



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TEINFORMÁTICA
CURSO DE GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO

DIEGO MELO DO NASCIMENTO

**UMA ANÁLISE COMPARATIVA ENTRE DIFERENTES PARADIGMAS DE
CONSTRUÇÃO DE MODELOS PREDITIVOS PARA VAZÃO DE RESERVATÓRIOS.**

FORTALEZA

2026

DIEGO MELO DO NASCIMENTO

UMA ANÁLISE COMPARATIVA ENTRE DIFERENTES PARADIGMAS DE
CONSTRUÇÃO DE MODELOS PREDITIVOS PARA VAZÃO DE RESERVATÓRIOS.

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Computação do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia de Computação.

Orientador: Prof. Dr. Guilherme de Alencar Barreto.

FORTALEZA

2026

DIEGO MELO DO NASCIMENTO

UMA ANÁLISE COMPARATIVA ENTRE DIFERENTES PARADIGMAS DE
CONSTRUÇÃO DE MODELOS PREDITIVOS PARA VAZÃO DE RESERVATÓRIOS.

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Computação do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia de Computação.

Aprovada em: 27 de Janeiro de 2026

BANCA EXAMINADORA

Prof. Dr. Guilherme de Alencar Barreto (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Francisco de Assis de Souza Filho
Universidade Federal do Ceará (UFC)

Dr. Renan Bessa
Universidade Federal do Ceará (UFC)

AGRADECIMENTOS

Primeiramente, agradeço a Deus por me permitir chegar até aqui, concedendo-me saúde, força e perseverança.

Ao meu pai, pelos ensinamentos de vida e pelo apoio permanente em minha trajetória pessoal e acadêmica.

À minha mãe, por ser exemplo de perseverança, dedicação e apoio incondicional aos meus objetivos.

Aos meus familiares, pelo suporte constante e pela compreensão nos momentos de ausência.

Aos meus amigos e colegas, pelo apoio, companheirismo e contribuição ao longo de toda a graduação.

Ao Prof. Dr. Guilherme de Alencar Barreto, pelo apoio, orientação e disponibilidade durante o percurso acadêmico.

Ao Dr. Renan Bessa, pela confiança, contribuições técnicas e comprometimento com o desenvolvimento deste trabalho.

Ao Prof. Dr. Francisco de Assis de Souza Filho, pela atenção, disponibilidade e contribuições acadêmicas que fundamentaram este trabalho.

À Universidade Federal do Ceará (UFC), pelo suporte institucional e pela qualidade do ensino oferecido.

A todos os professores da UFC, pela contribuição decisiva à minha formação acadêmica e profissional.

Aos colegas de turma, pelas trocas de conhecimento e experiências compartilhadas.

Aos colegas e amigos de laboratório, pelo apoio, discussões acadêmicas e aprendizado coletivo.

Aos funcionários e colaboradores da UFC, incluindo os terceirizados, pelo profissionalismo e contribuição ao funcionamento da instituição.

A todos que, direta ou indiretamente, contribuíram para minha formação pessoal, acadêmica e profissional, registro meu sincero agradecimento.

RESUMO

Este estudo investiga a previsão de um passo à frente da vazão natural afluyente média mensal em reservatórios hidrelétricos brasileiros, atividade diretamente relacionada ao planejamento e à operação do Sistema Interligado Nacional (SIN). É realizada uma análise comparativa entre três abordagens *data-driven*: um modelo autorregressivo regularizado, uma rede neural recorrente baseada em memória de longo e curto prazo, *long short-term memory* (LSTM), e uma abordagem baseada em *prompt* com grande modelo de linguagem, *large language model* (LLM), especificamente Gemini 2.5 Flash (GEMINI-2.5-F) e Gemini 3 Flash Preview (GEMINI-3-FP). Adota-se um protocolo experimental reprodutível, com validação por janelas deslizantes e divisão cronológica dos dados, aplicado a quatro reservatórios com regimes hidrológicos distintos: Furnas, Sobradinho, Itaipu e Ilha Solteira. O desempenho dos modelos é avaliado na escala original por meio da raiz do erro quadrático médio (RMSE) e do coeficiente de determinação (R^2). Os resultados indicam que não há um método com desempenho superior em todos os reservatórios analisados. A LSTM apresenta menores erros de previsão em Furnas e Sobradinho, enquanto o GEMINI-3-FP obtém melhores resultados em Itaipu e Ilha Solteira. Esses resultados indicam que abordagens baseadas em LLMs, mesmo sem treinamento paramétrico direto sobre séries numéricas, podem produzir previsões consistentes quando o problema é formulado como inferência condicionada a contextos textuais estruturados. O estudo evidencia a relevância de tratar LLMs como uma abordagem complementar à modelagem hidrológica tradicional, cuja aplicação depende das características da série e requer validação empírica cuidadosa.

Palavras-chave: previsão de vazões; séries temporais; modelo autorregressivo; long short-term memory; grande modelo de linguagem.

ABSTRACT

This study investigates one-step-ahead forecasting of mean monthly natural inflow in Brazilian hydroelectric reservoirs, an activity directly related to the planning and operation of the National Interconnected System (SIN). A comparative analysis is conducted among three *data-driven* approaches: a regularized autoregressive model, a recurrent neural network based on long and short-term memory, *long short-term memory* (LSTM), and a *prompt*-based approach using a large language model (LLM), specifically Gemini 2.5 Flash (GEMINI-2.5-F) and Gemini 3 Flash Preview (GEMINI-3-FP). A reproducible experimental protocol is adopted, employing sliding-window validation and chronological data splitting, applied to four reservoirs with distinct hydrological regimes: Furnas, Sobradinho, Itaipu, and Ilha Solteira. Model performance is evaluated on the original scale using the root mean square error (RMSE) and the coefficient of determination (R^2). The results indicate that no single method achieves superior performance across all analyzed reservoirs. The LSTM yields lower forecasting errors in Furnas and Sobradinho, whereas GEMINI-3-FP attains better results in Itaipu and Ilha Solteira. These findings indicate that LLM-based approaches, even without direct parametric training on numerical time series, can produce consistent forecasts when the problem is formulated as inference conditioned on structured textual contexts. The study highlights the relevance of treating LLMs as a complementary approach to traditional hydrological modeling, whose application depends on the characteristics of the time series and requires careful empirical validation.

Palavras-chave: streamflow forecasting; time series; autoregressive model; long short-term memory; large language model.

LISTA DE FIGURAS

Figura 1 – Decomposição clássica do hidrograma de vazão durante um evento de chuva	17
Figura 2 – Esquema conceitual de predição de um passo e multi-passos em séries temporais	25
Figura 3 – Componentes da célula (neurônio) LSTM básico.	27
Figura 4 – Comparação conceitual entre um framework baseado em transformer e o paradigma <i>prompt-based</i> do PromptCast.	32
Figura 5 – Diagrama geral do fluxo metodológico adotado no estudo.	35
Figura 6 – Série temporal, funções de autocorrelação e histograma da vazão natural afluente do reservatório de Furnas.	37
Figura 7 – Série de Furnas após normalização z-score global.	38
Figura 8 – Série de Furnas após normalização z-score mensal.	39
Figura 9 – Série temporal, funções de autocorrelação e histograma da vazão natural afluente do reservatório de Sobradinho.	40
Figura 10 – Série de Sobradinho após normalização z-score global.	41
Figura 11 – Série de Sobradinho após normalização z-score mensal.	42
Figura 12 – Série temporal, funções de autocorrelação e histograma da vazão natural afluente do reservatório de Itaipu.	43
Figura 13 – Série de Itaipu após normalização z-score global.	44
Figura 14 – Série de Itaipu após normalização z-score mensal.	45
Figura 15 – Série temporal, funções de autocorrelação e histograma da vazão natural afluente do reservatório de Ilha Solteira.	46
Figura 16 – Série de Ilha Solteira após normalização z-score global.	47
Figura 17 – Série de Ilha Solteira após normalização z-score mensal.	48
Figura 18 – Visão esquemática do pipeline de execução RS para modelos supervisionados de aprendizado de máquina.	53
Figura 19 – Pipeline metodológica adotada para previsão de vazão média mensal com LLMs.	54

LISTA DE TABELAS

Tabela 1 – Estatísticas descritivas das séries de vazão (1993-2025), em m^3/s	36
Tabela 2 – Espaço de busca dos hiperparâmetros do modelo AR	50
Tabela 3 – Espaço de busca dos hiperparâmetros da arquitetura LSTM	51
Tabela 4 – Parâmetros experimentais utilizados na interação com os LLMs	55
Tabela 5 – Hiperparâmetros selecionados e desempenho no conjunto de validação para o modelo AR.	61
Tabela 6 – Desempenho do modelo AR no conjunto de teste (RMSE em m^3/s e R^2). . .	62
Tabela 7 – Hiperparâmetros selecionados e desempenho no conjunto de validação para o modelo LSTM.	63
Tabela 8 – Desempenho da LSTM no conjunto de teste (RMSE em m^3/s e R^2), agrupado por reservatório.	64
Tabela 9 – Desempenho do GEMINI-2.5-F no conjunto de teste.	64
Tabela 10 – Desempenho do GEMINI-3-FP no conjunto de teste.	65
Tabela 11 – Comparação geral no conjunto de teste entre modelos supervisionados (AR e LSTM) e abordagens baseadas em LLM.	65

LISTA DE ABREVIATURAS E SIGLAS

R^2	coeficiente de determinação
API	<i>application programming interface</i>
AR	modelo autorregressivo
ARIMA	<i>autoregressive integrated moving average</i>
FAC	função de autocorrelação
FACP	função de autocorrelação parcial
GEMINI-2.5-F	Gemini 2.5 Flash
GEMINI-3-FP	Gemini 3 Flash Preview
LLM	<i>large language model</i>
LSTM	<i>long short-term memory</i>
MQO	mínimos quadrados ordinários
MSE	mean squared error
RMSE	<i>root mean squared error</i>
RNN	<i>recurrent neural network</i>
RS	<i>random search</i>
SAR	Sistema de Acompanhamento de Reservatórios
SIN	Sistema Interligado Nacional

LISTA DE SÍMBOLOS

$\hat{Q}_{t+h t}$	Vazão prevista para o instante $t + h$ condicionada à informação até t (m^3/s)
t	Instante temporal discreto
h	Horizonte de previsão
x_t	Observação da série temporal no instante t
N	Número total de observações da série
\bar{x}	Média amostral da série
μ	Média teórica do processo estocástico
σ^2	Variância do processo
σ	Desvio padrão do processo
$\gamma(k)$	Autocovariância no atraso k
$\rho(k)$	Autocorrelação no atraso k
$\alpha(k)$	Autocorrelação parcial no atraso k
k	Defasagem temporal
p	Ordem do modelo autorregressivo
ϕ_i	Coefficiente autorregressivo de ordem i
ε_t	Termo de erro aleatório no instante t
λ	Parâmetro de regularização
$c(t)$	Estado da célula de memória
$i(t)$	Porta de entrada
$f(t)$	Porta de esquecimento
$o(t)$	Porta de saída
$\sigma(\cdot)$	Função sigmoide
$g(\cdot)$	Função de ativação hiperbólica
\odot	Produto elemento a elemento
y_i	Valor observado
\hat{y}_i	Valor estimado pelo modelo

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Justificativa	14
1.2	Objetivos	15
<i>1.2.1</i>	<i>Objetivo geral</i>	<i>15</i>
<i>1.2.2</i>	<i>Objetivos específicos</i>	<i>15</i>
1.3	Organização do trabalho	15
2	O PROBLEMA DA PREDIÇÃO DE VAZÃO	16
2.1	Caracterização do problema	16
2.2	Séries temporais univariadas de vazão	16
<i>2.2.1</i>	<i>Propriedades estatísticas das séries de vazão</i>	<i>18</i>
<i>2.2.2</i>	<i>Dependência temporal e estruturas de correlação</i>	<i>19</i>
2.3	Modelagem linear autorregressiva	21
<i>2.3.1</i>	<i>Formulação do modelo</i>	<i>22</i>
<i>2.3.2</i>	<i>Seleção da ordem autorregressiva</i>	<i>23</i>
<i>2.3.3</i>	<i>Regularização de cumeeira (L2)</i>	<i>23</i>
<i>2.3.4</i>	<i>Predição de um passo à frente</i>	<i>24</i>
3	MODELO PREDITIVO BASEADO EM LSTM	27
3.1	Fundamentos da arquitetura	27
3.2	Mecanismo de memória	28
3.3	Capacidade de modelagem de dependências temporais	29
3.4	Uso em previsão de vazão	29
4	MODELOS DE LINGUAGEM DE GRANDE ESCALA NA PREVISÃO DE SÉRIES TEMPORAIS	31
4.1	LLMs como preditores de próximo token	31
4.2	Aprendizado em contexto para previsão temporal	32
4.3	Desafios teóricos na previsão com LLMs	33
5	METODOLOGIA	35
5.1	Conjunto de dados	36
<i>5.1.1</i>	<i>Reservatório de Furnas</i>	<i>37</i>
<i>5.1.1.1</i>	<i>Normalização global da série de Furnas</i>	<i>38</i>

5.1.1.2	<i>Normalização mensal da série de Furnas</i>	39
5.1.2	Reservatório de Sobradinho	40
5.1.2.1	<i>Normalização global da série de Sobradinho</i>	41
5.1.2.2	<i>Normalização mensal da série de Sobradinho</i>	42
5.1.3	Reservatório de Itaipu	43
5.1.3.1	<i>Normalização global da série de Itaipu</i>	44
5.1.3.2	<i>Normalização mensal da série de Itaipu</i>	45
5.1.4	Reservatório de Ilha Solteira	46
5.1.4.1	<i>Normalização global da série de Ilha Solteira</i>	47
5.1.4.2	<i>Normalização mensal da série de Ilha Solteira</i>	48
5.2	Modelos supervisionados de aprendizado de máquina	49
5.2.1	<i>Ambiente computacional e reprodutibilidade</i>	49
5.2.2	<i>Implementação do modelo AR</i>	49
5.2.3	<i>Implementação da rede neural LSTM</i>	50
5.3	Pré-processamento e estrutura experimental	51
5.3.1	<i>Estruturação temporal e janelamento</i>	52
5.3.2	<i>Protocolos de normalização</i>	52
5.3.3	<i>Pipeline de execução e otimização</i>	53
5.4	Modelos de linguagem de grande escala	53
5.4.1	<i>Configuração dos parâmetros dos modelos</i>	54
5.4.2	<i>Estruturação temporal e definição das janelas</i>	55
5.4.3	<i>Extração de estatísticas hidrológicas</i>	56
5.4.4	<i>Formulação dos prompts e interação com o LLM</i>	56
5.4.5	<i>Execução do pipeline preditivo</i>	56
5.4.6	<i>Template do prompt utilizado</i>	57
5.5	Métricas de avaliação de desempenho	59
5.5.1	<i>Raiz do erro quadrático médio</i>	59
5.5.2	<i>Coefficiente de determinação (R^2)</i>	59
6	RESULTADOS E DISCUSSÕES	61
6.1	Resultados do modelo autoregressivo	61
6.1.1	<i>Hiperparâmetros selecionados</i>	61
6.1.2	<i>Desempenho no conjunto de teste</i>	62

6.2	Resultados do modelo LSTM	62
6.2.1	<i>Hiperparâmetros selecionados</i>	63
6.2.2	<i>Desempenho no conjunto de teste</i>	63
6.3	Resultados da abordagem baseada em LLM	64
6.3.1	<i>Gemini 2.5 Flash</i>	64
6.3.2	<i>Gemini 3 Flash Preview</i>	64
6.4	Comparação geral	65
7	CONSIDERAÇÕES FINAIS	66
	REFERÊNCIAS	68

1 INTRODUÇÃO

A predição de vazões naturais afluentes em reservatórios hidrelétricos é fundamental para o planejamento e a operação de sistemas hídricos e de geração de energia, pois a vazão sintetiza os efeitos integrados de armazenamento e liberação em bacias hidrográficas (Chow *et al.*, 1988). Séries hidrológicas mensais apresentam persistência, sazonalidade e variabilidade interanual, o que torna a previsão sensível ao horizonte adotado, ao pré-processamento dos dados e à escolha do modelo (Hipel e Mcleod, 1994). Adicionalmente, não estacionariedades associadas à variabilidade climática e a intervenções antrópicas limitam a validade da hipótese clássica de estacionariedade, exigindo avaliação empírica cuidadosa e protocolos reprodutíveis (Milly *et al.*, 2008).

Do ponto de vista metodológico, a previsão de vazões pode ser realizada por modelos estatísticos lineares, redes neurais e, mais recentemente, por grande modelo de linguagem, *large language model* (LLM). A abordagem Box Jenkins oferece um arcabouço consolidado para representar dependência temporal por meio de defasagens e ruído de inovação, preservando interpretabilidade e eficiência computacional (Box e Jenkins, 1968). Em contraste, redes neurais recorrentes do tipo memória de curto longo prazo, *long short-term memory* (LSTM) incorporam mecanismos explícitos de memória, ampliando a capacidade de modelar dependências temporais longas e relações não lineares (Hochreiter e Schmidhuber, 1997; Houdt *et al.*, 2020). Paralelamente, estudos recentes investigam o uso de LLMs baseadas na reformulação do problema temporal como uma tarefa de geração condicionada a contexto textual (Gruver *et al.*, 2023; Xue e Salim, 2022; Jin *et al.*, 2023).

Neste trabalho, investiga-se em que medida um modelo autorregressivo (AR) regularizado, uma rede LSTM e uma abordagem baseada em LLM diferem em desempenho e robustez na previsão de vazão natural afluente média mensal, considerando reservatórios com regimes hidrológicos distintos. Para isso, analisam-se séries mensais de vazão natural afluente de reservatórios relevantes do sistema elétrico brasileiro, utilizando dados públicos da Agência Nacional de Águas e Saneamento Básico (2025), provenientes do Sistema de Acompanhamento de Reservatórios (SAR) do Sistema Interligado Nacional (SIN). Adota-se um horizonte de previsão de um passo à frente, permitindo comparação direta entre métodos por meio da raiz do erro quadrático médio, *root mean squared error* (RMSE), e do coeficiente de determinação (R^2).

A estratégia experimental segue uma abordagem comparativa estruturada. No eixo supervisionado clássico, emprega-se um modelo AR com regularização de cumeeira, alinhado à metodologia Box Jenkins, como forma de mitigar a multicolinearidade entre defasagens e preservar parcimônia paramétrica (Box e Jenkins, 1968). Em paralelo, utiliza-se uma rede LSTM, com configuração definida a partir de busca sistemática sobre tamanho de janela temporal, número de unidades ocultas, profundidade da arquitetura e termos de regularização, explorando sua capacidade de modelar dependências temporais não lineares e de maior alcance (Hochreiter e Schmidhuber, 1997; Houdt *et al.*, 2020). No eixo baseado em LLMs, avaliam-se modelos da família Gemini em regime de inferência direta, sem ajuste paramétrico sobre os dados numéricos, mas com incorporação explícita de estatísticas descritivas e histórico recente da série por meio de *prompting* estruturado, conforme abordagens recentes que reformulam a previsão temporal como geração condicionada a contexto textual (Gruver *et al.*, 2023; Xue e Salim, 2022; Jin *et al.*, 2023). Os experimentos são implementados em ambiente *Python*, com apoio das bibliotecas *Scikit-Learn* e *TensorFlow/Keras* (Pedregosa *et al.*, 2012; Abadi *et al.*, 2016).

1.1 Justificativa

A relevância científica deste estudo decorre da complexidade inerente às séries hidrológicas e da necessidade contínua de comparações rigorosas entre modelos lineares e não lineares (Box e Jenkins, 1968; Hochreiter e Schmidhuber, 1997). Evidências recentes indicam ganhos no uso de LSTM em aplicações hidrológicas, mas também apontam sensibilidade ao regime e ao desenho experimental, justificando análises comparativas específicas por reservatório (Kratzert *et al.*, 2018; Filho *et al.*, 2022). Além disso, o emprego de LLMs como previsores *zero-shot* representa um paradigma recente, cuja aplicabilidade a séries físicas requer avaliação criteriosa (Gruver *et al.*, 2023).

Sob o ponto de vista aplicado, os reservatórios analisados são estratégicos para o sistema elétrico e para a gestão de riscos hidrológicos. Assim, ganhos incrementais de acurácia e robustez podem apoiar decisões operacionais, mesmo em horizontes de curto prazo (Chow *et al.*, 1988; Nash e Sutcliffe, 1970). A padronização de procedimentos de aquisição, processamento e avaliação também contribui como base reprodutível para extensões futuras (Hipel e Mcleod, 1994; Milly *et al.*, 2008).

