

Inflow forecasting for the Tucuruí hydroelectric using LSTM and CNN neural networks

Raphael Teixeira¹, Ângello Gonçalves¹, Kevin Medeiros¹,
Ayla Santos¹, Gabriel Moreira¹, Cleison Silva¹

¹Universidade Federal do Pará - Campus Universitário de Tucuruí
Núcleo de Desenvolvimento Amazônico em Engenharia
Programa de Pós-Graduação em Computação Aplicada
Faculdade de Engenharia Elétrica
Laboratório de Inteligência Controle e Eletrônica
Tucuruí – PA – Brasil

raphaelbt@ufpa.br, cleison@ufpa.br

Abstract. *This paper investigates the application of LSTM neural networks, CNN and a hybrid CNN-LSTM model for modeling and predicting inflow of the Tucuruí Hydroelectric Plant. Forecasts of inflows to hydroelectric plant reservoirs are extremely important in optimizing energy operation planning. The historical data of the reservoirs of upstream plants were accessed by the National Agency of Water – (ANA), being the target of work to predict daily inflows in a horizon of 1 to 7 days. The results show the viability of the models proposed, with emphasis on the high precision of CNN networks above 96% accuracy.*

Resumo. *Este estudo investiga a aplicação de redes neurais do tipo LSTM, CNN e um modelo híbrido CNN-LSTM para modelar e prever a vazão afluente da Usina Hidrelétrica de Tucuruí (UHE). As previsões das vazões afluentes aos reservatórios das usinas hidrelétricas são de extrema importância na otimização do planejamento de operação energética. Os dados históricos dos reservatórios das usinas à montante foram acessados pela Agência Nacional de Água – ANA, sendo alvo do trabalho para prever as afluências diárias em um horizonte de 1 a 7 dias. Os resultados mostram a viabilidade dos modelos propostos, com ênfase na alta precisão das redes CNN acima de 96% de precisão.*

1. Introdução

A dinâmica da água nas grandes bacias hidrográficas afeta diversos processos naturais, sociais e econômicos, como a geração de energia elétrica. Neste cenário, o Brasil se destaca pela sua matriz energética reconhecidamente renovável [EPE 2024], com forte presença de Usinas Hidrelétricas (UHE) de grande porte, como é o caso da UHE de Tucuruí, no Estado do Pará, situada na área de abrangência da Bacia Hidrográfica Araguaia-Tocantins (BAT), no estuário da Amazônia. A gestão dos reservatórios de água destas usinas é uma tarefa complexa, que demanda a previsão do comportamento de suas variáveis para se buscar uma estratégia ótima de operação sob as perspectivas econômicas e ambientais.

Entre as variáveis fundamentais para explicar o comportamento deste sistema destaca-se a vazão afluente, entendida como a água que chega aos reservatórios pelo escoamento natural dos rios da bacia, cuja dinâmica não-linear e multivariável torna a

modelagem e a consequente previsão desta variável um problema não trivial, que tem desafiado a hidrologia ao longo do tempo [Guilhon et al. 2007].

Neste trabalho busca-se realizar a modelagem da dinâmica da vazão afluente ao lago da UHE-Tucuruí a partir de dados reais de vazão em usinas hidrelétricas a montante. Os modelos avaliados são redes neurais artificiais recorrentes (RNN) do tipo LSTM (Long Short-Term Memory) e as redes convolucionais (CNN - Convolutional neural network), algoritmos de aprendizagem profunda que têm se destacado na modelagem onde os dados são processados em sequência [Zhang et al. 2023], como é o caso das séries temporais de vazão.

Estes algoritmos tem sido explorados em muitos contextos de modelagem hidrológicas, superando resultados atingidos por abordagens convencionais [da Silva et al. 2024]. Em [Kabir et al. 2020] redes CNN são utilizadas para previsão da vazão e enchentes, [Janbain et al. 2023] utiliza LSTM para reconstrução de séries temporais de vazão. O sucesso particular destas técnicas têm estimulado a combinação delas em estruturas de modelos mais sofisticados, como é o caso de [Li et al. 2022a] que apresenta o modelo LSTM-CNN para a predição de vazão.

Este trabalho se insere nesta vertente e busca tanto explorar a capacidade das redes LSTM de modelar relações do tipo sequência para sequência (*seq2seq*), a habilidade das redes convolucionais no reconhecimento de padrões temporais expressos nos dados, e a combinação dos dois, replicando a estratégia LSTM-CNN para a modelagem da vazão afluente ao lago da UHE de Tucuruí.

