

XXVI SIMPÓSIO BRASILEIRO DE RECURSOS HIDRÍCOS

ACOPLAMENTO ENTRE MODELAGEM CONCEITUAL CHUVA-VAZÃO E APRENDIZADO DE MÁQUINAS PARA PREVISÃO DE VAZÃO HORÁRIA EM TEMPO REAL

Felipe Treistman¹ ; Lucas de Souza Khenayfis²; Rodrigo José Coelho Pereira³; William Cossich Marcial de Farias⁴; Bibiana Rodrigues Colossi⁵; Érica Couto Pereira Dos Santos⁶

Abstract – Operators responsible for real-time operation of the Brazilian National Interconnected System – SIN are often required to make critical choices within extremely limited timeframes. Although the daily programming phase provides valuable support, final decisions still rely heavily on human interpretation and judgment. In Southern Brazil, the low regulation capacity of reservoirs, rapid fluctuations in streamflows, and the occurrence of extreme floods pose additional challenges to real-time operation. To support decision-making under these conditions, a real-time streamflow forecasting model was developed as a pilot application for the Foz do Areia Hydropower Plant. The model provides hourly-discretized forecasts with a two-day lead time, designed for implementation in control rooms. It employs state-of-the-art machine learning techniques—specifically the LightGBM algorithm—with daily natural streamflow forecasts from the SMAP/ONS model, as well as hourly observed and forecasted precipitation and streamflow data from gauging stations. The model was trained on data from Jan/2020 to Dec/2023 and validated using data from the full year of 2024. Validation results demonstrate the model's performance gains, achieving a mean MAPE (NSE) of 5% (0,98) for the first six hours and 14% (0,86) for the final six hours of the forecast horizon. The model reliably anticipates flood events and maintains high accuracy during recession periods.

Resumo – Os operadores responsáveis pela programação da operação em tempo real do Sistema Interligado Nacional – SIN frequentemente precisam tomar decisões, e possuem pouquíssimo tempo para tal. Apesar do auxílio e confiabilidade que a programação da operação diária pode fornecer, as decisões finais dependem muitas vezes da análise e intervenção humana. No Sul do Brasil, a baixa capacidade de regularização de vazões dos reservatórios, as variações rápidas das vazões afluentes e a ocorrência de cheias extremas impõe dificuldades adicionais à operação em tempo real na região. Visando auxiliar a tomada de decisão dos operadores foi desenvolvido um modelo de previsão de vazão com aplicação piloto na bacia da UHE Foz do Areia, capaz de fornecer previsões em tempo real com discretização horária e horizonte de dois dias, voltado ao uso na sala de controle do Centro de Operação Sul do ONS. O modelo proposto combina o estado-da-arte de aprendizado de máquinas, empregando o modelo *LightGBM*, em conjunto com previsões de vazão natural diárias proveniente do modelo conceitual chuva-vazão SMAP/ONS, além de dados de vazão observada de estações fluviométricas, precipitação observada e prevista em base horária. Para calibração do modelo foram utilizados dados entre jan/2020 e dez/2023, e para validação, dados de 2024. Os resultados da etapa de validação demonstram o potencial do modelo proposto, onde foi obtido um MAPE (NSE) médio para as primeiras seis horas de 5% (0,98), e para as últimas seis horas do horizonte de 14% (0,86). O

1) Operador Nacional do Sistema Elétrico - ONS, rua Júlio do Carmo, 251 – Centro, Rio de Janeiro – RJ. < felipe.treistman@ons.org.br >

2) Operador Nacional do Sistema Elétrico - ONS, rua Júlio do Carmo, 251 – Centro, Rio de Janeiro – RJ. < lkhenayfis@ons.org.br >

3) Operador Nacional do Sistema Elétrico - ONS, rua Júlio do Carmo, 251 – Centro, Rio de Janeiro – RJ. < rodrigo.pereira@ons.org.br >

4) Operador Nacional do Sistema Elétrico - ONS, rua Júlio do Carmo, 251 – Centro, Rio de Janeiro – RJ. < william.farias@ons.org.br >

5) Operador Nacional do Sistema Elétrico - ONS, Rod. José Carlos Daux, 4756, Saco Grande, Florianópolis – SC. < bibiana.colossi@ons.org.br >

6) Operador Nacional do Sistema Elétrico - ONS, rua Júlio do Carmo, 251 – Centro, Rio de Janeiro – RJ. < erica.santos@ons.org.br >

modelo é capaz de antecipar as cheias com previsibilidade adequada, além de possuir precisão elevada durante as recessões.

Palavras-Chave – Previsão hidrológica; Operação em tempo real; Aprendizado de máquina.

INTRODUÇÃO

O planejamento e a operação de sistemas elétricos de grande porte, como o Sistema Interligado Nacional (SIN), apoiam-se fortemente em previsões e geração de cenários sintéticos precisos de vazão às usinas hidrelétricas. Apesar do auxílio e confiabilidade que a planejamento e programação da operação diária, proveniente da cadeia de modelos de planejamento energético (Maceira *et al.*, 2018), pode fornecer, as decisões finais dependem sempre da análise e intervenção humana. No contexto do Centro de Operação do Sistema Sul (OSL), o baixo nível de regularização das cascatas de usinas Hidrelétricas (UHE), as variações rápidas das vazões afluentes, e cheias extremas são alguns fatores extras que impõe dificuldades à operação em tempo real nessa região.

