

Previsão de Vazão em Cenário de Mudanças Climáticas: uma Análise Comparativa de Modelos LSTM e TFT na Sub-bacia do Rio Canoas (Bacia do Rio Uruguai)

Jeffley Garçon¹, Fernando Grison¹, Giancarlo D. Salton¹

¹Universidade Federal da Fronteira Sul
Campus Chapecó

jeffley.garcon@estudante.uffs.edu.br
{fernando.grison,gian}@uffs.edu.br

Resumo. A modelagem hidrológica na Bacia do Rio Uruguai enfrenta desafios decorrentes da não-estacionariedade climática, que fragiliza a validade de padrões históricos. Este trabalho avalia as arquiteturas Temporal Fusion Transformer (TFT) e Long Short-Term Memory (LSTM) na previsão de vazão diária e na detecção de estiagens, comparando um regime histórico (1980–1990) e um regime recente (2011–2018) com o índice ONI como covariável. O LSTM mostrou desempenho competitivo no nowcasting ($t + 1$), enquanto o TFT foi superior em horizontes subsazonais (7, 15 e 30 dias), com $\text{LogNSE} > 0,86$ e maior estabilidade em baixas vazões (RMSE_{dry}). A interpretabilidade do TFT sustenta sua capacidade de capturar forçantes físicas relevantes. Conclui-se que a arquitetura deve ser escolhida conforme o horizonte de previsão, sendo o TFT mais resiliente ao planejamento hídrico sob mudanças climáticas.

Palavras-chave: Não-estacionariedade; Temporal Fusion Transformer; LSTM; Previsão de vazão; Bacia do Rio Uruguai.

Abstract. Hydrological modeling in the Uruguay River Basin faces challenges due to climate non-stationarity, which weakens the validity of historical patterns. This study evaluates the Temporal Fusion Transformer (TFT) and Long Short-Term Memory (LSTM) architectures for daily streamflow forecasting and drought detection, comparing a historical regime (1980–1990) and a recent regime (2011–2018) while using the Oceanic Niño Index (ONI) as a covariate. The LSTM achieved competitive performance for nowcasting ($t + 1$), whereas the TFT outperformed it at subseasonal horizons (7, 15, and 30 days), with $\text{LogNSE} > 0.86$ and greater stability under low-flow conditions (RMSE_{dry}). The TFT's interpretability supports its ability to capture relevant physical drivers. Overall, model choice should be guided by the forecasting horizon, with the TFT offering greater resilience for water planning under climate change.

Keywords: Non-stationarity; Temporal Fusion Transformer; LSTM; Streamflow Forecasting; Uruguay River Basin.

1. Introdução

A Bacia Hidrográfica do Rio Uruguai (BHRU) desempenha um papel estratégico para a segurança energética e alimentar da Região Sul do Brasil, pois seus afluentes respondem

por uma parcela significativa da capacidade de geração hidrelétrica regional e sustentam uma agroindústria intensiva [Marcuzzo 2017].

A BHRU apresenta vulnerabilidade a rupturas hidrológicas, com tendências de redução na disponibilidade hídrica que ameaçam a estabilidade da geração hidrelétrica e da agroindústria regional. Embora a previsão de vazão em todo o espectro seja relevante, a literatura aponta que a maior criticidade para o planejamento reside na capacidade de antecipar períodos de baixa vazão. Este trabalho foca na avaliação da robustez das arquiteturas *Temporal Fusion Transformer* (TFT) e *Long Short-Term Memory* (LSTM) especificamente na detecção e previsão de eventos de seca, onde a precisão é fundamental para a gestão de riscos e segurança hídrica [Marcuzzo 2017, Rosa et al. 2022].

Historicamente, o planejamento de recursos hídricos baseou-se na hipótese de estacionariedade — a premissa de que as propriedades estatísticas das séries hidrológicas, como média e variância, oscilam dentro de um envelope fixo e previsível ao longo do tempo. Contudo, essa premissa tem sido desafiada pelo conceito de não-estacionariedade, amplamente discutido após o manifesto seminal de [Milly et al. 2008]. No contexto atual, a não-estacionariedade manifesta-se através de rupturas hidrológicas (*hydrological shifts*): mudanças abruptas no regime de vazão que alteram a distribuição de probabilidade dos eventos extremos, invalidando modelos calibrados exclusivamente em regimes passados.

A BHRU é particularmente vulnerável a esse fenômeno. Estudos recentes, como o de [Rosa et al. 2022], indicam uma tendência de redução significativa na vazão média de diversos rios no Sul do Brasil, com sub-bacias apresentando quedas de vazão entre 15% e 30% após 1990. A gênese dessa alteração é multifatorial, impulsionada pela interação entre mudanças climáticas globais e alterações regionais no uso do solo [Rosa et al. 2022]. Além disso, a variabilidade hidrológica local é fortemente modulada por teleconexões climáticas, especificamente o *El Niño-Southern Oscillation* (ENSO), cujas fases negativas estão diretamente associadas aos períodos de estiagem prolongada que motivam esta pesquisa.

Para capturar essa complexidade, a modelagem moderna deve incorporar indicadores climáticos robustos. Este trabalho utiliza o Índice Oceânico Niño (ONI), o principal indicador da NOAA (*National Oceanic and Atmospheric Administration*) para monitorar o ENSO, definido pela média móvel trimestral das anomalias da temperatura da superfície do mar no Pacífico Equatorial. A inclusão do ONI visa permitir que os modelos identifiquem padrões de variabilidade que antecedem eventos críticos de escassez hídrica.

A hipótese da estacionariedade, que fundamenta modelos hidrológicos convencionais, mostra-se insuficiente frente às novas distribuições de probabilidade observadas. Modelos preditivos calibrados exclusivamente sob padrões passados subestimam a severidade das estiagens atuais. Este trabalho investiga a capacidade de arquiteturas de *Deep Learning* para lidar com essa variabilidade, não apenas como uma alternativa, mas como uma ferramenta capaz de aprender dinâmicas complexas sob condições de regime climático em transição.

O objetivo central desta pesquisa é avaliar o desempenho das arquiteturas *Temporal Fusion Transformer* (TFT) e *Long Short-Term Memory* (LSTM) na previsão de vazão e na detecção de eventos de seca em bacias hidrográficas sujeitas a essas rupturas. Diferente de abordagens que buscam a precisão média em todo o espectro hidrológico, este trabalho

foca especificamente na resiliência preditiva em regimes de baixa vazão, onde os impactos socioeconômicos e a degradação da performance dos modelos são mais críticos para a segurança hídrica.

Além do diagnóstico do problema, a pesquisa propõe uma investigação sobre a adequação de novas tecnologias. Enquanto modelos tradicionais e redes recorrentes simples (LSTM) podem apresentar limitações em capturar dinâmicas não-lineares sob mudança de regime, a literatura recente [Amanambu et al. 2022, Lim et al. 2021] sugere que arquiteturas baseadas em *Transformers* oferecem maior robustez. O *Temporal Fusion Transformer* (TFT), em particular, destaca-se teoricamente por sua capacidade de interpretar a importância variável de *inputs* ao longo do tempo. Testar se essa vantagem teórica se traduz em ganho real de performance na previsão de secas no Brasil é a lacuna que este estudo visa preencher.

O trabalho está organizado da seguinte forma: a Seção 2 apresenta os fundamentos de modelagem hidrológica e a discussão sobre estacionariedade e não-estacionariedade; a Seção 3 descreve os conceitos de *Transformers* e detalha a arquitetura Temporal Fusion Transformer; a Seção 4 revisa os trabalhos relacionados; a Seção 5 apresenta a metodologia, incluindo bases de dados, pré-processamento, configuração experimental e métricas de avaliação; a Seção 6 apresenta e discute os resultados, com ênfase na comparação entre LSTM e TFT sob diferentes regimes climáticos; por fim, a Seção 7 sintetiza as conclusões e os potenciais desdobramentos para pesquisas futuras.

2. Modelagem Hidrológica e Não-Estacionariedade

A modelagem hidrológica constitui uma ferramenta importante para a gestão de recursos hídricos, pois possibilita a simulação e a previsão do comportamento de sistemas complexos, como bacias hidrográficas. No caso da BHRU, sistema transfronteiriço de relevância para a agroindústria e para o setor de geração hidrelétrica no Sul do Brasil, a previsão de vazões com níveis adequados de acurácia representa um elemento central para o planejamento energético e para a segurança hídrica regional [Marcuzzo 2017].