O diferencial da proposta aqui explorada é a concepção da topologia destes modelos a partir da estrutura expressa na geografia da bacia hidrográfica e no conjunto de usinas hidrelétricas do rio Tocantins. Assim são utilizados dados de vazão das usinas de Lajeado e Estreito, a montante de Tucuruí, sendo estas tomadas como variáveis de entrada do modelo, que deve explicar a vazão afluente de Tucuruí, variável de saída. A estrutura é então concebida como um sistema dinâmico e o processo de aprendizagem dos modelos faz uso de conceitos de identificação de sistemas, em particular ao considerar dados auto-regressivos da própria vazão de Tucuruí na entrada do modelo.

A base de dados, com amostragem diária, foi acessada do portal do Operador Nacional do Sistema Elétrico do Brasil (ONS) para uma janela de 01 de janeiro de 1999 a 01 de junho de 2023, perfazendo 22 anos de dados, com 8918 amostras para cada variável. Os modelos treinados são avaliados por métricas de erro entre suas previsões e dados de validação, típicas em problemas de regressão, tais como o erro médio quadrático (MSE) e do coeficiente de determinação R^2 .

O restante do artigo é sequenciado com a seção 2, que apresenta os fundamentos de modelagem de vazão afluente com a perspectiva das técnicas propostas, a seção 3 que aborda a metodologia utilizada no estudo, a seção 4 que apresenta os resultados e discussões e a seção 5 com a conclusão do trabalho.

2. Previsão de afluentes

A previsão da vazão afluente aos reservatórios hidrelétricos é uma demanda crítica para a otimização e gestão operacional. A antecipação da dinâmica da água disponível para a geração permite o controle de níveis, volumes e vazões adequados à gestão dos reser-

vatórios das usinas.

A BAT possui importantes aproveitamentos hidrelétricos no Rio Tocantins, como as usinas de Lajeado, Estreito e Tucuruí, a mais relevante. A vazão nestes sistemas hidrológicos é uma variável monitorada diariamente, de modo que os dados históricos coletados ao longo dos últimos anos de operação podem ser acessados e utilizados para a composição de modelos matemáticos que representam a dinâmica da vazão afluente, objeto de interesse deste trabalho.

Esta perspectiva se insere no contexto de identificação de sistemas onde representações matemáticas explícitas são utilizadas para modelar comportamentos dinâmicos em contextos onde métodos analíticos baseados em princípios físicos fundamentais do sistema encontram dificuldade, como é o caso dos sistemas hidrológicos [Billings 2013]. Estes sistemas apresentam uma dinâmica multivariável, não linear, variável no tempo, fortemente influenciada por fatores climáticos, geográficos e mesmo por processos antrópicos. Esta complexidade constitui um desafio a proposição de modelos explícitos de identificação, o que justifica o interesse deste trabalho por representações implícitas de aprendizagem profunda baseadas em RNAs, cuja estratégia de treinamento supervisionado associa-se fortemente à perspectiva de identificação de sistemas [Ljung et al. 2020].

A Figura 1 ilustra a área de abrangência da BAT, destacando os principais rios e UHEs de interesse deste. Neste caso, seguindo o fluxo do Rio Tocantins, a vazão de Lajeado $u_L(t)$ alimenta a vazão de Estreito $u_E(t)$ que gera a vazão afluente ao lago de Tucuruí $y_T(t)$. Nesta perspectiva, propõe-se tomar $u_L(t)$ e $u_E(t)$ como variáveis de entrada de um modelo para explicar $y_T(t)$, a saída do sistema.

A vazão afluente de Tucuruí predita por este modelo $\hat{y}_T(t)$ pode ser então concebida como uma variável no tempo discreto t , que é função não-linear das vazões de Estreito e Lajeado. Assim sendo, esta estrutura pode ser abstraída genericamente pela operação expressa no diagrama da Figura 2, de onde se define $\mathcal{H}\{\cdot\}$ como um operador que modela as relações não-lineares entre os atrasos temporais das variáveis. Neste caso,

$$\mathbf{X}_i = \begin{bmatrix} y_T(t) & y_T(t-1) & y_T(t-2) & \dots & y_T(t-n_i) \\ u_E(t) & u_E(t-1) & u_E(t-2) & \dots & u_E(t-n_i) \\ u_L(t) & u_L(t-1) & u_L(t-2) & \dots & u_L(t-n_i) \end{bmatrix} \quad (1)$$

