

A Comparative Analysis of Artificial Neural Networks on River Level Forecasting for the Rio Madeira Basin

João F. S. Brandão, Francis Wagner Silva Correia, Elloá B. Guedes

¹Escola Superior de Tecnologia
Universidade do Estado do Amazonas
Av. Darcy Vargas, 1200 – Manaus – Amazonas
{jfdsb.eng19, fcorreia, ebgcosta}@uea.edu.br

Abstract. *This paper presents a comparative analysis of two Artificial Neural Network approaches for predicting river levels using time series data from four hydrological stations in the Rio Madeira basin. The study utilizes data collected between 2001 and 2014 and employs a time series split validation strategy combined with a grid search over various network topologies. The findings reveal that Long-Short Term Recurrent Neural Networks exhibit a slight performance advantage over Multilayer Perceptron Networks, although both approaches demonstrate comparable performance in most scenarios. The proposed autoregressive approach for river quota prediction proves to be feasible, particularly in situations with limited historical data.*

Resumo. *Este artigo apresenta uma análise comparativa de duas abordagens de Redes Neurais Artificiais para prever os níveis de cota dos rios usando dados de séries temporais de quatro estações hidrológicas na bacia do Rio Madeira. O estudo utiliza dados coletados entre 2001 e 2014 e emprega uma estratégia de validação com divisão da série temporal combinada com uma busca em grade por diversas topologias de redes. Os resultados revelam que as Redes Neurais Recorrentes LSTM apresentam uma leve vantagem de desempenho em relação às Redes MLP, embora ambas as abordagens demonstrem desempenho comparável na maioria dos cenários. A abordagem autorregressiva proposta para previsão de cota dos rios mostra-se viável, especialmente em situações com dados históricos limitados.*

1. Introdução

Uma série temporal é uma sequência ordenada de observações de uma variável em intervalos igualmente espaçados por Δt representada como um conjunto de valores discretos $\mathbf{X} = \{X_1, X_2, X_3, \dots\}$. Em muitos problemas práticos, essa sequência de valores é obtida a partir de observações de grandezas físicas por sensores. A análise de séries temporais é um termo amplo que compreende atividades como: (i) definir, classificar e descrever a série temporal em si; (ii) elaborar modelos que utilizam valores coletados da série; e (iii) prever valores futuros da série [Palit and Popovic 2005].

Para a previsão de séries temporais, em particular, existem duas perspectivas proeminentes. A primeira fundamenta-se na utilização de técnicas de suavização exponencial e análise de regressão, enquanto a segunda perspectiva baseia-se no emprego de modelos para capturar a característica do comportamento de longo prazo do sistema dinâmico subjacente [Nielsen 2020]. Nesta última, em particular, destacam-se os modelos de *Machine Learning*, uma subárea da Inteligência Artificial que lida com a elaboração de soluções empíricas com a utilização de modelos e métodos computacionais que aprendem a partir de dados, permitindo assim a elaboração de inferências a partir de amostras [Russel and Norvig 2010, Alpaydin

2010]. A utilização de *Machine Learning* para previsão de séries temporais iniciou-se por volta dos anos 1980 e atualmente tem se beneficiado do rápido crescimento na quantidade e qualidade dos conjuntos de dados disponíveis; do aumento da capacidade de processamento dos computadores modernos; e também do desenvolvimento de métodos inteligentes para extração de padrões complexos nas séries temporais até mesmo na presença de ruído. [Palit and Popovic 2005, Nielsen 2020].

O conjunto de modelos de *Machine Learning* é abrangente e, para a previsão de séries temporais, destacam-se as Redes Neurais Artificiais (RNAs). Uma RNA é um modelo distribuído massivo e paralelo composto por unidades de processamento simples, que possui uma propensão natural para armazenar conhecimento experimental e disponibilizá-lo para uso [Haykin 2009]. As RNAs destacam-se na previsão de séries temporais por (i) não demandarem hipóteses sobre a dinâmica do processo que está sendo observado; (ii) não necessitarem de pré-requisito de estacionariedade na série temporal; (iii) não demandarem conhecimento altamente especializado sobre ajuste de parâmetros; e (iv) por oferecerem grande flexibilidade na manipulação de arquiteturas para previsão de séries temporais [Nielsen 2020].

O uso de RNAs para previsão de séries temporais é tipicamente modelado como uma tarefa autoregressiva, na qual um valor futuro da série, diga-se X_{t+h} em que h é denominado o horizonte de previsão, deve ser inferido a partir de observações anteriores, tais como $X_t, X_{t-1}, \dots, X_{t-k}$, em que k é denominado janela de tempo. Dessa forma, deseja-se inferir uma função $f : \mathbb{R}^k \rightarrow \mathbb{R}$ mediante Aprendizado Supervisionado tal que: $f(X_{t-k}, \dots, X_{t-1}, X_t) \mapsto X_{t+h}$. Nos métodos tradicionais de previsão de séries temporais, essa tarefa é abordada exclusivamente de forma linear, o que não é uma restrição para as RNAs, que possuem alta capacidade de adaptabilidade e modelagem de relações não-lineares [Lazzeri 2021].

As RNAs *Multi-Layer Perceptron* (MLP) compõem diversas soluções para previsão de séries temporais. Tais modelos são caracterizados pela disposição de múltiplos neurônios em camadas subsequentes, conforme a seguinte arquitetura: uma camada de entrada, uma ou duas camadas ocultas, e uma camada de saída. Neste tipo de RNA, os neurônios são do tipo *Perceptron* e conectam-se à todos os neurônios da camada subsequente. Ademais, o aprendizado dá-se pelo algoritmo *backpropagation*. As RNAs MLP são aproximadoras universais de funções [Haykin 2009]. Apesar disso, as RNAs MLP mostram-se limitadas no que diz respeito à realização de mapeamentos não-lineares de entrada-saída dependentes do tempo, essencial para problemas com séries temporais em que um certo valor X_t no tempo t está intimamente relacionado aos valores da série na janela anterior X_{t-j}, \dots, X_{t-1} .

Para superar a limitação na previsão de séries temporais observadas nas RNAs MLP, reconheceu-se a necessidade de incorporar a dimensão temporal na topologia das redes neurais, o que foi alcançado por meio da inclusão de conexões recorrentes (*loops*) entre os neurônios. Isso resultou no desenvolvimento das Redes Neurais Recorrentes (RNR). As RNRs fazem uso de uma memória do estado interno ao lidar com sequências de entradas e, em vez de simplesmente mapear apenas as entradas para as saídas como no caso das RNAs MLPs, são capazes de inferir funções de mapeamento para as entradas ao longo do tempo para então produzirem uma saída [Lazzeri 2021]. As RNRs possuíam, entretanto, dois problemas específicos no tocante ao gradiente do erro: explosão ou desvanecimento. Isso ocorria em razão do uso dos mesmos parâmetros, de tal modo que o gradiente rapidamente convergia para zero ou para o infinito, ambas situações não desejáveis para cenários práticos, fazendo com que a retropropagação de erros fosse difícil ou até mesmo impossível quando a RNR era desenrolada [Nielsen 2020]. Para contornar esse problema foram propostas as RNRs LSTM

(*Long-Short Term Memory*), as quais possuem mecanismos internos (*gates* e estados de célula) que podem regular o fluxo de informações relevantes. Esses mecanismos permitem preservar o erro que pode ser retropropagado ao longo do tempo e das camadas, tornando as LSTMs capazes de inferir dependências de longo prazo [Lazzeri 2021].

