



UNIVERSIDADE FEDERAL DO PARÁ
NÚCLEO DE DESENVOLVIMENTO AMAZÔNICO EM ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA

**Modelagem da dinâmica hidrológica para previsão do nível de jusante da UHE Tucuruí
por redes neurais artificiais**

Wanderley Pereira dos Santos

Tucuruí-PA

2026



UNIVERSIDADE FEDERAL DO PARÁ
NÚCLEO DE DESENVOLVIMENTO AMAZÔNICO EM ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA

Wanderley Pereira dos Santos

**Modelagem da dinâmica hidrológica para previsão do nível de jusante da UHE Tucuruí
por redes neurais artificiais**

Dissertação de mestrado apresentada ao colegiado do Programa de Pós-Graduação em Computação Aplicada, do Núcleo de Desenvolvimento Amazônico em Engenharia da Universidade Federal do Pará, como requisito necessário para obtenção do título de Mestre em Computação Aplicada.

Orientador: Prof. Dr. Raphael Barros Teixeira

Tucuruí-PA

2026

UNIVERSIDADE FEDERAL DO PARÁ
NÚCLEO DE DESENVOLVIMENTO AMAZÔNICO EM ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA

**Modelagem da dinâmica hidrológica para previsão do nível de jusante da UHE Tucuruí
por redes neurais artificiais**

AUTOR: Wanderley Pereira dos Santos

DISSERTAÇÃO DE MESTRADO SUBMETIDO À BANCA EXAMINADORA DO COLEGIADO DA PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA, SENDO JULGADO **APROVADA** PARA OBTENÇÃO DO TÍTULO DE MESTRE EM COMPUTAÇÃO APLICADA.

BANCA EXAMINADORA:

Prof. Dr. Raphael Barros Teixeira
Orientador: UFPA-NDAE-PPCA

Prof^a. Dr^a. Evanice Pinheiro Gomes
Avaliadora externa: UFPA - PEBGA

Prof. Dr. Fabrício de Souza Farias
Avaliadora interno: UFPA - PPCA

Resumo

Este trabalho investiga a aplicação de redes neurais artificiais tipo Temporal Convolutional Network (TCN), recorrentes como a Long Short-Term Memory (LSTM) e a *Gated Recurrent Unit* (GRU) na previsão do nível de jusante da UHE de Tucuruí, cuja natureza dinâmica, não linear e multivariável demanda por formalismos de modelagem baseados em aprendizagem profunda. A necessidade do estudo está atrelada ao fato da cidade de Tucuruí estar a jusante da Usina e possuir áreas habitadas próximas às margens do Rio Tocantins e este cenário exige controle e acompanhamento dos níveis na cidade. Para que isso aconteça, é necessária a obtenção de previsões assertivas do nível de jusante em tempo hábil. Desse modo, as ações de intervenções, quando necessárias, podem ser realizadas pelos órgãos competentes no tempo certo. Com esse intuito, neste trabalho, vários modelos foram desenvolvidos para serem treinados por redes neurais artificiais. As redes treinadas buscam modelar as complexas relações entre as vazões afluente e defluente e a variável-alvo, o nível de jusante. São utilizadas séries com amostragem horária de dados reais do período de 2010 a 2025, totalizando 140.279 amostras. Os modelos são treinados para predição de até 10 horas à frente, janela considerada apropriada para o cenário real desejado. O pré-processamento dos dados desempenhou papel fundamental na qualidade das previsões, sendo aplicada a média móvel, seguida de normalização por valor máximo, estratégia que contribuiu para a redução de ruídos de alta frequência e para a estabilização do processo de treinamento. Posteriormente, a desnormalização permitiu a análise dos resultados em unidades físicas reais, permitindo (ou proporcionando) interpretações hidrológicas consistentes. A avaliação quantitativa foi conduzida por meio das métricas Erro Médio Quadrático (MSE), Erro Quadrático Médio Normalizado (FitNRMSE) e Erro Quadrático do Ajuste (FitR²), que possibilitaram uma análise complementar entre precisão, erro absoluto e capacidade explicativa da variância do sinal real. Os resultados apresentam valores elevados de FitR² (superiores a 99%) para todas as redes, evidenciando que os modelos foram capazes de esclarecer a maior parte da variabilidade do nível de jusante. Além disso, os valores de FitNRMSE permaneceram também elevados, demonstrando que os erros de previsão são pequenos quando comparados à intercorrência natural do sistema. Após análise comparativa entre os modelos, observou-se que a Gated Recurrent Unit (GRU) apresentou o melhor desempenho global, menor degradação das previsões e leve superioridade nas métricas, sobretudo no modelo \mathcal{H}_3 . A LSTM apresentou desempenho próximo ao da GRU, porém com maior sensibilidade em regiões de maior variabilidade do sinal, já a TCN demonstrou bom desempenho médio, porém com tendência à suavização das predições em transições mais abruptas. Assim, conclui-se que os modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 , conjuntamente com as arquiteturas LSTM, GRU e TCN propostas, são capazes de

representar adequadamente a dinâmica do nível de jusante, com destaque para \mathcal{H}_3 e a GRU, que juntas mostraram desempenho superior às demais combinações. Nas figuras 42 e 43, destacam-se os resultados obtidos na análise de previsão, denominada de "alertas reais", cuja simulação mostra dois eventos em sequência registrados nos horários das 11:00 e 12:00 horas do dia 10/03/2025, com janelas deslizantes para as 10 amostras futuras, incluindo os valores reais comprovados para validar o desempenho das previsões. Tais resultados constituem uma base sólida para aplicações futuras em previsão operacional, apoio à tomada de decisão e integração com estratégias de controle preditivo, como o Model Predictive Control (MPC).

Palavras-chaves: Modelagem hidrológica; Vazão afluente; Vazão defluente; Nível de jusante; Identificação de sistemas; Métricas estatísticas e; Redes neurais LSTM, GRU e TCN.

Abstract

This work investigates the application of LSTM, GRU, and TCN artificial neural networks in forecasting the downstream level of the Tucuruí Hydroelectric Plant, whose dynamic, nonlinear, and multivariable nature requires modeling formalisms based on deep learning. The need for the study is linked to the fact that the city of Tucuruí is located downstream from the plant and has inhabited areas near the banks of the Tocantins River, and this scenario requires control and monitoring of water levels in the city. For this to happen, there is a need for accurate downstream level forecasts in a timely manner so that intervention measures, when necessary, can be carried out by the competent authorities at the right time. To this end, in this work, several models were developed to be trained by artificial neural networks. The trained networks seek to model the complex relationships between inflow and outflow and the target variable, the downstream level. Hourly sampling series of real data from the period 2010 to 2025 are used, totaling 140,279 samples. The models are trained to predict up to 10 hours ahead, a window considered appropriate for the desired real scenario. Data preprocessing played a key role in the quality of the forecasts, with the moving average being applied, followed by normalization by maximum value, a strategy that contributed to the reduction of high-frequency noise and the stabilization of the training process. Subsequently, denormalization allowed the analysis of the results in real physical units, enabling (or providing) consistent hydrological interpretations. The quantitative evaluation was conducted using the MSE, FitNRMSE, and FitR² metrics, which enabled a complementary analysis of precision, absolute error, and explanatory power of the real signal variance. The results show high FitR² values (above 99%) for all networks, showing that the models were able to explain most of the variability in the downstream level. In addition, FitNRMSE values also remained high, demonstrating that forecast errors are small when compared to the natural occurrence of the system. After comparative analysis between the models, it was observed that GRU presented the best overall performance, less degradation of forecasts, and slight superiority in metrics, especially in the \mathcal{H}_3 model. LSTM performed similarly to GRU, but with greater sensitivity in regions of greater signal variability, while TCN demonstrated good average performance, but with a tendency to smooth predictions in more abrupt transitions. Thus, it is concluded that the \mathcal{H}_1 , \mathcal{H}_2 , and \mathcal{H}_3 models, together with the proposed LSTM, GRU, and TCN architectures, are capable of adequately representing the dynamics of the downstream level, with emphasis on \mathcal{H}_3 and GRU, which together showed superior performance to the other combinations. Figures 42 and 43 highlight the results obtained in the forecast analysis, called “real alerts,” whose simulation shows two events in sequence

recorded at 11:00 a.m. and 12:00 pm on March 10, 2025, with sliding windows for the 10 future samples, including the actual values proven to validate the performance of the forecasts. These results constitute a solid basis for future applications in operational forecasting, decision support, and integration with predictive control strategies, such as MPC.

Keywords: Hydrological modeling; Influent flow; Effluent flow; Downstream level; System identification; Statistical metrics; LSTM, GRU, and TCN neural networks.

LISTA DE FIGURAS

Figura 1 – Histórico de vazão defluente - UHE Tucuruí. Fonte: Elaborada pelo autor.	2
Figura 2 – Usinas em cascata no Rio Tocantins.	7
Figura 3 – Bacia Tocantins-Araguaia.	8
Figura 4 – Diagrama de blocos.	11
Figura 5 – Malha de controle.	11
Figura 6 – UHE Tucuruí.	13
Figura 7 – Curva Referencial de Deplecionamento da UHE Tucuruí 2023.	15
Figura 8 – Estruturas simplificadas de neurônios biológico e artificial.	18
Figura 9 – Rede Neurais em Multicamadas - MLP.	19
Figura 10 – Rede Neural Convolutacional - RNC.	20
Figura 11 – Estrutura Interna de Rede Neural TCN.	21
Figura 12 – célula Recorrente a cada instante de tempo.	22
Figura 13 – Rede Recorrente a cada instante de tempo - RNN.	23
Figura 14 – Diagrama da Estrutura interna de uma Rede LSTM.	24
Figura 15 – Diagrama da estrutura interna de uma célula da rede LSTM.	24
Figura 16 – Estrutura interna de uma célula GRU.	26
Figura 17 – Variáveis hidrológicas de interesse para treino, validação e teste dos modelos.	28
Figura 18 – Log função custo - Modelos LSTM \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3	33
Figura 19 – Log da função custo dos Modelos GRU \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3	34
Figura 20 – Log da função custo dos Modelos TCN \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3	34
Figura 21 – FitNRMSE para os Modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 - Rede LSTM.	35
Figura 22 – FitNRMSE para os Modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 - Rede GRU.	36
Figura 23 – FitNRMSE para os Modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 - Rede TCN.	36
Figura 24 – FitR ² da rede LSTM para os Modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3	37
Figura 25 – FitR ² da rede GRU para os Modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3	38
Figura 26 – FitR ² da rede TCN para os Modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3	38
Figura 27 – Gráfico detalhado LSTM 1 ^a , 5 ^a e 10 ^a amostras de \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3	39
Figura 28 – Gráfico detalhado GRU 1 ^a , 5 ^a e 10 ^a amostras de \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3	40
Figura 29 – Gráfico detalhado TCN 1 ^a , 5 ^a e 10 ^a amostras de \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3	40
Figura 30 – Níveis Previstos Desnormalizados dos modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 - LSTM	41
Figura 31 – Níveis Previstos Desnormalizados dos modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 - GRU	41
Figura 32 – Níveis Previstos Desnormalizados dos modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 - TCN	42
Figura 33 – Níveis Previstos, Normalizados e Testados por LSTM - Modelo \mathcal{H}_3	43

Figura 34 – Níveis Previstos, Normalizados e Testados por GRU - Modelo \mathcal{H}_3	43
Figura 35 – Níveis Previstos, Normalizados e Testados por TCN - Modelo \mathcal{H}_3	44
Figura 36 – Valor Real x Previsto Testados por LSTM, GRU e TCN - Modelo \mathcal{H}_3	44
Figura 37 – Análise dos Resíduos dos Valores Previsto e Testados por LSTM, GRU e TCN - Modelo \mathcal{H}_3	45
Figura 38 – Histograma dos Resíduos dos valores Previstos Testados por LSTM, GRU e TCN - Modelo \mathcal{H}_3	45
Figura 39 – Níveis de Jusante Desnormalizados dos Testes Preditos pelo Modelo \mathcal{H}_3 - LSTM.	46
Figura 40 – Níveis de Jusante Desnormalizados - Testes Preditos pelo Modelo \mathcal{H}_3 - GRU. .	46
Figura 41 – Níveis de Jusante Desnormalizados - Testes Preditos pelo Modelo \mathcal{H}_3 - TCN. .	47
Figura 42 – Previsão x Nivel de Alerta - Evento Real em 10.03.2025, 11 horas AM.	48
Figura 43 – Previsão x Nivel de Alerta - Evento Real em 10.03.2025, às 12:00 AM.	48

LISTA DE TABELAS

Tabela 1 – Dados comuns aplicados as Redes TCN, LSTM e GRU.	30
Tabela 2 – Estrutura da rede neural dos Modelos LSTM \mathcal{H}_1 e \mathcal{H}_2	31
Tabela 3 – Estrutura da rede neural dos Modelos GRU \mathcal{H}_1 e \mathcal{H}_2	31
Tabela 4 – Estrutura da rede neural dos Modelos TCN \mathcal{H}_1 e \mathcal{H}_2	32

SUMÁRIO

Resumo	iv
Abstract	vi
Lista de figuras	ix
Lista de tabelas	xi
Sumário	xii
1	INTRODUÇÃO 1
1.1	Objetivos 3
1.2	Organização do trabalho 4
2	FUNDAMENTAÇÃO TEÓRICA 5
2.1	Introdução 5
2.2	Bacias hidrográficas 5
2.2.1	<i>A Bacia Tocantins-Araguaia 6</i>
2.3	Hidrologia - Níveis e Vazões 8
2.4	Controle de reservatórios das UHE's 10
2.4.1	<i>Vertedouro 12</i>
2.4.2	<i>Processo de controle do reservatório da UHE Tucuruí 14</i>
2.5	<i>Machine Learning 16</i>
2.5.1	<i>Deep Learning 17</i>
2.5.1.1	<i>MLP - Multilayer Perceptron) 17</i>
2.5.1.2	<i>Redes Neurais Convolucionais 19</i>
2.5.1.2.1	<i>Temporal Convolutional Network (TCN) 20</i>
2.5.1.3	<i>Redes Neurais Recorrentes 22</i>
2.5.1.3.1	<i>Long Short-Term Memory (LSTM) 23</i>
2.5.1.3.2	<i>Gated Recurrent Unit (GRU) 25</i>
3	MODELAGENS DE REDES NEURAI ARTIFICIAIS 27
3.1	Metodologia 27
3.2	Levantamento, organização e pré-processamento de dados 27

3.3	Proposição dos modelos	29
3.3.1	<i>Modelo LSTM</i>	30
3.3.2	<i>Modelo GRU</i>	31
3.3.3	<i>Modelo TCN</i>	31
4	RESULTADOS	33
4.1	Treinamento e validação dos modelos H1, H2, H3	33
4.1.1	<i>Função de Perda para as redes neurais LSTM, GRU e TCN</i>	33
4.1.1.1	<i>Rede LSTM</i>	33
4.1.1.2	<i>Rede GRU</i>	34
4.1.1.3	<i>Rede TCN</i>	34
4.1.2	<i>Desempenho dos modelos obtidos pelo FitNRMSE</i>	34
4.1.2.1	<i>Rede LSTM</i>	35
4.1.2.2	<i>Rede GRU</i>	35
4.1.2.3	<i>Rede TCN</i>	36
4.1.3	<i>Desempenho dos modelos por FitR²</i>	37
4.1.3.1	<i>Rede LSTM</i>	37
4.1.3.2	<i>Rede GRU</i>	37
4.1.3.3	<i>Rede TCN</i>	38
4.1.4	<i>Amostras detalhadas dos modelos</i>	39
4.1.4.1	<i>Rede LSTM</i>	39
4.1.4.2	<i>Rede GRU</i>	39
4.1.4.3	<i>Rede TCN</i>	40
4.1.5	<i>Níveis de Jusante Previstos e Desnormalizados</i>	40
4.2	Teste do modelo H3 adotado	42
4.2.1	Normalizado	42
4.2.1.1	<i>Valor Real x Valor Previsto</i>	43
4.2.1.2	<i>Análise dos resíduos (erro temporal)</i>	44
4.2.1.3	<i>Histograma dos Resíduos</i>	45
4.2.2	Desnormalizados	45
4.2.3	Alertas Reais	47
5	CONCLUSÃO	49
	REFERÊNCIAS	51

INTRODUÇÃO

Os reservatórios das Usinas Hidroelétricas (UHE) são sistemas hidrológicos complexos que comportam grandes volumes de água para uma geração perene e sob demanda de energia (AKKAYA, 2024). A gestão de sistemas hidrológicos requer uma modelagem dinâmica que possibilita o controle adequado de suas variáveis (DIAS *et al.*, 2018). Esse cenário é particularmente crítico em períodos chuvosos, quando ocorrem vazões e níveis elevados, exigindo uma capacidade de antecipação nas operações para evitar danos ambientais e sociais, como inundações que atingem populações ribeirinhas, quando os níveis acima (montante) e abaixo (jusante) das barragens tendem a extrapolar limites operacionais (LOUCKS; BEEK, 2017). A natureza não linear, multivariável, variante no tempo, aleatória e sazonal destes sistemas torna sua modelagem uma tarefa não trivial, o que desafia os métodos convencionais a oferecer modelos com previsões que permitam a otimização da geração e a mitigação de danos (SAHOO *et al.*, 2019).

Vale ressaltar que as bacias são unidades fundamentais para o planejamento e o gerenciamento dos recursos hídricos, para serem utilizados no abastecimento e uso da água e na geração de energia elétrica (ROSS; PRETTE, 1998). Para que tudo isso ocorra, a funcionalidade das bacias hidrográficas deve ser equilibrada com as estações climáticas acontecendo de forma regular. Por exemplo, o 'período úmido' nas bacias hidrográficas brasileiras, um conceito ligado à disponibilidade de água e à ocorrência de chuvas, não tem uma data fixa, mas varia conforme a região do país, refletindo-se nos regimes de rios, importantes para a criação de reservatórios, quer seja para uso da água, quer seja para geração de energia (TUCCI, 2017; Agência Nacional de Águas, 2013).

A UHE de Tucuruí, localizada no Estado do Pará, é uma das maiores e mais importantes geradoras de energia elétrica do Brasil, cujo reservatório, no Rio Tocantins, recebe vazões próximas de $40.000m^3/s$ durante o chamado inverno amazônico, o que pode ser potencializado por efeitos das mudanças climáticas (Wikipedia contributors, 2023). Esse grande volume de água provoca alagamentos a montante e a jusante de sua barragem. O nível de jusante é considerado uma variável crítica por estar sujeito a fortes variabilidades, uma vez que é influenciado tanto

pelas vazões turbinadas, correlacionadas à demanda de energia, quanto às vazões vertidas que asseguram as limitações máximas impostas aos níveis de montante (SILVA, 2023). Na figura 1, fica evidenciado (período entre 2010 a 2024), que enchentes no rio Tocantins não são esporádicas, visto que a vazão comportada dentro do leito, sem inundações, é em torno de 25.000 m³/s (MANUAL..., Ano de publicação), e na maioria dos anos às vazões estão acima dessa referencia, conforme destacado com a linha vermelha tracejada na figura 1.

Desta forma, conclui-se que, praticamente, todos os anos, áreas são alagadas a jusante da usina, consequentemente, famílias atingidas com necessidade de remanejamento para abrigos improvisados, o que gera custos e preocupações constantes aos órgãos governamentais, ao agente proprietário da usina e às populações locais.

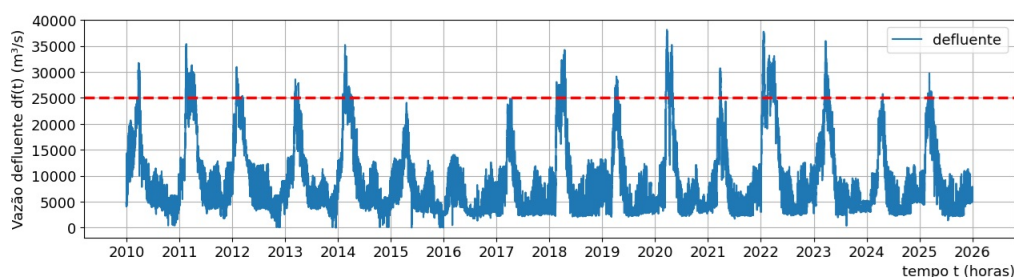


Figura 1 – Histórico de vazão defluente - UHE Tucuruí. Fonte: Elaborada pelo autor.

Para modelar e prever níveis de jusante de reservatórios ou usinas hidrelétricas, as melhores técnicas são aquelas capazes de lidar com séries temporais multivariadas, não linearidades e dependências de longo prazo (ZHAO *et al.*, 2024). Neste contexto, estratégias baseadas em inteligência computacional tem se destacado para esse problema (APAYDIN *et al.*, 2020). Este trabalho insere-se nesta vertente e visa investigar técnicas de aprendizagem profunda *Deep learning* (DL), baseadas em redes neurais artificiais tipo TCN, recorrentes como a LSTM e a GRU por ser uma representação mais enxuta da LSTM. Essas redes são reconhecidas por aprender padrões em dados processados sequencialmente em sistemas que exibem memória (BRUNTON; NOACK; KOUMOUTSAKOS, 2020), o que permite sua utilização na modelagem de sistemas dinâmicos, como os hidrológicos. O uso das redes LSTM, GRU e TCN se justificam pelo fato deste estudo, além de prever nível de jusante, busca também, criar modelo para ser usado futuramente em um Controlador Preditivo de Modelo MPC (KASHIMA, 2023), que faça o controle de nível automaticamente. Rede LSTM tem aplicação recente ao problema de modelagem de níveis e vazões em reservatórios e tem conduzido a resultados exitosos, como por exemplo, em (QI *et al.*, 2019) onde são utilizadas estratégias de engenharia de features por decomposição de modo empírico para treinar LSTM para a previsão de vazões diárias em um reservatório na China, o que demonstram bom desempenho, embora demande por estratégias para além das de DL. Já (KHORRAM; JEHBEZ, 2023) faz uma combinação de LSTM com as Redes Convolucionais (CNN) na previsão de vazões afluentes de rios iranianos atingindo precisão de 92%.

Neste trabalho são avaliadas diferentes topologias de modelos baseados em inteligência

artificial para predição de múltiplos passos a frente do nível de jusante da UHE-Tucuruí, os quais aproveitam a estrutura de processamento tipo sequência para sequência destas redes (SUTSKEVER; VINYALS; LE, 2014). No caso das redes LSTM, GRU e TCN selecionadas como as mais adequadas a este estudo, recebem uma sequência de dados e prevê uma nova sequência de 10 passos a frente. Estes modelos são desenvolvidos tomando como variáveis de entrada, a vazão afluente, a vazão defluente ou uma combinação das duas. Busca-se com isso aproveitar as reconhecidas correlações expressas nos dados entre estas variáveis e o nível a jusante, variável alvo de saída.

O estudo de previsão tem como premissa prever o comportamento do nível de Jusante a barragem da UHE Tucuruí, que tem como nível meta, 10 metros em frente a Cidade de Tucuruí, a partir do qual se tem áreas com alagamentos. Com possibilidade de se ter tais previsões, estas poderão servir de informes para os órgãos competentes tomarem ações antecipadas junto as comunidades que serão afetadas.

O modelo de previsão é a parte mais importante no desenvolvimento deste trabalho e para o qual serão traçadas estratégias visando a construção de algoritmos baseados em Inteligência artificial (IA) utilizando linguagem de programação Python (BOX *et al.*, 2015; RASCHKA; MIRJALILI, 2022).

A estratégia é realizada em cima de um grande histórico de dados hidrológicos de vazões e níveis, coletados do Site aberto do ONS (Operador Nacional do Sistema Elétrico (ONS), 2024). São utilizados neste trabalho, dados dos anos de 2010 a 2025, coletados hora a hora, perfazendo 140.279 mil amostras. Para se obter o desempenho dos modelos aplicou-se a métrica do Erro Quadrático Médio Normalizado ($FitNRMSE$) e para validá-los aplicou-se a métrica do coeficiente de determinação percentual ($FitR^2$) (ENGINEERING; MANAGEMENT, 2025). Os resultados mostraram a viabilidade da aplicação de modelos treinados pelas redes neurais LSTM, GRU e TCN na previsão de nível de jusante para uma janela relevante de tempo, para as quais foram atingidos níveis de ajuste entre 98,1% para a primeira e de 85,2% para a décima amostra, resultados que podem ser considerados adequados em cenários reais de operação.