1.2 Objetivos

1.2.1 *Objetivo geral*

Avaliar comparativamente o desempenho preditivo de modelos AR, LSTM e abordagem baseada em LLM na previsão da vazão em reservatórios brasileiros com regimes hidrológicos distintos, afim de investigar em que condições cada abordagem se mostra mais adequada.

1.2.2 *Objetivos específicos*

1. Caracterizar séries mensais de vazão quanto à dependência temporal e sazonalidade.
2. Ajustar um modelo AR como linha de base estatística.
3. Ajustar uma rede LSTM para previsão de um passo à frente.
4. Aplicar LLMs como preditores indiretos.
5. Comparar o desempenho dos modelos no conjunto de teste por meio das métricas RMSE e R^2 , analisando a adequação de cada abordagem em diferentes regimes hidrológicos.

1.3 Organização do trabalho

Este capítulo apresenta a contextualização, a delimitação do problema, a justificativa, os objetivos e a estrutura do trabalho. O Capítulo 2 discute fundamentos de séries temporais aplicados à predição de vazão. O Capítulo 3 descreve a arquitetura LSTM e sua aplicação em séries hidrológicas. O Capítulo 4 aborda o uso de LLMs em previsão de séries temporais. O Capítulo 5 detalha dados, metodologia experimental e métricas. O Capítulo 6 apresenta e discute os resultados. Por fim, o Capítulo 7 sintetiza as conclusões e indica trabalhos futuros.

2 O PROBLEMA DA PREDIÇÃO DE VAZÃO

Este capítulo apresenta o problema de predição de vazão natural afluyente, delimitando o enquadramento estatístico e metodológico adotado para a modelagem de séries temporais hidrológicas.

2.1 Caracterização do problema

A predição de vazão de reservatórios constitui um problema central da hidrologia aplicada, uma vez que a vazão representa a resposta integrada de uma bacia hidrográfica às condições climáticas e fisiográficas ao longo do tempo, sintetizando, em escala temporal discreta, os efeitos de armazenamento, liberação e transferência de água no sistema de drenagem, o que a torna a principal grandeza para planejamento, operação de reservatórios e avaliação de riscos hidrológicos (Chow *et al.*, 1988). Do ponto de vista da modelagem, esse problema envolve a representação da transformação de processos físicos complexos, como precipitação e evaporação, em escoamento fluvial sob condições de contorno dificilmente observáveis, caracterizando-se como intrinsecamente estocástico e dependente da memória do sistema, o que justifica o emprego de modelos conceituais e estatísticos baseados na análise de séries temporais históricas de vazão (Nash e Sutcliffe, 1970).

2.2 Séries temporais univariadas de vazão

Uma série temporal univariada pode ser definida como uma sequência ordenada de observações de uma única variável ao longo do tempo, coletadas em intervalos regulares. No contexto hidrológico, essa variável corresponde tipicamente à vazão natural afluyente de um rio ou reservatório, expressa em unidades de volume por tempo, como m^3/s . A característica fundamental desse tipo de dado é que a ordem temporal das observações carrega informação essencial sobre o processo físico subjacente, de modo que a permutação dos valores invalida a interpretação estatística da série (Chow *et al.*, 1988).

Formalmente, uma série temporal univariada de vazão pode ser representada como a sequência discreta $\{x_t\}_{t=1}^N$, em que $x_t \in \mathbb{R}$ denota a vazão observada no instante t e N representa o número total de observações disponíveis. Ao interpretar $\{x_t\}$ como uma realização amostral de um processo estocástico, a caracterização estatística do comportamento do sistema pode ser conduzida por medidas no domínio do tempo, como média, variância e correlação serial,

estimadas a partir da realização observada. Para uma série temporal, estimadores usuais da média e da variância amostrais são dados por

$$\bar{x} = \frac{1}{N} \sum_{t=1}^N x_t \quad (2.1)$$

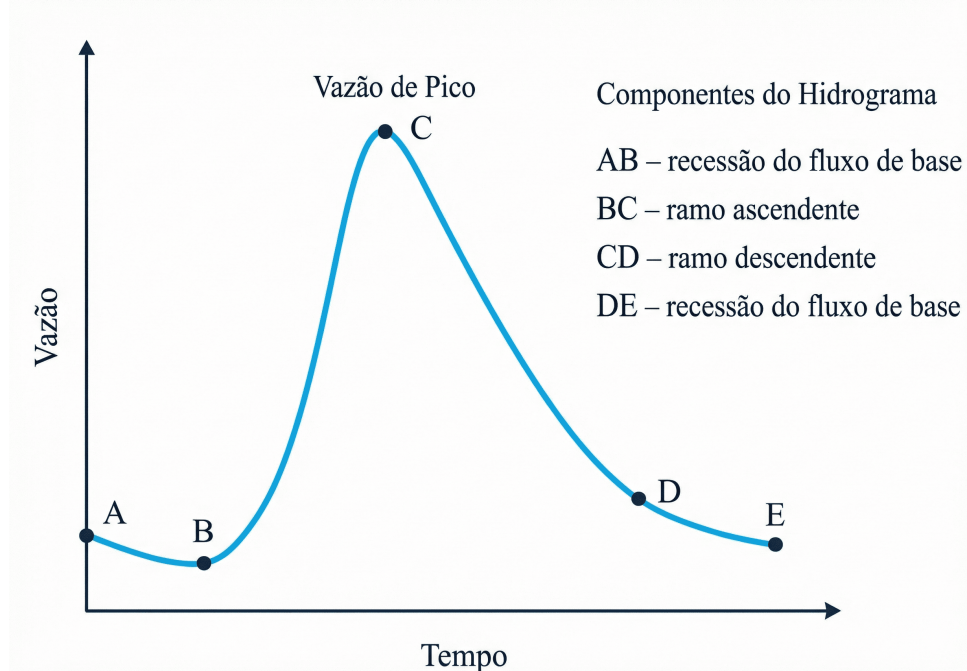
e (2.2)

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{t=1}^N (x_t - \bar{x})^2, \quad (2.3)$$

sendo \bar{x} a média amostral e $\hat{\sigma}^2$ a variância amostral, as quais constituem estatísticas fundamentais para descrever o nível médio e a dispersão da série (Hipel e Mcleod, 1994).

A Figura 1 apresenta a decomposição clássica do hidrograma de vazão durante um evento de chuva, evidenciando as principais componentes da resposta hidrológica de uma bacia. Conforme sistematizado por (Chow *et al.*, 1988), o hidrograma pode ser dividido em ramos bem definidos: o trecho AB representa a recessão do escoamento de base pré-evento; o segmento BC corresponde ao ramo ascendente, associado à rápida resposta do sistema à precipitação efetiva; o ponto C indica a vazão de pico; o trecho CD caracteriza o ramo descendente, no qual cessam gradualmente as contribuições diretas da chuva; e o segmento DE representa novamente a recessão do escoamento de base, quando a vazão passa a ser sustentada predominantemente pelos armazenamentos subterrâneos.

Figura 1 – Decomposição clássica do hidrograma de vazão durante um evento de chuva



Fonte: Adaptado de (Chow *et al.*, 1988).

A fase de recessão do escoamento de base é descrita por meio da curva normal de depleção, ou curva mestra de recessão, obtida pela sobreposição de múltiplas curvas de recessão observadas em um mesmo curso d'água, conforme descrito por Chow *et al.* (1988). Essa curva é frequentemente representada por uma função exponencial do tipo

$$x_t = x_0 e^{-(t-t_0)/k}, \quad (2.4)$$

na qual x_0 representa a vazão no instante inicial da recessão e k é uma constante positiva de decaimento, com dimensão temporal, relacionada às propriedades de armazenamento e drenagem da bacia. Tal formulação decorre do conceito de reservatório linear, em que a vazão de saída s_t é proporcional ao volume armazenado no sistema, expresso por

$$s_t = kx_t. \quad (2.5)$$

A identificação dos trechos do hidrograma coincidentes com a curva de recessão permite separar o escoamento direto do escoamento de base, fundamentando métodos clássicos de separação de *baseflow* amplamente utilizados em análises chuva-vazão.

2.2.1 Propriedades estatísticas das séries de vazão

A estacionariedade pode ser interpretada como uma condição de equilíbrio estatístico, na qual as propriedades probabilísticas do processo não dependem explicitamente do tempo. Em termos formais, a estacionariedade estrita exige invariância da distribuição conjunta a deslocamentos temporais; contudo, em aplicações de modelagem, é comum adotar a estacionariedade fraca (ou estacionariedade no sentido amplo), na qual se requer constância da média e da variância ao longo do tempo e dependência da estrutura de covariância apenas da defasagem (Hipel e Mcleod, 1994). Assim, a série é estacionária em segunda ordem se

$$\mathbb{E}[x_t] = \mu, \quad \forall t, \quad (2.6)$$

$$\text{Var}(x_t) = \sigma^2, \quad \forall t, \quad (2.7)$$

e se a autocovariância teórica no atraso k é definida por

$$\gamma(k) = \text{Cov}(x_t, x_{t-k}) = \mathbb{E}[(x_t - \mu)(x_{t-k} - \mu)], \quad (2.8)$$

dependendo apenas de k .

Nessas expressões, x_t representa a vazão natural afluyente média observada no instante discreto t ; $\mathbb{E}[\cdot]$ denota o operador esperança matemática; μ corresponde ao valor médio constante

da série ao longo do tempo; $\text{Var}(\cdot)$ indica a variância do processo; σ^2 é a variância constante associada à série estacionária; $\text{Cov}(\cdot, \cdot)$ representa o operador de covariância; $\gamma(k)$ é a função de autocovariância teórica no atraso k ; e k denota a defasagem temporal inteira entre duas observações da série.

Em séries hidrológicas mensais, a sazonalidade anual introduz variações sistemáticas do nível médio entre os meses, de modo que a estacionariedade pode ser adotada apenas como uma aproximação válida em subconjuntos sazonais, com médias aproximadamente constantes dentro de cada mês (Hipel e Mcleod, 1994). A persistência, por sua vez, expressa a influência do passado sobre o presente por meio de correlações seriais associadas aos mecanismos de armazenamento do sistema hidrológico (Hipel e Mcleod, 1994). Ademais, mudanças estruturais e tendências decorrentes de alterações climáticas ou intervenções antrópicas podem comprometer essa hipótese, justificando seu emprego restrito a janelas temporais sob condições relativamente estáveis (Milly *et al.*, 2008).

2.2.2 Dependência temporal e estruturas de correlação

A dependência temporal em séries de vazão manifesta-se, no domínio do tempo, por meio de medidas que quantificam a associação estatística entre observações separadas por diferentes defasagens. Nesse contexto, a análise conjunta da função de autocorrelação e da função de autocorrelação parcial constitui um instrumento central para a caracterização da memória do processo e para a identificação de estruturas autorregressivas relevantes em aplicações hidrológicas.

A função de autocorrelação (FAC) teórica é definida como a autocovariância normalizada do processo estacionário, expressa por

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} = \frac{\text{Cov}(x_t, x_{t-k})}{\sigma^2}, \quad -1 \leq \rho(k) \leq 1, \quad (2.9)$$

em que $\gamma(0) = \sigma^2$ representa a variância do processo. Na prática, a FAC é estimada a partir de uma realização finita da série. Considerando uma série temporal com média amostral \bar{x} e comprimento N , a autocovariância amostral no atraso k é dada por

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (x_t - \bar{x})(x_{t+k} - \bar{x}), \quad k = 0, 1, \dots, N-1, \quad (2.10)$$

e a correspondente autocorrelação amostral é obtida por

$$r_k = \frac{c_k}{c_0}. \quad (2.11)$$

A inspeção do decaimento de r_k ao longo das defasagens permite avaliar o alcance da dependência temporal e fornece subsídios para a escolha da ordem de modelos autorregressivos. Em séries hidrológicas mensais, é prática consolidada restringir a análise gráfica a defasagens positivas até aproximadamente $N/4$, de modo a reduzir efeitos espúrios associados ao tamanho finito da amostra. Sob hipóteses assintóticas, limites de confiança aproximados para r_k podem ser construídos com base na variância estimada da FAC amostral, conforme aproximações clássicas da literatura (Hipel e McLeod, 1994).

Complementarmente, a função de autocorrelação parcial (FACP) desempenha papel fundamental na identificação direta da estrutura autorregressiva do processo, ao isolar a correlação entre x_t e x_{t-k} após a remoção do efeito linear das defasagens intermediárias. Formalmente, a FACP no atraso k , denotada por ϕ_{kk} , é definida como o último coeficiente do modelo autorregressivo de ordem k ajustado à série, isto é,

$$x_t = \sum_{j=1}^k \phi_{kj} x_{t-j} + a_t, \quad (2.12)$$

em que a_t representa um termo de inovação não correlacionado. Nesse contexto, a FACP é dada por

$$\alpha(k) = \phi_{kk}, \quad (2.13)$$

correspondendo à correlação parcial entre x_t e x_{t-k} condicionada às observações intermediárias $x_{t-1}, \dots, x_{t-k+1}$.

Do ponto de vista computacional, os coeficientes ϕ_{kk} podem ser obtidos recursivamente por meio das equações de Yule–Walker, resolvidas pelo algoritmo de Levinson–Durbin, o que confere eficiência numérica à estimação da FACP mesmo para séries longas. Em termos operacionais, a FACP fornece uma indicação mais clara do número de defasagens relevantes, sendo amplamente utilizada na seleção de entradas em modelos estatísticos e em arquiteturas de aprendizado de máquina aplicadas à previsão de vazões. Estudos recentes empregam de forma combinada FAC e FACP para definir janelas temporais e lags informativos em modelos baseados em redes recorrentes e abordagens híbridas, evidenciando sua utilidade tanto em contextos lineares quanto não lineares (Ayana *et al.*, 2023; Iqbal e Siddiqi, 2025; Workneh e Jha, 2025).

A distinção entre processos de memória curta e de memória longa pode ser formalizada a partir da somabilidade da FAC teórica. Um processo estacionário em covariância

apresenta memória curta quando

$$M = \sum_{k=-\infty}^{\infty} |\rho(k)| < \infty, \quad (2.14)$$

ao passo que a violação dessa condição caracteriza dependência de longo alcance, frequentemente observada em séries hidrológicas influenciadas por persistência climática e variabilidade interanual.

Por fim, a adoção de modelos lineares estacionários para a descrição da dependência temporal é teoricamente sustentada por resultados de representação, em particular pela decomposição de Wold. Segundo esse resultado, um processo estacionário pode ser decomposto em um componente determinístico e em um componente puramente não determinístico, gerado por inovações não correlacionadas, admitindo a representação

$$x_t = \mu_t + a_t + \sum_{i=1}^{\infty} \psi_i a_{t-i}, \quad (2.15)$$

em que a_t denota a sequência de inovações não correlacionadas. Essa formulação fornece a base conceitual para o uso de estruturas autorregressivas e de média móvel na modelagem e previsão de séries de vazão (Hipel e Mcleod, 1994; Wold, 1939).

2.3 Modelagem linear autorregressiva

Em séries temporais hidrológicas, como a vazão natural afluente, a predição de curto prazo é comumente fundamentada na persistência do processo, isto é, na dependência estatística entre observações passadas e futuras. Na abordagem clássica de modelagem estocástica, essa dependência é representada por modelos autorregressivos e, de forma mais geral, por modelos autorregressivos integrados de médias móveis, nos quais a diferenciação é empregada para induzir estacionariedade quando necessário. No âmbito da modelagem estocástica clássica, essa dependência é representada por modelos autorregressivos e, de forma mais geral, pelo modelo autorregressivo integrado de médias móveis, *autoregressive integrated moving average* (ARIMA). Formalmente, um modelo $ARIMA(p, d, q)$ combina um componente autorregressivo de ordem p , um componente de médias móveis de ordem q e um operador de diferenciação de ordem d , responsável por remover tendências estocásticas da série original, constituindo uma classe de modelos amplamente utilizada como referência metodológica na previsão de vazões mensais e como base comparativa em estudos hidrológicos contemporâneos (Mamudu *et al.*, 2025; Hammond *et al.*, 2025; Filho *et al.*, 2022).

2.3.1 Formulação do modelo

Considere a série de vazões $\{x_t\}_{t=1}^N$, amostrada em intervalos regulares. A formulação autorregressiva de ordem p , $\text{AR}(p)$, assume que a melhor aproximação linear de x_t pode ser construída como combinação linear das p vazões imediatamente anteriores. Essa formulação é obtida como caso particular do modelo ARIMA quando $d = 0$ (isto é, não se aplica diferenciação para induzir estacionariedade) e $q = 0$, de modo que a dinâmica é descrita apenas por um polinômio autorregressivo (Box e Jenkins, 1968). Assim, escreve-se

$$x_t = c + \sum_{i=1}^p \phi_i x_{t-i} + \varepsilon_t, \quad (2.16)$$

em que c é um termo de nível, ϕ_i são os coeficientes autorregressivos e ε_t é um ruído não correlacionado, coerente com a suposição de que a série pode ser vista como resultado de um filtro linear excitado por ruído branco (Box e Jenkins, 1968).

Para explicitar o procedimento de estimação no contexto de predição de vazão, reorganiza-se a Equação 2.16 em forma matricial. Define-se o vetor de respostas

$$\mathbf{y} = [x_{p+1}, x_{p+2}, \dots, x_N]^\top,$$

e a matriz de regressores $\mathbf{X} \in \mathbb{R}^{(N-p) \times (p+1)}$, cuja primeira coluna é composta por termos unitários associados ao coeficiente constante c , e cujas demais colunas correspondem às defasagens da série. De forma explícita,

$$\mathbf{X} = \begin{bmatrix} 1 & x_p & x_{p-1} & \cdots & x_1 \\ 1 & x_{p+1} & x_p & \cdots & x_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N-1} & x_{N-2} & \cdots & x_{N-p} \end{bmatrix},$$

em que cada linha está associada a um instante temporal $t = p + 1, \dots, N$, e as colunas de defasagens estão ordenadas do atraso mais recente (x_{t-1}) ao mais distante (x_{t-p}).

Com essas definições, o modelo $\text{AR}(p)$ pode ser expresso de maneira compacta como

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\beta} = [c, \phi_1, \dots, \phi_p]^\top, \quad (2.17)$$

em que $\boldsymbol{\varepsilon} = [\varepsilon_{p+1}, \dots, \varepsilon_N]^\top$ é o vetor de resíduos.

A forma matricial 2.17 torna explícito que um $\text{AR}(p)$ é um modelo de regressão linear com variáveis explicativas altamente estruturadas (as próprias defasagens da série de vazão)

conectando diretamente a modelagem autorregressiva às questões de seleção de defasagens, multicolinearidade e regularização discutidas nas seções seguintes.

2.3.2 Seleção da ordem autorregressiva

A seleção de p controla simultaneamente a capacidade de capturar memória hidrológica e o risco de superajuste. No formalismo Box Jenkins, a escolha de ordens integra o ciclo de *identificação, estimação e verificação* do modelo, no qual estruturas de dependência temporal são testadas, parâmetros são ajustados e resíduos são avaliados quanto à ausência de autocorrelação remanescente (Box e Jenkins, 1968). Em vazões, esse passo deve refletir a escala temporal relevante do fenômeno (persistência diária, semanal ou sazonal) e a finalidade preditiva (horizonte $t + h$), evitando que p seja inflacionado sem evidência estatística.

Um procedimento operacional e passo a passo, consistente com a prática em séries hidrológicas e com abordagens de seleção de atrasos em modelos preditivos, pode ser descrito como: (i) definir um conjunto candidato $\mathcal{P} = \{1, \dots, P_{\max}\}$; (ii) para cada $p \in \mathcal{P}$, ajustar 2.17 no conjunto de treino; (iii) avaliar desempenho preditivo em validação e, em paralelo, verificar a parcimônia do modelo; (iv) escolher p que maximize desempenho fora da amostra e preserve estabilidade numérica. Em trabalhos recentes, a própria seleção de atrasos é tratada como etapa explícita de *feature selection*, na qual se busca detectar defasagens estatisticamente relevantes antes de expandir o conjunto de entradas, o que reduz dimensionalidade e custo computacional (Samadi-Koucheksaraee e Chu, 2024). Em linha semelhante, ao discutir seleção de entradas, menciona-se que técnicas clássicas incluem auto-correlação e correlação, entre outras, com o objetivo de evitar sobrecarga de parâmetros e degradação de desempenho (Kordani *et al.*, 2025).

2.3.3 Regularização de cumeeira (L2)

A regularização de cumeeira (*Ridge*), ou penalização L2, é uma estratégia clássica de estabilização de modelos lineares em presença de multicolinearidade, fundamentada na redução da variância dos estimadores por meio da introdução de viés controlado (Hoerl e Kennard, 1970; Hoerl, 2020). Em modelos autorregressivos, essa condição ocorre naturalmente devido à forte correlação entre defasagens sucessivas.