Historicamente, tanto os modelos conceituais baseados em processos físicos quanto os modelos estatísticos orientados por dados foram desenvolvidos sob a premissa de que o sistema hidrológico é estacionário. A estacionariedade pressupõe que os mecanismos que governam a dinâmica fluvial permanecem constantes ao longo do tempo. Do ponto de vista técnico, implica que propriedades estatísticas de séries históricas de vazão — como média, variância e padrões sazonais — mantêm-se estáveis ou previsíveis.

Esse pressuposto, entretanto, vem sendo invalidado por evidências empíricas. A não-estacionariedade não é apenas uma hipótese teórica, mas uma realidade quantificável na região de estudo. Análises recentes sobre a dinâmica hidrológica no Sul do Brasil, como as conduzidas por [Rosa et al. 2022], apontam para uma retração sistemática na disponibilidade hídrica. Em sub-bacias do Rio Uruguai, observou-se uma redução na vazão média anual da ordem de 15% a 25% nas últimas três décadas, exacerbada por eventos de seca mais frequentes e severos.

Essa magnitude de variação desafia a calibração de modelos tradicionais, pois altera as curvas de permanência e a relação chuva-vazão, sugerindo que parâmetros ajustados em décadas passadas podem subestimar a severidade das estiagens atuais. Além dessas

tendências de longo prazo, a variabilidade interanual é fortemente modulada por fenômenos climáticos de grande escala, como o *El Niño-Southern Oscillation* (ENSO), monitorado através do Índice Oceânico Niño (ONI) [NOAA Climate Prediction Center 2025].

Nesse contexto, a modelagem hidrológica baseada em processos físicos e modelos conceituais permanece como um pilar fundamental da ciência hídrica, conferindo transparência aos processos de transformação chuva-vazão. O objetivo deste trabalho não é substituir tais ferramentas consolidadas, mas sim complementá-las. A literatura recente indica que arquiteturas de *Deep Learning* oferecem flexibilidade estatística para representar não-linearidades complexas que, ocasionalmente, desafiam as parametrizações conceituais fixas quando as premissas de estacionariedade são violadas e o regime de forçantes climáticas se desloca significativamente do período de calibração [Reichstein et al. 2021].

3. Previsão de Séries Temporais com *Deep Learning*

A modelagem baseada em dados compreende um vasto espectro de técnicas de *Machine Learning* (ML). Enquanto métodos clássicos de ML (como regressão ou florestas aleatórias) frequentemente dependem de uma engenharia de atributos (*feature engineering*) manual para lidar com a não-linearidade, o *Deep Learning* — subárea na qual se inserem as arquiteturas LSTM e TFT — caracteriza-se pela extração hierárquica e automatizada de padrões complexos a partir de séries temporais brutas. Esta capacidade de aprendizado profundo é crucial para identificar as dinâmicas não-lineares da bacia sob regimes hidrológicos instáveis e desafia as premissas lineares dos modelos hidrológicos tradicionais. Nesse cenário, abordagens orientadas por dados (*data-driven*) têm adquirido destaque, impulsionadas pela disponibilidade de extensos conjuntos de dados — como o *CAMELS-BR* [Chagas et al. 2020] — e pelo avanço do poder computacional.

Ressalta-se que o uso do *CAMELS-BR* padroniza a entrada de dados e facilita a reprodutibilidade científica [Chagas et al. 2020]. Embora o uso de dados pré-processados possa abstrair certas particularidades locais, a consistência deste *dataset* já foi validada em diversos estudos hidrológicos nacionais, permitindo focar a análise na arquitetura dos modelos preditivos.

3.1. Long Short-Term Memory (LSTM)

Segundo [Hochreiter and Schmidhuber 1997], embora as *Recurrent Neural Networks* (RNNs) sejam redes utilizadas para modelar séries temporais, elas apresentam dificuldades para aprender dependências de longo prazo devido ao problema do desaparecimento de gradiente. Para mitigar essa limitação, foi proposta a arquitetura *Long Short-Term Memory* (LSTM).

O LSTM constitui um tipo especializado de RNN que emprega uma estrutura interna composta por células de memória e mecanismos de controle denominados *gates* (portões), responsáveis por regular o fluxo de informação. Essa estrutura é composta por três mecanismos principais:

1. Forget gate (f_t): decide qual parcela da memória anterior deve ser descartada.
2. Input gate (i_t): controla a incorporação de novas informações à célula de memória.
3. Output gate (o_t): determina qual parte do estado interno será utilizada para compor a saída.

Formalmente, o *forget gate* é definido por:

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

A capacidade de preservar e manipular informações ao longo de sequências extensas torna a arquitetura LSTM adequada para aplicações hidrológicas. Estudos realizados no contexto brasileiro, como [Doninelli et al. 2017], validaram seu uso na Bacia do Rio Apuaê (sub-bacia do Uruguai), estabelecendo-o como um *baseline* robusto para a região.

3.2. Transformers e Interpretabilidade

Embora o LSTM represente um avanço, sua natureza intrinsecamente sequencial limita a eficiência computacional e a capacidade de estabelecer relações diretas entre eventos muito distantes no tempo. Para superar isso, a arquitetura *Transformer* [Vaswani et al. 2017] introduziu o mecanismo de *self-attention* (autoatenção), que permite ao modelo considerar simultaneamente todos os elementos de uma sequência, ponderando a importância de cada ponto temporal passado para a previsão atual.

Para adaptar o poder dos *Transformers* às especificidades de séries temporais multivariadas, [Lim et al. 2021] propuseram o *Temporal Fusion Transformer* (TFT). Esta arquitetura híbrida combina a capacidade de aprendizado sequencial de camadas LSTM com a visão global dos mecanismos de atenção.

Uma crítica recorrente à aplicação de *Deep Learning* na hidrologia refere-se à natureza “caixa-preta” (*black-box*) das redes neurais, onde o processo de decisão é opaco. O TFT foi projetado especificamente para mitigar essa limitação através de componentes de interpretabilidade intrínseca:

- **Variable Selection Networks:** Seleccionam automaticamente quais variáveis de entrada (ex: precipitação, evapotranspiração, ONI) são mais relevantes para a previsão em cada passo de tempo, descartando ruídos.
- **Interpretable Multi-Head Attention:** Permite visualizar os “pesos de atenção”, indicando em quais momentos do passado o modelo focou para gerar a previsão atual (ex: identificando se o modelo está reagindo a uma chuva de ontem ou a um padrão sazonal de meses atrás).

Essa capacidade de interpretação responde à necessidade de transparência na modelagem hidrológica, permitindo verificar se o modelo está aprendendo relações fisicamente coerentes ou apenas correlações espúrias [Reichstein et al. 2021]. Estudos recentes confirmam a superioridade dessa abordagem para previsão de extremos hidrológicos [Amanambu et al. 2022, Fayer et al. 2023].

4. TRABALHOS RELACIONADOS

A aplicação de redes neurais *Long Short-Term Memory* (LSTM) para a previsão de vazão no Brasil já foi validada em diversos estudos, que atestam a robustez da arquitetura para as condições hidrológicas do país e oferecem uma base de comparação para a presente pesquisa. Contudo, a literatura ainda carece de análises que avaliem o impacto da não-estacionariedade na confiabilidade de modelos de *deep learning* treinados sob regimes climáticos passados — lacuna que este trabalho se propõe a investigar.

Para situar a contribuição desta pesquisa frente à literatura existente, a Tabela 1 sintetiza os principais estudos sobre modelagem hidrológica baseada em dados. A análise comparativa demonstra que, enquanto a literatura foca majoritariamente na validação de acurácia global dentro de regimes hidrológicos homogêneos, este trabalho introduz um "teste de estresse temporal". Esta abordagem permite distinguir se a eficácia do modelo deriva da memorização de padrões pretéritos ou da capacidade real de generalização frente à não-estacionariedade.