No escopo deste estudo é realizada uma análise comparativa entre RNAs MLP e RNRs LSTM para a previsão do nível dos rios na bacia do Rio Madeira. A avaliação desses métodos é de grande importância para o desenvolvimento de um sistema de previsão hidroclimática para a bacia do Rio Madeira. Esse sistema contribuirá no apoio aos tomadores de decisão (poder público) no planejamento de ações para redução e mitigação dos efeitos das secas e enchentes sobre os povos e comunidades na bacia do Rio Madeira. Ressalta-se que tal avaliação dos métodos foi desafiador, pois (i) a tarefa é de natureza empírica com dados oriundos de um contexto complexo e multi-fatorial; (ii) houve a necessidade de determinar um espaço de busca de topologias e parametrização para cada tipo de rede; e (iii) ao passo que se almejava o melhor desempenho na tarefa pelos modelos, era necessário ponderar também o ônus dos recursos computacionais para tanto. Embora não tenha sido possível determinar de maneira definitiva o imperativo de desempenho de um tipo de modelo sobre o outro, este trabalho colabora para ressaltar um melhor desempenho das RNRs LSTM em horizontes mais distantes, o menor esforço computacional na busca em grade de topologias neste tipo de modelo e também o delineamento de uma tarefa autoregressiva viável que pode ser utilizada em outros contextos para previsão de variáveis hidrológicas, especialmente quando há poucos dados disponíveis para aprendizado.

Dessa forma, este trabalho foi organizado da seguinte maneira: A Seção 2 desenvolverá uma análise teórica sobre o uso de Redes Neurais Artificiais, com ênfase nas arquiteturas *Long Short Term Memory*. A Seção 3 aborda os materiais e métodos utilizados no escopo desse trabalho, enquanto a Seção 4, faz a análise e comparação dos resultados obtidos. Por fim, a Seção 5 destina-se as considerações finais desse artigo.

2. Trabalhos Relacionados

Considerando as aplicações de *Machine Learning* na previsão de séries temporais, esta seção aborda os trabalhos relacionados com aplicações na previsão de variáveis associadas à rios em todo o mundo. Hussain e Khan (2020) avaliaram o desempenho de três algoritmos de Machine Learning para a previsão de um mês de antecedência para a descarga do Rio Hunza (Paquistão). Os autores consideraram RNAs MLPs e dois outros modelos de *Machine Learning*, a citar: Vetores de Suporte de Regressão e Florestas Aleatórias. Os resultados obtidos mostraram um desempenho competitivo dentre os modelos escolhidos, uma dificuldade comum na previsão de alguns picos da vazão e a melhor performance das Florestas Aleatórias no cenário em questão. Os autores ressaltam a importância de tal previsão para atividades locais como agricultura, produção de energia elétrica e decisões de infraestrutura, por exemplo.

O Brasil possui um forte compromisso com a geração de energia a partir de fontes renováveis. As hidrelétricas, em particular, compõem 60,2 % do total da matriz energética do País [Brasil 2022]. Nesse contexto, a previsão de variáveis hidrológicas colabora diretamente para o planejamento e o manejo de tais recursos com impactos diretos na economia, na sociedade e na natureza, pois podem mitigar a utilização de energia oriunda de termelétrica. Neste contexto, Melo et al. (2019), avaliaram a previsão da vazão mensal nas usinas hidrelétricas de Furnas e Camargos na bacia do Rio Paraná. Os autores utilizaram dados históricos de 82 anos, focaram na previsão mensal e construíram modelos regressores multivariados, isto é, capazes de receber dados simultaneamente das duas séries temporais das usinas em questão e fazer uma única previsão de natureza simultânea para as mesmas [de Melo et al. 2019].

Nessa tarefa, os autores consideraram RNAs MLP e RNRs dos tipos LSTM e GRU (do inglês, *Gated Recurrent Unit*), bem como o modelo estocástico ARIMA (do inglês, *Autoregressive Integrated Moving Average*). Os resultados experimentais destacaram que as RNAs possuíram melhor desempenho que os métodos clássicos de previsão de séries temporais, e também que as RNRs produziram melhores generalizações que as RNAs MLP, especialmente em um período em que houve seca. Quando as RNRs LSTM e GRU foram incrementadas com variáveis meteorológicas na entrada, ambas foram capazes de reduzir o erro nas previsões, capturando relações não-lineares de natureza mais complexa e tornando-se mais aptas para fins práticos da tarefa almejada.

Para a vazão diária do Rio Tâmesa, no Reino Unido, Zounemat-Kermani utilizaram nove estratégias distintas de previsão inteligente, as quais compreenderam desde as tradicionais RNAs MLP, RNRs LSTM bem como modelos de RNAs híbridos, combinando Lógica Fuzzy e Computação Evolutiva, por exemplo [Zounemat-Kermani et al. 2021]. De maneira geral, os autores observaram que em todos os cenários experimentais propostos houve bom desempenho dos modelos, com diferença percentual de 1,17% entre a melhor e a pior performance, o que sugere alta robustez no conjunto de RNAs avaliadas.

Em um comparativo na previsão de vazão nas bacias hidrográficas dos rios Kelantan e Muda na Malásia, autores utilizaram RNRs LSTM, CNNs (do inglês, *Convolutional Neural Networks*) e ConvLSTM (do inglês, *Convolutional Long Short-Term Memory*), em que essas duas últimas possuem natureza convolutiva, maior magnitude no número de parâmetros e também compõem abordagens contemporâneas para previsão de séries temporais. Os resultados experimentais demonstraram que o modelo LSTM proveu melhor resposta em bacias hidrográficas de menor escala enquanto os modelos convolutivos produziram respostas mais aproximadas em cenários com bacias hidrográficas de maior escala e com maiores valores de vazão [Dehghani et al. 2023].

Previsões para o Rio Acre, localizado na Região Amazônica, foram feitas com RNAs MLP e horizontes de previsão variando de um a quatro meses. Também foram consideradas variáveis preditoras externas, como a temperatura na superfície do mar nos Oceanos Pacífico e Atlântico tropicais e também valores de pressão atmosférica. Os resultados obtidos mostraram-se satisfatórios para antecipar a ocorrência de cheias e minimizar seus danos, mas observou-se uma dificuldade em captar sinais de ordem meteorológica na escala sub-mensal [Santos Neto et al. 2021]. Ainda na Região Amazônica, um trabalho na literatura comparou diferentes tipos de RNAs (MLP, CNN e LSTM) para previsão de cota dos Rios Solimões e Amazonas a partir de dados oriundos de satélites altimétricos. Para o Rio Solimões, a arquitetura de LSTM obteve melhor desempenho, enquanto para o Rio Amazonas a arquitetura de RNA MLP mostrou-se mais eficiente [da Silva Cardoso 2022].

A partir da análise das contribuições dos trabalhos relacionados foi possível destacar o uso bem sucedido de RNAs em diferentes contextos, a ascensão na consolidação de bons resultados empíricos das RNRs LSTM na previsão de séries temporais de variáveis hidrológicas, a necessidade de efetuar análises comparativas e também a lacuna de trabalhos que abordem tais métodos de previsão na Bacia do Rio Madeira. Estes fatores justificam a importância da contribuição proposta no sentido de minimizar a lacuna identificada, bem como contribuíram para nortear decisões do *design* experimental, o qual se encontra detalhado na seção a seguir.

3. Material e Métodos

A previsão do nível (cota) dos rios da bacia do Rio Madeira será abordada no escopo deste trabalho como uma tarefa de regressão univariada mediante Aprendizado Supervisionado. Para cada série temporal, as variáveis independentes são 15 observações diárias de valores anteriores e o atributo-alvo é um valor da série num momento futuro, com horizonte de previsão h . De maneira formal, deseja-se propor e avaliar RNAs capazes de induzir funções $f : \mathbb{R}^{15} \rightarrow \mathbb{R}$ tal que:

$$f(X_{t-14}, \dots, X_{t-1}, X_t) \mapsto X_{t+h}, h \in \{15, 30\}.$$

O objetivo é avaliar a capacidade preditiva dos modelos em questão para dois horizontes distintos e distantes ($h = 15$ ou $h = 30$) a partir das mesmas variáveis independentes e de natureza auto-regressiva, isto é, almeja-se aferir o grau de aprendizado dos padrões implícitos na própria série temporal e sua capacidade e utilidade na previsão de eventos posteriores com RNAs. Para alcançar este objetivo, as subseções a seguir detalham a área de estudo, os dados experimentais, a proposição e parametrização de modelos bem como as estratégias de avaliação de desempenho.