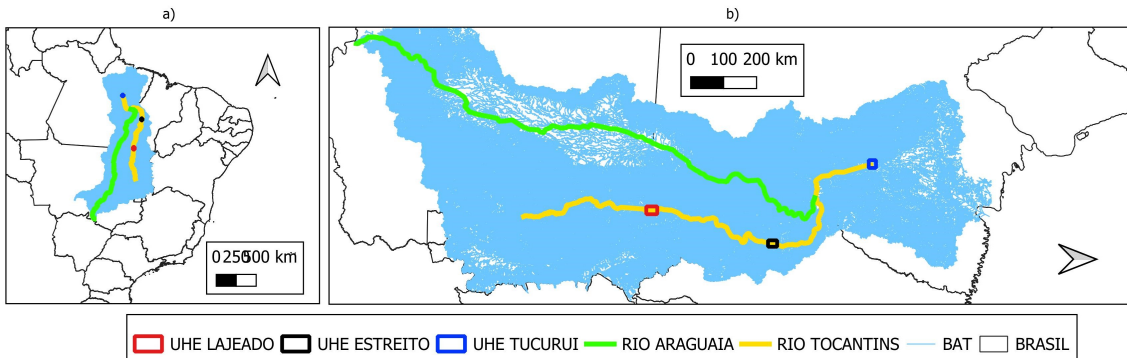


Figura 1. (a) Área da Bacia Araguaia-Tocantins, (b) UHE-Lajeado, UHE-Estreito, UHE-Tucuruí.

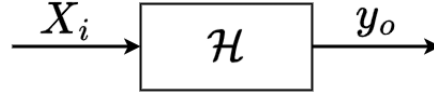


Figura 2. Diagrama da perspectiva dinâmica do modelo.

é uma matriz com um conjunto de variáveis de entrada, cada uma com n_i amostras atrasadas. Estes dados são processados por $\mathcal{H}(\cdot)$ para a geração de

$$\mathbf{y}_o = [y_T(t+1) \quad y_T(t+2) \quad \dots \quad y_T(t+n_o)] \quad (2)$$

que é uma sequência de dados com n_o amostras de predição de saída. Nenhuma estrutura a priori é concebida quanto a estas relações e a natureza do operador. O modelo assim proposto embute na entrada atrasos da própria variável alvo, o que caracteriza uma componente auto-regressiva.

Neste trabalho o interesse é em realizar $\mathcal{H}(\cdot)$ por RNAs que modelem relações dinâmicas não lineares do tipo sequência para sequência. Isto é particularmente importante em situações onde o modelo é utilizado para gerar cenários de predições com múltiplos passos a frente, como ocorre em contextos de modelagem de sistemas hidrológicos.

2.1. Modelagem Dinâmica com Redes Neurais

As redes neurais artificiais (RNAs) são reconhecidas pela sua capacidade de generalização e por aprenderem padrões complexos a partir de dados. No contexto de aprendizagem dinâmica, duas topologias tem se destacado na literatura: as redes recorrentes LSTM e as redes convolucionais CNN [Zhang et al. 2023].

2.1.1. Redes LSTM

Uma rede LSTM é um tipo de RNN capaz de aprender dependências de longo prazo entre um sinal de entrada e suas predições futuras, no contexto de série temporal [Goodfellow et al. 2016]. O módulo de uma RNN padrão tem uma única camada *tanh* que regula como a informação combinada da entrada x_t e saída anterior h_{t-1} produz a saída atual h_t . LSTMs têm uma estrutura sequencial similar a esta RNN, mas o módulo LSTM tem quatro *camadas de rede neural (feedforward)* controlando a informação que entra, sai e define o estado interno do módulo, conforme destacado na Figura 3.

A componente central de uma LSTM é o *estado da célula*. Uma célula LSTM recebe o estado C_{t-1} da célula em tempo anterior como entrada e entrega o estado C_t para a célula em tempo posterior como saída. É possível que a informação de estado que uma célula recebeu passe inalterada até a saída da célula, ou seja, $C_t = C_{t-1}$. Mas a LSTM é capaz alterar o estado da célula usando estruturas chamadas *portas*. As portas *regulam* toda informação que é adicionada ou removida do estado da célula. Uma porta é composta por uma camada de rede neural tipo *sigmoid* e um operador de multiplicação.

Uma camada *sigmoid* produz um número entre zero e um, regulando quanto cada componente de informação pode passar adiante. A LSTM aprende qual informação deve

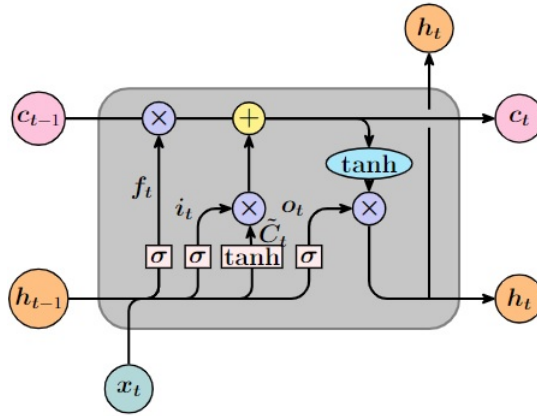


Figura 3. Célula de uma LSTM.

sair do estado da célula através da sua porta de esquecimento. Esta, observa h_{t-1} (saída anterior) e x_t (entrada atual) e gera um número entre 0 e 1 para “ponderar” cada número no estado da célula C_t , onde 1 sinaliza “lembrar” e 0 sinaliza “esquecer”.