Entre esses trabalhos, destaca-se [Doninelli et al. 2017], que analisou a Bacia do Rio Apuaê (sub-bacia do Rio Uruguai) e utilizou uma LSTM para prever vazões diárias a partir de séries históricas de vazão e precipitação (1976–2016). Além de confirmar o bom desempenho do LSTM, o estudo valida sua aplicabilidade à região de interesse deste trabalho; contudo, o treinamento sobre toda a série pressupõe estacionariedade e não testa a degradação da performance diante de mudanças de regime. Esta é a primeira lacuna que esta pesquisa busca endereçar: a robustez do LSTM em um cenário de não-estacionariedade.

De modo geral, a literatura nacional tende a otimizar arquiteturas e comparar o LSTM a modelos estatísticos e de *machine learning*, corroborando sua capacidade de capturar relações não-lineares do processo chuva–vazão, mas avaliando a acurácia dentro de um único regime, com a série histórica tratada como bloco homogêneo. Assim, o foco aqui não é revalidar a eficácia do LSTM, e sim testar sua confiabilidade quando a hipótese de estacionariedade é relaxada.

Com o avanço dos *Transformers*, a previsão hidrológica tem explorado arquiteturas capazes de superar limitações de modelos puramente sequenciais. O *Temporal Fusion Transformer* (TFT) é uma proposta híbrida voltada à previsão multivariada [Lim et al. 2021], cuja relevância tem sido confirmada por estudos recentes.

No estudo de [Amanambu et al. 2022], um modelo baseado em *Transformer* superou o LSTM na previsão de seca hidrológica em uma bacia nos Estados Unidos, sobretudo em eventos de baixa vazão, reforçando sua pertinência para o problema central desta pesquisa. Ainda assim, o desenho experimental não mede o impacto de dados de treino temporalmente obsoletos, isto é, o “teste de estresse” temporal proposto aqui.

No Brasil, [Fayer et al. 2023] validou o TFT na previsão de afluência em um reservatório hidroelétrico, destacando sua flexibilidade com múltiplas variáveis e desempenho superior a *baselines* como AutoARIMA e *DeepAR*. Entretanto, assim como nos demais trabalhos, a validação ocorre em um único regime climático, sem examinar como a não-estacionariedade afeta a performance do modelo.

Em síntese, a literatura aponta que modelos baseados em *Transformer* tendem a ser mais robustos que o LSTM em eventos extremos (como secas) e que o TFT já apresenta resultados promissores no cenário brasileiro, justificando sua seleção e a comparação direta com o LSTM. Em escala global, [Koya and Roy 2024] reforça essa escolha ao mostrar o TFT superando consistentemente tanto o LSTM quanto um *Transformer* puro em 2.610 bacias.

Por fim, a motivação desta pesquisa é ancorada por evidências de que a não-estacionariedade já é observável no Brasil. Em [Chagas et al. 2022], uma análise de 886 estações hidrométricas indica mudanças no ciclo hidrológico nas últimas quatro décadas,

com intensificação de secas e cheias e sinais de aceleração do ciclo em parte substancial da área estudada (29%). O estudo confirma que “as regras do sistema” mudaram, mas não quantifica a consequência direta para a prática de modelagem preditiva; é exatamente esse elo — o impacto da não-estacionariedade na confiabilidade de modelos de *deep learning* treinados sob regimes anteriores — que este trabalho se propõe a investigar.

Tabela 1. Síntese Comparativa dos Trabalhos Relacionados

Autor (Ano)	Método / Contexto	Resultado Chave	Limitação
Doninelli et al. (2017)	LSTM Bacia do Rio Uruguai (ANA)	Valida o LSTM como eficaz para a região hidrográfica de estudo.	Assume estacionariedade e não testa impacto de mudanças climáticas.
Lim et al. (2021)	TFT Benchmark Teórico	Propõe a arquitetura TFT para séries temporais multivariadas complexas.	Foco na arquitetura geral, sem validação específica na hidrologia brasileira.
Amanambu et al. (2022)	Transformer vs. LSTM Bacias USGS (EUA)	Demonstra que Transformers superam LSTM na previsão de secas.	Foca na comparação de modelos, mas não no impacto da obsolescência dos dados.
Fayer et al. (2023)	TFT vs. Baselines Reservatório (Brasil)	Valida o TFT no Brasil com performance superior ao ARIMA/DeepAR.	Aplicação em regime único, sem realizar o "teste de estresse" temporal.
Koya & Roy (2024)	TFT (Híbrido) Estudo Global	Confirma que o TFT supera tanto LSTM quanto Transformer puro.	Análise global de média, sem foco na degradação por não-estacionariedade local.
Chagas et al. (2022)	Análise Estatística CAMELS-BR	Comprova fisicamente a não-estacionariedade nos rios brasileiros.	Diagnostica o problema físico, mas não avalia o impacto nos modelos de IA.

Fonte: Elaborado pelo autor (2025).

A análise dos trabalhos relacionados permite posicionar a presente pesquisa em uma intersecção única. Enquanto [Doninelli et al. 2017] validam o LSTM localmente e [Fayer et al. 2023] demonstram o potencial do TFT no Brasil, nenhum estudo implementa um desenho experimental para medir explicitamente o impacto da não-estacionariedade diagnosticada por [Chagas et al. 2022] na confiabilidade dessas ferramentas.

Este trabalho propõe uma metodologia para avaliar o impacto da não-estacionariedade climática na confiabilidade de modelos hidrológicos. Ao comparar modelos treinados com dados antigos (1980-1990) contra modelos atualizados (pós-1991), este trabalho preenche a lacuna entre a detecção do problema físico e a validação das ferramentas computacionais, oferecendo uma análise crítica sobre a validade das práticas de modelagem preditiva atuais.

5. Metodologia

Este estudo adota o framework CRISP-DM para estruturar a previsão multi-horizonte da vazão na bacia 71200000. O objetivo central é avaliar como modelos de Deep Learning respondem à não-estacionariedade hidrológica, um fenômeno que desafia as premissas

clássicas de modelagem estatística [Milly et al. 2008]. O fluxo metodológico, detalhado na Figura 1, articula-se em quatro pilares: seleção da área, integração de dados, modelagem preditiva e validação experimental.

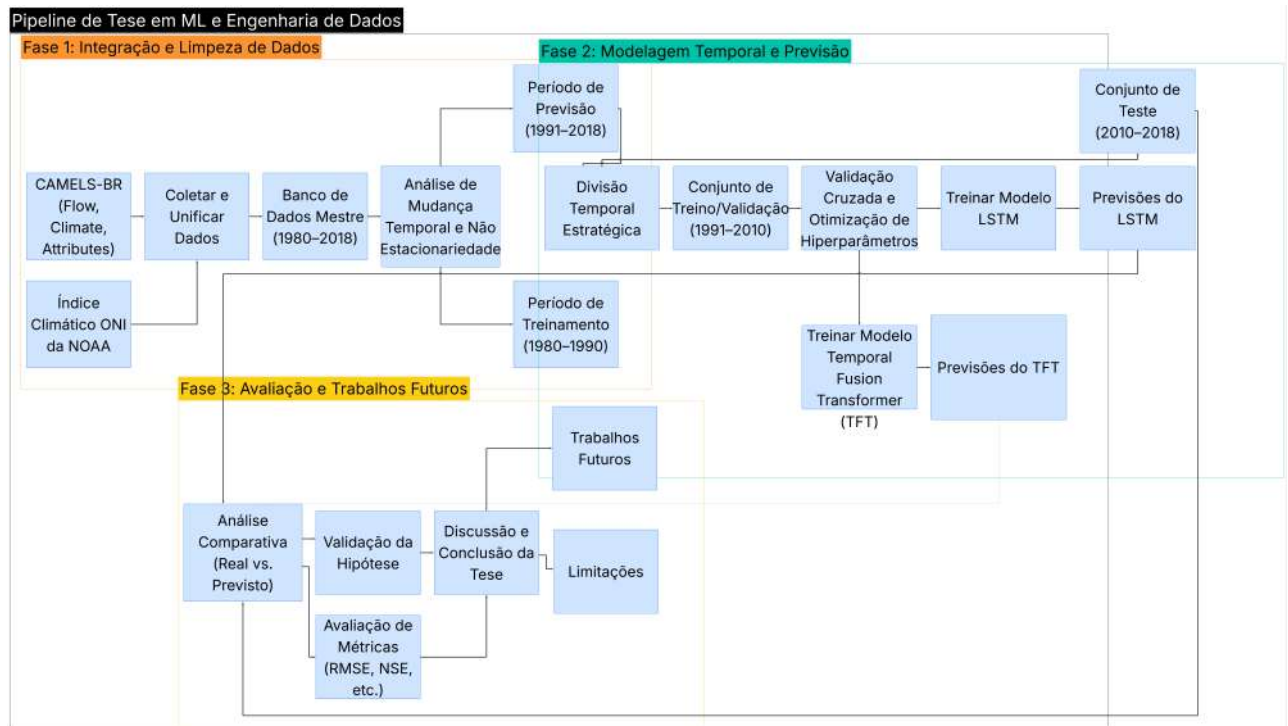


Figura 1. Fluxo metodológico do pipeline de previsão de vazão.