3.1. Área de Estudo

A bacia hidrográfica do Rio Madeira, ilustrada na Figura 1, possui uma área total de aproximadamente $1,42 \text{ mi km}^2$, representando aproximadamente 23 % da bacia Amazônica e com área de drenagem de $1,37 \text{ mi km}^2$. A bacia ocupa os territórios da Bolívia (51 %), Brasil (42 %) e Peru (7 %). Em território brasileiro, ocupa uma área de 548.960 km^2 , na qual estão compreendidos, total ou parcialmente, 88 municípios, sendo 52 no Estado de Rondônia, 18 no Mato Grosso, 12 no Amazonas e 6 no Acre. A pluviosidade média anual na bacia é de 1940 mm, a descarga média anual é de $31.200 \text{ m}^3 \text{ s}^{-1}$ e o coeficiente de escoamento igual a 0,38 [Molinier et al. 1994]. A bacia do Madeira possui baixa precipitação, coeficiente de escoamento e vazão específica, o que faz com que sua contribuição para a descarga total seja de apenas 15 %, apesar de sua área corresponder a 23 % do total da bacia Amazônica. [de Brito Gomes et al. 2022].

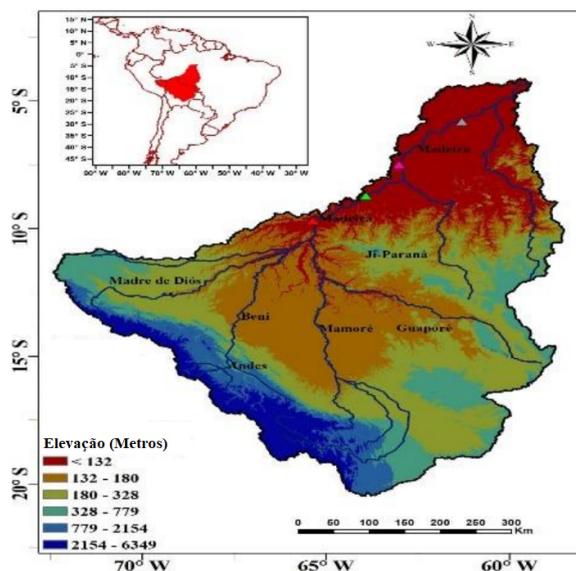


Figura 1: Bacia Hidrográfica do Rio Madeira. Fonte: [de Brito Gomes et al. 2022].

Em território brasileiro, a bacia do Rio Madeira contém duas usinas hidrelétricas de grande porte, as quais localizam-se no Estado de Rondônia e possuem capacidade conjunta de 6318 MW; abriga povos indígenas; possui importância para atividades de natureza pesqueira, principalmente em caráter artesanal de pequena escala; é uma moderna rota de exportação de cereais; e também é utilizada como rede de transporte [Brasil 2023, Silva and Costa 2014, Doria et al. 2012, do Rio et al. 2015].

3.2. Dados Experimentais e Análise Exploratória

Os dados de cota utilizados neste estudo foram coletados por quatro estações fluviométricas localizadas ao longo do Rio Madeira, sob responsabilidade da Agência Nacional de Águas (ANA) em <https://www.gov.br/ana/pt-br>. A série temporal da cota (nível ou altura) do rio em cada estação contempla registros diários entre 2001 e 2014, totalizando 20220 exemplos totais. As estações fluviométricas localizam-se nos municípios de Abunã, Humaitá, Manicoré e Porto Velho, conforme ordem alfabética.

Uma síntese da análise exploratória de cada série temporal encontra-se disponível na Tabela 1. Dada a dimensão da bacia hidrográfica em questão, percebe-se que há variações no tocante à magnitude da cota, em que a estação de Porto Velho registrou os menores valores em contrastate com os registros da estação de Manicoré, que mostraram-se os maiores dentre todos observados. A Figura 2 ilustra as séries temporais de cota das estações contempladas. Ao decompor as séries temporais de maneira aditiva, foi possível detectar a sazonalidade anual da dinâmica subjacente de cheia e vazão do rio e também uma tendência crescente de intensidade desses eventos.

Tabela 1: Análise exploratória dos valores da cota nas estações consideradas.

Estação	Coordenadas		Média	Mediana	Desvio Padrão	Máximo	Mínimo
	Latitude	Longitude					
Abunã	-09° 42' 11"	-65° 21' 53"	1319,31	1332,50	464,72	2554,00	522,00
Humaitá	-07° 30' 19"	-63° 01' 13"	1639,75	1654,00	452,81	2563,00	894,00
Manicoré	-05° 49' 00"	-61° 18' 07"	1913,80	1926,00	456,24	2887,00	1113,00
Porto Velho	-08° 44' 12"	-63° 55' 13"	943,74	960,00	443,48	1969,00	163,00

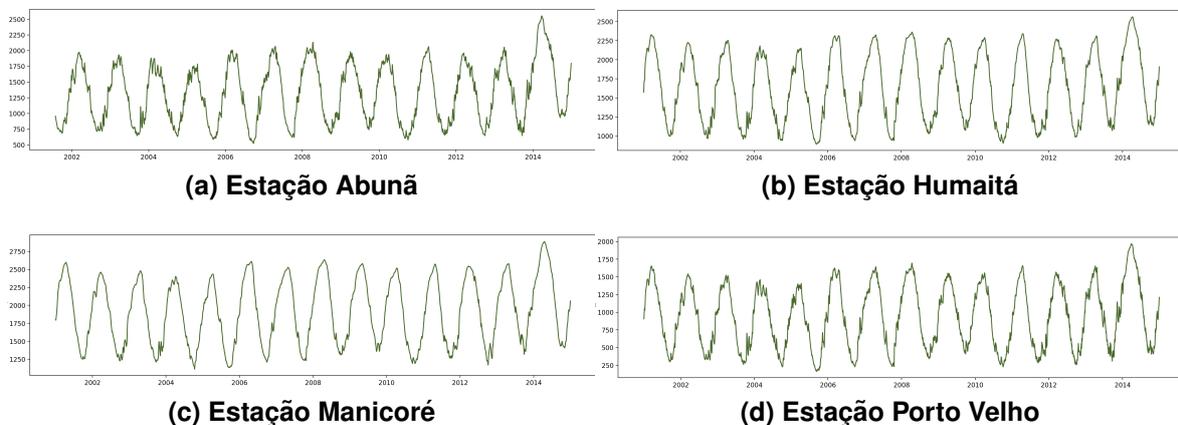


Figura 2: Série temporal da cota para as quatro estações consideradas. O eixo y denota os valores da cota (em centímetros) e o eixo x denota o intervalo de tempo da série.