A saída da porta de esquecimento f é:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f). \quad (3)$$

Esta estrutura é típica de uma camada *feedforward* com parâmetros de treinamento (W_f e b_f), dados de entrada (h_{t-1} e x_t) e função de ativação (σ). A próxima etapa é decidir qual nova informação guardar no estado da célula. Uma camada *sigmoid*, a “porta de entrada” i , decide quais valores serão atualizados e uma camada *tanh* cria um vetor de novos valores candidatos \tilde{C}_t que pode ser adicionado ao estado. Ou seja, f_t regula quanto do estado anterior C_{t-1} deve ser “esquecido” e i_t regula quanto de novo estado \tilde{C}_t deve ser “lembrado” pela célula atual. A saída da célula h_t é um valor filtrado do estado atualizado C_t . Uma camada *sigmoid* o_t decide quais partes do estado da célula irá para a saída e um filtro *tanh* é aplicado ao estado C_t .

A regulação da saída é:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (4)$$

de onde se pode determinar da célula como:

$$h_t = o_t \times \tanh C_t. \quad (5)$$

Esta composição de células LSTM, que mantém um estado atualizável dependendo do que é relevante para computar a saída atual dados a entrada atual e a saída da célula em tempo anterior, forma uma estrutura apta a aprender as relações de dependência entre os itens da série de dados.

2.1.2. Redes CNN

A convolução é um processo de deslocamento de filtros ao longo do conjunto de dados de entrada cujo objetivo é abstrair características fundamentais do trecho em que se

desloca o filtro. Este processo permite a extração de características locais particulares [Zhang et al. 2023]. É comum que as CNNs estruturam-se em camadas convolucionais, de *pooling* e camadas de *flattening*.

As camadas de *pooling* reduzem as dimensões da matriz de características mantendo as informações essenciais. A camada de *flattening*, responsável por mudança os tensores multidimensionais em vetores unidimensionais, adequa as informações para as camadas densas, as quais processam as informações que traduzem as características extraídas para tarefas específicas. Algumas dessas tarefas são, por exemplo, a classificação, detecção de objetos em imagens e regressão com destaque para a modelagem dinâmica de séries temporais [Ljung et al. 2020]. A Figura 4 ilustra uma rede CNN típica e seus operadores.

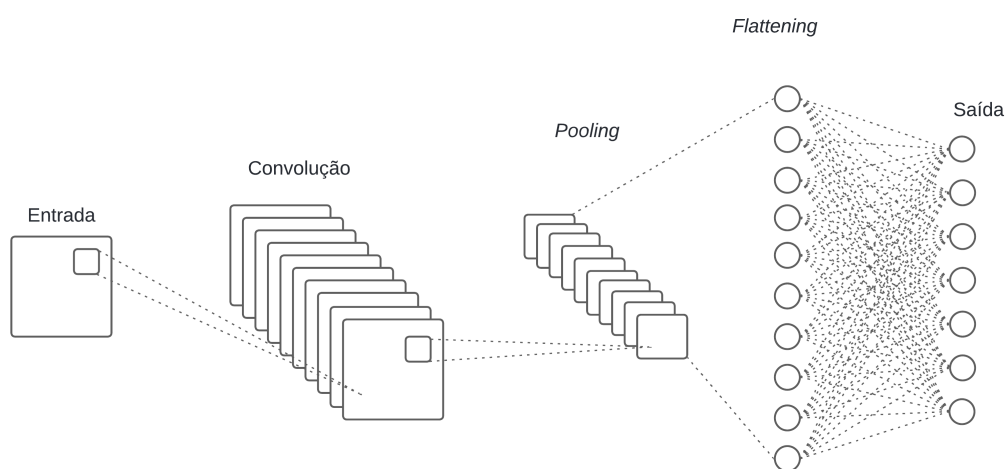


Figura 4. Rede CNN.

2.1.3. Redes CNN-LSTM

Além de LSTM e CNN individualmente, este trabalho busca avaliar a integração destas duas RNAs para modelar a predição da vazão afluente. Nessa arquitetura híbrida, a CNN tem a função de extrair automaticamente recursos fundamentais do conjunto de dados fornecido e a LSTM busca modelar a relação temporal que melhor modele a série temporal.

Esta estratégia tem sido explorada em trabalhos recentes [Li et al. 2020], [Li et al. 2022b], os quais demonstram que a CNN é responsável por capturar variações espaciais nos dados de entrada, enquanto a LSTM modela as dependências temporais complexas. Esta combinação tem o potencial de extrair vantagens das duas arquiteturas e se constituem em um objeto importante de investigação para o contexto de modelagem dinâmica, como a previsão de aflúências, onde as variáveis ambientais e hidrológicas apresentam padrões espaciais e temporais com relações não-triviais.