Partindo do estimador de mínimos quadrados ordinários (MQO),

$$\hat{\beta}_{\text{MQO}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}, \quad (2.18)$$

observa-se que o mau condicionamento de $\mathbf{X}^\top \mathbf{X}$ compromete a estabilidade dos coeficientes. A Regressão com regularização de cumeira reformula o problema como uma minimização penalizada,

$$\hat{\beta}_{\text{Ridge}} = \arg \min_{\beta} \{ \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda \|\beta\|_2^2 \}, \quad (2.19)$$

cuja solução fechada é dada por

$$\hat{\beta}_{\text{Ridge}} = (\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^\top \mathbf{y}. \quad (2.20)$$

A penalização L2 desloca o espectro da matriz normal, reduzindo a sensibilidade a flutuações amostrais e permitindo estimativas mais estáveis, com menor erro quadrático médio em comparação ao MQO sob colinearidade (Hoerl e Kennard, 1970). A formulação original também introduz o conceito de *ridge trace*, que evidencia a estabilização progressiva dos coeficientes à medida que λ aumenta, reforçando o caráter analítico do método (Hoerl, 2020).

No contexto da previsão de vazões com modelos $\text{AR}(p)$, a regularização de cumeira atua como mecanismo de controle da complexidade paramétrica, evitando coeficientes extremos e preservando a robustez preditiva em séries hidrológicas altamente correlacionadas.

2.3.4 *Predição de um passo à frente*

A predição de um passo à frente consiste na estimação da vazão futura x_{t+1} condicionada exclusivamente ao conjunto de informações disponíveis até o instante t . No contexto de modelos $\text{AR}(p)$, esse regime de previsão corresponde à inferência condicional imediata, na qual apenas observações efetivamente conhecidas da série histórica são utilizadas como entradas do modelo.

Matematicamente, a predição de um passo à frente decorre diretamente da formulação autorregressiva apresentada na Equação 2.16. Ao tomar a esperança condicional dessa equação dado o conjunto de informações

$$\mathcal{F}_t = \{x_t, x_{t-1}, \dots, x_{t-p+1}\},$$

obtém-se a previsão pontual

$$\hat{x}_{t+1|t} = \mathbb{E}[x_{t+1} | \mathcal{F}_t] = c + \sum_{i=1}^p \phi_i x_{t+1-i}, \quad (2.21)$$

uma vez que o termo de inovação ε_{t+1} associado ao instante futuro possui esperança condicional nula e, portanto, não contribui para a previsão (Box e Jenkins, 1968). Todos os outros símbolos empregados em 2.21 já foram previamente definidos na formulação do modelo $\text{AR}(p)$ em 2.16.

A Equação 2.21 evidencia que a predição de um passo à frente utiliza exclusivamente valores observados da série temporal como variáveis explicativas, configurando um problema estatisticamente bem condicionado. Nesse regime, não ocorre realimentação de valores previstos no conjunto de entradas, o que reduz a propagação de erros e torna esse tipo de previsão particularmente adequado para avaliação empírica e comparação de modelos.

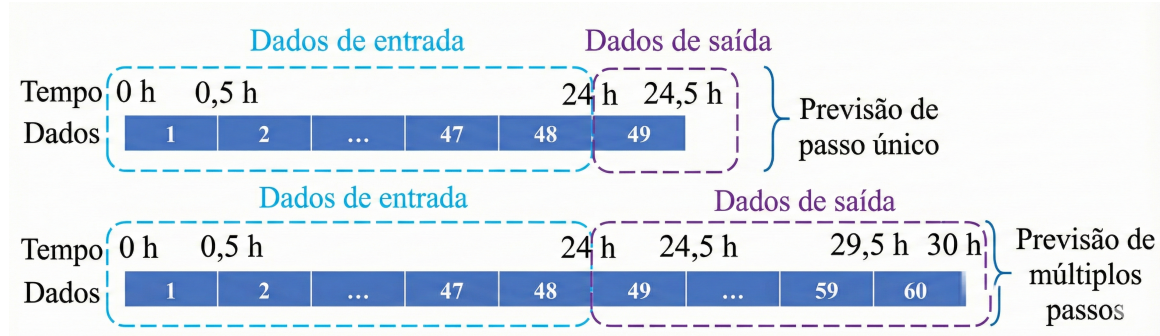
A distinção fundamental entre a predição de um passo à frente e a predição multi-passos não reside na estrutura do modelo autorregressivo, mas no regime operacional de aplicação da Equação 2.16. No caso de um passo à frente, todas as defasagens correspondem a observações disponíveis. Em contraste, na predição multi-passos, previsões previamente geradas passam a ser reutilizadas como entradas para horizontes futuros. Para $h > 1$, a previsão assume a forma recursiva

$$\hat{x}_{t+h|t} = c + \sum_{i=1}^p \phi_i \hat{x}_{t+h-i|t}, \quad (2.22)$$

caracterizando um processo de simulação dinâmica no qual erros de estimação e incertezas estruturais tendem a se propagar e a se acumular à medida que o horizonte temporal se estende.

A Figura 2 ilustra essa distinção de maneira conceitual. Enquanto a predição de um passo à frente envolve uma única inferência condicionada ao histórico observado, a predição multi-passos corresponde a uma sequência de inferências encadeadas, nas quais cada nova estimativa depende, parcial ou totalmente, de valores previamente previstos, ampliando a sensibilidade a erros e instabilidades do modelo (Liao *et al.*, 2022).

Figura 2 – Esquema conceitual de predição de um passo e multi-passos em séries temporais



Fonte: Adaptado de (Liao *et al.*, 2022).

No contexto hidrológico, a predição de um passo à frente é amplamente adotada como referência metodológica, pois reflete diretamente a persistência do escoamento associada aos mecanismos de armazenamento da bacia e minimiza efeitos de propagação de erro. Por essa razão, esse regime de previsão é frequentemente utilizado para validação empírica, comparação

de modelos e análise de desempenho preditivo em estudos de previsão de vazão (Nash e Sutcliffe, 1970; Samadi-Koucheksaraee e Chu, 2024; Kordani *et al.*, 2025).

Em síntese, este capítulo estabeleceu a base conceitual e estatística do problema de predição de vazão natural afluente, articulando a interpretação física do processo hidrológico com sua formulação como série temporal univariada. Foram discutidos os principais conceitos teóricos associados ao comportamento temporal da vazão, incluindo estacionariedade, dependência temporal, estruturas de correlação e regimes de memória, bem como as formulações clássicas de modelos autorregressivos lineares amplamente adotados na literatura hidrológica. Ao reunir esses elementos, o capítulo constrói o referencial teórico necessário para compreender os mecanismos de persistência e previsibilidade da vazão. Com base nesse enquadramento, o capítulo seguinte aprofunda a discussão ao introduzir abordagens não lineares com LSTM, nas quais a representação explícita da memória temporal amplia o escopo de modelagem em relação aos modelos lineares clássicos.

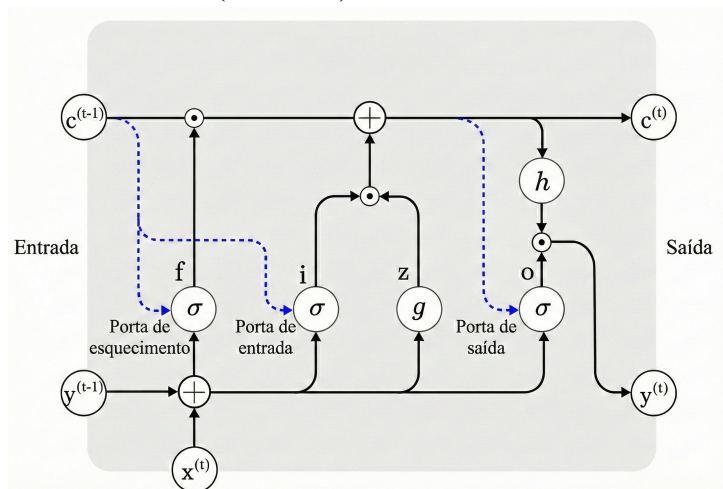
3 MODELO PREDITIVO BASEADO EM LSTM

3.1 Fundamentos da arquitetura

A arquitetura LSTM foi introduzida como modificação estrutural de *recurrent neural network* (RNN) para tratar a instabilidade do gradiente no aprendizado de dependências temporais extensas, por meio da incorporação de um estado de memória e de mecanismos explícitos de controle do fluxo de informação (Hochreiter e Schmidhuber, 1997). Nessa formulação, a retenção temporal deixa de depender apenas de recorrências multiplicativas e passa a ser mediada por uma célula de memória cuja dinâmica é modulada por unidades de porta, permitindo separar armazenamento interno de exposição da informação na saída (Houdt *et al.*, 2020). Essa separação constitui o elemento organizador da LSTM e sustenta sua formulação como modelo recorrente com memória explícita.

A Figura 3 apresenta um bloco básico LSTM, destacando a propagação do estado de célula $c(t - 1)$ ao longo do tempo e a atuação das portas de entrada, esquecimento e saída como operadores seletivos implementados por funções sigmóides (Houdt *et al.*, 2020). A leitura do diagrama é conceitualmente necessária, pois explicita, de forma verificável, os caminhos do estado de memória, as combinações entre a entrada $x(t)$ e a saída anterior $y(t - 1)$, bem reconhecidas nas implementações computacionais contemporâneas. Dessa forma, a figura funciona como especificação estrutural do bloco e como suporte direto à interpretação das equações da propagação direta.

Figura 3 – Componentes da célula (neurônio) LSTM básico.



Fonte: (Houdt *et al.*, 2020).

3.2 Mecanismo de memória

O mecanismo de memória pode ser formalizado pelo conjunto de equações do propagação direta, que relaciona a entrada corrente, a recorrência via saída anterior e a atualização do estado de célula, com funções de ativação não lineares e portas sigmóides (Houdt *et al.*, 2020). Considerando a formulação básica amplamente adotada em bibliotecas *Python* de aprendizado profundo, o sinal candidato de atualização é computado por

$$z(t) = g(W_z x(t) + R_z y(t-1) + b_z), \quad (3.1)$$

enquanto as portas de entrada e esquecimento definem, respectivamente, a incorporação do candidato e a preservação do conteúdo anterior,

$$i(t) = \sigma(W_i x(t) + R_i y(t-1) + b_i), \quad (3.2)$$

$$f(t) = \sigma(W_f x(t) + R_f y(t-1) + b_f), \quad (3.3)$$

em que $\sigma(\cdot)$ representa a função sigmoide e $g(\cdot)$ é tipicamente a função tangente hiperbólica (Houdt *et al.*, 2020). A atualização do estado interno assume a forma aditiva

$$c(t) = f(t) \odot c(t-1) + i(t) \odot z(t), \quad (3.4)$$

em que \odot denota o produto elemento a elemento (produto de Hadamard), operacionalizando a preservação controlada do erro ao longo do tempo, associada ao princípio do *constant error carousel* discutido na proposição original da LSTM (Hochreiter e Schmidhuber, 1997). Por fim, a porta de saída controla a projeção do estado de célula para a saída recorrente,

$$o(t) = \sigma(W_o x(t) + R_o y(t-1) + b_o), \quad (3.5)$$

$$y(t) = g(c(t)) \odot o(t), \quad (3.6)$$

estabelecendo uma dinâmica em que memória e exposição da informação são parametrizadas separadamente, em conformidade com as implementações computacionais mais difundidas (Houdt *et al.*, 2020).

3.3 Capacidade de modelagem de dependências temporais

A capacidade de modelar dependências temporais em redes LSTM está associada à forma como o estado de célula é atualizado recursivamente ao longo do tempo. O termo $f(t) \odot c(t-1)$ permite que informações passadas sejam preservadas de maneira controlada, enquanto o termo $i(t) \odot z(t)$ regula a incorporação de novas informações, evitando que a retenção dependa apenas da propagação instável de gradientes em sequências longas (Hochreiter e Schmidhuber, 1997; Houdt *et al.*, 2020). Esse mecanismo define uma memória explícita com controle seletivo, tornando a LSTM adequada para séries temporais em que a informação relevante não está restrita a poucas defasagens.

Em estudos comparativos com modelos autorregressivos lineares, essa característica é frequentemente utilizada como base conceitual para explicar diferenças de desempenho. Trabalhos como o de Siami-Namini *et al.* (2018) são citados para ilustrar que a presença de memória explícita e não linearidades internas permite às LSTM representar padrões temporais mais complexos do que aqueles capturados por modelos lineares tradicionais. Assim, a literatura reforça, de forma aplicada, a coerência entre a estrutura da LSTM e sua capacidade de lidar com dependências temporais de maior alcance.

3.4 Uso em previsão de vazão

Na modelagem hidrológica, a predição exige a representação de efeitos de armazenamento e retardos entre forçantes meteorológicas e a resposta do sistema, associados a processos como umidade antecedente do solo, acumulação e derretimento de neve e escoamento subterrâneo. Nesse contexto, a formulação recorrente da LSTM, baseada em um estado de célula com atualização controlada, fornece um arcabouço conceitual adequado para modelar dependências temporais inerentes a processos hidrológicos (Hochreiter e Schmidhuber, 1997; Houdt *et al.*, 2020). A atualização do estado interno em 3.6 estabelece o mecanismo matemático mínimo para representar persistência hidrológica por meio da acumulação controlada de informação ao longo do tempo.

A adequação desse arcabouço ao problema hidrológico é discutida por Kratzert *et al.* (Kratzert *et al.*, 2018), que aplicam redes LSTM à modelagem chuva–vazão em larga escala e relatam desempenho comparável ou superior a modelos hidrológicos conceituais consolidados, particularmente em bacias nas quais a memória temporal exerce papel central. De forma

complementar, Filho *et al.* (2022) analisam o uso de LSTM na regionalização de vazões em bacias não monitoradas no semiárido brasileiro, evidenciando que a vantagem estrutural da LSTM depende do regime hidrológico e da escala temporal considerada. Em conjunto, esses trabalhos fundamentam conceitualmente o enquadramento da LSTM como modelo *data-driven* para previsão de vazão, ao mesmo tempo em que delimitam suas condições de aplicabilidade em função da memória hidrológica do sistema.

Por fim, este capítulo apresentou os fundamentos teóricos do modelo preditivo baseado em redes neurais recorrentes do tipo LSTM, detalhando sua arquitetura, o mecanismo de memória e a forma como a atualização controlada do estado interno permite a modelagem explícita de dependências temporais de longo alcance. Discutiu-se, ainda, a adequação conceitual da LSTM à previsão de vazão, à luz dos processos hidrológicos de armazenamento e persistência, bem como evidências empíricas reportadas na literatura especializada. Esse arcabouço estabelece um referencial no qual a dependência temporal é representada diretamente no domínio numérico por meio de estados recorrentes. A partir desse ponto, o capítulo seguinte amplia essa discussão ao introduzir modelos LLM, que abordam a previsão temporal sob um paradigma distinto, no qual séries históricas são reformuladas como sequências simbólicas e a dependência temporal emerge do processo de predição de próximo token condicionado ao contexto linguístico.

4 MODELOS DE LINGUAGEM DE GRANDE ESCALA NA PREVISÃO DE SÉRIES TEMPORAIS

4.1 LLMs como preditores de próximo token

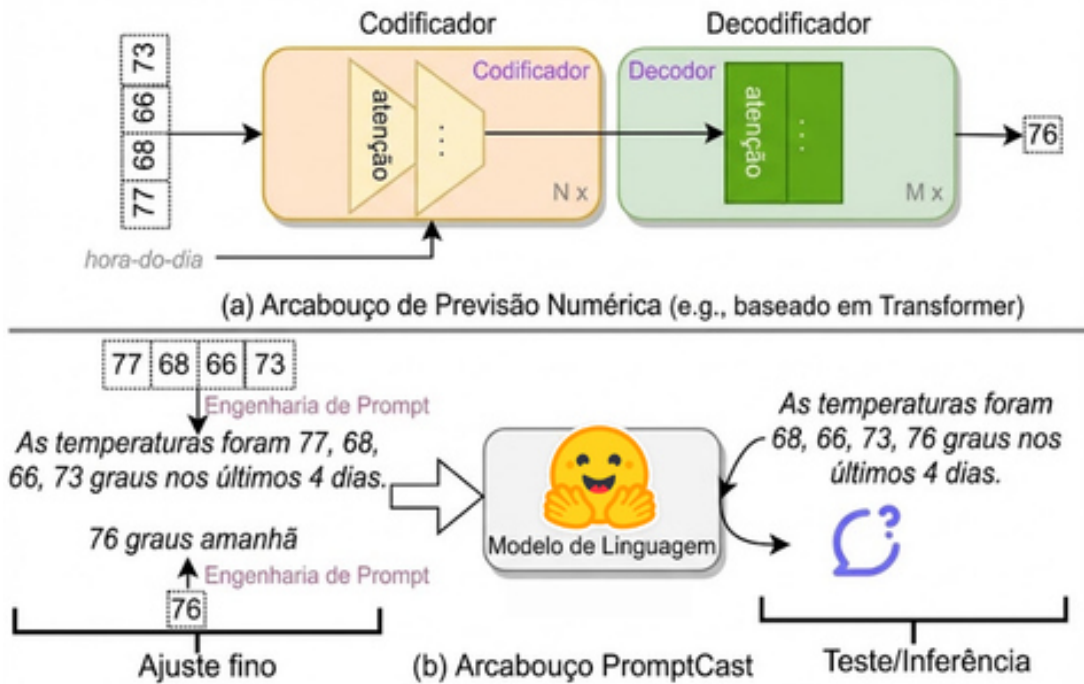
Os LLMs são treinados segundo o princípio da predição autoregressiva de próximo token, no qual a geração de uma sequência é modelada como a maximização da probabilidade condicional de cada *token* dado um contexto precedente. Ao serem aplicados à previsão de séries temporais, esses modelos permitem reformular o problema de forecasting como geração sequencial, na qual observações históricas são codificadas como tokens textuais e os valores futuros são inferidos de forma iterativa. Formalmente, esse processo pode ser representado por

$$p(\mathbf{y}_{t+1:t+H} \mid \mathbf{c}) = \prod_{h=1}^H p(y_{t+h} \mid \mathbf{c}, y_{t+1:t+h-1}), \quad (4.1)$$

em que \mathbf{c} representa o contexto fornecido ao modelo. Gruver et al. (Gruver *et al.*, 2023) demonstram que, mesmo sem qualquer ajuste supervisionado, LLMs podem atuar como previsores *zero-shot* ao tratar a série temporal como um problema de *next-token prediction*, explorando diretamente os vieses indutivos adquiridos durante o pré-treinamento em larga escala.

A Figura 4 ilustra de forma conceitual essa reformulação ao contrastar dois paradigmas distintos de previsão. No painel (a), apresenta-se o *framework* numérico tradicional, no qual modelos de previsão baseados em redes neurais profundas e mecanismos de atenção operam diretamente sobre sequências numéricas contínuas, utilizando camadas empilhadas para extrair dependências temporais e realizar projeções explícitas no domínio do tempo. Em contraste, o painel (b) representa o paradigma *prompt-based*, no qual a série histórica deixa de ser tratada como uma sequência numérica e passa a ser convertida em uma descrição textual estruturada. Nesse caso, a previsão não é produzida por uma função explícita de mapeamento temporal, mas emerge como uma saída linguística gerada pelo modelo. Essa mudança de abordagem evidencia que, nos LLMs, a representação direta da dependência temporal é substituída por um processo de inferência probabilística sequencial no espaço simbólico da linguagem, conforme proposto no PromptCast (Xue e Salim, 2022).

Figura 4 – Comparação conceitual entre um framework baseado em transformer e o paradigma *prompt-based* do PromptCast.



Fonte: Adaptado de (Xue e Salim, 2022).

4.2 Aprendizado em contexto para previsão temporal

O *in-context learning* constitui um mecanismo central para a aplicação de LLMs na previsão de séries temporais ao permitir que o comportamento preditivo do modelo seja condicionado exclusivamente pelo contexto fornecido no *prompt*, sem qualquer atualização paramétrica. Nesse enquadramento, a tarefa de previsão é reformulada como um problema de inferência condicional, no qual exemplos históricos, instruções formais e estatísticas descritivas da série especificam implicitamente a função preditiva a ser executada pelo modelo. Essa formulação desloca o foco da aprendizagem supervisionada tradicional para a construção do contexto, que passa a concentrar a informação relevante para a geração das previsões.