3.3. Modelos, Parâmetros e Hiperparâmetros

A proposição dos modelos para o problema em questão considerou as RNAs MLP e as RNRs LSTM pois, conforme identificado em trabalhos relacionados, tais modelos já haviam sido utilizados de forma bem sucedida na previsão de séries temporais hidrológicas em diversos lugares do mundo. Apesar de tal vantagem, permanece o problema em aberto na determinação da melhor arquitetura de RNAs para um dado problema. Neste sentido, determinou-se uma grade de busca de topologias para cada tipo de modelo, cujo detalhamento é efetuado a seguir:

1. **RNAs MLP.** A proposição de arquiteturas levou em conta $N_i = 15$ neurônios na camada de entrada, $N_o = 1$ neurônio na camada de saída, e uma ou duas camadas ocultas com N_h neurônios, em que N_h foi definido conforme a Regra da Pirâmide Geométrica:

$$N_h = \alpha \left\lceil \sqrt{N_i \cdot N_o} \right\rceil = \alpha \left\lceil \sqrt{15} \right\rceil = 4\alpha, \quad (1)$$

com $\lceil \cdot \rceil$ denotando o inteiro mais próximo e $0.5 \leq \alpha \leq 2$, o que implica em $2 \leq N_h \leq 8$ [Palit and Popovic 2005]. Considerando todas as possíveis disposições de N_h neurônios em uma ou duas camadas ocultas sob a forma de um par ordenado $(b, a - b)$ com $2 \leq a \leq 8$, $1 \leq b \leq a$ e com a restrição de que $b \neq 0$, tem-se 35 arquiteturas. Além disso, considerou-se também duas funções de ativação na grade de busca de RNAs MLP (identidade e ReLU), resultando em 70 modelos desse tipo a serem treinados e testados para cada um dos horizontes de previsão. Deste total, apenas os três com melhor desempenho serão elencados nos resultados;

2. **RNRs LSTM.** A proposição das topologias para LSTM levou em conta uma única camada oculta recorrente, a qual será responsável pelo aprendizado de padrões temporais presentes nas entradas. O número de neurônios, entretanto, será igual a 7, 15 ou 30, em que esses dois últimos valores foram escolhidos em consonância com os horizontes de previsão. A variação na quantidade de unidades ocultas na camada recorrente compõe uma grade de busca pequena e considera a necessidade de explorar diferentes níveis de complexidade e capacidade de representação de cada topologia, permitindo equilibrar capacidade, generalização, e custo computacional para o problema específico em questão.

Em ambos os tipos de RNAs, as seguintes decisões sobre treinamento e determinação de hiperparâmetros foram adotadas: (i) os pesos foram inicializados de maneira aleatória, sem transferência de aprendizado ou aproveitamento de pesos de outra tarefa; (ii) o treinamento considerou um número máximo de 100 épocas; (iii) adotou-se a técnica de *Early Stopping* com uma paciência de 20 épocas para evitar *overfitting*, em que foram monitoradas métricas no conjunto de validação; e (iv) a escolha do hiperparâmetro relativo ao tamanho dos *batches* deu-se de forma empírica para cada arquitetura, balanceando a quantidade de parâmetros e a utilização dos recursos computacionais disponíveis, a exemplo da memória principal.

3.4. Avaliação de Desempenho

Para cada tarefa de regressão em questão foi levada em consideração a métrica de desempenho conhecida como R^2 -Score, a qual indica a proporção da variância na variável dependente, que é previsível, a partir da variável independente. Esta métrica é obtida como segue:

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

em que y é a variável a ser prevista, \hat{y} é uma previsão para y , e \bar{y} é a média de y [Cameron and Windmeijer 1997]. Os valores para esta métrica residem no intervalo $(-\infty, +1]$, em que $R^2 = +1$ aponta um regressor perfeito. Um estudo recente considera esta métrica mais intuitivamente informativa que o MAE, MAPE, MSE e RMSE na avaliação de tarefas de regressão, uma vez que pode ser representada como uma porcentagem, enquanto as alternativas possuem intervalos arbitrários [Chicco et al. 2021].

A estratégia de validação cruzada utilizada foi o *Time Series Split*, em que o total de observações foi particionado em 5 intervalos fixos e disjuntos, respeitando a sequência temporal. Na primeira etapa, os modelos foram treinados com o primeiro *split* e avaliados com o segundo. Em seguida, efetuou-se um treinamento com os dois primeiros *splits* e avaliou-se com o terceiro, e assim sucessivamente, de modo que os índices de teste sempre foram posteriores que os de treino. Neste cenário não se utilizou aleatoriedade, para evitar que dados do futuro fossem utilizados para prever eventos do passado. Essa é uma estratégia de validação de modelos para previsão de séries temporais em que se afere a robustez perante poucas amostras do fenômeno e sua evolução à medida que mais dados tornam-se disponíveis [Bergmeir and Benítez 2012].

Após a validação cruzada, elencou-se o melhor modelo de cada tipo segundo o critério do maior R^2 médio com o propósito de avaliar a hipótese nula (H_0) de que ambos seriam equivalentes perante a tarefa de previsão considerada. Como hipótese alternativa (H_A), tem-se que um tipo de modelo é melhor que outro, diga-se aquele com maior R^2 médio nos *splits* da validação cruzada. Na impossibilidade de assegurar que as amostras do R^2 nos *splits* seguiu a distribuição normal, partiu-se para utilização do Teste Não-Paramétrico de Wilcoxon para amostras pareadas com nível de confiança de 95 % ($\alpha = 0,05$) [Walpole et al. 2016].

3.5. Tecnologias Utilizadas

A infraestrutura computacional utilizada para os experimentos consistiu em um computador com processador Intel(R) Core(TM) i7-10875 CPU @ 2,30 GHz, 32 GB de memória principal, 512 GB de memória secundária e 1 GPUs NVIDIA GeForce RTX 2070 Super com 8 GB de VRAM para aceleração em *hardware* do treinamento das RNRs LSTM.

A linguagem de programação Python foi adotada para implementação no ambiente de desenvolvimento Jupyter Notebooks. As bibliotecas Pandas, Numpy, Matplotlib, Scipy, Sci-kit Learn e Keras foram utilizadas para manipulação dos dados das séries temporais, geração de visualizações, testes estatísticos, proposição, treino e teste das redes neurais consideradas, respectivamente. Os artefatos produzidos e o código-fonte para reprodução dos experimentos realizados encontram-se pública e gratuitamente disponíveis no repositório <https://github.com/jfsbrandao/ENIAC2023>.

4. Resultados e Discussão

Os experimentos computacionais foram executados respeitando os passos propostos e a síntese dos resultados encontra-se nas Tabelas 2 a 5, as quais denotam o desempenho das 3 melhores RNAs MLP e das 3 RNRs LSTM para cada estação sob cada horizonte de previsão considerado. A última coluna destas tabelas denota a média e o desvio padrão do R^2 resultante dos cinco *splits* da estratégia de validação cruzada adotada.

Segue-se então uma análise detalhada e comparativa dos modelos propostos para cada série temporal das estações hidrológicas consideradas. No caso de Abunã, conforme Tabela 2, é possível perceber que no primeiro *split* todos os modelos registraram o menor valor de R^2 , o que é justificável em razão da menor disponibilidade de dados para treinamento. Para o

horizonte $h = 15$, ao observar detalhadamente os modelos MLP e LSTM com melhor desempenho médio, vide Figura 3a, percebe-se que houve um decréscimo abrupto no quarto *split* para a primeira, ao passo que a segunda mostrou-se resiliente perante os desafios de tal cenário de previsão. No caso do horizonte $h = 30$, nenhum dos dois modelos elencados foi estável ao longo dos *splits*, conforme Figura 3b. Ao sobrepor as previsões dos melhores modelos para a Estação de Abunã, tem-se os gráficos ilustrados nas Figuras 4a e 4b. Conforme os testes estatísticos realizados para o desempenho dos modelos de previsão da Estação Abunã, os quais encontram-se reportados na Tabela 6, não é possível afirmar que houve diferença significativa no desempenho das melhores RNA MLP e da RNR LSTM. Isto pode ter decorrido em virtude da variabilidade dos valores da série temporal e da assimetria na sua distribuição, como justificado pelo maior desvio padrão observado e pela diferença entre média e mediana, conforme previamente apresentado na Tabela 1.

Tabela 2: Desempenho dos modelos na previsão para a Estação de Abunã.