3. Metodologia

3.1. Dados e pré-processamento

O trabalho foi conduzido a partir do acesso aos dados no repositório do ONS para as três variáveis de interesse. As séries históricas abrangem um período de 22 anos, de 01 de

janeiro de 1999 a 01 de junho de 2023, um total de 8918 amostras para cada variável. O pré-processamento incluiu a normalização das amostras que foram divididas pela máxima vazão da UHE Tucuruí, visando eliminar discrepâncias numéricas que pudessem incorrer em problemas de convergência para as RNAs. Os dados normalizados entre 0 e 1 são apresentados na Figura 5 que evidencia a divisão de 75% das amostras para treinamento e 25% para a validação dos modelos, procedimento típico em contextos de aprendizagem supervisionada.

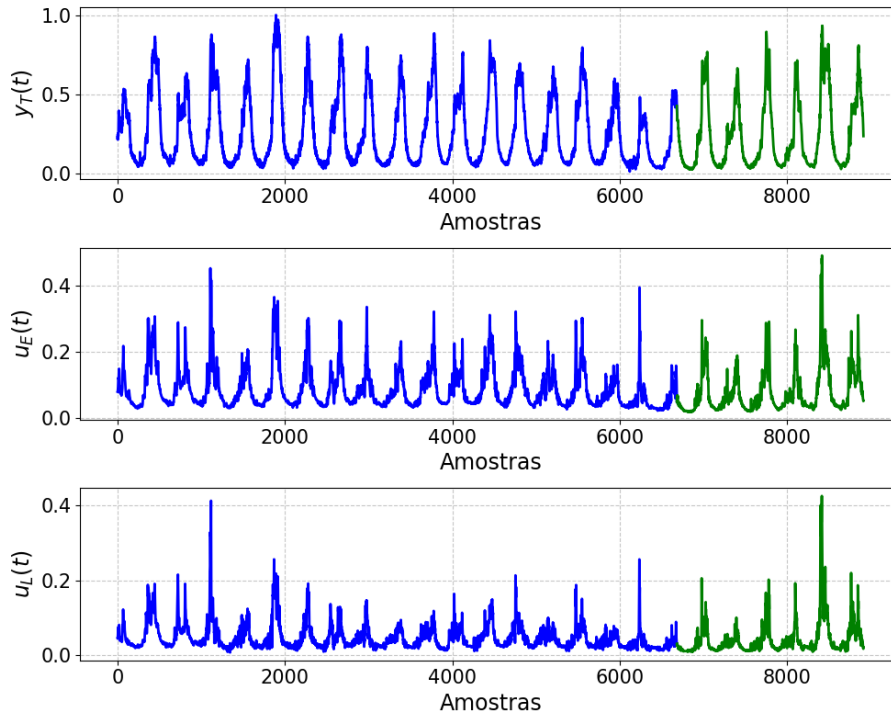


Figura 5. Identificação (75%) x Validação(25%) - 3 UHE's

Considerando os modelos propostos, os dados foram ainda pré-processados para gerar padrões de entrada e saída a serem aprendidos pelas redes, onde foram geradas matrizes com as três variáveis de entrada. Foram definidos valores para os atrasos das variáveis e para o tamanho da sequência de saída, respectivamente como $n_i = 30$ e $n_o = 7$. Assim, os modelos devem processar dados de 30 dias de atraso das variáveis de entrada visando prever os próximos 7 dias da vazão afluente em Tucuruí.

3.2. Arquitetura dos Modelos

São propostas e avaliadas três arquiteturas de modelos, com as RNAs: uma LSTM (modelo 1), uma CNN (modelo 2) e uma estrutura híbrida CNN-LSTM (modelo 3). A Figura 6 detalha a estrutura de cada modelo, sua sequência de camadas e operadores.

Para a arquitetura híbrida, camadas unidimensionais convolucionais foram adicionadas, em sequência, uma camada de *MaxPooling*. Afim de processar os dados no formato proposto pela LSTM, uma camada de *Flatten* foi adicionada. Em aprendizado profundo, temos um fenômeno comum que é o sobreajuste (*overfitting*), e há diversas maneiras para mitigar este problema, nesse caso foi utilizado o método de *Dropout* que faz o papel de retirar temporariamente algumas células aleatórias da rede neural. Logo em