Atualmente, o modelo SMAP/ONS (ONS, 2024; Treistman *et al.* 2023) fornece previsões de vazão natural diária para a programação da operação. Esta abordagem de previsão foi concebida com o objetivo de uso na cadeia de modelos de planejamento. No entanto, a previsão de vazão natural em escala diária nem sempre atende plenamente às demandas específicas da operação em tempo real.

Apesar dos avanços em modelos hidrológicos conceituais e sistemas de previsão numérica de tempo, a latência entre as previsões e a capacidade de capturar dinâmicas não lineares e variações locais em bacias complexas, onde relações simples chuva–vazão podem se degradar em eventos de curta duração ou regimes de escoamento atípicos, continuam sendo desafios para operação em tempo real. Por outro lado, técnicas de aprendizado de máquina, especialmente algoritmos de *gradient boosting*, têm demonstrado alto desempenho em cenários hidrológicos ao capturar estruturas não lineares e explorar grandes volumes de dados históricos (ZOUNEMAT-KERMANI *et al.*, 2021).

OBJETIVOS

O presente artigo tem como objetivo apresentar o modelo de previsão de vazão para salas de controle, em etapa piloto para UHE Governador Bento Munhoz (anteriormente UHE Foz do Areia), destinado ao uso no Centro de Operação do Sistema Sul. O modelo possui discretização horária, sendo executado a cada hora, e horizonte de dois dias, visando apoiar a tomada de decisão na operação em tempo real, auxiliar na detecção antecipada de cheias e na otimização da geração hidroenergética em atendimento à demanda.

METODOLOGIA

Nesta seção, são descritos em detalhes os procedimentos adotados para a construção, calibração e validação do modelo predictor de vazão horária. A metodologia engloba a aquisição e pré-processamento dos dados hidrológicos e meteorológicos, a engenharia de variáveis, e a configuração da topologia de modelos de aprendizado de máquina (LGBM – *LightGBM*).

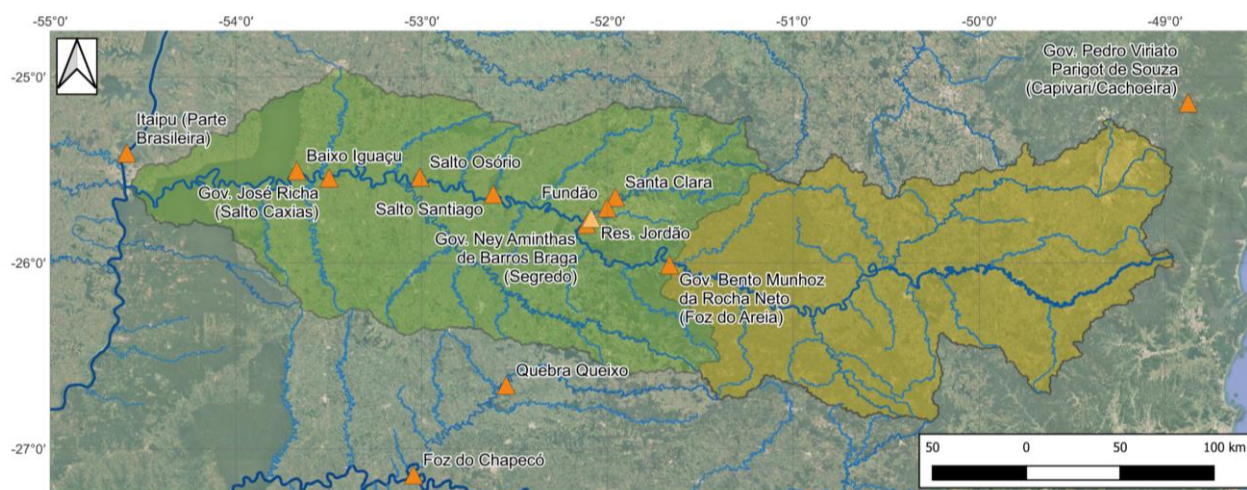
Caracterização geral da bacia incremental da UHE Foz do Areia

A bacia hidrográfica do rio Iguaçu cobre parte dos territórios dos estados do Paraná e Santa Catarina e da Argentina, e possui aproximadamente 67.500 km² de área de drenagem até a confluência com o rio Paraná logo a jusante da UHE Itaipu Binacional. Na bacia do rio Iguaçu estão localizados 9 aproveitamentos hidroelétricos componentes do SIN, cujas principais informações podem ser vistas na Tabela 1. A bacia hidrográfica do rio Iguaçu está apresentada na Figura 1, com destaque para os aproveitamentos hidrelétricos de interesse do SIN.

Tabela 1 – Características dos aproveitamentos hidroelétricos do SIN na bacia do rio Iguaçu.