5.1. Área de Estudo e Critérios de Seleção

A área de estudo compreende a sub-bacia hidrográfica monitorada pela estação fluviométrica 71200000, situada no domínio da Bacia do Rio Uruguai, conforme ilustrado na Figura 2. Esta unidade hidrográfica possui uma área de drenagem de 1.010 km², altitude média de 1.188,47 metros e cobertura florestal predominante de 86,3%. Tais características indicam uma bacia com elevado grau de preservação ambiental, onde a cobertura vegetal desempenha um papel fundamental na regulação da resposta hidrológica e na manutenção do fluxo de base.

A seleção desta unidade foi conduzida por meio de um protocolo automatizado, implementado via script `rank_camels_br_basins.py`, que avaliou as bacias do conjunto CAMELS-BR [Chagas et al. 2020] com base em critérios rigorosos de elegibilidade. Foram priorizadas bacias com séries temporais superiores a 30 anos (1980–2018) e elevada completude de dados, visando garantir a consistência estatística e minimizar os impactos de eventuais falhas telemétricas. Adicionalmente, exigiu-se a disponibilidade integral dos oito atributos fisiográficos estáticos — como permeabilidade geológica, profundidade do solo e declividade —, requisito essencial para a integração do contexto físico nas arquiteturas de *Temporal Fusion Transformer* (TFT).

A estação 71200000 destacou-se com um índice de recomendação de 0,96, consolidando-se como o objeto de estudo ideal. O ranqueamento considerou a sensi-

bilidade da bacia à variabilidade da precipitação e sua importância estratégica para a segurança hídrica regional. Os dados indicam que a unidade apresenta uma resposta hidrológica representativa tanto de eventos extremos de cheia quanto de períodos críticos de estiagem. Dessa forma, a bacia selecionada oferece um cenário experimental robusto para avaliar a performance preditiva dos modelos em horizontes variados, assegurando a validade das análises de não-estacionariedade propostas neste estudo.

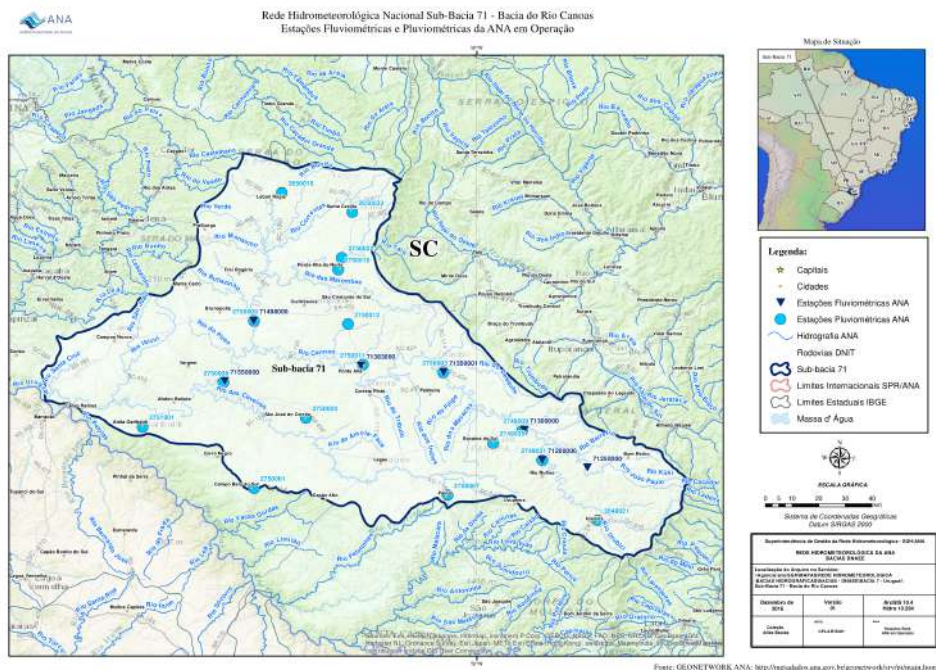


Figura 2. Mapa de localização da sub-bacia 71 (Bacia do Rio Canoas).

5.2. Integração dos Dados

Após a definição da área de estudo, os dados foram consolidados no arquivo `.dist/71200000_master_dataset.csv`, abrangendo uma série histórica de 38 anos (1980–2018), composta por 14.245 observações diárias. O conjunto foi estruturado em três eixos informacionais fundamentais para capturar a dinâmica hidrológica da bacia.

O primeiro eixo é a dinâmica hídrica, representada pelas séries de vazão diária (*streamflow_m3s*), que servem como variável alvo para o treinamento dos modelos, permitindo o aprendizado dos processos de recessão e fluxo de base. O segundo eixo compreende as causas climáticas, que incluem precipitação, temperatura e evapotranspiração real. Além destas, foi incorporado o Índice Oceânico Niño (ONI), aplicado com defasagens temporais (*lags*) de ± 3 meses. A escolha deste intervalo baseia-se na literatura hidrológica regional, que aponta um tempo de resposta defasado entre o estabelecimento das anomalias de temperatura no Pacífico Equatorial e seus efeitos sobre a precipitação e a vazão no Sul do Brasil [Grimm and Silva 2013]. Essa defasagem de até um trimestre é essencial para permitir que modelos preditivos capturem a propagação do sinal climático do ENSO como uma forçante exógena, auxiliando na antecipação de períodos de estiagem que caracterizam a não-estacionariedade da Bacia do Rio Uruguai.

A integração desses dados permite que as arquiteturas de *Deep Learning* compreendam não apenas a variação temporal das variáveis meteorológicas, mas também o contexto

físico inerente à bacia hidrográfica. Essa combinação cria uma base de dados robusta e consistente, essencial para a análise comparativa entre diferentes regimes climáticos e para a avaliação da resiliência dos modelos frente às mudanças observadas no comportamento hidrológico.

5.3. Preparação dos Dados e Modelagem

O *pipeline* de tratamento de dados, implementado no módulo `optuna_hydro_utils.py`, priorizou a qualidade e a integridade da série temporal. Lacunas de até cinco dias foram corrigidas por interpolação linear, garantindo a continuidade do registro. A variável alvo foi submetida a uma transformação logarítmica, $\log(y + 1)$, visando mitigar a influência de valores extremos e aproximar a distribuição da normalidade. Para evitar o vazamento de informações do futuro para o passado *data leakage*, a normalização das variáveis foi realizada via `StandardScaler`, com parâmetros de média e desvio padrão ajustados exclusivamente sobre o conjunto de treinamento.

A estratégia de validação adotada foi o *teste de estresse temporal*. Os períodos de dados foram divididos em *Regime Passado* (1980–1990) e *Regime Recente* (1991–2010), com um conjunto de teste independente *hold-out* entre 2011 e 2018. Esta segmentação não é aleatória; ela foi desenhada especificamente para quantificar a degradação de performance que ocorre quando um modelo, treinado com padrões de variabilidade do século passado, é submetido a dados recentes que refletem um clima alterado. Ao confrontar o modelo com essa "variabilidade obsoleta", este desenho experimental permite distinguir se a acurácia do modelo é fruto de uma compreensão física real do processo hidrológico ou apenas uma memorização de regimes climáticos que não representam mais a realidade atual da bacia.

