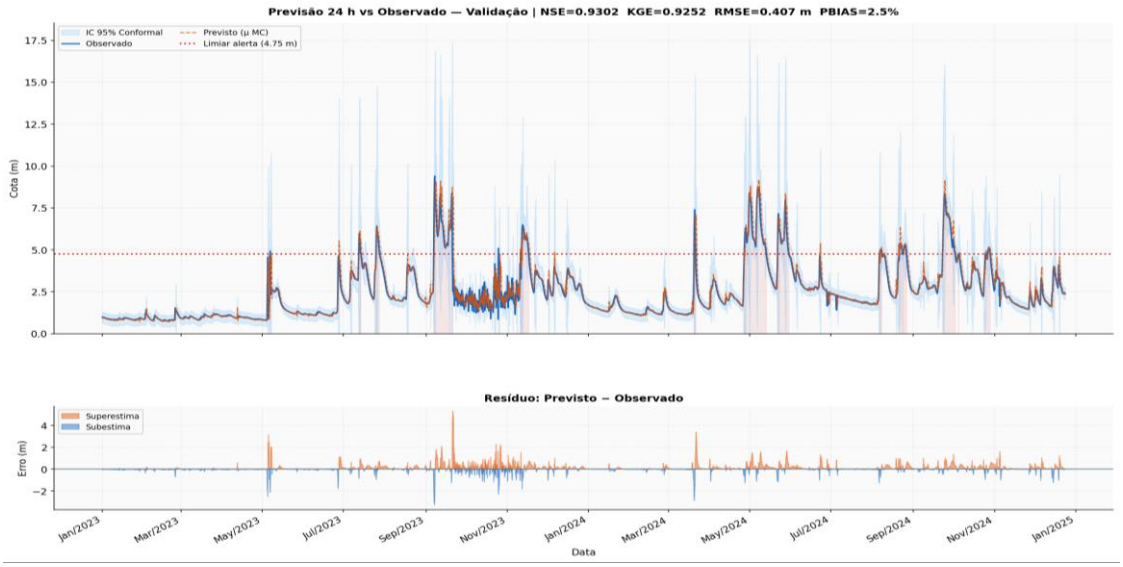


Figura 1 - Hidrograma observado e previsto com intervalo de confiança de 95% para o período de validação (jan/2023 – dez/2024), horizonte de previsão de 1 dia. Fonte: autoria própria.

CONCLUSÕES

A utilização de dados de precipitação por satélite do produto MERGE/GPM-IMERG como forçante principal do modelo demonstrou ser uma alternativa viável para a previsão hidrológica em bacias com monitoramento fluviométrico esparso, contornando a dependência de redes de pluviômetros terrestres de alta densidade. Os mecanismos de correção de inércia implementados, sendo eles o alvo de variação da cota, o *dropout* assimétrico e a função de perda com penalidade assimétrica, combinados ao Conformal Prediction estratificado por regime de cota, resultaram em um sistema de previsão probabilística capaz de antecipar eventos de cheia com elevada precisão temporal e incerteza calibrada. O modelo de horizonte de 1 dia alcançou NSE de 0,93 e probabilidade de detecção de 93,8%, com *lead time* médio de 23,1 horas, demonstrando desempenho consistente com *benchmarks* da literatura para bacias com monitoramento fluviométrico esparso.

O modelo apresentou subestimação da cota de pico no evento mais extremo da série (setembro de 2023, 9,39 m), indicando dificuldades na extrapolação de eventos além dos padrões históricos de treino. As métricas de validação foram obtidas com previsão de precipitação futura derivada de *proxy* histórico com base na própria série do MERGE, o que tende a superestimar o desempenho esperado em operação real com previsão meteorológica numérica, contexto em que os dados de satélite apresentam incertezas próprias, especialmente em eventos convectivos de alta intensidade.

Como perspectivas imediatas, destacam-se a consolidação do modelo para o horizonte de 3 dias, com POD de 84,3% e FAR de 18,1% já obtidos, e a integração com previsões numéricas de tempo do INMET e CEMADEN para operacionalização do sistema de alerta. Em perspectiva de longo prazo, a delimitação da área de contribuição real da bacia para ponderação espacial dos pixels de satélite, a incorporação de variáveis físicas de solo e geologia e a extensão da metodologia para outras bacias hidrográficas brasileiras com escassez de monitoramento representam caminhos promissores para aprimoramento e ampliação do sistema.

REFERÊNCIAS

ANA. AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO (2020). Catálogo de Metadados — Rede Hidrometeorológica Nacional (RHN). Disponível em: https://metadados.snirh.gov.br. Acesso em: 13 jun. 2026.

CUI, Z.; GUO, S.; ZHOU, Y.; WANG, J. (2023). "Exploration of dual-attention mechanism-based deep learning for multi-step-ahead flood probabilistic forecasting". Journal of Hydrology, v.622, 129688.

KRATZERT, F.; KLOTZ, D.; BRENNER, C.; SCHULZ, K.; HERRNEGGER, M. (2018). "Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks". Hydrology and Earth System Sciences, v.22, n.11, p.6005–6022.

LI, D.; MARSHALL, L.; LIANG, Z.; SHARMA, A.; ZHOU, Y. (2021). "Bayesian LSTM with stochastic variational inference for estimating model uncertainty in process-based hydrological models". Water Resources Research, v.57, e2021WR029772.

ROZANTE, J. R. (2024). "Improvements in the MERGE Precipitation Product for South America". Remote Sensing, v.16, n.24, 4722.

ROZANTE, J. R.; MOREIRA, D. S.; GONÇALVES, L. G. G.; VILA, D. A. (2010). "Combining TRMM and Surface Observations of Precipitation: Technique and Validation over South America". Weather and Forecasting, v.25, n.3, p.885–894.

SZELĄG, B.; KOWAL, P.; KICZKO, A.; BIAŁEK, A.; WAŁEK, G.; MAJEREK, D.; SIWICKI, P.; FATONE, F.; BOCZKAJ, G. (2023). "Integrated model for the fast assessment of flood volume: Modelling – management, uncertainty and sensitivity analysis". Journal of Hydrology, v.625, 129967.

SZELĄG, B.; MAJEREK, D.; EUSEBI, A. L.; KICZKO, A.; DE PAOLA, F.; MCGARITY, A.; WAŁEK, G.; FATONE, F. (2024). "Tool for fast assessment of stormwater flood volumes for urban catchment: A machine learning approach". Journal of Environmental Management, v.355, 120214.

WANG, S.; XU, O. (2024). "Mountain flood level forecasting in small watersheds based on recurrent neural networks and multi-dimensional data". IEEE Access, v.12, p.82706–82719.

WANG, S.; XU, O. (2025). "Confidence interval forecasting model of small watershed flood based on compound recurrent neural networks and Bayesian". PLOS ONE, v.20, n.4, e0321583.

AGRADECIMENTOS

O autor agradece aos professores Bruno Brentan e André Ferreira pela orientação e suporte ao longo do desenvolvimento deste trabalho, e ao grupo de pesquisa Hydro-UAI pelo ambiente colaborativo. Agradece também ao Instituto Nacional de Pesquisas Espaciais (INPE) pela disponibilização do produto MERGE V07B. Por fim, agradece à sua mãe Erika Soares e à sua avó Lucia Silva, e a todos os familiares e amigos pelo apoio e incentivo durante essa jornada.