UFPA

PPGEC

Universidade Federal do Pará

Leonardo Melo de Mendonça

REDES NEURAIS RECORRENTES PARA MODELAGEM CHUVA-VAZÃO DE PEQUENAS BACIAS HIDROGRÁFICAS DA AMAZÔNIA

DISSERTAÇÃO DE MESTRADO

Instituto de Tecnologia

Programa de Pós-Graduação em Engenharia Civil

Orientador Prof. Claudio José Cavalcante Blanco, PhD.

Belém – Pará – Brasil

2022

Leonardo Melo de Mendonça

REDES NEURAIS RECORRENTES PARA MODELAGEM CHUVA-VAZÃO DE PEQUENAS BACIAS HIDROGRÁFICAS AMAZÔNICAS

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Civil da Universidade Federal do Pará, como requisito para obtenção do Título de Mestre em Engenharia Civil.

Área de concentração: Engenharia Hídrica.

Orientador Professor Claudio José Cavalcante Blanco, PhD.

Belém – Pará – Brasil
2022

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)

M539r Mendonça, Leonardo Melo de.
Redes Neurais Recorrentes para modelagem chuva-vazão de pequenas bacias hidrográficas da Amazônia / Leonardo Melo de Mendonça. — 2022.
109 f. : il. color.

Orientador(a): Prof. Dr. Claudio José Cavalcante Blanco
Dissertação (Mestrado) - Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-Graduação em Engenharia Civil, Belém, 2022.

1. Aprendizado de máquina. 2. Modelo chuva-vazão. 3. Modelagem hidrológica. 4. Pequenas bacias hidrográficas da Amazônia. I. Título.

CDD 333.91

SERVIÇO PÚBLICO FEDERAL
UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

REDES NEURAIS RECORRENTES PARA MODELAGEM CHUVA-VAZÃO DE PEQUENAS BACIAS HIDROGRÁFICAS DA AMAZÔNIA.

AUTOR:

LEONARDO MELO DE MENDONÇA

DISSERTAÇÃO SUBMETIDA À BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL DO INSTITUTO DE TECNOLOGIA DA UNIVERSIDADE FEDERAL DO PARÁ, COMO REQUISITO PARA OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA CIVIL NA ÁREA DE RECURSOS HÍDRICOS E SANEAMENTO AMBIENTAL.

APROVADO EM: 29 / 04 / 2022.

BANCA EXAMINADORA:

Prof. Claudio José Cavalcante Blanco, Ph.D.
Orientador (UFPA)

Prof. Dr. Frede de Oliveira Carvalho
Membro Externo (UFAL)

Profa. Dra. Evanice Pinheiro Gomes
Membro Externo (UFPA)

Prof. Dr. Nelio Moura de Figueiredo
Membro Interno (UFPA)

Visto:

Prof. Dr. Marcelo de Souza Picanço
Coordenador do PPGEC / ITEC / UFPA

AGRADECIMENTOS

A Deus e Santo Antônio de Pádua pela dádiva da vida e por me permitir realizar tantos sonhos. Obrigado pela oportunidade de errar, aprender e crescer. Obrigado por tanto quanto mereço.

A minha Mãe Socorro, meu Pai Carlindo, e minha Irmã Camilla. Toda construção pessoal e profissional que tenho é proveniente de vocês. Pretendo honrar seus valores até meu último suspiro de vida. Amo vocês!

A toda minha Família, que sempre apoia minhas decisões. Em especial, agradeço aos meus avós, Benedito e Maria, pelas orações e torcida. Também, agradeço imensamente aos meus padrinhos, Leila e Aldivan, por serem pessoas tão presentes. Agradeço muito ter nascido entre vocês.

Ao meu orientador, Prof. Claudio Blanco. Muito obrigado pela oportunidade, confiança e profissionalismo. Tenho muito respeito e apreço pela sua pessoa. És um grande exemplo!

Aos docentes do PPGEC, em especial à Prof. Mariane, Prof. Lindemberg, e Prof. Francisco, que se dedicaram em transmitir conhecimento nos anos que se passaram, mesmo no cenário da pandemia de COVID-19.

Aos meus amigos Deyverson e Natalia, que em muitos momentos dispuseram tempo e atenção em meu favor. Aos meus amigos Igor e Jaime, que a UFRA me proporcionou, por me incentivarem no caminho acadêmico. Obrigado pela parceria.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de mestrado.

A todos aqueles que, mesmo discretamente, estão envolvidos na construção deste Título.

“O que eu penso não muda nada além do meu pensamento. O que eu faço a partir disso muda tudo” (Leandro Karnal)

RESUMO

Modelos chuva-vazão podem auxiliar o gerenciamento de recursos hídricos, principalmente na Amazônia, região marcada pela baixa densidade de monitoramento hidrológico, e assim beneficiar os usos múltiplos da água e o adequado aproveitamento dos recursos hídricos. Este trabalho busca simular vazões diárias de cinco pequenas bacias hidrográficas da Amazônia, através da Rede Neural Recorrente Não linear Autorregressiva com Variável Exógena (RNN-NARX). Dados pluviométricos e fluviométricos diários foram utilizados para simulação. As funções de correlação cruzada e autocorrelação parcial auxiliaram a determinação de dados defasados, entradas relevantes, com nível de significância de 5%. Além disso, foi utilizado o algoritmo de retropropagação do erro Levenberg-Marquardt para treinamento supervisionado das RNN-NARX. Também foram utilizados cinco índices estatísticos e a contribuição relativa de Garson de cada variável de entrada para avaliação das simulações. Deste modo, as vazões simuladas foram classificadas entre o insatisfatório e o muito bom, além de apresentar tendência, geral, a subestimar vazões de cheia. A característica autorregressiva de cada bacia é fundamental para melhores resultados, qualidade atribuída à capacidade de armazenamento de água. Uma explicação plausível para as principais fontes de incerteza é devida a variabilidade espacial da precipitação entre as estações de monitoramento e as precipitações ocorridas na bacia, anomalias meteorológicas e aspectos de discretização. A análise de sensibilidade dos modelos frente a diferentes intervalos de treinamento mostrou que a implementação de 2 anos, para o treinamento supervisionado das RNN-NARX, são suficientes para se obterem simulações eficientes em quatro das cinco pequenas bacias hidrográficas da Amazônia analisadas.

Palavras-chave: Aprendizado de máquina, modelo chuva-vazão, modelagem hidrológica, pequenas bacias hidrográficas da Amazônia.

ABSTRACT

Rainfall-runoff models can help the management of water resources, especially in the Amazon, a region marked by the low density of hydrological monitoring, and thus benefit the multiple uses of water and the adequate use of water resources. This work seeks to simulate daily streamflows of five small catchments in the Amazon, through the Autoregressive Recurrent Nonlinear Neural Network with Exogenous Variable (RNN-NARX). Daily rainfall and streamflow data were used for simulation. The cross-correlation and partial auto-correlation functions helped to determine lagged data, relevant inputs, with a significance level of 5%. In addition, the Levenberg-Marquardt error backpropagation algorithm was used for supervised training of RNN-NARX. Five statistical indices and Garson's relative contribution of each input variable were also used to evaluate the simulations. Thus, the simulated flows were classified between unsatisfactory and very good, in addition to showing a general tendency to underestimate floods. The autoregressive characteristic of each catchment is fundamental for better results, quality attributed to the water storage capacity. A plausible explanation for the main sources of uncertainty is due to the spatial variability of precipitation between monitoring stations and the precipitations occurring in the catchment, meteorological anomalies and discretization aspects. The sensitivity analysis of the models against different training intervals showed that the implementation of 2 years, for the supervised training of the RNN-NARX, is sufficient to obtain efficient simulations in four of the five small Amazon catchment analyzed.

Keywords: Machine learning, rainfall-runoff model, hydrological modeling, small Amazon catchments.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE SIGLAS E ABREVIATURAS

ANA – Agência Nacional de Águas

CR – Contribuição relativa de Garson

EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária

ENOS – El Niño Oscilação Sul

IBGE – Instituto Brasileiro de Geografia e Estatística

INPE – Instituto Nacional de Pesquisas Espaciais

MAPE – Erro absoluto percentual médio

NARX – Não-linear Autorregressiva com Variável Exógena

NSE – Coeficiente de eficiência de Nash-Sutcliffe

P – Precipitação observada

PBIAS – Viés percentual

Q – Vazão observada

Qe – Vazão estimada

RMSE – Raiz do erro quadrático médio

RNAs – Redes Neurais Artificiais

RNNs – Redes Neurais Recorrentes

RSR – Raiz do erro quadrático médio padronizado

USEPA – United States Environmental Protection Agency

USGS – United States Geological Survey

α – Nível de significância

SUMÁRIO

1 INTRODUÇÃO

A ausência parcial ou total de séries temporais de dados hidrológicos pode comprometer o adequado gerenciamento e aproveitamento dos recursos hídricos, principalmente em uma época marcada pelo aumento do uso dos recursos naturais e da degradação ambiental. Por isso, a concepção de novas ferramentas de simulação de variáveis hidrológicas, cada vez mais exatas, é de fundamental importância para auxiliar o monitoramento, e promover o desenvolvimento sustentável (BEHRANGI et al., 2011).

Ao longo da história, houve várias metodologias para simulação de vazões, como o método racional (MULVANEY, 1851), o hidrograma unitário (SHERMAN, 1932), além do desenvolvimento de modelos conceituais (LINSLEY e CRAWFORD, 1960), distribuídos (HUGGINS e MONKE, 1968), físicos (BEVEN et al., 1980) e modelos baseados em inteligência artificial (DANIELL, 1991), sendo favorecidos a partir do desenvolvimento computacional ocorrido na década de 50 (PEEL e MCMAHON, 2020). Segundo Peel e McMahon (2020), entre os anos de 1960 e 2010 houve a intensificação da busca por otimizações dos modelos de simulação, que apresentassem resultados satisfatórios e demandassem poucos parâmetros implementados. No entanto, ainda é controversa a abordagem ideal, frente à heterogeneidade das bacias hidrográficas, das características meteorológicas em diferentes locais, em contraste com a disponibilidade de dados (BEVEN et al., 1988; COULTHARD et al., 2005; BEVEN, 2012).

A simulação de vazões é complexa devido à alta variabilidade espacial e temporal em que diversos fatores podem desenvolver-se ao longo de uma bacia hidrográfica, conforme a concepção do balanço hídrico (BEVEN et al., 1988; BEVEN, 2012). As variabilidades climáticas, hidráulica dos solos e de seus respectivos usos, e das características fisiográficas das bacias são exemplos de variáveis que causam, muitas vezes, um comportamento não linear nas vazões afluentes de um reservatório, ou bacia hidrográfica (RIBEIRO NETO et al., 2008; QUINTAS et al., 2011; BLANCO et al., 2013; MIRUS E LOAGUE, 2013).

O modelo de Redes Neurais Recorrentes (RNN) denominado de Autorregressivo Não Linear com Entradas Exógenas (NARX) tem mostrado grande potencial de modelar sistemas entrada-saída com características não lineares (MENEZES JR. e BARRETO, 2008). Desse modo, tal técnica vem sendo aplicada na simulação e previsão de variáveis hidrológicas (MAIER et al. 2010; SHEN e CHANG,

2013; LEE e RESDI, 2016; WUNSCH et al., 2018; GUZMAN et al., 2019; BANIHABIB et al., 2019; SIT et al., 2020; WUNSCH et al., 2021; MENDONÇA et al., 2021; WANG e CHEN, 2022), com bons resultados de simulação em regiões com baixa disponibilidade de dados (GUZMAN et al., 2019). Além disso, RNN-NARX convergem mais rapidamente e generalizam melhor do que outras técnicas de RNAs (LIN et al., 1998; IZADY et al., 2013).

A Amazônia apresenta um dos mais complexos ecossistemas, e a maior bacia hidrográfica do mundo (PAIVA et al., 2012). Porém, o monitoramento hidrometeorológico neste local apresenta distribuição espacial precária e baixa resolução temporal de dados observados, prejudicando a simulação de fenômenos chuva-vazão (COLLISCHONN et al. 2008; BLANCO et al. 2013). No contexto das pequenas bacias, este cenário é agravado, pois, o interesse econômico é menor (BLANCO et al. 2013), entretanto, isto não destitui a necessidade de promover o gerenciamento das mesmas. Assim, as simulações de vazões em pequenas bacias via RNN-NARX tem como propósito auxiliar o planejamento de atividades vinculadas ao uso da água e análise de impactos ambientais de obras hidráulicas, disponibilidade hídrica, etc.

Embasado nisto, este trabalho busca simular, primeiramente, hidrogramas, e consequentemente, curvas de permanência de vazões diárias de pequenas bacias hidrográficas da Amazônia através de modelos fundamentados em RNN-NARX. Foi realizado também a análise de sensibilidade quanto a dimensão das séries temporais utilizadas para treinamento dessas estruturas, buscando o melhor desempenho para o objetivo do estudo.

1.1 OBJETIVOS

1.1.1 Geral

Simular hidrogramas e curvas de permanência de vazões diárias médias em pequenas bacias hidrográficas amazônicas através do treinamento supervisionado de Redes Neurais Recorrentes (RNN), com a arquitetura Autorregressiva Não Linear com Entradas Exógenas (NARX).

1.1.2 Específicos

– Obter dados de vazões e chuvas diárias em bancos de dados nacionais;

– Realizar uma análise exploratória baseada em correlações lineares entre os dados obtidos e determinar vetores de entrada significantes;

– Treinar e testar modelos de simulação chuva-vazão, avaliando o desempenho estatístico entre as vazões observadas e simuladas;

– Verificar a sensibilidade do modelo frente a diferentes intervalos de tempo de dados de entrada.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 CICLO HIDROLÓGICO

A importância da água é indiscutível. O equilíbrio ecológico de ecossistemas, além do fundamental papel no desenvolvimento socioeconômico dos seres humanos, está ligado à disponibilidade de água, quer seja em quantidade, quer seja em qualidade. Porém, a água é considerada um bem não renovável, devido à desigual distribuição de água potável no globo terrestre, e as complexas interações hidroclimáticas do ciclo hidrológico (VON SPERLING, 2006). Logo, o conhecimento e o monitoramento dos recursos hídricos são importantíssimos para a manutenção do bem-estar social.

Os fluxos de água na porção continental do ciclo hidrológico apresentam-se como um sistema dinâmico. Conceitualmente, ao adotar a bacia hidrográfica como a área de drenagem de um rio e seus afluentes, é possível analisar os fluxos de entrada e saída de água, que atua simultaneamente. Os principais componentes deste sistema dinâmico são a precipitação, a evapotranspiração, a interceptação, a infiltração, e as vazões efluentes e afluentes (Figura 1) (WOOLHISER, 1973; TUCCI, CLARKE, 1997). Destaca-se, também, que estas variáveis são submetidas à variabilidade espacial e temporal, o que torna a estes fenômenos não lineares, e altamente complexos (BLANCO et al., 2013).

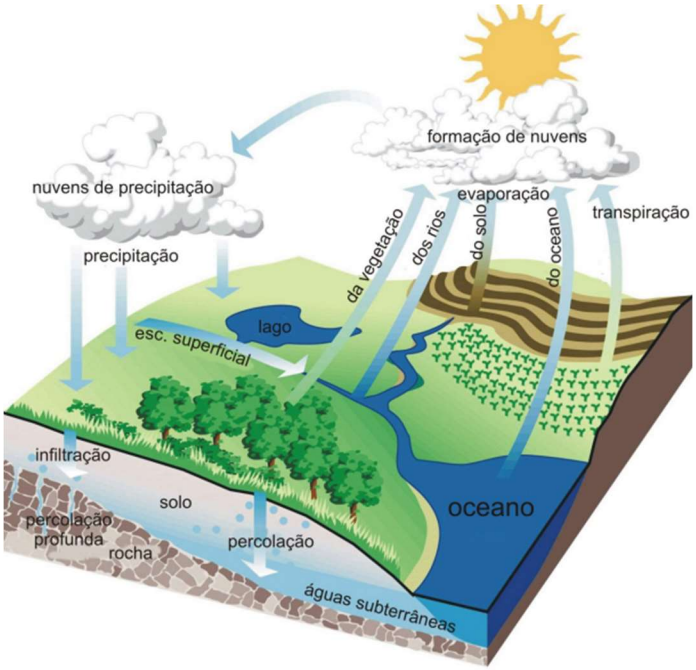

Figura 1: Concepção dos fluxos de água no ciclo hidrológico

Fonte: Adaptado de USEPA (1998)

Quando a intensidade da precipitação excede a taxa de infiltração, o escoamento começa imediatamente. Esse escoamento depende da capacidade de infiltração, da declividade e da porosidade do solo, que também apresenta resistência ao fluxo de água para as camadas mais profundas. Neste sentido, solos com altos teores de argila têm menor capacidade de infiltração e indiretamente produzem mais escoamento superficial, em comparação com os solos arenosos. Além disso, ações antrópicas como a alteração do uso e cobertura do solo podem favorecer o escoamento superficial (CHANDWANI et al., 2015).

As bacias hidrográficas de primeira ordem, que são normalmente de pequena escala, ocupam aproximadamente 80% da área de drenagem de um rio (SANTANA e BLANCO, 2020), por isso, é conveniente considera-las como maior fonte de geração de vazões afluentes (ROGERS, 1972). Em pequenas escalas espaciais e temporais, os fenômenos hidrológicos apresentam um comportamento mais heterogêneo. Enquanto que, em escalas espaciais e temporais maiores, a complexidade do processo geralmente se reduz através da média de complexidades de menor escala (WOOD et al. 1988; BLÖSCHL et al., 1995; GRAYSON, et al., 1992; SAVENIJE, 2001). Por exemplo, o balanço hidrológico de pequenas bacias pode ser considerado menos complexo, visto que o monitoramento de resposta aos eventos de precipitação é mais rápido e mais evidente (BUTTLE, 1998). Por outro lado, Hewlett e Hibbert (1967) ressaltam a relevância de processos subsuperficiais e subterrâneos, comparado aos processos superficiais, na produção das vazões afluentes de pequenos rios perenes de áreas úmidas. Enquanto que Estrany et al. (2010) comentam que pequenas bacias do mediterrâneo tem como principal variável influenciadora das vazões afluentes a evapotranspiração. Logo, fica evidente que a intensidade de diferentes componentes das bacias varia de local para local, principalmente em pequenas escalas.

2.2 MODELOS CHUVA-VAZÃO

Um modelo chuva-vazão é uma representação virtual do sistema dinâmico que divide a precipitação de chuva em escoamento superficial, subsuperficial, evapotranspiração, umidade do solo ou escoamento subterrâneo (PEEL e MCMAHON, 2020). Esta abordagem é fundamental devido às limitações de técnicas de medições, ocasionadas pela complexidade da natureza. Com a aplicação de modelos chuva-vazão, espera-se melhorar a tomada de decisões sobre um

problema hidrológico, seja no planejamento de recursos hídricos, cobertura de falhas históricas, proteção contra eventos extremos, mitigação de poluição hídrica, licenciamento de captações, ou outras aplicações (BEVEN, 2012).

Ao longo da história, houve cerca de 280 modelos chuva-vazão, além de pequenas modificações dos modelos oriundos destes modelos (PEEL e MCMAHON, 2020). Segundo Peel e McMahon (2020), a primeira descrição formal de um modelo chuva-vazão veio através do método racional de Mulvaney (1851), que calcula a vazão máxima de pequenas bacias hidrográficas a partir de uma intensidade máxima de precipitação e de um tempo de concentração. Outro modelo matemático relevante, que inspirou diversas outras modificações, é o hidrograma unitário de Sherman (1932), que utiliza a precipitação efetiva unitária, com intensidade constante no tempo e uniformemente distribuída no espaço e sobre a bacia hidrográfica.

Há uma evidente busca por um modelo que represente de maneira satisfatória os fenômenos chuva-vazão, mas que também o faça a partir do menor número de parâmetros. Segundo Clarck et al. (2011), a superabundância de modelos chuva-vazão mostra que há uma compreensão científica insuficiente da dinâmica ambiental do ciclo hidrológico, que pode ser atribuída às dificuldades em medir e representar a heterogeneidade encontrada nos sistemas naturais. Sitterson et al. (2018) falam que é preciso identificar as prioridades da modelagem e as limitações da disponibilidade de dados, tempo e orçamento para modelos ajudarem a restringir as escolhas e garantir que o modelo seja o melhor para o propósito pretendido. Os autores classificam os modelos chuva-vazão quanto aos tipos de dados abordados, sejam eles determinísticos ou estocásticos; quanto a estrutura do modelo, que pode ser empírico, conceitual ou físico; quanto à variabilidade espacial, podendo ser modelos agrupados, semi-distribuídos e distribuídos.

A estrutura do modelo define como a vazão afluente é calculada (SITTERSON et al., 2018). Para Melsen et al. (2016), a principal característica limitante de modelos chuva-vazão está na escolha da estrutura do modelo. As estruturas desses modelos mais simples são os empíricos, que mapeiam a relação entrada-saída, ou seja, fazem uma modelagem caixa-preta ao não utiliza qualquer relação física. Modelos conceituais utilizam equações simplificadas que representam o armazenamento de água na bacia hidrográfica, analisando parcialmente os fluxos de água através de compartimentos, logo, realizam uma modelagem caixa-cinza. E,

modelos físicos são aqueles que utilizam as leis e equações físicas baseadas em respostas hidrológicas reais, ou modelos caixa-branca. Logo, os modelos conceituais e físicos necessitam de uma compreensão completa da física dos fluxos de água envolvidos. Cada estrutura apresenta vantagens e desvantagens, a depender da finalidade do modelo e do nível de detalhamento que se deseja alcançar (BEVEN, 2012, SITTERSON et al., 2018, PEEL e MCMAHON, 2020).

Quando o modelo não utiliza alguma forma de representar a variabilidade espacial é chamado de agrupado, ou seja, considera a manifestação de fenômenos hidrológicos como uma unidade homogênea. Os semi-distribuídos refletem alguma variabilidade espacial. E, os modelos totalmente distribuídos processam variabilidade espacial por células de grade (SITTERSON et al., 2018). Estas diferentes formas de representar a variabilidade espacial das variáveis hidrológicas são mostradas na Figura 2. Os modelos semi-distribuídos levam em consideração a variabilidade espacial em escalas menores do que os modelos agrupados, mas não calculam o escoamento em cada célula da grade.

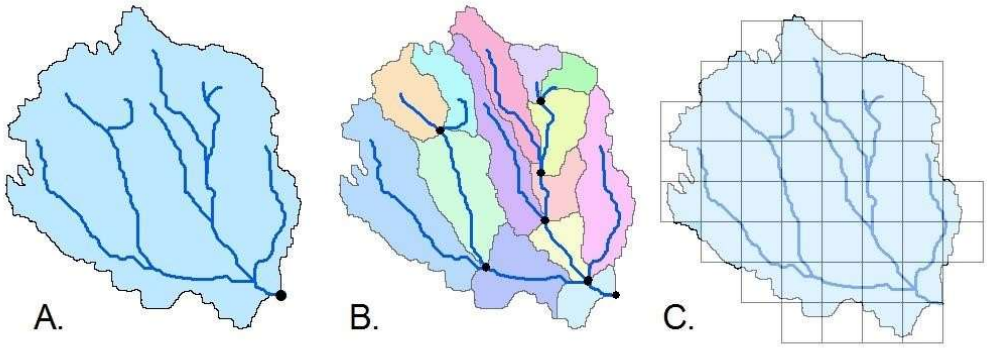

Figura 2: Abordagem espacial agrupada (a), semi-distribuída (b) e distribuída (c) de modelos hidrológicos

Fonte: Sitterson et al. (2018)

Por ser uma representação simplificada de um sistema dinâmico e complexo, todos os modelos chuva-vazão são incertos até certo ponto. A incerteza dos modelos pode vir dos dados observados, incertezas naturais, estimativa de parâmetros, calibração ou suposições do modelo, e nas mudanças de variáveis hidrometeorológicas (SITTERSON et al., 2018). Segundo Pechlivanidis et al. (2011), os dados de entrada são uma importante fonte de incerteza, pois a calibração dos parâmetros de qualquer modelo será baseado na qualidade do monitoramento. Neste sentido, o aspecto mais desafiador da validação de modelos com dados

observados está contido em saber até onde os modelos são confiáveis e o quão bem podem representar o sistema abordado. Além disso, a comparação entre diferentes modelos chuva-vazão pode ser contraditória e de difícil interpretação, pois podem diferenciar-se em muitos aspectos estruturais. Alguns modelos podem ser considerados satisfatórios em determinadas bacias, enquanto outros podem não compreender razoavelmente os fenômenos intrínsecos de outras bacias (FEKETE, 2002; ANDRÉASSIAN et al., 2004).

Diversos autores relatam a dificuldade em realizar simulações na Amazônia devido à falta de dados (BLANCO et al., 2007; COLLISCHONN et al., 2008; BLANCO et al., 2013; ISHIRARA et al., 2013). Segundo Blanco et al. (2007), a abordagem ideal frente às pequenas bacias são por dados horários, devido ao curto tempo de concentração. No entanto, os únicos dados disponíveis nesta escala espacial são as precipitações e vazões diárias. Apesar de serem modelos empíricos, modelos chuva-vazão são baseados em aprendizado de máquina (DEVIA et al., 2015) e trazem um conjunto de benefícios para a hidrologia, pois são adaptáveis à heterogeneidade de cada bacia, demandam poucos dados e são tolerantes a possíveis falhas. Logo, a fundamentação deste trabalho surge na tentativa de estimar hidrogramas e curvas de permanência de vazões diárias de pequenas bacias hidrográficas através desta abordagem.