Horizonte	Tipo	Características		R ²					$\mu \pm \sigma$
		Topologia	Ativação	Split 1	Split 2	Split 3	Split 4	Split 5	
$h = 15$	MLP	(5, 2)	Identidade	0,8076	0,8873	0,9238	0,7917	0,9129	$0,8647 \pm 0,0546$
	MLP	(4, 4)	Identidade	0,7768	0,8908	0,9257	0,7954	0,9007	$0,8579 \pm 0,0599$
	MLP	(6, 2)	Identidade	0,7177	0,8608	0,9031	0,8642	0,9166	$0,8525 \pm 0,0707$
	LSTM	(30)	Identidade	0,8247	0,9115	0,8937	0,9482	0,8959	$0,8948 \pm 0,0401$
	LSTM	(15)	Identidade	0,8376	0,8980	0,8843	0,9077	0,9004	$0,8856 \pm 0,0251$
	LSTM	(7)	Identidade	0,8076	0,8995	0,8911	0,9192	0,8992	$0,8833 \pm 0,0389$
$h = 30$	MLP	(3, 2)	Identidade	0,6343	0,6462	0,8105	0,7041	0,8413	$0,7273 \pm 0,0844$
	MLP	(4, 4)	Identidade	0,6790	0,6176	0,8059	0,6113	0,7849	$0,6997 \pm 0,0818$
	MLP	(5, 2)	Identidade	0,6282	0,6643	0,7252	0,6582	0,7817	$0,6915 \pm 0,0550$
	LSTM	(30)	Identidade	0,7281	0,8326	0,7734	0,8788	0,7765	$0,7979 \pm 0,0523$
	LSTM	(15)	Identidade	0,7899	0,8278	0,7123	0,8749	0,7820	$0,7974 \pm 0,0537$
	LSTM	(7)	Identidade	0,7037	0,7518	0,8158	0,8232	0,7252	$0,7639 \pm 0,0478$

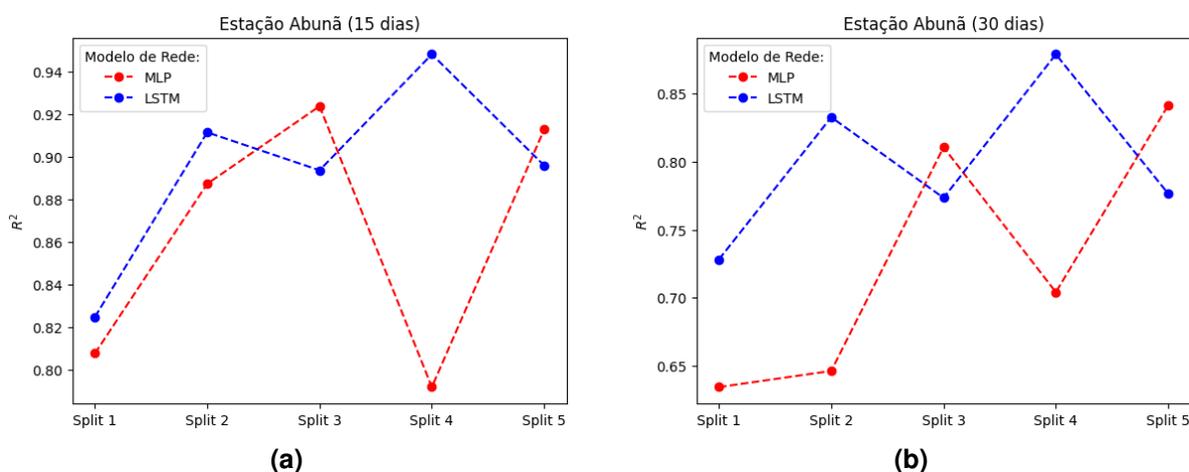
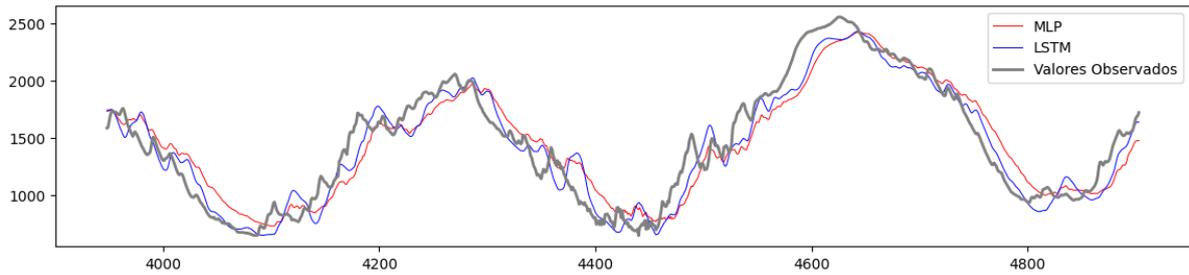
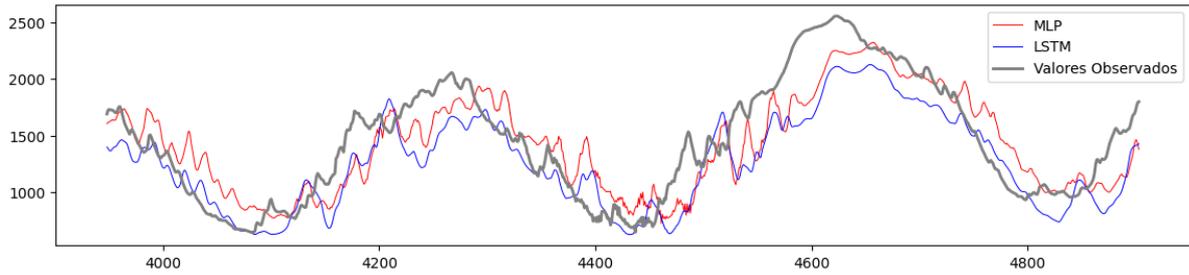


Figura 3: Detalhamento do desempenho dos modelos na previsão da Estação Abunã.

No caso da série oriunda da Estação Humaitá, as previsões efetuadas pelos melhores modelos encontram-se dispostas na Tabela 3. Para o horizonte $h = 15$, o desempenho médio de todas as RNRs LSTM foi superior ao de todas as RNAs MLP, com crescimento contínuo e



(a)



(b)

Figura 4: Previsões dos melhores modelos para a série da Estação Abunã.

estável ao longo dos *splits*, conforme pode ser visto na Figura 3a. Esta diferença é estatisticamente significativa, conforme Tabela 6. Embora uma inspeção visual tenda a levar às mesmas conclusões para o horizonte $h = 30$, vide Figura 3b, tal fato não é atestado de maneira estatisticamente significativa. O comparativo entre valores previstos e observados para a evolução da série temporal nos dois horizontes considerados é mostrado nas Figuras 6a e 6b.

Tabela 3: Desempenho dos modelos na previsão para a Estação de Humaitá.

Horizonte	Tipo	Características		R ²					$\mu \pm \sigma$
		Topologia	Ativação	Split 1	Split 2	Split 3	Split 4	Split 5	
$h = 15$	MLP	(5, 1)	Identidade	0,8694	0,8763	0,8929	0,9178	0,9310	$0,8975 \pm 0,0236$
	MLP	(5, 2)	Identidade	0,8110	0,9002	0,9140	0,9362	0,8874	$0,8898 \pm 0,0425$
	MLP	(7, 2)	Identidade	0,8461	0,8595	0,9253	0,9079	0,8886	$0,8855 \pm 0,0294$
	LSTM	(30)	Identidade	0,9256	0,9393	0,9440	0,9459	0,9295	$0,9368 \pm 0,0080$
	LSTM	(7)	Identidade	0,8987	0,9251	0,9445	0,9466	0,9145	$0,9259 \pm 0,0181$
	LSTM	(15)	Identidade	0,9197	0,9295	0,8619	0,9536	0,9344	$0,9198 \pm 0,0310$
$h = 30$	MLP	(5, 3)	Identidade	0,7492	0,7259	0,8131	0,8248	0,8333	$0,7893 \pm 0,0433$
	MLP	(3, 5)	Identidade	0,6266	0,7960	0,8212	0,7577	0,8505	$0,7704 \pm 0,0780$
	MLP	(3, 3)	Identidade	0,6468	0,8239	0,6965	0,8433	0,7982	$0,7617 \pm 0,0765$
	LSTM	(15)	Identidade	0,7316	0,8663	0,8872	0,8745	0,8651	$0,8449 \pm 0,0572$
	LSTM	(7)	Identidade	0,7454	0,8536	0,8838	0,8948	0,7824	$0,8320 \pm 0,0584$
	LSTM	(30)	Identidade	0,5575	0,8306	0,7944	0,9079	0,8462	$0,7873 \pm 0,1206$