1.1 Objetivos

Estudar e avaliar modelos de algoritmos baseado em *Deep Learning* que identifiquem a dinâmica do nível de jusante da UHE Tucuruí e gerem previsões dez passos a frente para que possam ser implementadas e acompanhadas pelo sistema de supervisão da usina. Neste contexto, destacam-se os seguintes objetivos específicos:

- Levantar, organizar, analisar e tratar dados reais de amostras hidrológicas horárias;
- Estudar as técnicas e metodologias de aprendizado de máquina mais empregadas no estudo de vazões e níveis hidrológicos;

- Aplicar modelos de algoritimos utilizando Redes Neurais do tipo CNN e RNN para previsão de nível de jusante de hidrelétricas;
- Avaliar e realizar teste de desempenho dos modelos produzidos comparando resultados dos mesmos;

1.2 Organização do trabalho

O presente trabalho está estruturado de modo a conduzir o leitor desde a contextualização do problema até a consolidação dos resultados obtidos.

O Capítulo 1 apresenta a contextualização do tema, evidenciando a relevância do problema investigado sob a perspectiva técnica e científica. Contrmpla, também, os objetivos geral e específicos.

O Capítulo 2 reúne a fundamentação teórica que sustenta o desenvolvimento do trabalho. Inicialmente, são abordados os conceitos relacionados dados de bacias hidrogeáficas, com ênfase para a bacia Tocantins-Araguaia, da hidrologia, do controle de reservatório e do vertedouro. São apresentados também à modelagem da dinâmica implementada e os princípios das redes neurais artificiais, destacando-se as arquiteturas TNC, LSTM e GRU, bem como suas características estruturais.

O Capítulo 3 descreve a metodologia adotada na pesquisa. São apresentados os dados utilizados, os procedimentos de tratamento e normalização das variáveis, bem como a estratégia de divisão dos conjuntos de treino, validação e teste. Também são explicitadas as configurações das arquiteturas neurais implementadas, os critérios de treinamento e as métricas de avaliação empregadas, com ênfase na análise de desempenho em previsão multi-step.

No Capítulo 4, são apresentados e discutidos os resultados obtidos na etapa de modelagem. Realiza-se uma análise comparativa entre os modelos desenvolvidos, considerando o comportamento do erro ao longo do horizonte de predição e a capacidade de generalização das diferentes arquiteturas. Os resultados são examinados de forma crítica, buscando identificar vantagens, limitações e implicações práticas associadas a cada abordagem.

Por fim, o Capítulo 5 apresenta as conclusões do trabalho, destacando as principais contribuições técnicas e científicas alcançadas, as limitações identificadas ao longo da pesquisa e as recomendações para o desenvolvimento de estudos futuros.

FUNDAMENTAÇÃO TEÓRICA

2.1 Introdução

Neste capítulo, serão abordados conceitos gerais de Bacias Hidrográficas, em especial as características da Bacia Tocantins-Araguaia (([ANA](#), 2015)), com foco na Usina Hidrelétrica de Tucuruí, que está presente no Rio Tocantins. Serão tratados, também, aspectos sobre o controle de reservatórios e as variáveis hidrológicas (níveis e vazões), dando-se destaque à variável nível de jusante, que representa a variável a ser controlada.

Este trabalho tem como foco realizar a previsão do nível de jusante da UHE Tucuruí, logo, será necessária uma abordagem sobre modelos dinâmicos deste sistema. Tendo em conta a dificuldade de se modelar este comportamento a partir de leis fundamentais da Física ([BEVEN](#), 2012), o trabalho segue com abordagens baseadas em aprendizado de máquina, que fazem uso de dados históricos do comportamento do sistema ([LJUNG](#), 1999). Essas avaliações permitirão escolher um modelo que melhor represente as previsões de níveis de jusante, visando ações preventivas que possam reduzir impactos causados por inundações de áreas habitáveis ([CLOKE](#); [PAPPENBERGER](#), 2009).

2.2 Bacias hidrográficas

As bacias hidrográficas do Brasil são extensas e formadas por nascentes, lençóis freáticos, área de drenagem, mas, principalmente por rios caudalosos e por outros rios ao seu redor, denominados de afluentes e subafluentes, importantes para suas respectivas regiões, contribuindo para que o país tenha a maior reserva de água doce/potável do mundo, correspondendo a 12% do total ([SILVA](#), 2012).

Uma bacia hidrográfica é delimitada pelo relevo que está ao seu redor ou pela vegetação

que pode ser um fator de divisão entre uma bacia e outra. As bacias podem desaguar no mar, no subsolo e/ou em lagos. Segundo a ([Agência Nacional de Águas e Saneamento Básico \(ANA\), 2021](#)), no Brasil existem doze grandes bacias hidrográficas principais, das quais destaca-se as principais:

- Bacia Amazônica: maior bacia hidrográfica do mundo. Rio principal Amazonas;
- Bacia do São Francisco: Fundamental para o abastecimento do Nordeste. Rio Principal são Francisco;
- Bacia do Paraná: Abriga usinas importantes com a Hidrelétrica de Itaipu. Rio principal Paraná;
- Bacia do Tocantins-Araguaia: Importante para a geração de energia e navegação. Rio principal Tocantins

Destas, a Bacia Tocantins-Araguaia, será mais detalhada, por estar inserida no contexto deste trabalho.

2.2.1 A Bacia Tocantins-Araguaia

A Bacia Tocantins-Araguaia é considerada a maior bacia hidrográfica em território brasileiro e leva o nome de seus dois rios principais: o rio Tocantins e o rio Araguaia. Ambos nascem na região Centro-Oeste do Brasil e escoam rumo à Região Norte ([\(ANA\), 2021](#)). O Rio Tocantins tem sua nascente em Goiás, no extremo norte do estado, e recebe as águas do Rio Araguaia na divisa com o estado do Tocantins. Esta bacia é uma das principais regiões hidrográficas do Brasil, destacando-se pela importância econômica, energética e ambiental. A bacia apresenta grande diversidade ecológica, abrangendo biomas como o Cerrado e a Amazônia ([CARVALHO, 2025](#)). A região possui biodiversidade, com espécies aquáticas, aves e mamíferos típicos desses biomas. Em seu curso estão dispostas as principais usinas hidrelétricas que abastecem a região, que são as de Tucuruí, Lajeado, Peixe Angical, São Salvador, Cana Brava e Serra da Mesa, dispostas em cascata, ou seja, construídas em sequência ao longo do curso de um mesmo rio, conforme dispostas na figura 2 e representadas geograficamente na figura 3. De acordo com a legenda a figura 2, tem-se: UHEs com regularização, que são aquelas com reservatório em que se tem geração de energia, mesmo em período de estiagem e; as UHEs a fio d'água são as que não possuem reservatórios, cuja geração de energia depende única e exclusivamente da vazão natural do rio ([Empresa de Pesquisa Energética, 2022](#))

O Rio Tocantins atravessa importantes áreas produtivas do Centro-Oeste e Norte do país, cujo transporte hidroviário seria relevante e estratégico, mas os trechos do rio que antecedem a confluência com o Araguaia não são navegáveis ([GUITARRARA, s.d](#)).

Já o Rio Araguaia tem sua nascente localizada nas proximidades do Parque Nacional das Emas, na divisa entre Goiás e Mato Grosso, em uma altitude de 850 metros ([\(ANA\), 2009](#)). No seu

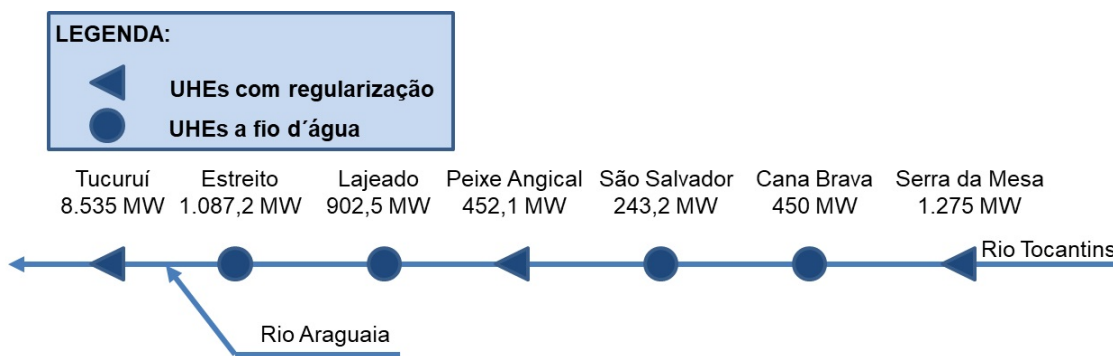


Figura 2 – Usinas em cascata no Rio Tocantins.

Fonte: Adaptado de ONS, 2022.

percurso está incluso a Ilha do Bananal, a maior ilha fluvial do mundo e de grande valor turístico para a região (IBGE, 2004). O Rio Araguaia não possui hidrelétricas, mas possui um extenso trecho navegável, utilizado principalmente para o transporte de cargas (GUITARRARA, s.d).

Cabe ressaltar que a Bacia Tocantins-Araguaia tem uma área de aproximadamente 920.000 Km², equivalente a 10,8% do território nacional, englobando partes dos estados de Tocantins, Pará, Maranhão, Goiás, Mato Grosso e Distrito Federal. A Figura 3 apresenta um mapa da região com destaque para a rica hidrografia que é drenada pelos seus dois principais rios (AMBIENTE, 2006). A Bacia tem grande importância para a região, pois é responsável pelo abastecimento de áreas com uma grande diversidade ecológica e produtiva. Em muitos municípios do centro norte do Brasil, algumas atividades como a extração mineral, vegetal e agropecuária, desempenham papel principal e de grande relevância para essa região. Como destaque no setor econômico temos a mineração na Serra dos Carajás no Pará, e os plantios de commodities agrícolas, como a soja, o milho e o algodão nas áreas do cerrado e que agora se estendem para áreas da floresta amazônica (SILVEIRA *et al.*, 2022). Destaca-se também o uso da água destinada à irrigação e para uso industrial. Vale ressaltar que algumas dessas atividades estão diretamente associadas ao processo de poluição das águas da bacia e degradação ambiental decorrentes disso (GUITARRARA, s.d).

O clima predominante da Bacia do Tocantins-Araguaia é tropical, tornando-o mais úmido e quente, conforme se aproxima da foz, localizada na Ilha do Marajó, Estado do Pará. Ao longo do ano são evidenciadas duas características climáticas bem distintas: um período seco de estiagem, entre os meses de setembro e outubro, e um outro mais úmido no período de fevereiro a abril, quando as chuvas são bem mais frequentes (BARRAGENS, 2000).

Quanto à vegetação, o Cerrado é o bioma dominante, dado que grande parte da Bacia está situada na Região Centro-Oeste. Já nas regiões mais ao Norte, a Bacia é marcada pela presença da floresta Amazônica. Nas últimas décadas esta região tem sido alvo de grandes desmatamentos, dada a intensa presença da agropecuária, da mineração e da exploração madeireira, o que gera impactos negativos ao solo e rios da Bacia (AMBIENTE, 2006).

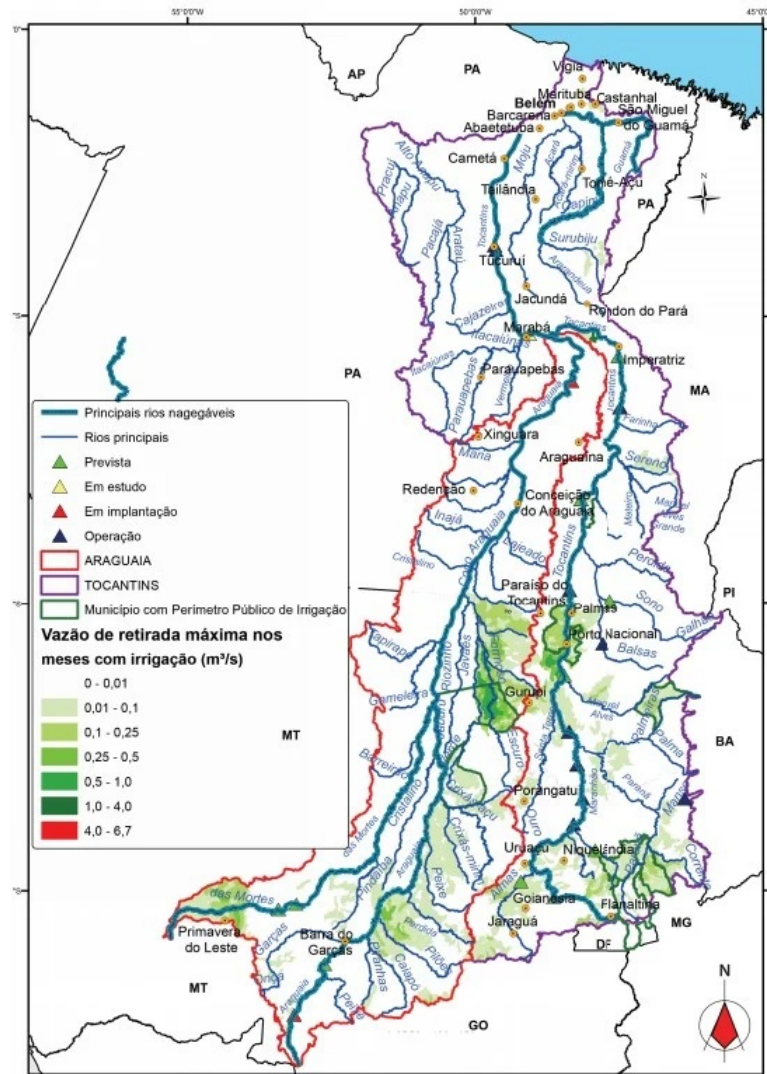


Figura 3 – Bacia Tocantins-Araguaia.

Fonte: Escola Educação, 2019.

2.3 Hidrologia - Níveis e Vazões

As variáveis hidrológicas nível e vazão são de suma importância para descrever e gerenciar o comportamento dos recursos hídricos de uma bacia hidrográfica (PORTO; PORTO, 2012). Para que se tenha uma gestão eficaz dos recursos hídricos e mitiguem impactos de eventos e mudanças ambientais nesta bacia, faz-se necessário a medir, monitorar e acompanhar as variáveis nível e vazão ((ANA), 2012). A natureza dinâmica da hidrologia está relacionada à variação temporal e espacial do fluxo de água em um sistema hidrológico (GUILHON; ROCHA; MOREIRA, 2007). Em bacias com extensa área de drenagem, a equação de balanço hídrico é empregada na análise quantitativa das demandas e/ou disponibilidades de água, com o objetivo de subsidiar ações de gestão, controle e preservação dos recursos hídricos (JÚNIOR, 2022). Compreender essa dinâmica é fundamental para a realização de uma gestão eficaz e sustentável dos recursos hídricos e para o

controle dos reservatórios nos rios que os têm.

A vazão hidrológica em um local específico pode variar ao longo do tempo devido a vários fatores como sazonalidade das chuvas, mudanças na vegetação, armazenamento de águas em reservatórios e, dependendo da região, fusão da neve ([World Meteorological Organization \(WMO\), 2008](#)). Tais variações ocorrem ao longo do tempo e são influenciadas por condições meteorológicas, pelas características hidrográficas da bacia e por interferências humanas, algo visível e bem discutido nas últimas décadas, principalmente pela imprensa.

Para que se tenha a representação quantitativa dos processos e fenômenos que ocorrem dentro de um sistema hidrológico, precisa-se entender a natureza matemática e a dinâmica do mesmo. Algumas das características da natureza matemática de um sistema hidrológico ([DOOGE; KANE, 1997](#)) são:

- Equações de conservação: São equações fundamentais na modelagem matemática que descrevem o comportamento da água e energia ao longo do tempo e do espaço levando-se em consideração o balanço de água na superfície do solo, escoamento superficial e subterrâneo;
- Equações de fluxo: Representadas por equações diferenciais que descrevem a variação dos fluxos em resposta aos gradientes de pressão e elevação, condutibilidade hidráulica e outras propriedades físicas do sistema;
- Modelos de precipitação: Podem ser modelados por métodos estatísticos, de regressão, de interpolação espacial, de simulação estocástica, entre outros, para prever padrões de precipitação com base em dados observados ou previsões meteorológicas;
- Modelos de escoamento: Podem ser descritos por equações diferenciais parciais, modelos de elementos finitos ou baseados em redes neurais artificiais ou modelos baseados em sistemas de informação geográfico (SIG) para prever os processos de escoamento superficial e subterrâneo da água;
- Modelos de Evaporação: Podem ser modelados por equações que levam em consideração fatores como umidade relativa, temperatura, radiação solar, velocidade do vento e características da vegetação.

Ainda neste contexto, ([NAGHETTINI, 2017](#)), destaca também características intrínsecas aos dados que surgem quando se trabalha com observações de chuva, nível, vazão, etc, que são:

- Estacionalidade: São dados hidrológicos que costumam ter padrões sazonais (cheias anuais, períodos secos e outros). Estaticamente, surgem como componentes periódicos nas séries temporais.

- Autocorrelação temporal: são valores atuais que dependem de valores passados (chuva acumulada, nível do reservatório, etc), que gera memória longa nas séries;
- Não linearidade nos dados: A relação entre as variáveis não é linear, como por exemplo chuva x vazão, que pode ser detectada por métodos estatísticos ou por necessidade de modelos não lineares (LSTM, GRU e TCN);
- Interdependência entre variáveis: Diferentes series são correlacionadas. Por exemplo: Chuva x vazão afluente x nível do reservatório x vazão defluente x nível de jusante, etc.
- Distribuição não normais: Dados de chuva e vazão geralmente têm distribuições assimétricas e com caudas grandes (enchentes); item Dados incompletos e ruídos: Séries hidrológicas frequentemente apresentam falhas de medição, lacunas e ruído instrumental, exigindo pré-processamento (interpolação, filtragem, normalização).
- Eventos extremos: São dados que apresentam outliers físicos significativos (enchentes, secas), que do ponto de vista estocástico, são eventos difíceis, mas de grande impacto.

Dessa forma pode se dizer que a modelagem matemática é uma ferramenta de grande relevância para se compreender e prever o comportamento dos sistemas hidrológicos (BRUNNER, 2023). Do ponto de vista baseado em dados, um sistema hidrológico é estocástico, não linear, sazonal, autocorrelacionado, multivariado, com heteroscedasticidade e sujeito a extremos (DI; YANG; WANG, 2014), (WILKS, 2011). Assim, tanto as equações matemáticas, baseadas em física, quanto os dados são bastante utilizados na pesquisa científica, na gestão de recursos hídricos e na engenharia de sistemas.

2.4 Controle de reservatórios das UHE's

No reservatório da UHE, atuam diversas variáveis hidrológicas que expressam o estado do sistema a partir da sua evolução dinâmica no tempo t . A vazão afluente $v_a(t)$, gerada pelas chuvas na bacia hidrográfica e que escoam naturalmente pelos rios, é interpretada como a variável de entrada. O acúmulo desta água gera um volume $u_v(t)$ que determina o nível $n_m(t)$ a montante da barragem. A água armazenada é utilizada para geração de energia através da vazão turbinada $v_t(t)$. Para respeitar as cotas operativas máximas, o excedente hídrico é drenado pelo vertedouro, configurando a vazão vertida $v_v(t)$. Portanto, a vazão defluente $v_d(t)$ é composta pela soma $v_t(t) + v_v(t)$. Por fim, esta defluência determina o nível a jusante $n_j(t)$, variável-alvo deste trabalho. A Figura 4 ilustra as variáveis no processo hidrológico e suas respectivas relações.

O diagrama inclui uma malha de controle, mostrada na Figura 5, que permite aos operadores decidirem a atuação sobre as vazões e com base nos níveis $n_m(t)$ e $n_j(t)$ observados. Na UHE Tucuruí, durante o inverno amazônico, a abertura dos vertedouros é o principal meio de atuação

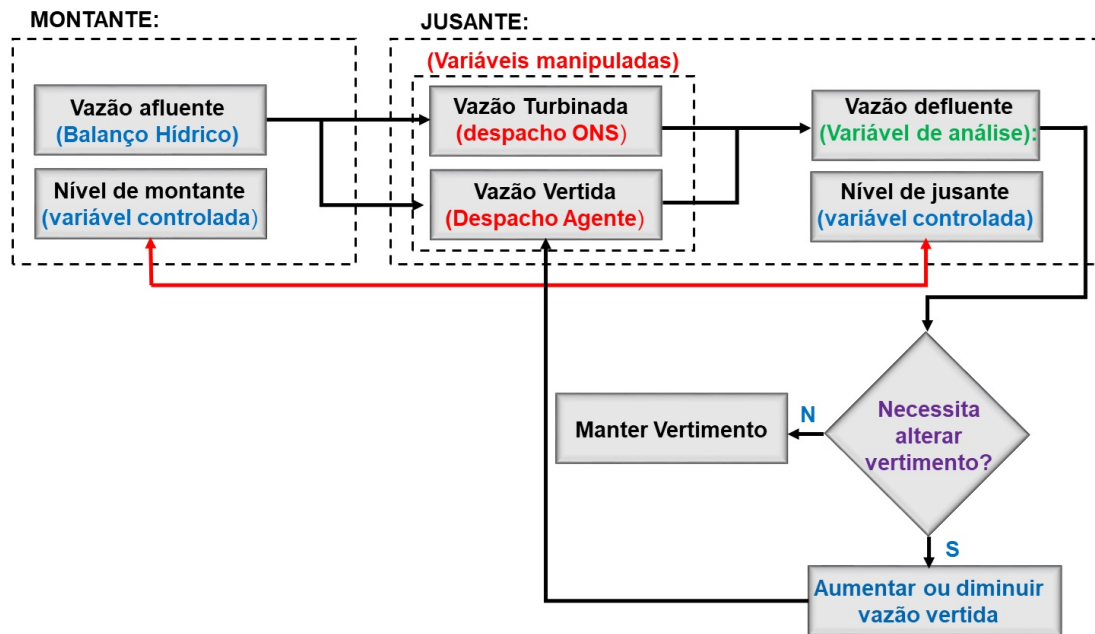


Figura 4 – Diagrama de blocos.

Fonte: Elaborada pelo autor.

para o controle do nível de montante. Contudo, a tomada de decisão deve respeitar o comportamento dinâmico do nível de jusante para evitar danos por elevações abruptas e alagamentos. Quando o vertedouro está aberto, a vazão defluente, bem como os níveis de montante e jusante, tornam-se as variáveis-chave para decidir sobre a abertura ou fechamento das comportas. A vazão turbinada é outra variável de impacto direto, considerada manipulada por ser operada pela UHE Tucuruí, sob coordenação do Operador Nacional do Sistema (ONS) e vinculada às demandas do Sistema Interligado Nacional (SIN).

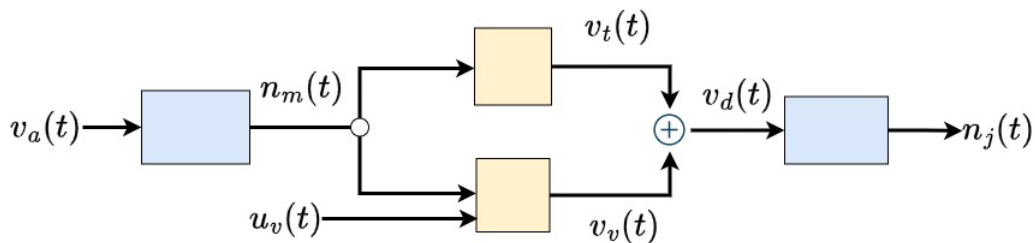


Figura 5 – Malha de controle.

Fonte: Elaborada pelo autor.

Por meio do diagrama da Figura 5 é possível verificar a interdependência entre as variáveis. Assim sendo, este trabalho avalia algoritmos inteligentes visando modelar relações dinâmicas entre as vazões $v_a(t)$ e $v_d(t)$, que são tomadas como variáveis de entrada, e o nível de jusante

$n_j(t)$, a saída do modelo, de modo que se pode considerar:

$$n_j(t) = \mathcal{H}_1\{v_a(t)\} \quad (2.1)$$

$$n_j(t) = \mathcal{H}_2\{v_d(t)\} \quad (2.2)$$

$$n_j(t) = \mathcal{H}_3\{v_a(t), v_d(t)\} \quad (2.3)$$

Neste caso $\mathcal{H}\{\cdot\}$ é um operador não linear, multivariável, que representa a dinâmica do sistema.