As arquiteturas LSTM e TFT, implementadas com a biblioteca *pytorch-forecasting*, foram otimizadas via *Optuna* utilizando o algoritmo de amostragem bayesiana *Tree-structured Parzen Estimator* (TPE). Esta etapa buscou explorar o espaço de hiperparâmetros de forma eficiente, executando 30 *trials* por configuração experimental, orquestrados pelos *scripts* `tune_lstm_v2_optuna.py` e `tune_tft_v2_optuna.py`. Para prevenir o sobreajuste *overfitting*, integrou-se o critério de *Early Stopping*, interrompendo o treinamento caso o desempenho na validação estagnasse.

Quanto à janela de observação *lookback*, para o cenário de *nowcasting* ($t + 1$), testaram-se intervalos de 30, 60, 90 e 180 dias. Já para os horizontes de previsão subsazonal (7, 15 e 30 dias), a janela base foi calibrada em 48 dias — estimativa derivada do ciclo médio de resposta hidrológica da bacia — sendo as janelas subsequentes de 96, 144 e 192 dias definidas como múltiplos diretos ($2\times$, $3\times$ e $4\times$) dessa base, garantindo que o modelo capturasse escalas temporais coerentes com a memória da bacia.

5.4. Experimentos

A avaliação de desempenho dos modelos baseou-se em um conjunto de métricas complementares, selecionadas para capturar diferentes nuances da resposta hidrológica: o Índice de Eficiência de Nash-Sutcliffe (NSE) [Nash and Sutcliffe 1970], a Eficiência de Kling-Gupta (KGE) [Gupta et al. 2009], o LogNSE [Reichstein et al. 2021] e o $RMSE_{dry}$. Enquanto o NSE e o KGE oferecem uma visão global do ajuste do modelo, o LogNSE foi empregado para conferir maior sensibilidade aos fluxos de baixa vazão, atenuando o peso

desproporcional que grandes eventos de cheia exercem sobre as métricas tradicionais. Adicionalmente, o $RMSE_{dry}$ foi calculado focando exclusivamente nas amostras abaixo do percentil histórico Q_{90} , permitindo quantificar a resiliência dos modelos em períodos de estiagem.

Todas as métricas foram processadas no espaço original da variável alvo, após a aplicação da transformação inversa e a devida correção de viés. Este protocolo comparativo, aplicado sobre os regimes climáticos distintos, permite mensurar de forma quantitativa o impacto da não-estacionariedade sobre a confiabilidade das previsões.

O custo computacional foi monitorado durante a execução dos *trials* de otimização no *Optuna*. O tempo médio de treinamento variou conforme a arquitetura: enquanto o LSTM, devido à sua natureza sequencial, convergia em aproximadamente 15 minutos por *trial* em GPU NVIDIA CUDA 11.8, o TFT demandou um tempo cerca de 3 vezes superior, justificado pela complexidade dos mecanismos de atenção *multi-head* e pela carga computacional acrescida pela busca exaustiva no espaço de hiperparâmetros. Esta disparidade reflete o custo inerente à maior capacidade de modelagem hierárquica do TFT em comparação à arquitetura recorrente simples do LSTM.

Os experimentos foram executados em ambiente Python 3.10+, utilizando as bibliotecas *PyTorch*, *pytorch-forecasting* e *Optuna*. A infraestrutura computacional baseou-se em processamento acelerado por GPU (CUDA 11.8), sendo todas as sementes aleatórias (*seeds*) fixadas em 42 para garantir a replicabilidade integral dos resultados e a consistência das comparações entre as arquiteturas propostas¹. A transparência metodológica, materializada através do versionamento e abertura do código, visa não apenas a validação dos achados deste estudo, mas também o incentivo ao reuso da metodologia em outras bacias hidrográficas do conjunto CAMELS-BR, contribuindo para a evolução das técnicas de previsão hidrológica baseadas em inteligência artificial.

6. Resultados e Discussão

Os resultados apresentados a seguir detalham o desempenho das arquiteturas TFT e LSTM frente aos desafios da não-estacionariedade hidrológica. A análise foi estruturada comparando o comportamento dos modelos sob dois cenários temporais distintos: o *Regime Passado* (1980–1990), utilizado como base de calibração, e o *Regime Recente* (2011–2018), que representa o conjunto de teste sob condições climáticas alteradas. O panorama geral de desempenho, consolidado na Figura 3, evidencia a trajetória de eficiência das arquiteturas frente aos diferentes horizontes temporais."

¹https://github.com/Jeffleyg/TCC_Previsao_Vazao_LSTM_TFT

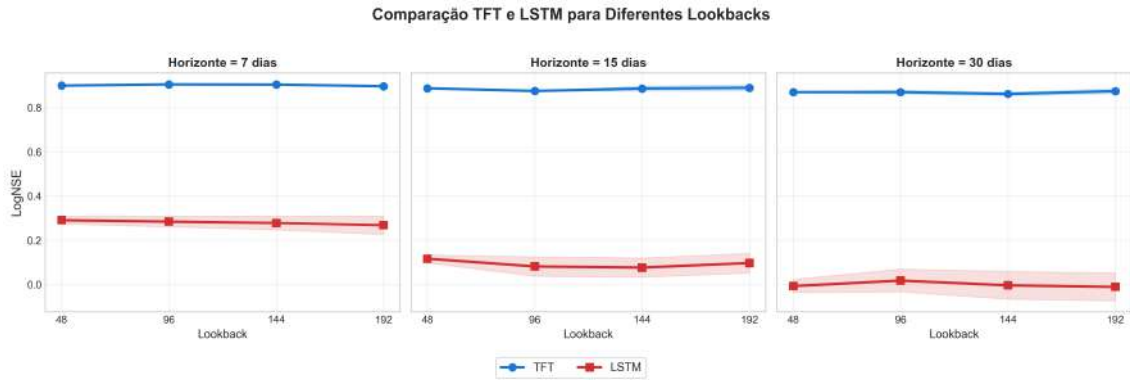


Figura 3. Desempenho comparativo (LogNSE) entre as arquiteturas TFT e LSTM ao longo dos horizontes de previsão.

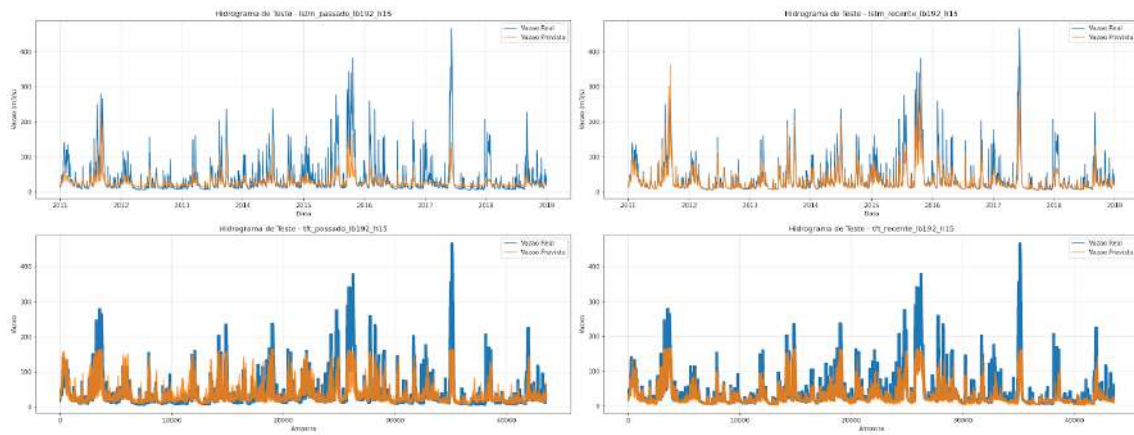


Figura 4. Hidrogramas comparativos das previsões dos modelos LSTM e TFT sob os regimes Passado (esquerda) e Recente (direita). A visualização permite observar a estabilidade das arquiteturas frente às variações hidrológicas do conjunto de teste.