Na série temporal da Estação Manicoré, conforme Tabela 4, tem-se conclusões similares para os modelos de previsão em analogia ao que foi observado na Estação Humaitá: diferença estatística em favor da RNR LSTM no caso de $h = 15$ e desempenho equiparável entre os dois melhores modelos de cada tipo para o horizonte $h = 30$, conforme reportado na Tabela 6. O comparativo entre valores previstos e observados para a evolução da série temporal pelos dois melhores modelos nos horizontes considerados é mostrado nas Figuras 7a e 7b.

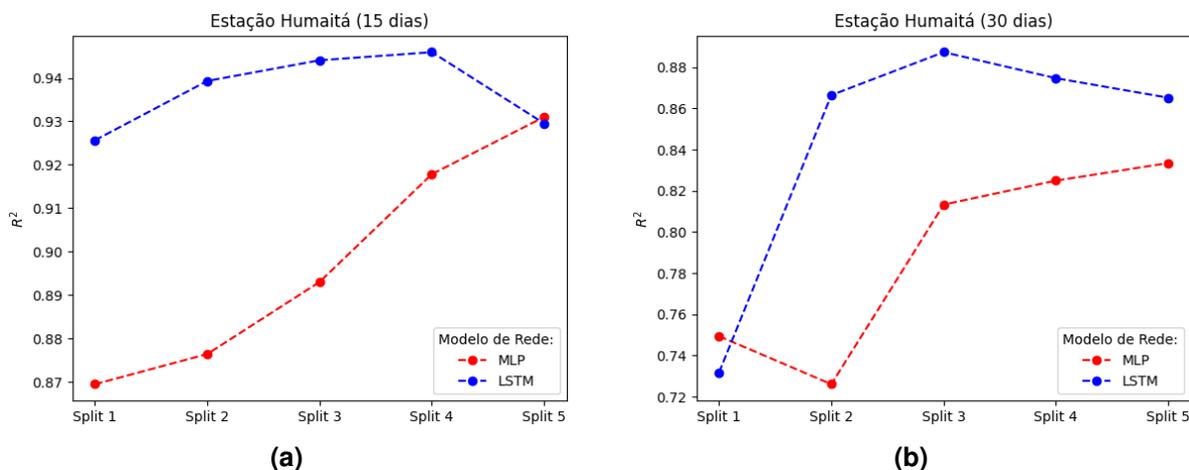


Figura 5: Detalhamento do desempenho dos modelos na previsão da Estação Humaitá.

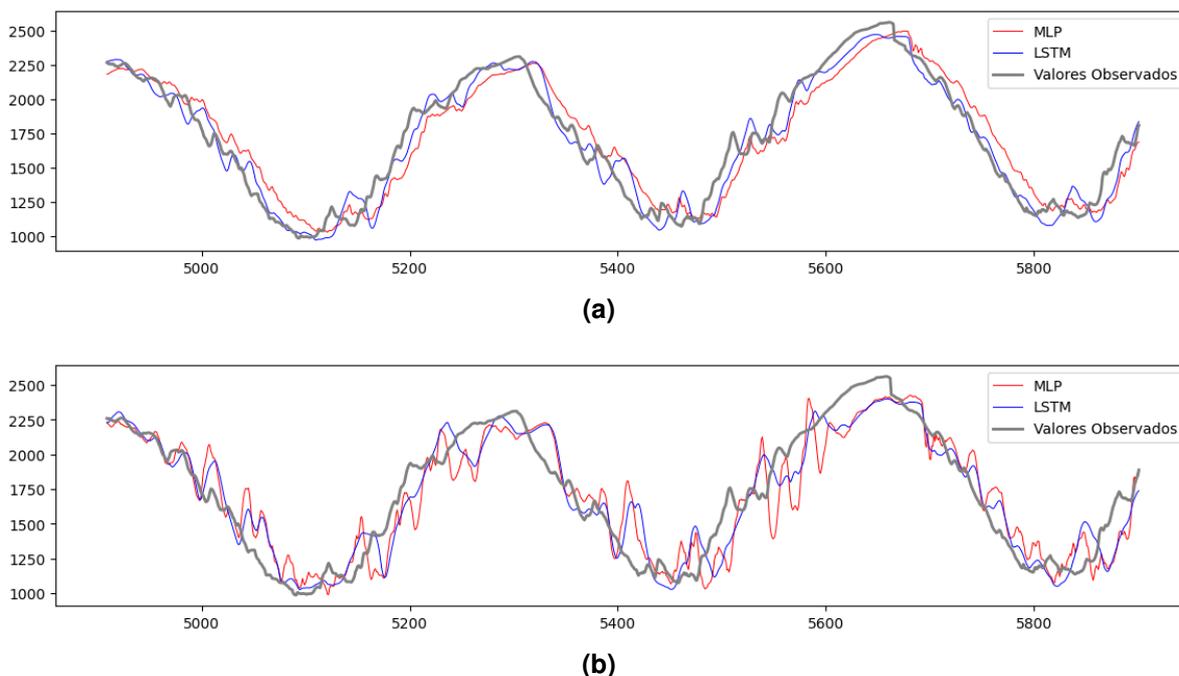


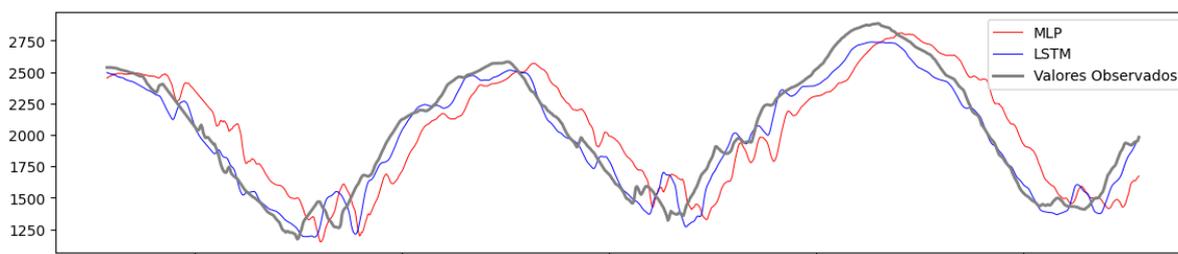
Figura 6: Previsões dos melhores modelos para a série da Estação Humaitá.

Nos experimentos realizados para Estação Porto Velho, conforme reportado detalhadamente na Tabela 5, tem-se que o desempenho de ambos os tipos de modelos foi bastante satisfatório no horizonte $h = 15$ e que, no horizonte $h = 30$, a LSTM com melhor desempenho médio mostrou-se 8,45% mais eficiente na previsão que a melhor MLP avaliada. Essas conclusões são embasadas nos testes estatísticos detalhados na Tabela 6. O comparativo entre valores previstos e observados nos horizontes considerados é mostrado nas Figuras 8a e 8b. Uma hipótese para o melhor desempenho neste cenário de previsão é decorrente da menor magnitude nos valores da cota e também do menor desvio padrão observado, conforme disposto previamente na Tabela 1.