Nesta conjuntura, a modelagem de sistemas hidrológicos desempenha um papel importante na gestão de recursos hídricos, previsão de cheias, controle de reservatórios e outros processos relacionados ao ciclo da água (QIU *et al.*, 2019). Tradicionalmente, esses sistemas são modelados com base na hidrodinâmica clássica (NAVARRO; PLAYÁN, 2008), que utiliza princípios físicos, equações de continuidade, conservação de energia e as equações de Navier-Stokes. Esses modelos fundamentam-se em conjecturas rigorosas, parâmetros específicos e descrições físicas, exigindo, muitas vezes, medições de campo precisas e alta capacidade computacional. No entanto, a complexidade intrínseca aos sistemas hidrológicos, que apresentam comportamentos não lineares, interações multiescalares e dependência de condições iniciais e de contorno, desafia as abordagens convencionais (KOUTSOYIANNIS, 2013). Outrossim, fatores como mudanças climáticas e a escassez de dados podem comprometer a capacidade de previsão dos modelos hidrodinâmicos

Neste contexto, o Aprendizado de Máquina (AM) surge como uma alternativa ou complemento aos modelos tradicionais. Ao invés de priorizar as relações físicas explícitas entre as variáveis AM, esse modelo baseia-se em dados observados para aprender padrões e inferir a dinâmica do sistema (REICHSTEIN *et al.*, 2019). Paralelamente, modelos baseados em redes neurais têm se destacado na modelagem de sistemas hidrológicos, dada a sua alta capacidade de capturar dependências temporais e padrões não lineares (KRATZERT *et al.*, 2018).

Na primeira etapa, serão avaliados os comportamentos do nível de jusante em relação às vazões de entrada (afluente e defluente). Com base nessa observação, será modelada a dinâmica que explica o nível de jusante em função das demais variáveis do sistema. No contexto operacional de uma usina hidrelétrica, a alteração do nível de jusante é realizada principalmente pelas vazões turbinada e vertida, com ênfase nesta última, operacionalizada pela abertura e fechamento do vertedouro.

2.4.1 Vertedouro

O vertedouro ou vertedor é uma estrutura física construída em barragens, principalmente de hidrelétricas, destinada ao controle de vazão ou nível. Ele transpõe o excedente de água não utilizada na geração de energia quando o nível de montante atinge a cota de segurança, evitando que a barragem transborde e sofra danos (Agência Nacional de Águas (ANA), 2016). A Figura 6 apresenta os elementos da usina de Tucuruí, com destaque para o vertedouro composto por 23 comportas

(manobradas com aberturas de 0 a 12 metros), com capacidade para uma vazão vertida máxima de $110.000 \text{ m}^3/\text{s}$.



Figura 6 – UHE Tucuruí.

Fonte: Adaptado de Ferdinando de Sousa 2019.

As manobras do vertedouro são realizadas individualmente, de modo remoto, ou através de comando elétrico local nas proximidades de cada comporta. Tais comandos chegam até servo motores CA de cada comporta realizando abertura ou fechamento da mesma de acordo com o valor em metros estipulado a cada comporta.

No período de cheia, de janeiro a maio, normalmente, é necessário abrir as comportas do vertedouro para escoar certa quantidade de água que chega no reservatório, com o objetivo de controlar e tornar mais suave a elevação deste e evitar, também, a subida mais abrupta do nível a jusante da barragem. Para evitar tal situação, a Operação da UHE Tucuruí estabeleceu premissas para início do vertimento baseadas em relações que envolvem o nível de montante e a diferença entre as vazões afluente e defluente, conforme a seguir:

- Nível do reservatório igual a $70,00\text{m}$ e diferença entre vazões afluente e defluente => $20.000\text{m}^3/\text{s}$, ou;
- Nível do reservatório igual a $71,00\text{m}$ e diferença entre vazões afluente e defluente => $15.000\text{m}^3/\text{s}$, ou;
- Nível do reservatório igual a $72,00\text{m}$ e diferença entre vazões afluente e defluente => $10.000\text{m}^3/\text{s}$, ou;
- Nível do reservatório igual a $73,00\text{m}$ e diferença entre vazões afluente e defluente => $5.000\text{m}^3/\text{s}$.

Tais premissas estão definidas nas instruções internas de operação da própria usina e buscam limitar vertimentos superiores a 3000 m³/s diários e evitar elevações ou reduções abruptas no nível de jusante o que pode causar transtornos as comunidades ribeirinhas.

Além das premissas, outras variantes são consideradas:

- Período de Cheia:
 - Dezembro a maio: Nível de Água (NA) em torno de 73,80 m para permitir receber variações de vazões não esperadas provocadas por pancadas de chuva sobre o reservatório e evitar aberturas do vertedouro com vazões que possam, a curto prazo, atingir moradores a jusante;
 - Final da cheia – junho e julho NA 74,00 m (nível máximo normal do reservatório);
- Período de seca:
 - Junho a dezembro: Segue-se a curva referencial de deplecionamento estabelecido em comum acordo entre ONS e agente gestor da usina conforme figura 7.

A UHE Tucuruí é parte da concepção integrada do aproveitamento hidrelétrico das bacias dos rios Tocantins e Araguaia e de grande relevância para o atendimento ao SIN e a própria região norte, não somente no período úmido, mas, principalmente, no período seco, de junho a dezembro, em que a maioria das bacias do norte perdem suas capacidades hidroelétricas ([Secretaria de Estado de Planejamento e Desenvolvimento \(SEPLAD\) - PA, 2015](#)). Assim sendo, o ONS, junto com o agente responsável pela gestão do reservatório, mediante estudos de cenários hidrológicos e de consumo de energia previstos para os meses de junho a dezembro ([Operador Nacional do Sistema Elétrico \(ONS\), 2018](#)), traçam a chamada “Curva Referencial de Deplecionamento” do reservatório da UHE Tucuruí, mostrada na figura 7 para o ano de 2023. Esta curva visa a exploração adequada dos recursos hídricos da UHE Tucuruí e otimizá-los em face à necessidade de atendimento às demandas do Sistema Interligado Nacional – SIN no período seco.

2.4.2 Processo de controle do reservatório da UHE Tucuruí

A área de operação em tempo real da usina realiza leituras horárias nas régua limnimétricas de montante e jusante e as insere no software de supervisão operacional, que recebe, também, automaticamente, do Sistema de Medição e Faturamento (SMF), os valores de geração de cada Unidade Geradora (UG) e realiza cálculos das vazões hidrológicas (afluente, turbinada e defluente). A Equipe de pré operação acompanha os dados de níveis e vazões disponibilizados pelo sistema e baseados nas premissas descritas no item 2.4.1, avalia a necessidade de alterar ou não o vertimento. Hoje, este cálculo da vazão a ser vertida é realizado manualmente pela equipe de pré operação, utilizando tabelas e planilhas em excel.

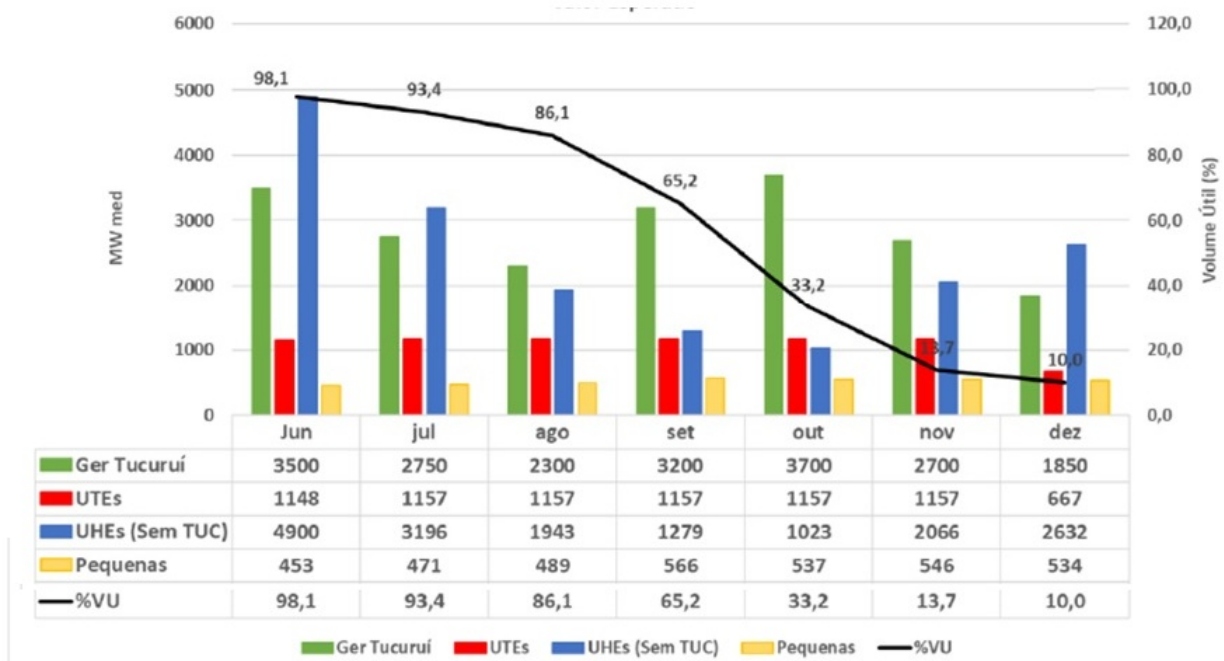


Figura 7 – Curva Referencial de Deplecionamento da UHE Tucuruí 2023.

Fonte: NT-ONS DOP 0059/2023.

O foco deste trabalho é ter como resultado um modelo de previsão que determine, hora a hora, de maneira automática e preventiva a evolução ou involução das vazões e níveis para as próximas 10 horas subsequentes, indicando a necessidade de se abrir ou fechar vertedouro de forma a garantir o controle efetivo do reservatório dentro do nível meta para o período e minimizar impactos a jusante.

A quantidade de água a ser vertida é estabelecida a partir da obtenção das previsões de vazão afluente, determinada através de equações de regressão, e da geração estimada para o período de sete (7) dias a frente. No entanto, a geração programada para o dia seguinte é consensada entre agente e ONS na noite que antecede a sua realização, o que torna o processo de controle do reservatório, através do vertimento, muito dinâmico e exigindo, algumas vezes, redefinição do valor a ser vertido a curto prazo, questões de horas. Ainda neste contexto, tem-se mais um agravante, a geração pode variar em tempo real (executada fora do programado), atendendo condições sistêmicas, tornando o controle de vazão e nível ainda mais difícil.

Para se obter o valor da vazão a ser vertida $v_v(t)$ utiliza-se a expressão:

$$v_v(t) = v_a(t) - v_t(t) \tag{2.4}$$

onde $v_a(t)$ é a vazão afluente e $v_t(t)$ a vazão turbinada.

A vazão vertida obtida em (2.4) é distribuída nas comportas, cuja quantidade e o valor da abertura em metro (m) dependem da vazão vertida calculada. Contudo, para se chegar ao valor da vazão a ser vertida — prevendo eventos do tipo alagamentos de áreas a montante ou

jusante — é necessário que se saiba a vazão afluente que está chegando no reservatório, bem como a vazão que está sendo defluída (turbinada + vertida). A partir dessas vazões, busca-se prever níveis de jusante com intuito de, futuramente, realizar seu controle preditivo de forma automática automaticamente.

Para isso, serão utilizados modelos de redes neurais que se adequam melhor aos sistemas hidrológicos que são temporais e com dinâmica lenta. Entre as redes neurais mais apropriadas para modelagem dinâmica e aplicação futura em MPC, destacam-se as arquiteturas LSTM, GRU e TCN. As LSTM têm se mostrado eficazes como modelos substitutos em MPC (KUMAR *et al.*, 2021), enquanto GRUs vêm sendo aplicadas para Controle Preditivo por Modelo Não Linear (NMPC) com bom desempenho (CHEN *et al.*, 2020). Já as TCNs apresentam vantagens comprovadas na modelagem de sequências temporais e vêm sendo estudadas para MPC não linear (BAI; KOLTER; KOLTUN, 2018) e (ZHI; HUANG; WANG, 2022). Apesar do controlador MPC não ser foco direto deste trabalho, está sendo citado, por ter grande probabilidade de ser aplicado como um sistema de controle automático dos níveis do reservatório e a jusante, tendo como modelo, o melhor resultado dos modelos aqui estudados. A estrutura geral do MPC e sua integração com modelos não lineares pode ser encontrada em (CAMACHO; ALBA, 2013).

2.5 Machine Learning

Aprendizado de máquina é uma vertente específica da Inteligência Artificial (IA) que se concentra no desenvolvimento de algoritmos e modelos estatísticos que permitem a máquina aprender com uso de dados (MITCHELL, 1997). Em ML, os algoritmos passam a ideia de que as máquinas são capazes de aprender com os dados, sem serem explicitamente programados (SAMUEL, 1959).

A característica do ML é aprender padrões relevantes nos dados de treino e aplicar uma hipótese que estima a função desejada. A qualidade de generalização do problema permite que o sistema apresente um bom desempenho ao ser submetido a novos dados que nunca viu antes (AWAD; KHANNA, 2015).

Segundo (BISHOP; NASRABADI, 2006), as principais características, e componentes do *machine learning* são:

- Algoritmos de aprendizado: Um algoritmo de aprendizado de máquina é projetado para aprender com os dados disponíveis, identificando padrões para realizar tarefas ou tomar decisões com base nesses padrões. Eles são capazes de ajustar seus parâmetros internos com base nas informações passadas durante o treinamento;
- Dados de treinamentos: São informações fornecidas a um algoritmo de aprendizado de máquina durante o processo de treino. Esses dados são essenciais e normalmente incluem características que descrevem objetos a serem aprendidos durante o treinamento para melhorar suas capacidades de realizar previsões sobre novos dados não vistos anteriormente;

- Aprendizado Supervisionado e Não supervisionado: Existem duas abordagens principais em ML:
 - a) Aprendizado Supervisionado: Consiste no aprendizado formado pelos conjuntos de atributos de entrada e de saída, chamados também de rótulos ou seja o valor de saída é conhecido e induzido por ação humana. Tem como objetivo treinar o modelo com esses dados para que ele possa fazer previsões precisas em dados não rotulados;
 - b) Aprendizado não Supervisionado: Nesse aprendizado o algoritmo analisa e agrupa conjuntos de dados não rotulados, portanto, não há a solução desejada ao algoritmo. O algoritmo encontra padrões nos dados e agrupa-os em diferentes segmentos com base nesses padrões
- Validação e teste: Todo conjunto de dados para ter a melhor performance necessita de ser separado em três (03) sub-conjuntos: O primeiro sub-conjunto é utilizado para treinamento do modelo; o segundo sub-conjunto é o de validação usado para avaliar a precisão do modelo e ajustar os parâmetros, buscando evitar-se o overfitting. O overfitting ocorre quando o modelo se ajusta muito bem aos dados de treinamento, mas não consegue generalizar para novos dados; enquanto o sub-conjunto de teste é usado para avaliar a precisão do modelo com dados não vistos anteriormente.

2.5.1 *Deep Learning*

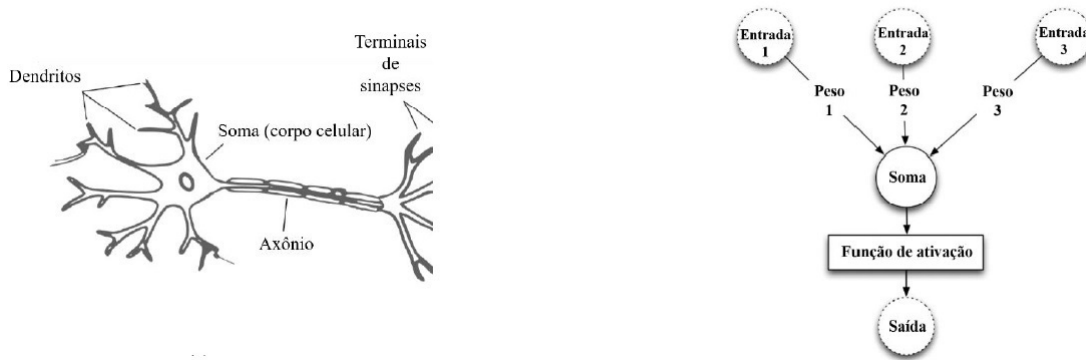
DL é uma divisão de *Machine Learning* (ML). DL é uma rede neural profunda, com três ou mais camadas, que tem como objetivo simular o comportamento do cérebro humano, para que a máquina “aprenda”, conceitos complicados descritos por uma lista de regras matemáticas formais (GOODFELLOW; BENGIO; COURVILLE, 2016). Uma rede neural simples, com uma única camada, consegue fazer previsões aproximadas, no entanto, camadas ocultas adicionais podem ajudar a otimizar e melhorar a precisão (BENGIO; GOODFELLOW; COURVILLE, 2017).

Na atualidade, os algoritmos de *Deep Learning* tem vasta aplicação e dentre essas redes tem-se: Multilayer Perceptron (MLP), CNN, Recurrent Neural Network (RNN) e Temporal Fusion Transformer (TFT), que utilizam esses algoritmos para resolver problemas de regressão estática e dinâmica, com destaque para aqueles que trata da análise de problemas com dependência temporal (GOODFELLOW; BENGIO; COURVILLE, 2016).

2.5.1.1 *MLP - Multilayer Perceptron*

As RNA's são estruturas representadas por modelos matemáticos que buscam simular características de neurônios biológicos, equivalentes ao cérebro de animais, em ambiente computacional, com capacidade de aprender e reconhecer padrões (FURTADO, 2019). Os neurônios biológicos são compostos de dendritos (canais de entrada das informações), responsáveis pela recepção dos

sinais, corpo celular, onde a informação é processada e, axônios, que transmitem os sinais gerados aos terminais de sinapse - Figura 8a. Os terminais de sinapse se conectam a outros dendritos formando uma estrutura interconectada de rede (KRUG *et al.*, 2008).



(a) Estrutura celular de um neurônio biológico.

(b) Neurônio artificial simplificado (perceptron).

Figura 8 – Estruturas simplfcadas de neurônios biológico e artificial.

Fonte: Silva (2018) e Heaton (2012).

A estrutura modelada com representação matemática de um neurônio biológico é chamada de perceptron (ROSENBLATT, 1958), conforme Figura 8ab. Observa-se que, as entradas fornecidas são multiplicadas por pesos, cujo resultado da soma é inserida em uma função de ativação. Na Equação (2.5) as variáveis x e w representam as entradas e os pesos, respectivamente, enquanto a variável i representa a quantidade de entradas e dos pesos (GOODFELLOW; BENGIO; COURVILLE, 2016). Cada entrada é multiplicada pelo seu peso correspondente e a soma dos resultados alimenta a função de ativação σ . Essa operação não-linear que atua sobre o potencial de ativação produz uma única saída do neurônio (HAYKIN, 1998).

$$f(x_i, w_i) = \varphi(\sum(x_i \cdot w_i)) \quad (2.5)$$

O perceptron, embora fundamental, demonstrava baixa acurácia em problemas lógicos simples, algumas limitações para os objetivos mais audaciosos das RNA's, logo, surgiu a proposta da união de vários perceptrons para a formação de uma Rede Neural Artificial chamada Perceptrons em Multicamadas (MLP - Multilayer Perceptron) (SANTOS *et al.*, 2020).

De acordo com (FLECK *et al.*, 2016) em uma MLP, os neurônios estão dispostos em várias camadas, classificadas como sendo: uma camada de entrada, uma ou mais camadas intermediárias (ocultas) e uma camada de saída. Cada neurônio de uma determinada camada está conectado a todos os outros na camada seguinte por meio de conexões ponderadas, ver figura 9.

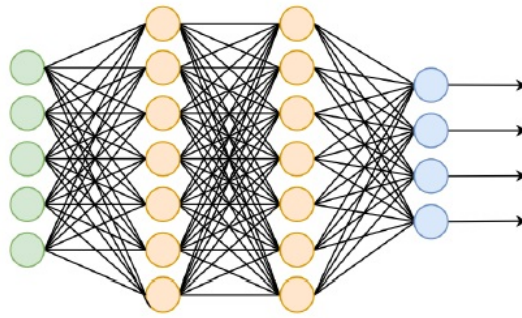


Figura 9 – Rede Neurais em Multicamadas - MLP.

Fonte: Silva (2018).

Cada conexão tem um peso associado que é ajustado na fase de treinamento para contribuir na aprendizagem dos padrões nos dados de entrada. A função de ativação é aplicada em todas as camadas intermediárias e de saída, permitindo que ela aprenda relações não lineares nos dados (BISHOP; NASRABADI, 2006).

Na fase de treinamento, os pesos são ajustados utilizando algoritmo de retropropagação que calcula valores de gradientes afim de minimizar a diferença entre as saídas e os valores desejados da rede (RUMELHART; HINTON; WILLIAMS, 1986). As redes MLP tem a capacidade de aprender representações complexas e extrair atributos dos dados de entrada, com intuito de adequá-los a uma grande variedade de problemas tipo classificação e regressão (BISHOP, 1995). No entanto, em problemas onde informações dinâmicas são necessárias, como é o caso da modelagem de dados hidrológicos, as redes MLP não possuem elementos de memória em sua estrutura, que deem ao modelo capacidade de representar o processamento sequencial de informação dos dados citados (GOODFELLOW; BENGIO; COURVILLE, 2016). Neste caso, é comum utilizar arquiteturas mais complexas, como as redes convolucionais (RNC) ou redes neurais recorrentes (RNN).

2.5.1.2 Redes Neurais Convolucionais

A Convolutional Neural Network (Rede Neural Convolucional - RNC) é um tipo de rede neural projetada para extrair automaticamente padrões espaciais ou temporais dos dados usando operações de convolução (VAN *et al.*, 2020). A ideia principal é usar operações de convolução em conexões densas para conseguir extrair padrões hierárquicos e locais, ou seja, uma RNC aprende filtros automáticos que detectam padrões em dados como: bordas; formas e tendências em séries (GOODFELLOW; BENGIO; COURVILLE, 2016). O funcionamento da RNC é baseado em camadas de convolução e pooling que reduzem a dimensionalidade mantendo as informações mais relevantes e utiliza camadas densas finais que combinam os padrões aprendidos para gerar a saída (HA; YUN; CHOI, 2015), conforme figura 10.

Matematicamente tem-se a convolução em 1D ou 2D dada por:

$$s(t) = (x * w)(t) = \sum_{\tau} x(\tau)w(t - \tau) \quad (2.6)$$

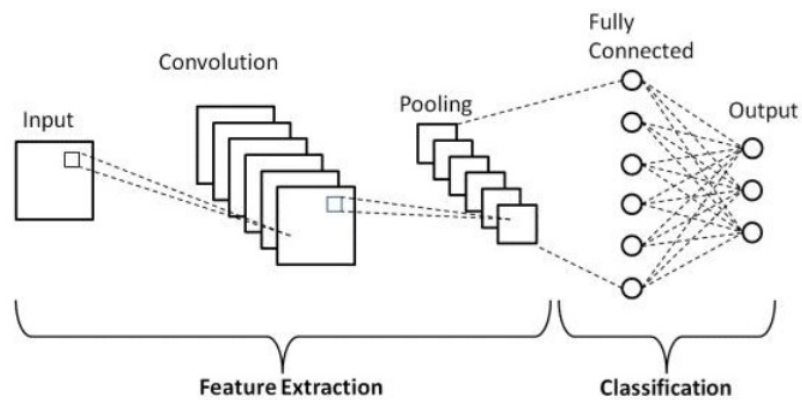


Figura 10 – Rede Neural Convolutacional - RNC.

Fonte: Couto, Gabriel. O Poder das Redes Neurais Convolucionais: [Linkedin.com](https://www.linkedin.com/pulse/o-poder-das-redes-neurais-convolucionais-um-guia-pratico-couto-j7omf/), 2023. Disponível em: www.linkedin.com/pulse/o-poder-das-redes-neurais-convolucionais-um-guia-pratico-couto-j7omf/

onde:

- x = entrada (sinal, série ou imagem),
- w = filtro (kernel) aprendido,
- s = saída (feature map).

As principais características das RNC são:

- Localidade: aprende padrões locais em: imagens, bordas; em séries: tendências locais),
- Compartilhamento de pesos: o mesmo filtro é aplicado em diferentes regiões utilizando menos parâmetros que redes densas,
- Hierarquia de características: as camadas iniciais detectam padrões simples e as camadas profundas detectam estruturas mais complexas.

Ainda que famosas em visão computacional (FARIAS; ROSSI; SUL-MS-BRAZIL, 2021), as RNC são aplicáveis, também, em séries temporais e hidrologia. Desse grupo se dará ênfase a TCN.