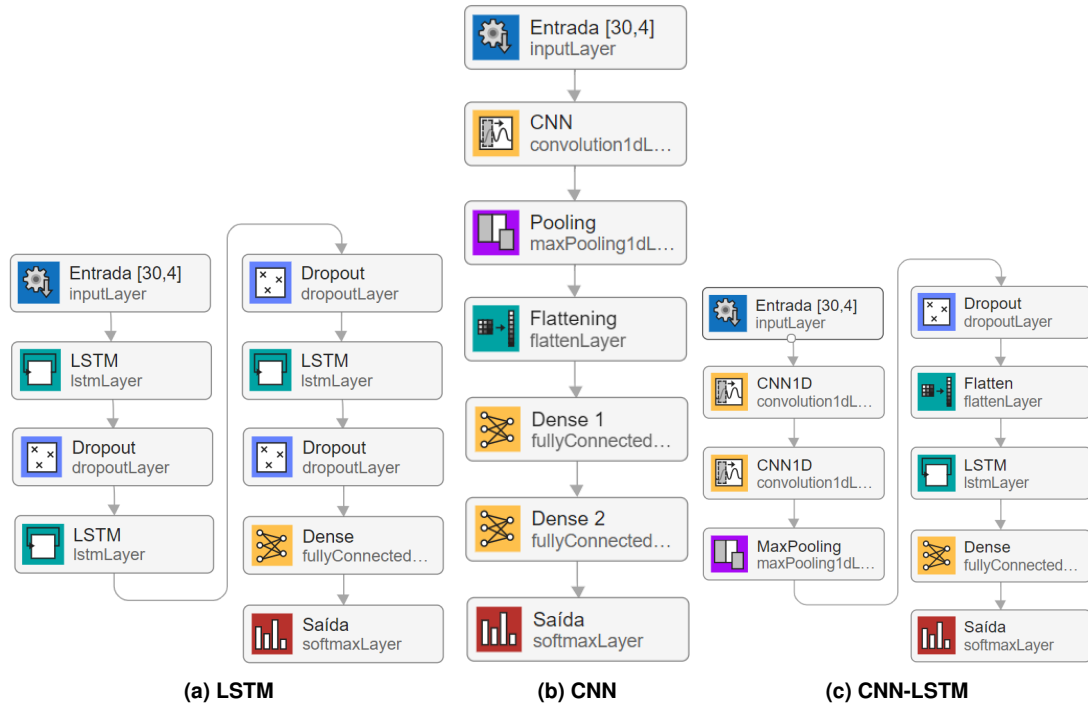


Figura 6. Diagramas das arquiteturas dos modelos

seguida, temos uma camada de LSTM contendo 100 unidades de neurônios e finalmente conectada a uma camada *Dense*.

Cada uma dessas arquiteturas foram projetadas para capturar diferentes aspectos dos dados de séries temporais e fornecer previsões precisas da afluência para a usina hidrelétrica de Tucuruí.

Os modelos foram implementados em linguagem Python com métodos das bibliotecas Keras, Sklearn e TensorFlow. A implementação e execução dos experimentos foram feitas no Google Colab, aproveitando os recursos de computação em nuvem para acelerar o processo de treinamento.

3.3. Métricas de avaliação

A função objetivo utilizada ao avaliar o desempenho dos modelos deve preferencialmente incluir múltiplas métricas que possam medir diferentes aspectos do desempenho dos modelos. Sendo assim, no presente trabalho foram utilizadas, o erro médio quadrático (MSE) definido por:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (6)$$

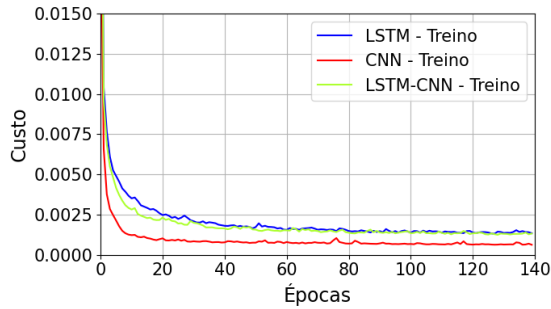
e o coeficiente de determinação (R^2), dado por:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}. \quad (7)$$

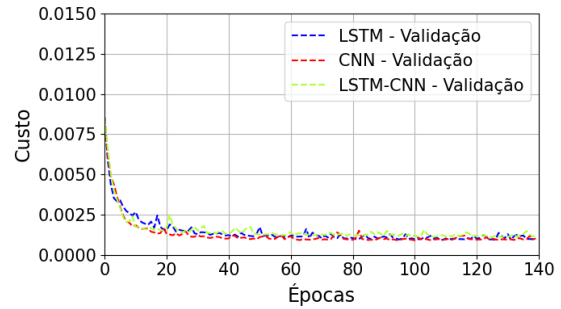
Onde \bar{y} é a média dos valores preditos. Tipicamente $R^2 \in [0, 1]$, onde valores próximos de 1 indicam melhores previsões.