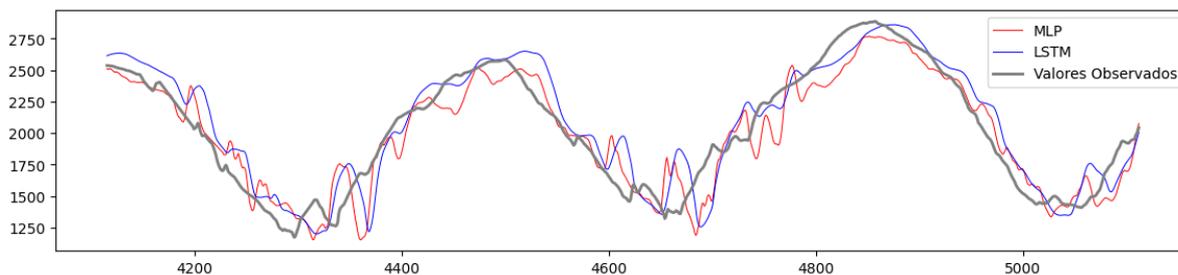
A análise comparativa dos resultados, que abrangeu uma busca em grade com mais de 300 arquiteturas de MLPs e LSTMs para as estações hidrológicas investigadas e testes

Tabela 4: Desempenho dos modelos na previsão para a Estação de Manicoré.

Horizonte	Tipo	Características		R ²					$\mu \pm \sigma$
		Topologia	Ativação	Split 1	Split 2	Split 3	Split 4	Split 5	
$h = 15$	MLP	(4, 3)	Identidade	0,8522	0,8488	0,9057	0,8912	0,8437	0,8683 \pm 0,0251
	MLP	(3, 5)	Identidade	0,7753	0,8517	0,8537	0,8950	0,9211	0,8594 \pm 0,0494
	MLP	(5, 1)	Identidade	0,8053	0,8911	0,9163	0,7947	0,8875	0,8590 \pm 0,0492
	LSTM	(30)	Identidade	0,8707	0,9045	0,9365	0,9240	0,9186	0,9109 \pm 0,0225
	LSTM	(15)	Identidade	0,8474	0,8523	0,9343	0,9304	0,9177	0,8964 \pm 0,0384
	LSTM	(7)	Identidade	0,8069	0,8828	0,9321	0,9169	0,9188	0,8915 \pm 0,0453
$h = 30$	MLP	(6, 2)	Identidade	0,7114	0,7133	0,8029	0,7579	0,8176	0,7606 \pm 0,0440
	MLP	(3, 4)	Identidade	0,7193	0,7472	0,6197	0,7062	0,8122	0,7209 \pm 0,0624
	MLP	(3, 1)	Identidade	0,6183	0,6787	0,7555	0,7303	0,7888	0,7143 \pm 0,0599
	LSTM	(30)	Identidade	0,7168	0,8227	0,8437	0,8510	0,7964	0,8061 \pm 0,0485
	LSTM	(15)	Identidade	0,7558	0,7485	0,8720	0,8172	0,8276	0,8042 \pm 0,0463
	LSTM	(7)	Identidade	0,6918	0,7661	0,7927	0,8568	0,8212	0,7857 \pm 0,0558



(a)



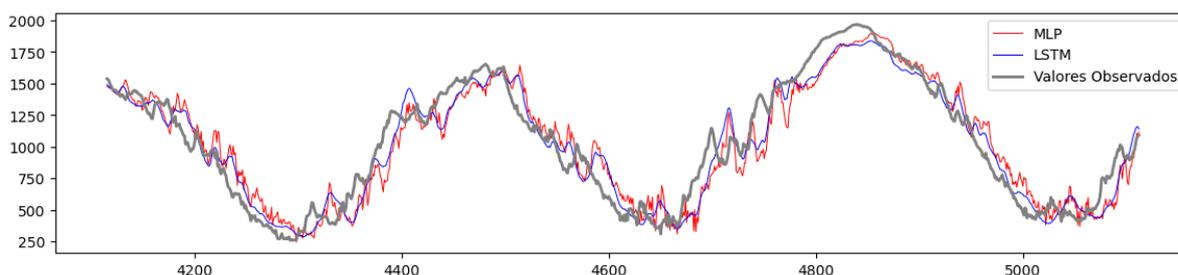
(b)

Figura 7: Previsões dos melhores modelos para a série da Estação de Manicoré.

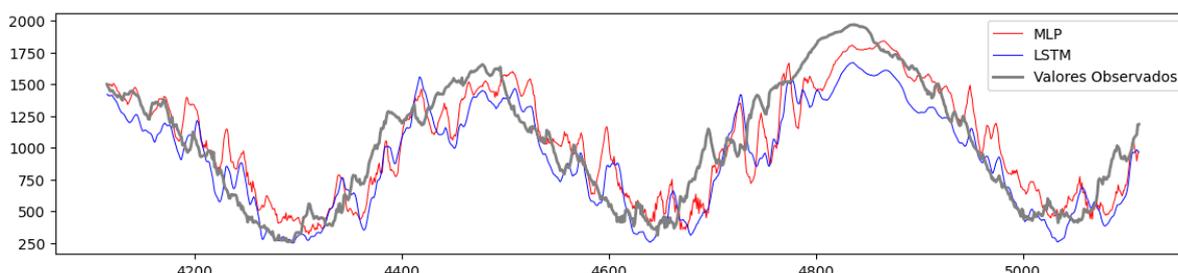
estatísticos ao nível de confiança de 95 %, revela a viabilidade da abordagem autoregressiva de previsão para os dois horizontes temporais considerados e para os dois tipos de modelos de RNAs analisados. Observou-se uma discreta superioridade das RNRs LSTM em algumas situações, a qual pode ser justificada pela melhor capacidade desses modelos em capturar a memória do processo subjacente afim de refleti-la em previsões mais precisas. É importante ressaltar que, embora este estudo não tenha contemplado variáveis preditoras relacionadas a outros fatores que influenciam o nível dos rios, o uso das RNAs demonstrou-se oportuno para avaliar a capacidade de generalização em um problema realístico. Além disso, esses resultados sugerem a aplicação desses modelos em cenários de previsão de séries temporais hidrológicas com disponibilidade limitada de dados.

Tabela 5: Desempenho dos modelos na previsão para a Estação de Porto Velho.

Horizonte	Tipo	Características		R ²					$\mu \pm \sigma$
		Topologia	Ativação	Split 1	Split 2	Split 3	Split 4	Split 5	
$h = 15$	MLP	(5, 3)	Identidade	0,8697	0,9449	0,9464	0,9294	0,9537	0,9288 \pm 0,0306
	MLP	(2, 2)	Identidade	0,8701	0,9334	0,9637	0,8769	0,8647	0,9018 \pm 0,0395
	MLP	(4, 2)	Identidade	0,8812	0,9169	0,9087	0,8760	0,9173	0,9000 \pm 0,0178
	LSTM	(15)	Identidade	0,9382	0,9602	0,9606	0,9674	0,9497	0,9553 \pm 0,0102
	LSTM	(30)	Identidade	0,9329	0,9598	0,9444	0,9654	0,9431	0,9491 \pm 0,0118
	LSTM	(7)	Identidade	0,8873	0,9458	0,9565	0,9659	0,9364	0,9384 \pm 0,0273
$h = 30$	MLP	(5, 2)	ReLU	0,7301	0,7562	0,8983	0,8591	0,8856	0,8259 \pm 0,0692
	MLP	(5, 1)	Identidade	0,6543	0,8596	0,8726	0,8701	0,7896	0,8093 \pm 0,0832
	MLP	(5, 2)	Identidade	0,6334	0,7535	0,8891	0,8643	0,8452	0,7971 \pm 0,0937
	LSTM	(15)	Identidade	0,8704	0,9170	0,9324	0,9022	0,8892	0,9022 \pm 0,0214
	LSTM	(7)	Identidade	0,8003	0,8864	0,9166	0,9012	0,9012	0,8670 \pm 0,0442
	LSTM	(30)	Identidade	0,8268	0,8906	0,8385	0,8960	0,8366	0,8577 \pm 0,0293



(a)



(b)

Figura 8: Previsões dos melhores modelos para a série da Estação Porto Velho.