2.2.1 Análise de sensibilidade

Na modelagem hidrológica, a análise de sensibilidade é definida como a investigação da variação do desempenho dos modelos, conforme a mudança nas variáveis de entrada ou nos parâmetros dos mesmos. Isto permite a determinação de diferentes fontes de incerteza para a variação da saída, utilizando abordagens qualitativas ou quantitativas sob um determinado conjunto de suposições e objetivos (SONG et al., 2015).

Os métodos de análise de sensibilidade são divididos, majoritariamente, em globais e locais. Segundo Song et al. (2014), os métodos globais têm certas vantagens em relação aos locais, pois têm a capacidade de incorporar a influência dos parâmetros de entrada sobre toda a gama de variação, e são bem adequados para modelos não lineares e não monotônicos. Dentre os principais métodos utilizados na modelagem hidrológica, é possível citar o método de triagem, análise de regressão, método baseado na variância e método de meta-modelagem (SONG

et al., 2014). O método de triagem é o mais comum, em que, basicamente, os parâmetros de entrada são variados individualmente, enquanto que os demais parâmetros são mantidos constantes.

Alguns dos parâmetros do modelo geralmente representam processos que só importam durante períodos de tempo específicos, ou seja, modos específicos do sistema, como constantes de recessão ou parâmetros que controlam a extensão das áreas saturadas durante um evento de inundação. Esses parâmetros, provavelmente, só serão identificáveis se esses períodos puderem ser isolados ou se tiverem impactos suficientes em uma função objetivo global. Em geral, os métodos de análise de sensibilidade usados para análise variante no tempo incluem abordagens locais e globais. Independentemente do método aplicado, eles geralmente são usados para estimar a sensibilidade em cada passo de tempo ou para uma janela de execução (MASSMAN et al., 2014).

2.3 REDES NEURAIS ARTIFICIAIS

As Redes Neurais Artificiais (RNAs) fazem parte do conjunto de técnicas de aprendizado de máquina, uma subcategoria da Inteligência Artificial. É, essencialmente, uma forma de estatística aplicada, com ênfase no uso de computadores para estimar funções não lineares. Deste modo, torna-se capaz de resolver problemas complexos de classificação ou regressão (HAYKIN, 2001).

Inspiradas no sistema nervoso central dos seres humanos, as RNAs foram idealizadas através da observação do funcionamento dos neurônios biológicos (Figura 3) (HAYKIN, 2001). Nestas células do cérebro humano, os impulsos nervosos são provenientes das árvores dendríticas e um único neurônio humano pode receber milhares destes sinais de entrada, podendo ser inibitórios ou excitatórios. A propagação dos impulsos nervosos através do axônio[1] ocorrerá somente quando esta membrana sofrer uma hiperpolarização, através da distribuição assimétrica de íons de sódio e potássio, conhecido como limiar de disparo. Alcançado este cenário, a propagação dos impulsos nervosos é efetivada, levando informações a outros neurônios (KOVÁCS, 2002).

[1] Axônio é um prolongamento da célula nervosa, que tem como função propagar um pulso elétrico a outras células.

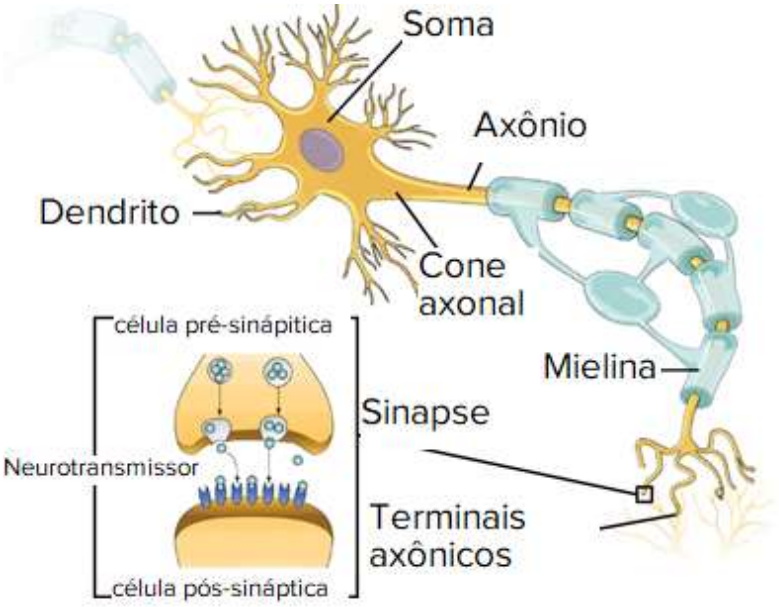

Figura 3: Neurônio biológico humano

Fonte: Adaptado de Clark te al. (2018)

Semelhante ao sistema nervoso humano, a principal unidade de processamento em RNAs são os neurônios artificiais (Figura 4 e Equação 1). Nessas estruturas, os dados de entrada x_n são direcionados ao neurônio artificial, para gerar uma resposta y_n. Cada x_n é ponderado por um peso sináptico w_{kn}, e posteriormente todos são somados juntamente com o parâmetro bias b_k, que se torna variável independente de uma função de ativação f(.) para gerar sua saída y_n (HAYKIN, 2001).

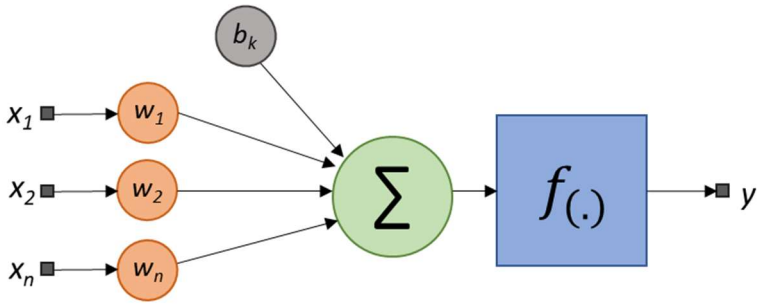

Figura 4: Neurônio artificial sob a forma de diagrama de blocos

$$y_k = f\left(b_k + \sum_{i=1}^{n} w_{ki} x_i\right) \quad (1)$$

O início das pesquisas científicas acerca da técnica de RNAs é frequentemente datado a partir do trabalho de McCulloch e Pitts (1943), que descreveram neurônios artificiais binários, para determinar o estado de ativação do neurônio. Através de provas formais, os autores mostraram que tais neurônios

poderiam implementar qualquer função lógica ou aritmética (HAYKIN, 2001; YADAV et al., 2015).

Hebb (1949) formulou a primeira regra formal de aprendizagem associativa de neurônios biológicos, que resultaria em uma modificação permanente do padrão de atividade de um grupo destas células nervosas espacialmente distribuídas. Assim, esta lei de aprendizagem foi utilizada para construir uma explicação qualitativa de alguns resultados experimentais da psicologia. Este foi o fundamento necessário para o desenvolvimento de regras de aprendizagens específicas das RNAs (YADAV et al., 2015).

O primeiro modelo computacionalmente prático foi proposto por Rosenblatt (1958), o Perceptron. É considerada a forma mais simples de configuração de uma RNAs, pois este tipo de rede é composto por apenas um neurônio artificial, sendo utilizada principalmente para classificação binária (YADAV et al., 2015).

Widrow e Hoff (1960) desenvolveram o modelo Adaline, que foi equipado com o aprendizado via Regra Delta, baseada no gradiente descendente da função do erro. Os autores aplicaram com sucesso a um grande número de problemas. No entanto, Minsky e Papert (1969) provaram que as redes que empregam funções de ativação linear (como o Perceptron e o Adaline da época) são incapazes de resolver problemas de classificação não linear (YADAV et al., 2015).

Com o desenvolvimento da Regra Delta Generalizada, hoje muito conhecida como Retropropagação, ou *Backpropagation* (RUMELHART et al., 1986), a aplicação de RNAs se consolidou, pois, esta técnica possibilitou a aproximação de funções não-lineares de arquiteturas *feedforward* (arquiteturas em que as informações se movem em uma única direção) (YADAV et al., 2015).

A clássica arquitetura RNA *feedforward*, chamada de Perceptron de Múltiplas Camadas, é composta por muitos Perceptrons simples, que apresenta uma ou mais camadas ocultas entre as camadas de entrada e saída (Figura 5). Uma das tarefas iniciais para usar tais é determinar qual topologia se empregar, ou seja, definir o número de camadas ocultas, e de neurônios em tais camadas. Na maioria das vezes, estes são definidos heuristicamente ou pelo método das aproximações sucessivas, o que pode resultar em estruturas hiperparametrizadas e não-parcimoniosas (ZHANG e MORRIS; 1998; STATHAKIS, 2009). No entanto, segundo Hornik et al. (1989), RNAs com uma simples camada escondida, provida de função

de ativação sigmóide logística, podem aproximar qualquer função contínua arbitrária, dando o número suficiente de neurônios.

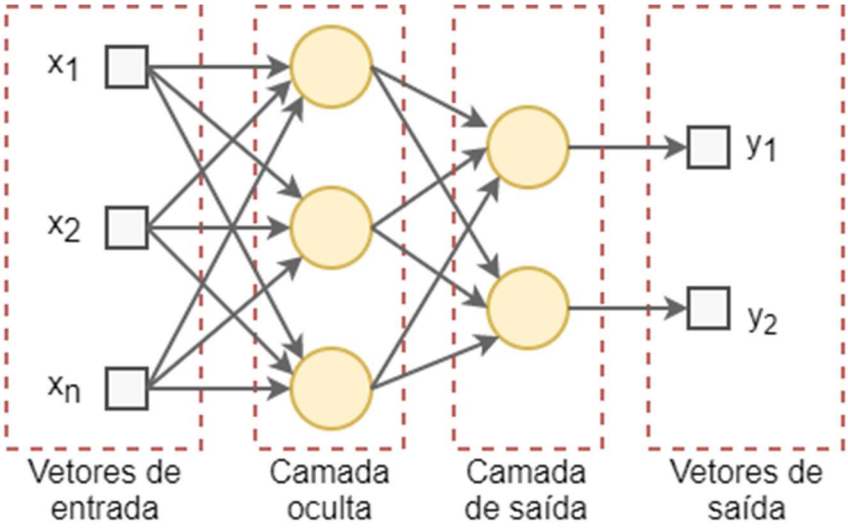

Figura 5: Arquitetura Multilayer Perceptron

2.3.1 Funções de ativação

As funções de ativação, também chamadas de funções transferência, são responsáveis por modular a resposta y_k do neurônio artificial, em função do produto entre as entradas e os parâmetros treináveis das RNAs. Nessas circunstâncias, a escolha das funções de ativação dos neurônios artificiais dependerá, essencialmente, do sistema abordado e a finalidade que se quer atribuir a cada neurônio. As principais funções de ativação são a degrau, linear, sigmóide e tangente hiperbólica (Figura 6) (YAZID et al., 2018).

A função de transferência degrau (*hard-limiter*) força um neurônio artificial a emitir um valor β se a variável independente atingir um limite θ, caso contrário, emite-se o valor δ (Equação 2) (Figura 6a). Isto permite que um neurônio tome uma decisão ou classificação, em vista da característica binária (BEALE et al., 2016).

$$Hardlim(x) = \begin{cases} \beta \; se \; x \geq \theta \\ \delta \; se \; x < \theta \end{cases} \quad (2)$$

A função de ativação linear (*purelin*) faz com que o neurônio artificial aplique uma transformação linear, conforme a Equação 3 (Figura 6b). Esta função geralmente está associada à camada de saída das RNAs, pois podem assumir qualquer valor, tanto positivos quanto negativos (BEALE et al., 2016).

$$Purelin(x) = ax \quad (3)$$

A função de ativação log-sigmóide (*logsig*) é utilizada para propagação de valores positivos, porque a resposta varia entre os valores de 0 a 1, conforme a Equação 4 (GHOSE et. al, 2018) (Figura 6c).

$$Logsig(x) = \frac{1}{e^{-x} + 1} \quad (4)$$

A função de ativação tangente hiperbólica (*tanh*) é utilizada para propagação de valores tanto positivos como negativos, pois a resposta varia entre os valores de -1 a 1, conforme a Equação 5 (GHOSE et. al, 2018) (Figura 6d). As funções de ativação log-sigmoide e tangente hiperbólica são fundamentais na aproximação numérica de funções não lineares, pois são diferenciáveis (HAYKIN, 2001).

$$Tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \quad (5)$$

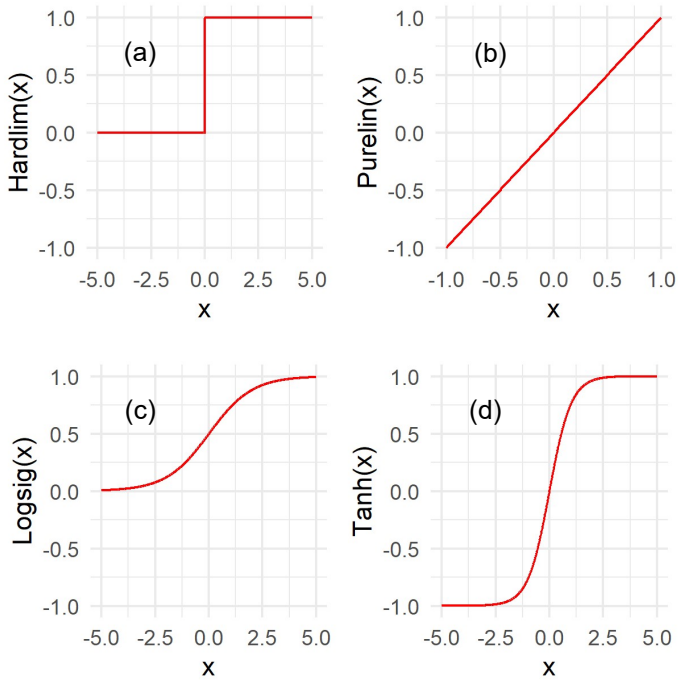

Figura 6: Funções de ativação degrau (a), linear (b), log-sigmoide (c) e tangente hiperbólica (d)

2.3.2 Algoritmos de treinamento

O processo de treinamento de RNAs consiste em ajustar os pesos sinápticos das RNAs, de tal modo que as entradas x_n convirjam para o valor mais próximo à saída desejada y_n, ou seja, com o menor erro possível. De modo geral, estes algoritmos podem ser classificados em treinamentos não supervisionados, aprendizado por reforço e supervisionados (HAYKIN, 2001, YADAV et al., 2015).

Um dos grandes avanços no desenvolvimento de RNAs ocorreu na década de 1980, com a proposição do algoritmo de retropropagação (do inglês

backpropagation), ou Regra Delta Generalizada, apresentado por Rumelhart et al. (1986). É uma técnica de treinamento supervisionada, que consiste em utilizar a regra da cadeia para calcular a derivada da função de erro em relação a cada peso sináptico, assim, estes serão ajustados pelo gradiente descendente do erro (LIPTON et al., 2015). O ajuste dos pesos ocorre em duas etapas: a primeira, chamada de propagação, o vetor de entrada é aplicado nas camadas ocultas e de saída até gerar uma resposta, sendo que todos os pesos são fixos; na segunda, a retropropagação, os pesos são todos ajustados com a regra de aprendizado, que é baseada no erro (HAYKIN, 2001). A partir disso, um grande volume de trabalho em técnicas heurísticas de pré-treinamento e otimização levou a um sucesso empírico impressionante em muitas tarefas de aprendizado supervisionado (WYTHOFF, 1993).

O algoritmo de retropropagação Levenberg-Marquardt é o método supervisionado mais rápido para treinar RNAs até várias centenas de pesos sinápticos (BEALE et al, 2016). Neste algoritmo, que opera em função do erro quadrático médio, os pesos são modificados através da matriz Hessiana aproximada, oriunda da matriz Jacobiana transposta, conforme expresso na Equação 6 (WILAMOWSKI et al., 1999; SAHOO e JHA, 2013).

$$w_{(k+1)} = w_k - (J^T J + \mu I)^{-1} J^T Er \qquad (6)$$

Em que J^T é a matriz Hessiana aproximada, Er é o erro quadrático, J é a matriz Jacobiana, w_k e $w_{(k+1)}$ são os pesos sinápticos, μ é um parâmetro de ajuste de taxa de convergência, e I é a matriz identidade.

2.3.3 Parada antecipada e validação cruzada

Quando se treina as RNAs, geralmente se está interessado em obter desempenho de generalização muito bom, ou seja, que o erro em exemplos não utilizados durante o treinamento seja baixo. No entanto, em algum momento do treinamento, o erro deste conjunto pode estar diminuindo enquanto que o erro em conjuntos de dados diferentes, não utilizados no treinamento, começa a aumentar. Isto é conhecido como *overtraining* (PRECHELT, 1998).

Segundo Prechelt (1998), as formas de se evitar *overtraining* são as técnicas regularização e parada antecipada (do inglês *early stopping*). Conforme Finnoff et al. (1993), esta última é amplamente utilizada porque é simples de compreender, fácil

de implementar e tem sido relatada como superior aos métodos de regularização em muitos casos. A parada antecipada por validação cruzada consiste na avaliação iterativa de um subconjunto de dados, que é independente do utilizado para treinamento da RNA, de modo que quando alcançado algum critério específico, tal como o menor erro pré-definido, seja finalizado o treinamento.

2.3.4 Redes Neurais Recorrentes

Redes Neurais Recorrentes (RNN) são RNAs que recuam informações de neurônios de camadas posteriores para camadas anteriores (*feedback*), de tal forma que pode haver conexões arbitrárias entre quaisquer neurônios (HAYKIN, 2001; SIT et al., 2020). Segundo Diaconescu (2008) isto dá a capacidade de reter informações de estados anteriores, úteis para transferir informações para simulações posteriores. Isto se torna útil quando os sistemas abordados apresentam dependência temporal e espacial, tais como trocadores de calor, sistemas de deforma catalítica, séries temporais etc. (LIN et al., 1998).

A arquitetura Autorregressiva Não Linear com Entradas Exógenas (NARX) é uma rede dinâmica recorrente, com conexões de realimentação do valor resultante do neurônio de saída diretamente para a camada de entrada. Assim como encontrado em outras RNAs, as RNN-NARX também são divididas em camadas de entrada, oculta e de saída (WANG e CHEN, 2022). Nesta perspectiva, o processamento de dados nas camadas ocultas e de saída é igual às redes *feedforward* (CHEN et al. 1990). O modelo NARX é baseado no modelo linear autorregressivo, que é comumente utilizado na modelagem de séries temporais (IZADY et al., 2013). Além disso, conforme Lin et al. (1998), as RNN-NARX convergem mais rapidamente, e tem melhor habilidade de generalização que outras arquiteturas recorrentes.

RNN-NARX é uma importante ferramenta para simulação de sistemas não lineares discretizados no tempo, que pode ser descrito conforme a Equação 7 (MENEZES JR. e BARRETO, 2008). É realizada a aproximação da função não linear desconhecida f, onde o valor atual $y_{(t)}$ é definido por seus próprios valores passados $y_{(t-n)}$, por valores atuais $x_{(t)}$ e passado $x_{(t-n)}$ de entradas exógenas, e por $\varepsilon_{(t)}$, que é o erro associado a cada simulação.

$$y_{(t)} = f\left(y_{(t-1)}, y_{(t-2)}, \dots, y_{(t-n)}, x_{(t)}, x_{(t-1)}, \dots, x_{(t-n)}\right) + \varepsilon_{(t)} \qquad (7)$$

Duas configurações RNN-NARX são importantes e úteis na modelagem. Na forma paralela (*closed loop*) (Figura 7), é possível considerar o valor estimado de saída do modelo (*output*), realimentando diretamente na camada de entrada. Na forma série-paralela (*open loop*) (Figura 8), são utilizados os valores observados (*target*), sendo implementados manualmente na camada de entrada de uma *feedforward* padrão (MENEZES JR. e BARRETO, 2008; BEALE et al., 2016).

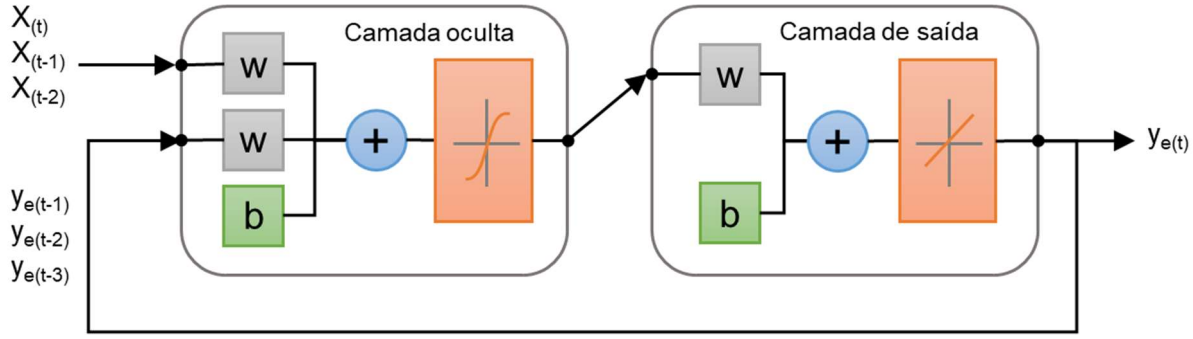

Figura 7: Arquitetura NARX paralela

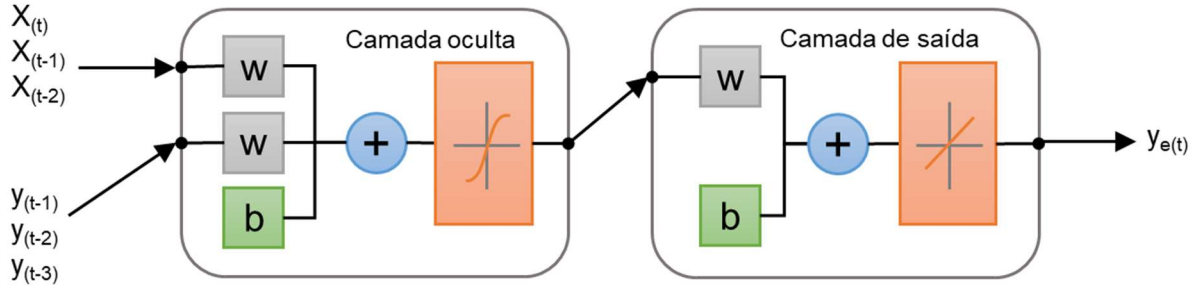

Figura 8: Arquitetura NARX série-paralelo

Quando se está interessado em um horizonte de simulação mais longo, tais como a simulação de múltiplos passos (em inglês *multi-step-ahead*) ou de longo prazo (em inglês *long-term*), deve-se optar pela conversão da estrutura série-paralelo em paralelo (SORJAMAA et al., 2007, CHANG et al., 2016). Neste caso, os componentes da camada de entrada, anteriormente composto de pontos de amostra reais da série observada, são gradualmente substituídos por anteriores valores estimados pelo próprio regressor, tornando-se um sistema autônomo. A simulação com várias etapas posteriores e a modelagem dinâmica são problemas muito mais complexos do que se lidar com simulações de uma etapa de antecedência. Desta maneira, acredita-se que estas são tarefas nas quais as RNN-NARX são relevantes (MENEZES JR. e BARRETO, 2008).

As RNN-NARX são aplicadas em várias áreas do conhecimento. Na hidrologia, são evidenciados estudos para simulação e previsões de vazões

afluentes, de níveis de águas subterrâneas, de parâmetros de qualidade da água, etc (MOSAVI et al., 2018).

Lee e Resdi (2016) analisaram a capacidade das RNN-NARX para previsão simultânea horária em múltiplas estações de medição em uma bacia hidrográfica do distrito de Kemaman em Terengganu, na Malásia. A bacia tem 2.535 km^2, e é propensa a inundações, que causam sérios problemas socioambientais com relativa frequência. Os dados utilizados foram medições de quatro estações meteorológicas, contendo precipitação, temperatura, umidade relativa do ar média e evaporação. Foram obtidos bons resultados quando comparado com observações de campo, utilizando um conjunto mínimo de dados de treinamento de 120 dias.

Debastiani et al. (2016) simularam as vazões diárias em uma sub-bacia hidrográfica do Rio Canoas, Estado de Santa Catarina, Brasil, elaborando RNN-NARX. A bacia possui 1.980 km^2, está bem preservada em termos ambientais e é considerada importante zona de recarga do Aquífero Guarani. Foram utilizadas a precipitação e evapotranspiração de quatro estações meteorológicas, duas inseridas na bacia e duas próximas. Após o treinamento em série-paralelo, o modelo foi convertido para paralelo, a fim de simular um ano de vazões diárias. Os autores exaltaram o potencial da arquitetura, alcançando resultados significantes.