6.1. Análise Visual dos Hidrogramas

A análise dos hidrogramas apresentados na Figura 4 complementa as métricas de erro discutidas anteriormente. É possível notar que, no Regime Recente (coluna da direita), o LSTM apresenta uma maior amplitude de erro nos picos de cheia, enquanto o TFT (linha inferior) mantém uma aderência mais consistente tanto em vazões médias quanto em períodos de estiagem. Esta estabilidade visual corrobora a superioridade do TFT identificada pelo $LogNSE$, confirmando que a arquitetura baseada em atenção é menos suscetível às rupturas hidrológicas presentes na série histórica.

6.2. Avaliação Comparativa: Regimes Passado e Recente

As Tabelas 2 e 3 sintetizam o desempenho preditivo das arquiteturas, com ênfase no $LogNSE$ como métrica principal para a avaliação em regimes de baixa vazão. Observa-se que o TFT apresenta superioridade clara nos horizontes subsazonais (7, 15 e 30 dias), mantendo índices de $LogNSE$ acima de 0,86 e um controle mais eficaz do $RMSE_{dry}$.

No *Regime Passado*, o LSTM apresentou dificuldades de convergência em horizontes mais longos, refletidas em métricas negativas. Já no *Regime Recente*, embora ambos os modelos apresentem uma degradação de performance consistente com a hipótese de não-estacionariedade [Milly et al. 2008], o TFT demonstrou maior estabilidade preditiva. Essa disparidade reforça a capacidade do TFT em adaptar-se a mudanças no comportamento hidrológico da bacia, validando sua adequação para o planejamento hídrico frente às incertezas climáticas atuais.

Tabela 2. Desempenho agregado (média $\pm$ desvio padrão) — Regime Passado.

Modelo	Horiz.	NSE	LogNSE	RMSE	RMSE _{dry}
TFT	7	0.855 ± 0.01	0.902 ± 0.01	16.48 ± 0.51	0.81 ± 0.21
	15	0.845 ± 0.01	0.893 ± 0.01	17.12 ± 0.42	1.09 ± 0.19
	30	0.832 ± 0.01	0.876 ± 0.01	17.96 ± 0.24	1.38 ± 0.41
LSTM	7	-0.026 ± 0.02	0.251 ± 0.02	44.09 ± 0.62	5.01 ± 0.46
	15	-0.191 ± 0.03	0.054 ± 0.03	47.54 ± 0.69	6.86 ± 0.86
	30	-0.227 ± 0.01	-0.053 ± 0.02	48.28 ± 0.33	11.33 ± 1.50

Tabela 3. Desempenho agregado (média $\pm$ desvio padrão) — Regime Recente.

Modelo	Horiz.	NSE	LogNSE	RMSE	RMSE _{dry}
TFT	7	0.845 ± 0.01	0.901 ± 0.00	13.41 ± 0.20	1.38 ± 0.20
	15	0.817 ± 0.01	0.877 ± 0.00	14.61 ± 0.28	1.95 ± 0.30
	30	0.813 ± 0.00	0.863 ± 0.01	14.87 ± 0.10	2.45 ± 0.20
LSTM	7	0.102 ± 0.00	0.309 ± 0.00	32.19 ± 0.06	5.55 ± 0.15
	15	-0.058 ± 0.01	0.130 ± 0.01	34.95 ± 0.17	8.25 ± 0.17
	30	-0.146 ± 0.01	0.051 ± 0.02	36.37 ± 0.12	9.33 ± 0.63

6.3. Robustez em Regimes de Baixa Vazão ($RMSE_{dry}$)

A avaliação da performance preditiva em regimes de baixa vazão é um componente crítico para a segurança hídrica, uma vez que modelos de *Deep Learning* tendem, frequentemente, a apresentar viés em direção à média, subestimando eventos extremos de estiagem. Para quantificar essa capacidade, utilizamos o *LogNSE* — que confere maior peso estatístico aos erros em baixas vazões — e o $RMSE_{dry}$, calculado especificamente para amostras abaixo do percentil histórico Q_{90} .

Ao analisar os resultados obtidos (Tabelas 2 e 3), observa-se que o TFT mantém valores de *LogNSE* consistentemente superiores aos do LSTM em ambos os regimes hidrológicos. Enquanto o LSTM exhibe uma queda acentuada na performance conforme o horizonte de previsão aumenta — chegando a apresentar valores de *LogNSE* próximos a zero ou negativos no Regime Recente — o TFT sustenta índices acima de 0.86, demonstrando maior estabilidade na representação da curva de recessão do escoamento.

O comportamento do $RMSE_{dry}$ reforça essa percepção: embora ambos os modelos apresentem um aumento no erro em regimes de seca quando submetidos ao Regime Recente, o TFT apresenta uma taxa de degradação mais controlada. Esta superioridade

estatística é detalhada na Figura 5, que ilustra a distribuição dos valores de *LogNSE* para cada regime, confirmando a resiliência do TFT na representação de eventos de estiagem. Este robustez do TFT frente ao LSTM pode ser atribuída à sua arquitetura de *Variable Selection Networks*, que permite ao modelo identificar e priorizar, dinamicamente, as covariáveis climáticas mais relevantes para a manutenção do fluxo de base, atenuando a propagação de incertezas que afeta arquiteturas recorrentes puras. Consequentemente, o TFT revela-se como a ferramenta mais apta para a gestão de recursos hídricos em cenários de estiagem prolongada, onde a precisão na predição de baixas vazões é imperativa.

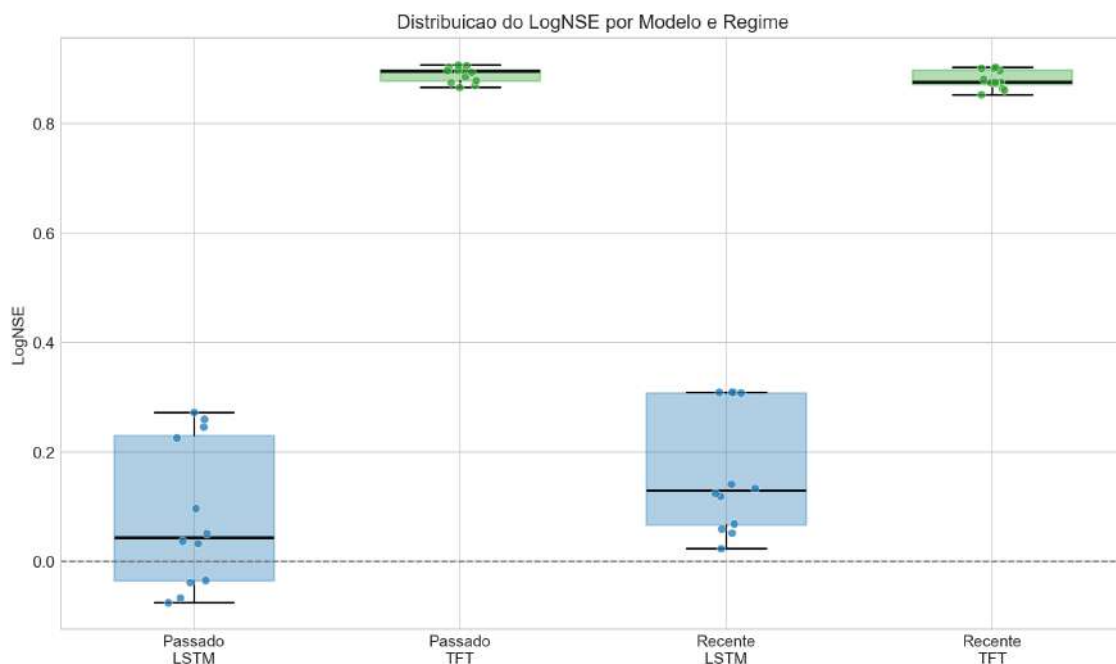


Figura 5. Distribuição do LogNSE para os modelos LSTM e TFT nos regimes Passado e Recente.

6.4. Interpretabilidade e Relevância de Variáveis

O TFT destaca-se pela sua interpretabilidade intrínseca, permitindo abrir a "caixa-preta" e validar quais informações o modelo prioriza para compor suas previsões. A Figura 6 ilustra a importância das variáveis de entrada para diferentes horizontes de previsão ($H = 1, 7, 15, 30$ dias). Nota-se que as variáveis meteorológicas clássicas, como precipitação e evapotranspiração, detêm o maior peso, o que é hidrológicamente coerente dada a resposta rápida da bacia.