2.5.1.2.1 Temporal Convolutional Network (TCN)

A TCN é uma rede convolutacional adaptada especificamente para dados sequenciais temporais, diferentemente das CNN's tradicionais que utilizam imagens, tendo como vantagem a paralelização das RNC's, o que evita problemas típicos como gradiente desvanecente, problema comum no qual os gradientes diminuem ao longo do tempo durante o processo de retropropagação dificultando o aprendizado de dependências de longo prazo. A Temporal Convolutional

Network (TCN) é uma arquitetura de rede neural voltada para o processamento de séries temporais propostas por (BAI; KOLTER; KOLTUN, 2018). Nesse estudo, os autores a compararam com redes recorrentes, como LSTM e GRU, para modelagem de sequência, e provaram que a TCN, frequentemente, superava as RNNs citadas em estabilidade e precisão. A TCN foi projetada para: aprender dependências de longo prazo em series temporais; produzir previsões univariadas ou multivariadas de forma robusta; ser eficiente e paralelizável e; ser estável no treinamento sem apresentar problemas de gradientes (BAI; KOLTER; KOLTUN, 2018).

A figura 11 representa a Arquitetura de uma TCN com camadas correspondentes a fatores de dilatação exponencialmente crescentes $d = 1, 2, 4$, onde os círculos verdes correspondem a camada de entrada, os círculos azuis correspondem as camadas ocultas e os círculos vermelhos correspondem a camada de saída (ENKHBAT *et al.*, 2024). Em geral, observa-se que o fator de dilatação é duplicado a cada camada, onde a unidade 2, primeira camada oculta, processa as informações das observações 1 e 2 e a unidade 4, segunda camada oculta, processa as informações 1, 2, 3 e 4. E assim por diante. Para evitar que a estrutura de rede mais profunda tenha dificuldade com o processo de aprendizagem, uma conexão residual é adicionada à saída (HE *et al.*, 2016). Como a primeira e a segunda camadas convolucionais do primeiro e ultimo bloco residual, respectivamente, podem ter larguras de canal de entrada e saída diferentes, a largura do tensor residual pode precisar ser ajustada, usando uma convolução 1×1 . (POLLHAMMER *et al.*, 2017).

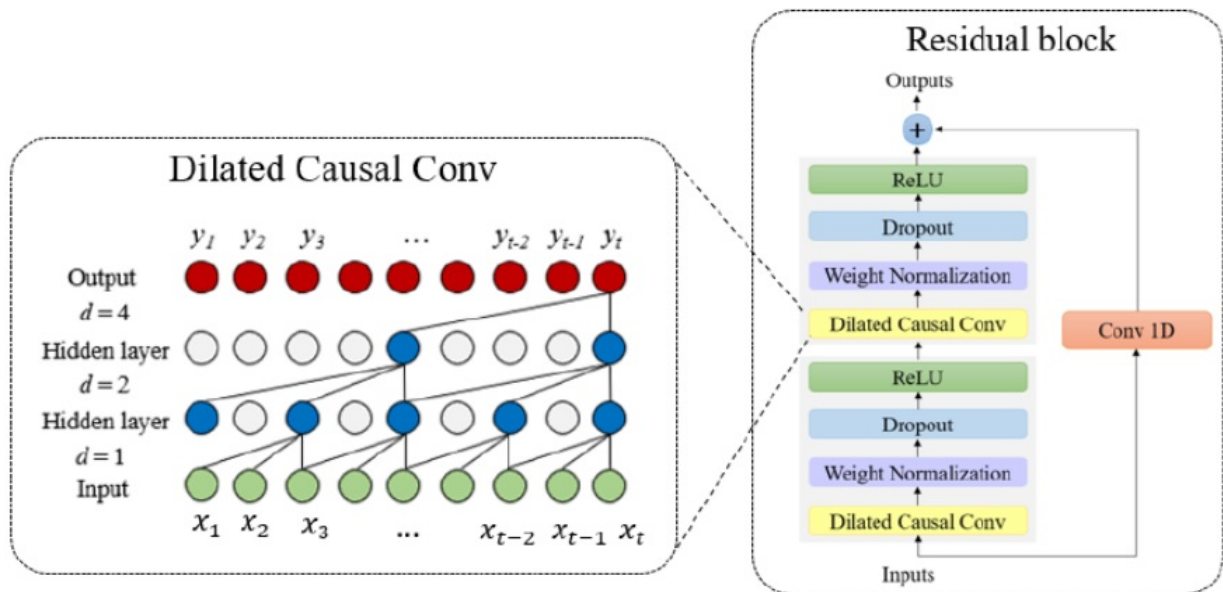


Figura 11 – Estrutura Interna de Rede Neural TCN.

Fonte: Enkhbat (2024).

Mais precisamente, uma sequência unidimensional de entrada $\{x\} \in R^n$ e um filtro $f : \{0, \dots, k - 1\} \rightarrow R^n$, o operador $\{F\}$ da convolução dilatada aplicado a sequência $\{s\}$ pode ser

definido como:

$$F(s) = \sum_{i=1}^{k-1} f(i) \cdot x_{s-d.i} \quad (2.7)$$

Onde, d é o fator de dilatação, k é o tamanho do filtro, $s-d.i$ explica a direção que a convolução será aplicada. Se d for igual a 1 a convolução dilatada se reduz a uma convolução regular (LIMA, 2024). Na TCN a profundidade da rede está relacionada ao campo receptivo, bem como ao tamanho do filtro e do fator de dilatação.

Para tornar a TCN mais do que um modelo de regressão linear excessivamente complexo, funções de ativação ReLU são adicionadas aos blocos residuais, após ambas as camadas convolucionais, para introduzir não linearidades (BAI; KOLTER; KOLTUN, 2018), (NAIR; HINTON, 2010).

Para normalizar a entrada de camadas ocultas, a normalização de peso é aplicada a cada camada convolucional para neutralizar o problema do gradiente explosivo, entre outras coisas.

Para evitar overfitting, a regularização é introduzida via dropout, após cada camada convolucional, em cada bloco residual (BAI; KOLTER; KOLTUN, 2018), conforme bloco a direita da figura 11.

2.5.1.3 Redes Neurais Recorrentes

Redes Neurais Recorrentes (RNNs) são extraordinariamente versáteis e foram elaboradas para lidar com dados sequenciais ou temporais (ELMAN, 1990). Ela é particularmente útil quando a ordem dos dados importa. Diferentemente das redes neurais convencionais, as RNNs possuem conexões retroalimentadas, o que lhes dá condições de manter um estado interno ou uma memória do passado ou seja possuem a capacidade de aprender relações dinâmicas envolvidas nos dados ao longo do tempo (GOODFELLOW; BENGIO; COURVILLE, 2016). Na figura 12, tem-se a estrutura de uma RNN, cuja as saídas da estrutura de aprendizado dos instantes anteriores $y_{(t-1)}$ servem de entrada nas próximas células junto com os dados de entrada $x(t)$ a cada instante de tempo, transmitindo informações de toda a cadeia de dados (dependência temporal) (BABIĆ; MARTINČIĆ-IPŠIĆ; MEŠTROVIĆ, 2020).

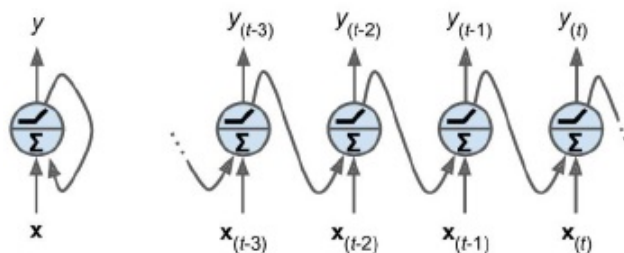


Figura 12 – célula Recorrente a cada instante de tempo.

Fonte: Babic (2020).

Nas RNNs, cada neurônio é composto por dois conjuntos de pesos: W_x para as entradas x_t e W_y para a saída do instante anterior $y_{(t-1)}$. A equação (2.8) representa a saída da rede no instante t . Sendo φ a função de ativação, W_x e W_y são matrizes de pesos das entradas e saídas da rede e b sendo o vetor com bias (SANTOS *et al.*, 2020).

$$y_t = \varphi(W_x^T x_t + W_y^T y_{t-1} + b) \quad (2.8)$$

$$h_t = \tanh(W_x x_t + (W_y h_{t-1} + b)) \quad (2.9)$$

Na estrutura de uma RNN a memória é uma propriedade intrínseca dos neurônios e de suas relações, onde cada neurônio processa informações vindas das amostras da entrada $x(t)$ e também, da estrutura de memória das saídas $h(t)$ das células adjacentes representada pela equação (2.9), em diferentes instantes de tempo (t) (GOODFELLOW; BENGIO; COURVILLE, 2016), (ELMAN, 1990), conforme figura 13.

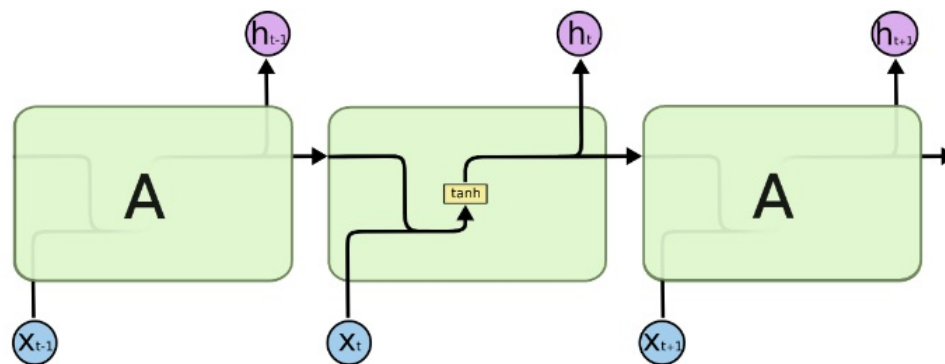


Figura 13 – Rede Recorrente a cada instante de tempo - RNN.

Fonte: 2022 Data Science Academy.

Assim como qualquer rede neural, as RNNs possuem a forma de uma cadeia de módulos repetidos e interconetados. Na RNN padrão, cada módulo de repetição tem uma estrutura simples, com uma única camada utilizando a função *tanh* (GOODFELLOW; BENGIO; COURVILLE, 2016).

2.5.1.3.1 Long Short-Term Memory (LSTM)

As redes LSTM foram introduzidas em 1997 (HOCHREITER; SCHMIDHUBER, 1997) para superar problemas de convergência dos pesos no treinamento das RNN convencionais, que exibem desaparecimento dos sinais de erro entre as unidades internas da rede. Isto comprometia a utilização dessas redes em problemas sequenciais de longa dependência temporal (CORREIA *et al.*, 2021). O projeto das LSTM foi estruturado para capturar estes padrões, particularmente em séries temporais (SAHOO *et al.*, 2019). Uma RNN tradicional funciona de tal forma que a ativação do estado de um neurônio é influenciado pelas outras ativações locais mais próximas, o que corresponde a uma memória de curto prazo, enquanto os pesos da rede são influenciados pelos cálculos que

ocorrem ao longo de todo o tempo, sequencialmente, que correspondem à memória de longo prazo (NOGUEIRA *et al.*, 2020). Deste modo, a LSTM surgiu como um redesenho da RNN para que se tenha um estado de ativação que também possa atuar como pesos e preservar informações em longas distâncias (CHAUHAN, 2020).

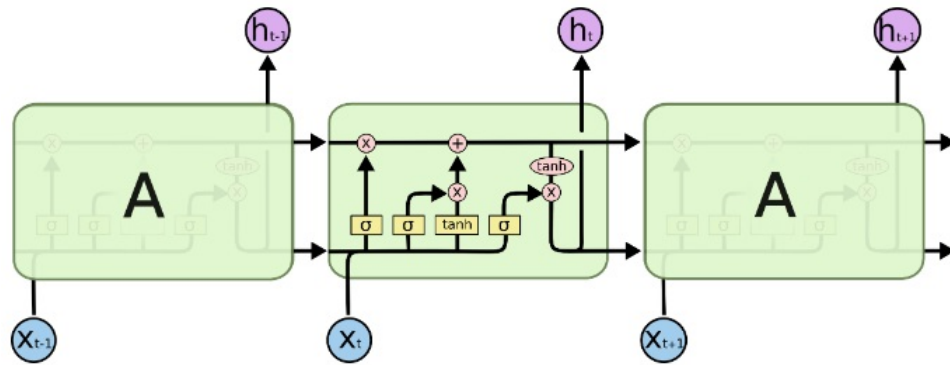


Figura 14 – Diagrama da Estrutura interna de uma Rede LSTM.

Fonte: 2022 Data Science Academy.

Na figura 14 é mostrada uma rede LSTM, cujo o valor da saída da célula atual é determinado pelas entradas anteriores, que influencia no resultado do estado atual e assim sucessivamente (HOCHREITER; SCHMIDHUBER, 1997). Nessa estrutura tem-se uma sequencia de observações sendo uma por vez, que pode avaliar quais informações vistas anteriormente pela rede podem ser importante para se fazer uma previsão.

Uma rede LSTM é composta por camadas onde cada célula é responsável pelo processamento de uma amostra sequencialmente. A Figura 15 apresenta uma célula LSTM, correspondente ao seu neurônio artificial (HOCHREITER; SCHMIDHUBER, 1997), (OLAH, 2015).

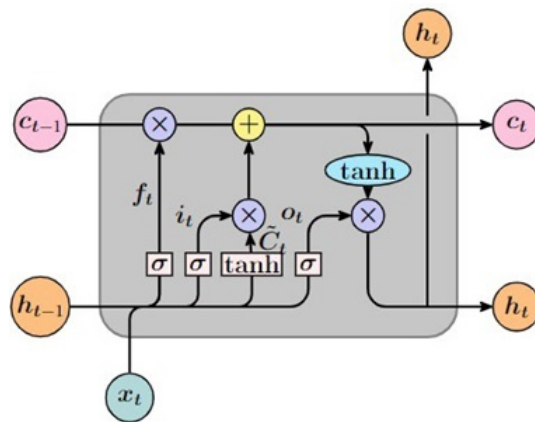


Figura 15 – Diagrama da estrutura interna de uma célula da rede LSTM.

Fonte: Fu (2022).

Esta unidade de processamento recebe a entrada genérica x_t além da saída h_{t-1} e do estado c_{t-1} da célula anterior, para atualizar seu próprio estado c_t e saída h_t . O estado representa a

memória interna da célula, que armazena informações de curto e de longo prazo (FU *et al.*, 2022). Segundo (FU *et al.*, 2022), essas informações são processadas por funções de ativação sigmoidais σ que geram os chamados sinais de *gate*, os quais ponderam a incorporação i_t , esquecimento f_t e transmissão o_t de informação na célula, sendo definidos respectivamente por:

$$i_t = \sigma(\mathbf{W}_{ix}x_t + \mathbf{W}_{ih}h_{t-1} + \mathbf{b}_i), \quad (2.10)$$

$$f_t = \sigma(\mathbf{W}_{fx}x_t + \mathbf{W}_{fh}h_{t-1} + \mathbf{b}_f) \text{ e} \quad (2.11)$$

$$o_t = \sigma(\mathbf{W}_{ox}x_t + \mathbf{W}_{oh}h_{t-1} + \mathbf{b}_o). \quad (2.12)$$

Neste caso as matrizes \mathbf{W} e os vetores de *bias* \mathbf{b} representam os pesos a serem aprendidos pela rede em cada função. Note que a atualização do estado da célula depende fator candidato à memória \tilde{C}_t , dado por:

$$\tilde{C}_t = \tanh(\mathbf{W}_{cx}x_t + \mathbf{W}_{ch}h_{t-1} + \mathbf{b}_c). \quad (2.13)$$

De onde se tem

$$c_t = f_t c_{t-1} + i_t \tilde{C}_t. \quad (2.14)$$

Enquanto a saída resulta

$$h_t = o_t \tanh(c_t). \quad (2.15)$$

2.5.1.3.2 Gated Recurrent Unit (GRU)

Em 2014 foi proposta por (CHO *et al.*, 2014) uma nova arquitetura de rede neural recorrente, denominada de *Gated Recurrent Unit (GRU)*, que foi introduzida como uma alternativa mais simples visando reduzir a complexidade em relação a arquitetura LSTM mantendo a eficiência e o desempenho com menor complexidade computacional. Outro aspecto positivo da GRU é a de permitir capturar dependências de forma adaptativa de grandes sequencias de dados sem descartar informações relevantes de partes anteriores da sequencia (AGGARWAL *et al.*, 2018).

Assim como a LSTM, a GRU busca resolver o problema da dissipação do gradiente de uma RNN tradicional com a utilização de apenas dois portões, o de reset gate (porta de redefinição) e o de update gate (porta de atualização) (ABID; COMPAGNON; LEFEBVRE, 2021), conforme mostra a figura 16. Esses portões, de fato, são vetores que decidem quais informações devem ser passadas para a saída, pois, contem valores entre 0 e 1, que serão multiplicados pelos dados de entrada e / ou do estado oculto. Se o valor do vetor for zero, por exemplo, as informações na entrada ou do estado oculto não são relevantes e retornarão como zero, ou se for 1, significa que todos os dados são relevantes e que serão utilizados (CHUNG *et al.*, 2014). Durante o treinamento da rede, através

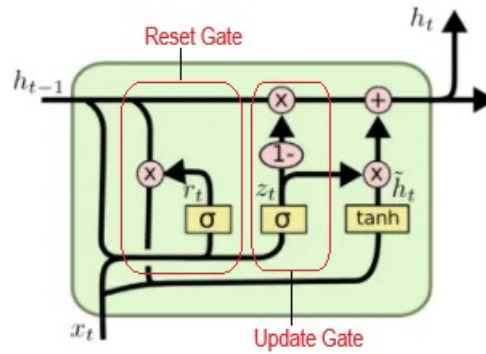


Figura 16 – Estrutura interna de uma célula GRU.

Fonte: Abid (2021).

do método *backpropagation*, os pesos são ajustados de maneira que os vetores aprendam a reter apenas as informações pertinentes.

Segundo (ABID; COMPAGNON; LEFEBVRE, 2021), a estrutura interna da GRU descrita acima, pode ser representada através das equações (ou do):

- Porta de Atualização ou *update gate* que define o quanto do estado anterior deve ser mantido:

$$z_t = \sigma(\mathbf{W}_z \cdot [h_{t-1}, \mathbf{x}_t]), \quad (2.16)$$

- Porta de Reinicilização ou *reset gate* que controla quanto da informação passada é esquecida:

$$r_t = \sigma(\mathbf{W}_r \cdot [h_{t-1}, \mathbf{x}_t]), \quad (2.17)$$

- Estado Oculto Atualizado que combina as portas com a nova entrada:

$$\tilde{\mathbf{h}}_t = \tanh(\mathbf{W} \cdot [r_t * h_{t-1}, \mathbf{x}_t]), \quad (2.18)$$

$$\mathbf{h}_t = (1 - z_t) * h_{t-1} + z_t * \tilde{\mathbf{h}}_t \quad (2.19)$$

Em resumo, a rede GRU é uma opção eficiente e eficaz para lidar com dados sequenciais, entregando um desempenho satisfatório em tarefas que requerem a aprendizagem com dependência de longo prazo. Em relação a LSTM, a simplicidade da GRU a torna mais rápida nos treinamentos e exige menos recursos computacionais (GREFF *et al.*, 2017).

MODELAGENS DE REDES NEURAIS ARTIFICIAIS

3.1 Metodologia

A estratégia desenvolvida neste trabalho adota uma abordagem típica de problemas de aprendizado baseado em dados, utilizando redes neurais tipo LSTM, GRU e TCN. Esta escolha justifica-se pela capacidade de tais redes lidarem com séries temporais multivariadas, não linearidades e dependências de longo prazo. A metodologia compreende os seguintes passos:

- (i) Levantamento, organização e pré processamento de dados;
- (ii) proposição da estrutura dos modelos;
- (iii) Escolha da técnica de Deep learning;
- (iv) codificação e treinamento dos modelos;
- (v) validação e testes dos modelos;
- (vi) Resultados

3.2 Levantamento, organização e pré-processamento de dados

A base de dados foi coletada do Site ONS ([Operador Nacional do Sistema Elétrico, 2025](#)). Os dados possuem amostragem horária de medições das variáveis de vazões e níveis de interesse do trabalho, abrangem um histórico de 14 anos de operação, totalizando aproximadamente 140.280 (Cento e

quarenta mil, duzentos e oitenta amostras). A qualidade dos dados é considerada boa, não há valores discrepantes ou ausentes que exijam de revisão.

Para o pré-processamento, os dados foram normalizados pela aplicação da média móvel dividida pelo valor máximo da variável. Esta técnica é fundamental para impedir que a rede processe dados de alta magnitude, o que geraria instabilidade nos pesos e dificultaria a convergência do treinamento. A Figura 17 exibe as séries de vazão afluyente, defluyente e nível de jusante normalizadas. O conjunto de dados foi dividido em: treinamento (2010-2021, 105.191 amostras), validação (2022-2024, 26.304 amostras) e teste (2025, 8.785 amostras). O conjunto reflete os perfis sazonais típicos dos reservatórios, com períodos distintos de estiagem e cheia.

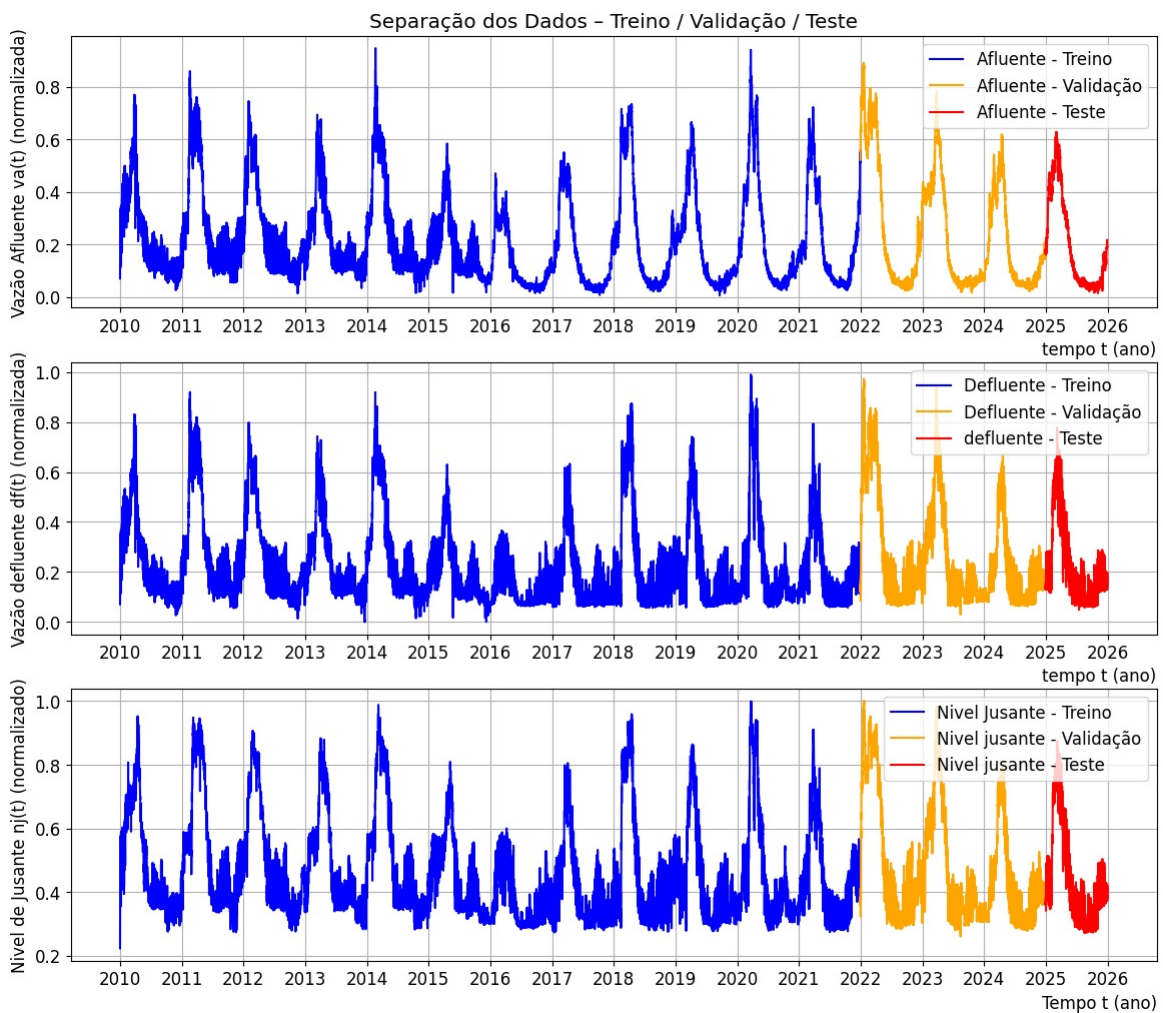


Figura 17 – Variáveis hidrológicas de interesse para treino, validação e teste dos modelos.