4. Resultados

Com o treinamento das três estruturas propostas pode-se avaliar a partir das curvas da função custo, tanto para dados de treinamento quanto de validação, conforme apresentado nas Figuras 7a e 7b. Estas curvas exibem uma convergência uniforme do custo ao



(a) Custo de Treinamento: LSTM X CNN X LSTM-CNN

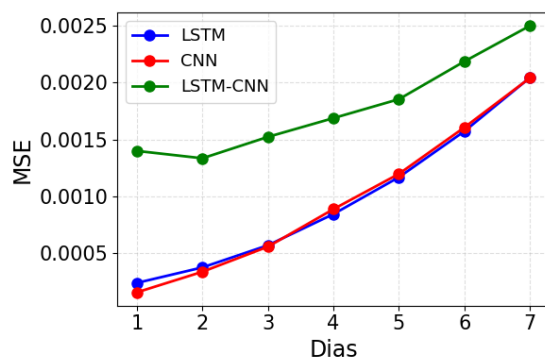


(b) Custo de Validação: LSTM X CNN X LSTM-CNN

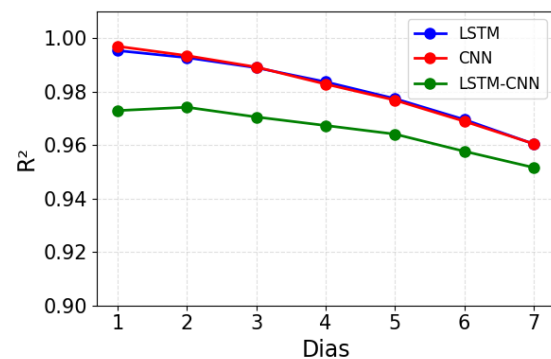
longo das épocas, com valores finais e bastante similares tanto entre os modelos quanto entre os dados de treinamento e de validação, com a rede CNN exibindo o melhor desempenho. As curvas não sugerem anomalias como super-ajuste ou tendências, de modo que se considera que o número de épocas adotadas adequado para a proposta.

Os modelos treinados foram então validados com o conjunto de dados de validação. A avaliação fez uso inicialmente das medições das métricas MSE e R^2 , cujos valores e tendência são apresentados para os 7 dias previstos por todos os modelos nas Figuras 8a e 8b. Os valores de R^2 superando 0,95 para todas as amostras sugerem que a estratégia proposta é apropriada para a modelagem da vazão afluente. Nota-se que os modelos exibem desempenho notáveis para o primeiro dia da sequência predita, com R^2 superando 0,99 e decaindo à medida que o horizonte de predição aumenta, o que é justificável pelas realimentações dos erros de predição entre as amostras, o que é evidenciado pelo aumento uniforme do MSE para cada modelo.

Estes dados são apresentados com valores numéricos na Tabela 1. Estas curvas e valores evidenciam que os modelos com redes LSTM e CNN exibem desempenho muito similar, com uma vantagem sutil para as redes CNN. As métricas apontam para uma discrepância de desempenho destas para o modelo híbrido, que exibiu resultados inferiores na estratégia adotada.



(a) Métrica MSE



(b) Métrica R^2

Modelos	Métricas	Dias						
		1	2	3	4	5	6	7
LSTM	MSE	0.0002	0.0004	0.0006	0.0008	0.0012	0.0016	0.0020
	R^2	0.9954	0.9927	0.9890	0.9837	0.9774	0.9696	0.9604
CNN	MSE	0.0002	0.0003	0.0006	0.0009	0.0012	0.0016	0.0020
	R^2	0.9970	0.9935	0.9892	0.9828	0.9768	0.9689	0.9605
CNN-LSTM	MSE	0.0014	0.0013	0.0015	0.0017	0.0019	0.0022	0.0025
	R^2	0.9729	0.9742	0.9705	0.9673	0.9641	0.9577	0.9516

Tabela 1. Performance dos modelos conforme as métricas.

Uma hipótese para este desempenho inferior é a pequena quantidade de dados disponíveis para o treinamento de redes de aprendizagem profunda com estrutura mais complexas como as do tipo CNN-LSTM.

A Figura 9 apresenta curvas de simulações para uma faixa dos dados de validação para diferentes dias, a saber: (i) o primeiro dia na Figura 9a que possui o melhor ajuste para os três modelos, com destaque para o excelente ajuste para as redes CNN em toda faixa de dados e a dificuldade dos dois outros modelos particularmente nos picos das curvas de vazão; (ii) o quarto dia na Figura 9b que ressalta os mesmos detalhes e (iii) o sétimo dia na Figura 9c que aponta para uma dificuldade de ajuste dos três modelos que passam a expressar uma variabilidades e atrasos crescentes em relação aos valores reais.