Previsões de níveis diários das águas subterrâneas em um poço agrícola foram realizadas por Guzman et al. (2019), no noroeste do Mississipi, nos Estados Unidos, com o intuito de ajudar os agricultores e as partes interessadas a gerenciar e planejar com eficácia o uso eficiente dos recursos de águas subterrâneas. O local de estudo faz parte de um dos aquíferos com maiores volumes outorgados do país. Os autores relataram boas simulações de RNN-NARX, com dados de precipitações diárias. Além disso, dissertaram sobre as vantagens das simulações baseadas em dados, por oferecerem uma alternativa menos dispendiosa e mais eficiente em escala regional, em comparação com os modelos conceituais.

Wunsch et al. (2018) realizaram a previsão de níveis de águas subterrâneas de seis meses em seis poços no sudoeste da Alemanha, nos estados de Hessen, Baden-Wurttemberg e Baviera. No trabalho, foram utilizadas as precipitações diárias e a temperatura diária média de estações próximas. Os autores mostraram que as RNN-NARX são notavelmente adequadas para realizar previsões de águas subterrâneas para poços de observação não influenciados em todos os tipos de aquíferos, mesmo que o número de parâmetros de entrada seja limitado. Nesse

caso, foram comparadas as previsões de níveis de lençol freático provenientes de diferentes algoritmos de aprendizado de máquina, incluindo RNN-NARX, em 17 poços de águas subterrâneas dentro da área do Alto Reno, que é o maior recurso de águas subterrâneas da Europa Central (WUNSCH et al., 2021). Foi relatado que o RNN-NARX apresenta um melhor desempenho em comparação com os modelos de Redes Neurais Convolucionais (CNN) e Long-Short Term Memory (LSTM), quando apenas poucos dados de treinamento estão disponíveis.

Mendonça et al. (2021) aplicaram as RNN-NARX em uma sub-bacia do Rio Guamá, no Estado do Pará, Brasil, que apresenta área de contribuição de 5.032 km^2. A área de estudo é marcada pela agropecuária e presença de comunidades tradicionais. Foram utilizadas precipitações diárias de quatro estações pluviométricas em torno da bacia, para simular hidrogramas de vazões diárias médias. Os resultados mostraram um coeficiente de determinação de 0,99 e baixos erros de simulação. Os autores destacam, mesmo que conceitualmente, a relação entre o potencial autorregressivo de bacias planas, que naturalmente favorece a infiltração de águas pluviais.

Wang e Chen (2022) utilizaram a RNN-NARX para preencher a lacuna de dados entre os projetos *Gravity Recovery and Climate Experiment* e o *Gravity Recovery and Climate Experiment Follow On,* sobre a Bacia do Rio Yangtzé, na China. Devido a problemas de bateria, o primeiro foi substituído pelo segundo, não havendo monitoramento entre o período de outubro de 2017 até maio de 2018. Conforme os autores, este preenchimento de dados é fator primordial para análise de extremos climáticos, pois são necessárias séries temporais com alta resolução espaço-temporal. Para isso, foram utilizados dados de precipitação, temperatura e anomalias de armazenamento de água terrestre como entradas exógenas. Na etapa de treinamento, foi utilizada a RNN-NARX série-paralelo. A topologia ideal foi encontrada através de aproximações sucessivas, sendo a ideal com nove neurônios na camada oculta. Posteriormente, na fase de simulação do período não observado, foi utilizada a estrutura RNN-NARX paralelo. Comparando os valores simulados e observados após a falha histórica, os critérios estatísticos obtidos, foram de 1,34 cm para a raiz do erro quadrático médio (RMSE), e 0,95 de coeficiente de eficiência de Nash-Sutcliffe (NSE). Os autores concluíram que foi obtido sucesso na previsão de eventos de inundação durante o período de lacuna de dados, sugerindo que a RNN-

NARX é promissora para prever extremos hidrológicos de curto prazo sobre a área de estudo.

3 MATERIAL E MÉTODOS

3.1 ÁREA DE ESTUDO

Foram selecionadas cinco pequenas bacias hidrográficas da Amazônia. Esta seleção foi baseada no trabalho de Santana e Blanco (2020), que definiram pequenas bacias hidrográficas na Amazônia sendo aquelas com área de drenagem inferior a 500 km^2. Para essas bacias, os dados de precipitações e vazões diárias foram obtidos no banco de dados da Agencia Nacional de Águas (ANA), através do portal HidroWeb, sob o endereço eletrônico https://www.snirh.gov.br/hidroweb/serieshistoricas. A Tabela 1 mostra os dados das estações meteorológicas e a resolução temporal das séries temporais utilizadas. Como critério de escolha das resoluções temporais, buscaram-se intervalos que não apresentassem falhas históricas. A Figura 9 mostra a localização das estações utilizadas no estudo.

As simulações dos hidrogramas e curvas de permanência das bacias foram construídas a partir da abordagem de RNN-NARX, que são modelos empíricos baseados em séries temporais (DEVIA et al., 2015; GHOSE et al., 2018). Estes modelos foram desenvolvidos a partir da biblioteca *Deep Learning Toolbox*, com o software MATLAB, da Mathworks. As bases fundamentais dos algoritmos utilizados constam no Apêndice A. Supondo que essas pequenas bacias têm precipitação homogênea em toda a extensão da bacia, espera-se que a partir da capacidade de aprendizado no erro e aproximação do sistema não linear, seja possível estimar vazões diárias médias de uma série temporal de 365 dias, a fim de proporcionar uma ferramenta de gestão de recursos hídricos em pequenas escalas.

Tabela 1: Estações hidrometeorológicas utilizadas

Nome da estação	Código	Latitude (º)	Longitude (º)	Tipo	Tempo (anos)	Sigla
Ourém	147016	-1,5517	-47,1172	Pluviométrica	2009-2014	P1
Abreulândia	949000	-9,6244	-49,1553	Pluviométrica	2009-2014	P2
Arraial	147000	-1,55	-47,1167	Pluviométrica	1966-1971	P3
Colinas do Tocantins	848000	-8,0528	-48,4817	Pluviométrica	1993-1998	P4
KM 947 BR-163	855000	-8,1872	-55,1194	Pluviométrica	2015-2018	P5
Marambaia	31600000	-1,65220	-47,1167	Fluviométrica	2009-2014	Q1
Fazenda Craveiro	27370000	-9,60360	-48,9708	Fluviométrica	2009-2014	Q2
Arraial	32300000	-1,31667	-47,11667	Fluviométrica	1966-1971	Q3
Próximo Colinas do Tocantins	23130000	-8,0692	-48,4517	Fluviométrica	1993-1998	Q4
Base do Cachimbo	17345000	-9,3589	-54,9039	Fluviométrica	2015-2018	Q5

Fonte: Agência Nacional de Águas (2021)

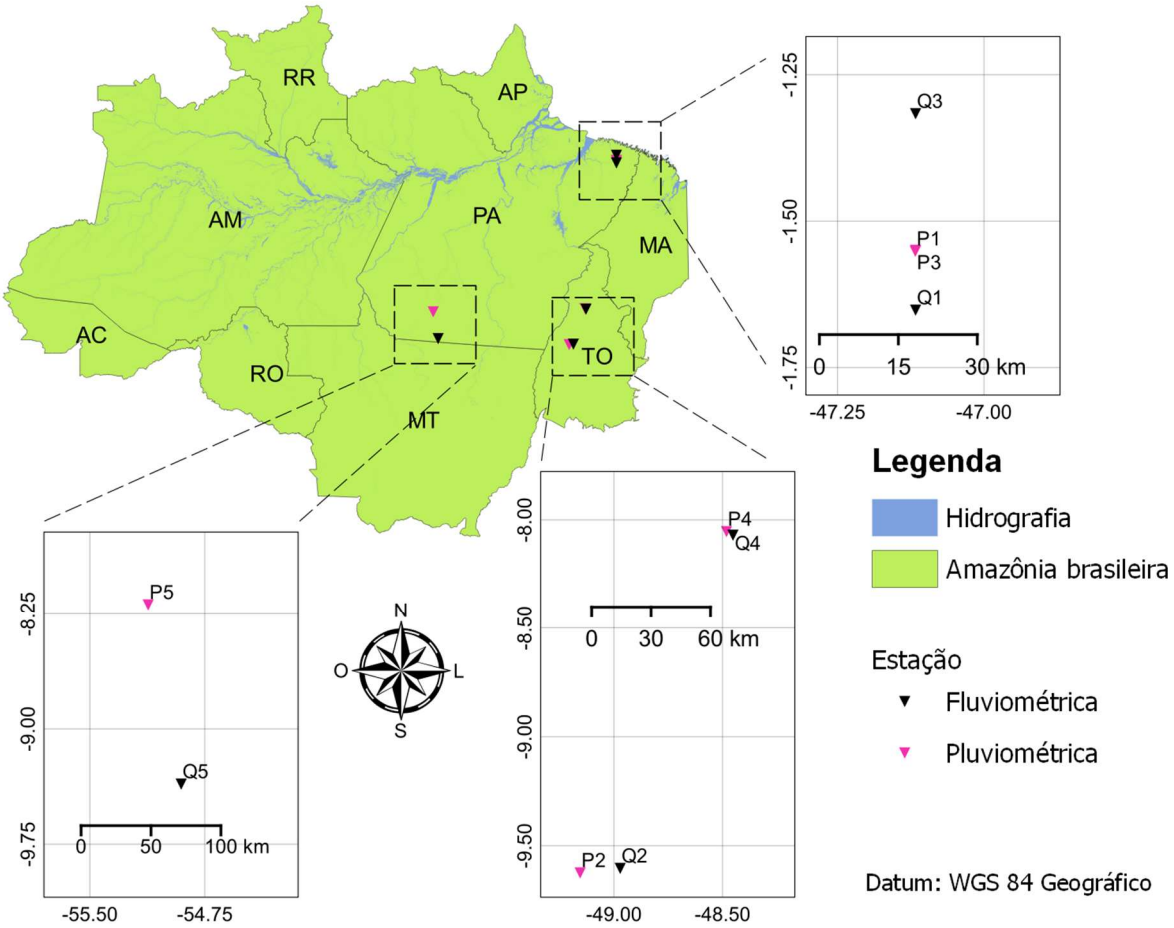

Figura 9: Localização das estações hidrometeorológicas de estudo

3.2 ANÁLISE EXPLORATÓRA

As séries temporais utilizadas foram submetidas a uma análise exploratória, baseadas nas funções de correlação cruzada e autocorrelação parcial, para identificar vetores de precipitação e vazões defasadas significantes. Segundo Debastiani et al. (2016), estas análises têm a capacidade de identificar processos-chaves ligados às variáveis de entrada que influenciam diretamente as vazões no exultório de uma bacia hidrográfica, especialmente quando se desconhece o sistema a ser modelado (OYEBODE E STRETCH, 2018). Logo, espera-se identificar potencialidades nas variáveis disponíveis, que serão utilizadas como vetores de entrada nas RNN-NARX.

A correlação cruzada é o método utilizado para analisar o comportamento da correlação linear entre duas variáveis em função das defasagens no tempo das séries temporais (BAYER et al., 2012). Calcula-se este índice através das Equações 9 e 10.

$$cov_{(y,x_{(t-k)})} = \frac{1}{n}\left[\sum_{t=k+1}^{n}(y-\bar{y})(x_{t-k}-\bar{x})\right] \quad (8)$$

$$r_{x(t-k)y} = \frac{cov_{(y,x_{(t-k)})}}{\sqrt{(y-\bar{y})}} \quad (9)$$

Em que *cov* é a covariância amostral, y e x são as variáveis abordadas, k são os valores defasados do instante t, $\bar{y}$ e $\bar{x}$ são médias amostrais, e n é o número de observações.

A função de autocorrelação é uma medida da correlação linear entre as observações de uma série temporal que são separadas por k unidades de tempo (BAYER et al., 2012), calculada conforme a Equação 11. No entanto, evidencia-se em muitas séries temporais que os coeficientes de autocorrelação de defasagens de menor intervalo de tempo podem influenciar nos coeficientes de defasagens de maior intervalo de tempo. Diante disso, surge o coeficiente de autocorrelação parcial que tem a capacidade de filtrar tais influências, para que possa ser representada a autêntica dependência temporal de cada defasagem (BAYER et al., 2012). Este índice é calculado conforme a Equação 12.

$$\rho_k = \left[\sum_{t=k+1}^{n}(y_t-\bar{y})(y_{t-k}-\bar{y})\right]\left[\sum_{t=1}^{n}(y_t-\bar{y})^2\right]^{-1} \quad (10)$$

$$\phi_{kk} = \left[\rho_k - \sum_{j=1}^{k-1}\left(\phi_{k-1j}-\rho_{k-j}\right)\right]\left[1-\sum_{j=1}^{k-1}\left(\phi_{k-1j}\rho_k\right)\right]^{-1} \quad (11)$$

Em que, ρ_k é o coeficiente de autocorrelação, y é a variável autocorrelacionada nos instantes t e t-k, $\bar{y}$ é a média de y, ϕ_{kk} é o coeficiente de autocorrelação parcial e ϕ_{k-1j} são os coeficientes de autocorrelação parcial filtrados das defasagens anteriores. Com base nessas correlações lineares, foi identificado diferentes vetores de entrada, com nível de significância de 5% (α = 0,05), para serem implementados na camada de entrada dos modelos RNN-NARX.

3.3 MODELOS PROPOSTOS

A primeira etapa para aplicação dos modelos RNN-NARX está relacionada à escolha dos conjuntos de dados. A divisão dos conjuntos de treinamento, validação e teste foi definida, conforme a disponibilidade dos dados de precipitação diária e vazões diárias médias em cada bacia, sendo que a unidade utilizada para a divisão entre os conjuntos é a cada ano de dados completo, sem qualquer falha histórica.

A topologia das RNN-NARX utilizadas é descrita a seguir. O número de entradas é definido pela quantidade de precipitações e vazões defasadas implementadas. Foi utilizada apenas uma camada oculta, pois é suficiente para aproximar qualquer função não linear (HORNIK et al.,1989). E, o número de neurônios ocultos foi definido pelo método de tentativa e erro, com busca do número de neurônios entre 2 e 20, sendo aferido erro quadrático médio a cada acréscimo de neurônio. Cada combinação possível é treinada várias vezes, e a configuração com o menor erro é a escolhida para discussão.

Ao utilizar redes de algoritmos de retropropagação, dependendo da função de ativação dos neurônios, é necessário realizar o pré-tratamento dos dados utilizados para o treinamento. As funções de ativação comumente utilizadas para abordagem de sistemas não lineares são as funções sigmoides, tal como a logística e a tangente hiperbólica (HAYKIN, 2001; MAIER e DANDY, 2000). Conforme Rezaeian Zadeh et al. (2010) mostraram que a função de ativação tangente hiperbólica teve um desempenho melhor do que a ativação sigmóide logística na previsão de vazões diárias. O mesmo foi evidenciado por Yonaba et al (2010). Embasado nisso, foi utilizada a função de ativação tangente hiperbólica (*tanh*) nas camadas ocultas da RNN-NARX. A partir disso, os dados de precipitações e vazões foram submetidos a uma função de normalização. Sola e Sevilla (1997) afirmaram que a normalização dos dados de entrada é crucial para obter bons resultados, bem como agilizar substancialmente os cálculos. Guzman et al. (2019) comentou que este pré-processamento reduz o ruído no processo de calibração. Desta forma, os dados foram reajustados, conforme a Equação 12, ao intervalo de -1 a 1. Este intervalo corresponde à faixa de funcionamento da função de ativação tangente hiperbólica.

$$x_n = i_{min} + \frac{x - x_{min}}{x_{máx} - x_{min}}(i_{máx} - i_{min}) \qquad (12)$$

Em que x é a variável, que se deseja normalizar, x_n é a variável normalizada, x_{min} e $x_{máx}$ são os valores mínimos e máximos encontrados nos dados, e i_{min} e $i_{máx}$ são, respectivamente, os limites inferior e superior do intervalo, que se deseja normalizar. Ao fim das simulações, as vazões diárias foram desnormalizadas, obedecendo a Equação descrita anteriormente, para comparar os resultados obtidos com os dados observados.

Para ajuste dos parâmetros treináveis das RNN-NARX, foi utilizado o algoritmo de retropropagação do erro Levenberg-Marquardt (*trainlm*), com o auxílio

da função do erro quadrático médio em cada iteração. Foi utilizado, também, o método de parada antecipada por validação cruzada para evitar *overtraining* cujos critérios de finalização do treinamento estão expostos na Tabela 2.

Tabela 2: Condições de parada antecipada por validação cruzada

Parâmetro	Valor
Máximo número de iterações	1500
Desempenho desejado (erro quadrático médio máximo)	0,005
N° máx. de aumento de desempenho na validação	300
Tempo máximo de treinamento (segundos)	3000

Após a fase de ajuste dos parâmetros, ocorrida no treinamento, as RNN-NARX foram convertidas para a estrutura paralela, para simulação de longo período (*long-term*), de forma que fosse possível estimar os próximos 365 dias somente com as precipitações observadas, tal como ocorre em Debastiani et al. (2016), Wunsch et al. (2018), e Wang e Chen (2021).

Foi efetuada, também, a análise de sensibilidade do modelo frente a diferentes tamanhos de amostras de precipitações diárias e vazões diárias, de modo a elucidar o tempo mínimo necessário para realizar simulações adequadas. As amostras destinadas ao treinamento das RNN-NARX dependem do tamanho da série temporal encontrada em cada bacia. No entanto, foi adotada a redução das amostras em 1 ano como critério para fracionamento.

3.4 CRITÉRIOS DE DESEMPENHO

Para verificar o desempenho estatístico dos modelos propostos, e identificar os modelos mais adaptados, os hidrogramas e curvas de permanência observadas e simuladas foram submetidos aos seguintes critérios: o coeficiente de eficiência de Nash-Sutcliffe (NSE), a raiz do erro quadrático médio (RMSE), a raiz do erro quadrático médio padronizado (RSR), o viés percentual (PBIAS), e o erro absoluto percentual médio (MAPE).

O índice NSE (Equação 14) é uma estatística normalizada, que determina a intensidade relativa da variância residual, comparada à variância dos dados medidos (NASH e SUTCLIFFE, 1970). Este é um dos critérios estatísticos mais utilizados para avaliar a eficácia dos modelos hidrológicos (ARNOLD et al. 2012). Valores de NSE iguais a 1 indicam um ajuste perfeito entre os dados observados e simulados, e valores menores ou iguais a 0 indicam que a média dos dados observados é um

preditor mais preciso que os simulados pelo modelo (MORIASI et al., 2007; PRAHARAJ et al., 2018).

$$\mathrm{NSE} = 1 - \left[\frac{\sum_{i=1}^{n}(\mathrm{Q} - \mathrm{Q_e})^2}{\sum_{i=1}^{n}(\mathrm{Q} - \mathrm{Q_m})^2}\right] \quad (13)$$

O RMSE (Equação 15) mostra os erros de simulação na mesma unidade da variável abordada, no entanto, como é proveniente do erro quadrático, destaca erros com maiores magnitudes (MORIASI et al., 2007). O valor zero indica que o erro de simulação é inexistente ou desprezível.

$$RMSE = \sqrt{\frac{\sum_{1}^{n}(Q_e - Q)^2}{n}} \quad (14)$$

Embora seja comumente aceito que quanto menor o RMSE melhor o desempenho do modelo, Singh et al. (2005) apresentam uma diretriz para qualificar o que é considerado um RMSE baixo com base no desvio padrão das observações. Diante disso, surge o índice RSR (Equação 16) que é a razão entre o índice RMSE e o desvio padrão das vazões observadas, ou seja, é uma versão padronizada do RMSE. Segundo Moriasi et al. (2007) o RSR varia do valor ótimo, 0 (zero), até um grande valor positivo. Além disso, os autores comentam que quanto menor o RSR, menor o RMSE, e melhor o desempenho da simulação do modelo.

$$RSR = \frac{\sqrt{\sum_{i=1}^{n}(Q - Q_e)^2}}{\sqrt{\sum_{i=1}^{n}(Q - Q_m)^2}} \quad (15)$$

O índice PBIAS (Equação 17) avalia a tendência média dos valores simulados, que podem ser maiores ou mais baixos do que seus equivalentes observados. Seu valor ideal é 0, indicando simulação precisa. Valores positivos indicam viés de superestimação do modelo, enquanto que valores negativos indicam viés de subestimação do modelo (MORIASI et al. 2007).

$$\mathrm{PBIAS} = \left(\frac{\sum_{i=1}^{n}\mathrm{Q_e} - \mathrm{Q}}{\sum_{i=1}^{n}\mathrm{Q}}\right)100 \quad (16)$$

O índice MAPE mostra os erros absolutos sob a forma percentual (GYASI-AGYEI et al., 1995), logo, é possível observar o panorama geral dos erros, independente se há subestimação ou superestimação. O valor zero indica que o erro de simulação é inexistente ou desprezível. É calculado conforme a Equação 18 (XU e LIU, 2013).

$$\mathrm{MAPE} = \frac{\sum_1^{\mathrm{n}} \left(\frac{|\mathrm{Q}_e - \mathrm{Q}|}{\mathrm{Q}} \right)}{\mathrm{n}} 100 \quad (17)$$

Nas equações 12-18, Q são as vazões observadas, Q_e são as vazões estimadas, Q_m é a média das vazões observadas e n é o número amostral.

Além da importância de calcular os índices estatísticos dos modelos, é fundamental adotar critérios específicos para classificação quanto à qualidade dos modelos. Neste estudo foi adotada a classificação proposta por Moriasi et al. (2007), tal qual classificaram os índices conforme a Tabela 3.

Tabela 3: Classificação do desempenho estatístico dos modelos hidrológicos

Índice de desempenho	RSR	NSE	PBIAS
Muito bom	0,00 ≤ RSR ≤ 0,50	0,75 < NSE ≤ 1,00	PBIAS < ± 0,10
Bom	0,50 < RSR ≤ 0,60	0,65 < NSE ≤ 0,75	± 0,10 ≤ PBIAS < ± 0,15
Satisfatório	0,60 < RSR ≤ 0,70	0,50 < NSE ≤ 0,65	± 0,15 ≤ PBIAS < ± 0,25
Insatisfatório	RSR > 0,70	NSE ≤ 0,50	PBIAS ≥ ± 0,25

Por fim, foi realizada a análise de pesos sinápticos das RNN através do coeficiente de contribuição relativa de Garson (1991), CR (Equação 19). O método consiste em avaliar tais parâmetros de cada variável e atribuir um percentual de relevância à estimativa dos valores de saída (DEBASTIANI et al., 2016; MENDONÇA et al., 2021).

$$\mathrm{CR} = \frac{\sum_1^{n_o} \left(w_{jo} \frac{w_{ij}}{\sum_1^{n_i} w_{ij}} \right)}{\sum_1^{n_i} \left[\sum_1^{n_o} \left(w_{jo} \frac{w_{ij}}{\sum_1^{n_i} w_{ij}} \right) \right]} \quad (18)$$

Em que w_{ij} é o peso atribuído ao vetor de entrada i em direção ao neurônio oculto j, w_{jo} é o peso atribuído à resposta do neurônio oculto j em direção ao neurônio de saída o, n_o é o número de neurônios ocultos e n_i é o número de variáveis de entrada. O fluxograma mostrado na Figura 10 sumariza a metodologia da pesquisa.