Aproveitamento	Corpo Hídrico	Agente	Potência Instalada (MW)	Início do Enchimento	Início da Operação	Área de Drenagem (km²)
UHE Santa Clara - PR	Jordão	COPEL	120,2	21/04/2005	31/07/2005	3908
UHE Fundão	Jordão	COPEL	120,2	20/05/2006	23/06/2006	4092
RES Jordão	Jordão	COPEL	-	24/04/1996	09/05/1996	4672
UHE Gov. Bento Munhoz da Rocha Netto (Foz do Areia)	Iguaçu	COPEL	1676,0	01/04/1980	22/09/1980	30150
UHE Gov. Ney Aminthas de Barros Braga (Segredo)	Iguaçu	COPEL	1260,0	05/07/1992	26/09/1992	34387
UHE Salto Santiago	Iguaçu	Engie	1420,0	29/09/1979	03/01/1981	43907
UHE Salto Osório	Iguaçu	Engie	1078,0	05/06/1975	14/10/1975	45839
UHE Gov. José Richa (Salto Caxias)	Iguaçu	COPEL	1240,0	01/10/1998	01/03/1999	57077
UHE Baixo Iguaçu	Iguaçu	Neoenergia	350,1	14/12/2018	31/01/2019	61644

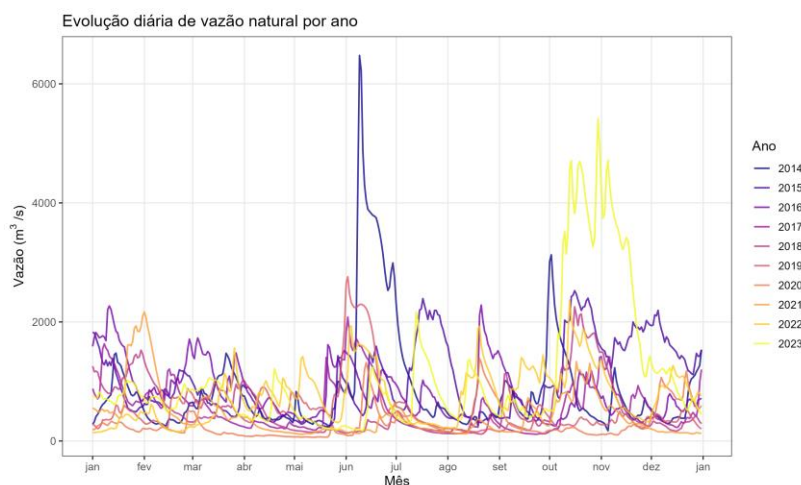
Figura 1 – Bacia hidrográfica do rio Iguaçu.



Na cabeceira desta bacia do Rio Iguaçu está situada a região metropolitana de Curitiba, com predominância de atividades industriais e comerciais. As formações geológicas, relevos e solos são diversificados. O clima da região é subtropical, apresentando chuvas bem distribuídas ao longo do ano (GOVERNO DO PARANÁ, 2025). Consequentemente, esta bacia hidrográfica não apresenta comportamento hidrológico sazonal, como pode ser observado nas bacias dos subsistemas Sudeste, Norte-Centro Oeste e Nordeste.

A Figura 2 mostra os hidrogramas ano a ano de vazões diárias naturalizadas a partir do processo de reconstituição realizado pelo ONS (2025b). Pode-se observar a ausência de sazonalidade das vazões e a presença de eventos de cheia considerados rápidos se comparados com outras bacias de interesse do SIN. Por este motivo se faz necessário o acompanhamento em tempo real das vazões afluentes e o aprimoramento da previsão para a operação das usinas hidrelétricas desta bacia.

Figura 2 – Hidrograma de vazões naturais diárias da UHE Governador Bento Munhoz (Foz do Areia).



Dados fluviométricos

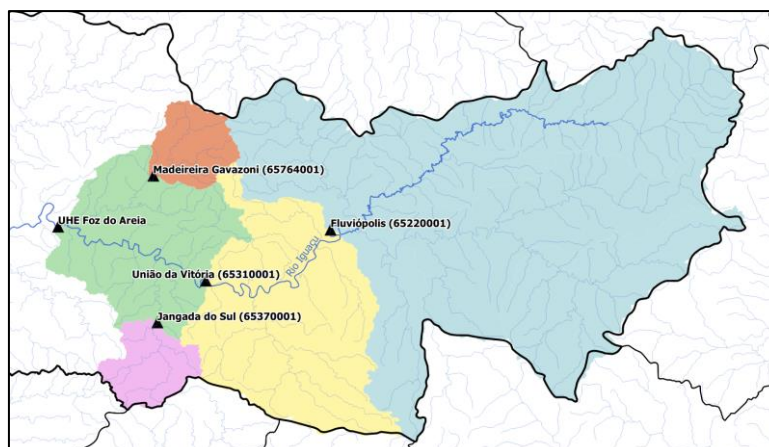
Para a modelagem hidrológica da bacia do rio Iguaçu até a UHE Foz do Areia, foi realizada inicialmente uma análise da rede hidrometeorológica disponível, considerando critérios como localização, qualidade e dados disponíveis em tempo real. A partir dessa triagem, foi definido um conjunto final de quatro estações fluviométricas localizadas dentro da área de contribuição da bacia. Os dados fluviométricos passaram por um processo de tratamento com o intuito de se obter séries de vazões horárias com o maior grau possível de consistência. Inicialmente, foi realizada uma análise visual das séries históricas de leituras de régua e vazões disponíveis para eliminar inconsistências evidentes. Na sequência, avaliou-se a representatividade das medições de descarga líquida sobre a relação cota-vazão, verificando a aderência às curvas-chave oficiais de cada estação.