O PromptCast formaliza essa abordagem ao reformular o problema de previsão como uma tarefa de mapeamento entre descrições textuais, na qual séries históricas, metadados e horizontes de previsão são organizados por meio de estruturas linguísticas predefinidas. Essa estratégia permite o uso direto de modelos de linguagem como geradores condicionais, nos quais valores numéricos são inferidos a partir de saídas textuais produzidas pelo modelo (Xue e Salim, 2022). De forma complementar, o TIME-LLM propõe a reprogramação de séries temporais, em que segmentos da série são convertidos em representações compatíveis com o espaço interno do

modelo por meio de protótipos textuais. Adicionalmente, o mecanismo denominado Prompt-as-Prefix (PaP) insere, no prefixo da entrada, informações relativas à tarefa, ao conjunto de dados e a estatísticas da série, explorando a capacidade de condicionamento contextual do modelo para estabilizar a inferência e reduzir a dependência de ajustes paramétricos (Jin *et al.*, 2023). Nesse mesmo eixo, abordagens recentes de previsão temporal com modelos LLM destacam que o desempenho depende criticamente das escolhas de codificação e de contexto, o que torna a construção do texto de entrada um componente estrutural do modelo efetivo, e não apenas um elemento superficial de interface (Gruver *et al.*, 2023).

Esse panorama é coerente com a evolução recente de modelos fundacionais multimodais, nos quais o *prompting* é tratado como mecanismo de controle e especialização *on-the-fly*. Em particular, a família Gemini reforça a centralidade do condicionamento por contexto em modelos de larga escala e indica que a generalização por instruções e exemplos no *prompt* é um princípio de projeto que se estende a diferentes modalidades, o que sustenta o uso de LLMs como preditores condicionais quando a série é adequadamente representada e contextualizada (Anil *et al.*, 2023).

4.3 Desafios teóricos na previsão com LLMs

Apesar do desempenho empírico observado em regimes *zero-shot* e *few-shot*, o uso de LLMs em previsão de séries temporais apresenta limitações teóricas relevantes. A discretização de valores contínuos em tokens impõe restrições de representação e torna o desempenho sensível a escolhas de tokenização, reescala, precisão numérica e amostragem, pois pequenas variações na codificação podem alterar a distribuição sobre sequências plausíveis e, conseqüentemente, as estatísticas preditivas (Gruver *et al.*, 2023). Além disso, a previsão pode passar a depender da estrutura de codificação e do contexto fornecido ao modelo, deslocando parcialmente a modelagem do domínio estatístico explícito para mecanismos implícitos de linguagem e amostragem, o que tende a aumentar a variabilidade metodológica entre estudos e dificulta a formulação de garantias teóricas de consistência sob mudanças de escala, ruído e regime.

No TIME-LLM, a necessidade de módulos explícitos de reprogramação evidencia que a correspondência entre séries temporais e linguagem não é intrínseca aos LLMs, exigindo mecanismos artificiais de alinhamento para compatibilizar modalidade e preservar informação temporal (Jin *et al.*, 2023). De modo análogo, no PromptCast, a generalização é mediada pela

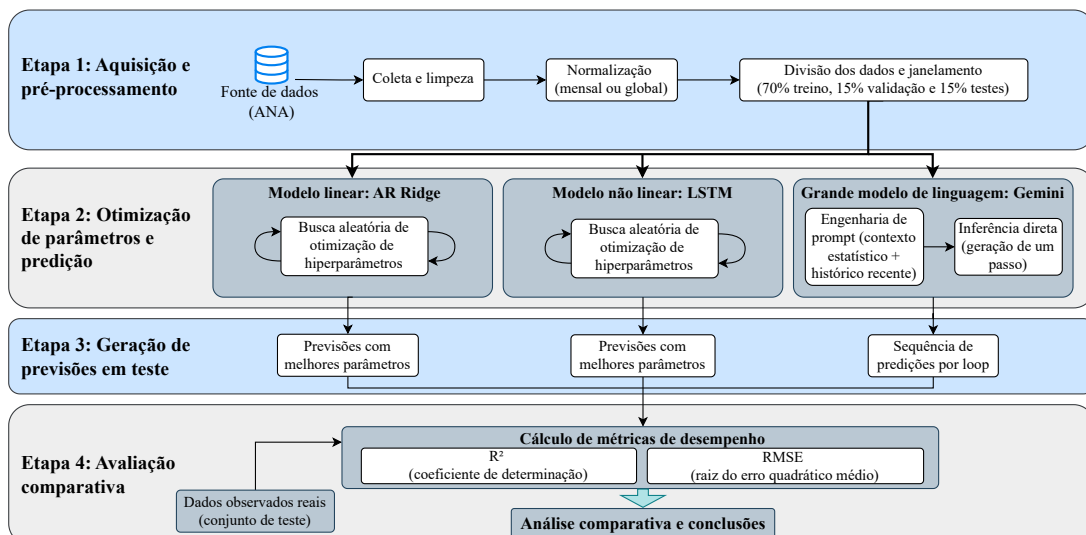
expressividade dos *templates* textuais, o que condiciona o espaço de hipóteses acessível ao modelo e torna o desempenho dependente do desenho do esquema de serialização e extração numérica (Xue e Salim, 2022). Em modelos fundacionais amplos, como Gemini, o treinamento e o pós-treinamento voltados a seguimento de instruções e segurança podem introduzir vieses comportamentais que não são otimizados para calibração probabilística e precisão numérica fina, reforçando a necessidade de protocolos de avaliação específicos para séries temporais, sobretudo quando se pretende quantificar incerteza e extrapolar regimes (Anil *et al.*, 2023). Assim, embora conceitualmente unificadores, LLMs devem ser compreendidos como ferramentas complementares à modelagem estatística e neural tradicional, especialmente em contextos físicos, como a previsão hidrológica de vazões, nos quais estrutura, conservação e ruído têm papel determinante.

5 METODOLOGIA

Este capítulo apresenta o delineamento metodológico adotado no estudo, contemplando de forma integrada as etapas de aquisição e tratamento dos dados hidrológicos, a configuração dos experimentos computacionais e os critérios empregados para avaliação do desempenho preditivo. A estratégia metodológica fundamenta-se em duas linhas complementares de investigação. A primeira explora modelos supervisionados de aprendizado de máquina, incluindo um modelo autorregressivo linear com regularização de cumeeira (*ridge regression*) e uma rede neural recorrente do tipo LSTM, ambos submetidos a procedimentos sistemáticos de ajuste e seleção de hiperparâmetros. A segunda linha investiga o uso de modelos LLM, como mecanismos de inferência preditiva, operacionalizados por meio de engenharia de *prompt* orientada por contexto estatístico. A Figura 5 sintetiza o fluxo metodológico do estudo, oferecendo uma representação esquemática das principais etapas envolvidas, desde a preparação dos dados até a avaliação comparativa dos modelos.

Sob a perspectiva epistemológica, a pesquisa caracteriza-se como quantitativa, aplicada e de natureza descritivo-exploratória, conduzida segundo uma abordagem hipotético-dedutiva. Partem-se de hipóteses relacionadas ao desempenho relativo de distintas classes de modelos na tarefa de previsão de vazões, as quais são examinadas por meio de experimentação computacional controlada e análise quantitativa dos resultados obtidos. Esse enquadramento metodológico permite avaliar, de forma objetiva, tanto a capacidade preditiva quanto a robustez das abordagens consideradas em diferentes cenários experimentais.

Figura 5 – Diagrama geral do fluxo metodológico adotado no estudo.



Fonte: Elaborado pelo autor.

5.1 Conjunto de dados

O conjunto de dados é composto por séries temporais de vazão natural afluyente média mensal (m^3/s) de reservatórios hidrelétricos pertencentes ao SIN. A análise concentra-se exclusivamente nessa variável por representar diretamente a dinâmica hidrológica das bacias, sem influência de decisões operativas associadas à geração de energia.

Foram consideradas quatro séries temporais correspondentes aos reservatórios de Furnas, Sobradinho, Itaipu e Ilha Solteira, selecionados em função de sua relevância para o SIN e da disponibilidade de registros históricos longos e consistentes. Para cada reservatório, utilizou-se uma série mensal compreendendo o período de janeiro de 1993 a agosto de 2025.

As séries apresentam diferenças marcantes de escala e variabilidade, refletindo as características hidrológicas específicas de cada bacia contribuinte. Com o objetivo de caracterizar o domínio dos dados e subsidiar as etapas de modelagem, a Tabela 1 apresenta estatísticas descritivas resumidas das séries, incluindo medidas de tendência central, dispersão e valores extremos.

Tabela 1 – Estatísticas descritivas das séries de vazão (1993-2025), em m^3/s

Estatística	Furnas	Sobradinho	Itaipu	Ilha Solteira
Número de observações	392	392	392	392
Média	763.55	1 883.03	11 149.58	4 914.77
Desvio padrão	561.80	1 485.18	4 628.12	2 928.13
Mínimo	96.94	206.98	3 403.04	1 166.03
1º quartil	377.26	782.90	7 640.71	2 736.15
Mediana	578.48	1 291.53	10 147.82	4 126.93
3º quartil	977.03	2 598.82	13 573.73	6 275.48
Máximo	3 620.64	7 121.43	28 591.85	17 850.52

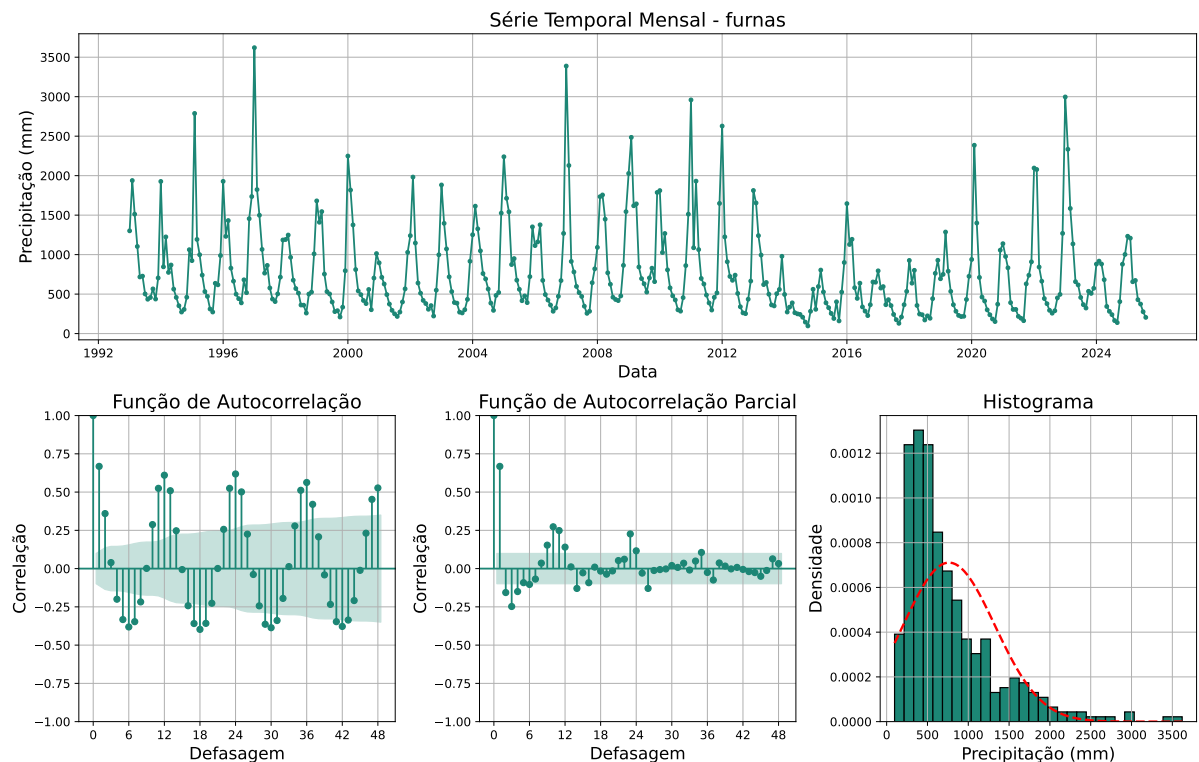
Fonte: Elaborado pelo autor.

A seguir, são apresentadas as visualizações descritivas das séries temporais mensais de vazão natural afluyente dos reservatórios analisados. Para cada reservatório, as figuras incluem a série temporal mensal, a FAC, a FACP e o histograma, com o objetivo de caracterizar a variabilidade temporal, a dependência serial e a distribuição empírica dos dados utilizados nas etapas subsequentes do estudo.

5.1.1 Reservatório de Furnas

A Figura 6 apresenta a análise gráfica da série de vazão natural afluyente do reservatório de Furnas, cujos valores se concentram predominantemente entre aproximadamente 200 e 1.000 m^3/s , com recorrentes picos que ultrapassam esse intervalo. A FAC evidencia dependência temporal significativa, com coeficiente elevado no lag 1, indicando persistência de curto prazo entre meses consecutivos, e picos bem definidos nos lags 12, 24 e 36, caracterizando uma sazonalidade anual característica. A FACP reforça esse padrão ao mostrar que a dependência direta da série está concentrada em poucos atrasos iniciais, com dominância no lag 1 e picos secundários nos lags 3, 6, 10, 14, 23 e 26, todos entre os valores mais relevantes da FACP. O histograma confirma que a distribuição da vazão é fortemente assimétrica à direita, com maior densidade de observações entre aproximadamente 300 e 800 m^3/s , em torno da mediana (578,48 m^3/s) e da média (763,55 m^3/s), enquanto a cauda superior se estende até valores superiores a 3.000 m^3/s , evidenciando períodos de maior aporte de água nos reservatórios.

Figura 6 – Série temporal, funções de autocorrelação e histograma da vazão natural afluyente do reservatório de Furnas.

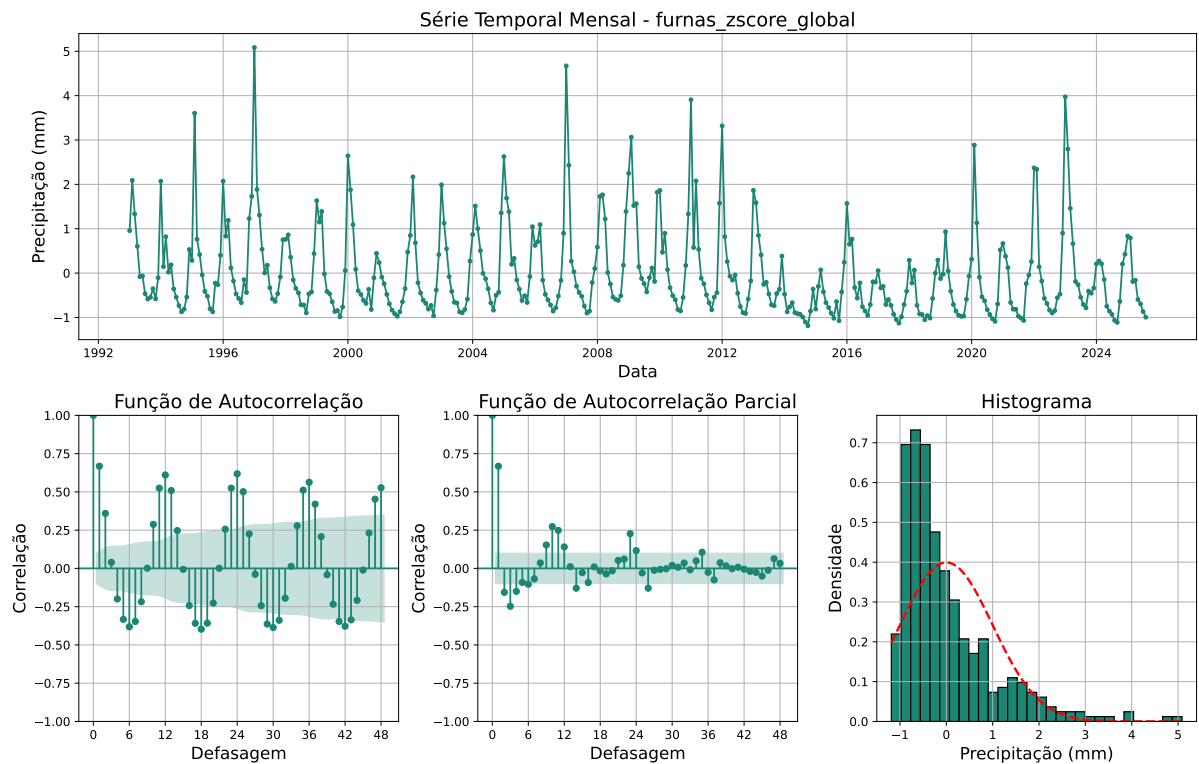


Fonte: Elaborado pelo autor.

5.1.1.1 Normalização global da série de Furnas

A Figura 7, após a normalização global por z -score, a série de Furnas passa a ser expressa em desvios relativos à média histórica total, eliminando a escala em m^3/s . A sazonalidade anual e a persistência temporal permanecem inalteradas, enquanto a magnitude absoluta das cheias deixa de dominar a representação, permitindo comparação direta entre períodos hidrológicos distintos.

Figura 7 – Série de Furnas após normalização z -score global.

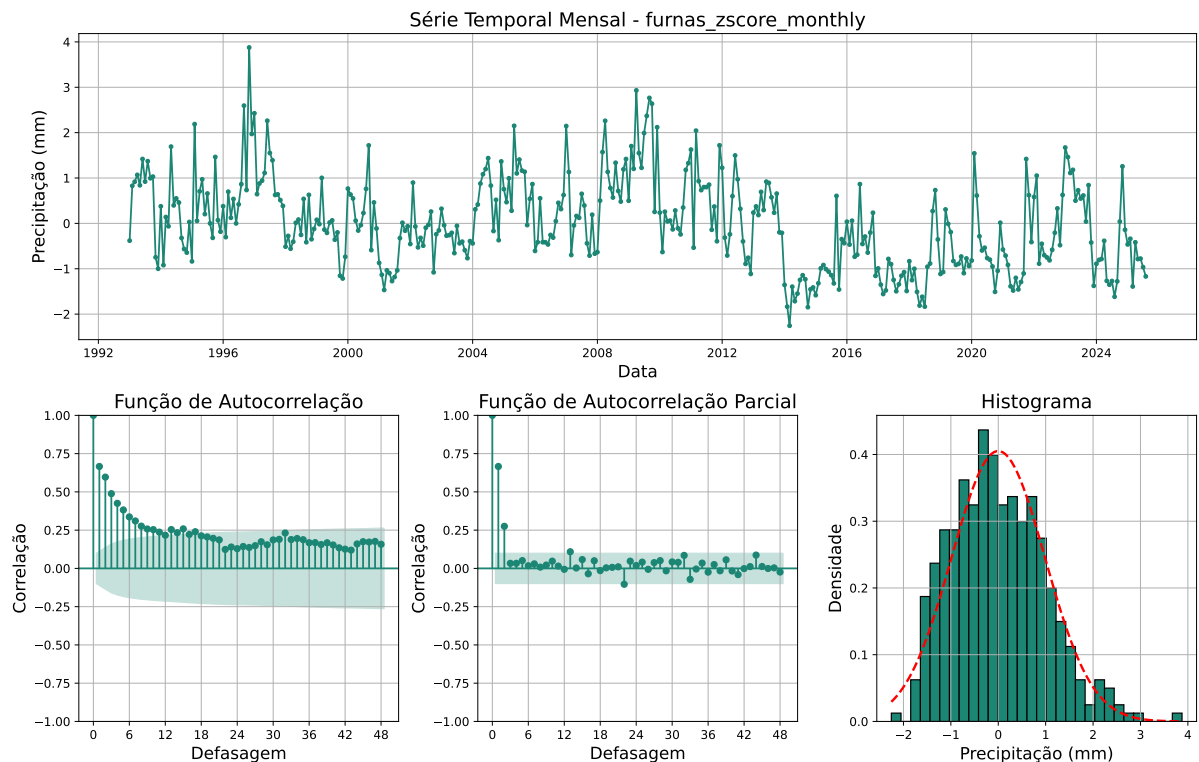


Fonte: Elaborado pelo autor.

5.1.1.2 Normalização mensal da série de Furnas

Na Figura 8, apresenta-se a série de Furnas após a normalização mensal por z -score, na qual remove-se o ciclo sazonal médio da série, reduzindo a recorrência visual dos picos anuais. Com isso, a dependência temporal associada à sazonalidade é atenuada, e a série passa a enfatizar desvios hidrológicos específicos de cada mês em relação ao seu comportamento climatológico típico.

Figura 8 – Série de Furnas após normalização z -score mensal.