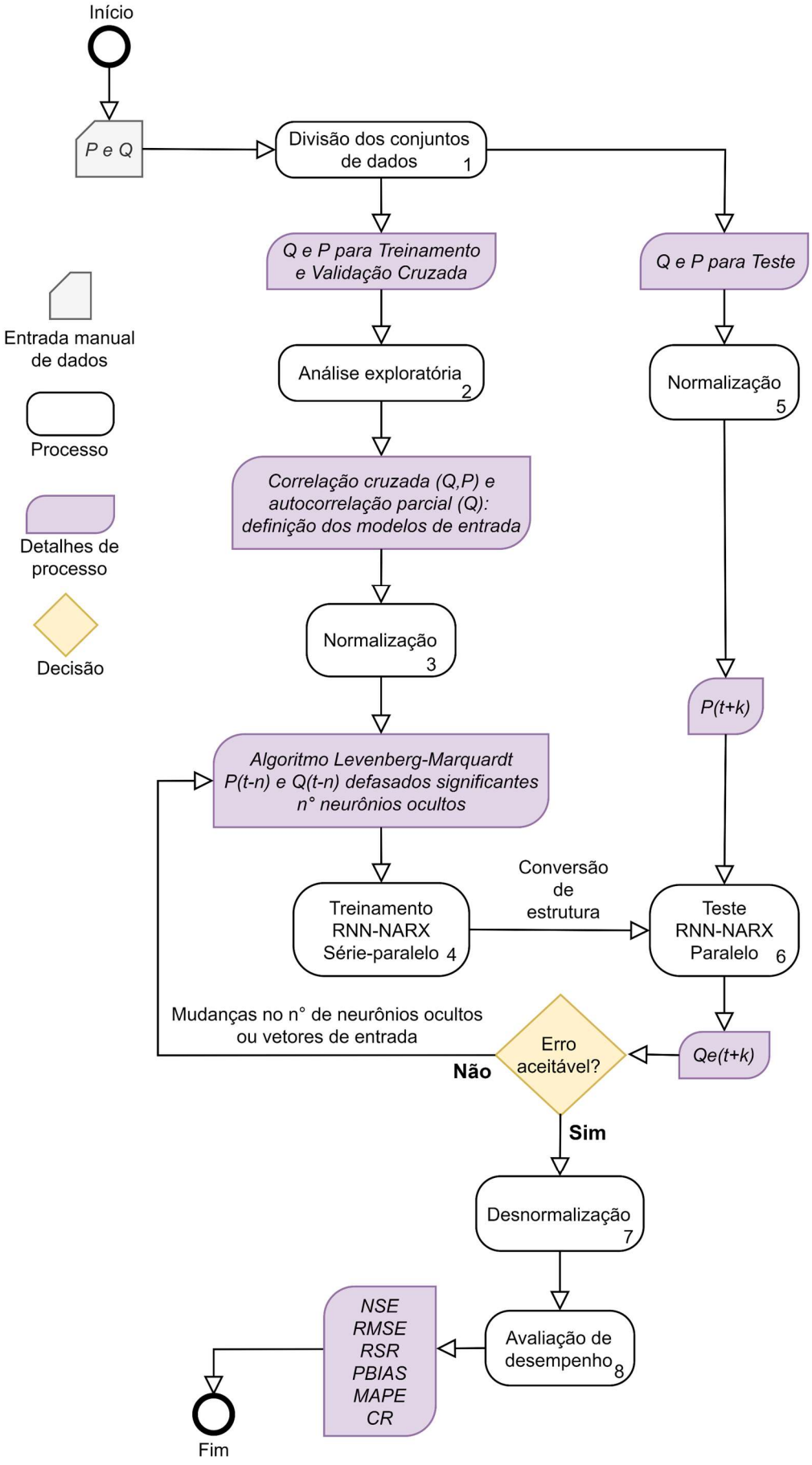

Figura 10: Fluxograma da metodologia empregada

4 RESULTADOS E DISCUSSÃO

4.1 IGARAPÉ DA PRATA

A estação fluviométrica Marambaia (Q1) (Cód. 31600000) monitora as vazões diárias de uma pequena sub-bacia hidrográfica do Igarapé da Prata, que tem área de drenagem de 32,7 km^2 e está situada no município de Capitão-Poço, Estado do Pará (ANA, 2021). A altitude média desta bacia hidrográfica é de 79 m, mínima de 46 m, e máxima de 104 m (Figura 11). A declividade é classificada como suave ondulada, pois a média é de 3,45% (USGS, 2021; EMBRAPA, 1999). O tempo de concentração é estimado em 2,2 horas (CARIELLO, 2014). Além disso, a comunidade rural de Nova Colônia está inserida nesta área, sob a latitude 1°42'02"S e longitude 47°06'27"W.

As classes de solos inseridas na bacia Q1 são o latossolo amarelo distrófico e o argissolo vermelho-amarelo distrófico (IBGE, 2021). Sartori et al. (2005) comentam que esta classe pedológica é marcada por altas profundidades e condutividade hidráulica. Porém, o uso do solo é predominantemente destinado à pastagem, com vegetação forrageira herbácea de espécies cultivadas (INPE, 2014). Logo, a dispensa do manejo sustentável do solo ao longo das atividades desenvolvidas pode comprometer as características físicas do solo, ou seja, pode ocasionar mudanças significativas no comportamento hidrológico da bacia ao longo do tempo.

A 15 km de distância da estação Q1, na direção norte, está localizada a estação pluviométrica Ourém (P1) (Cód. 147016). Blanco et al. (2005), confirmaram que os regimes pluviométricos da estação meteorológica P_1 e as vazões de Q_1 são razoavelmente correlacionados. O clima desta área é definido como Am, conforme a classificação Köppen-Geiger (KÖPPEN, 1936; GEIGER, 1954), o que representa um clima tropical com regime de monções, altos índices pluviométricos, e baixa variabilidade na temperatura, que sempre está acima dos 18°C. A variabilidade intra-anual da precipitação na bacia Q_1 é caracterizado por dois períodos distintos (MORAES et al., 2005). O período chuvoso é compreendido entre os meses de janeiro e julho, e o período de estiagem é compreendido entre os meses de agosto a dezembro, conforme a normal climatológica observada na estação P1 (Figura 12).

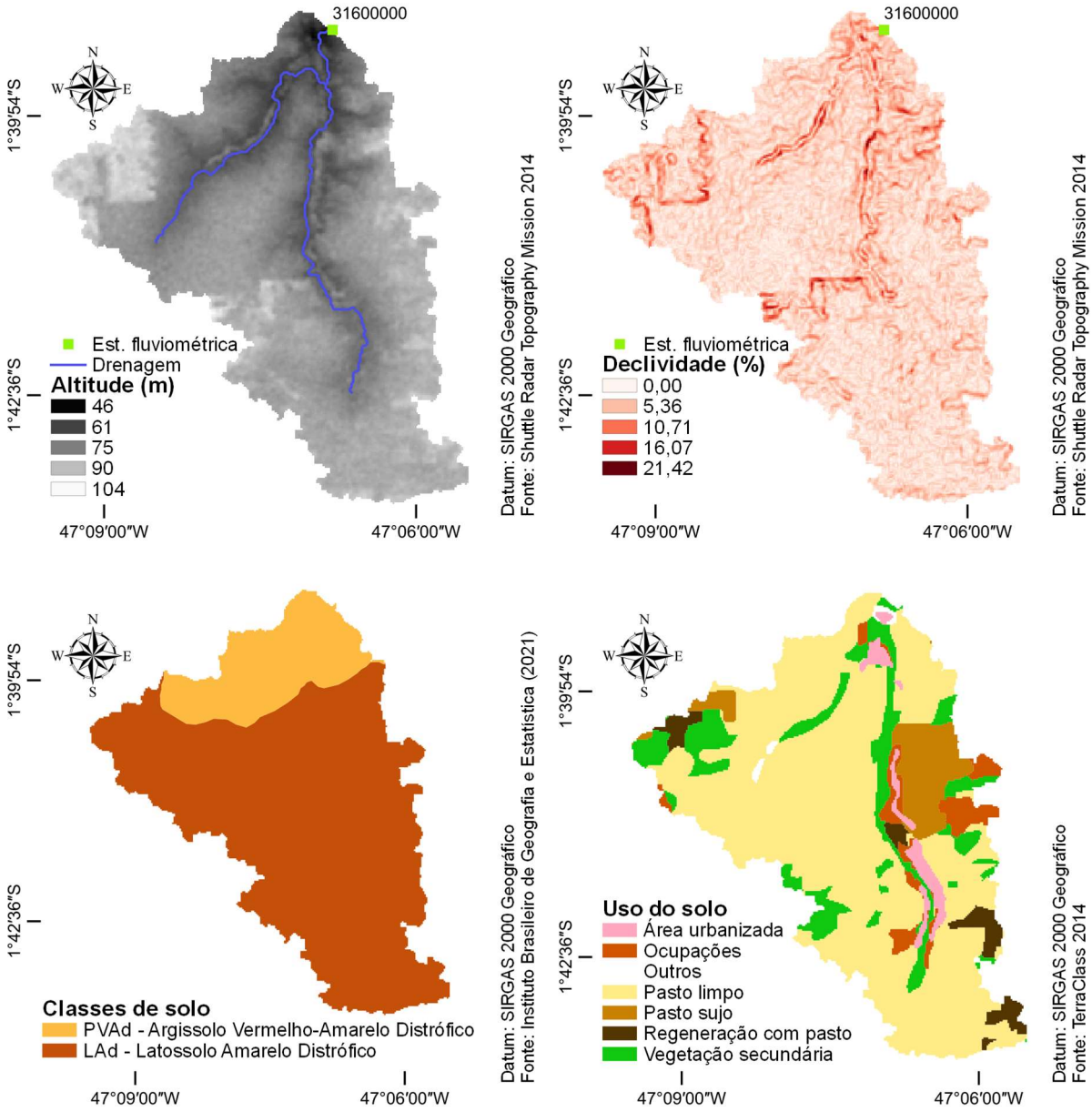

Figura 11: Caracterização da bacia da estação Q1

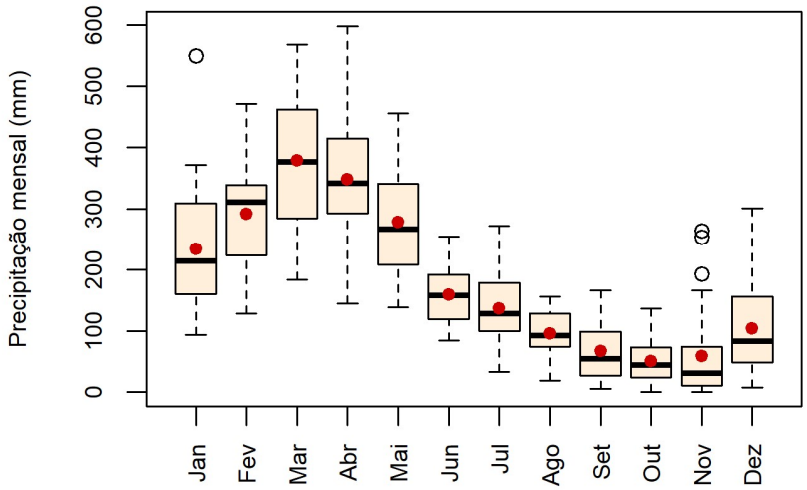

Figura 12: Normal climatológica da estação P1 entre 1991-2020

As séries temporais P1 e Q1 selecionadas têm seis anos hidrológicos completos, desde 2009 até 2014. A divisão dos conjuntos de dados ocorreu da seguinte maneira: fazem parte do conjunto de treinamento os anos hidrológicos entre 2009 e 2011 (1961 dados), de validação os anos hidrológicos de 2012 e 2013 (731 dados) e do conjunto de teste o ano de 2014 (365 dados). É possível observar a divisão dos conjuntos na Figura 13, sob a forma de hidrogramas de vazões.

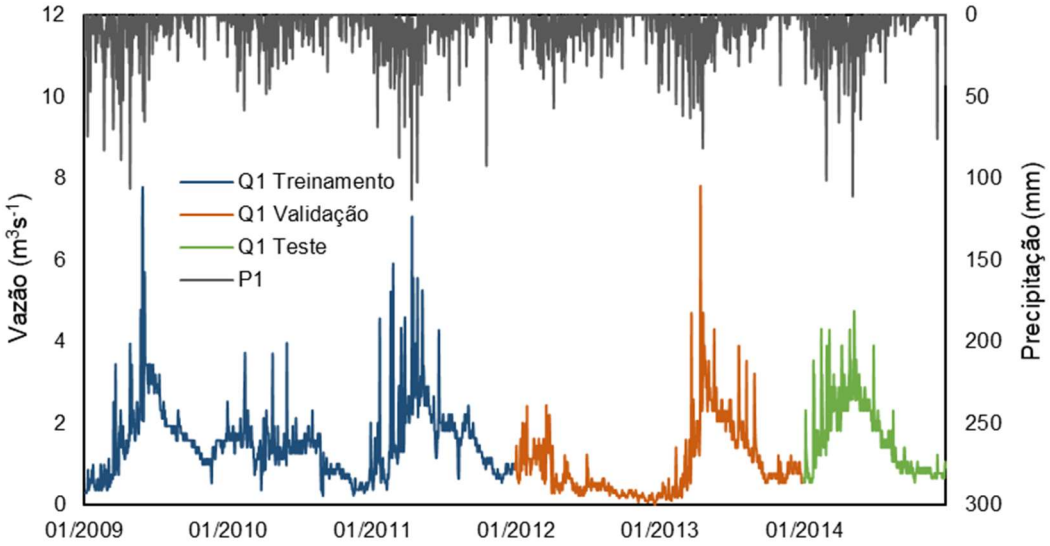

Figura 13: Hidrogramas de vazões da estação Q1 entre 2009-2014

A Tabela 4 mostra a estatística descritiva das variáveis por ano hidrológico. Observa-se que os anos com o menor e o maior índice pluviométrico ocorrem em 2010 e 2011, respectivamente, com total anual de 1.738 mm e 2.716 mm. No entanto, o menor e o maior índice fluviométrico ocorrem, respectivamente, nos anos de 2012 (vazão média de 0,58 m^3s^{-1} e a vazão de referência Q_{95} de 0,13 m^3s^{-1}) e 2011 (vazão média de 1,68 m^3s^{-1} e a vazão de referência Q_{95} de 0,63 m^3s^{-1}). Uma das principais causas dessa diferença pode ser a variabilidade espacial e temporal de anomalias meteorológicas, tais como o fenômeno El Niño Oscilação Sul (ENOS). Segundo dados do Instituto Nacional de Pesquisas Espaciais (INPE, 2021), os anos de 2009-2010 foram alvos de El Niño moderado, fase positiva do fenômeno ENOS, que tem a capacidade de ocasionar diminuição da precipitação em baixas latitudes, como na bacia de Q1. Em contrapartida, no ano de 2010-2011 houve fenômeno La Niña de magnitude, o que pode acarretar aumento dos índices pluviométricos em regiões tropicais (INPE, 2021). Além disso, o ano de 2012 sugere variabilidade espacial entre a precipitação efetivamente ocorrida em Q1 e a observada em P1. Espinoza et al. (2021) mostraram que os meses de dezembro-2011 a março-2012

apresentaram anomalias negativas no comportamento da precipitação no nordeste paraense, com reduções entre -50 e -100 mm/mês. Segundo os mesmos autores, houve a concentração de precipitação na parte ocidental amazônica, resultando em eventos extremos.

Tabela 4: Estatística descritiva das variáveis P1 e Q1

Variável hidrológica	Estatística	Treinamento			Validação		Teste
		2009	2010	2011	2012	2013	2014
P1	Média[1] (mm/dia)	7,34	4,76	7,44	5,38	6,32	7,19
	Anual (mm/ano)	2.678,10	1.738,00	2.715,90	1.969,70	2.306,60	2.623,70
	Máx (mm/dia)	106,20	58,40	113,30	56,70	81,70	111,20
	Min[2] (mm/dia)	0,20	0,30	0,20	0,30	0,30	0,40
	Coef. de variação (%)	194,45	179,65	200,02	161,79	182,89	193,94
Q1	Média (m³/s)	1,63	1,20	1,68	0,58	1,34	1,62
	Máx (m³/s)	7,76	3,93	7,03	2,42	7,73	4,69
	Min (m³/s)	0,27	0,21	0,35	0,02	0,10	0,54
	Coef. de variação (%)	59,86	46,34	57,17	76,29	73,78	55,05
	$Q_{50\%}$ (m³/s)	1,56	1,25	1,62	0,44	1,05	1,39
	$Q_{95\%}$ (m³/s)	0,35	0,35	0,63	0,13	0,23	0,64

[1]Considerando dias sem chuva
[2]Acima de zero

A Figura 14 mostra as funções de correlação cruzada entre Q_1 e P_1, em diferentes resoluções temporais. Considerando os anos entre 2009-2011, e o nível de significância de 5%, destacam-se as precipitações defasadas entre 0-6 dias, 9-11 dias e 14 dias. Considerando os anos entre 2010 e 2011, e o nível de significância de 5%, destacam-se as precipitações diárias defasadas entre 0-19 dias, e entre 22-28 dias defasados. E considerando o somente ano de 2011, e o nível de significância de 5% ($\alpha = 0,05$), destacam-se as precipitações diárias defasadas entre 0-11 dias, entre 13-16 dias, 18 dias e 27 dias. Todas as correlações são consideradas fracas (menores que 0,5), conforme Peck et al. (2015). Assim, verifica-se que a função de correlação cruzada referente aos anos entre 2009-2011 apresenta comportamento diferenciado, com a diminuição dos coeficientes em toda a extensão do gráfico. Além disso, as funções de correlação cruzada de 2010-2011 e 2011 apresentam um aumento dos coeficientes principalmente entre as precipitações defasadas em 3-14 dias. Logo, a redução dos coeficientes de correlação pode estar relacionada à diminuição do regime pluviométrico ocasionado pelo El Niño, enquanto que o aumento pode estar associado ao efeito inverso, proveniente do fenômeno La Niña.

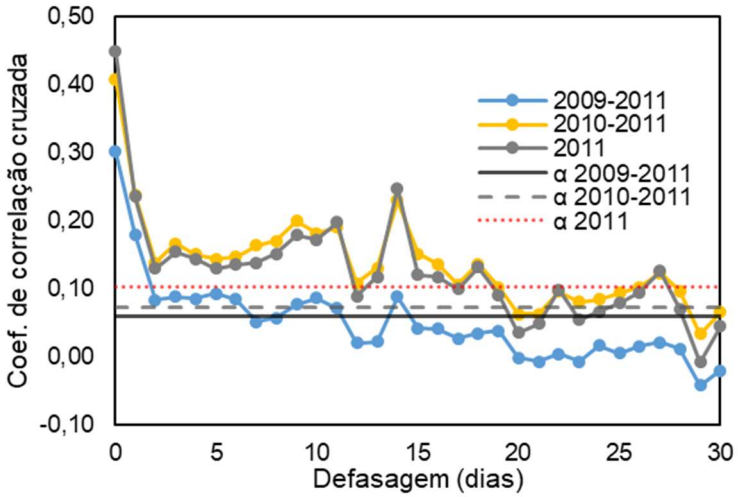

Figura 14: Função de correlação cruzada entre as estações P1 e Q1

As Figuras 15 e 16 mostram as funções de autocorrelação e autocorrelação parcial de Q_1, respectivamente. Considerando os anos entre 2009-2011, e o nível de significância de 5%, destacam-se as vazões diárias entre 1-6 dias, 8-9 dias, 14 dias, 24 dias, 26 e 27 dias defasados. Considerando os anos entre 2010-2011, e um nível de significância de 5%, destacam-se as vazões diárias entre 1-6 dias, 9 dias, 14 dias, 26 e 27 dias, e 29 dias defasados. E, considerando o ano de 2011, e o nível de significância de 5% (α = 0,05), destacam-se as vazões diárias entre 1-6 dias, 14 dias, 26 e 27 dias defasados. A vazão defasada em 1 dia tem correlação linear forte, e as demais variáveis apresentam correlação fraca (PECK et al., 2015). Diferente do que ocorre na função de correlação cruzada entre P_1 e Q_1, as funções de autocorrelação sofrem diminuição dos coeficientes nos anos de ocorrência de La Niña e aumento no ano de El Niño.

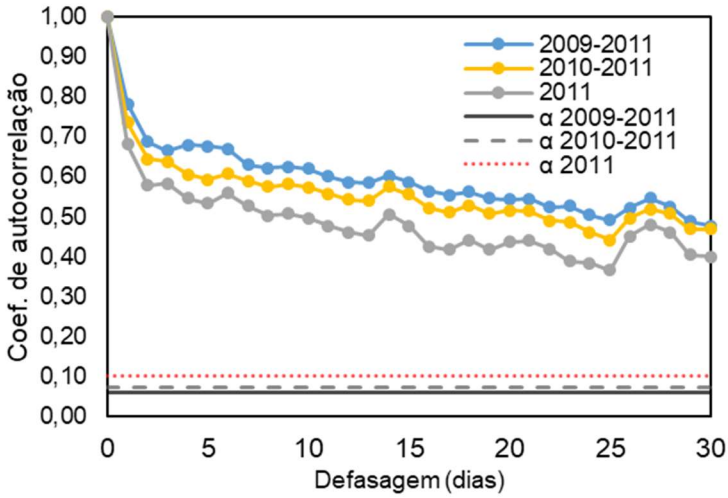

Figura 15: Função de autocorrelação da estação Q1

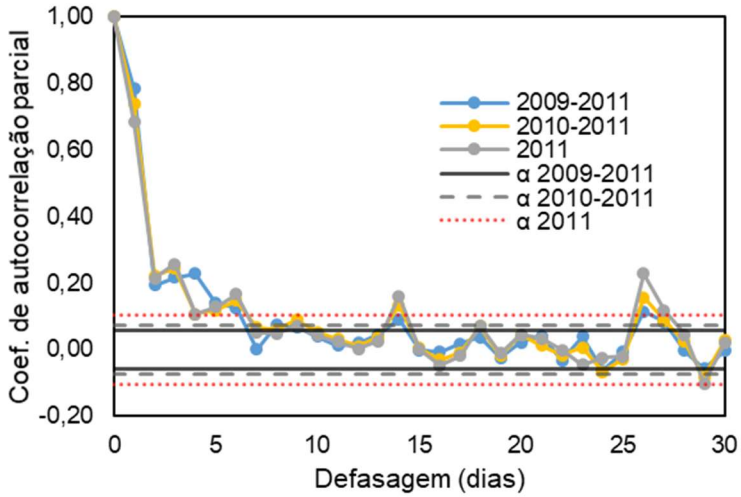

Figura 16: Função de autocorrelação parcial da estação Q1

A partir das correlações obtidas anteriormente, optou-se por avaliar cinco diferentes vetores de entrada. Assim, os modelos propostos para simulação de Q1 e os respectivos índices de desempenho são mostrados na Tabela 5. O modelo que apresenta o melhor desempenho tem como entradas as precipitações defasadas entre 0-14 dias, e vazões defasadas entre 1-6 dais ($P1_{0:14}Q1_{1:6}$). Analisando os critérios de desempenho, foram obtidos os índices NSE de 0,658, RSR de 0,58, PBIAS de -8,72%, RMSE de 0,52 m^3s^{-1} e MAPE de 22,35%, o que o classifica como um bom modelo.

Tabela 5: Desempenhos dos modelos propostos para simulação de Q1

Modelo	NO	NSE	RSR	RMSE	PBIAS	MAPE
$P1_{0:6}Q1_{1:6}$	6	0,635	0,60	0,54	-8,58	22,51
$P1_{0:14}Q1_{1:6}$	3	0,658	0,58	0,52	-8,72	22,35
$P1_{0:14}Q1_{1:14}$	7	0,633	0,61	0,54	-11,99	21,16
$P1_{0:27}Q1_{1:14}$	7	0,601	0,63	0,56	0,11	27,98
$P1_{0:27}Q1_{1:27}$	5	0,542	0,68	0,6	-3,79	26,74

NO-Neurônios Ocultos; NSE-Coeficiente de Eficiência de Nash-Sutcliffe; RSR-Raiz do erro quadrático médio padronizado; RMSE-Raiz do erro quadrático médio; PBIAS-Viés percentual; MAPE-Erro absoluto percentual médio

As Figuras 17 e 18 mostram as vazões observadas e simuladas pela RNN-NARX para simulação de Q1, modelo $P1_{0:14}Q1_{1:6}$. Os maiores erros percentuais são observados nos meses de janeiro (MAPE de 42%) e fevereiro (MAPE de 41%), período em que começam as chuvas na região. O maior erro absoluto percentual corresponde à vazão observada com 0,643 m^3s^{-1}, no dia 25/01/2014, que foi simulada com 1,44 m^3s^{-1} (superestimação de 119%). Outro exemplo é a vazão observada no dia 03/03/2014, com 2,04 m^3s^{-1}, que foi simulada com 4,21 m^3s^{-1}

(superestimação de 106%). Por outro lado, os menores erros percentuais estão no mês de junho (MAPE de 10,18%) e o mês de julho (MAPE de 12,00%), correspondentes ao período de recessão do hidrograma. O menor erro absoluto percentual foi de 0,23%, com a vazão observada de 0,89 m^3s^{-1}, no dia 08/11/2014, que foi simulada com 0,89 m^3s^{-1}.