Embora o Índice Oceânico Niño (ONI) apresente uma relevância numérica inferior em comparação às variáveis meteorológicas de alta frequência, sua inclusão é estratégica para o objetivo deste trabalho. Enquanto a precipitação explica a dinâmica imediata do escoamento, o ONI atua como um modulador de baixa frequência, permitindo ao TFT capturar a variabilidade climática de longo prazo que impulsiona rupturas hidrológicas. A presença constante dessa variável, mesmo que com pesos reduzidos, sinaliza que o modelo integra o contexto de teleconexões climáticas para ajustar suas previsões à não-estacionariedade do sistema. Essa capacidade de ponderação dinâmica permite que o TFT

priorize os drivers físicos mais relevantes para cada horizonte temporal, validando que a arquitetura aprendeu relações hidrológicamente fundamentadas em vez de correlações espúrias.

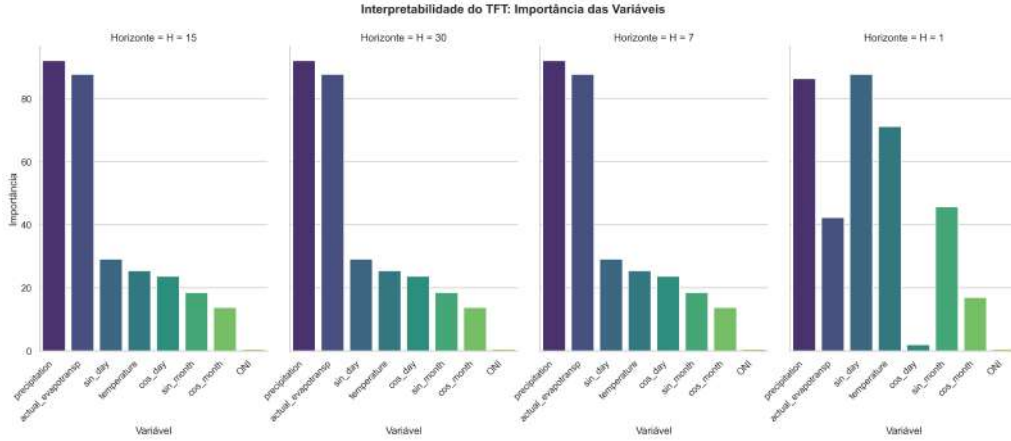


Figura 6. Importância das variáveis (Feature Importance) identificada pelo TFT no horizonte de previsão.

6.5. Análise de Nowcasting ($t + 1$)

A previsão imediata, ou *nowcasting* ($t + 1$), representa o cenário de menor complexidade temporal, onde a memória recente da série hidrológica atua como fator determinante. As Tabelas 4 e 5 sintetizam o desempenho dos modelos variando a janela de *lookback* de 30 a 180 dias.

Nota-se que, para este horizonte, ambas as arquiteturas atingem elevados índices de acurácia global. Contudo, ao analisar a estabilidade através do *LogNSE*, o TFT revela uma superioridade marginal e mais consistente. Enquanto o LSTM exhibe maior volatilidade nas métricas de erro — particularmente ao expandir a janela de observação — o TFT mantém uma estabilidade notável, com valores de *LogNSE* praticamente invariantes mesmo ao elevar a janela de entrada de 30 para 180 dias. Esta resiliência do TFT sugere que seu mecanismo de atenção atua como um filtro eficiente, isolando o sinal hidrológico relevante do ruído contido em janelas de observação estendidas.

O desempenho do $RMSE_{dry}$ reforça esta conclusão: o TFT apresenta erros mais controlados em situações de baixa vazão, mantendo a confiabilidade da predição em escalas de curtíssimo prazo. Em contraste, as flutuações observadas no LSTM, especialmente no *Regime Recente*, indicam uma maior vulnerabilidade a mudanças nas propriedades estatísticas da série, o que pode comprometer a confiança em aplicações críticas de tempo real.

Conforme ilustrado na Figura 7, a estabilidade do TFT nas métricas de desempenho sob diferentes janelas temporais reafirma sua robustez. Em suma, para aplicações onde a precisão imediata é vital, o TFT revela-se uma escolha arquitetônica mais resiliente, capaz de integrar o histórico recente da bacia com maior eficácia do que as abordagens recorrentes tradicionais, garantindo uma resposta consistente mesmo sob condições de incerteza climática.

Tabela 4. Desempenho comparativo para o *nowcasting* ($t + 1$) — Regime Passado.

Modelo	LB	NSE	LogNSE	KGE	MAE	RMSE	$RMSE_{dry}$
TFT	30	0.939	0.957	0.963	5.343	10.635	0.533
LSTM	30	0.919	0.953	0.901	6.076	12.327	0.874
TFT	60	0.943	0.958	0.956	5.427	10.289	1.286
LSTM	60	0.924	0.951	0.947	5.947	11.951	0.472
TFT	90	0.943	0.957	0.938	5.286	10.318	0.612
LSTM	90	0.925	0.950	0.920	6.316	11.885	0.529
TFT	120	0.942	0.957	0.925	5.260	10.393	0.802
LSTM	120	0.927	0.945	0.950	6.119	11.730	0.421
TFT	180	0.944	0.958	0.947	5.124	10.151	0.615
LSTM	180	0.918	0.951	0.902	6.374	12.529	0.498

Tabela 5. Desempenho comparativo para o *nowcasting* ($t + 1$) — Regime Recente.

Modelo	LB	NSE	LogNSE	KGE	MAE	RMSE	$RMSE_{dry}$
TFT	30	0.933	0.960	0.964	4.900	8.809	0.950
LSTM	30	0.940	0.958	0.956	4.788	8.349	0.596
TFT	60	0.937	0.962	0.965	4.776	8.546	1.045
LSTM	60	0.927	0.955	0.886	4.981	9.169	0.778
TFT	90	0.937	0.962	0.965	4.776	8.546	1.045
LSTM	90	0.939	0.957	0.939	4.714	8.381	0.552
TFT	120	0.936	0.961	0.960	4.770	8.571	1.006
LSTM	120	0.924	0.952	0.881	4.728	9.356	0.452
TFT	180	0.933	0.960	0.919	4.755	8.822	0.963
LSTM	180	0.937	0.959	0.967	4.745	8.546	0.614

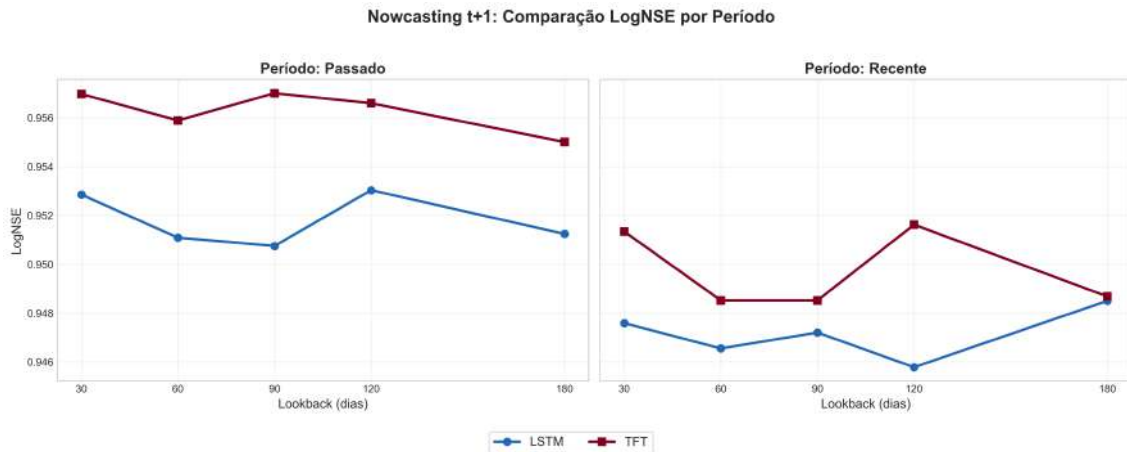


Figura 7. Comparativo de performance (LogNSE) entre TFT e LSTM no *nowcasting* ($t + 1$) para diferentes janelas de *lookback*.