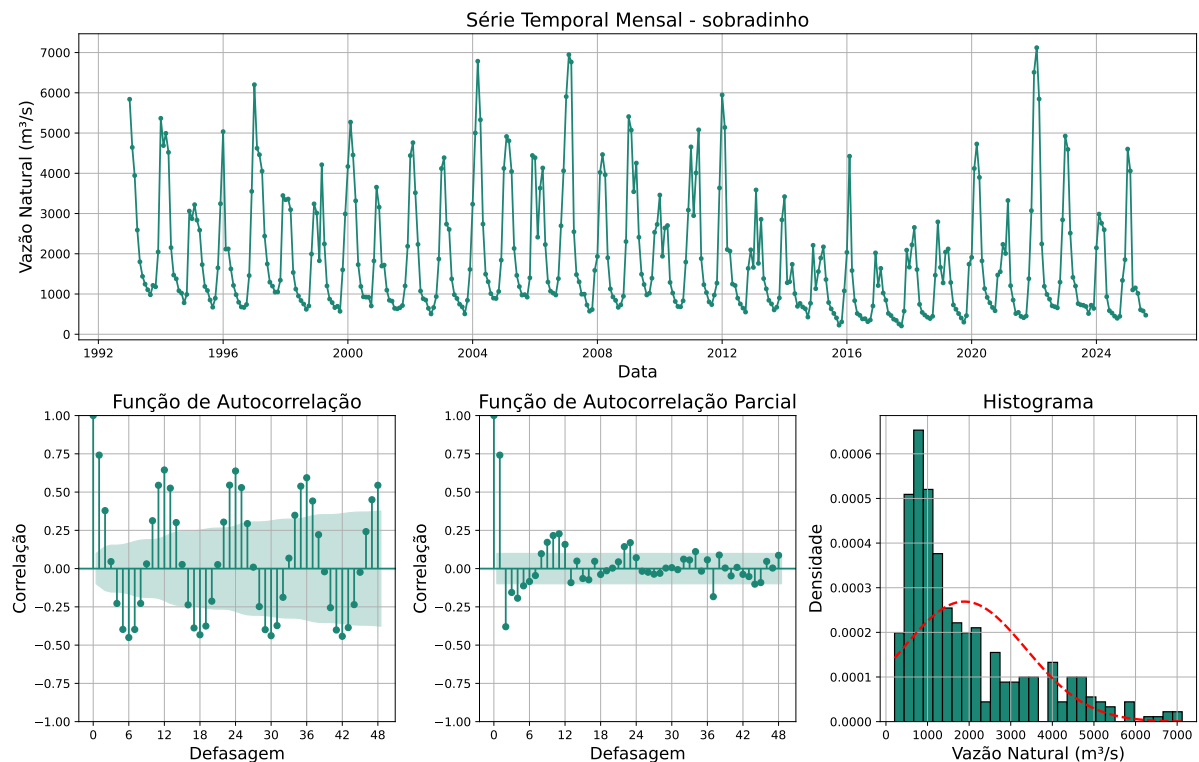


Fonte: Elaborado pelo autor.

5.1.2 Reservatório de Sobradinho

Em Sobradinho, seguindo a Figura 9, a série temporal mensal, apresenta elevada variabilidade hidrológica, com valores mínimos recorrentes próximos de $200 \text{ m}^3/\text{s}$ e picos sazonais que ultrapassam $6.000 \text{ m}^3/\text{s}$. Observa-se um padrão cíclico regular, com cheias concentradas em intervalos anuais bem definidos e longos períodos de recessão. A FAC exibe coeficiente elevado no lag 1, indicando forte persistência entre meses consecutivos, além de picos pronunciados nos lags 12, 24 e 36, confirmando uma sazonalidade anual robusta. A FACP evidencia dependência direta dominante nos lags 1 e 2, com contribuições relevantes adicionais nos lags 11 e 23. O histograma indica concentração de observações em faixas inferiores a aproximadamente $2.000 \text{ m}^3/\text{s}$, com maior densidade entre cerca de 500 e $1.500 \text{ m}^3/\text{s}$, enquanto a cauda superior se estende até valores superiores a $6.000 \text{ m}^3/\text{s}$, caracterizando uma distribuição assimétrica à direita, com mediana ($1.291,53 \text{ m}^3/\text{s}$) visualmente inferior à média ($1.883,03 \text{ m}^3/\text{s}$) e presença de eventos extremos pouco frequentes.

Figura 9 – Série temporal, funções de autocorrelação e histograma da vazão natural afluyente do reservatório de Sobradinho.

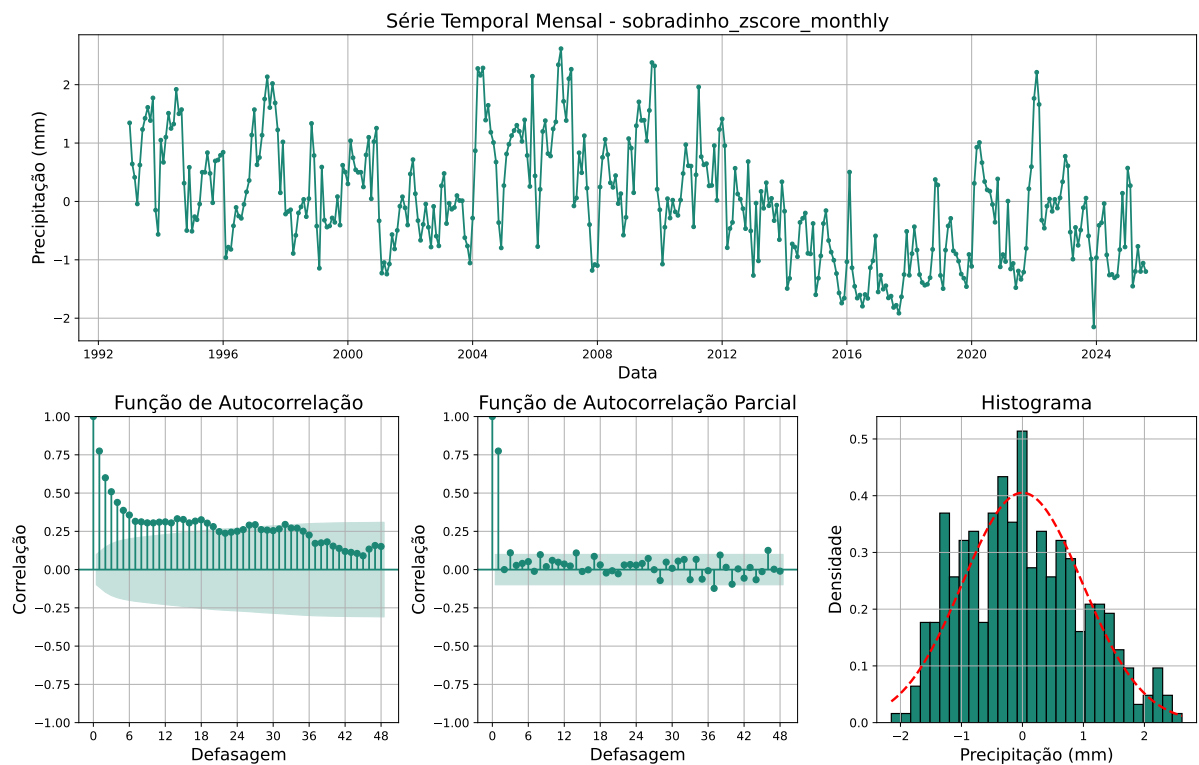


Fonte: Elaborado pelo autor.

5.1.2.1 Normalização global da série de Sobradinho

A Figura 10 apresenta a série de vazão natural afluyente do reservatório de Sobradinho após a normalização global, que mantém sua estrutura temporal e o padrão sazonal anual, agora expressos em unidades padronizadas. A elevada variabilidade hidrológica permanece evidente, porém sem a influência direta da magnitude absoluta dos picos de cheia.

Figura 10 – Série de Sobradinho após normalização z-score global.

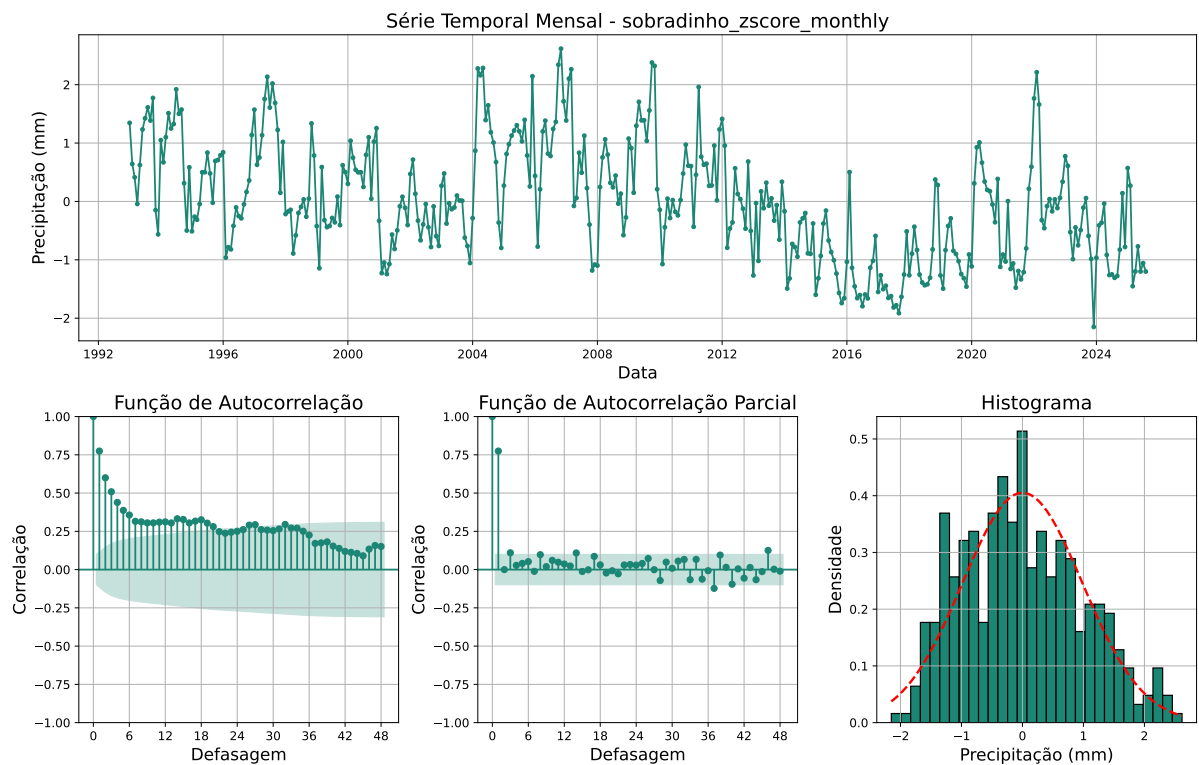


Fonte: Elaborado pelo autor.

5.1.2.2 Normalização mensal da série de Sobradinho

Na Figura 11, apresenta-se a normalização mensal, que suprime o ciclo sazonal médio característico de Sobradinho, reduzindo a persistência anual observada na série original. A variabilidade restante passa a refletir predominantemente anomalias hidrológicas relativas ao regime típico de cada mês.

Figura 11 – Série de Sobradinho após normalização z-score mensal.

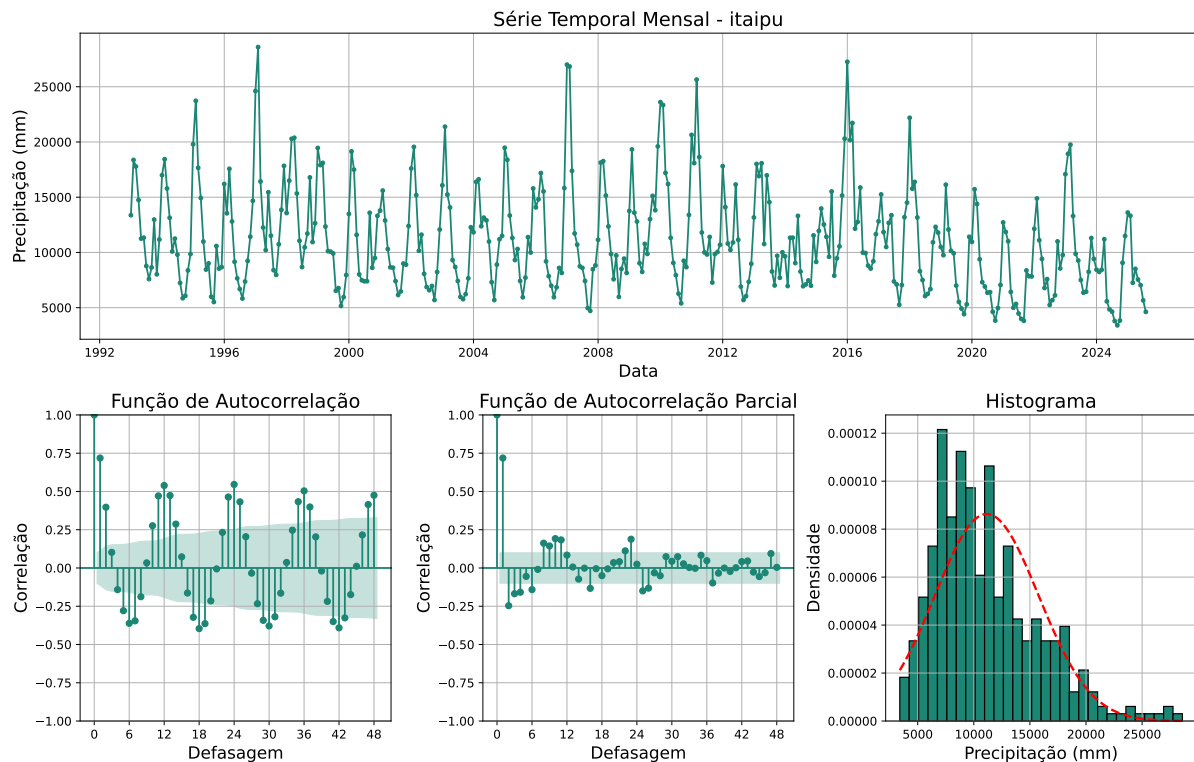


Fonte: Elaborado pelo autor.

5.1.3 Reservatório de Itaipu

Na Figura 12, apresentam-se os dados referentes ao reservatório de Itaipu, caracterizadas por elevada magnitude hidrológica, com valores predominantemente concentrados entre aproximadamente 5.000 e 15.000 m^3/s e picos recorrentes que ultrapassam 25.000 m^3/s . A FAC indica dependência temporal persistente, com coeficiente elevado no lag 1 e picos bem definidos nos lags 12, 24 e 36, além de atrasos adjacentes, caracterizando uma sazonalidade anual robusta. A FACP reforça esse comportamento ao evidenciar dependência direta concentrada nos primeiros atrasos, com dominância nos lags 1 e 2 e picos secundários nos lags 3, 10, 11, 23, 25 e 26. O histograma confirma uma distribuição assimétrica à direita, coerente com a assimetria positiva observada, apresentando maior densidade de observações em níveis intermediários de vazão e uma cauda superior longa associada à ocorrência de eventos extremos hidrológicos.

Figura 12 – Série temporal, funções de autocorrelação e histograma da vazão natural afluente do reservatório de Itaipu.

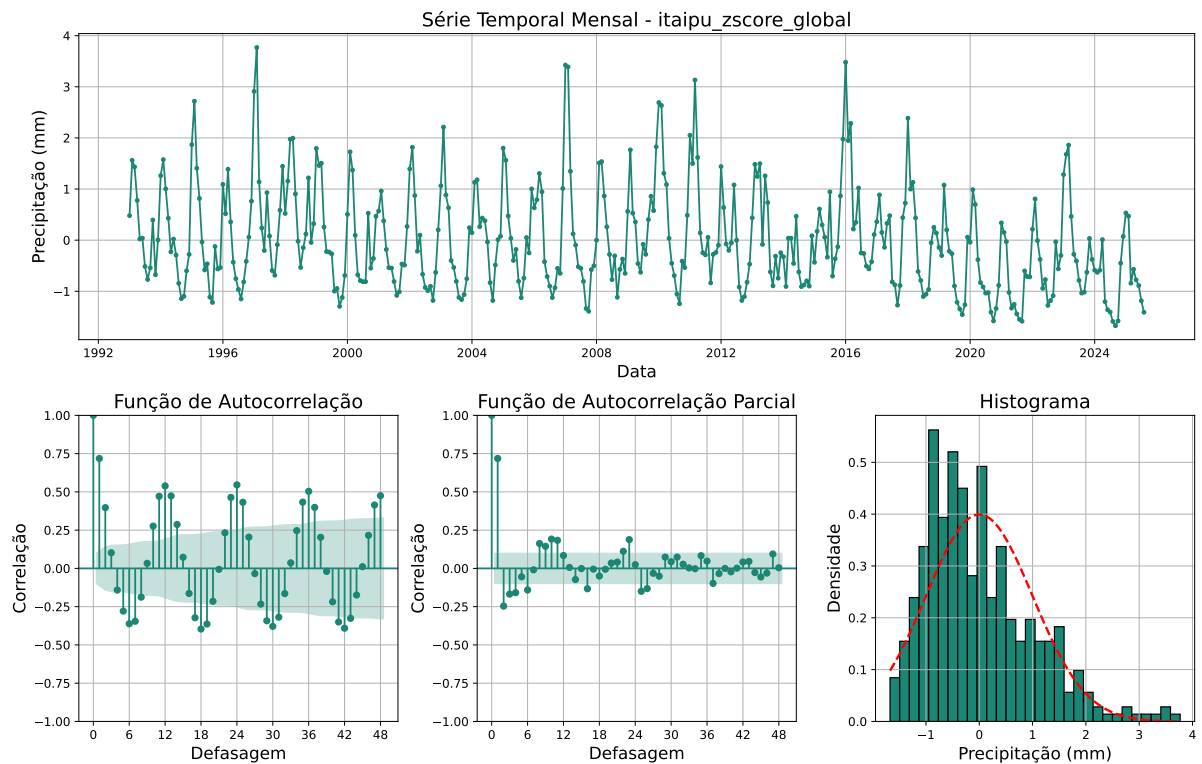


Fonte: Elaborado pelo autor.

5.1.3.1 Normalização global da série de Itaipu

A padronização global A Figura 13 apresenta a série de vazão natural afluente do reservatório de Itaipu com a normalização global, assim, a série deixa de ser dominada por sua elevada magnitude hidrológica, mantendo intactas a sazonalidade anual e a dependência temporal. Cheias e estiagens passam a ser interpretadas como desvios relativos ao comportamento médio da própria série.

Figura 13 – Série de Itaipu após normalização z-score global.

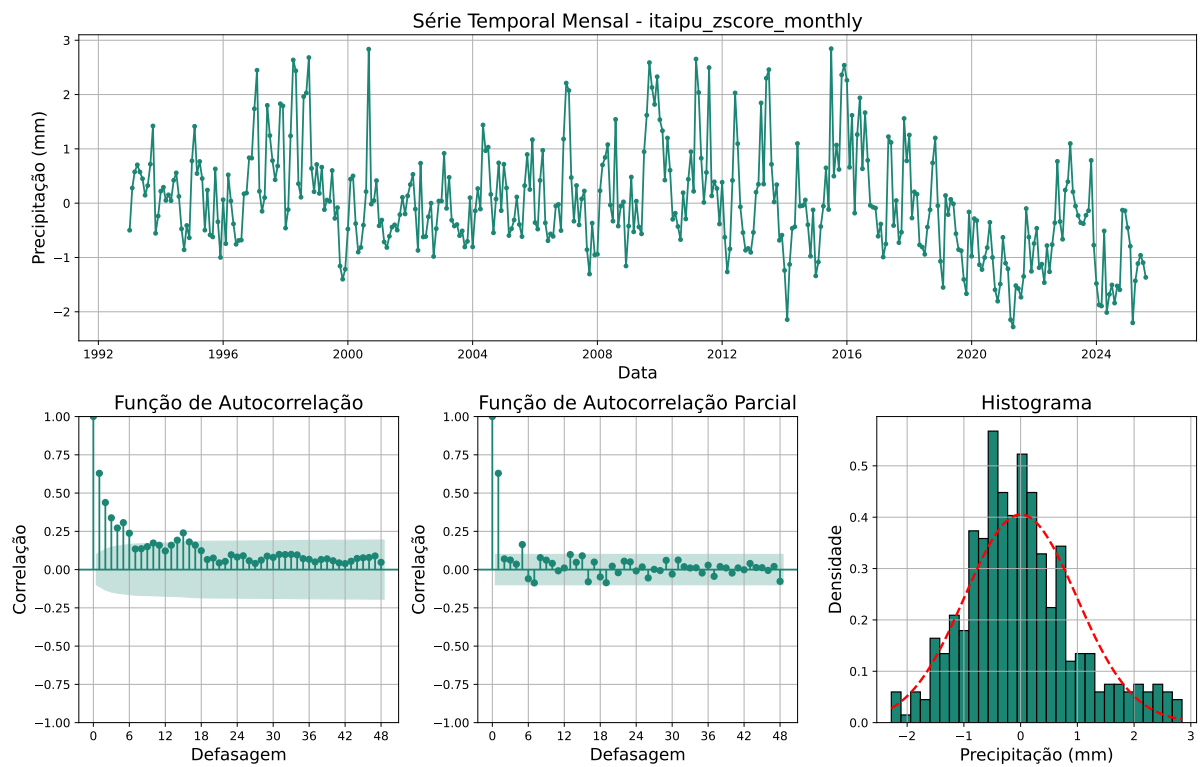


Fonte: Elaborado pelo autor.