Após essa etapa preliminar, foram aplicadas técnicas de correlação cruzada entre as estações selecionadas para preenchimento de falhas de observação, e analisadas as defasagens temporais entre estações consecutivas de modo a obter um melhor entendimento da hidrodinâmica dos trechos de rio analisados. Finalmente, foram comparados sistematicamente os fluviogramas em pares de estações montante-jusante a fim de eliminar erros grosseiros, como incrementais negativas. Essas etapas consolidaram um conjunto final de dados fluviométricos consistentes e contínuos que foram utilizados na modelagem. A Figura 3 ilustra esquematicamente a localização da seleção final de estações na bacia, apresentando também a correlação com a UHE G. B. Munhoz.

Tabela 2 - Estações fluviométricas utilizadas no modelo.

Código	Nome	Curso d'água	Área de drenagem (km²)	Latitude (°)	Longitude (°)	Responsável/ Operador	Correlação com a UHE G. B. Munhoz
65220001	UHE Gov. Bento Munhoz Fluiópoli	Rio Iguaçu	18600	-26,0192	-50,5925	F.D.A.	0.89
65310001	UHE Gov. Bento Munhoz União da Vitória	Rio Iguaçu	24200	-26,2281	-51,0803	F.D.A.	0.95
65370001	UHE Gov. Bento Munhoz Jangada do Sul	Rio Jangada	1010	-26,3869	-51,2719	F.D.A.	0.73
65764001	UHE Gov. Bento Munhoz Madeireira Gavazoni	Rio da Areia	976	-25,8075	-51,2886	F.D.A.	0.77

Figura 3— Localização das quatro estações fluviométricas selecionadas para o estudo.



Dados Meteorológicos

As precipitações observadas foram extraídas do produto MERGE (Rozante, *et al.* 2020), disponibilizado pelo INPE/CPTEC. Esse dado possui uma latência de aproximadamente 3 h 30 min. As previsões numéricas de precipitação foram obtidas do modelo global ECMWF, com a mesma grade espacial do MERGE, disponibilizadas às 00 UTC e 12 UTC, com latência de cerca de 6 h. Todas essas informações são levadas em consideração na hora de estimar os modelos, evitando que seja feita a estimação de modelos considerando dados que não estariam disponibilizados levando em consideração o atraso do recebimento dessas informações para a operação em tempo real.

Dados de vazão natural diária prevista

Com base nos dados operativos hidráulicos das UHEs do SIN, o ONS realiza a naturalização das vazões nos pontos dos aproveitamentos, conforme explicitado no Submódulo 2.4 dos Procedimentos de Rede (ONS, 2025a). Estas vazões são aquelas que ocorreriam no trecho de rio se não houvesse o efeito dos usos consuntivos, diferenças na evaporação do reservatório em relação àquela que existiria sem o aumento da superfície de água, e sem os efeitos da operação dos reservatórios.

As vazões naturais, reconstituídas em escala diária, são diariamente publicadas e alimentam o modelo hidrológico concentrado SMAP/ONS (ONS, 2024). A previsão de vazões naturais para os aproveitamentos do SIN é insumo para a programação da operação em base diária e semanal.

Modelo de previsão de vazão horária

Em contextos hidrológicos, modelos de aprendizado de máquina baseados em reforço de gradiente (*Gradient Boosting*) têm demonstrado robustez para capturar relações não lineares entre precipitação, vazões antecedentes e regime de cheia. *Gradient Boosting* é uma técnica de aprendizado por *ensemble* que tem se tornado cada vez mais popular, onde funções simples (h_m), chamadas de “aprendiz fraco”, são sucessivamente estimadas, até que algum critério de convergência seja satisfeito. Entre os modelos desse tipo, o modelo *Light Gradient Boost Machine* - LGBM (Ke, et al. 2017) tem se destacado em diversos contextos, tanto pela qualidade das previsões fornecidas, quanto por sua eficiência em termos de velocidade e uso de memória (SZCZEPANEK, 2022). No modelo LGBM, utilizam-se como aprendizes fracos árvores de decisão.