Fonte: Elaborada pelo autor.

Visto que as RNAs são algoritmos de aprendizagem supervisionada, os dados foram pré-processados para a geração de padrões entrada-saída, utilizados tanto no treinamento quanto na validação de cada modelo proposto. Estes padrões foram estruturados como uma matriz X_i ,

dada por:

$$\mathbf{X}_i = \begin{bmatrix} v_a(t) & v_a(t-1) & v_a(t-2) & \dots & v_a(t-n_{si}) \\ v_d(t) & v_d(t-1) & v_d(t-2) & \dots & v_d(t-n_{si}) \end{bmatrix} \quad (3.1)$$

e os de saída pelo vetor alvo $\mathbf{n}_j(t)$

$$\mathbf{n}_j(t) = [n_j(t+1) \quad n_j(t+2) \quad \dots \quad n_j(t+n_{so})], \quad (3.2)$$

aqui exemplificadas para o caso particular do modelo \mathcal{H}_3 , definida na equação (2.3). Note-se que n_{si} e n_{so} são hiperparâmetros associados ao máximo atraso das variáveis de entrada e a máxima janela de predição buscada pelo modelo respectivamente.

3.3 Proposição dos modelos

Nesta etapa serão apresentadas as três topologias selecionadas para as três redes TCN, LSTM e GRU, definidas pelas equações (2.1), (2.2) e (2.3), para representar o modelo dinâmico do nível de jusante $n_j(t)$. Os modelos considerados para os três cenários são:

- Modelo \mathcal{H}_1 : tendo como única entrada a vazão afluyente;
- Modelo \mathcal{H}_2 : tendo como única entrada a vazão defluente;
- Modelo \mathcal{H}_3 : tendo como entradas as vazões afluyente e defluente;

As topologias de modelos consideradas foram codificadas por meio da biblioteca Keras, utilizando a linguagem Python, nas quais foi possível avaliar hiperparâmetros principais, como o tipo de função de ativação, o otimizador e a função de custo. Além destes, foi necessário definir, em particular, o número de atrasos das variáveis. Esses dados estão sumarizados na Tabela 1, que inclui também o número de amostras de identificação e validação, o número de épocas para treinamento e o número de passos à frente. Os valores dos hiperparâmetros foram definidos por ajuste empírico, ao executar diferentes rotinas de treinamento e promovendo pequenos ajustes visando a melhoria dos resultados. Cita-se, como exemplo, o número de épocas, cujo erro, a partir de 200 épocas, se torna tão ínfimo que não há necessidade de aumentá-lo. Outro parâmetro, cuja escolha foi baseada em testes, é a janela de atrasos na entrada, que**, ** acima de 12 passos, traz resultados com pequenas melhorias e alto custo computacional, não agregando qualidade ao resultado esperado.

Ressalta-se que os modelos \mathcal{H}_1 e \mathcal{H}_2 possuem a mesma topologia, diferindo de \mathcal{H}_3 apenas pela dimensionalidade da entrada.

Para medir o ajuste da predição dos modelos aos dados reais de validação, avalia-se o quanto o valor de predição se aproxima do valor esperado, para treinamento, para validação, ou

Dados de treinamento (75%):	105.191 amostras (horas)
Dados de validação (19%):	26.304 amostras (horas)
Dados de teste (6%):	8.785 amostras (horas)
Função de custo:	MSE
Função de ativação:	ReLU
Otimizador:	Adam
Épocas	200
Batch-size	256
Atrasos na entrada, N_{si} :	12
Numero de passos a frente:	10

Tabela 1 – Dados comuns aplicados as Redes TCN, LSTM e GRU.

Fonte: Elaborada pelo autor.

teste, aplica-se a métrica do Erro Médio Quadrático MSE , definidor por:

$$MSE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2. \quad (3.3)$$

Para avaliar o grau de ajuste do modelo, normalizando o erro em relação à amplitude do sinal, ou seja, indicar qual percentual da dinâmica foi bem representado, aplica-se a métrica do Erro Quadrático Médio Normalizado $FitNRMSE$

$$FitNRMSE(y, \hat{y}) = \left(1 - \frac{\sqrt{\sum_{i=1}^n (y - \hat{y})^2}}{\sqrt{\sum_{i=1}^n (y - \bar{y})^2}}\right) \quad (3.4)$$

Para reproduzir a tendência e a dinâmica dos modelos e avaliar comportamento global do sistema, aplicou-se o coeficiente de determinação R^2 , definido por:

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y - \hat{y})^2}{\sum_{i=1}^n (y - \bar{y})^2} \quad (3.5)$$

Onde os valores genéricos y , \bar{y} e \hat{y} representam o valor real, médio e predito.

3.3.1 Modelo LSTM

"A rede criada expressa uma LSTM empilhada com duas camadas ocultas de vinte neurônios cada, função de ativação ReLU e uma camada densa com dez neurônios. A escolha final da quantidade de camadas ocultas e do número de neurônios foi resultado de várias etapas de treinamento, nas quais essas variáveis foram alteradas e testadas visando o melhor custo-benefício. Analisou-se o ganho de desempenho e o custo computacional versus o melhor resultado. A camada de entrada recebe dados no formato $3 \times N_{ni}$, enquanto a camada densa agrega a saída de todos os neurônios da segunda camada para gerar a saída da rede.

Além da escolha acima que definiu a melhor estrutura da rede para realização do treinamento, foram definidos também os hiperparâmetros, que em parte foram concebidos a partir de uma estratégia de tentativa e erro na busca de um modelo com desempenho satisfatório. Os

valores utilizados no modelo LSTM estão resumidos na tabela 1 e são os mesmos apresentados para os três cenários avaliados. Na tabela 2 esta representada a composição estrutural dos modelos \mathcal{H}_1 e \mathcal{H}_2 já treinados, sendo a única diferença para o modelo \mathcal{H}_3 a dimensionalidade dos dados de entrada, sendo para \mathcal{H}_1 e \mathcal{H}_2 (None, 12, 1) e para \mathcal{H}_3 (None, 12, 2).

Idx	Nome da camada	Tipo	Forma de saída	Ativação	Parâmetros
0	Input	InoutLayer	(None, 12, 1)	-	0
1	lstm_1	LSTM	(None, 12, 20)	relu	1840
2	lstm_2	LSTM	(None, 20)	relu	3280
3	Dense	Dense	(None, 10)	linear	210

Tabela 2 – Estrutura da rede neural dos Modelos LSTM \mathcal{H}_1 e \mathcal{H}_2

Fonte: Elaborada pelo autor.

3.3.2 Modelo GRU

A GRU apresenta estrutura semelhante à LSTM descrita no item anterior tanto no treinamento quanto na definição dos seus hiperparâmetros, ou seja, composta por três camadas: duas ocultas com vinte neurônios cada e uma densa com dez neurônios. Os valores utilizados estão resumidos na Tabela 1 e são os mesmos para os três cenários avaliados. A Tabela 3 representa a composição estrutural dos modelos \mathcal{H}_1 e \mathcal{H}_2 (também baseada em LSTM). Em relação ao \mathcal{H}_3 , a diferença está apenas na dimensionalidade: (None, 12, 1) para \mathcal{H}_1 e \mathcal{H}_2 , e (None, 12, 2) para \mathcal{H}_3 , respectivamente.

Idx	Nome da camada	Tipo	Forma de saída	Ativação	Parâmetros
0	Input	InputLayer	(None, 12, 1)	-	0
1	gru_1	GRU	(None, 12, 20)	relu	1440
2	gru_2	GRU	(None, 20)	relu	960
3	Dense	Dense	(None, 10)	linear	110

Tabela 3 – Estrutura da rede neural dos Modelos GRU \mathcal{H}_1 e \mathcal{H}_2

Fonte: Elaborada pelo autor.

3.3.3 Modelo TCN

A Rede TCN está como as demais, na forma empilhada, com duas camadas ocultas convolucionais temporais causais contendo: vinte filtros convolucionais por camada; *kernel-size* igual a 3, para capturar dependências em 3 passos de tempo; *padding* tipo "causal", para garantir que somente informações passadas sejam usadas e; 2 dilatações para expandir o campo receptivo, o que ajuda a capturar dependências de curto e médio prazo. Contem ainda, função de ativação ReLu e camada densa com dez neurônios. A camada densa agrega a saída de todos os neurônios da segunda camada, na geração da saída da rede. A seguir representa-se o modelo aplicado nos cenários \mathcal{H}_1 e \mathcal{H}_2 que se assemelham aos das redes LSTM e GRU, inclusive com a mesma observação para o cenário \mathcal{H}_3 .

Idx	Nome da camada	Tipo	Forma de saída	Ativação	Parâmetros
0	Input	InputLayer	(None, 12, 1)	-	0
1	tcn_1	TCN	(None, 12, 20)	relu	3860
2	tcn_2	TCN	(None, 20)	relu	4880
3	Dense	Dense	(None, 10)	linear	210

Tabela 4 – Estrutura da rede neural dos Modelos TCN \mathcal{H}_1 e \mathcal{H}_2

Fonte: Elaborada pelo autor.

Observando os dados das tabelas 2, 3 e 4 o principal ponto a ser destacado é a quantidade de parâmetro das redes LSTM, GRU e TCN que são bem diferentes, o que gera diferenças quanto ao tempo de treinamento e resultados que serão apresentados no próximo capítulo.

RESULTADOS

4.1 Treinamento e validação dos modelos H1, H2, H3

4.1.1 Função de Perda para as redes neurais LSTM, GRU e TCN

A utilização das Redes LSTM, GRU e TCN para a previsão do $n_j(t)$, mostrou resultados satisfatórios do ponto de vista computacional, apesar das complexas relações existentes entre os dados de entrada e saída dos sistemas. Os modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 propostos nesse trabalho apresentaram exímios resultados, quanto ao custo de treinamento e validação, com uma acentuada convergência via métrica MSE.

Nas figuras 18, 19 e 20 é possível visualizar custo de treinamento e validação em escala logarítmica para as redes aplicada aos três modelos.

4.1.1.1 Rede LSTM

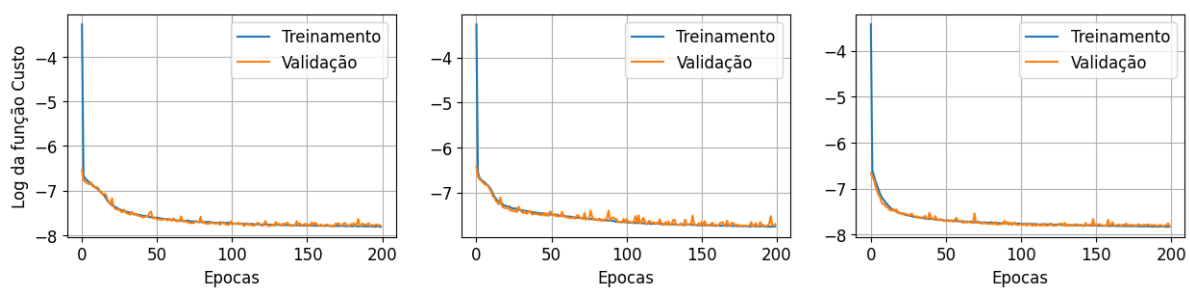


Figura 18 – Log função custo - Modelos LSTM \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 .

Fonte: Elaborada pelo autor.

Na função custo dos modelos da LSTM, mostrado na figura 18, observa-se que: em \mathcal{H}_1 há boa convergência, mas uma pequena oscilação entre a validação e treino; enquanto em \mathcal{H}_2 tem-se

curvas bem próximas e estáveis apresentando melhor equilíbrio entre treino e validação; já em \mathcal{H}_3 as curvas estão próximas, mas com alguns picos de instabilidade na validação.

4.1.1.2 Rede GRU

Para a função custo dos modelos da GRU, figura 19, tem-se: em \mathcal{H}_1 excelente convergência, treino e validação praticamente sobrepostas indicando ótima capacidade de generalização; \mathcal{H}_2 os resultados também são muito bons, mas ligeiramente menos estável que \mathcal{H}_1 ; enquanto \mathcal{H}_3 tem ótimo ajuste, porém há pequenas oscilações na validação foi menos suave que em \mathcal{H}_1 e \mathcal{H}_2 .

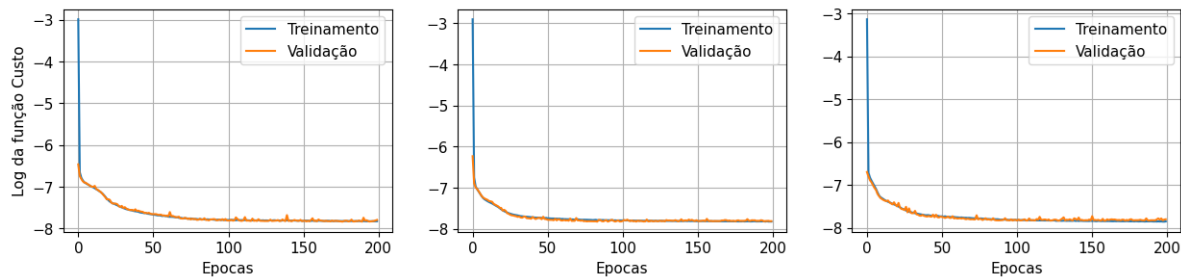


Figura 19 – Log da função custo dos Modelos GRU \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 .

Fonte: Elaborada pelo autor.

4.1.1.3 Rede TCN

Ainda sobre a função custo, agora para os modelos da TCN, figura 20, observa-se que: o treinamento converge bem até por volta da época 50 no modelo \mathcal{H}_1 , mas apresenta flutuações indicando alguma instabilidade, embora não haja sobreajuste forte; para \mathcal{H}_2 as curvas de treino e validação próximas e suaves; enquanto em \mathcal{H}_3 o treino converge bem, mas a validação oscila um pouco mais que em \mathcal{H}_2 , ainda assim, o desempenho é firme.

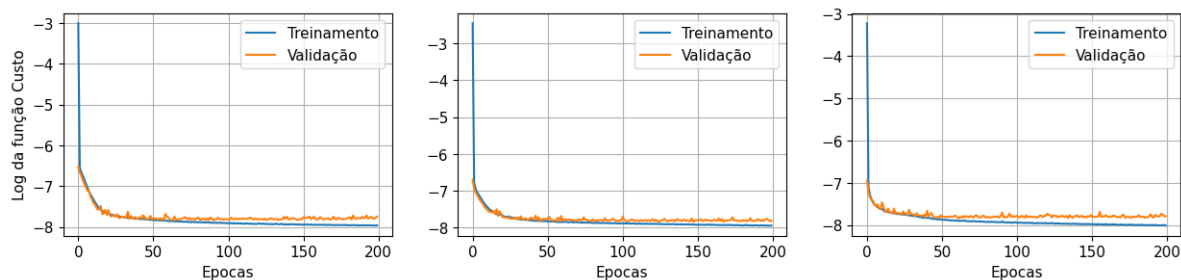


Figura 20 – Log da função custo dos Modelos TCN \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 .

Fonte: Elaborada pelo autor.

4.1.2 Desempenho dos modelos obtidos pelo FitNRMSE

A seguir apresenta-se os resultados obtidos pela métrica FitNRMSE para os modelos propostos para cada rede neural, aplicada na previsão do $n_j(t)$. Através das figuras 21, 22 e 23 observa-se o

quão próximo os resultados preditos estão dos valores reais ou de referência. Em outras palavras, é correto afirmar que a qualidade percentual das previsões expressam o quanto os modelos reproduzem a dinâmica do sinal real.

Em resumo os três modelos simulados pelas redes LSTM, GRU e TCN exibem resultados satisfatórios de predição e muito próximos entre si.

4.1.2.1 Rede LSTM

Na figura 21 tem-se a precisão dos valores previstos em relação aos reais obtidos pela rede LSTM. Observa-se um bom desempenho nas primeiras amostras, convergindo para uma faixa de 84,6 a 85,3%. Analisando todas as previsões, o melhor cenário se dá para \mathcal{H}_3 . Como ponto forte da LSTM destaca-se uma boa estabilidade e consistência mantendo seu desempenho alto em todo o horizonte.

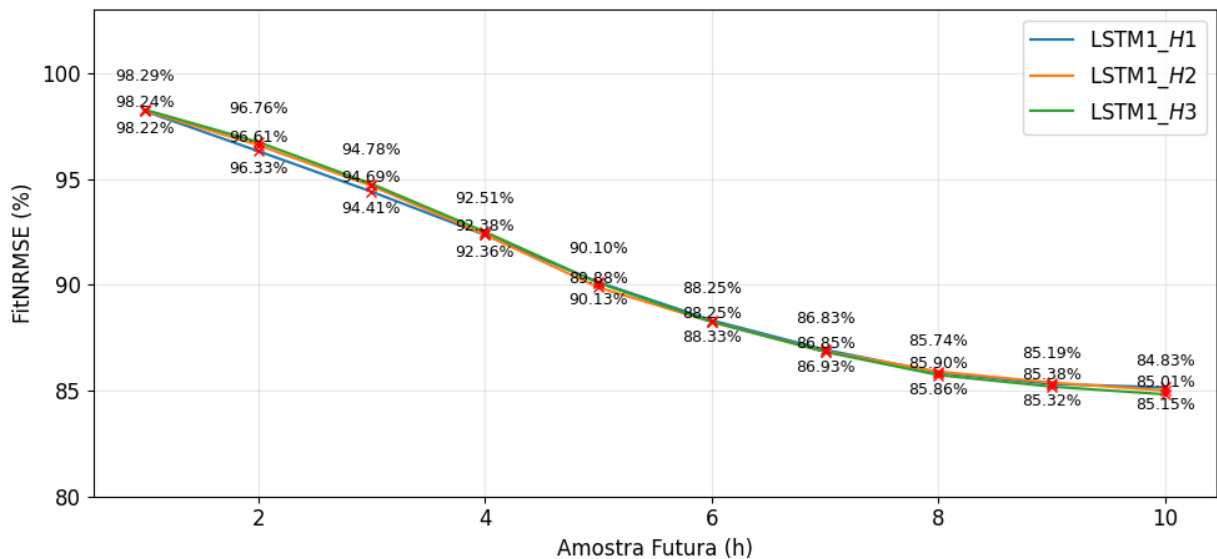


Figura 21 – FitNRMSE para os Modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 - Rede LSTM.

Fonte: Elaborada pelo autor.

4.1.2.2 Rede GRU

Para a GRU se observa um início ligeiramente superior para o modelo \mathcal{H}_2 , chegando a 98,5% e valores finais entre 85,22 e 85,33%. Como melhores cenários destaca-se a primeira amostra de \mathcal{H}_2 (98,5%) última amostra para \mathcal{H}_2 (85,12%). Como ponto forte destaca-se um melhor equilíbrio nas primeiras e últimas amostras.

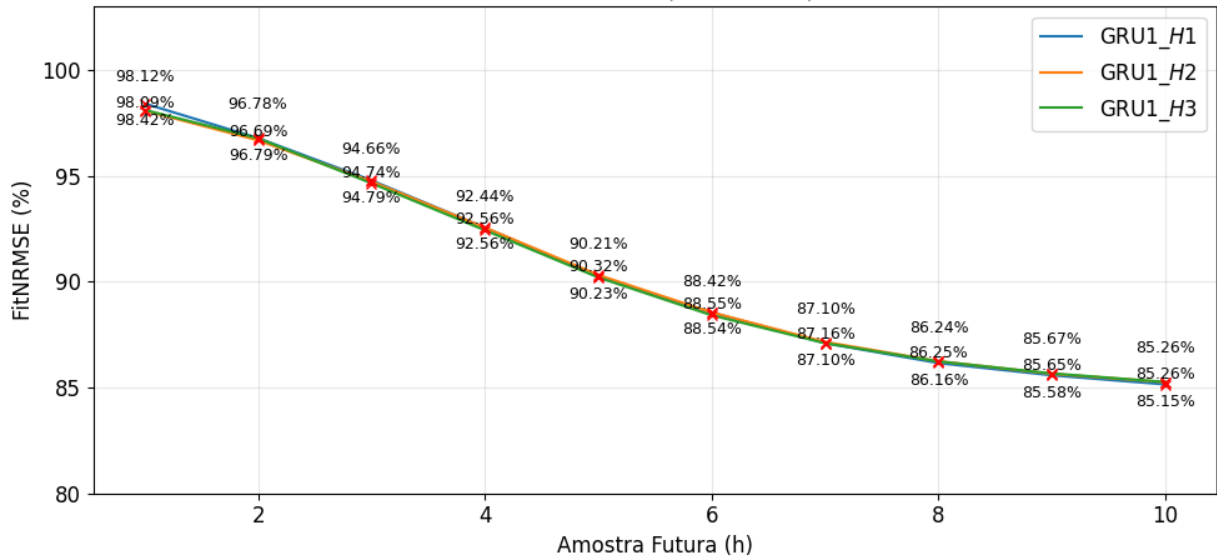


Figura 22 – FitNRMSE para os Modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 - Rede GRU.

Fonte: Elaborada pelo autor.

4.1.2.3 Rede TCN

Já para a TCN, os valores são ligeiramente inferiores a LSTM e GRU, mas ainda com bom ajuste. Como ponto forte destaca-se um desempenho sólido com melhor cenário para a última amostra de \mathcal{H}_3 (85,12%) e para a primeira amostra de \mathcal{H}_1 (97,3%), mas, menos competitivo nos extremos em comparação com a LSTM e GRU.

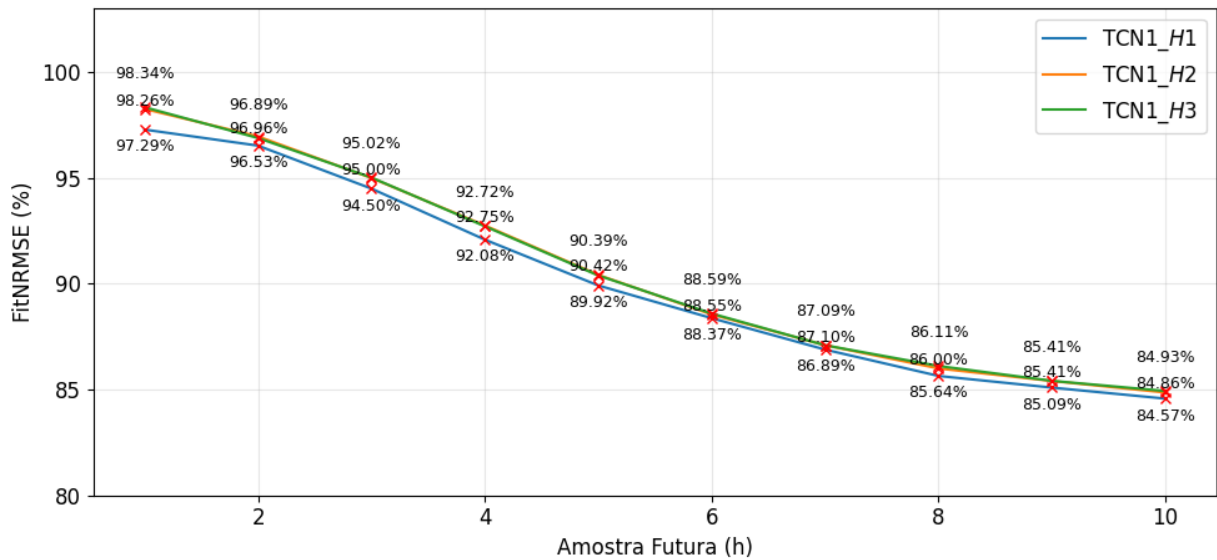


Figura 23 – FitNRMSE para os Modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 - Rede TCN.

Fonte: Elaborada pelo autor.

4.1.3 Desempenho dos modelos por $FitR^2$

Nas figuras 24, 25 e 26 expressa-se via métrica R^2 , a qualidade dos resultados de cada rede e modelos associados. Observa-se com clareza que os três modelos tem a mesma tendência, iniciando em torno de 99,97%, caindo gradativamente até próximo a 97,5%, com variabilidade bem pequena, o que denota que os hiperparâmetros das redes capturaram bem a dinâmica da série real, indicando o quão bem os modelos se ajustam aos dados observados. Tais resultados apontam que as redes foram bem treinadas e que os modelos conseguiram aprender bem tanto a tendência global quanto as pequenas variações locais.