5.1.3.2 Normalização mensal da série de Itaipu

Na Figura 14, apresenta-se a série de Itaipu após a normalização mensal, reduzindo significativamente o componente sazonal médio da série, enfraquecendo a recorrência anual sistemática. O sinal resultante destaca variações hidrológicas intra-anuais, associadas a condições anômalas em relação ao padrão mensal esperado.

Figura 14 – Série de Itaipu após normalização z-score mensal.

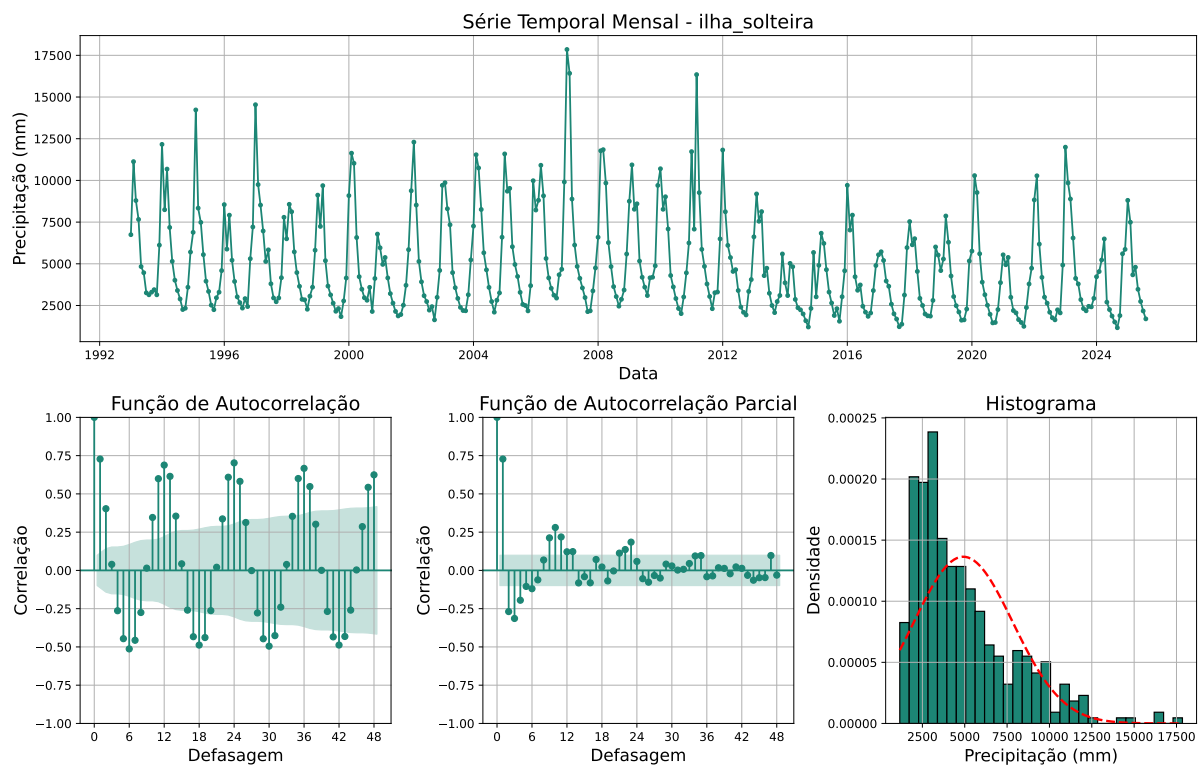


Fonte: Elaborado pelo autor.

5.1.4 Reservatório de Ilha Solteira

A Figura 15 ilustra o comportamento da série mensal de vazão natural afluente do reservatório de Ilha Solteira, cuja escala se posiciona em nível intermediário em relação aos demais reservatórios analisados. Visualmente, observa-se maior concentração de valores no intervalo aproximado entre 2.500 e 10.000 m^3/s , região que engloba tanto a mediana (4.126,93 m^3/s) quanto a média (4.914,77 m^3/s), enquanto episódios de cheia se manifestam por picos acima de 12.500 m^3/s , refletindo elevada dispersão ao longo do tempo. A FAC evidencia persistência temporal significativa, com destaque para o primeiro atraso e para picos regulares nos lags 12, 24 e 36, caracterizando um padrão sazonal anual bem definido. A FACP complementa essa análise ao indicar que a dependência direta da série se concentra majoritariamente nos atrasos iniciais, com maior relevância nos lags 1, 3, 10 e 23, enquanto os demais apresentam influência mais limitada. O histograma confirma a assimetria positiva da distribuição, com maior densidade em faixas intermediárias de vazão e uma cauda superior extensa, associada à ocorrência de eventos extremos de maior magnitude.

Figura 15 – Série temporal, funções de autocorrelação e histograma da vazão natural afluente do reservatório de Ilha Solteira.

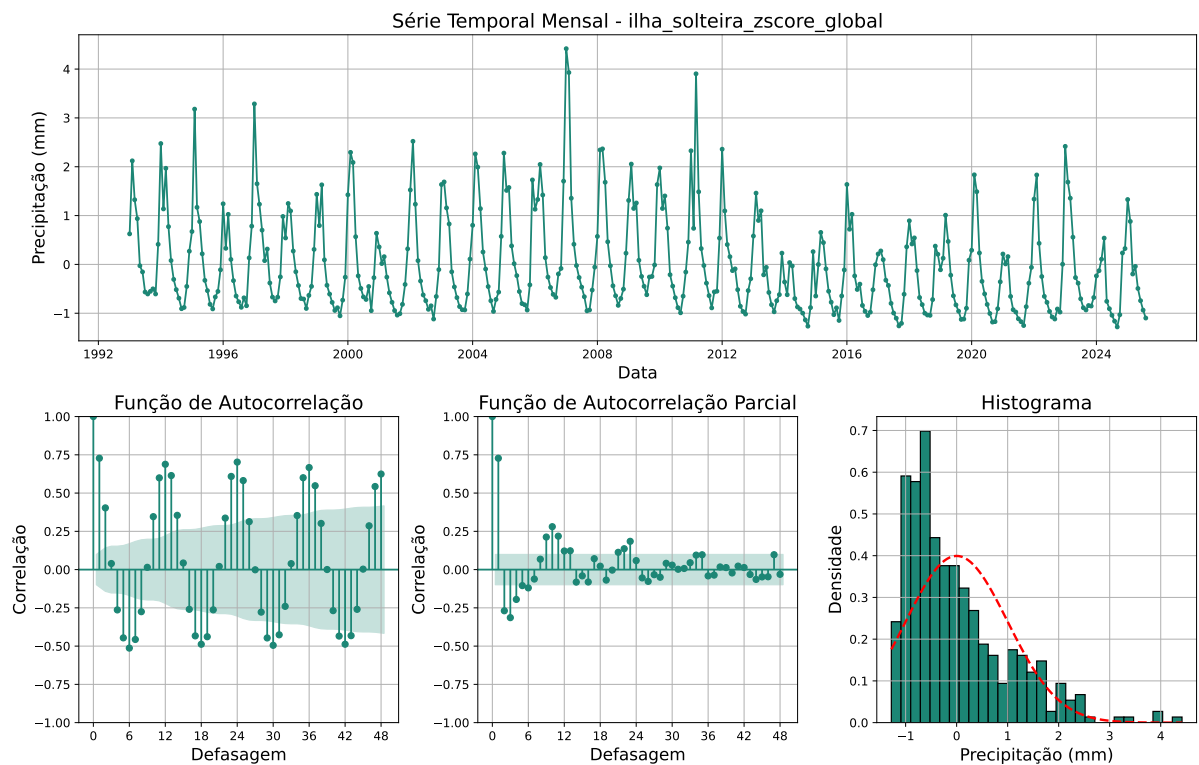


Fonte: Elaborado pelo autor.

5.1.4.1 Normalização global da série de Ilha Solteira

A Figura 16 apresenta a série de vazão natural afluyente do reservatório de Ilha Solteira, que após a normalização global, a série mantém sua estrutura temporal e sazonal, agora representada em termos relativos. A padronização reduz o peso da escala intermediária original e facilita a interpretação comparativa dos eventos hidrológicos ao longo do tempo.

Figura 16 – Série de Ilha Solteira após normalização z-score global.

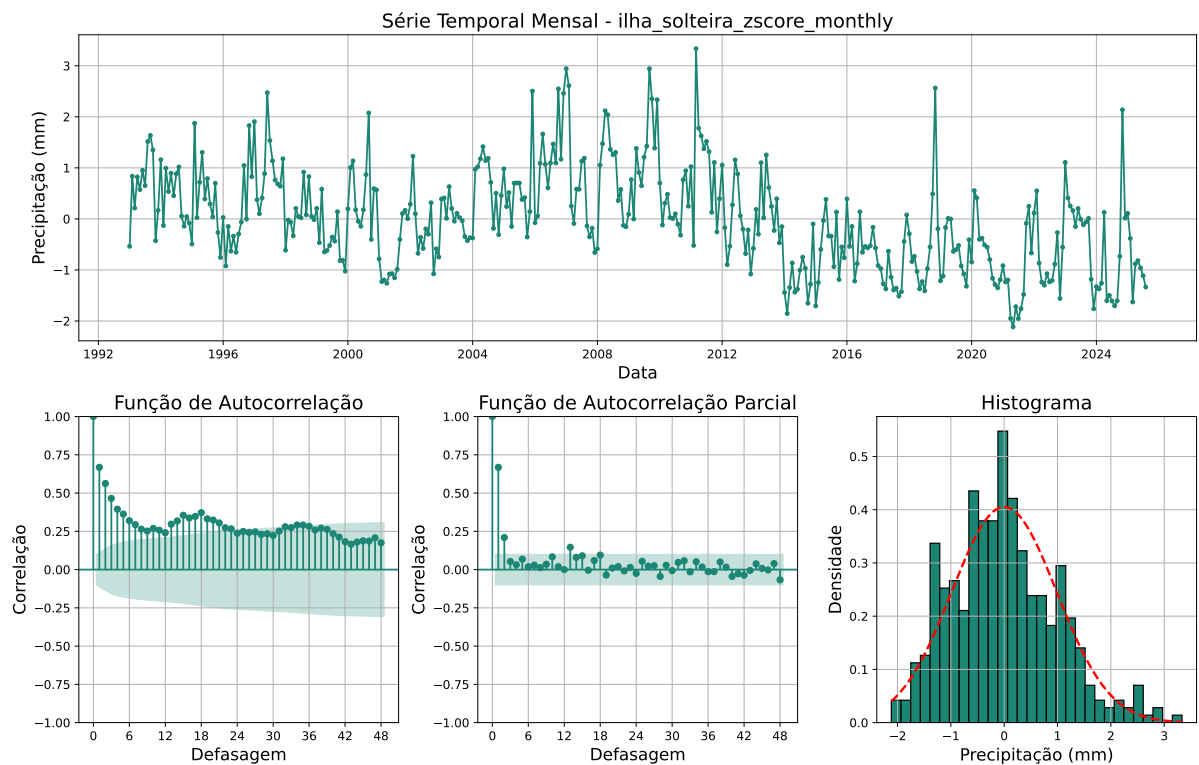


Fonte: Elaborado pelo autor.

5.1.4.2 Normalização mensal da série de Ilha Solteira

Na Figura 17, apresenta-se a série de Ilha Solteira após a normalização mensal, que atenua o ciclo sazonal médio, reduzindo a persistência anual observada na série original. O comportamento restante passa a evidenciar anomalias hidrológicas específicas de cada mês, em relação ao regime climatológico típico.

Figura 17 – Série de Ilha Solteira após normalização z-score mensal.



Fonte: Elaborado pelo autor.

5.2 Modelos supervisionados de aprendizado de máquina

A etapa de modelagem preditiva supervisionada fundamentou-se na aplicação e na comparação de duas arquiteturas distintas para a previsão da vazão média mensal, considerando um horizonte de um passo à frente ($h = 1$). A estratégia experimental consistiu na confrontação de um método estatístico linear com regularização (AR) e de uma abordagem baseada em aprendizado profundo (LSTM), sendo ambos submetidos a um protocolo de otimização de hiperparâmetros e avaliados sob dois regimes de normalização dos dados: global e mensal.

5.2.1 Ambiente computacional e reprodutibilidade

Os experimentos foram desenvolvidos na linguagem Python 3.12, utilizando as bibliotecas *Scikit-Learn* para os modelos lineares e *TensorFlow/Keras* para as redes neurais. Para garantir a reprodutibilidade estrita dos resultados e da inicialização dos pesos sinápticos, fixou-se a semente aleatória (*random seed*) global em 42 em todas as bibliotecas estocásticas (*Numpy*, *Random* e *TensorFlow*).

5.2.2 Implementação do modelo AR

Como modelo estatístico de referência, foi adotada AR. O vetor de entrada do modelo foi composto por p defasagens temporais consecutivas da própria série de vazão, caracterizando um modelo autorregressivo de ordem p , com horizonte de predição fixado em um passo à frente.

A configuração do modelo AR foi determinada por meio de um procedimento de busca aleatória, *random search* (RS) com 1.000 iterações independentes para cada reservatório. Em cada iteração, o número de defasagens temporais foi amostrado no intervalo discreto de 2 a 12 meses, permitindo avaliar diferentes níveis de memória temporal do sistema hidrológico. O hiperparâmetro de regularização α foi amostrado a partir de uma distribuição log-uniforme no intervalo $[10^{-6}, 10^0]$. Adicionalmente, avaliou-se a inclusão ou não do termo de intercepto no ajuste do modelo.

Antes do treinamento, as séries foram organizadas em janelas deslizantes de acordo com o número de defasagens selecionado. A normalização dos dados foi realizada por meio de padronização do tipo *z-score*, ajustada mensalmente com base exclusivamente no conjunto de treino, de modo a preservar a sazonalidade hidrológica e evitar vazamento de informação temporal. As previsões foram posteriormente reconvertidas para a escala original da vazão para

fins de avaliação.

O treinamento do modelo foi conduzido exclusivamente sobre o conjunto de treino, enquanto a seleção do melhor conjunto de hiperparâmetros baseou-se no desempenho obtido no conjunto de validação. O critério de seleção considerou o erro quadrático médio, mean squared error (MSE) calculado na escala original da vazão, sendo o R^2 utilizado como métrica complementar de desempenho. Esse procedimento assegura uma avaliação imparcial da capacidade preditiva do modelo e estabelece uma linha de base consistente para a comparação com abordagens não lineares mais complexas.

A Tabela 2 sintetiza o espaço de busca dos hiperparâmetros considerados na implementação do modelo AR, evidenciando o caráter parcimonioso, interpretável e reproduzível da abordagem estatística adotada.

Tabela 2 – Espaço de busca dos hiperparâmetros do modelo AR

Hiperparâmetro	Valores avaliados
Ordem do modelo (<i>lags</i>)	$\{2, 3, \dots, 12\}$
Parâmetro de regularização α	$[10^{-6}, 10^0]$ (log-uniforme)
Intercepto	{ativado, desativado}
Horizonte de predição	1 passo à frente
Método de otimização	Solução fechada (Regulação de cumeeira)
Função objetivo	MSE

Fonte: Elaborado pelo autor.

5.2.3 Implementação da rede neural LSTM

A modelagem baseada em redes neurais recorrentes do tipo LSTM foi adotada com o objetivo de capturar padrões não lineares e dependências temporais de longo alcance presentes nas séries hidrológicas analisadas.

A implementação considerou uma amostra experimental composto por janelas deslizantes da série temporal, previamente normalizadas por *z-score*. Cada instância de entrada consiste em uma sequência temporal de comprimento fixo (*window size*), utilizada para a predição da vazão no horizonte de um passo à frente.

A definição da arquitetura e dos hiperparâmetros da LSTM foi conduzida por meio de um procedimento de RS com 1.000 experimentos independentes por reservatório. Foram avaliadas configurações contendo uma ou duas camadas recorrentes LSTM empilhadas, seguidas por uma camada densa intermediária e uma camada de saída linear.

O treinamento dos modelos utilizou o otimizador Adam, em conjunto com a função objetivo do MSE. Foram incorporados mecanismos de regularização explícitos, incluindo o *dropout* nas camadas recorrentes, técnica que consiste na desativação aleatória de unidades durante o treinamento com o objetivo de reduzir o sobreajuste. Adicionalmente, empregaram-se estratégias adaptativas de controle do treinamento, como o *Early Stopping*, que interrompe o processo quando o desempenho no conjunto de validação deixa de melhorar, evitando o ajuste excessivo ao conjunto de treino, e o *ReduceLROnPlateau*, que reduz automaticamente a taxa de aprendizado quando a função de perda entra em regime de estagnação, favorecendo uma convergência mais estável.

A Tabela 3 resume o espaço de busca dos principais hiperparâmetros avaliados no procedimento experimental.

Tabela 3 – Espaço de busca dos hiperparâmetros da arquitetura LSTM

Hiperparâmetro	Valores avaliados
Número de camadas LSTM	{1, 2}
Unidades por camada LSTM	{16, 32, 64, 128, 256}
Tamanho da janela temporal	{6, 12, 18, 24}
Unidades da camada densa	{16, 32, 64}
Função de ativação	{relu, tanh, sigmoid}
Taxa de <i>dropout</i>	{0,2;0,3;0,4;0,5}
Taxa de aprendizado	{ 10^{-4} ; 5×10^{-4} ; 10^{-3} }
Tamanho do lote	{8, 16, 32, 64}
Número máximo de épocas	200
Horizonte de predição	1 passo à frente
Otimizador	Adam
Função objetivo	MSE

Fonte: Elaborado pelo autor.

5.3 Pré-processamento e estrutura experimental

A confiabilidade dos modelos preditivos depende intrinsecamente da qualidade do tratamento dos dados. Esta etapa engloba a conversão da série temporal em matrizes de aprendizado supervisionado, a normalização das variáveis para estabilização numérica e a definição de um pipeline rigoroso de validação.

5.3.1 Estruturação temporal e janelamento

O janelamento adotado segue estritamente o regime de predição de um passo à frente já definido anteriormente. Cada amostra é construída apenas com observações efetivamente disponíveis no instante da previsão, e o valor alvo corresponde ao próximo passo temporal da série. Não ocorre realimentação de previsões no conjunto de entrada, o que assegura consistência temporal, evita vazamento de informação entre treino, validação e teste, e garante que o desempenho obtido reflita exclusivamente a capacidade dos modelos em capturar dependências temporais a partir de dados observados.

Os dados foram particionados cronologicamente, destinando-se 70% da série ao treinamento, 15% à validação e 15% ao teste.

5.3.2 Protocolos de normalização

A normalização das séries de vazão é adotada como etapa de pré-processamento para estabilização numérica e tratamento da não estacionariedade sazonal típica de séries hidrológicas mensais. Foram avaliados dois regimes de padronização do tipo *z-score*: global e sazonal (mensal).

A normalização sazonal é motivada pelo caráter periódico das propriedades estatísticas das séries mensais de vazão, em especial média e variância, conforme discutido por Filho *et al.* (2023) no contexto de modelos autorregressivos periódicos.

Na normalização global, a série é padronizada a partir de uma única média μ e desvio padrão σ , estimados exclusivamente no conjunto de treinamento:

$$z_t = \frac{x_t - \mu}{\sigma} \quad (5.1)$$

Na normalização sazonal, assume-se ciclo-estacionariedade anual, e cada observação é padronizada com base nas estatísticas do mês m ao qual pertence:

$$z_t = \frac{x_t - \mu_m}{\sigma_m}, \quad m \in \{1, \dots, 12\} \quad (5.2)$$

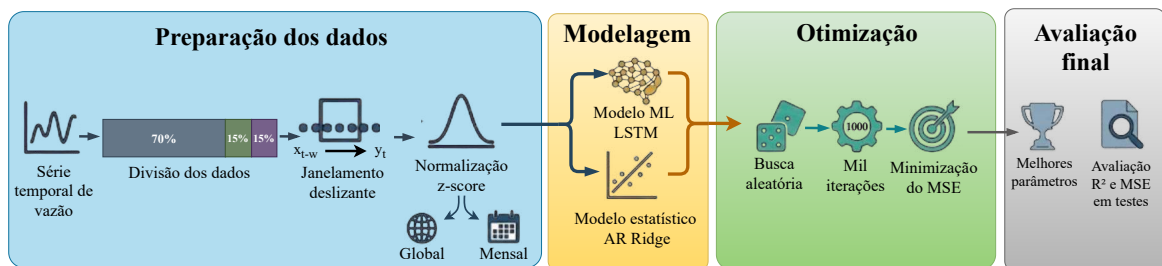
em que μ_m e σ_m representam, respectivamente, a média e o desvio padrão do mês m , calculados a partir do conjunto de treinamento.