6.6. Discussão: A Superioridade do TFT sob a Lente da Hidrológica

A análise dos resultados revela que a escolha entre as arquiteturas TFT e LSTM não deve ser guiada por uma superioridade absoluta, mas sim pela natureza do horizonte de previsão e pela criticidade do evento hidrológico em questão.

No *nowcasting* ($t + 1$), o LSTM demonstrou um desempenho altamente competitivo, alcançando métricas de acurácia (NSE e RMSE) equivalentes ou superiores às do TFT em diversas configurações do *Regime Recente*. Este comportamento confirma a robustez da arquitetura recorrente em cenários de curtíssimo prazo, onde a inércia da vazão e a correlação sequencial imediata são os fatores determinantes, tornando-a uma alternativa eficiente para operações de tempo real.

Contudo, a superioridade do TFT torna-se evidente nos horizontes subsazonais (7, 15 e 30 dias), onde a acurácia global (NSE > 0,81) é acompanhada por uma estabilidade superior em condições de seca, conforme indicado pelo **LogNSE** e $RMSE_{dry}$. Esta transição na performance reforça que o mecanismo de atenção e a seleção dinâmica de variáveis conferem ao TFT uma capacidade de generalização mais robusta frente às oscilações da série.

A arquitetura TFT mantém sua resiliência baseada em três pilares:

1. **Seleção Dinâmica de Variáveis:** O TFT prioriza variáveis climáticas (como o ONI) de forma dinâmica, adaptando-se às forçantes que realmente importam em cada regime.
2. **Mecanismo de Atenção Multi-Horizonte:** A atenção permite identificar padrões passados relevantes que o LSTM tende a atenuar em horizontes estendidos.
3. **Resiliência à Não-Estacionariedade:** O TFT integra covariáveis físicas globais, tornando-o menos suscetível às oscilações que degradam o desempenho recorrente puro em regimes recentes.

Em suma, enquanto o LSTM é uma ferramenta valiosa para previsões operacionais de curtíssimo prazo ($t + 1$), o TFT revela-se como uma solução superior para o planejamento

hídrico em um contexto de mudanças climáticas, onde a transparência e a capacidade de aprender com forçantes físicas globais são tão cruciais quanto a precisão imediata.

A superioridade do TFT em capturar as dinâmicas de eventos de baixa vazão está em consonância com os achados de [Amanambu et al. 2022], que também identificaram nos mecanismos de atenção uma maior capacidade de isolar sinais hidrológicos relevantes em comparação às redes recorrentes tradicionais."

6.7. Síntese dos Resultados

Em suma, a investigação conduzida confirma que a transição de arquiteturas recorrentes puras para modelos baseados em atenção (TFT) representa um avanço estratégico para a modelagem hidrológica na Bacia do Rio Uruguai. Os resultados demonstram que, enquanto o LSTM mantém eficácia em cenários de curto prazo, o TFT supera as limitações de memória e interpretabilidade que historicamente restringiram a aplicação de modelos de *Deep Learning* no setor hídrico. A robustez do TFT perante o *Regime Recente* reafirma a necessidade de modelos que integrem fatores climáticas globais, provendo uma ferramenta não apenas precisa, mas também transparente e adaptável à não-estacionariedade do clima global.

7. Conclusão

Este trabalho demonstrou que a arquitetura *Temporal Fusion Transformer* (TFT) apresenta uma resiliência superior para o planejamento hídrico e a detecção de eventos de seca em bacias sujeitas à não-estacionariedade. A análise comparativa revelou uma especialização funcional das arquiteturas: enquanto o LSTM demonstrou um desempenho altamente competitivo no horizonte de curtíssimo prazo (*nowcasting*, $t + 1$) — apresentando métricas de acurácia e $RMSE_{dry}$ equivalentes ou superiores em condições de baixa vazão — a arquitetura TFT consolidou-se como a solução mais robusta nos horizontes subsazonais (7, 15 e 30 dias).

A superioridade do TFT nos horizontes de planejamento não é meramente estatística, mas fundamentada na sua capacidade de integrar forçantes climáticas globais de forma interpretável. O desenho experimental de teste de estresse temporal evidenciou que a arquitetura recorrente tradicional, embora eficaz na inércia imediata da série, é mais suscetível às rupturas hidrológicas ao longo do tempo. Em contrapartida, o TFT, através de suas *Variable Selection Networks* e mecanismos de atenção multi-horizonte, provou ser capaz de isolar sinais climáticos relevantes — como o Índice Oceânico Niño (ONI) — para ajustar suas previsões à nova realidade climática da bacia, superando o paradigma de "caixa-preta" associado a modelos recorrentes puras.

Portanto, este estudo não propõe a substituição das arquiteturas tradicionais, mas sim a adoção estratégica de modelos baseados em atenção para a gestão de recursos hídricos. Enquanto o LSTM permanece como uma ferramenta de alta eficiência para operações de tempo real, o TFT revela-se a ferramenta mais apta para o planejamento hídrico de médio prazo. Ao preencher a lacuna entre a detecção do problema físico da não-estacionariedade e a validação de ferramentas computacionais avançadas, este arcabouço técnico contribui diretamente para a tomada de decisão em cenários de estiagem prolongada, oferecendo uma metodologia transparente e resiliente para a segurança hídrica na Bacia do Rio Uruguai e além.

Referências

- [Amanambu et al. 2022] Amanambu, A. C., Mossa, J., and Chen, Y. C. E. (2022). Hydrological drought forecasting using a Deep Transformer Model. *Water*, 14(17):2623.
- [Chagas et al. 2022] Chagas, V. B. P., Chaffe, P. L. B., and Blöschl, G. (2022). Climate and land management accelerate the Brazilian water cycle. *Nature Communications*, 13(1).
- [Chagas et al. 2020] Chagas, V. B. P., Chaffe, P. L. B., Fleischmann, A. S., Paiva, R. C. D., and Collischonn, W. (2020). CAMELS-BR: a large-sample dataset for hydrology and machine learning in Brazil. *Earth System Science Data*, 12(3).
- [Doninelli et al. 2017] Doninelli, T. L. et al. (2017). Previsão de vazões médias diárias com uso de redes neurais artificiais LSTM na bacia do Rio Apuaê-Inhandava. In *XXII Simpósio Brasileiro de Recursos Hídricos*, Florianópolis, SC. ABRH.
- [Fayer et al. 2023] Fayer, S. et al. (2023). Temporal Fusion Transformer for reservoir inflow forecasting in the Brazilian hydropower system. *Energy and AI*, 12.
- [Grimm and Silva 2013] Grimm, A. M. and Silva, D. A. (2013). The impact of el niño-southern oscillation on the hydrological cycle in south america. *Journal of Climate*, 26(15):5328–5347.
- [Gupta et al. 2009] Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F. (2009). Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377:80–91.
- [Hochreiter and Schmidhuber 1997] Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8):1735–1780.
- [Koya and Roy 2024] Koya, S. R. and Roy, T. (2024). Temporal Fusion Transformers for streamflow prediction: Value of combining attention with recurrence. *Journal of Hydrology*, 637:131301.
- [Lim et al. 2021] Lim, B., Arik, S. O., Loeff, N., and Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4):1748–1764.
- [Marcuzzo 2017] Marcuzzo, F. F. (2017). Bacia Hidrográfica do Rio Uruguai: Altimetria e áreas. In *XXII Simpósio Brasileiro de Recursos Hídricos*, Florianópolis, SC. ABRH.
- [Milly et al. 2008] Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R. J. (2008). Stationarity is dead: Whither water management? *Science*, 319(5863):573–574.
- [Nash and Sutcliffe 1970] Nash, J. E. and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part i — a discussion of principles. *Journal of Hydrology*, 10:282–290.
- [NOAA Climate Prediction Center 2025] NOAA Climate Prediction Center (2025). Oceanic Niño index (ONI). https://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php. Acessado em 29 de setembro de 2025.
- [Reichstein et al. 2021] Reichstein, M. et al. (2021). Early warning of complex climate risk with integrated artificial intelligence. *Nature*, 599(7883):39–41.

- [Rosa et al. 2022] Rosa, I. M. D. et al. (2022). Widespread streamflow decline in Brazil linked to deforestation. *Nature Communications*, 13(1).
- [Vaswani et al. 2017] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems*, pages 5998–6008.