4.1.3.1 Rede LSTM

O desempenho da rede LSTM tem início excelente, maior que 99,7%, decaindo progressivamente até 97,6%, o que mostra uma leve queda de precisão. Todos os testes mostram o mesmo padrão com variação mínima entre rodadas. A consistência entre os modelos mostra que a rede é estável e confiável, não dependendo do acaso do treinamento.

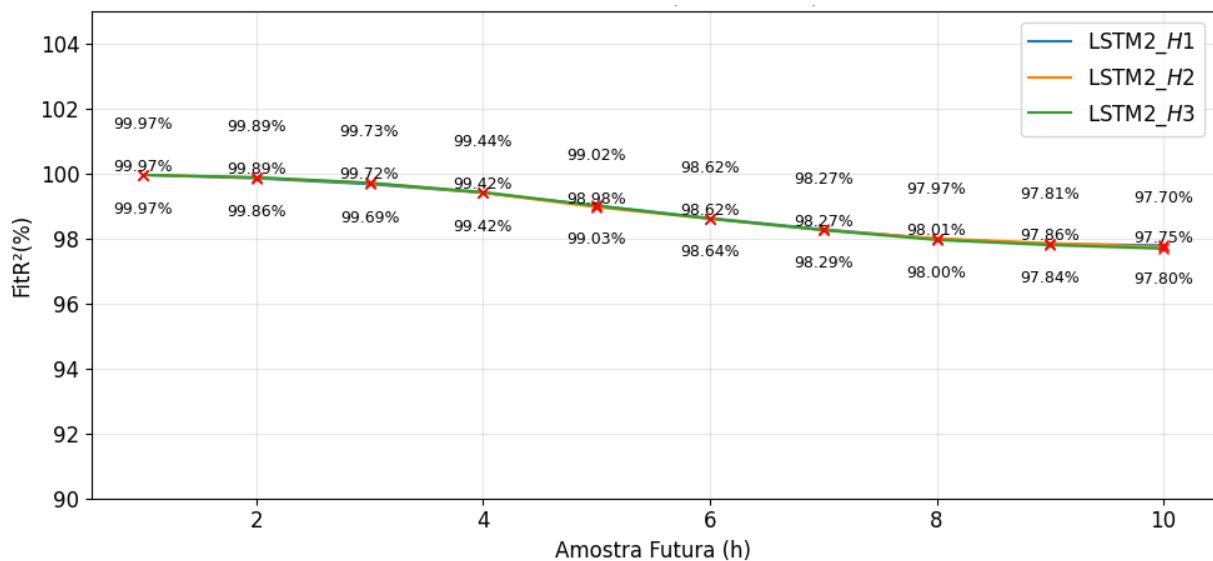


Figura 24 – $FitR^2$ da rede LSTM para os Modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 .

Fonte: Elaborada pelo autor.

4.1.3.2 Rede GRU

A GRU apresenta a maior precisão média, melhor R^2 em relação a LSTM e TCN, principalmente nas últimas amostras onde apresenta menor degradação.

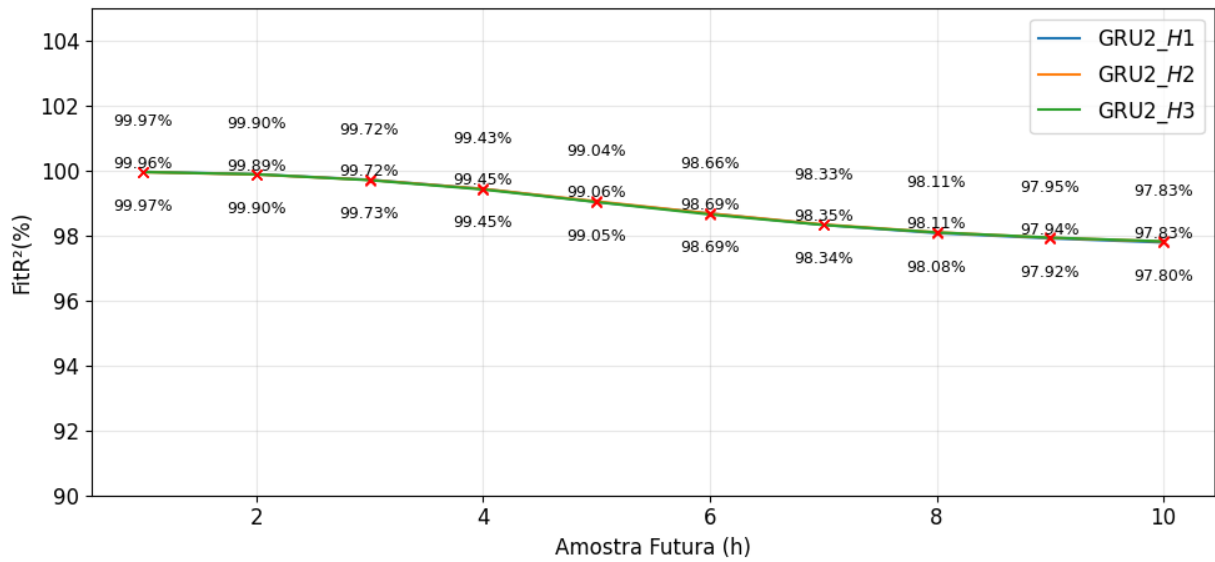


Figura 25 – FitR² da rede GRU para os Modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 .

Fonte: Elaborada pelo autor.

4.1.3.3 Rede TCN

A TCN se mostra mais competitiva nas amostras intermediárias, mas finaliza ligeiramente abaixo dos outros.

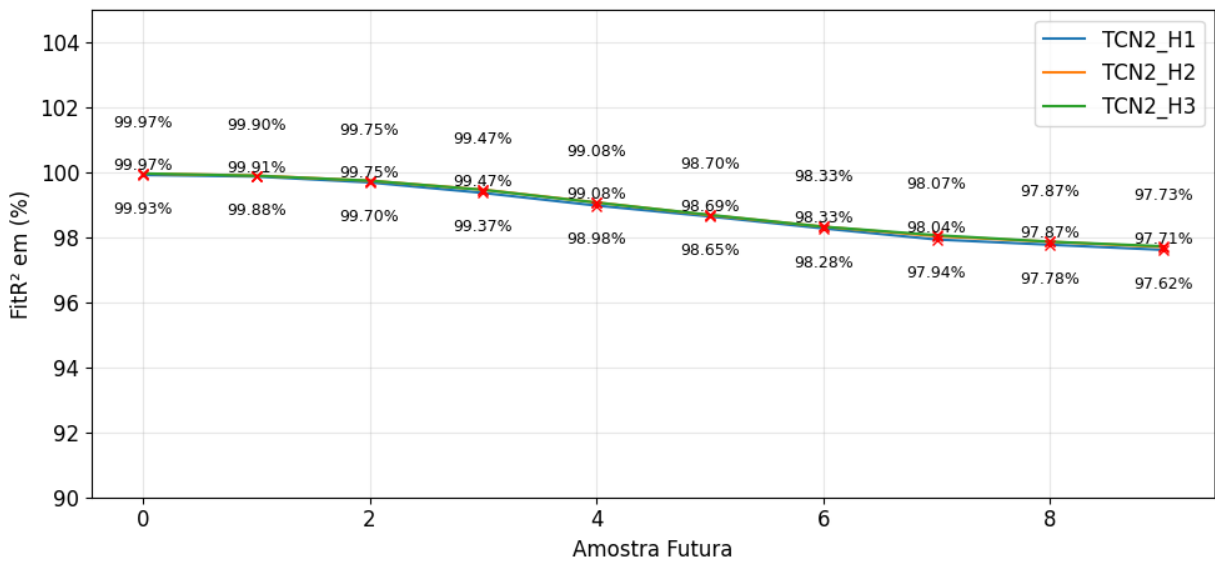


Figura 26 – FitR² da rede TCN para os Modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 .

Fonte: Elaborada pelo autor.

Pelos resultados que se observa nas três redes representadas pelas figuras 24, 25 e 26, pode-se concluir que a GRU entrega o melhor R^2 médio com pequena vantagem no meio e no final. Enquanto a LSTM fica bem próxima da GRU, performando levemente abaixo no final da

previsão. Já a TCN tem R^2 global praticamente empatado com a LSTM, perdendo um pouco no início e no final da serie de previsão.

4.1.4 Amostras detalhadas dos modelos

De maneira geral os modelos apresentam qualidade excelente no 1º passo, praticamente sem erro visível, boa qualidade até o 5º passo, quando começam a surgir pequenas diferenças, mas mantendo a estrutura da série. No 10º passo maior suavização e defasagem entre validação e as predições, mas ainda com boa aderência ao sinal.

4.1.4.1 Rede LSTM

A figura 27 mostra as curvas de predição dos modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 preditas pela rede LSTM. Para a 1ª amostra, observa-se que elas estão praticamente sobrepostas à validação, mostrando que a rede LSTM consegue capturar muito bem a dinâmica imediata da serie a curto prazo. Os erros são mínimos e não há desvios sistemáticos com valores de R^2 . Na 5ª amostra já aparece uma pequena dispersão entre as predições \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 , ainda assim, a correlação é alta e o modelo mantém boa qualidade. Já na 10ª amostra o afastamento é mais visível onde a curva de validação (azul) apresenta picos e vales mais perceptíveis enquanto as predições ficam mais suavizadas.

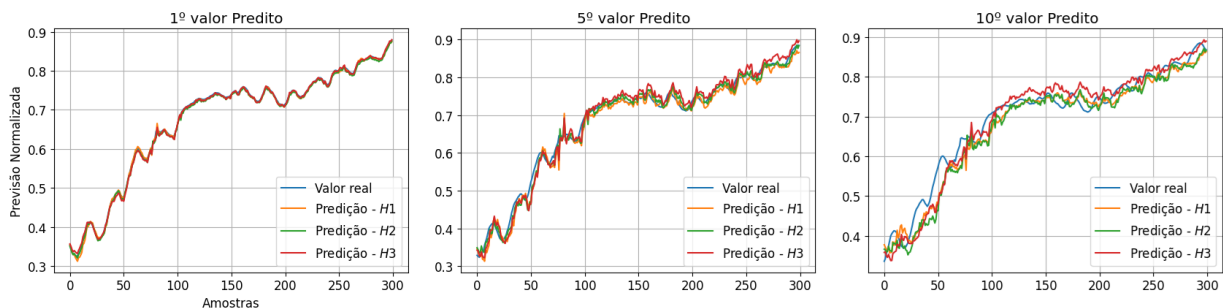


Figura 27 – Gráfico detalhado LSTM 1ª, 5ª e 10ª amostras de \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3

Fonte: Elaborada pelo autor.

4.1.4.2 Rede GRU

Para GRU os resultados são bem similares a LSTM, conforme se observa na figura 28 cuja as curvas de predição \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 estão, também, praticamente sobrepostas à curva de validação para a 1ª amostra (curto prazo). A GRU captura muito bem a dinâmica da série e as três variantes do modelo apresentam comportamento quase idêntico. Na 5ª amostra, de médio prazo, as predições ainda seguem perto da validação, mas já apresentam pequenos desvios, com diferenças bem discretas entre \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 , mas já se destacam quando comparadas com as da 1ª amostra. Com relação a 10ª amostra (longo prazo), a curva de validação (azul) apresenta variações mais acentuadas, enquanto as predições ficam ligeiramente suavizadas.

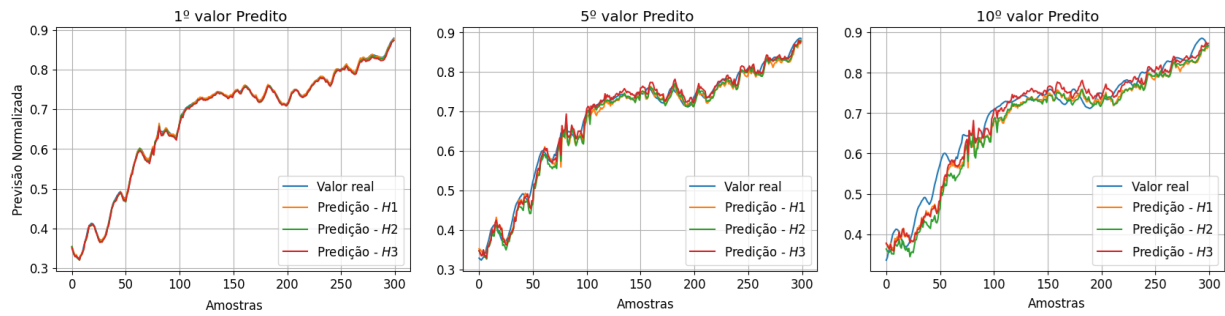


Figura 28 – Gráfico detalhado GRU 1ª, 5ª e 10ª amostras de \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3

Fonte: Elaborada pelo autor.

4.1.4.3 Rede TCN

Para a TCN pode-se dizer que o resultado segue a tendência da LSTM e GRU para a 1ª amostra. Para a 5ª amostra, de médio prazo, mantém-se a comparação de resultado, com destaque para as três predições cujas curvas continuam bem próximas entre si, mostrando consistência e estabilidade do modelo. Para a 10ª amostra, destaca-se que nesta de mais longo prazo surgem maiores divergências, principalmente em \mathcal{H}_2 em que os desvios são mais visíveis. Ainda assim, as três predições conseguem capturar a tendência geral.

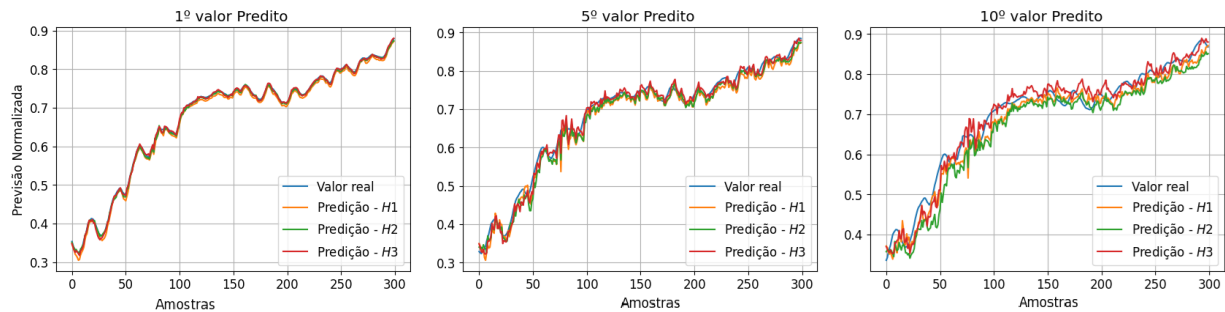


Figura 29 – Gráfico detalhado TCN 1ª, 5ª e 10ª amostras de \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3

Fonte: Elaborada pelo autor.

Em resumo, destaca-se que: \mathcal{H}_1 e \mathcal{H}_3 são os modelos de maior qualidade em todas as arquiteturas; \mathcal{H}_2 é sistematicamente o mais fraco, principalmente em horizontes maiores; a TCN se destacou em previsões de curto e médio prazo, mostrando maior aderência à validação e menor suavização em relação a LSTM e GRU e; para finalizar, conclui-se que a LSTM e GRU têm desempenho semelhante, com ligeira vantagem para a arquitetura GRU pela maior estabilidade em horizontes longos.

4.1.5 Níveis de Jusante Previstos e Desnormalizados

As figuras 30, 31 e 32, mostram com detalhes o comportamento dos valores de níveis previstos, após validação dos modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 , treinados pelas redes LSTM, GRU e TCN.

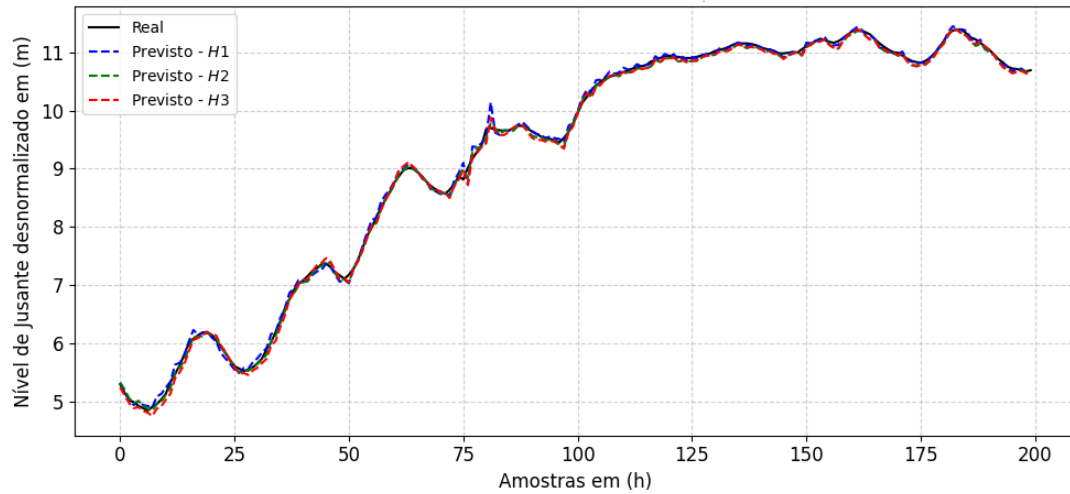


Figura 30 – Níveis Previstos Desnormalizados dos modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 - LSTM
 Fonte: Elaborada pelo autor.

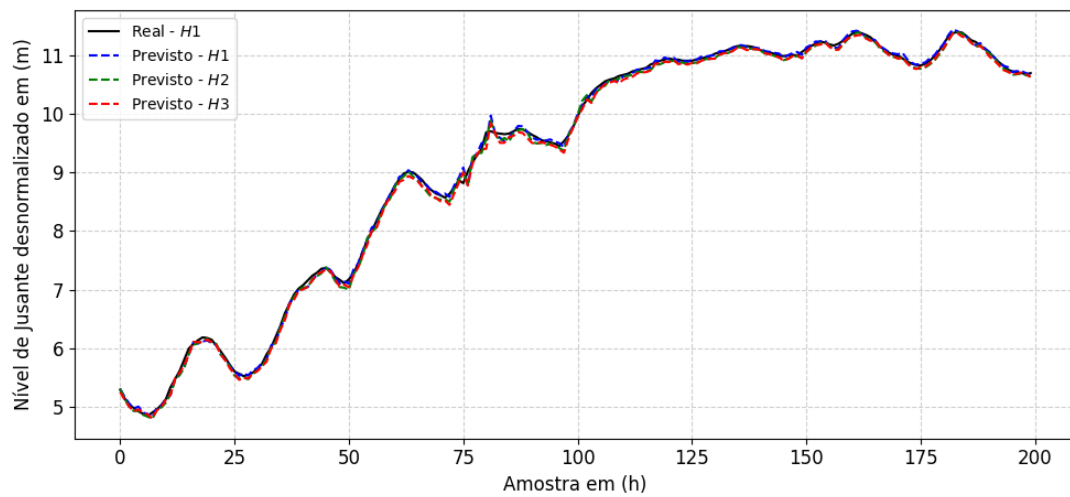


Figura 31 – Níveis Previstos Desnormalizados dos modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 - GRU
 Fonte: Elaborada pelo autor.

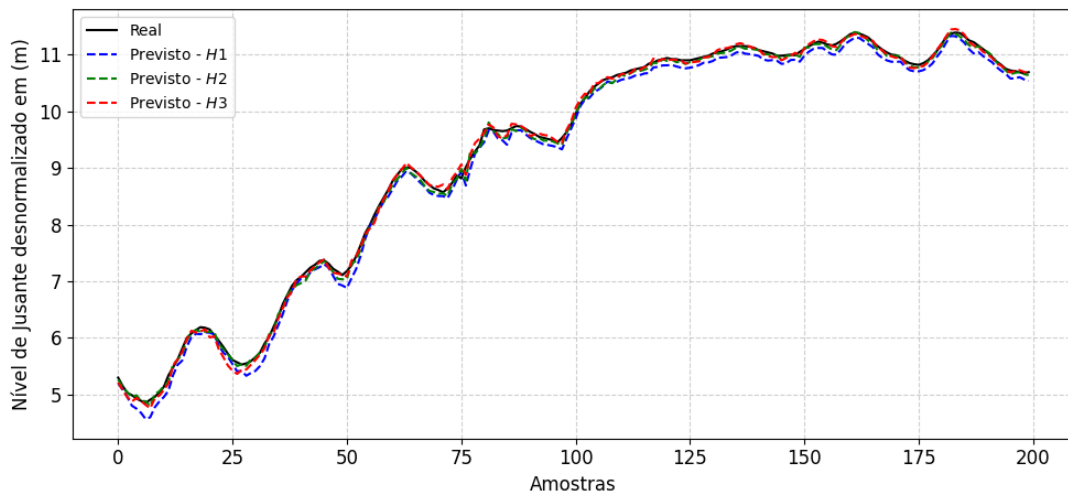


Figura 32 – Níveis Previstos Desnormalizados dos modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 - TCN

Fonte: Elaborada pelo autor.

Para fins de análise visual e interpretação física, as séries previstas foram desnormalizadas utilizando os parâmetros do conjunto original de dados e dos valores preditos pelos modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 .

Ao fazer uma análise gráfica comparativa entre valores reais e previstos se evidencia uma boa capacidade dos modelos em reproduzir a dinâmica do nível de jusante, o que permite avaliar de forma direta a aderência temporal e a suavização das previsões. Embora o modelo H1 apresente desempenho satisfatório, o modelo H3, incorpora simultaneamente múltiplas variáveis de entrada, apresenta maior robustez e menor acúmulo de erro, preservando a coerência temporal das previsões, logo será o modelo a ser aplicado ao conjunto de teste a seguir.

4.2 Teste do modelo H3 adotado

4.2.1 Normalizado

Os resultados normalizados de teste obtidos para o modelo H3, estão apresentados nas figuras 33, 34 e 35 e indicam que as três arquiteturas avaliadas LSTM, GRU e TCN foram capazes de reproduzir adequadamente a dinâmica do sinal real, apresentando baixos valores de erro absoluto e elevada capacidade explicativa.

De forma a complementar as análises dos níveis de jusante testados e normalizados apresentados pelas figuras 33, 34 e 35 são apresentados a seguir: O valor real x valor previsto; análise dos resíduos (erro temporal) e; histograma dos resíduos.

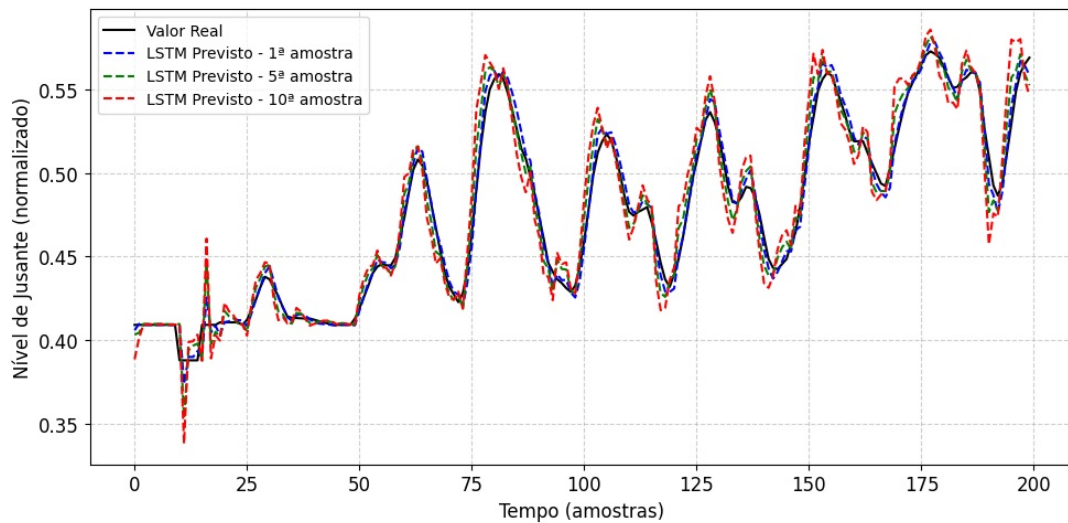


Figura 33 – Níveis Previstos, Normalizados e Testados por LSTM - Modelo \mathcal{H}_3 .

Fonte: Elaborada pelo autor.