Modelos de árvores de decisão particionam o espaço de entrada \mathbb{R}^p em regiões \mathbb{R}_m por meio de divisões binárias sequenciais, de modo a minimizar uma função de custo (por exemplo, o erro quadrático médio) em cada nó interno. Cada partição é definida por um teste do tipo:

$$x_j \leq s \quad (1)$$

onde x_j é a j -ésima feature e s um valor de corte. A predição final de uma árvore simples (aprendiz fraco) é dada por constantes γ_m atribuídas a cada região \mathbb{R}_m :

$$f(x) = \sum_{m=1}^M \gamma_m \mathbf{1}\{x \in \mathbb{R}_m\} \quad (2)$$

O *Gradient Boosting* combina em série K dessas árvores, ajustando cada nova árvore ao gradiente da função de perda $L(y, \hat{y})$ residual do ensemble anterior. A função de perda empregada no nosso caso é o erro quadrático médio (MSE). Em cada iteração k , estima-se:

$$h_k(x) = \arg \min_h \sum_{i=1}^n \left[-g_i^{(k-1)} \cdot h(x_i) + \frac{1}{2} h(x_i)^2 \right], \quad (3)$$

onde $g_i^{(k-1)} = \partial L(y_i, \hat{y}_i^{(k-1)})$ é o gradiente para a i -ésima observação. A predição do ensemble é então atualizada por:

$$\hat{y}_i^{(k)} = \hat{y}_i^{(k-1)} + \nu h_k(x_i), \quad (4)$$

sendo ν a taxa de aprendizado.

No modelo desenvolvido para a previsão de vazão horária em tempo real, estima-se um modelo LGBM para cada horizonte de previsão da hora $h+1$ até $h+48$. Ao treinar modelos específicos por horizonte, consegue-se capturar nuances distintas de importâncias de variáveis em horizontes curtos versus horizontes longos. Para o processo de estimação dos modelos, foram utilizados todos os dados entre 2020 e 2023, separando-se o ano de 2024 para a etapa de validação. Como variáveis de entrada, além dos dados brutos descritos anteriormente, vazão observada na UHE e nos postos fluviométricos, precipitação observada e prevista, e previsão de vazão natural diária, foi realizado um trabalho de “engenharia de variáveis”, que contempla a extração e transformação das variáveis brutas hidrológicas e meteorológicas em insumos (*features*) robustos para o algoritmo de aprendizado de máquina.

Foram utilizadas como transformações defasagens temporais ($h-1$ a $h-6$) e médias móveis, para capturar a dinâmica autorregressiva da série de vazão, bem como o tempo de resposta hidrológico da bacia, somas cumulativas (de 3 horas a 72 horas), para representar o total de chuva acumulada recente e prevista. Adicionalmente, foi utilizada a transformada de *wavelet* discreta (Heil & Walnut, 1989), para extrair componentes de frequência específicas da série de vazão, tais como oscilações sazonais de baixa frequência ou ruídos de alta frequência. Ao todo, o modelo possui cerca de 230 variáveis de entrada. A Tabela 3 apresenta exemplos de variáveis finais geradas após as transformações.

Tabela 3 – Exemplos de variáveis finais utilizadas no treinamento dos modelos.

Tipo de Variável	Notação	Descrição
Vazão defasada (1 h)	$Q_{\text{GBM}}(t-1)$	Vazão observada 1 hora antes na UHE
Média móvel (4 h)	$Q_{\text{posto}}(t-4:t-1)$	Média das 4 últimas horas de vazão do posto fluviométrico
Somatório de precipitação observada (24 h)	$P_{\text{obs}}(t-1:t-24)$	Soma das precipitações observadas nas últimas 24 horas
Somatório de precipitação prevista (48 h)	$P_{\text{prev}}(t+1:t+48)$	Soma das precipitações previstas para as próximas 48 horas
Coefficiente wavelet nível 2	$W_{2,\text{GBM}}$	Coefficiente de wavelet da DWT de nível 2 para GBM
Previsão diária SMAP/ONS (dia $d+1$)	$Q_{\text{SMAP/ONS}}(d+1)$	Vazão natural prevista para o dia seguinte pelo SMAP/ONS

RESULTADOS

Avaliação da estimação dos modelos

Nesta subseção, apresentamos a análise comparativa da contribuição dos cinco principais grupos de variáveis – vazão observada (vazao), vazão prevista pelo modelo SMAP/ONS, vazão incremental observada nas estações fluviométricas (posto_flu_obs), precipitação observada (prec_obs) e precipitação prevista (prec_prev) – aos modelos ajustados para cada horizonte de previsão ($h = 1 \dots 48$). A interpretação baseia-se em três métricas extraídas do *LightGBM*: ganho percentual: soma do ganho de informação por variável, normalizada em cada horizonte; cobertura percentual: proporção de observações afetadas pelas divisões que utilizam a variável; frequência de splits: percentual de vezes que a variável foi usada para particionar nós nas árvores.

Deve ser ressaltado que mesmo que alguma variável tenha um ganho relativamente pequeno, isso não quer dizer necessariamente que tenha uma baixa relevância nas previsões. Algumas variáveis podem, por exemplo, ser mais importantes em contextos específicos como picos de cheia, mas ter pouca relevância em outras situações, e assim possuir um ganho pequeno quando avalia-se todo o período simulado. A Figura 4 apresenta o ganho para cada conjunto de variáveis ao longo dos 48 horizontes. De maneira geral, o comportamento do ganho confirma a natureza híbrida do modelo: autodependência forte no curtíssimo prazo e transição para dependência em previsões externas (SMAP/ONS) em horizontes médios e longos.