Em ambos os regimes, os parâmetros de normalização são estimados exclusivamente no conjunto de treino, e a transformação inversa é aplicada às previsões para avaliação dos resultados na escala original da vazão.

5.3.3 Pipeline de execução e otimização

O fluxo experimental foi estruturado como um pipeline modular e totalmente automatizado, organizado desde a preparação dos dados até a avaliação final no conjunto de teste, conforme ilustrado na Figura 18.

Figura 18 – Visão esquemática do pipeline de execução RS para modelos supervisionados de aprendizado de máquina.



Fonte: Elaborado pelo autor.

5.4 Modelos de linguagem de grande escala

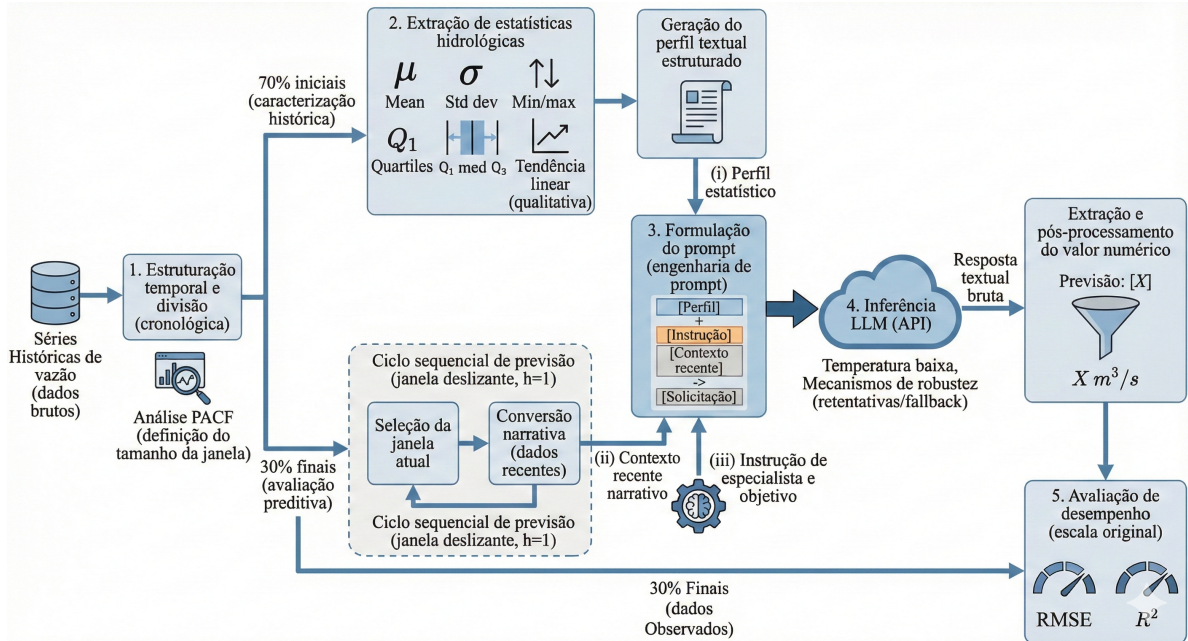
Nesta etapa, investigou-se o potencial dos LLMs para a previsão de vazão média mensal em reservatórios hidrelétricos, explorando sua capacidade de raciocínio contextual, generalização a partir de descrições narrativas e incorporação de informações estatísticas agregadas. Diferentemente das abordagens supervisionadas tradicionais, os LLMs não foram treinados diretamente sobre os dados numéricos, sendo utilizados como modelos preditivos indiretos, operando exclusivamente a partir de descrições textuais derivadas da série histórica.

A Figura 19 apresenta uma visão geral da *pipeline* metodológica adotada, sintetizando as etapas de estruturação temporal das séries, extração de estatísticas hidrológicas, formulação dos *prompts*, inferência via LLM e avaliação de desempenho na escala original.

Foram avaliadas duas famílias de modelos da plataforma Google Gemini: Gemini 2.5 Flash (GEMINI-2.5-F) e Gemini 3 Flash Preview (GEMINI-3-FP). O acesso aos modelos foi realizado por meio do *Google Generative AI SDK* (Google Cloud, 2025), utilizando infra-

estrutura gerenciada via *Vertex AI*, plataforma de serviços em nuvem do Google voltada ao desenvolvimento, implantação e operação de modelos de aprendizado de máquina. Todas as interações com os modelos foram conduzidas por chamadas diretas à interface de programação de aplicações, *application programming interface* (API), respeitando limites de tokens, políticas de *retry* e controle explícito de aleatoriedade.

Figura 19 – Pipeline metodológica adotada para previsão de vazão média mensal com LLMs.



Fonte: Elaborado pelo autor.

5.4.1 Configuração dos parâmetros dos modelos

Buscando minimizar a influência de cadeias internas extensas de raciocínio e priorizar respostas diretas e determinísticas, o nível de raciocínio interno (*thinking*) foi explicitamente controlado. Para o GEMINI-2.5-F, o *THINKING_LEVEL* foi totalmente desativado, enquanto para o GEMINI-3-FP foi configurado no modo mínimo, com orçamento de pensamento nulo.

A Tabela 4 apresenta os principais parâmetros adotados no *pipeline* preditivo.

Tabela 4 – Parâmetros experimentais utilizados na interação com os LLMs

Parâmetro	Valor / Configuração
Modelo	GEMINI-2.5-F; GEMINI-3-FP
Random Seed (GEMINI-3-FP)	42
THINKING_LEVEL (GEMINI-2.5-F)	Desativado
THINKING_LEVEL (GEMINI-3-FP)	Mínimo
MAX_OUTPUT_TOKENS	2000
TEMPERATURE	0.1
MAX_RETRIES	3
MIN_WINDOW	13
MAX_WINDOW	30
PACF_MAX_LAG	36

Fonte: Elaborado pelo autor.

O parâmetro `THINKING_LEVEL` controla o nível de raciocínio interno do modelo durante a geração das respostas, sendo ajustado para limitar explicitações intermediárias e favorecer saídas mais diretas. O parâmetro `MAX_OUTPUT_TOKENS` define o comprimento máximo da resposta gerada, enquanto `TEMPERATURE` regula o grau de aleatoriedade do processo de amostragem, com valores baixos induzindo previsões mais determinísticas. O parâmetro `MAX_RETRIES` estabelece o número máximo de novas tentativas em caso de falhas de execução ou comunicação com a API. Os parâmetros `MIN_WINDOW` e `MAX_WINDOW` determinam, respectivamente, os tamanhos mínimo e máximo da janela temporal utilizada para compor o contexto histórico fornecido ao modelo. Por fim, `PACF_MAX_LAG` especifica a defasagem máxima considerada na análise da função de autocorrelação parcial, empregada como apoio à seleção de atrasos temporalmente informativos.

5.4.2 Estruturação temporal e definição das janelas

As séries históricas de vazão média mensal foram organizadas em janelas temporais deslizantes, considerando um horizonte de previsão fixo de um passo à frente ($h = 1$). A divisão temporal respeitou estritamente a ordem cronológica dos dados, destinando 70% das observações iniciais para a extração de estatísticas descritivas e caracterização do regime hidrológico, e os 30% finais para a avaliação preditiva. Não foi empregado um conjunto de validação explícito, uma vez que não há ajuste paramétrico direto dos modelos de linguagem.

O tamanho da janela temporal utilizada como entrada foi definido de forma adaptativa para cada reservatório, com base na análise da FACP. Foram selecionadas defasagens estatisticamente relevantes dentro de um intervalo pré-estabelecido, de modo a capturar a memória

temporal dominante do sistema hidrológico analisado.

5.4.3 *Extração de estatísticas hidrológicas*

A partir do subconjunto inicial da série temporal, foram extraídas estatísticas descritivas globais, incluindo média, desvio padrão, valores mínimo e máximo, além dos quartis da distribuição. Adicionalmente, foi estimada a tendência linear de longo prazo da série, classificada de forma qualitativa como crescente, decrescente ou aproximadamente estável.

Essas informações foram posteriormente convertidas em uma descrição textual estruturada, representando o perfil hidrológico histórico do reservatório.

5.4.4 *Formulação dos prompts e interação com o LLM*

Para cada janela temporal avaliada, foi construído um *prompt* narrativo contendo: (i) a caracterização estatística do regime hidrológico do reservatório; (ii) a descrição cronológica das vazões observadas na janela recente; e (iii) uma instrução explícita para a previsão da vazão média mensal no horizonte subsequente.

O modelo foi instruído a atuar como um especialista em hidrologia, enfatizando a consideração da sazonalidade, da persistência temporal e da coerência física das previsões.

5.4.5 *Execução do pipeline preditivo*

A avaliação foi conduzida de forma sequencial, simulando um cenário operacional de previsão. Para cada janela do conjunto de teste, o *prompt* correspondente foi submetido ao LLM por meio de chamadas à API, com temperatura reduzida, visando minimizar a variabilidade estocástica das respostas e induzir um comportamento mais determinístico, condição desejável em tarefas de previsão numérica.

Foram implementados mecanismos de robustez, incluindo múltiplas tentativas de geração e uma estratégia de *fallback* baseada na média das previsões válidas obtidas. Esse procedimento assegura a continuidade do experimento mesmo diante de respostas inválidas ou inconsistentes, além de reduzir o impacto de flutuações residuais associadas ao processo de amostragem do modelo.

5.4.6 *Template do prompt utilizado*

O Quadro 1 apresenta a estrutura do *prompt* empregado na interação com os modelos de linguagem. Trata-se de um artefato metodológico descritivo, utilizado para organizar e padronizar o contexto fornecido ao modelo, não correspondendo a código executável.

Quadro 1: Estrutura do *prompt* utilizado na interação com os modelos de linguagem

```

### Perfil Hidrológico Histórico
Com base nos dados de treinamento, o reservatório apresenta um
↳ regime
caracterizado por uma vazão média de <MEDIA>, com uma
↳ variabilidade
(desvio padrão) de <DESVIO_PADRAO>.
A amplitude histórica é significativa, oscilando entre uma
↳ mínima
absoluta de <MINIMO> e picos de cheia atingindo <MAXIMO>.
Em termos de distribuição de probabilidade:
Os períodos de estiagem (25% inferiores) apresentam vazões de
↳ até <Q1>,
enquanto a mediana da série se estabiliza em <MEDIANA>.
No quartil superior, 75% dos registros históricos mantêm-se
↳ abaixo de <Q3>.
A análise linear da série sugere uma tendência <TENDENCIA>,
indicando continuidade dos padrões observados.
## Instrução
Atue como um especialista em hidrologia. Preveja a VAZÃO MÉDIA
↳ MENSAL
do reservatório de <NOME_DO_RESERVATORIO> (m3/s).
## Objetivo
Máxima acurácia para os próximos <HORIZONTE> meses.
## Contexto Histórico
Período de <DATA_INICIAL> até <DATA_FINAL>:
<DESCRICAO_NARRATIVA_DAS_VAZOES_MENSAIS>.
### Previsão Solicitada
Quais serão as vazões nos meses: <DATAS_FUTURAS>?

```

Fonte: Elaborado pelo autor.

5.5 Métricas de avaliação de desempenho

O desempenho preditivo dos modelos foi avaliado no domínio original da variável de interesse, após a aplicação da transformação inversa das normalizações utilizadas no treinamento. Neste estudo, distingue-se a métrica empregada no processo interno de ajuste daquela utilizada para a apresentação e comparação dos resultados finais.

O MSE foi adotado como função objetivo durante o treinamento e a otimização dos modelos, em razão de suas propriedades matemáticas favoráveis à convergência. Para a análise comparativa dos resultados, utilizaram-se a RMSE e o R^2 , métricas consolidadas na literatura de previsão hidrológica e séries temporais.

5.5.1 Raiz do erro quadrático médio

O RMSE quantifica a magnitude média do erro de previsão, penalizando de forma mais severa desvios de grande amplitude, característica particularmente relevante em séries hidrológicas marcadas pela ocorrência de eventos extremos de cheia. Diferentemente do MSE, o RMSE é expresso na mesma unidade da vazão (m^3/s), o que facilita sua interpretação física e a comparação direta entre modelos.

O RMSE é definido por

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (5.3)$$

em que y_i representa o valor observado, \hat{y}_i a previsão correspondente e N o número total de observações avaliadas.

5.5.2 Coeficiente de determinação (R^2)

O R^2 avalia a proporção da variância observada que é explicada pelo modelo, fornecendo uma medida adimensional da qualidade do ajuste. Valores próximos de 1 indicam elevada capacidade explicativa, enquanto valores próximos ou inferiores a 0 sugerem desempenho comparável ou inferior a um modelo ingênuo baseado na média histórica.

O R^2 é calculado como

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (5.4)$$

onde \bar{y} denota a média dos valores observados no período de avaliação.

Em conjunto, o RMSE e o R^2 permitem uma análise objetiva e complementar do desempenho dos modelos. Enquanto o RMSE fornece uma medida direta da magnitude absoluta dos erros de previsão, o R^2 expressa a capacidade do modelo em explicar a variabilidade hidrológica observada.

De modo geral, este capítulo consolidou o arcabouço metodológico do trabalho, detalhando as bases de dados utilizadas, os procedimentos de pré-processamento, a estrutura experimental e as estratégias de modelagem adotadas para a previsão de vazão. Foram apresentadas, de forma integrada, as abordagens supervisionadas e o uso de modelos de linguagem de grande escala, bem como os critérios empregados para avaliação do desempenho preditivo. Com essa fundamentação estabelecida, o capítulo seguinte dedica-se à apresentação e à discussão dos resultados obtidos, permitindo analisar comparativamente o comportamento dos modelos e interpretar seus desempenhos à luz das dinâmicas hidrológicas específicas de cada reservatório.

6 RESULTADOS E DISCUSSÕES

Este capítulo apresenta os resultados obtidos na tarefa de previsão de vazão natural afluente média mensal, e os protocolos descritos no Capítulo 5. São avaliados três modelos preditivos, distribuídos em duas linhas metodológicas distintas: (i) modelos supervisionados, representados pelo modelo AR e pela rede neural recorrente do tipo LSTM; e (ii) uma abordagem baseada em LLM, empregada como modelo preditivo indireto via engenharia de *prompt*. O desempenho final é reportado na escala original da variável (m^3/s), utilizando RMSE e R^2 .

6.1 Resultados do modelo autoregressivo

Esta seção apresenta os resultados obtidos com o modelo AR, adotado como linha de base estatística para a previsão de vazão natural afluente média mensal. O objetivo central desta análise é avaliar a capacidade de um modelo linear regularizado em representar a dependência temporal da série e servir como referência comparativa para abordagens de maior complexidade.

6.1.1 Hiperparâmetros selecionados

A Tabela 5 sintetiza, de forma integrada, os hiperparâmetros selecionados e os respectivos resultados obtidos no conjunto de validação para o modelo AR. Para cada reservatório e regime de normalização, são apresentados simultaneamente a configuração ótima do modelo e as métricas de desempenho associadas, expressas pelo R^2 e pela RMSE.

Tabela 5: Hiperparâmetros selecionados e desempenho no conjunto de validação para o modelo AR.

Reservatório	Normalização	Defasagens (p)	α	Intercepto	R^2	RMSE
Furnas	Global	12	$2,5528 \times 10^{-6}$	Não	0,5221	264,1098
	Mensal	5	$8,7706 \times 10^{-1}$	Não	0,5127	286,6165
Sobradinho	Global	12	$8,7508 \times 10^{-1}$	Não	0,5446	664,1917
	Mensal	3	$1,1969 \times 10^{-4}$	Não	0,5227	735,0411
Itaipu	Global	4	$7,4706 \times 10^{-1}$	Sim	0,5543	3 194,0499
	Mensal	5	$8,7706 \times 10^{-1}$	Sim	0,7085	2 583,1715
Ilha Solteira	Global	12	$9,9472 \times 10^{-6}$	Não	0,7339	1 095,1387
	Mensal	5	$8,1105 \times 10^{-1}$	Não	0,5983	1 414,5794

Fonte: Elaborado pelo autor.

De modo geral, sob normalização global observa-se a predominância de ordens autorregressivas mais elevadas em três dos reservatórios analisados, indicando a necessidade de

capturar dependências temporais de maior alcance, uma vez que a sazonalidade anual permanece explicitamente presente nas séries; em contraste, a normalização mensal conduz, na maioria dos casos, à seleção de modelos com menor número de defasagens, sugerindo que a remoção do padrão sazonal médio reduz a complexidade temporal exigida, ao passo que, no conjunto de validação, os ganhos mais expressivos em termos de R^2 e RMSE ocorrem para Itaipu sob normalização mensal, enquanto para Furnas e Sobradinho as diferenças entre os regimes são menos acentuadas, com ligeira vantagem para a configuração global, refletindo a importância das dependências de longo prazo nesses reservatórios.

6.1.2 Desempenho no conjunto de teste

O desempenho do AR no conjunto de teste é apresentado na Tabela 6. Os resultados indicam, de forma consistente, que a normalização mensal proporciona ganhos de acurácia, refletidos na redução do RMSE e no aumento do coeficiente de determinação R^2 em todos os reservatórios avaliados. Esse comportamento corrobora a premissa de que modelos lineares se beneficiam da maior estacionariedade induzida pela remoção dos ciclos sazonais.

Tabela 6: Desempenho do modelo AR no conjunto de teste (RMSE em m^3/s e R^2).

Reservatório	Normalização	RMSE (m^3/s)	R^2
Furnas	Global	345,81	0,6317
	Mensal	301,23	0,7174
Sobradinho	Global	883,79	0,6983
	Mensal	822,83	0,7318
Itaipu	Global	2 974,56	0,3544
	Mensal	2 487,55	0,5485
Ilha Solteira	Global	1 510,74	0,6372
	Mensal	1 428,40	0,6730

Fonte: Elaborado pelo autor.

6.2 Resultados do modelo LSTM

Nesta seção, apresentam-se os resultados empíricos da LSTM no conjunto de teste, considerando os diferentes regimes de normalização e as configurações selecionadas no conjunto de validação.

6.2.1 Hiperparâmetros selecionados

A Tabela 7 apresenta, de forma integrada, as configurações completas da rede LSTM selecionadas no conjunto de validação e os respectivos resultados de desempenho associados.

Tabela 7: Hiperparâmetros selecionados e desempenho no conjunto de validação para o modelo LSTM.

Reservatório	Norm.	w	Batch	Drop.	Neur.	Cam.	Ativ.	Dense	LR	R^2	RMSE
Furnas	Global	12	8	0,2	128	1	ReLU	32	0,001	0,6108	238,3375
	Mensal	24	64	0,2	256	1	ReLU	32	0,001	0,5674	253,1048
Sobradinho	Global	24	8	0,3	256	1	ReLU	16	0,001	0,7087	537,7905
	Mensal	18	16	0,5	128	2	ReLU	16	0,001	0,7269	518,2393
Itaipu	Global	12	32	0,2	256	1	ReLU	32	0,001	0,6010	2 583,3039
	Mensal	6	32	0,2	256	1	Sigmoid	32	0,001	0,7112	2 508,5755
Ilha Solteira	Global	24	8	0,2	64	1	ReLU	64	0,0001	0,7773	986,1017
	Mensal	24	32	0,2	256	1	Sigmoid	32	0,0001	0,7734	994,7728

Fonte: Elaborado pelo autor.

Observa-se a predominância de janelas de entrada mais longas, entre 12 e 24 meses, e de arquiteturas com uma única camada LSTM, com exceção de Sobradinho sob normalização mensal, que requer maior capacidade representacional. Os hiperparâmetros indicam ainda o uso de *dropout* moderado, tamanhos de lote ajustados por reservatório e predominância da função ReLU, com uso pontual da função sigmoidal. No conjunto de validação, destacam-se os elevados valores de R^2 para Ilha Solteira e Sobradinho, além do melhor desempenho de Itaipu sob normalização mensal, evidenciado pela redução do erro quadrático médio.