5. Considerações Finais

Os experimentos realizados no escopo deste trabalho ressaltam uma análise comparativa entre modelos de RNAs MLP e RNRs LSTM para o problema prático de previsão de cota na Bacia do Rio Madeira a partir de séries temporais oriundas de quatro estações hidrológicas localizadas ao longo do seu curso, composta de valores diários coletados entre 2001 e 2014. Os resultados permitiram concluir que ambos os modelos são viáveis na previsão autoregressiva, resultando em boas métricas de desempenho na captura da dinâmica temporal da série. Enquanto para cada estação e horizonte foram avaliadas 70 topologias de RNAs MLP das quais as três melhores foram elencadas, foi efetuado o contraste com três topologias de RNRs LSTM. Embora não seja possível concluir de maneira definitiva que um modelo de RNAs é melhor que o outro, o ônus computacional para percorrer o espaço de busca das RNRs LSTM foi significativamente menor, com melhor desempenho distinguível em alguns dos cenários.

Tabela 6: Análise estatística comparativa do desempenho dos modelos.

Estação	Horizonte	p-valor	Comparativo	Conclusão
Abunã	15	0.5001	$p \not< \alpha$	Não é possível rejeitar H_0
Abunã	30	0.2249	$p \not< \alpha$	Não é possível rejeitar H_0
Humaitá	15	0.0796	$p \not< \alpha$	Não é possível rejeitar H_0
Humaitá	30	0.0796	$p \not< \alpha$	Não é possível rejeitar H_0
Manicoré	15	0.04311	$p < \alpha$	Rejeitar H_0 em favor de H_A
Manicoré	30	0.1380	$p \not< \alpha$	Não é possível rejeitar H_0
Porto Velho	15	0.0796	$p \not< \alpha$	Não é possível rejeitar H_0
Porto Velho	30	0.0431	$p < \alpha$	Rejeitar H_0 em favor de H_A

Essa vantagem é especialmente útil para fins práticos, especialmente em cenários com poucos dados disponíveis.

O trabalho apresentado colabora para previsão de séries temporais hidrológicas na Bacia do Rio Madeira, no qual ainda são escassas as aplicações envolvendo métodos de *Machine Learning*. Em trabalhos futuros pretende-se aumentar a quantidade de arquiteturas de RNRs LSTM a serem avaliadas, bem como incluir outros algoritmos de *Machine Learning* na previsão de tais séries temporais (GRU e *Prophet*, por exemplo). Ademais, almeja-se também considerar a viabilidade da previsão autoregressiva em outros rios da Bacia Amazônica.

6. Agradecimentos

Esse trabalho contou com apoio da Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM) - Código Financeiro: 01.02.016301.00268/2021 e do projeto ACADEMIA STEM financiado pela Samsung Eletrônica da Amazônia Ltda. de acordo com o artigo 48 do Decreto Nº 6.008/2006 no âmbito da Lei Federal Nº 8.387/1991. EBG agradece ao suporte material provido pelo Laboratório de Sistemas Inteligentes da Universidade do Estado do Amazonas.

Referências

- Alpaydin, E. (2010). *Introduction to Machine Learning*. The MIT Press, Cambridge, 2 edition.
- Bergmeir, C. and Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191:192–213.
- Brasil (2022). Empresa de pesquisa energética – balanço energético nacional 2022. Disponível em <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-2022>. Acesso em 25 de agosto de 2023.
- Brasil (2023). Ministério de minas e energia – secretaria de planejamento e desenvolvimento energético – empreendimentos do rio madeira. Disponível em <http://antigo.mme.gov.br/web/guest/secretarias/planejamento-e-desenvolvimento-energetico/publicacoes/empreendimentos-do-rio-madeira>. Acesso em 25 de agosto de 2023.
- Cameron, A. C. and Windmeijer, F. A. (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics*, 77(2):329–342.
- Chicco, D., Warrens, M. J., and Jurman, G. (2021). The coefficient of determination r-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7:623.

- da Silva Cardoso, C. C. (2022). Análise comparativa de redes neurais artificiais para previsão de séries temporais das cotas dos rios solimões e Amazonas. Monografia de Pós-Graduação *Lato Sensu* em Ciência dos Dados – Universidade do Estado do Amazonas.
- de Brito Gomes, W., Satyamurty, P., Correia, F. W. S., Chou, S. C., Vergasta, L. A., and de Arruda Lyra, A. (2022). Intraseasonal scale ensemble forecasts of precipitation and evapotranspiration for the Madeira river basin using different physical parameterizations. *Atmospheric Research*, 270:106086.
- de Melo, A., Sugimoto, G., Nakandakari, D., Tasinaffo, P. M., Santos, A. H. M., Cunha, A. M., and Dias, L. A. V. (2019). A new approach to river flow forecasting: LSTM and GRU multivariate models. *IEEE Lat. Am. Trans.*, 17(12):1978–1986.
- Dehghani, A., Moazam, H. M. Z. H., Mortazavizadeh, F., Ranjbar, V., Mirzaei, M., Mortezaei, S., Ng, J. L., and Dehghani, A. (2023). Comparative evaluation of lstm, cnn, and convlstm for hourly short-term streamflow forecasting using deep learning approaches. *Ecological Informatics*, 75:102119.
- do Rio, G. A. P., Coelho, M. C. N., and Wanderley, L. J. (2015). Rio Madeira: fronteiras, redes, rotas e regiões. *Novos Cadernos NAE*, 18(5):93–109.
- Doria, C. R. d. C., Ruffino, M. L., Hijazi, N. C., and Cruz, R. L. d. (2012). A pesca comercial na bacia do rio Madeira no estado de Rondônia, Amazônia Brasileira. *Acta Amazon.*, 42(1):29–40.
- Haykin, S. (2009). *Neural Networks and Learning Machines*. Pearson Prentice-Hall, Estados Unidos, 3 edition.
- Lazzeri, F. (2021). *Machine learning for time series forecasting with Python*. John Wiley & Sons, Nashville, TN.
- Molinier, M., Guyot, J.-L., Oliveira, E. d., Guimarães, V., and Chaves, A. (1994). Hidrologia da bacia do rio Amazonas. *A Água em Revista*, 2(3):31–36.
- Nielsen, A. (2020). *Practical Time Series Analysis – Prediction with Statistics & Machine Learning*. O’Reilly, Canada, 1 edition.
- Palit, A. K. and Popovic, D. (2005). *Computational Intelligence in Time Series Forecasting - Theory and Engineering Applications*. Springer, Londres, 1 edition.
- Russel, S. and Norvig, P. (2010). *Artificial Intelligence – A Modern Approach*. Prentice Hall, Nova Jersey, 3 edition.
- Santos Neto, L. A. d., Maniesi, V., Querino, C. A. S., Silva, M. J. G. d., and Brown, V. R. (2021). Modelagem hidroclimatológica utilizando redes neurais multi layer perceptron em bacia hidrográfica no sudoeste da Amazônia. *Rev. Bras. Climatol.*, 26.
- Silva, C. G. P. d. and Costa, A. F. (2014). Um quadro histórico das populações indígenas no alto rio Madeira durante o Século XVIII. *Amaz. - Rev. Antropol.*, 6(1):110.
- Walpole, R. E., Myers, R. H., Myers, S. L., and Ye, K. E. (2016). *Probability & Statistics for Engineers & Scientists*. Pearson, Estados Unidos, 9 edition.
- Zounemat-Kermani, M., Mahdavi-Meymand, A., and Hinkelmann, R. (2021). A comprehensive survey on conventional and modern neural networks: application to river flow forecasting. *Earth Sci. Inform.*, 14(2):893–911.