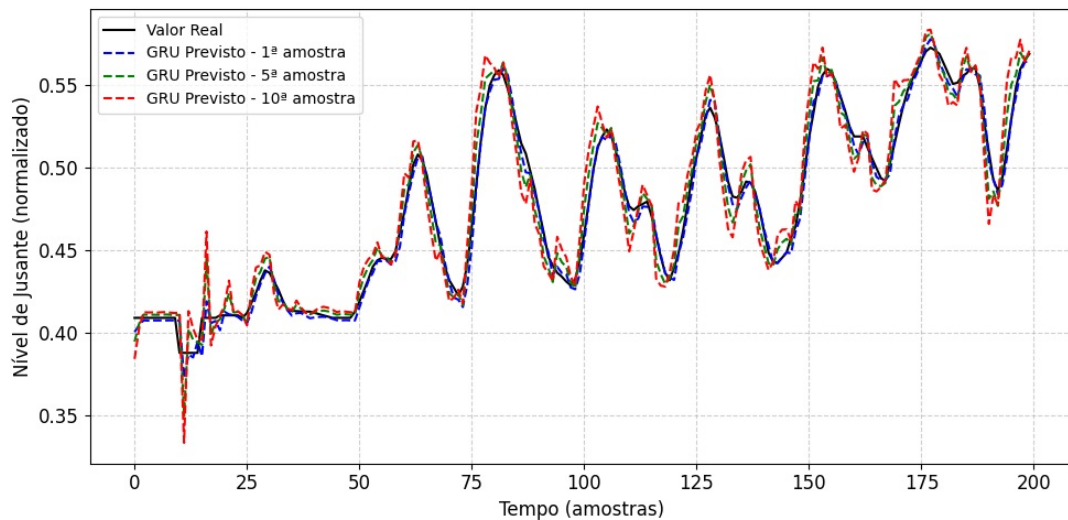


Figura 34 – Níveis Previstos, Normalizados e Testados por GRU - Modelo \mathcal{H}_3 .

Fonte: Elaborada pelo autor.

4.2.1.1 Valor Real x Valor Previsto

Através da figura 36, a análise visual da comparação entre o sinal real e as predições evidencia que as redes LSTM, GRU e TCN conseguem acompanhar o comportamento temporal do sistema, capturando as tendências de longo prazo. Contudo, observa-se que a GRU apresenta melhor alinhamento global entre os valores previstos e reais e menor dispersão ao longo da série temporal. A LSTM apresenta desempenho semelhante a GRU, mas com pequenas discrepâncias em algumas regiões, enquanto a TCN tende a apresentar desempenho ligeiramente inferior a LSTM e GRU.

Essas observações qualitativas são coerentes com os valores de FitR^2 , que se mantêm elevados para todas as redes, indicando que mais de 97% da variância do sinal real é explicada

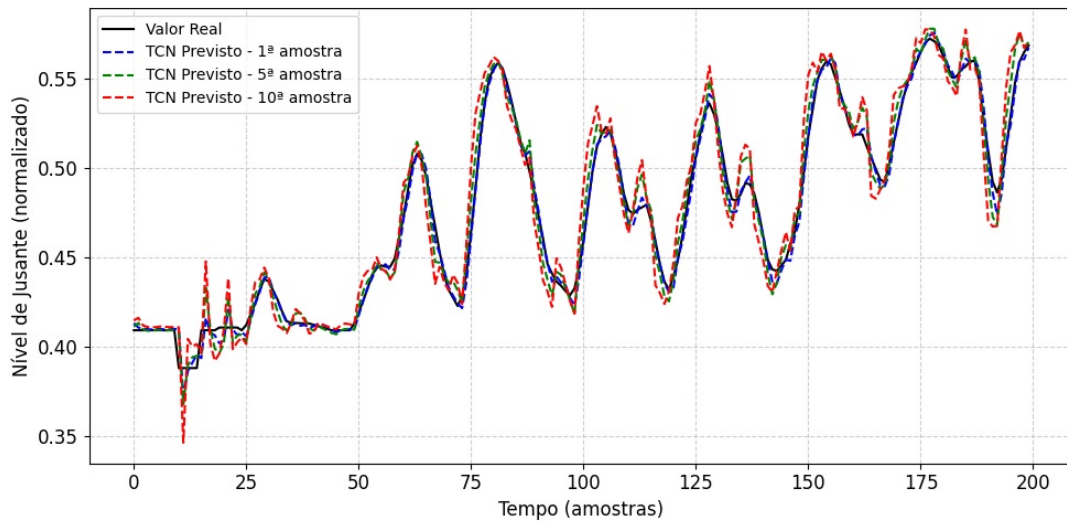


Figura 35 – Níveis Previstos, Normalizados e Testados por TCN - Modelo \mathcal{H}_3 .

Fonte: Elaborada pelo autor.

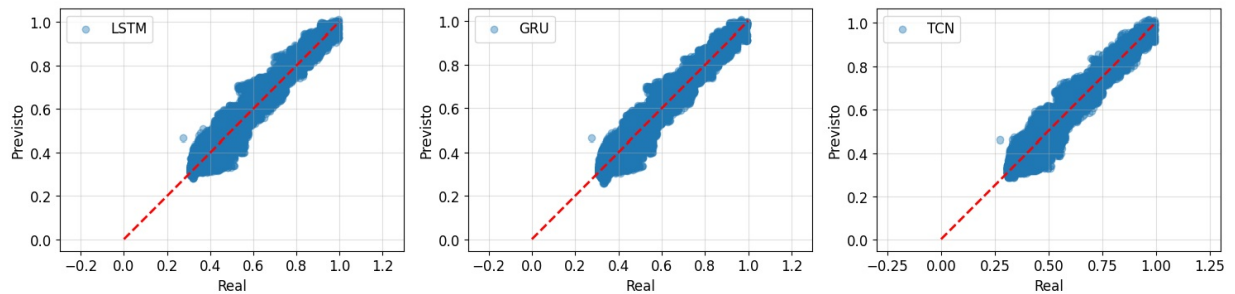


Figura 36 – Valor Real x Previsto Testados por LSTM, GRU e TCN - Modelo \mathcal{H}_3 .

Fonte: Elaborada pelo autor.

pelos modelos.

4.2.1.2 Análise dos resíduos (erro temporal)

A figura 37 mostra que, para as três redes, a análise dos resíduos apresenta erros centrados próximos de zero, sem a indicação de tendências sistemáticas ou padrões persistentes, o que indica ausência de viés significativo nas previsões. Como em outras análises a GRU apresenta resíduos ligeiramente menores em magnitude e com menor variabilidade, refletindo menor erro. A LSTM apresenta resíduos de amplitude um pouco maior, enquanto a TCN apresenta comportamento intermediário.

A ausência de autocorrelação nos resíduos sugere que os modelos capturam de forma pertinente a dinâmica do sistema.

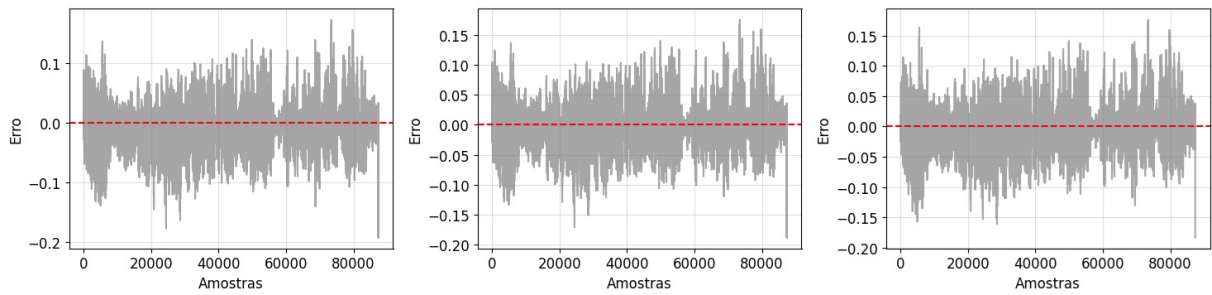


Figura 37 – Análise dos Resíduos dos Valores Previsto e Testados por LSTM, GRU e TCN - Modelo \mathcal{H}_3 .

Fonte: Elaborada pelo autor.

4.2.1.3 Histograma dos Resíduos

Os histogramas dos resíduos reforçam as análises anteriores, mostrando distribuições aproximadamente simétricas e centradas em torno de zero para todas as redes. Conforme figura 38, a distribuição associada à GRU é ligeiramente mais concentrada, indicando maior consistência das predições, conseqüentemente, menor dispersão. A LSTM apresenta uma distribuição um pouco mais espalhada, enquanto a TCN apresenta distribuição intermediária.

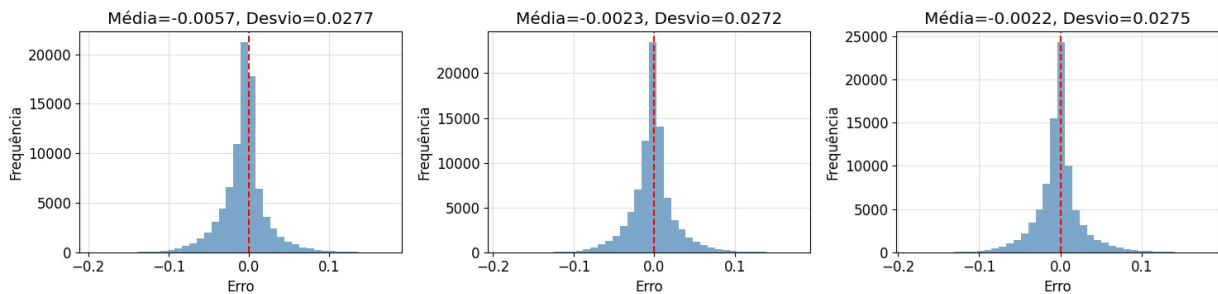


Figura 38 – Histograma dos Resíduos dos valores Previstos Testados por LSTM, GRU e TCN - Modelo \mathcal{H}_3 .

Fonte: Elaborada pelo autor.

De maneira geral, os resultados dos testes com o modelo \mathcal{H}_3 demonstram ser capazes de fornecer predições precisas e robustas independentemente da arquitetura utilizada, o que evidencia a importância da utilização de múltiplas variáveis de entrada para a modelagem do sistema

4.2.2 Desnormalizados

Com base nos gráficos das figuras 39, 40 e 41 dos valores desnormalizados do nível de jusante (Real \times Previsto) para as 1^a, 5^a e 10^a amostras implementados com as arquiteturas LSTM, GRU e TCN, é possível realizar uma análise qualitativa e comparativa de desempenho, com foco em

aderência temporal e degradação ao longo do horizonte de previsão, seguindo as tendências das predições demonstradas pelas validações da etapa dois deste estudo.

Nas três amostras do modelo \mathcal{H}_3 das arquiteturas LSTM, GRU e TCN, observa-se que os valores previstos preservam a dinâmica média do nível de jusante, não apresentando instabilidades numéricas após desnormalização, consequentemente, os erros permanecem fisicamente plausíveis. Contudo, há diferenças claras entre arquiteturas e horizontes.

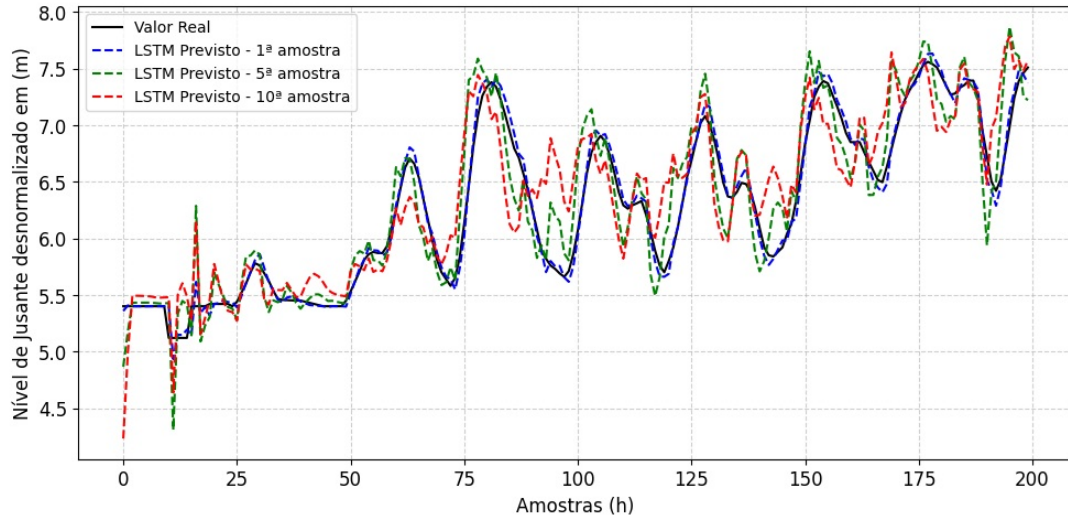


Figura 39 – Níveis de Jusante Desnormalizados dos Testes Preditos pelo Modelo \mathcal{H}_3 - LSTM.

Fonte: Elaborada pelo autor.

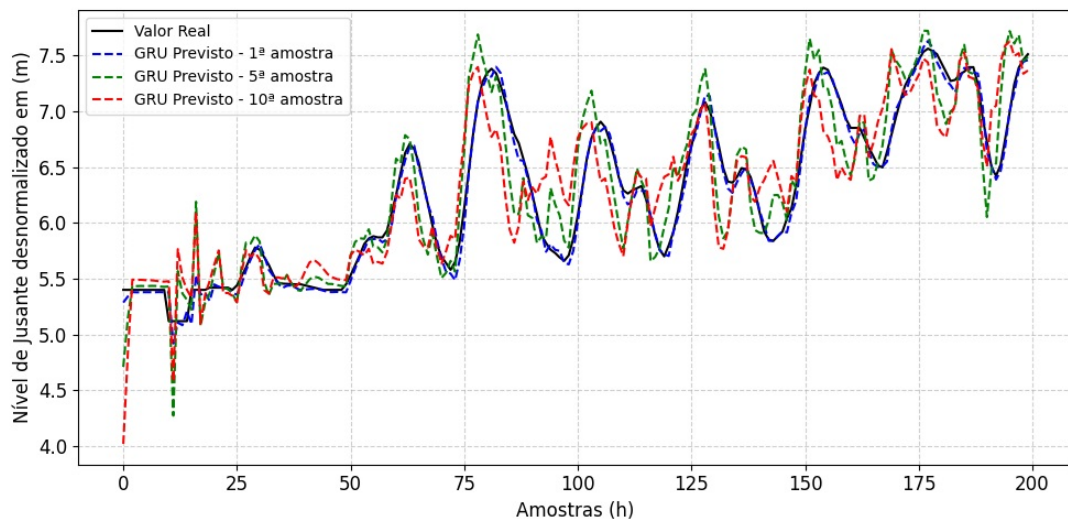


Figura 40 – Níveis de Jusante Desnormalizados - Testes Preditos pelo Modelo \mathcal{H}_3 - GRU.

Fonte: Elaborada pelo autor.

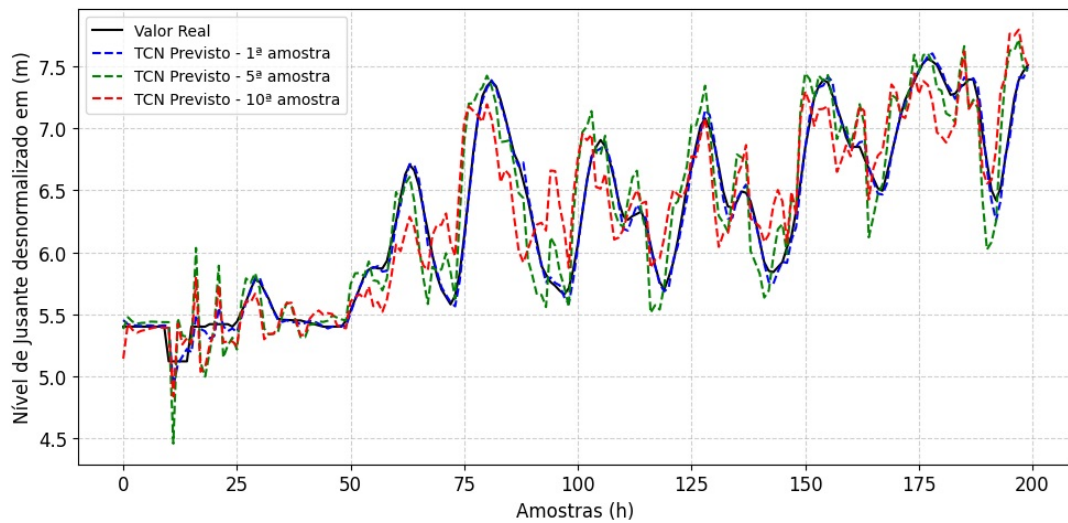


Figura 41 – Níveis de Jusante Desnormalizados - Testes Preditos pelo Modelo \mathcal{H}_3 - TCN.

Fonte: Elaborada pelo autor.

Numa análise mais global por horizonte de predição (1ª, 5ª e 10ª) tem-se:

- 1ª amostra: Sobreposição clara entre os sinais real e previsto, com erros predominantemente pequenos e com variações suaves.
- 5ª amostra: Começa surgir pequena defasagem temporal e aumento gradual do erro acumulado.
- 10ª amostra: Maior afastamento em relação ao valor real e suavização acentuada.

4.2.3 Alertas Reais

O modelo de previsão de alerta utiliza abordagem multivariada com entrada composta por vazão afluente e vazão defluente e como saída o nível de jusante, empregando arquitetura ensemble baseada em LSTM, GRU e TCN, com horizonte de previsão de 10 passos à frente a partir de 12 observações passadas. Para esta análise, tem-se nas figuras 42 e 43 os registros de eventos reais ocorridos no dia 10/03/2025, cuja previsão indica valores acima do nível de referência de 12,5 metros a jusante da usina.

Numa análise comparativa entre as previsões iniciadas às 11:00 e 12:00 permite avaliar a estabilidade dinâmica do modelo, onde Observou-se que:

- As curvas previstas apresentam elevada correlação estrutural.
- O instante estimado de ultrapassagem do nível de alerta permaneceu consistente.
- A amplitude prevista não sofreu variações abruptas entre janelas.

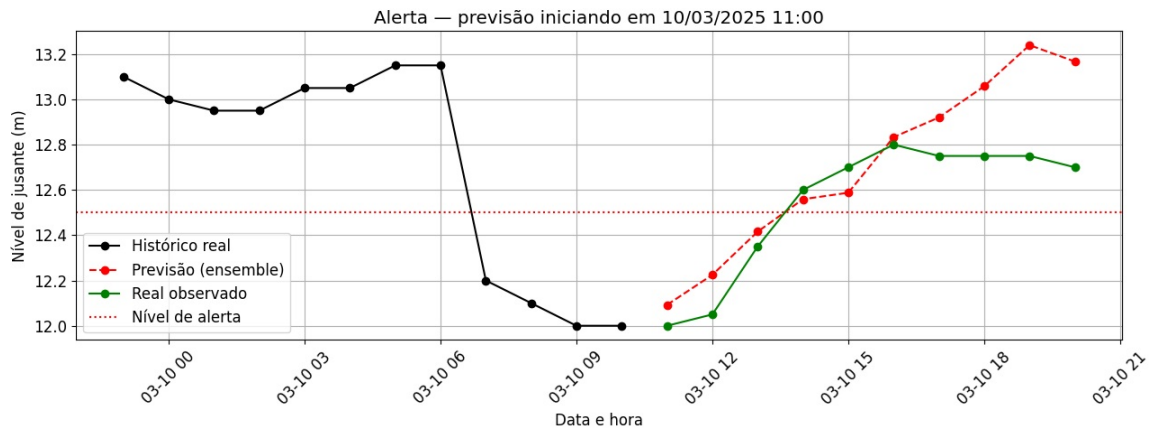


Figura 42 – Previsão x Nível de Alerta - Evento Real em 10.03.2025, 11 horas AM.
 Fonte: Elaborada pelo autor.

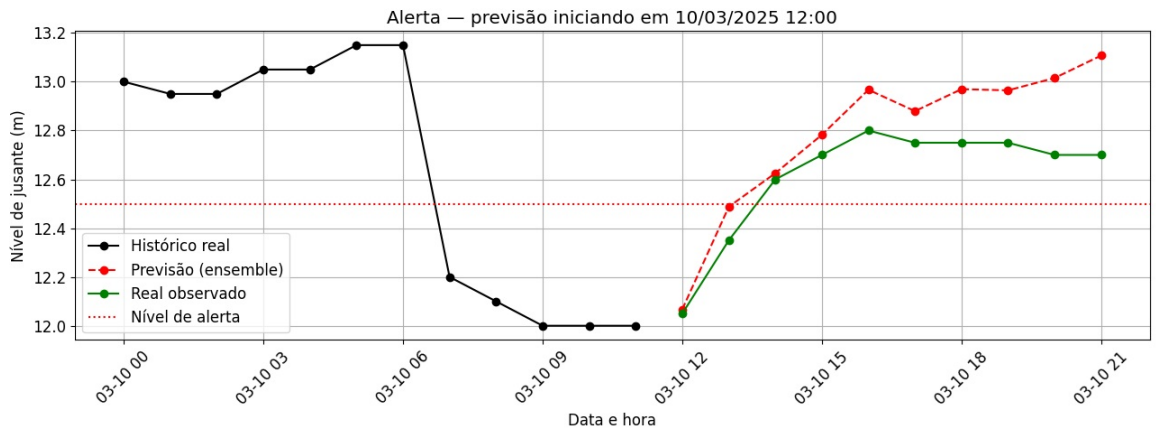


Figura 43 – Previsão x Nível de Alerta - Evento Real em 10.03.2025, às 12:00 AM.
 Fonte: Elaborada pelo autor.

Esse comportamento indica ausência de instabilidade numérica ou sensibilidade excessiva a pequenas variações na entrada, característica fundamental para aplicação operacional.

Observou-se também, comportamento típico de modelos multi-step, no qual:

- O erro é reduzido nos primeiros passos (t+1, t+2, t+3).
- O erro apresenta crescimento progressivo conforme o horizonte aumenta.
- Não foram identificadas divergências explosivas ou instabilidade acumulativa.

Esse padrão sugere que o modelo mantém capacidade preditiva adequada até o décimo passo, com degradação gradual e controlada.

CONCLUSÃO

Este trabalho, utilizou redes neurais tipo LSTM, GRU e TCN como formalismo de modelo para representação dinâmica aplicado à predição da série temporal do nível do Rio Tocantins à jusante da UHE Tucuruí. Os modelos propostos exibem ajustes com qualidade satisfatória aos dados e os resultados apontam para a sua viabilidade de aplicação de previsão em cenários reais. Merece destaque no ajuste dos hiperparâmetros, o valor do atraso das variáveis de entrada, em que se nota que, quanto maior a janela de tempo aplicada à rede, melhor o desempenho global, apesar de aumentar o custo computacional e, conseqüentemente, o aumento significativo do tempo de treinamento.

Os resultados obtidos indicam que a avaliação da influência das vazões afluente $v_a(t)$ e/ou defluente $v_d(t)$ em relação ao nível de jusante $n_j(t)$ apresentou desempenho satisfatório em todas as simulações realizadas pelos modelos \mathcal{H}_1 , \mathcal{H}_2 e \mathcal{H}_3 . Mesmo no experimento mais desafiador, com 10 horas à frente, o erro relativo aplicado ao desvio padrão dos resultados variou na faixa de 98% (1ª amostra) a 84% (10ª amostra), valores considerados satisfatórios para esse tipo de evento, dada a ordem de grandeza da variável medida.

Logo, a estratégia utilizada nas modelagens via LSTM, GRU e TCN resultou numa melhor precisão, prescrita pela métrica FitR^2 , principalmente envolvendo as variáveis de entrada $v_a(t)$ e $v_d(t)$, proposta no modelo \mathcal{H}_3 , dado que as duas features contribuem bem para os resultados, o que dá melhor objetividade na determinação do $n_j(t)$, que se espera para a janela de tempo de 10 horas à frente.

Sob uma ótica mais global, é pertinente considerar que ter um previsor multivariável aumenta a probabilidade de acerto. Assim, pode-se concluir que a melhor estratégia é a do modelo \mathcal{H}_3 que utiliza duas variáveis de entrada, cujo melhor desempenho foi, moderadamente, da rede GRU.

A análise dos gráficos das figuras 39, 40 e 41, dos valores desnormalizados do nível de

jusante, reforçam que os modelos baseados na arquitetura GRU apresentam superioridade em todos os horizontes de previsão avaliados (1^a, 5^a e 10^a amostras). Neles, observa-se maior aderência ao sinal real e menor degradação da performance ao longo do tempo, quando comparados às redes LSTM e TCN. Tais achados, aliados às métricas quantitativas, consolidam a arquitetura GRU como a mais adequada para o controle preditivo do nível de jusante.