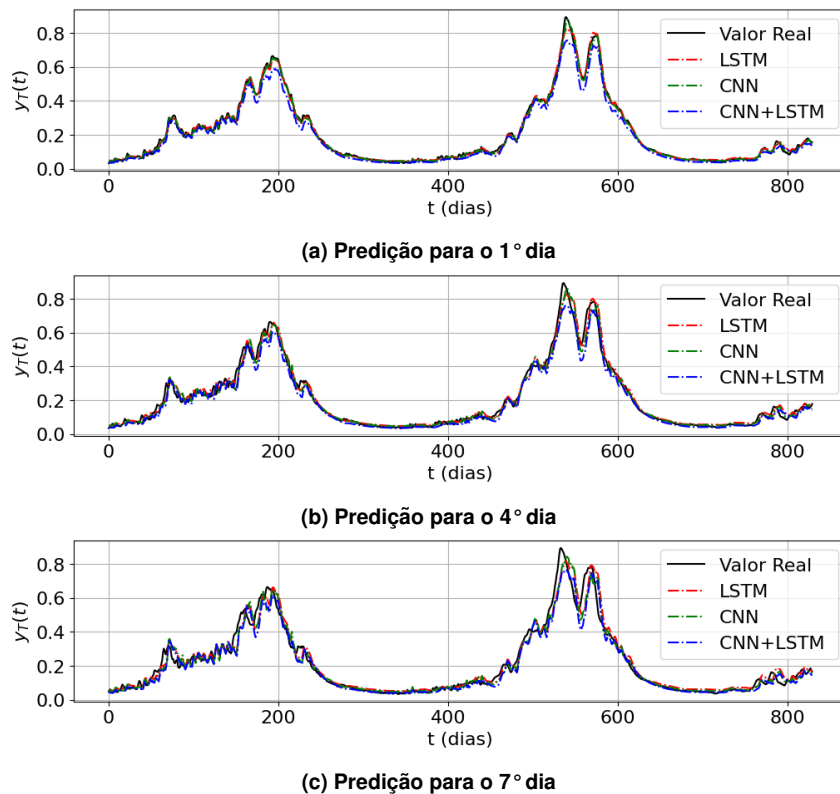


Figura 9. Predição de vazão afluente $y(t)$ para os três modelos, confrontadas com os dados reais.

5. Conclusão

O trabalho apresenta a aplicação de redes LSTM, CNN e uma híbrida CNN-LSTM na previsão da vazão afluente da UHE Tucuruí, na BAT. A partir de dados reais acessados pelo ONS, os modelos foram treinados e validados, e seu desempenho foi aferido utilizando métricas de erro como MSE e R^2 .

Os resultados mostraram que os três modelos propostos foram bem-sucedidos na previsão de vazão afluente. As redes CNN se destacaram em termos de melhores desempenhos, uma vez que o valor de R^2 supera 0,99 no primeiro dia, e acima de 0,96 para os outros dias. O desempenho do modelo híbrido, inferior às duas outras topologias, pode se dar em função da quantidade modesta de dados para treinamento. Em trabalhos onde esta estratégia é utilizada é comum a utilização de dezenas ou centenas de milhares de amostras.

A previsão de afluências com 7 dias de antecipação com desempenho acima de 90% de ajuste podem ser considerados apropriados para cenários de planejamento de operação da usina.

Referências

- Billings, S. (2013). *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*. Wiley.
- da Silva, E. C., Finardi, E. C., and Stefenon, S. F. (2024). Enhancing hydroelectric inflow prediction in the Brazilian power system: A comparative analysis of machine learning models and hyperparameter optimization for decision support. *Electric Power Systems Research*, 230:110275.
- EPE, M. (2024). *Balanco Energético Nacional*. Empresa Brasileira de Pesquisa Energética, Ministério de Minas e Energia do Brasil.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press.
- Guilhon, L. G. F., Rocha, V. F., and Moreira, J. C. (2007). Comparação de métodos de previsão de vazões naturais afluentes a aproveitamentos hidroelétricos. *Revista Brasileira de Recursos Hídricos*, 12(3):13–20.
- Janbain, I., Deloffre, J., Jardani, A., Vu, M. T., and Massei, N. (2023). Use of long short-term memory network (Lstm) in the reconstruction of missing water level data in the river seine. *Hydrological Sciences Journal*, 68(10):1372–1390.
- Kabir, S., Patidar, S., Xia, X., Liang, Q., Neal, J., and Pender, G. (2020). A deep convolutional neural network model for rapid prediction of fluvial flood inundation. *Journal of Hydrology*, 590:125481.
- Li, P., Zhang, J., and Krebs, P. (2022a). Prediction of flow based on a cnn-lstm combined deep learning approach. *Water*, 14(6):993.
- Li, T., Hua, M., and Wu, X. (2020). A hybrid cnn-lstm model for forecasting particulate matter (pm_{2.5}). *IEEE Access*, PP.
- Li, X., Xu, W., Ren, M., Jiang, Y., and Fu, G. (2022b). Hybrid cnn-lstm models for river flow prediction. *Water Supply*, 22(5):4902–4919.

- Ljung, L., Andersson, C., Tiels, K., and Schön, T. B. (2020). Deep learning and system identification. *IFAC-PapersOnLine*, 53(2):1175–1181.
- Zhang, A., Lipton, Z. C., Li, M., and Smola, A. J. (2023). *Dive into deep learning*. Cambridge University Press.