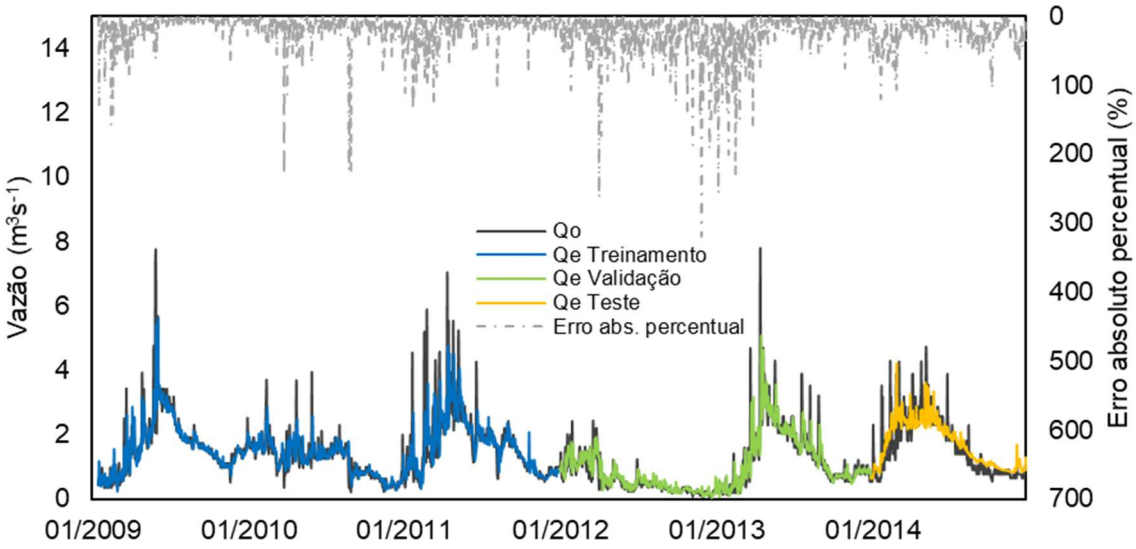

Figura 17: Hidrogramas observados e simulados para a estação Q1

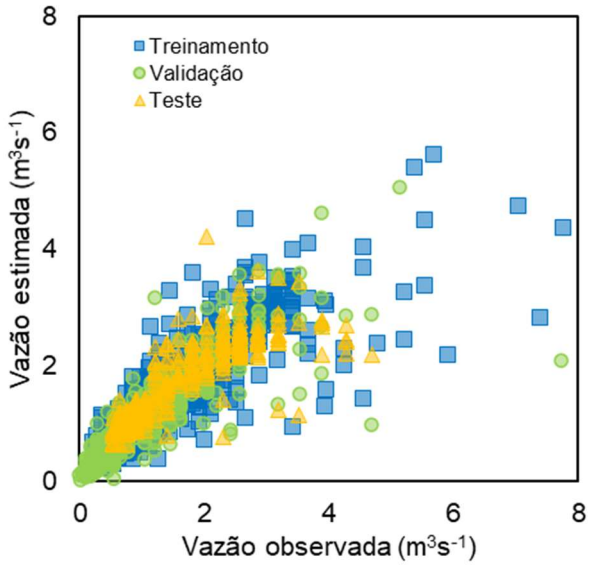

Figura 18: Dispersão entre os dados observados e simulados para a estação Q1

A função de autocorrelação do erro descreve a forma como os erros de previsão estão relacionados no tempo. A Figura 19 apresenta os resultados da função de autocorrelação de erro para as simulações de vazões em Q1. É verificado que há presença de autocorrelações significativas nos erros, inseridos no nível de significância de α=5%, com correlação linear do erro defasado de primeira ordem de

0,32. O erro apresenta média 0,06 m^3s^{-1} e desvio padrão de 0,49 m^3s^{-1}. Isto pode estar relacionado a variabilidade

Isto pode estar relacionado com a alteração da estrutura de autorregressão, que pode ser compreendida através das diferentes funções de autocorrelação, ao se utilizar dimensões da série de vazões (Figura 15). Modelos puramente autorregressivos usam a prerrogativa da estacionariedade das séries, o que dificilmente ocorre nas séries hidrológicas em longos períodos, uma vez que mudanças climáticas ou de uso e cobertura do solo podem alterar tais características, ou seja, pode haver mudanças na estrutura de autorregressão, média e variância das vazões ao longo do tempo.

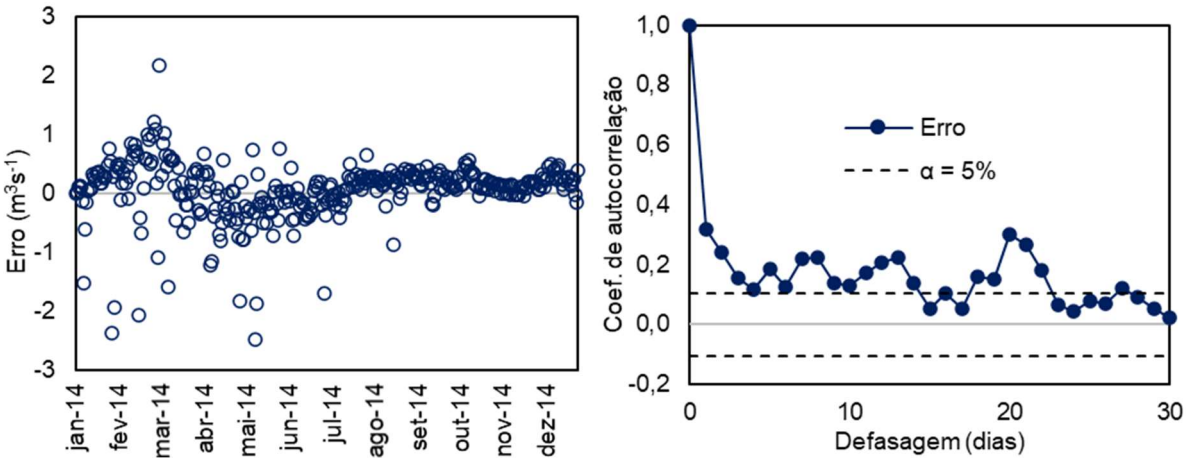

Figura 19: Série temporal e autocorrelação do erro em Q1

A Figura 20 exibe as curvas de permanência observadas e simuladas para a estação Q1. Os índices estatísticos mostram simulações muito boas, com índices NSE de 0,894, RSR de 0,32, RMSE de 0,29 m^3s^{-1}, PBIAS de -8,72%, e MAPE de 9,61%. As vazões de referência Q_{90} e Q_{95}, observadas com 0,64 m^3s^{-1} e 0,64 m^3s^{-1}, respectivamente, foram simuladas com 0,88 m^3s^{-1} e 0,80 m^3s^{-1}. Entretanto, as vazões acima de Q_{20} sofrem com maiores erros percentuais.

Os resultados insatisfatórios nas simulações de cheias é consequência da discretização de monitoramento para a pequena dimensão da bacia hidrográfica, uma vez que o tempo de concentração é pequeno (horas) para a discretização diária dos dados de simulação (SOARES et al. 2010). Em termos físicos, após a saturação da bacia hidrográfica através da infiltração de água no solo, toda a chuva é convertida em escoamento superficial. Assim, após poucas horas de precipitações, no caso de pequenas bacias, há grande variabilidade das vazões que podem não ser evidenciadas no comportamento médio diário, gerando incertezas para os

métodos estimadores. No entanto, as vazões mínimas são melhor simuladas, o que pode ser reflexo da água estocada nos solos durante os eventos de chuva, que são as principais fontes de manutenção dos regimes de vazão nos períodos em que não são observadas chuvas no local.

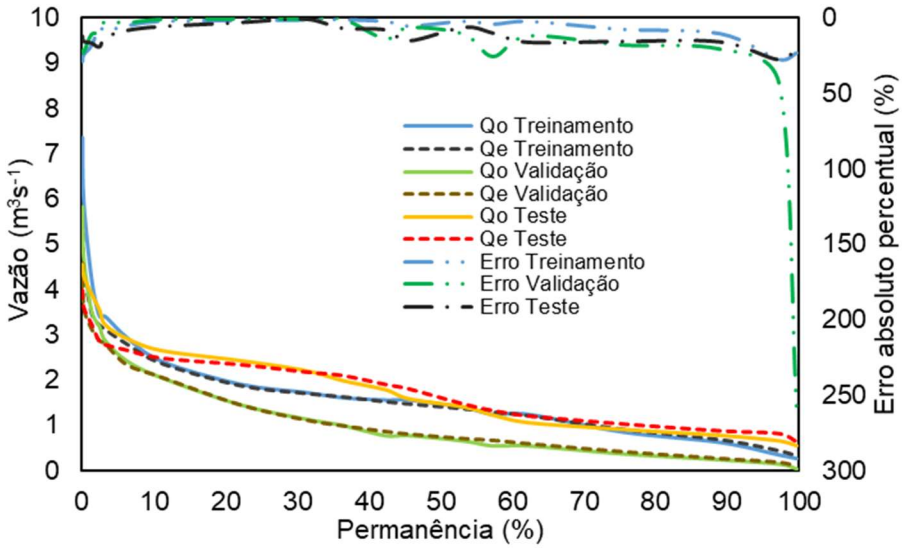

Figura 20: Curvas de permanência observadas e simuladas para a estação Q1

O comportamento do RMSE ao longo do treinamento supervisionado do modelo $P1_{0:14}Q1_{1:6}$ pode ser visualizado na Figura 21. O treinamento teve 213 iterações, no entanto a iteração 6 encontrou os melhores hiperparâmetros da RNN-NARX. Os RMSEs de treinamento, validação cruzada e teste foram de 0,42 m^3s^{-1}, 0,38 m^3s^{-1} e 0,45 m^3s^{-1}, respectivamente.

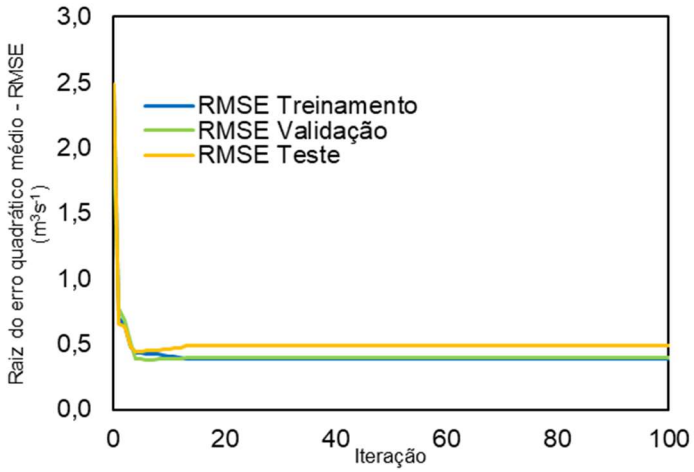

Figura 21: Raiz do erro quadrático médio em função das iterações do modelo para simulação de vazões da estação Q1

Observando a contribuição relativa de cada vetor (Figura 22), nota-se que os vetores de entrada obtiveram semelhantes relevâncias na simulação. Entretanto, os mais influentes são os que apresentam defasagens acima de dois dias. As

precipitações ocorridas em dois dias defasados obtiveram 6,83% de contribuição, seguido das vazões defasadas em três dias, com 5,65%, e das vazões defasadas em dois dias, com 5,50%. Resultados semelhantes foram alcançados por Blanco et al. (2007) e Blanco et al. (2005) para esta mesma bacia, ao identificar que as precipitações defasadas em 3 dias são as respostas impulsionais que apresentam menores erros em simulações lineares. Este resultado reforça a ideia de que as RNN-NARX obtiveram a sensibilidade às influências dos escoamentos subsuperficial e subterrâneo, que são mais lentos e de resposta gradual na saída da bacia, a partir da implementação das variáveis defasadas.

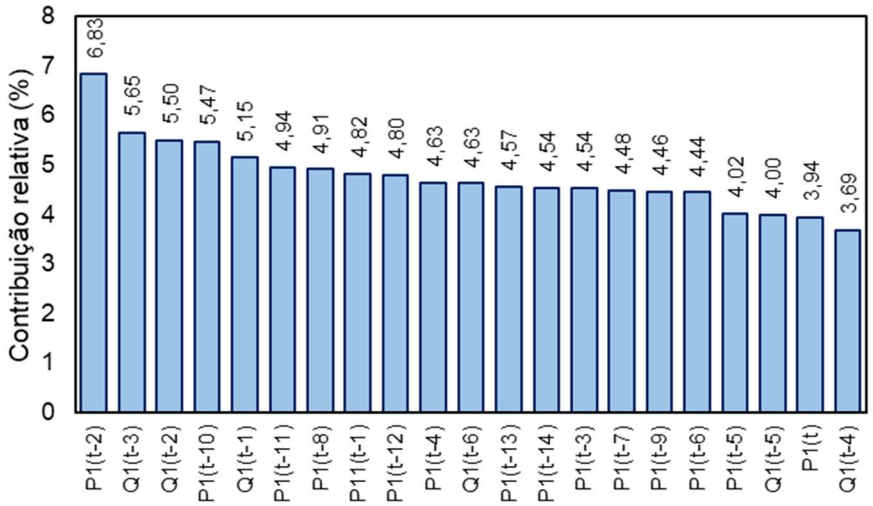

Figura 22: Contribuição relativa dos dados de entrada do modelo para simulação de vazões da estação Q1

A análise de sensibilidade do modelo $P1_{0:14}Q1_{1:6}$ diante de diferentes intervalos de treinamento é apresentada na Tabela 6. De modo geral, os índices estatísticos apontam boas simulações utilizando os anos de 2009-2011 e 2010-2011. Somente com o ano de 2011, porém, a simulação é classificada como satisfatória. Houve um discreto aumento nos índices estatísticos ao utilizar somente os anos 2010-2011 para treinamento das RNN-NARX, com índices NSE de 0,675, RSR 0,57, RMSE de 0,51 m^3s^{-1}, PBIAS de -3,81% e MAPE de 20,68%. Este fato pode estar relacionado ao fenômeno El Niño ocorrido em 2009 (INPE, 2021), pois as correlações cruzadas entre Q1 e P1 são menores, como evidenciado anteriormente. Logo, a utilização de dois anos hidrológicos apresentou melhores resultados para simular o hidrograma observado em 2014, na bacia de Q1. A literatura comenta que modelos empíricos têm melhores desempenhos com o maior número de observações disponíveis para treinamento (WUNSCH et al., 2021), porém, a questão

qualitativa dos hidrogramas implementados pode ser tão influente quanto, pois a simulação de hidrogramas com eventos extremos, por exemplo, pode ser melhorada a partir da implementação de observações semelhantes no momento de treinamento.

Tabela 6: Sensibilidade do modelo a diferentes resoluções temporais para simulações de Q1

Período	NO	NSE	RSR	RMSE	PBIAS	MAPE
2009-2011	3	0,658	0,58	0,52	-8,72	22,35
2010-2011	3	0,675	0,57	0,51	-3,81	20,68
2011	3	0,553	0,67	0,6	7,32	39,33

NO-Neurônios Ocultos; NSE-Coeficiente de Eficiência de Nash-Sutcliffe; RSR-Raiz do erro quadrático médio padronizado; RMSE-Raiz do erro quadrático médio; PBIAS-Viés percentual; MAPE-Erro absoluto percentual médio

As Figuras 23 e 24 mostram os hidrogramas e curvas de permanência simuladas para Q_1, respectivamente, partir dos diferentes intervalos de treinamento. Nota-se que a sazonalidade dos hidrogramas e a forma das curvas de permanência simuladas são semelhantes ao observado, e entre si. Porém, o alinhamento a longo prazo é comprometido, principalmente ao utilizar somente o ano de 2011. Isto está associado à transferência contínua do erro, uma vez que a simulação autorregressiva de longo período se apropria das vazões simuladas para auxiliar a descrever o próximo passo de regressão.

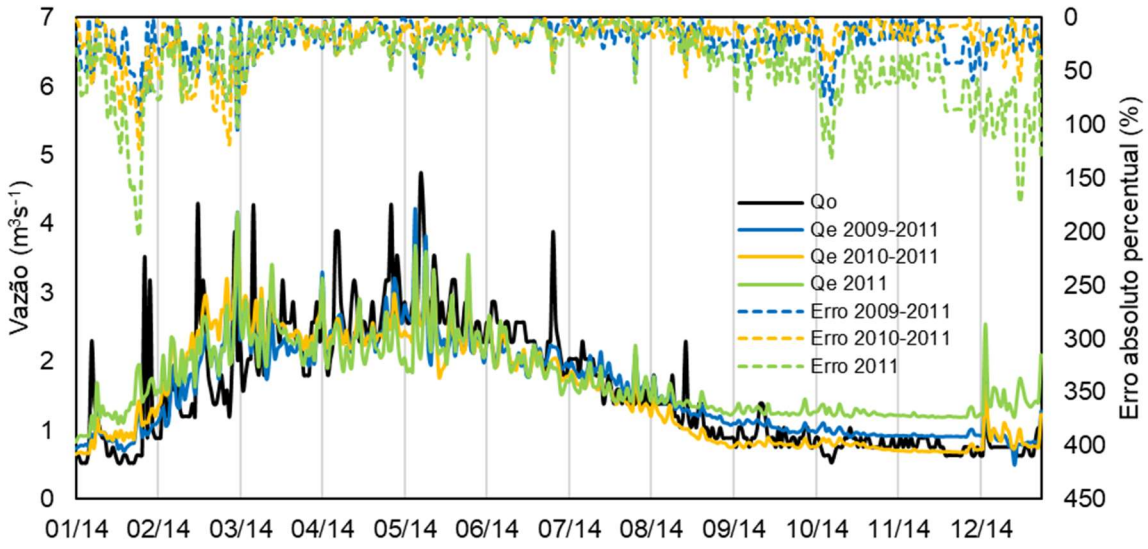

Figura 23: Hidrogramas observado e simulados para Q1, utilizando diferentes intervalos temporais

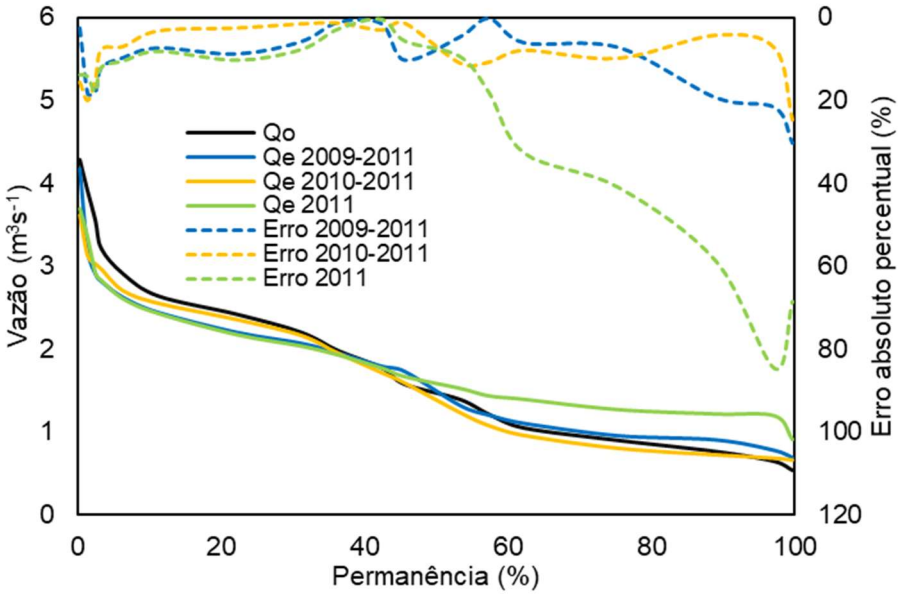

Figura 24: Curvas de permanência observada e simuladas para Q1, utilizando diferentes intervalos temporais

4.2 RIO PIRANHAS

Localizada no município de Abreulândia, no Estado do Tocantins, a estação fluviométrica Fazenda Craveiro (Q2) (Cód. 27370000) monitora a variabilidade das vazões diárias de uma sub-bacia do rio Piranhas, cuja área de contribuição de aproximadamente 186 km^2 (ANA, 2021). As altitudes mínima, máxima e média desta bacia hidrográfica são de 228 m, 679 m e 393 m, respectivamente. Conforme a classificação proposta por Embrapa (1999), a declividade da bacia é considerada suave ondulada, uma vez que a média é de 11,91%. A Figura 25 mostra as principais características desta área de contribuição.

As classes de solos presentes são o Neossolo Quartzarênico Órtico e, predominantemente, o Plintossolo Pétrico Concrecionário (IBGE, 2021). Cordeiro et al. (2021) constataram que a capacidade de infiltração de água do Plintossolo Pétrico Concrecionário está na faixa de alta a muito alta, em uma área de pastagem degradada no município de Brejinho de Nazaré, local próximo à bacia de Q2. A área tem grandes percentuais de vegetação natural florestal primária (61,51%), que podem ser descritas como estrato arbóreo-arbustivo com altura até 18m. E também, 30,34% da área é destinada a pastagem, compostas com vegetação forrageira herbácea (INPE, 2018).

A estação pluviométrica Abreulândia (cód. 949000) (P2) está situada há cerca de 20,4 km de distância de Q2. Neste local, o regime intra-anual das chuvas é

definido pelo período chuvoso, que se estende entre os meses de novembro a abril, e pelo período de estiagem, compreendido entre os meses de maio a outubro, como mostra a normal climatológica observada em P2 (Figura 26). O clima característico da região é o Aw, conforme a classificação de Köppen-Geiger, logo, é um clima tropical de savana, pois o mês mais seco tem precipitação inferior a 60 mm, além das temperaturas sempre acima dos 18°C (ALVARES, 2013; ROLDÃO E FERREIRA, 2019).

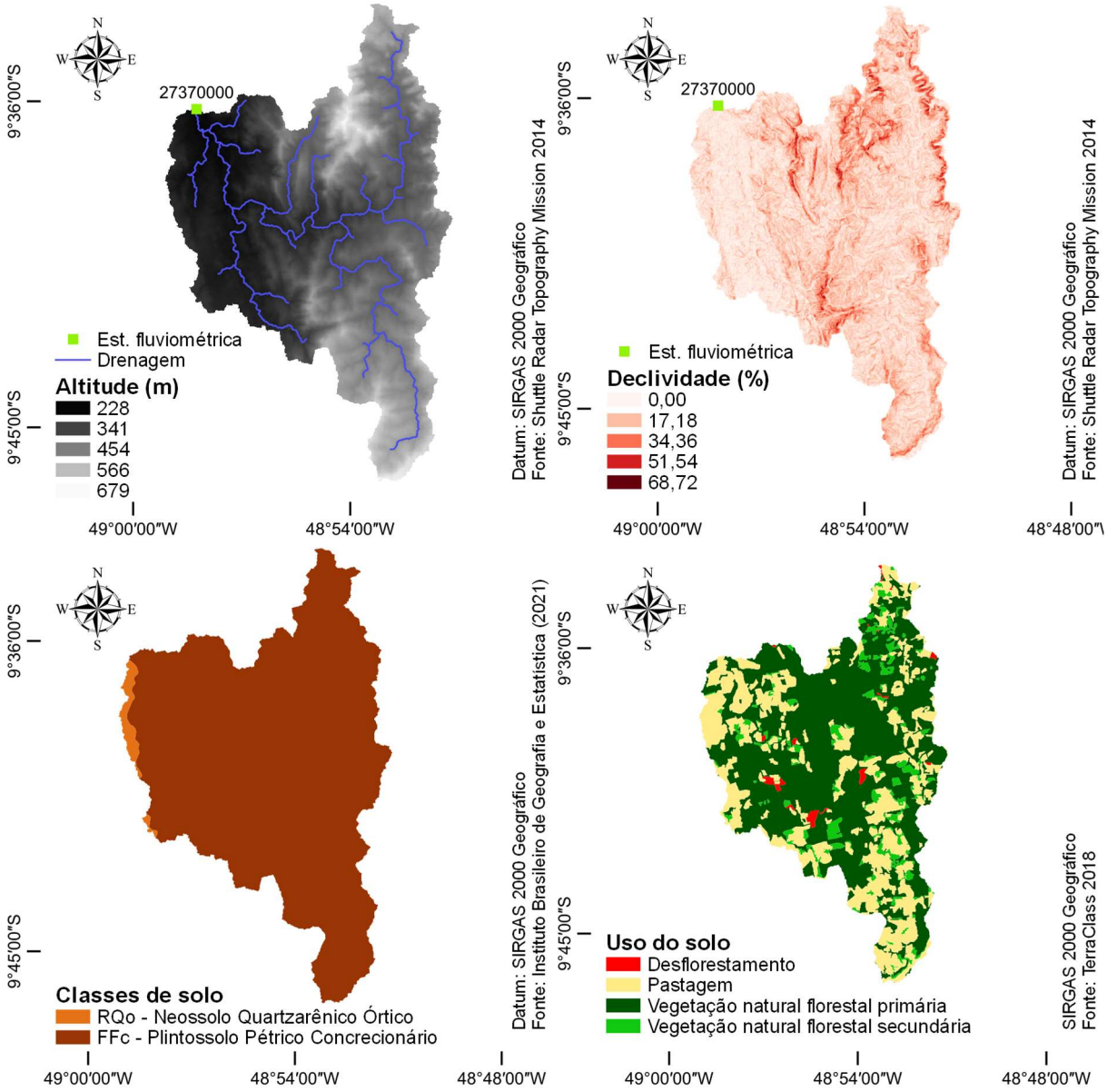

Figura 25: Caracterização da bacia da estação Q2

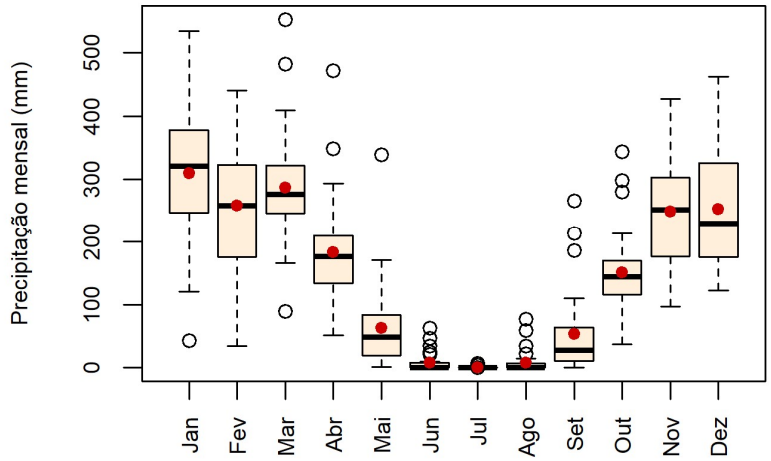

Figura 26: Normal climatológica da estação P2 entre 1991-2020

A resolução temporal utilizada para produzir as simulações de Q2 é compreendida entre os anos de 2009-2014. Logo, o conjunto de treinamento é composto pelos anos de 2009-2011; o de validação é definido pelos anos de 2012-2013; e o de teste é definido pelo ano de 2014. Esta divisão dos dados pode ser visualizada na Figura 27, sob a forma de hidrogramas de vazões.