6.2.2 Desempenho no conjunto de teste

Os resultados da LSTM no conjunto de teste constam na Tabela 8. A avaliação permite verificar em que medida a maior flexibilidade do modelo, em comparação à abordagem autorregressiva linear, resulta em ganhos efetivos de desempenho fora da amostra, bem como evidenciar que o impacto do regime de normalização não é uniforme entre os reservatórios, mostrando-se dependente das características hidrológicas de cada série e devendo, portanto, ser tratado como uma escolha específica do processo de modelagem.

Tabela 8: Desempenho da LSTM no conjunto de teste (RMSE em m^3/s e R^2), agrupado por reservatório.

Reservatório	Normalização	RMSE (m^3/s)	R^2
Furnas	Global	319,0385	0,6865
	Mensal	291,7205	0,7409
Sobradinho	Global	817,2055	0,7465
	Mensal	875,8484	0,7037
Itaipu	Global	2 802,9581	0,4209
	Mensal	2 575,9980	0,5158
Ilha Solteira	Global	1 501,0042	0,6481
	Mensal	1 554,0985	0,6228

Fonte: Elaborado pelo autor.

6.3 Resultados da abordagem baseada em LLM

Esta seção avalia o uso de modelos de linguagem de grande escala como ferramenta de previsão. Foram testadas duas versões: GEMINI-2.5-F e GEMINI-3-FP.

6.3.1 Gemini 2.5 Flash

A Tabela 9 apresenta os valores de desempenho obtidos pelo GEMINI-2.5-F no conjunto de teste. Os resultados indicam desempenho competitivo em termos de RMSE, embora a capacidade explicativa, medida pelo coeficiente de determinação R^2 , varie entre os reservatórios, refletindo a sensibilidade do modelo à dinâmica específica de cada série.

Tabela 9: Desempenho do GEMINI-2.5-F no conjunto de teste.

Reservatório	RMSE (m^3/s)	R^2
Furnas	320,31	0,6877
Ilha Solteira	1 540,34	0,6294
Itaipu	2 609,18	0,5054
Sobradinho	978,96	0,6362

Fonte: Elaborado pelo autor.

6.3.2 Gemini 3 Flash Preview

A Tabela 10 apresenta o desempenho do modelo GEMINI-3-FP no conjunto de teste, considerando a previsão de um passo à frente e a avaliação na escala original da vazão. Observa-se que o modelo alcança valores de R^2 elevados em todos os reservatórios, com destaque para Sobradinho e Furnas, indicando boa capacidade de explicação da variabilidade observada, mesmo em um cenário de previsão indireta mediada por linguagem natural.

Tabela 10: Desempenho do GEMINI-3-FP no conjunto de teste.

Reservatório	RMSE (m^3/s)	R^2
Furnas	299,32	0,7273
Ilha Solteira	1 392,42	0,6972
Itaipu	2 097,54	0,6804
Sobradinho	836,53	0,7344

Fonte: Elaborado pelo autor.

6.4 Comparação geral

A Tabela 11 sumariza, para cada reservatório, o melhor resultado supervisionado (menor RMSE nas Tabelas 6 e 8) e o compara com os resultados das abordagens baseadas em LLM (Tabelas 9 e 10). Em cada reservatório, destaca-se em negrito o menor RMSE entre todas as abordagens apresentadas.

Tabela 11: Comparação geral no conjunto de teste entre modelos supervisionados (AR e LSTM) e abordagens baseadas em LLM.

Reservatório	AR		LSTM		GEMINI-2.5-F		GEMINI-3-FP	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
Furnas	301,23	0,7174	291,72	0,7409	320,31	0,6877	299,32	0,7273
Sobradinho	822,83	0,7318	817,21	0,7465	978,96	0,6362	836,53	0,7344
Itaipu	2 487,55	0,5485	2 576,00	0,5158	2 609,18	0,5054	2 097,54	0,6804
Ilha Solteira	1 428,40	0,6730	1 501,00	0,6481	1 540,34	0,6294	1 392,42	0,6972

Fonte: Elaborado pelo autor, a partir das Tabelas 6, 8, 9 e 10.

Em síntese, a comparação geral evidencia que não há superioridade universal entre as abordagens avaliadas. Os modelos supervisionados mantêm vantagem em Furnas e Sobradinho, enquanto o GEMINI-3-FP apresenta melhor desempenho em Itaipu e Ilha Solteira, superando o melhor resultado supervisionado em termos de RMSE e R^2 . Esses resultados indicam que abordagens baseadas em LLM podem alcançar desempenho competitivo, e em alguns casos superior, mesmo operando sob um paradigma de previsão indireta. Contudo, as diferenças observadas reforçam o caráter dependente do reservatório e sugerem que o uso de LLMs deve ser compreendido como complementar às técnicas supervisionadas tradicionais, exigindo cautela na interpretação comparativa em função das particularidades do processo de validação e da natureza probabilística da geração de previsões.

7 CONSIDERAÇÕES FINAIS

Este trabalho investigou a previsão univariada da vazão natural afluyente média mensal em quatro reservatórios brasileiros (Furnas, Sobradinho, Itaipu e Ilha Solteira), comparando três paradigmas *data-driven* sob um protocolo reprodutível: um modelo linear autorregressivo regularizado, uma rede recorrente LSTM e uma abordagem *prompt-based* com LLMs (GEMINI-2.5-F e GEMINI-3-FP). A análise foi conduzida em horizonte de um passo à frente, com métricas avaliadas na escala original (RMSE e R^2), permitindo uma comparação direta entre métodos com diferentes vieses indutivos e distintos níveis de interpretabilidade, em consonância com aplicações no planejamento e na operação de sistemas hídricos e energéticos (Chow *et al.*, 1988; Nash e Sutcliffe, 1970).

Os objetivos propostos foram integralmente atingidos. Inicialmente, as séries temporais mensais de vazão foram organizadas e caracterizadas quanto à dependência temporal e à sazonalidade, subsidiando a definição dos regimes de normalização e dos protocolos experimentais. Em seguida, o modelo AR com regulação de cumeeira foi implementado como linha de base estatística interpretável, permitindo avaliar o desempenho de uma abordagem linear regularizada sob diferentes pré-processamentos. Na sequência, uma arquitetura LSTM foi ajustada e avaliada, evidenciando ganhos de desempenho em reservatórios cujas séries apresentam dinâmica temporal favorável à modelagem não linear. Por fim, foi desenvolvido e aplicado um pipeline de previsão baseado em LLMs, demonstrando que modelos de linguagem podem produzir previsões competitivas mesmo sem ajuste paramétrico supervisionado, quando o problema é formulado como inferência condicionada a contexto textual estruturado.

Os resultados indicaram que não existe um método dominante para todos os reservatórios analisados, sendo o desempenho dependente tanto das características hidrológicas locais quanto do regime de normalização adotado. Entre os modelos supervisionados, a LSTM apresentou os menores valores de RMSE em Furnas (291,72 m³/s; $R^2 = 0,7409$) e em Sobradinho (817,21 m³/s; $R^2 = 0,7465$), o que sugere maior adequação em contextos nos quais a dinâmica temporal favorece a aprendizagem de relações não lineares a partir de janelas históricas (Hochreiter e Schmidhuber, 1997). Por outro lado, o GEMINI-3-FP obteve melhor desempenho em Itaipu (2097,54 m³/s; $R^2 = 0,6804$) e em Ilha Solteira (1392,42 m³/s; $R^2 = 0,6972$), evidenciando que modelos de linguagem podem produzir previsões consistentes mesmo sem ajuste paramétrico direto, quando o problema é formulado como inferência condicionada a um contexto textual estruturado (Gruver *et al.*, 2023). Observou-se ainda que a normalização mensal favoreceu, de

modo geral, os modelos lineares e parte das configurações neurais ao atenuar a componente sazonal média, embora esse efeito não tenha sido uniforme entre os reservatórios, indicando a necessidade de calibração específica do pré-processamento.

As limitações do estudo decorrem principalmente do escopo controlado adotado. Destacam-se o uso estritamente univariado, sem incorporação de variáveis exógenas como precipitação, evapotranspiração, índices climáticos ou estados de armazenamento; a escala temporal mensal e o horizonte de um passo à frente, que não abrangem previsões multi-passos nem a análise da propagação de erros; o número restrito de reservatórios, que limita a generalização dos resultados; e a avaliação predominantemente determinística, sem tratamento explícito da incerteza preditiva. Adicionalmente, cada paradigma apresenta limitações próprias. Modelos lineares são sensíveis a mudanças de regime e à presença de não estacionaridades; redes LSTM dependem fortemente da escolha de hiperparâmetros e do desenho experimental; e LLMs apresentam dependência significativa da estrutura do *prompt*, além de não oferecerem garantias teóricas de consistência quando aplicados a séries físicas, especialmente em cenários de não estacionariedade hidrológica (Milly *et al.*, 2008).

Como perspectivas de continuidade, recomenda-se a ampliação do protocolo em quatro direções principais. A primeira consiste na incorporação de abordagens multivariadas e físico-informadas, integrando variáveis hidroclimáticas com o objetivo de reduzir ambiguidades e melhorar a representação do processo hidrológico. A segunda envolve a adoção de previsões probabilísticas, com intervalos e métricas adequadas à quantificação de incerteza, visando aplicações operacionais. A terceira refere-se à avaliação de generalização e transferência, por meio de treinamento em múltiplas bacias ou reservatórios e testes fora do domínio, conforme indicado por estudos recentes baseados em LSTM (Kratzert *et al.*, 2018; Filho *et al.*, 2022). Por fim, no contexto de LLMs, sugere-se uma investigação mais sistemática envolvendo padronização de esquemas de tokenização e escala, estudos de ablação dos componentes do *prompt* e comparações entre estratégias de *prompting*, ajustes leves de parâmetros e modelos especializados em previsão de séries temporais (Jin *et al.*, 2023; Xue e Salim, 2022). Em síntese, os resultados obtidos indicam que LLMs podem ser considerados como uma alternativa complementar à modelagem estatística e neural tradicional, desde que utilizados com critérios metodológicos claros e validação adequada ao regime hidrológico analisado (Chow *et al.*, 1988; Hochreiter e Schmidhuber, 1997; Gruver *et al.*, 2023).

REFERÊNCIAS

ABADI, M.; BARHAM, P.; CHEN, J.; CHEN, Z.; DAVIS, A.; DEAN, J.; DEVIN, M.; GHEMAWAT, S.; IRVING, G.; ISARD, M.; KUDLUR, M.; LEVENBERG, J.; MONGA, R.; MOORE, S.; MURRAY, D. G.; STEINER, B.; TUCKER, P.; VASUDEVAN, V.; WARDEN, P.; WICKE, M.; YU, Y.; ZHENG, X. **TensorFlow: A system for large-scale machine learning**. arXiv, 2016. Disponível em: <https://arxiv.org/abs/1605.08695>.

AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO. **Sistema de acompanhamento de reservatórios (SAR): dados históricos do Sistema Interligado Nacional**. Brasília: ANA, 2025. Disponível em: <https://www.ana.gov.br/sar0/MedicaoSin>. Acesso em: 17 ago. 2025.

ANIL, R.; BORGEAUD, S.; ALAYRAC, J.-B.; YU, J.; SORICUT, R.; SCHALKWYK, J.; DAI, A.; HAUTH, A.; MILLICAN, K.; SILVER, D.; ANTONOGLU, I.; SCHRITTWIESER, J.; GLAESE, A.; CHEN, J.; PITLER, E.; LILLICRAP, T.; LAZARIDOU, A.; FIRAT, O.; VINYALS, O. Gemini: A family of highly capable multimodal models. arXiv, 12 2023. Disponível em: <https://arxiv.org/abs/2312.11805>.

AYANA, O.; KANBAK, D. F.; KELEŞ, M. K.; TURHAN, E. Monthly streamflow prediction and performance comparison of machine learning and deep learning methods. **Acta Geophysica**, Springer Science and Business Media LLC, v. 71, n. 6, p. 2905–2922, jan. 2023. ISSN 1895-7455. Disponível em: <http://dx.doi.org/10.1007/s11600-023-01023-6>.

BOX, G. E. P.; JENKINS, G. M. Some recent advances in forecasting and control. **Applied Statistics**, JSTOR, v. 17, n. 2, p. 91, 1968. ISSN 0035-9254. Disponível em: <http://dx.doi.org/10.2307/2985674>.

CHOW, V. T.; MAIDMENT, D. R.; MAYS, L. W. **Applied Hydrology**. New York: Tata McGraw-Hill Education, 1988. (McGraw-Hill series in water resources and environmental engineering).

FILHO, F. d. A. d. S.; ROCHA, R. V.; ESTÁCIO, B.; ROLIM, L. Z. R.; FILHO, J. D. d. A. P.; PORTO, V. C.; GUIMARÃES, S. O. Enhancing streamflow forecasting for the brazilian electricity sector: a strategy based on a hyper-multimodel. **RBRH**, FapUNIFESP (SciELO), v. 28, 2023. ISSN 1414-381X. Disponível em: <http://dx.doi.org/10.1590/2318-0331.282320230120>.

FILHO, F. J. M. N.; FILHO, F. d. A. S.; PORTO, V. C.; ROCHA, R. V.; ESTÁCIO, B. S.; MARTINS, E. S. P. R. Deep learning for streamflow regionalization for ungauged basins: Application of long-short-term-memory cells in semiarid regions. **Water**, MDPI AG, v. 14, n. 9, p. 1318, abr. 2022. ISSN 2073-4441. Disponível em: <http://dx.doi.org/10.3390/w14091318>.

Google Cloud. **SDK do Google Gen AI**. Mountain View: [s.n.], 2025. Disponível em: <https://docs.cloud.google.com/vertex-ai/generative-ai/docs/sdks/overview?hl=pt-br>. Documentação técnica. Acesso em: 18 ago. 2025.

GRUVER, N.; FINZI, M.; QIU, S.; WILSON, A. G. **Large Language Models Are Zero-Shot Time Series Forecasters**. arXiv, 2023. Disponível em: <https://arxiv.org/abs/2310.07820>.

HAMMOND, J.; GOODLING, P.; DIAZ, J.; CORSON-DOSCH, H.; HELDMYER, A.; HAMSHAW, S.; MCSHANE, R.; ROSS, J.; SANDO, R.; SIMEONE, C.; SMITH, E.; STAUB, L.; WATKINS, W.; WIECZOREK, M.; WNUK, K.; ZWART, J. Machine learning generated

streamflow drought forecasts for the conterminous united states (conus): Developing and evaluating an operational tool to enhance sub-seasonal to seasonal streamflow drought early warning for gaged locations. California Digital Library (CDL), set. 2025. Disponível em: <http://dx.doi.org/10.31223/X56X77>.

HIPEL, K. W.; MCLEOD, A. I. **Time series modelling of water resources and environmental systems**. Amsterdam: Elsevier, 1994.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural Computation**, MIT Press, v. 9, n. 8, p. 1735–1780, nov. 1997. ISSN 1530-888X. Disponível em: <http://dx.doi.org/10.1162/neco.1997.9.8.1735>.

HOERL, A. E.; KENNARD, R. W. Ridge regression: Biased estimation for nonorthogonal problems. **Technometrics**, Informa UK Limited, v. 12, n. 1, p. 55–67, fev. 1970. ISSN 1537-2723. Disponível em: <http://dx.doi.org/10.1080/00401706.1970.10488634>.

HOERL, R. W. Ridge regression: A historical context. **Technometrics**, Informa UK Limited, v. 62, n. 4, p. 420–425, out. 2020. ISSN 1537-2723. Disponível em: <http://dx.doi.org/10.1080/00401706.2020.1742207>.

HOUDT, G. V.; MOSQUERA, C.; NÁPOLES, G. A review on the long short-term memory model. **Artificial Intelligence Review**, Springer Science and Business Media LLC, v. 53, n. 8, p. 5929–5955, maio 2020. ISSN 1573-7462. Disponível em: <http://dx.doi.org/10.1007/s10462-020-09838-1>.

IQBAL, A.; SIDDIQI, T. A. Enhancing seasonal streamflow prediction using multistage hybrid stochastic data-driven deep learning methodology with deep feature selection. **Environmental and Ecological Statistics**, Springer Science and Business Media LLC, v. 32, n. 2, p. 409–459, fev. 2025. ISSN 1573-3009. Disponível em: <http://dx.doi.org/10.1007/s10651-024-00642-6>.

JIN, M.; WANG, S.; MA, L.; CHU, Z.; ZHANG, J. Y.; SHI, X.; CHEN, P.-Y.; LIANG, Y.; LI, Y.-F.; PAN, S.; WEN, Q. **Time-LLM: Time Series Forecasting by Reprogramming Large Language Models**. arXiv, 2023. Disponível em: <https://arxiv.org/abs/2310.01728>.

KORDANI, M.; BAGHERITABAR, M.; AHMADIANFAR, I.; SAMADI-KOUCHEKSARAEI, A. Forecasting water quality indices using generalized ridge model, regularized weighted kernel ridge model, and optimized multivariate variational mode decomposition. **Scientific Reports**, Springer Science and Business Media LLC, v. 15, n. 1, maio 2025. ISSN 2045-2322. Disponível em: <http://dx.doi.org/10.1038/s41598-025-99341-w>.

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**, Copernicus GmbH, v. 22, n. 11, p. 6005–6022, nov. 2018. ISSN 1607-7938. Disponível em: <http://dx.doi.org/10.5194/hess-22-6005-2018>.

LIAO, Z.; HUANG, J.; CHENG, Y.; LI, C.; LIU, P. X. A novel decomposition-based ensemble model for short-term load forecasting using hybrid artificial neural networks. **Applied Intelligence**, Springer Science and Business Media LLC, v. 52, n. 10, p. 11043–11057, jan. 2022. ISSN 1573-7497. Disponível em: <http://dx.doi.org/10.1007/s10489-021-02864-8>.

MAMUDU, L.; ALDREES, A.; DAN'AZUMI, S.; YAHAYA, A. Evaluating the predictive accuracy of some regression models and artificial neural networks in streamflow forecasting (a

case study of the kaduna river, northwest nigeria). **Modeling Earth Systems and Environment**, Springer Science and Business Media LLC, v. 11, n. 2, fev. 2025. ISSN 2363-6211. Disponível em: <http://dx.doi.org/10.1007/s40808-025-02296-0>.

MILLY, P. C. D.; BETANCOURT, J.; FALKENMARK, M.; HIRSCH, R. M.; KUNDZEWICZ, Z. W.; LETTENMAIER, D. P.; STOUFFER, R. J. Stationarity is dead: Whither water management? **Science**, American Association for the Advancement of Science (AAAS), v. 319, n. 5863, p. 573–574, fev. 2008. ISSN 1095-9203. Disponível em: <http://dx.doi.org/10.1126/science.1151915>.

NASH, J.; SUTCLIFFE, J. River flow forecasting through conceptual models part i — a discussion of principles. **Journal of Hydrology**, Elsevier BV, v. 10, n. 3, p. 282–290, abr. 1970. ISSN 0022-1694. Disponível em: [http://dx.doi.org/10.1016/0022-1694\(70\)90255-6](http://dx.doi.org/10.1016/0022-1694(70)90255-6).

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; MÜLLER, A.; NOTHMAN, J.; LOUPPE, G.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, Scikit-learn: Machine learning in python. arXiv, 2012. Disponível em: <https://arxiv.org/abs/1201.0490>.

SAMADI-KOUCHEKSARAEI, A.; CHU, X. Development of a novel modeling framework based on weighted kernel extreme learning machine and ridge regression for streamflow forecasting. **Scientific Reports**, Springer Science and Business Media LLC, v. 14, n. 1, dez. 2024. ISSN 2045-2322. Disponível em: <http://dx.doi.org/10.1038/s41598-024-81779-z>.

SIAMI-NAMINI, S.; TAVAKOLI, N.; NAMIN, A. S. A comparison of arima and lstm in forecasting time series. In: **2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)**. IEEE, 2018. p. 1394–1401. Disponível em: <http://dx.doi.org/10.1109/ICMLA.2018.00227>.

WOLD, H. **A Study in the Analysis of Stationary Time Series. By Herman Wold. [Pp. 214 + viii. Almqvist and Wiksells Boktryckeri-A.-B., Uppsala. 1938. Price kr. 6.]**. Cambridge University Press (CUP), 1939. v. 70. 113–115 p. ISSN 2058-1009. Disponível em: <http://dx.doi.org/10.1017/S0020268100011574>.

WORKNEH, H. A.; JHA, M. K. Utilizing deep learning models to predict streamflow. **Water**, MDPI AG, v. 17, n. 5, p. 756, mar. 2025. ISSN 2073-4441. Disponível em: <http://dx.doi.org/10.3390/w17050756>.

XUE, H.; SALIM, F. D. **PromptCast: A New Prompt-based Learning Paradigm for Time Series Forecasting**. arXiv, 2022. Disponível em: <https://arxiv.org/abs/2210.08964>.