Quanto a avaliação das previsões realizadas às 11:00 e 12:00 do dia 10/03/2025, figuras 42 e 43, demonstra que o modelo ensemble apresenta:

- Capacidade de capturar tendência hidrológica;
- Robustez frente à atualização da janela de entrada;
- Potencial aplicabilidade em sistema de alerta antecipado.

A ideia final é de que o modelo escolhido, neste caso, o \mathcal{H}_3 , possa ser implementado num sistema real, para que faça previsões dos níveis $n_j(t)$ e sirva de referência para o controle do nível tanto do reservatório quanto da jusante e auxilie autoridades de órgãos de interesse na tomada de decisão quanto a ter ou não ações de remanejamento de possíveis famílias a serem atingidas por inundações provocadas por enchentes do Rio Tocantins.

A inclusão de variáveis exógenas, a avaliação em cenários hidrológicos extremos e a implementação em tempo real com controle preditivo pode ser explorado como trabalhos futuros para complementar ou aprimorar este.

REFERÊNCIAS

- ABID, M.; COMPAGNON, P.; LEFEBVRE, G. Improved cnn-based magnetic indoor positioning system using attention mechanism. In: IEEE. *2021 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. [S.l.], 2021. p. 1–8. 25, 26
- AGGARWAL, C. C. *et al.* *Neural networks and deep learning*. [S.l.]: Springer, 2018. v. 10. 25
- Agência Nacional de Águas. *Conjuntura dos Recursos Hídricos no Brasil*. Brasília: ANA, 2013. 1
- Agência Nacional de Águas (ANA). *Manual do Empreendedor sobre Segurança de Barragens (Volume e Guias associados)*. [S.l.], 2016. Guias e manuais nacionais que abordam inspeção, operação e segurança de barragens — inclui orientação sobre vertedouros e planos de operação de reservatórios. Disponível em: <https://biblioteca.ana.gov.br/sophia_web/Busca/Download?codigoArquivo=124738>. 12
- Agência Nacional de Águas e Saneamento Básico (ANA). *Divisão Hidrográfica Nacional (DHN250)*. Brasília: ANA, 2021. Acesso em: 16 out. 2025. Disponível em: <<https://www.gov.br/ana>>. 6
- AKKAYA, S. O. Operation of hydroelectric power plants, dam reservoirs and implications for energy planning. *Applied Sciences*, v. 15, n. X, p. 183, 2024. 1
- AMBIENTE, M. do M. *Caderno setorial de recursos hídricos: geração de energia hidrelétrica*. [S.l.]: Secretaria de Recursos Hídricos Brasília, 2006. 7
- (ANA), A. N. de Águas. *Atlas dos Recursos Hídricos do Brasil*. Brasília, 2009. Inclui descrição da nascente e características físicas do Rio Araguaia, nas proximidades do Parque Nacional das Emas a cerca de 850 m de altitude. 6
- _____. *Rede Hidrometeorológica Nacional: Diagnóstico e Propostas para Monitoramento Hidrológico no Brasil*. Brasília, 2012. Discute a importância da medição e monitoramento contínuo de variáveis hidrológicas, especialmente nível e vazão, para gestão e mitigação de impactos ambientais em bacias hidrográficas. 8
- _____. *Plano de Recursos Hídricos da Bacia Tocantins–Araguaia*. Brasília, 2015. Caracterização física, hidrológica e socioeconômica da Bacia Tocantins–Araguaia. 5
- (ANA), A. N. de Águas e S. B. *Divisão Hidrográfica Nacional (DHN250)*. 2021. <<https://metadados.ana.gov.br/geonetwork/srv/por/catalog.search#/metadata/909dfd00-39cb-47a2-b332-3b43695a2a03>>. Brasília: ANA. Acesso em: 16 out. 2025. 6
- APAYDIN, H.; FEIZI, H.; SATTARI, M. T.; COLAK, M. S.; SHAMSHIRBAND, S.; CHAU, K.-W. Comparative analysis of recurrent neural network architectures for reservoir inflow forecasting. *Water*, MDPI, v. 12, n. 5, p. 1500, 2020. 2
- AWAD, M.; KHANNA, R. *Efficient learning machines: theories, concepts, and applications for engineers and system designers*. [S.l.]: Springer nature, 2015. 16

- BABIĆ, K.; MARTINČIĆ-IPŠIĆ, S.; MEŠTROVIĆ, A. Survey of neural text representation models. *Information*, MDPI, v. 11, n. 11, p. 511, 2020. 22
- BAI, S.; KOLTER, J. Z.; KOLTUN, V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*, 2018. 16, 21, 22
- BARRAGENS, C. M. D. Estudos de caso da comissão mundial de barragens: Usina hidrelétrica de tucuruí (brasil). *RELATÓRIO FINAL*. Disponível em: http://www.lima.coppe.ufrj.br/files/projetos/ema/tucuruui_rel_final.pdf, 2000. 7
- BENGIO, Y.; GOODFELLOW, I.; COURVILLE, A. *Deep learning*. [S.l.]: MIT press Cambridge, MA, USA, 2017. v. 1. 17
- BEVEN, K. *Rainfall-Runoff Modelling: The Primer*. 2. ed. Chichester: Wiley-Blackwell, 2012. 5
- BISHOP, C. M. *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press, 1995. 19
- BISHOP, C. M.; NASRABADI, N. M. *Pattern recognition and machine learning*. [S.l.]: Springer, 2006. v. 4. 16, 19
- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. *Time Series Analysis: Forecasting and Control*. 5. ed. Hoboken: Wiley, 2015. 3
- BRUNNER, M. I. Multivariate hydrological extremes. *Hydrology and Earth System Sciences*, v. 27, p. 2479–2497, 2023. Aborda fenômenos extremos multivariados em hidrologia, incluindo dependências complexas entre variáveis, mudanças de regime e não-estacionaridade. Disponível em: <https://hess.copernicus.org/articles/27/2479/2023/>. 10
- BRUNTON, S. L.; NOACK, B. R.; KOUMOUTSAKOS, P. Machine learning for fluid mechanics. *Annual review of fluid mechanics*, Annual Reviews, v. 52, p. 477–508, 2020. 2
- CAMACHO, E. F.; ALBA, C. B. *Model Predictive Control*. 2. ed. London: Springer, 2013. 16
- CARVALHO, D. C. Tocantins and araguaia. In: *Rivers of South America*. [S.l.]: Elsevier, 2025. 6
- CHAUHAN, N. S. Introduction to artificial neural networks (ann). *Medium*, [Online]. Available: <https://towardsdatascience.com/introduction-to-artificial-neural-networks-ann-1aea15775ef9>. [Accessed 1 August 2020], 2020. 24
- CHEN, Y. *et al.* Nonlinear model predictive control using gru neural networks for dynamic systems. *IEEE Transactions on Industrial Electronics*, IEEE, v. 67, n. 9, p. 7896–7905, 2020. 16
- CHO, K.; MERRIËNBOER, B. V.; GULCEHRE, C.; BAHDANAU, D.; BOUGARES, F.; SCHWENK, H.; BENGIO, Y. Learning phrase representations using rnn encoder–decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014. 25
- CHUNG, J.; GULCEHRE, C.; CHO, K.; BENGIO, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, 2014. 25
- CLOKE, H. L.; PAPPENBERGER, F. Hydrological forecasting: Progress and future directions. *Progress in Physical Geography*, v. 33, n. 3, p. 306–330, 2009. 5
- CORREIA, P. V.; DANTASY, L.; GUEDES, L. A.; FERNANDES, M. Análise de desempenho de redes neurais lstm com técnicas de pruning para detecção de falhas em processos industriais. In: *Simpósio Brasileiro de Automação Inteligente-SBAI*. [S.l.: s.n.], 2021. v. 1, n. 1. 23

- DI, C.; YANG, X.; WANG, X. A four-stage hybrid model for hydrological time series forecasting. *PLoS One*, v. 9, n. 8, p. e104663, 2014. Modelo híbrido para séries hidrológicas com características não-lineares, não-estacionárias e multiescala, destacando a integração entre dados e modelagem. Disponível em: <<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0104663>>. 10
- DIAS, V. de S. *et al.* An overview of hydropower reservoirs in brazil: Current status and challenges. *Water*, v. 10, n. 5, p. 592, 2018. Discusses climatic impacts, storage and operational rules for Brazilian hydropower reservoirs. 1
- DOOGE, J. C. I.; KANE, P. *Deterministic Methods in Systems Hydrology*. Rotterdam, Netherlands: Taylor & Francis, 1997. (IHE Delft Lecture Notes). Apresenta a teoria de sistemas aplicada à hidrologia: a natureza dos sistemas hidrológicos, fundamentos matemáticos e aplicação à resposta de bacias de drenagem. Ideal para fundamentar aspectos como sistema dinâmico, resposta de bacia e modelagem conceitual (linear e não linear). ISBN 9781315275123. Disponível em: <<https://www.taylorfrancis.com/books/mono/10.1201/9781315275123/deterministic-methods-systems-hydrology-james-dooge-philip-kane>>. 9
- ELMAN, J. L. Finding structure in time. *Cognitive Science*, Wiley, v. 14, n. 2, p. 179–211, 1990. 22, 23
- Empresa de Pesquisa Energética. *Plano Decenal de Expansão de Energia 2032*. Brasília, 2022. 6
- ENGINEERING, J. of I. S.; MANAGEMENT. Models validation using normalized rmse and r^2 metrics. *Journal of Information Systems Engineering and Management*, v. 10, n. 4, p. 1–15, 2025. Discussão da aplicação prática de métricas como nRMSE e R^2 na comparação de modelos. 3
- ENKHBAT, A.; GOCHOO, M.; SHIH, T.; CHEEWAPRAKOBKIT, P.; ADITYA, W.; QUY, T. D.; LIN, H.; LIN, Y.-T. Using hybrid models for action correction in instrument learning based on ai. *IEEE Access*, PP, p. 1–1, 01 2024. 21
- FARIAS, T. S.; ROSSI, R. G.; SUL-MS-BRAZIL, M. G. do. *Estudo Comparativo de Arquiteturas de Redes Neurais em Análise de Sentimentos*. Tese (Doutorado) — Dissertação (Mestrado)—Universidade Federal do Mato Grosso do Sul, 2021 . . . , 2021. 20
- FLECK, L.; TAVARES, M. H. F.; EYNG, E.; HELMANN, A. C.; ANDRADE, M. d. M. Redes neurais artificiais: Princípios básicos. *Revista Eletrônica Científica Inovação e Tecnologia*, v. 1, n. 13, p. 47–57, 2016. 18
- FU, Y.; YANG, A.; YAO, Z.; LIU, Y.; LI, H.; CHEN, H.; WANG, X. Inversion of rayleigh wave dispersion curves via long short-term memory combined with particle swarm optimization. *Computational Intelligence and Neuroscience*, Wiley Online Library, v. 2022, n. 1, p. 2640929, 2022. 25
- FURTADO, M. I. V. Redes neurais artificiais: uma abordagem para sala de aula. *Ponta Grossa, PR. Atena Editora*, p. 19, 2019. 17
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. Cambridge, MA: MIT Press, 2016. Obra fundamental que define Deep Learning como subárea de Machine Learning baseada em redes neurais profundas capazes de aprender representações complexas. 17, 18, 19, 22, 23
- GREFF, K.; SRIVASTAVA, R. K.; KOUTNIK, J.; STEUNEBRINK, B. R.; SCHMIDHUBER, J. Lstm: A search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, IEEE, v. 28, n. 10, p. 2222–2232, 2017. 26

- GUILHON, L. G. F.; ROCHA, V. F.; MOREIRA, J. C. Comparação de métodos de previsão de vazões naturais afluentes a aproveitamentos hidroelétricos. *Revista Brasileira de Recursos Hídricos*, v. 12, n. 3, p. 13–20, 2007. 8
- GUITARRARA, P. "o que é sustentabilidade ambiental?"; brasil escola. <https://brasilecola.uol.com.br/o-que-e/geografia/o-que-e-sustentabilidade.htm>, s.d. 6, 7
- HA, S.; YUN, J.-M.; CHOI, S. Multi-modal convolutional neural networks for activity recognition. In: IEEE. *2015 IEEE International conference on systems, man, and cybernetics*. [S.l.], 2015. p. 3017–3022. 19
- HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. 2. ed. Upper Saddle River, NJ: Prentice Hall, 1998. Explica que a função de ativação não-linear atua sobre o potencial de ativação para produzir a saída do neurônio artificial. 18
- HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 770–778. 21
- HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. *Neural computation*, MIT press, v. 9, n. 8, p. 1735–1780, 1997. 23, 24
- (IBGE), I. B. de Geografia e E. *Geografia do Brasil: Região Centro-Oeste*. Rio de Janeiro, 2004. Inclui descrição do Rio Araguaia, sua nascente próxima ao Parque Nacional das Emas e informações sobre a Ilha do Bananal, a maior ilha fluvial do mundo. 7
- JÚNIOR, A. R. B. *Elementos de Hidrologia Aplicada*. São Paulo: Editora Blucher, 2022. 8
- KASHIMA, K. *Learning-based Model Predictive Control: Toward Safe, High-performance Control by Machine Learning*. [S.l.]: Springer, 2023. 2
- KHORRAM, S.; JEHBEZ, N. A hybrid cnn-lstm approach for monthly reservoir inflow forecasting. *Water Resources Management*, Springer, p. 1–25, 2023. 2
- KOUTSOYIANNIS, D. Hydrology: change and uncertainty—a needed change of perspective. *Hydrological Sciences Journal*, v. 58, n. 3, p. 765–778, 2013. Discussões sobre mudança, incerteza e memória de longo alcance em séries hidrológicas; relevante para não-linearidade e multiescala. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/02626667.2013.804626>>. 12
- KRATZERT, F.; KLOTZ, D.; BRENNER, C.; SCHULZ, K.; HERRNEGGER, M. Rainfall–runoff modelling using long short-term memory (lstm) networks. *Hydrology and Earth System Sciences*, v. 22, n. 11, p. 6005–6022, 2018. Introduz o uso de LSTM para previsão chuva-vazão, demonstrando a capacidade da rede em aprender dinâmicas temporais complexas sem equações físicas explícitas. Disponível em: <<https://hess.copernicus.org/articles/22/6005/2018/>>. 12
- KRUG, Á. B.; PARRAGA, A.; LORENZI, F.; NICOLAO, M.; MORALES, A. S. Análise e reconhecimento de padrões usando processamento de imagens e inteligência artificial. *Revista de Iniciação Científica da ULBRA*, v. 7, n. 7, 2008. 18
- KUMAR, K. R. *et al.* Lstm-based surrogate modeling for nonlinear model predictive control of complex systems. *Control Engineering Practice*, Elsevier, v. 115, p. 104880, 2021. 16
- LIMA, R. d. S. Previsão de temperaturas dentro de um forno de reaquecimento de placas de aço usando redes neurais. Serra, 2024. 22

- LJUNG, L. *System Identification: Theory for the User*. 2. ed. Upper Saddle River: Prentice Hall, 1999. 5
- LOUCKS, D. P.; BEEK, E. van. *Water Resources Systems Planning and Management: An Introduction to Methods, Models and Applications*. Cham: Springer / UNESCO, 2017. Comprehensive treatment of reservoir operation e planejamento de sistemas hídricos. 1
- MANUAL de Procedimentos da Operação (MPO) — Módulo 5: Submódulo 5.11. Brasil, Ano de publicação. Instrução do ONS referente ao Submódulo 5.11. 2
- MITCHELL, T. M. *Machine Learning*. New York: McGraw-Hill, 1997. Obra clássica que define Machine Learning como área da Inteligência Artificial focada em algoritmos que aprendem a partir de dados. 16
- NAGHETTINI, M. *Hidrologia Estatística*. Belo Horizonte: Editora UFMG, 2017. 9
- NAIR, V.; HINTON, G. E. Rectified linear units improve restricted boltzmann machines. In: *Proceedings of the 27th International Conference on Machine Learning (ICML)*. [S.l.: s.n.], 2010. p. 807–814. 22
- NAVARRO, P. G.; PLAYÁN, E. (Ed.). *Numerical Modelling of Hydrodynamics for Water Resources: Proceedings of the Conference on Numerical Modelling of Hydrodynamic Systems*. [S.l.]: Taylor & Francis / CRC Press, 2008. Reúne modelagem hidrodinâmica numérica aplicada a recursos hídricos, canalização, reservatórios, geometria de bacia. 12
- NOGUEIRA, T. d. C. *et al.* Modelo baseado em redes neurais profundas com unidades recorrentes bloqueadas para legendagem de imagens por referências. Universidade Federal de Goiás, 2020. 24
- OLAH, C. Understanding lstm networks. *colah's blog*, 2015. Disponível em: <<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>>. 24
- Operador Nacional do Sistema Elétrico. *Dados Hidrológicos de Reservatórios – Base Horária*. 2025. Portal de Dados Abertos do ONS. Acesso em: 06 jan. 2026. Disponível em: <https://dados.ons.org.br/dataset/dados_hidrologicos_ho>. 27
- Operador Nacional do Sistema Elétrico (ONS). *O planejamento da operação energética no Sistema Interligado Nacional: conceitos, modelagem matemática, previsão de geração e carga*. [S.l.], 2018. Documento do operador do SIN que descreve a operação integrada; útil para fundamentar afirmações sobre a integração de usinas (incluindo Tucuruí) ao SIN e seu papel na garantia de suprimento, especialmente em períodos secos. Disponível em: <<https://www.ons.org.br/AcervoDigitalDocumentosEPublicacoes/O%20Planejamento%20da%20Operacao%20Energ%C3%A9tica%20no%20Sistema%20Interligado%20Nacional%20conceitos%20modelagem%20matem%C3%A1tica%20previs%C3%A3o%20de%20gera%C3%A7%C3%A3o%20e%20carga.pdf>>. 14
- _____. *Dados Hidráulicos por Reservatório – Base Diária*. 2024. <<https://dados.ons.org.br/dataset/dados-hidrologicos-res>>. Acesso em: 16 out. 2025. Licença: CC BY. 3
- POLLHAMMER, W.; SPIJKER, C.; SIX, J.; ZOGLAUER, D.; RAUPENSTRAUCH, H. Modeling of a walking beam furnace using cfd-methods. *Energy Procedia*, Elsevier, v. 120, p. 477–483, 2017. 21
- PORTO, R. L. L.; PORTO, M. F. *Hidrologia Ambiental*. São Paulo: Editora da USP, 2012. Aborda a importância das variáveis hidrológicas nível e vazão para o monitoramento e gestão dos recursos hídricos de bacias hidrográficas. 8

- QI, Y.; ZHOU, Z.; YANG, L.; QUAN, Y.; MIAO, Q. A decomposition-ensemble learning model based on lstm neural network for daily reservoir inflow forecasting. *Water Resources Management*, Springer, v. 33, p. 4123–4139, 2019. 2
- QIU, J.; YANG, Q.; ZHANG, X.; HUANG, M.; ADAM, J. C.; MALEK, K. Implications of water management representations for watershed hydrologic modeling in the yakima river basin. *Hydrology and Earth System Sciences*, v. 23, p. 35–49, 2019. Trabalho que mostra como representações de gestão hídrica (reservatórios, abstrações, irrigação) afetam a modelagem hidrológica de bacias. 12
- RASCHKA, S.; MIRJALILI, V. *Machine Learning with Python and Scikit-Learn, Keras, and TensorFlow*. 3. ed. Birmingham: Packt Publishing, 2022. 3
- REICHSTEIN, M.; CAMPS-VALLS, G.; STEVENS, B.; JUNG, M.; DENZLER, J.; CARVALHAIS, N.; PRABHAT. Deep learning and process understanding for data-driven earth system science. *Nature*, v. 566, n. 7743, p. 195–204, 2019. Revisão ampla que discute como o deep learning (AM) pode extrair padrões espaço-temporais a partir de grandes bases de dados e como integrá-lo com modelos físicos (hybrid/physics-informed ML). 12
- ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, American Psychological Association, v. 65, n. 6, p. 386–408, 1958. Artigo seminal que introduz o perceptron como modelo matemático inspirado no neurônio biológico. 18
- ROSS, J. L. S.; PRETTE, M. E. D. Recursos hídricos e as bacias hidrográficas: âncoras do planejamento e gestão ambiental. *Revista do departamento de geografia*, v. 12, n. 12, p. 89–121, 1998. 1
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature*, Nature Publishing Group, v. 323, n. 6088, p. 533–536, 1986. 19
- SAHOO, B. B.; JHA, R.; SINGH, A.; KUMAR, D. Long short-term memory (lstm) recurrent neural network for low-flow hydrological time series forecasting. *Acta Geophysica*, Springer, v. 67, n. 5, p. 1471–1481, 2019. 1, 23
- SAMUEL, A. L. Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, IBM, v. 3, n. 3, p. 210–229, 1959. 16
- SANTOS, A. T. *et al.* Desenvolvimento de estratégia para projeto de controlador preditivo baseada em deep learning. Universidade Federal de Alagoas, 2020. 18, 23
- Secretaria de Estado de Planejamento e Desenvolvimento (SEPLAD) - PA. *PDRS — Plano de Desenvolvimento Regional Sustentável: Lago de Tucuruí (versão final)*. 2015. Relatório regional que descreve a UHE Tucuruí, sua construção (etapas 1992 e 2007), capacidade instalada e sua importância para a geração regional e integração ao SIN. Disponível em: <https://seplad.pa.gov.br/wp-content/uploads/2015/07/pdrs_lago_de_tucuruui.pdf>. 14
- SILVA, C. H. R. T. *Recursos hídricos e desenvolvimento sustentável no Brasil*. Brasília, 2012. Acesso em: 16 out. 2025. Disponível em: <<https://www2.camara.leg.br/atividade-legislativa/estudos-e-notas-tecnicas/publicacoes-da-consultoria-legislativa/areas-da-conle/tema11/recursos-hidricos-e-desenvolvimento-sustentavel-no-brasil>>. 5
- SILVA, e. a. 50 years of the water-flow variance in tucuruí reservoir related with brazilian energy consumption. *Heliyon*, v. 9, p. e12640, 2023. 2

- SILVEIRA, J. G. *et al.* Land use, land cover change and sustainable alternatives for agriculture and livestock in Brazilian biomes. *Sustainability*, v. 14, 2022. Acesso em: 18 out. 2025. Disponível em: <<https://www.mdpi.com/2071-1050/14/5/2563>>. 7
- SUTSKEVER, I.; VINYALS, O.; LE, Q. V. Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems*, v. 27, 2014. 3
- TUCCI, C. E. M. *Hidrologia: Ciência e Aplicação*. 4. ed. Porto Alegre: Editora da UFRGS, 2017. 1
- VAN, S. P.; LE, H. M.; THANH, D. V.; DANG, T. D.; LOC, H. H.; ANH, D. T. Deep learning convolutional neural network in rainfall-runoff modelling. *Journal of Hydroinformatics*, IWA Publishing, v. 22, n. 3, p. 541–561, 2020. 19
- Wikipedia contributors. *Usina Hidrelétrica de Tucuruí*. 2023. <https://pt.wikipedia.org/wiki/Usina_Hidrelétrica_de_Tucuruí>. Acesso em 10 jan. 2026. 1
- WILKS, D. S. *Statistical Methods in the Atmospheric Sciences*. 3rd. ed. [S.l.]: Academic Press, 2011. Livro-texto abrangente com métodos estatísticos relevantes (autocorrelação, heteroscedasticidade, testes de significância, modelos de séries temporais) aplicáveis a dados hidrometeorológicos. 10
- World Meteorological Organization (WMO). *Guide to Hydrological Practices: Volume I - Hydrology – From Measurement to Hydrological Information*. 6. ed. Geneva: WMO, 2008. Aborda os principais fatores que influenciam a variabilidade temporal da vazão, incluindo sazonalidade das chuvas, alterações de cobertura vegetal, efeitos de reservatórios e fusão de neve. 9
- ZHAO, X. *et al.* A comprehensive review of methods for hydrological forecasting using deep learning. *Water (MDPI)*, 2024. Review article. Disponível em: <<https://www.mdpi.com/2073-4441/16/10/1407>>. 2
- ZHI, Y.; HUANG, D.; WANG, S. Temporal convolutional networks for nonlinear model predictive control. *Journal of Process Control*, Elsevier, v. 112, p. 10–22, 2022. 16