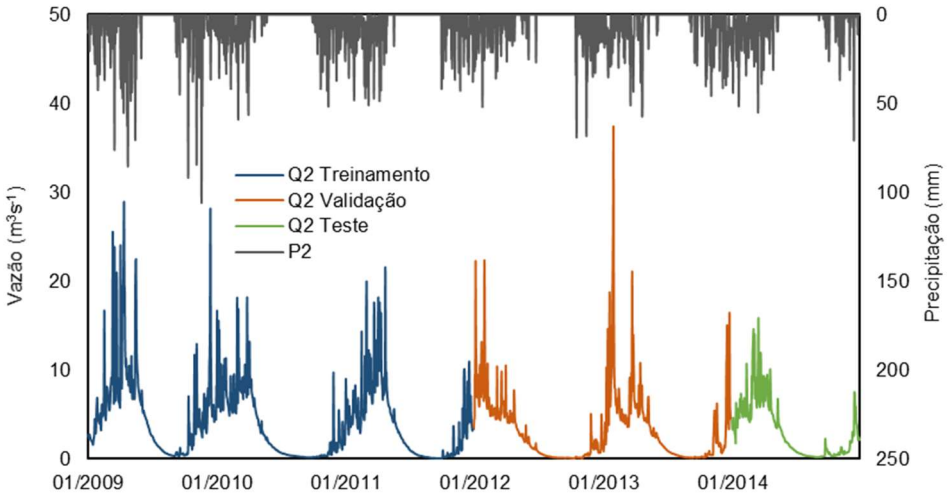

Figura 27: Hidrogramas de vazões da estação Q2 entre 2009-2014

A Tabela 7 mostra a estatística descritiva dos dados utilizados para simulação de vazões da estação Q2. O menor e o maior índice de chuva anual são observados nos anos de 2014 e 2009, respectivamente, com 1.508 mm e 2.557 mm. No entanto, o menor e o maior índice de vazões são observados nos anos de 2012 e 2014, respectivamente, com média de 2,82 m^3s^{-1} e 5,31 m^3s^{-1}. Assim como descrito no capítulo anterior, esta diferença pode estar associada a anomalias meteorológicas, que alteram padrões espaciais de distribuição da precipitação (INPE, 2021; ESPINOZA et al., 2021).

Tabela 7: Estatística descritiva das variáveis P2 e Q2

Variável hidrológica	Estatística	Treinamento			Validação		Teste
		2009	2010	2011	2012	2013	2014
P2	Média[1] (mm/dia)	7,01	4,43	5,13	4,36	5,59	4,13
	Anual (mm/ano)	2.557,20	1.618,30	1.872,50	1.597,00	2.040,80	1.508,00
	Máx (mm/dia)	105,50	58,90	50,80	68,90	56,50	70,60
	Min[2] (mm/dia)	0,60	0,60	0,60	1,90	1,50	0,30
	Coef. de variação (%)	215,48	211,17	199,38	215,56	184,00	223,30
Q2	Média (m³/s)	5,31	3,25	3,81	2,82	3,44	3,04
	Máx (m³/s)	28,86	18,17	21,60	22,20	37,23	15,81
	Min (m³/s)	0,26	0,15	0,04	0,08	0,14	0,16
	Coef. de variação (%)	96,26	113,44	106,07	118,61	123,53	101,42
	$Q_{50\%}$ (m³/s)	4,37	1,48	2,90	1,41	2,08	1,87
	$Q_{95\%}$ (m³/s)	0,38	0,17	0,07	0,12	0,18	0,20

[1]Considerando dias sem chuva
[2]Acima de zero

A Figura 28 mostra as funções de correlação cruzada entre Q2 e P2, em diferentes resoluções temporais. Considerando os anos entre 2009-2011, e um nível de significância de 5%, destacam-se as precipitações diárias defasadas entre 0-30 dias. Considerando os anos entre 2010-2011, e um nível de significância de 5%, destacam-se as precipitações diárias defasadas entre 0-29 dias. E, considerando o ano de 2011, e o nível de significância de 5% (α = 0,05), destacam-se as precipitações diárias defasadas entre 0-23 dias, 25-26 dias. Todas as correlações analisadas são consideradas fracas (PECK et al., 2015). Assim como evidenciado na bacia hidrográfica do capítulo anterior, a função de correlação cruzada tem uma diminuição nos anos de 2009-2011, e aumento com os anos de 2010-2011, podendo ser atribuído aos fenômenos ENOS, respectivamente, El Niño moderado e La Niña moderado (INPE, 2021).

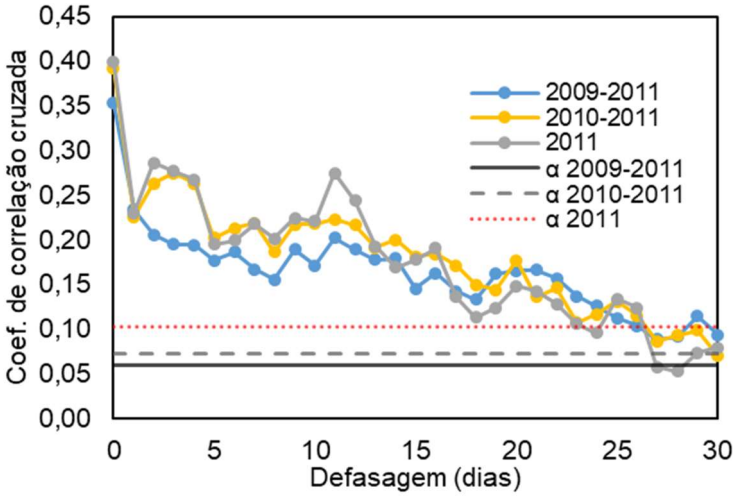

Figura 28: Função de correlação cruzada entre as estações P2 e Q2

As Figuras 29 e 30 mostram as funções de autocorrelação e autocorrelação parcial de Q2, respectivamente. Considerando os anos entre 2009 e 2011, e um

nível de significância de 5%, destacam-se as vazões diárias defasadas entre 1-5 dias, 9 e 10 dias, 12 dias, 14 dias, 17 dias, e 19 dias. Considerando os anos entre 2010 e 2011, e o nível de significância de 5% (α = 0,05), destacam-se as vazões diárias entre 1-4 dias, 6 e 9 dias, 12 e 14 dias, 17 dias, 19 dias, 23-24 dias defasados. E, considerando o ano de 2011, e um nível de significância de 5%, destacam-se as vazões diárias entre 0-1 dia, 3-4 dias, 9 dias, 12 dias, 14 dias, 17 dias e 24 dias defasados. Somente a vazão defasada em 1 dia apresenta correlação forte, e as demais são consideradas fracas (PECK et al., 2015). No ano de El Niño moderado (2009-2010), há a diminuição dos coeficientes em toda função de autocorrelação. Além disso, no ano de La Niña moderado (2010-2011), há o aumento de tais coeficientes (INPE, 2021). Este efeito é o inverso ao ocorrido na bacia de Q1, o que pode ser atribuído às características da bacia, uma vez que são solos diferentes e a declividade média da bacia de Q2 é 3,5 vezes maior que em Q1.

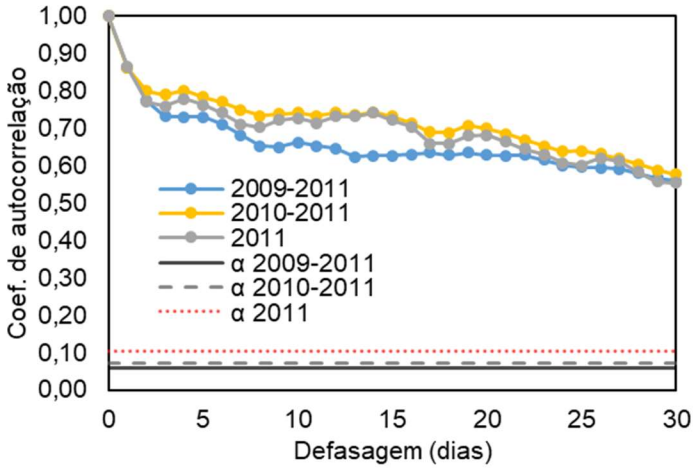

Figura 29: Função de autocorrelação da estação Q2

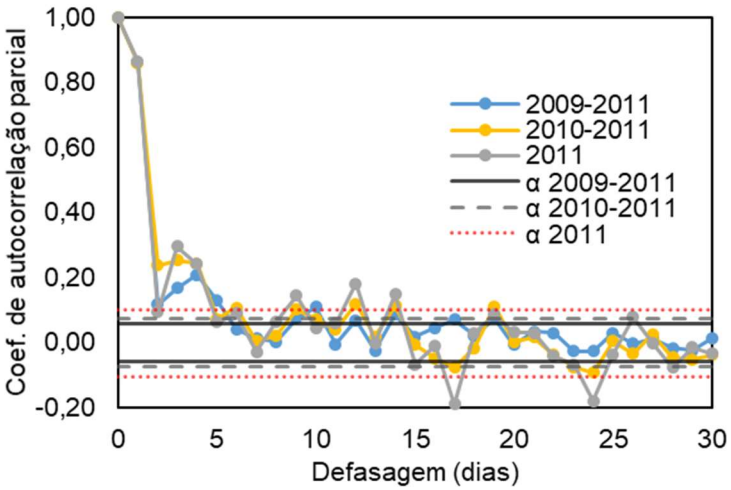

Figura 30: Função de autocorrelação parcial da estação Q2

Com base nos gráficos anteriores, foram propostos seis diferentes vetores de entrada para simulação de Q2. A Tabela 8 mostra tais vetores, bem como o desempenho estatístico de cada. Constata-se que alguns modelos alcançam desempenho satisfatório. Os melhores desempenhos foram evidenciados com as precipitações defasadas entre 0-19 dias e vazões defasadas em entre 1-12 dias ($P2_{0:19}Q2_{1:12}$). Este modelo obteve índices de desempenho NSE de 0,590, RSR de 0,64, PBIAS de -4,03%, RMSE de 1,97 m^3s^{-1} e MAPE de 77,76%, classificado com satisfatório, segundo os índices propostos por Moriasi et al. (2007).

Tabela 8: Desempenhos dos modelos propostos para simulação de Q2

Modelo	Série-Paralelo					
	NO	NSE	RSR	RMSE	PBIAS	MAPE
$P2_{0:11}Q2_{1:5}$	3	0,484	0,72	2,21	-30,00	110,28
$P2_{0:11}Q2_{1:5}$	3	0,470	0,73	2,24	32,46	112,15
$P2_{0:23}Q2_{1:5}$	3	0,458	0,73	2,27	35,72	69,89
$P2_{0:19}Q2_{1:12}$	3	0,590	0,64	1,97	-4,03	77,76
$P2_{0:23}Q2_{1:19}$	4	0,592	0,64	1,97	-2,64	62,07
$P2_{0:28}Q2_{1:19}$	4	0,558	0,66	2,05	-4,29	46,58

NO – Neurônios Ocultos; NSE – Coeficiente de Eficiência de Nash-Sutcliffe; RSR – Raiz do erro quadrático médio padronizado; PBIAS – Viés percentual; MAPE – Erro absoluto percentual médio

As Figuras 31 e 32 mostram as vazões observadas e simuladas pela RNN-NARX $Q2_{0:19}P2_{1:12}$ em série-paralelo sob a forma de hidrograma de vazões e gráfico de dispersão, respectivamente. Os maiores erros absolutos percentuais são evidenciados nos meses de outubro (MAPE de 153%) e novembro (173%), meses iniciais do período chuvoso local. O maior erro absoluto percentual corresponde à vazão observada com 0,438 m^3s^{-1}, no dia 04/11/2014, que foi simulada com 3,219 m^3s^{-1} (superestimação de 635%). Outro exemplo é a vazão observada no dia 27/10/2014, com 0,472 m^3s^{-1}, que foi simulada com 2,759 m^3s^{-1} (superestimação de 484%). Em contrapartida, os menores erros percentuais estão nos meses de abril (MAPE de 22,94%) e julho (MAPE de 15,59%), correspondentes ao início do período de estiagem. O menor erro absoluto percentual foi de 0,10%, com a vazão observada de 3,628 m^3s^{-1}, no dia 17/05/2014¸ que foi simulada com 3,624 m^3s^{-1}.

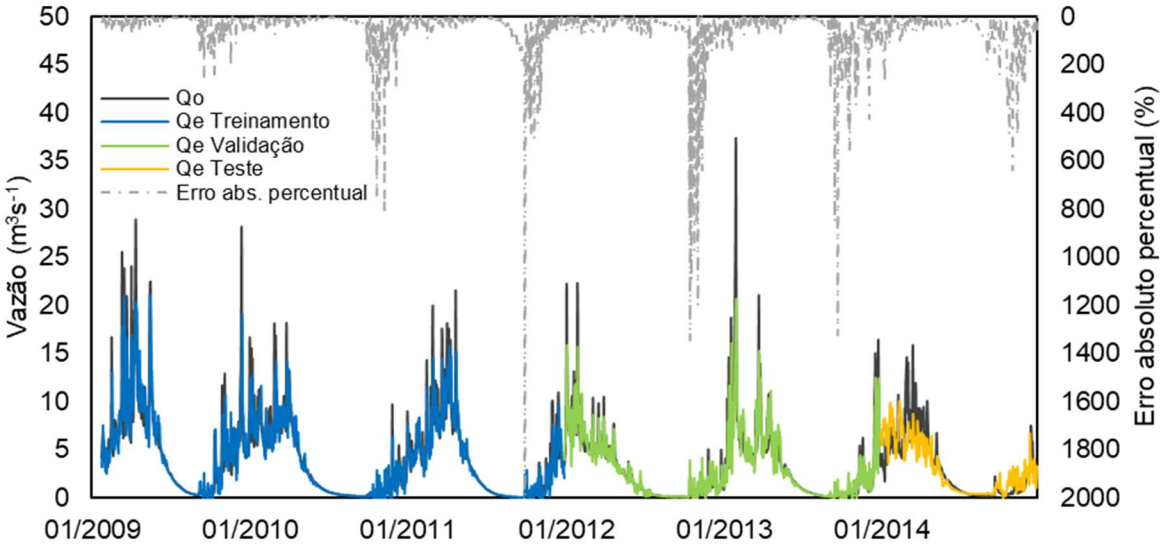

Figura 31: Hidrogramas observados e simulados para a estação Q2

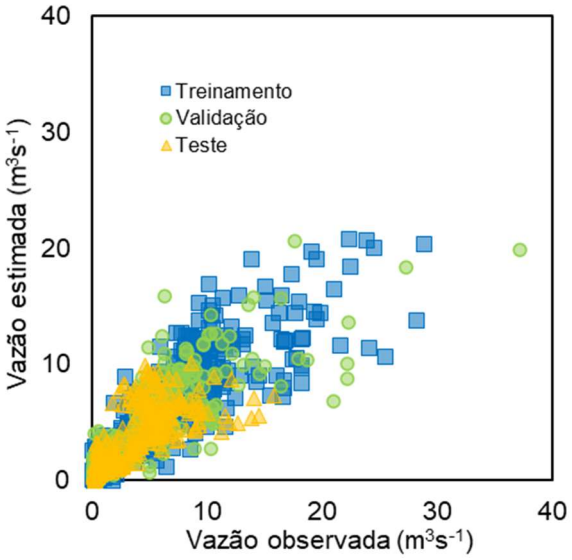

Figura 32: Dispersão entre os dados observados e simulados para a estação Q2

A Figura 33 apresenta os resultados da função de autocorrelação de erro para as simulações de vazões em Q2. É verificado que há presença de autocorrelações significativas nos erros, inseridos no nível de significância de α=5%, com correlação linear do erro defasado de primeira ordem de 0,74. O erro tem média de - 0,1 m^3s^{-1} e desvio padrão de 1,97 m^3s^{-1}.

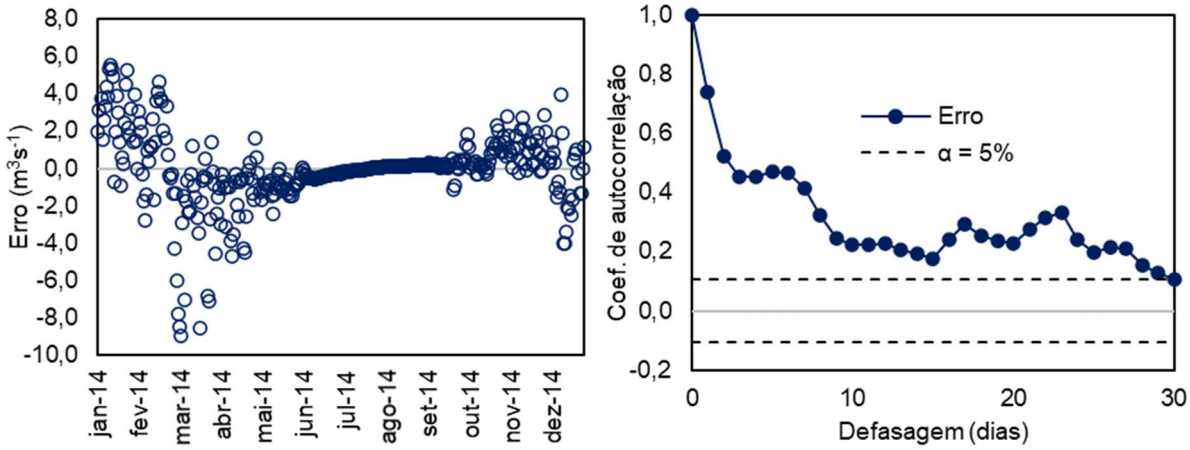

Figura 33: Série temporal e função de autocorrelação do erro em Q2

A Figura 34 exibe as curvas de permanência observadas e simuladas para a estação Q1. Os índices estatísticos mostram simulações muito boas, com índices NSE de 0,945, RSR de 0,23, RMSE de 0,72 m^3s^{-1}, PBIAS de -4,03%, e MAPE de 37,22%. Nota-se que a curva de permanência simulada tem comportamento semelhante até às vazões de referência Q_{15}, e acima disso há tendência de subestimação. As vazões de referência Q_{90} e Q_{95}, observadas com 0,28 m^3s^{-1} e 0,20 m^3s^{-1}, respectivamente, foram simuladas com 0,37 m^3s^{-1} e 0,35 m^3s^{-1}.

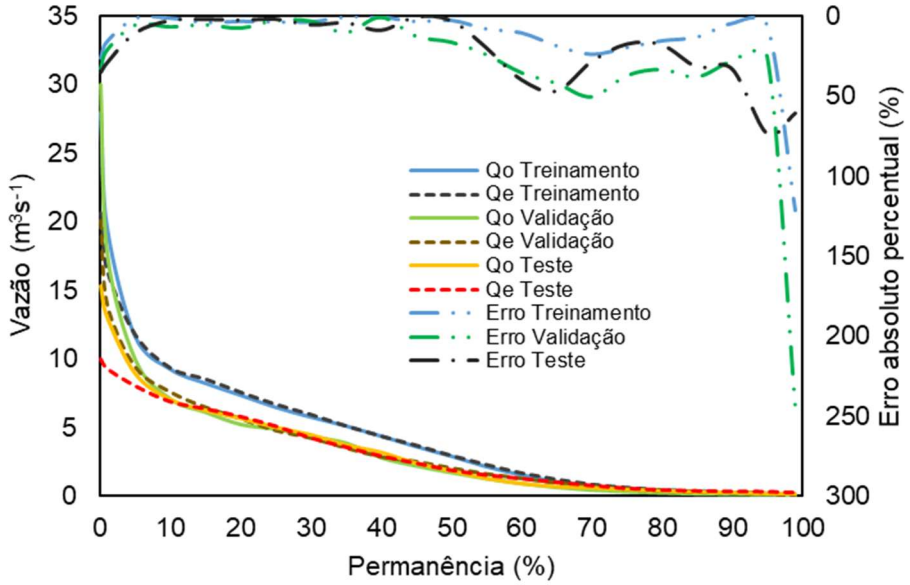

Figura 34: Curvas de permanência observadas e simuladas para a estação Q2

A Figura 35 exibe o RMSE em função do treinamento supervisionado do modelo $P2_{0:19}Q2_{1:12}$. O treinamento teve 213 iterações, no entanto a iteração 13 encontrou os pesos para a RNN-NARX. O treinamento foi finalizado pelo aumento contínuo do RMSE de validação cruzada. Os RMSEs de treinamento, validação cruzada e teste foram de 1,88 m^3s^{-1}, 1,80 m^3s^{-1} e 1,97 m^3s^{-1}, respectivamente.

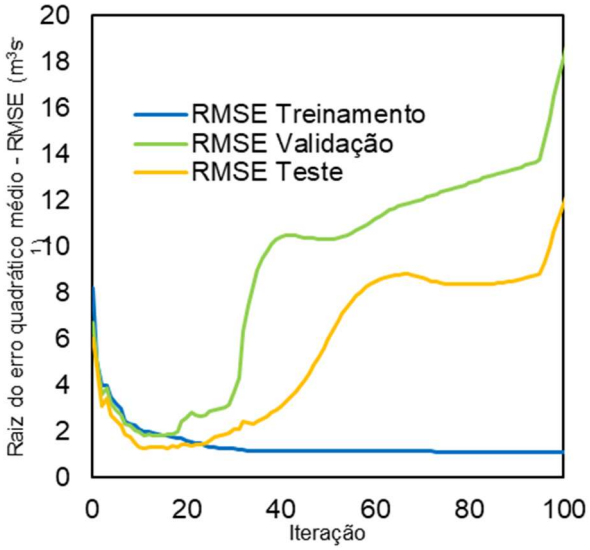

Figura 35: Raiz do erro quadrático médio em função das iterações do modelo para simulação de vazões da estação Q2

A soma das contribuições relativas das precipitações defasadas é de 52,50%, e das vazões defasadas foi de 47,50%. Como mostra a Figura 36, as duas variáveis com maiores percentuais são as vazões defasadas em 10 dias (9,10%) e 1 dia (8,04%). E, as precipitações defasadas com maiores percentuais são as ocorridas em 0 dia (7,11%) e 13 dias (6,71%). Debastiani et al. 2016 comentam que a dependência excessiva de vazões defasadas pode comprometer as simulações de longo período, uma vez que a forçante do sistema é a precipitação.

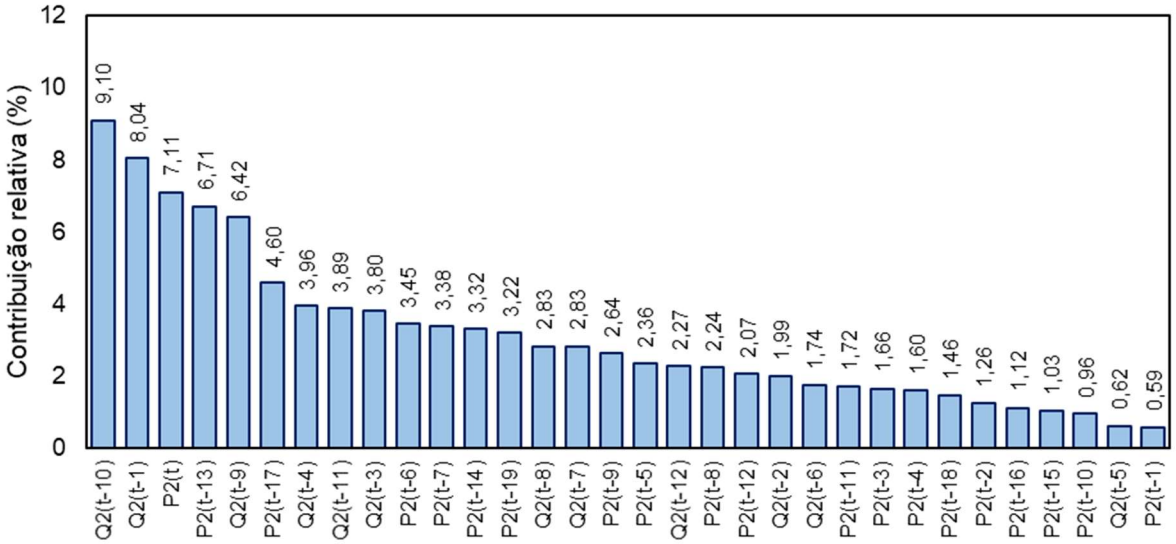

Figura 36: Contribuição relativa dos dados de entrada do modelo para simulação de vazões da estação Q2

A Tabela 9 apresenta a análise de sensibilidade do modelo $P2_{0:19}Q2_{1:12}$ a diferentes intervalos de treinamento. Os índices estatísticos apontam simulações

satisfatórias utilizando os anos de 2009-2011 e 2010-2011. Somente com o ano de 2011, porém, a simulação é classificada como insatisfatória (MORIASI et al., 2007). Assim, utilizando somente os anos de 2010-2011, houve a redução de apenas -0,03% no NSE, e aumento da subestimação do PBIAS em 0,38%, do MAPE em 13,52%, 0,03 no RSR, e 0,08 m^3s^{-1} no RMSE. Diante disso, é verificado que a utilização de dois anos hidrológicos são suficientes para apresentar resultados satisfatórios.