A Figura 5a e 5b, apresentam a cobertura das árvores de decisão, e a frequência de divisão das árvores para cada conjunto de variáveis ao longo dos 48 horizontes. De maneira geral, os postos fluviométricos detêm a maior cobertura em todos os horizontes (cerca de 30–40 %), indicando que, embora não seja o grupo de maior ganho, participa de divisões que abrangem grande parte dos dados. Já a previsão do SMAP/ONS e vazão observada na UHE apresentam coberturas intermediárias (aproximadamente 15–25 %), crescendo suavemente com o horizonte para SMAP/ONS e decaindo para vazão observada da UHE

Em resumo, os modelos de Curto Prazo ($h \leq 6$) possuem um comportamento fortemente autorregressivos. A prioridade é a vazão observada e as séries incrementais dos postos, justificada pela alta correlação temporal imediata. Para o Médio Prazo ($7 \leq h \leq 24$) há uma transição progressiva. O ganho de SMAP/ONS supera a vazão, indicando que incorporar previsões em base diária melhora a capacidade de antecipar variações hidrológicas além de algumas horas. Já para os modelos de Longo Prazo ($h > 24$), nota-se uma maior dependência do SMAP/ONS. A autodependência perde força, e a previsão de vazão natural diária torna-se a variável-chave. Embora de baixo ganho e cobertura, ambas precipitações (observada e prevista) mantêm influência constante – particularmente úteis para modelar picos repentinos em horizontes curtos e intermediários.

Figura 4— Ganho percentual por horizonte e conjunto de variáveis. Colunas vermelha, marrom, verde, azul e magenta representam os conjuntos de variáveis dos postos fluviométricos, precipitação observada, precipitação prevista, previsões do SMAP/ONS e a vazão observada da UHE, respectivamente.

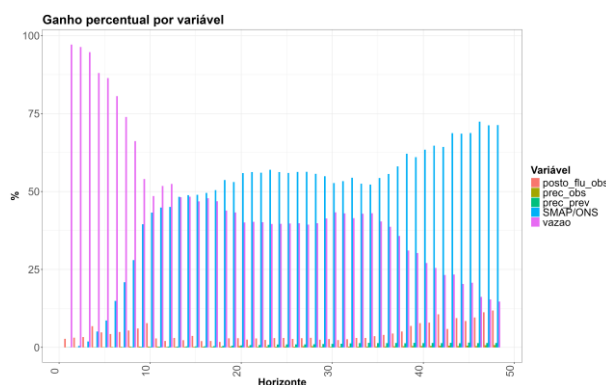
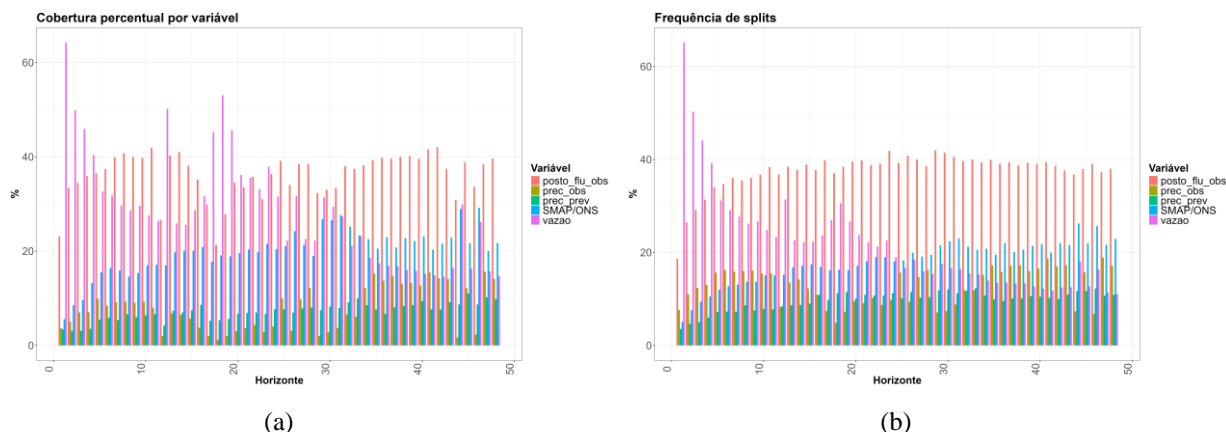


Figura 5– Frequência de divisão das árvores de decisão (a) e cobertura percentual por horizonte e conjunto de variáveis. Colunas vermelha, marrom, verde, azul e magenta representam os conjuntos de variáveis dos postos fluviométricos, precipitação observada, precipitação prevista, previsões do SMAP/ONS e a vazão observada da UHE, respectivamente.



Avaliação da etapa de validação

Para a etapa de validação foram executadas previsões para cada hora do ano de 2024, resultado em uma amostra de 8.774 casos. Para a avaliação serão utilizadas as seguintes métricas: erro médio percentual absoluto (MAPE), coeficiente de eficiência de Nash-Sutcliffe (NSE) e raiz do erro médio quadrático (RMSE).