As Figuras 37 e 38 mostram os hidrogramas de vazões e curvas de permanência simuladas pelos diferentes vetores de sensibilidade. Com o ano de 2011, houve a tendência de superestimar os eventos fluviométricos, tanto que as vazões acima da vazão de referência Q_{70} estão mais distantes do observado. No entanto, os modelos ajustam-se melhor as vazões menores que a Q_{70}.

Tabela 9: Sensibilidade do modelo à diferentes resoluções temporais para simulações de Q2

Período	NO	NSE	RSR	RMSE	PBIAS	MAPE
2009-2011	4	0,592	0,64	1,97	-2,64	62,07
2010-2011	4	0,555	0,67	2,05	-3,02	75,59
2011	4	0,370	0,79	2,45	33,62	76,76

NO – Neurônios Ocultos; NSE – Coeficiente de Eficiência de Nash-Sutcliffe; RSR – Raiz do erro quadrático médio padronizado; PBias – Viés percentual; MAPE – Erro absoluto percentual médio

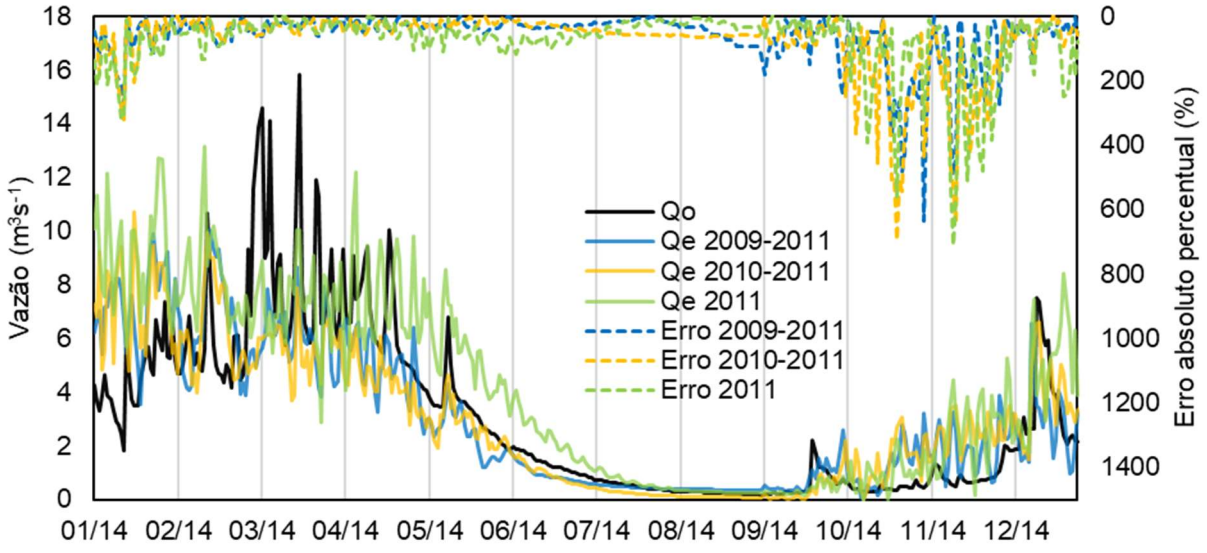

Figura 37: Hidrogramas observado e simulados para Q2, utilizando diferentes intervalos temporais

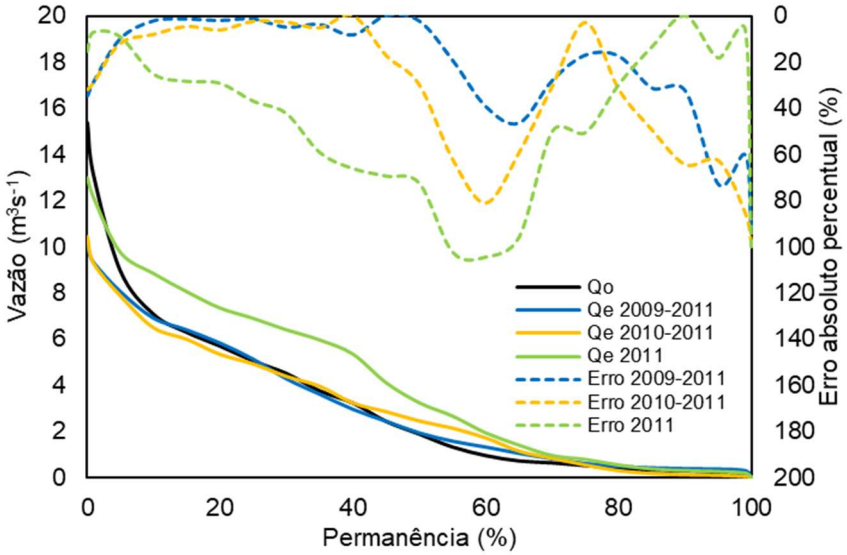

Figura 38: Curvas de permanência observada e simuladas para Q2, utilizando diferentes intervalos temporais

4.3 RIO CAETÉ

A estação fluviométrica Arraial (Q_3) (cód. 32300000) monitorou as vazões diárias do rio Caeté durante os anos de 1965 a 1972, em um trecho situado no município de Ourém (Figura 39). A área de contribuição desta sub-bacia é de 290 km^2 (ANA, 2021). As altitudes mínima, máxima e média são de 25 m, 106 m e 63 m, respectivamente. A declividade média é de 4,36%, e a máxima é de 28,39%, logo, a bacia é classificada como suave ondulada (EMBRAPA, 1999, USGS, 2021).

A classe de solo encontrada na bacia de Q_3 é o latossolo amarelo distrófico (IBGE, 2021). E, as classes de uso e ocupação do solo são a vegetação natural secundária (34,25%), a pastagem (30,09%) e vegetação natural primária (11,41%) (INPE, 2014). Além disso, a Vila Arraial do Caeté é evidenciada no interior desta bacia hidrográfica, próxima à rodovia PA-112, sob latitude 1°16'44.5"S e longitude 46°52'48.9"W. Diante dessas características, trata-se de uma bacia hidrográfica com baixa variabilidade de altitude, baixas declividades, moderado percentual de vegetação natural, o que pode favorecer os escoamentos subsuperficiais e subterrâneos.

A estação pluviométrica Arraial (P_3) (cód. 147000) está localizada a 25,8 km da estação Q3, na direção sudeste. Assim como relatado anteriormente na estação P1, o clima deste local é definido como Am, segundo a classificação Köppen-Geiger (KÖPPEN, 1936; GEIGER, 1954), portanto, é um clima tropical com regime de

monções, altos índices pluviométricos, e baixa variabilidade na temperatura, que sempre está acima dos 18°C (ALVARES et al., 2013). A variabilidade intra-anual da precipitação na bacia Q3 é caracterizada por dois períodos distintos (MORAES et al., 2005). O período chuvoso é compreendido entre os meses de janeiro e maio, e o período de estiagem é compreendido entre os meses de junho e dezembro, conforme a normal climatológica observada na estação P1 (Figura 40).

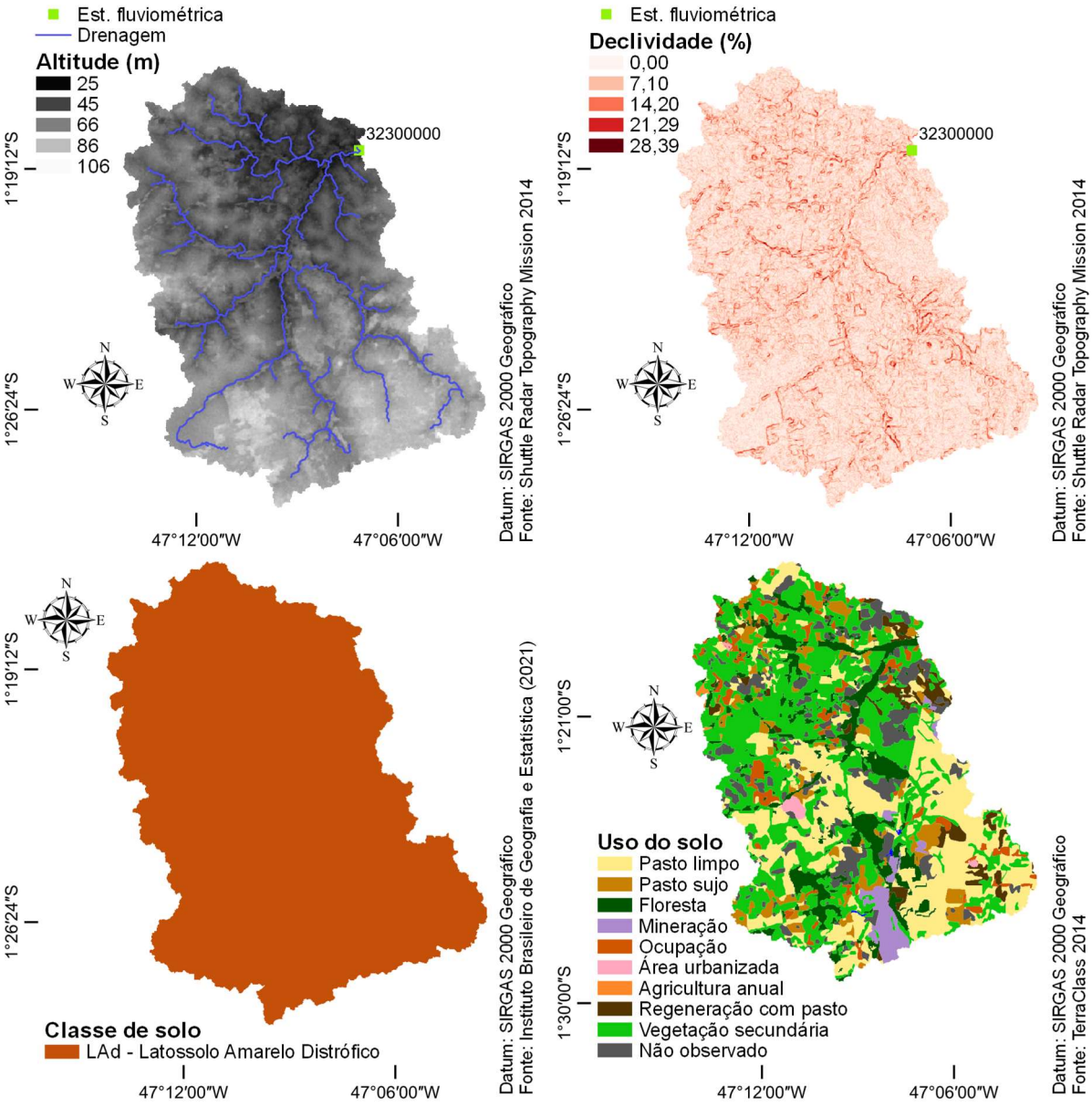

Figura 39: Caracterização da bacia da estação Q3

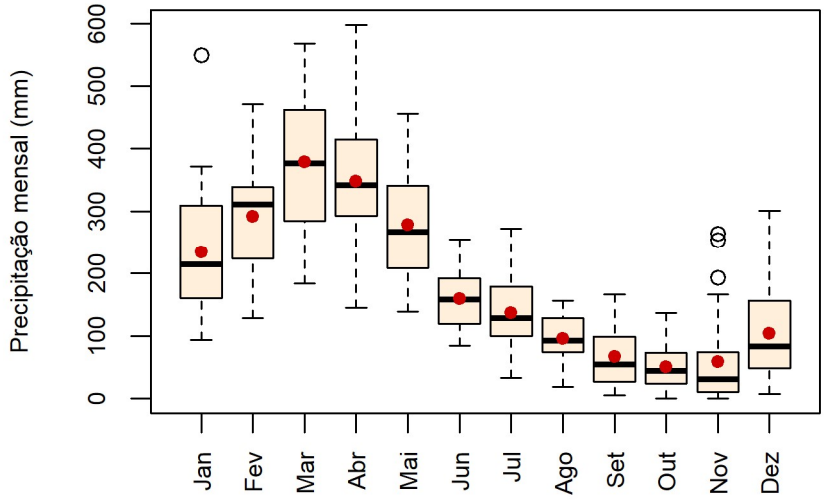

Figura 40: Normal climatológica da estação P3 entre 1991-2020

A Figura 41 mostra a divisão dos três conjuntos de dados para simulação de vazões em Q_3, sob a forma de hidrogramas. Os anos entre 1966-1968 definem o conjunto de treinamento, utilizado para treinar as RNN-NARX e análise exploratória. Os anos de 1969-1970 definem o conjunto de validação. E, o ano de 1971 é destinado à avaliação estatística do modelo.

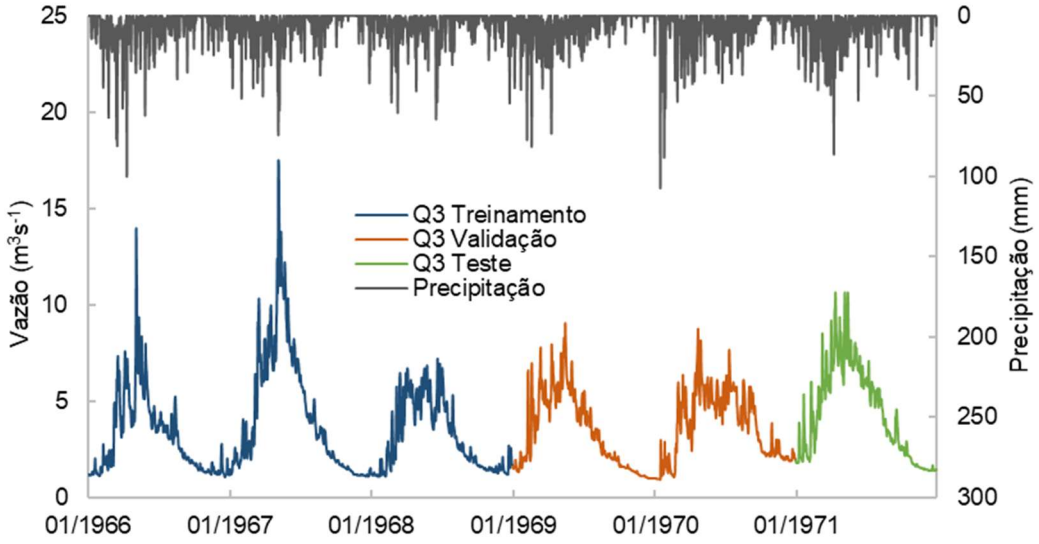

Figura 41: Hidrogramas de vazões da estação Q3 entre 1966-1971

A estatística descritiva das variáveis Q3 e P3 é mostrada na Tabela 10. O menor e o maior índice pluviométrico anual são encontrados nos anos de 1968 e 1970, respectivamente, com 2.187 mm e 2.633 mm. No entanto, nota-se que a variabilidade interanual dos anos abordados é baixa, com coeficiente de variação de apenas 6,98%. Embora este fato, os anos entre 1965-1971 apresentaram diferentes eventos magnitudes de El Niño e La Niña (INPE, 2021).

Tabela 10: Estatística descritiva das variáveis P3 e Q3

Variável hidrológica	Estatística	Treinamento			Validação		Teste
		1966	1967	1968	1969	1970	1971
P3	Média[1] (mm/dia)	7,11	6,31	5,98	6,61	7,21	6,61
	Anual (mm/ano)	2.593,60	2.303,20	2.187,60	2.411,70	2.633,00	2.413,60
	Máx (mm/dia)	100,20	74,30	64,50	81,60	107,40	86,00
	Min[2] (mm/dia)	0,10	0,10	0,10	0,50	0,60	0,20
	Tempo de retorno (anos)	6,03	2,34	2,23	3,16	7,12	3,18
	Coef. de variação (%)	171,42	169,88	160,79	167,75	168,98	170,24
Q3	Média (m³/s)	3,08	4,34	3,09	3,32	3,71	4,33
	Máx (m³/s)	13,80	17,50	7,19	9,00	8,70	10,60
	Min (m³/s)	1,03	1,07	1,07	1,02	0,94	1,39
	Coef. de variação (%)	65,19	74,68	56,48	57,57	46,19	54,14
	$Q_{50\%}$ (m³/s)	2,36	3,16	2,45	2,83	3,85	4,05
	$Q_{95\%}$ (m³/s)	1,17	1,13	1,13	1,07	1,12	1,44

[1]Considerando dias sem chuva
[2]Acima de zero

A Figura 42 mostra as funções de correlação cruzada entre Q3 e P3, considerando diferentes tamanhos de séries temporais. Considerando os anos entre 1966-1968, e um nível de significância de 5%, destacam-se as precipitações diárias defasadas entre 0-27 dias. Considerando os anos entre 1967-1968, e o nível de significância de 5% (α = 0,05), destacam-se as precipitações diárias defasadas entre 0-29 dias. E, considerando o ano de 1968, destacam-se as precipitações diárias defasadas entre 0-9 dias, 20 dias, 26 e 27 dias. Todas as correlações lineares são consideradas fracas, pois são menores que 0,5 (PECK et al., 2015). Assim como nas bacias anteriores, anos hidrológicos associados à La Niña apresentam maiores coeficientes de correlação linear entre a vazão e as precipitações defasadas, como o caso do La Niña fraco de 1967-1968 (INPE, 2021).

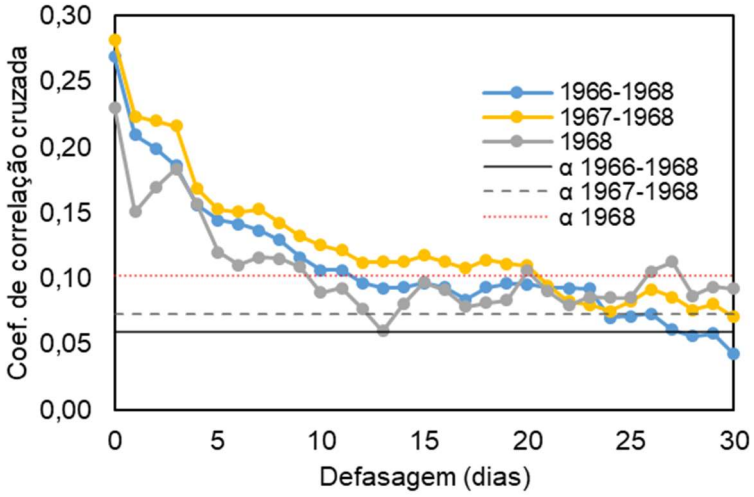

Figura 42: Função de correlação cruzada entre as estações P3 e Q3

As Figuras 43 e 44 mostram as funções de autocorrelação e autocorrelação parcial de Q3, respectivamente. Considerando os anos entre 1967-1969, e um nível

de significância de 5%, destacam-se as vazões diárias de 1 dia, entre 3-6 dias, 9 dias, 12 dias, 14 dias, 23 dias. Considerando os anos entre 1968-1969, e um nível de significância de 5%, destacam-se as vazões diárias entre 1 dia, 3 dias, 5 dias, 9 dias, e 14 dias defasados. E, considerando o ano de 2011, e o nível de significância de 5% (α = 0,05), destacam-se as vazões diárias entre 1 dia, 3 dias, 5 dias, 7 dias defasados. A vazão defasada em 1 dia é forte, e as demais são consideradas fracas (PECK et al., 2015). No entanto, é observado que as vazões defasadas em 2-3 dias têm coeficientes próximos ao limiar de 0,5, o que mostra potencialidades preditivas. A autorregressividade é proveniente de processos estocásticos estacionários, ou seja, a propriedade determinante do fenômeno gerador está presente em toda a série. Para Marengo e Alves (2005) e Mendonça et al. (2021), na hidrologia, isto está diretamente relacionado a maior capacidade de armazenamento de água no solo, ou a estruturas de regularização devido a estruturas hidráulicas.

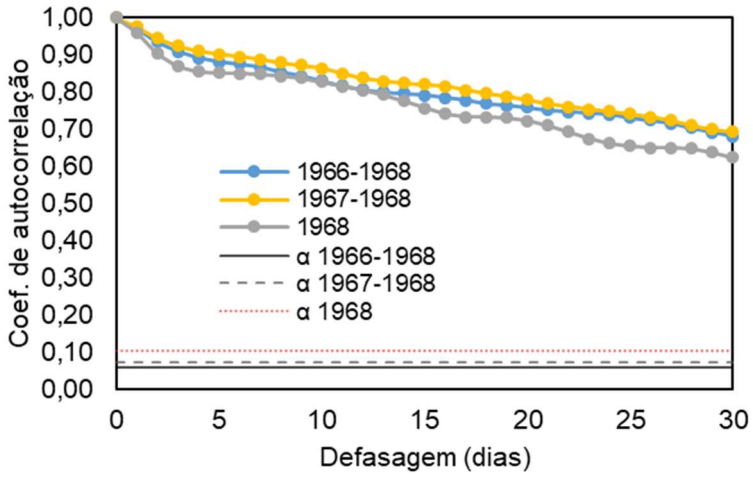

Figura 43: Função de autocorrelação da estação Q3

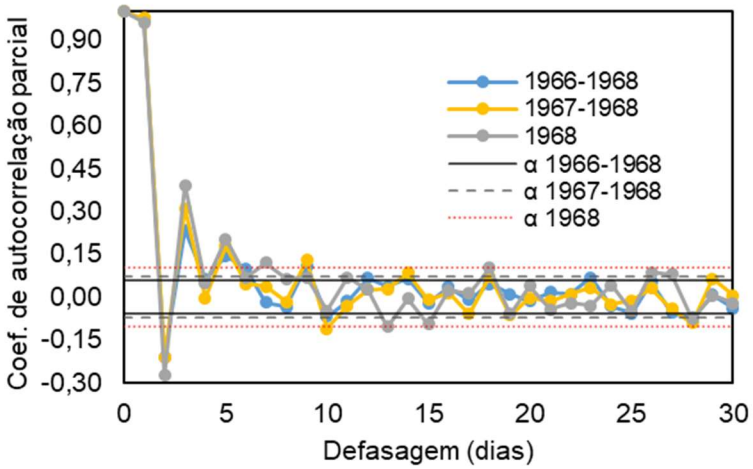

Figura 44: Função de autocorrelação parcial da estação Q3

A partir da análise exploratória, optou-se por avaliar seis diferentes configurações de dados de entrada. Assim, os modelos propostos para simulação de Q3 e os respectivos índices de desempenho estão contidos na Tabela 11. De modo geral, todos os modelos treinados em série-paralelo apresentaram desempenhos muito bons, enquanto que todos os modelos do tipo paralelo alcançaram um bom desempenho. O modelo que apresentou o melhor desempenho tem como entradas as precipitações defasadas entre 0-9 dias, e vazões defasadas de 1-3 dias ($P3_{0:9}Q3_{1:3}$). A RNN-NARX apresentou NSE de 0,877, RSR de 0,35 e PBIAS de 1,52%, logo, são índices considerados muito bons. Além disso, os índices RMSE e MAPE foram de 0,82 m^3s^{-1} e 12,84%, respectivamente, são considerados erros aceitáveis.

Tabela 11: Desempenhos dos modelos propostos para simulação de Q3

Modelo	NO	NSE	RSR	RMSE	PBIAS	MAPE
$P3_{0:9}Q3_{1:3}$	2	0,877	0,35	0,82	1,52	12,84
$P3_{0:9}Q3_{1:5}$	3	0,818	0,43	1,00	-9,58	17,07
$P3_{0:20}Q3_{1:5}$	2	0,838	0,40	0,94	-1,42	15,16
$P3_{0:27}Q3_{1:5}$	2	0,837	0,40	0,95	-1,42	15,95
$P3_{0:20}Q3_{1:9}$	2	0,804	0,44	1,04	-9,44	13,63
$P3_{0:27}Q3_{1:9}$	2	0,827	0,42	0,97	-6,66	15,29

NO – Neurônios Ocultos; NSE – Coeficiente de Eficiência de Nash-Sutcliffe; RSR – Raiz do erro quadrático médio padronizado; PBias – Viés percentual; MAPE – Erro absoluto percentual médio

Os hidrogramas de vazões e o gráfico de dispersão das vazões observadas e simuladas para a estação Q2 são apresentados nas Figuras 45 e 46, respectivamente. A RNN-NARX $P3_{0:9}Q3_{1:3}$ em série-paralelo apresentou os maiores erros percentuais inseridos no início do período chuvoso, meses de janeiro (MAPE de 20,99%) e fevereiro (MAPE de 22,25%). No dia 30/01/1971, a vazão observada foi de 2,15 m^3s^{-1}, e a vazão estimada foi de 3,512 m^3s^{-1} (superestimação de 63,36%). No dia 11/02/1971, a vazão foi observada com 1,61 m^3s^{-1}, e a vazão foi estimada com 2,39 m^3s^{-1} (superestimação de 48,42%). Por outro lado, os menores erros percentuais nos meses de junho (MAPE de 8,93%), julho (MAPE de 6,65%) e agosto (MAPE de 6,30%). O menor erro percentual foi de 0,10%, evidenciado no dia 04/11/1971 com vazão observada de 1,67 m^3s^{-1} e estimada de 1,668 m^3s^{-1}.