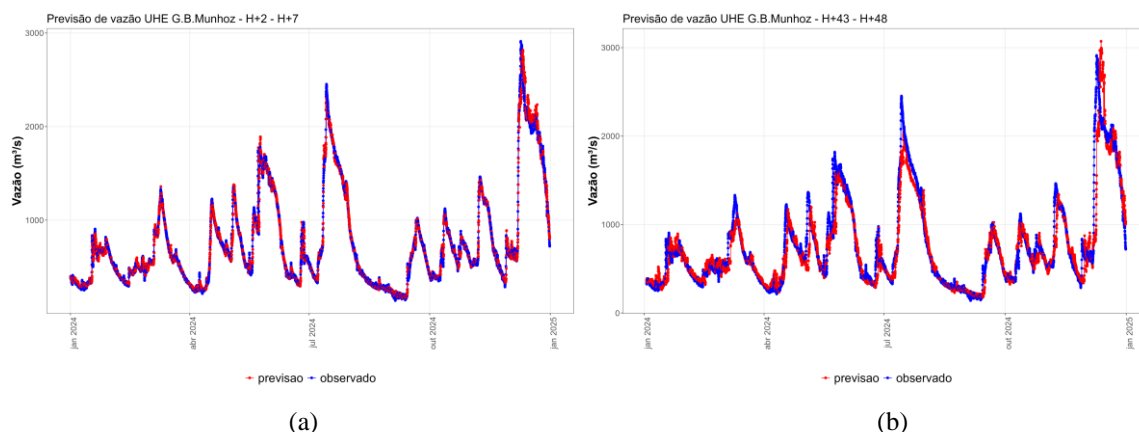
Para o cálculo das métricas, são considerados as médias das próximas 6 horas, tanto para a previsão quanto para a vazão observada, visando diminuir o ruído existente nessas séries. A Tabela 4 apresenta o resumo das métricas para os distintos horizontes de avaliação. Pode-se observar baixos níveis de erro nos horizontes iniciais, que vão se degradando, conforme o esperado, ao longo do tempo.

Tabela 4 – Métricas de erro da etapa de validação.

Horizonte	MAPE	NSE	RMSE
h+1 - h+6	4.2%	0.99	52.47
h+7 - h+12	5.8%	0.97	95.01
h+13 - h+18	6.8%	0.94	120.41
h+19 - h+24	7.9%	0.93	134.65
h+25 - h+30	9.4%	0.91	153.46
h+31 - h+36	10.2%	0.89	167.04
h+37 - h+42	11.1%	0.87	180.99
h+43 - h+48	12.2%	0.86	192.81

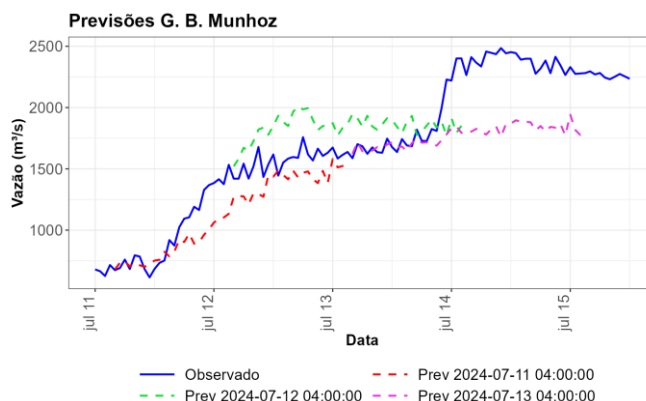
As Figura 6 (a) e (b) apresentam, respectivamente as previsões médias para as horas h+2 e h+7, e para as horas h+43 e h+48. Nota-se que para horizontes mais curtos, o modelo consegue prever adequadamente os picos das cheias. Já para previsões de mais longo prazo, em geral, cheias são subestimadas. Porém, é possível observar na cheia de dezembro de 2024 que o modelo indica em diversos casos, para h+43 e h+48, previsões bastante superestimadas durante o início da recessão. Ressalta-se que para o período de estimação do modelo, entre os anos de 2020 e 2023, houve apenas um evento desse porte. Esse fato possivelmente levou a um desempenho inferior do modelo, dado que a amostra utilizada na estimação não foi representativa o suficiente para “aprender” o comportamento da bacia nessa situação. Para as recessões, nota-se uma elevada precisão até dois dias à frente.

Figura 6— Previsão de vazão média (linha vermelha) em comparação com a vazão observada (linha azul) para o horizonte h+2 a h+7 (a) e horizonte h+43 a h+48 (b).



Por fim, a Figura 7 apresenta a segunda maior cheia ocorrida em 2024, e três previsões contínuas espaçadas por um dia de diferença. Pode ser observado que a primeira previsão consegue indicar adequadamente uma elevação das vazões, ainda que não na mesma velocidade ocorrida, enquanto a segunda previsão realizada indica uma pequena elevação da vazão nas próximas 48h. Por fim, a última previsão indica uma pequena elevação nas primeiras 24h, porém, não consegue indicar a última elevação da vazão nas últimas horas do dia 13 de julho.

Figura 7— Vazão observada na UHE G.B. Munhoz entre os dias 11 e 15 de julho (linha azul), em comparação com três previsões contínuas espaçadas por um dia (linhas tracejadas vermelha, verde e roxa).



CONCLUSÕES

No presente trabalho, foi desenvolvido e testado um modelo híbrido de previsão de vazão horária em tempo real, que integra previsões diárias naturais do modelo conceitual SMAP/ONS com técnicas de aprendizado de máquina baseadas em *Gradient Boosting* (LGBM). Em horizontes de curtíssimo prazo ($h \leq 6$), o modelo atingiu MAPE de 5% e NSE de 0,98, evidenciando alta acurácia ao reproduzir rapidamente as variações observadas. Para horizontes intermediários e longos ($h > 6$), observou-se degradação natural do desempenho, chegando a MAPE de 14% e NSE de 0,86 para as últimas seis horas do horizonte de 48 h. Ainda assim, o modelo manteve capacidade satisfatória de antecipar eventos extremos e recessões hidrológicas.

O modelo atende aos requisitos de latência e confiabilidade exigidos pelo Centro de Operação do Sistema Sul, fornecendo previsões horárias a cada hora com horizonte de 48 h. A robustez demonstrada em cenários de cheias e em períodos de recessão reforça seu potencial de incorporação

na rotina de apoio à decisão dos operadores do SIN, sobretudo em bacias de baixa regularização como a do rio Iguaçu.

Como aprimoramentos futuros, a incorporação de dados de precipitação proveniente de estações telemétricas, acabaria com a latência da observação dessa informação, o que poderia trazer uma resposta mais consistente ao modelo. Por conta da flexibilidade e rapidez de estimação, outras variáveis podem ser criadas a partir de distintas transformações levando ao aprimoramento do modelo.

Por fim, a integração de modelos hidrológicos conceituais com técnicas modernas de aprendizado de máquina mostrou-se promissora para enfrentar os desafios da operação em tempo real, oferecendo aos operadores do SIN uma ferramenta capaz de apoiar decisões críticas com elevada precisão e rapidez.

REFERÊNCIAS

- GOVERNO DO PARANÁ. Instituto Água e Terra. Disponível em: <https://www.iat.pr.gov.br/sites/agua-terra/arquivos_restritos/files/documento/2020-07/iguacu.pdf>. Acesso em: 18/06/2025.
- HEIL, C. E., & WALNUT, D. F. (1989). Continuous and discrete wavelet transforms. *SIAM review*, 31(4), 628-666.
- KE, G., MENG, Q., FINLEY, T., WANG, T., CHEN, W., MA, W., ... & LIU, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.
- MACEIRA, M. E. P. *et al.*, "Twenty Years of Application of Stochastic Dual Dynamic Programming in Official and Agent Studies in Brazil-Main Features and Improvements on the NEWAVE Model," *2018 Power Systems Computation Conference (PSCC)*, Dublin, Ireland, 2018, pp. 1-7, doi: 10.23919/PSCC.2018.8442754.
- OPERADOR NACIONAL DO SISTEMA ELÉTRICO (ONS) - (2024). Manual de Metodologia - Modelo SMAP/ONS.
- OPERADOR NACIONAL DO SISTEMA ELÉTRICO (ONS) - (2025a). Procedimentos de Rede. Disponível em: <<https://www.ons.org.br/paginas/sobre-o-ons/procedimentos-de-rede/vigentes>>. Acesso em: 18/06/2025
- OPERADOR NACIONAL DO SISTEMA ELÉTRICO (ONS) - (2025b). Séries Históricas de Vazões Naturais. Disponível em: <<https://sintegre.ons.org.br/sites/9>>. Acesso em: 18/06/2025.
- ROZANTE, J. R.; GUTIERREZ, E. R.; FERNANDES, A. A.; VILA, D. A. (2020) Performance of precipitation products obtained from combinations of satellite and surface observations, *International Journal of Remote Sensing*, 41:19, 7585-7604, DOI: 10.1080/01431161.2020.1763504
- SZCZEPANEK, R. Daily streamflow forecasting in mountainous catchment using XGBoost, LightGBM and CatBoost. *Hydrology*, v. 9, n. 12, p. 226, 2022.
- TREISTMAN, F., MACIEL, A.; BRAGA, R. G. S., BARROS, A. M. DE L., TORRACA, H. B. B., CAMARGO, H. J. (2023) "Análise do desempenho do modelo SMAP/ONS para a previsão de vazões no 1º mês operativo do programa mensal da operação energética utilizando previsões estendidas do modelo ECMWF" in *Anais XXV Simpósio Brasileiro de Recursos Hídricos*, Aracaju, Nov. 2023.
- ZOUNEMAT-KERMANI, M., BATELAAN, O., FADAEI, M., & HINKELMANN, R. (2021). Ensemble machine learning paradigms in hydrology: A review. *Journal of Hydrology*, 598, 126266.