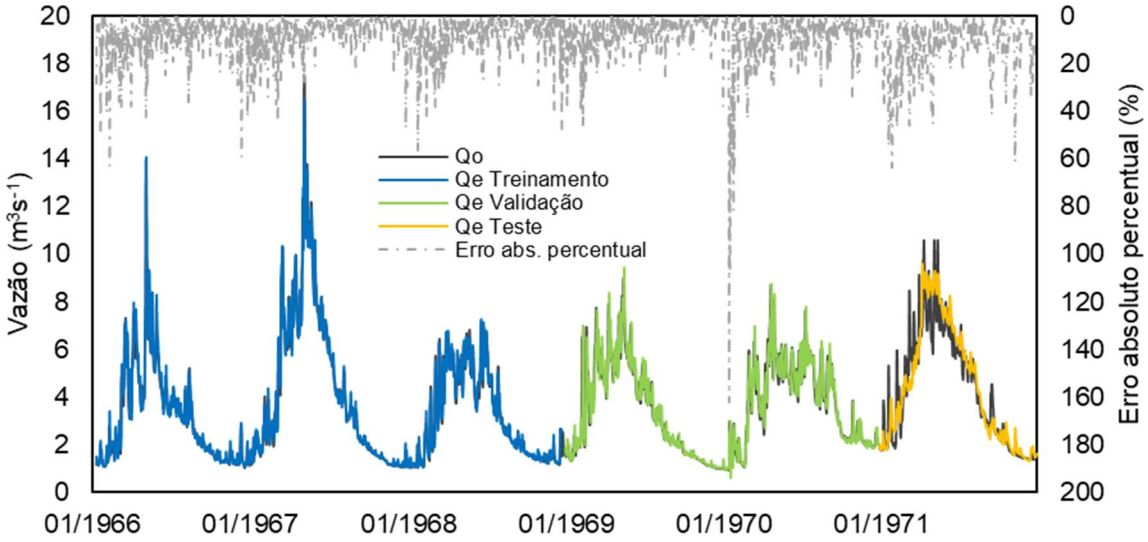

Figura 45: Hidrogramas observados e simulados para a estação Q3

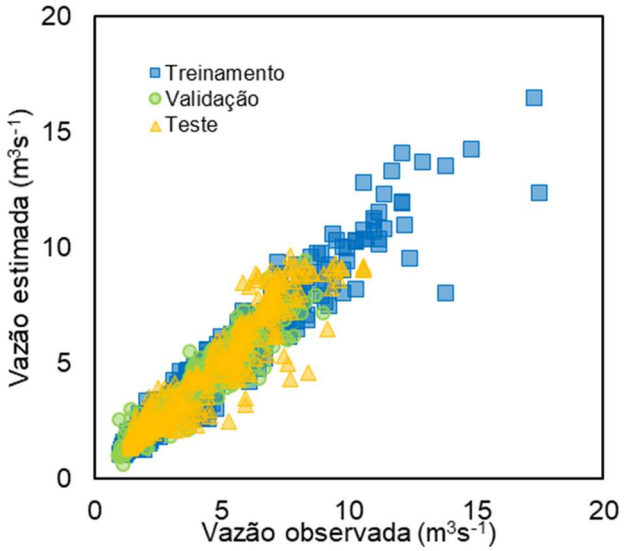

Figura 46: Dispersão entre os dados observados e simulados para a estação Q3

A Figura 47 apresenta os resultados da função de autocorrelação de erro para as simulações de vazões em Q3. É verificado que há presença de autocorrelações significativas nos erros, inseridos no nível de significância de α=5%, com correlação linear do erro defasado de primeira ordem de 0,76. O erro tem média de 0,07 m^3s^{-1} e desvio padrão de 0,82 m^3s^{-1}.

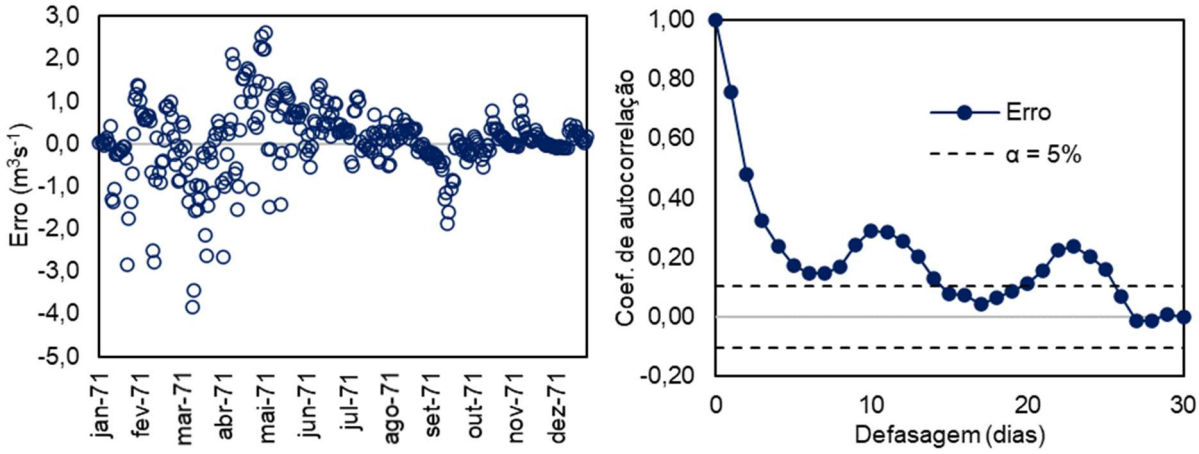

Figura 47: Série temporal e função de autocorrelação do erro em Q3

Analisando as curvas de permanência (Figura 48), foram obtidas boas simulações, com índices NSE de 97,54%, RSR de 0,16 e PBIAS de 1,52%, RMSE de 0,37 m^3s^{-1} e MAPE de 5,69%. Nota-se que a curva de permanência simulada tem baixos erros percentuais até às vazões de referência Q_{20}, e acima disso há tendência de superestimação de vazões. As vazões de referência Q_{90} e Q_{95}, observadas com 1,50 m^3s^{-1} e 1,44 m^3s^{-1}, respectivamente, foram simuladas com 1,66 m^3s^{-1} e 1,48 m^3s^{-1}.

Nestas simulações é observado que há melhores resultados de picos de vazões diárias. Esta propriedade deve ser atribuída à autorregressividade das vazões defasadas em 2 e 3 dias, o que significa que há maior capacidade de retenção de água no interior da bacia, quer seja pelas propriedades físicas do solo, quer seja por intervenção humana. Além disso, a hipótese de estacionariedade para o comportamento da vazão, ao longo do período de simulação, aparenta ser coerente, o que corrobora com a estatística descritiva, pois, há baixo coeficiente de variação e médias próximas entre os conjuntos de dados (MARENGO e ALVES, 2005).

O treinamento supervisionado teve 213 iterações, sendo que a configuração ideal para a RNN-NARX foi encontrada na iteração 22, como mostra a Figura 49. Sendo assim, os RMSE de treinamento, validação e teste foram de 0,19 m^3s^{-1}, 0,19 m^3s^{-1} e 0,30 m^3s^{-1} respectivamente.

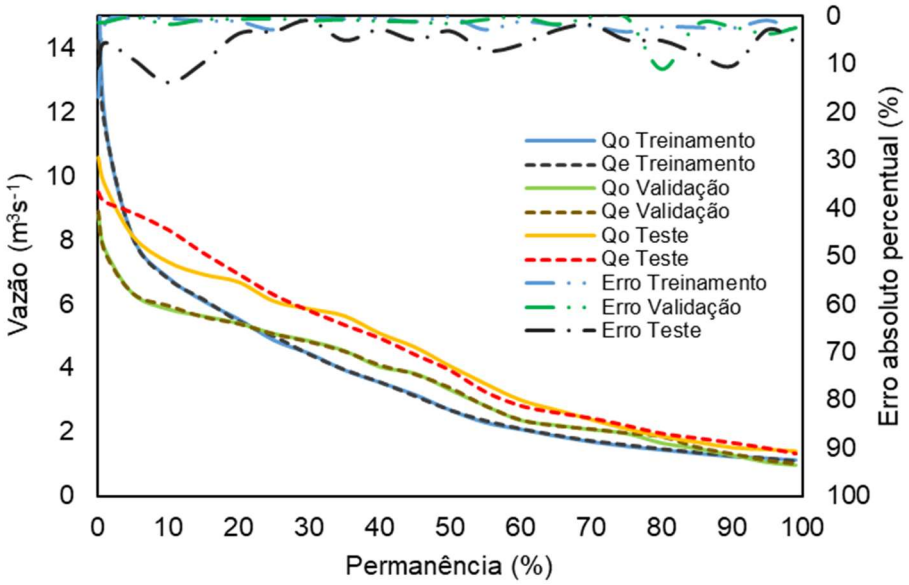

Figura 48: Curvas de permanência observadas e simuladas para a estação Q3

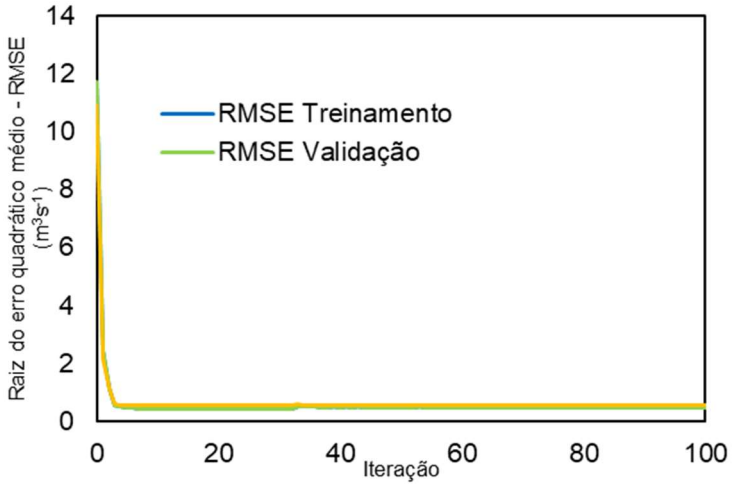

Figura 49: Raiz do erro quadrático médio em função das iterações do modelo para simulação de vazões da estação Q3

A contribuição relativa de cada variável preditora do modelo $P3_{0:9}Q3_{1:3}$ é expressa na Figura 50. As precipitações defasadas em 7 e 6 dias obtiveram contribuição relativa de 22,66% e 14,93%, respectivamente. E, a vazão defasada com maior contribuição é a ocorrida no dia anterior à simulação, com 12,01%. Dessa forma, o êxito das simulações, bem como os índices estatísticos expressivos, é proveniente da relação entre a vazão e as precipitações defasadas nesses dias em específico, permitindo construir a componente sazonal da série.

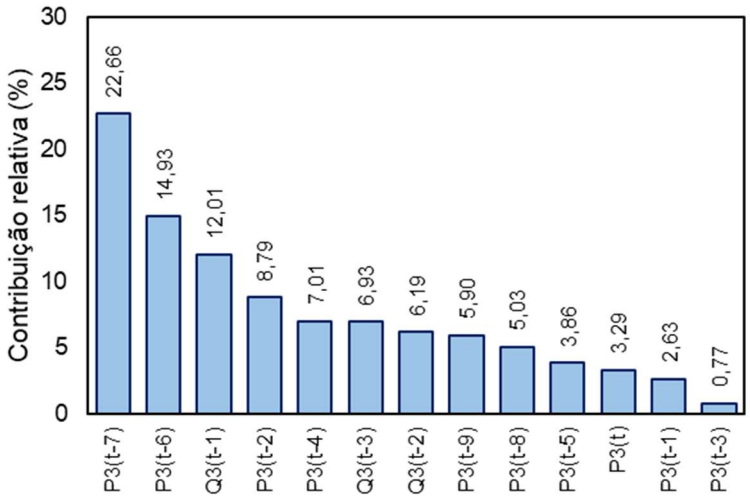

Figura 50: Contribuição relativa dos dados de entrada do modelo para simulação de vazões da estação Q3

A sensibilidade do modelo $P3_{0:9}Q3_{1:3}$ a diferentes períodos de treinamento supervisionado é apresentado na Tabela 12. Os melhores índices estatísticos são encontrados, utilizando-se a maior série disponível, com NSE de 0,877, RSR de 0,35, RMSE de 0,82 m^3s^{-1}, PBIAS de 1,52% e MAPE de 12,84%. Utilizando somente os anos de 1967-1968, há a diminuição de 0,09 do NSE, e aumento de 0,12 no RSR, 0,27 m^3s^{-1} de RMSE, 8,97% de PBIAS, e 4,59 de MAPE. E, com o ano de 1968, o índice NSE é reduzido em 0,126, e os erros aumentam consideravelmente, principalmente o PBIAS, que passa a ser -13,39%. Sendo assim, é constatada a degradação dos índices com a redução das séries temporais.

Tabela 12: Sensibilidade do modelo à diferentes resoluções temporais para simulações de Q3

Período	NO	NSE	RSR	RMSE	PBIAS	MAPE
1966-1968	2	0,877	0,35	0,82	1,52	12,84
1967-1968	3	0,781	0,47	1,09	10,49	17,43
1968	3	0,751	0,50	1,17	-13,39	16,00

NO – Neurônios Ocultos; NSE – Coeficiente de Eficiência de Nash-Sutcliffe; RSR – Raiz do erro quadrático médio padronizado; PBias – Viés percentual; MAPE – Erro absoluto percentual médio

As Figuras 51 e 52 mostram os diferentes hidrogramas e curvas de permanência da análise de sensibilidade do modelo para simulação de Q3. Ao utilizar somente o ano de 1968, percebe-se que há um declínio acentuado entre os meses de abril e julho. E, as simulações realizadas com os anos de 1966-1968 e 1967-1968 têm formas semelhantes. Além disso, as curvas de permanência simuladas têm melhor aderência às vazões menores que Q_{50}.

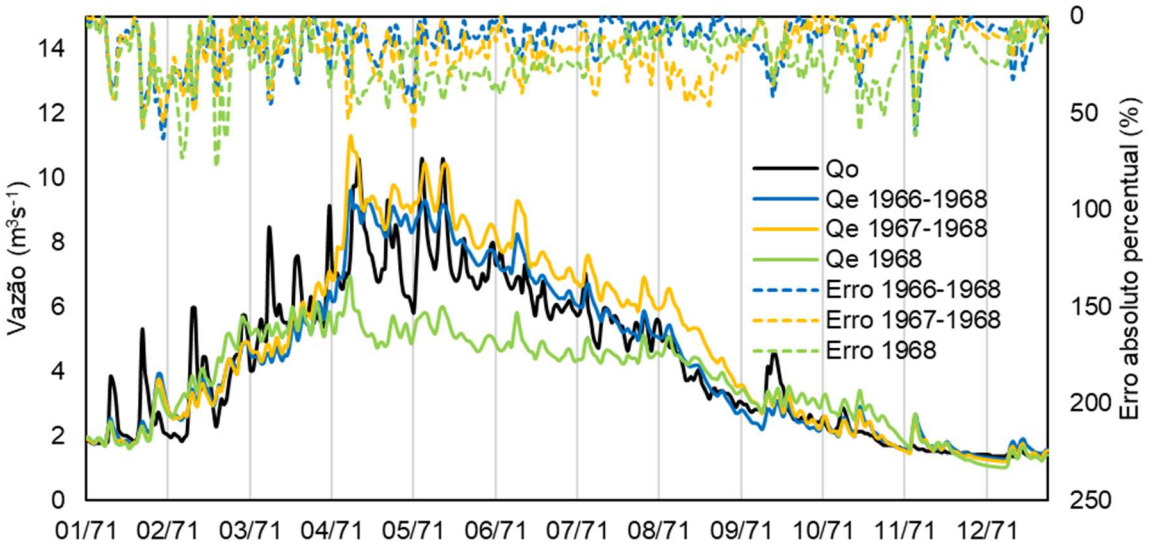

Figura 51: Hidrogramas observado e simulados para Q3, utilizando diferentes intervalos temporais

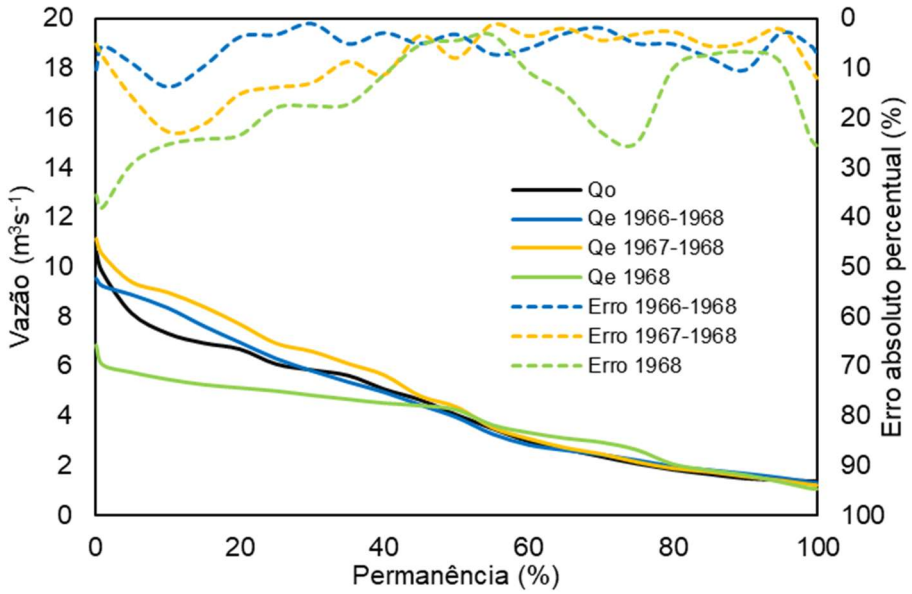

Figura 52: Curvas de permanência observada e simuladas para Q3, utilizando diferentes intervalos temporais

4.4 RIO CAPIVARA

A estação fluviométrica cujo nome é Próximo a Colinas do Tocantins (Q4) (cód. 23130000) realiza o monitoramento das vazões do rio Capivara, localizado no município de Colinas do Tocantins, norte do Estado do Tocantins. A bacia hidrográfica tem área de aproximadamente 386 km^2 (ANA, 2021), as altitudes mínima, máxima e média são de 189 m, 594 m e 269 m, respectivamente, e a declividade é predominantemente suave ondulada (média de 7,03%) (USGS, 2014; EMBRAPA, 1999). Os mapas de caracterização desta bacia são mostrados na Figura 53.

Há cinco diferentes classes de solos no interior da bacia de Q_4 (IBGE, 2021). O latossolo vermelho distrófico (37,39%), o neossolo quartzarênico órtico (35,75%) e o latossolo vermelho-amarelo distrófico (7,73%), que são solos de altas taxas de infiltração. Por outro lado, o argissolo vermelho-amarelo distrófico (9,73%) e o neossolo litólico distrófico (9,70%) são as classes que apresentam baixas taxas de infiltração, sendo causado principalmente pelo alto teor de argila e baixas profundidades, respectivamente (SARTORI et al, 2005). O principal uso do solo da bacia é a pastagem, com percentual de 70,92%, seguido da classe vegetação natural secundária, com 15,64% (INPE, 2018).

A estação pluviométrica Colinas do Tocantins (cód. 848000) (P4) está localizada na parte sul da bacia de Q4, na área urbana do município de Colinas do Tocantins. Neste local, o regime intra-anual das chuvas é definido pelo período chuvoso, que se estende entre os meses de novembro a abril, e pelo período de estiagem, compreendido entre os meses de maio a outubro, como mostra a normal climatológica observada em P4 (Figura 54). O clima característico da região é o Aw, conforme a classificação de Köppen-Geiger, logo, é um clima tropical de savana, pois o mês mais seco tem precipitação inferior a 60 mm, com temperatura média de aproximadamente 26°C (ALVARES, 2013; ROLDÃO e FERREIA, 2019).

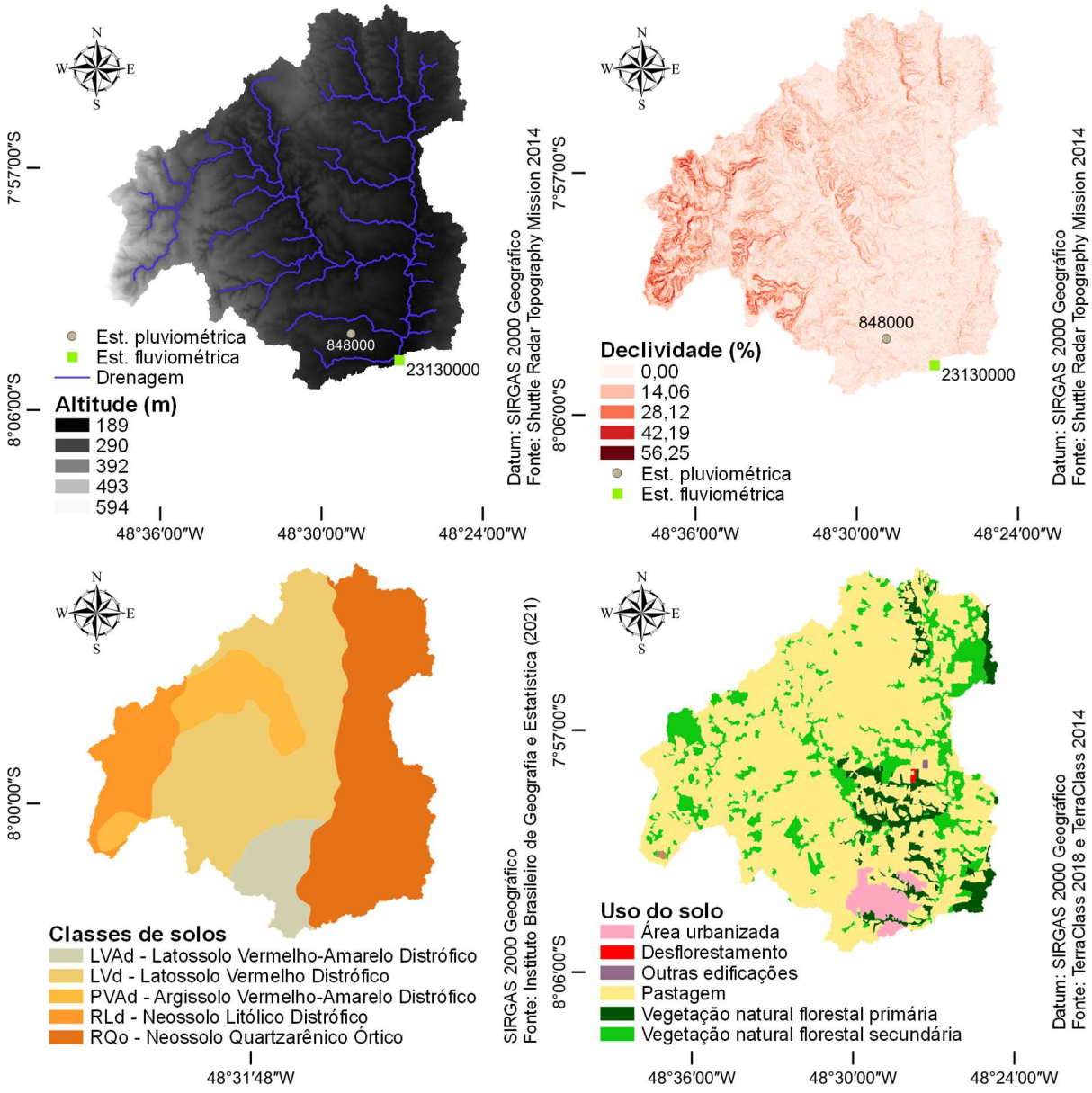

Figura 53: Caracterização da bacia da estação Q4

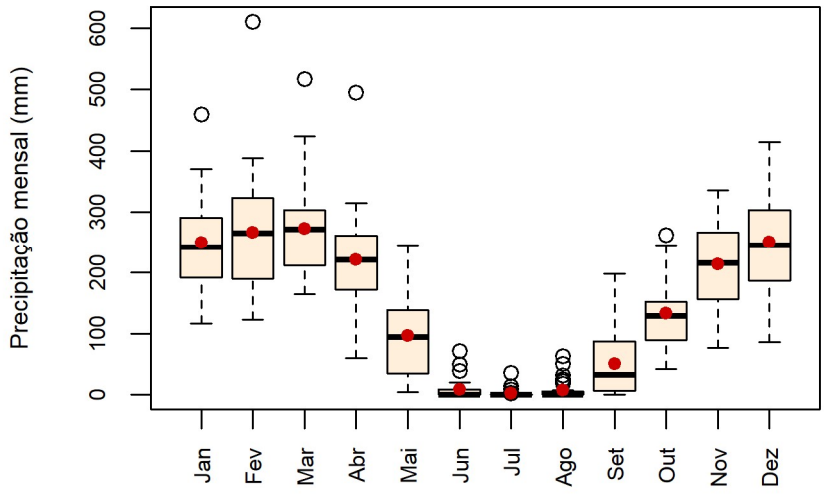

Figura 54: Normal climatológica da estação P4 entre 1991-2020

As séries históricas selecionadas para simulação de Q_4 possuem períodos de medição entre os anos de 1993-1998. A divisão dos conjuntos de dados pode ser visualizada na Figura 55. Foi definido que o conjunto de treinamento é referente aos dados observados nos anos entre 1993-1995 (1.095 dados); o de validação é referente aos anos de 1996-1997 (730 dados); e o de teste é referente ao ano de 1998 (365 dados).

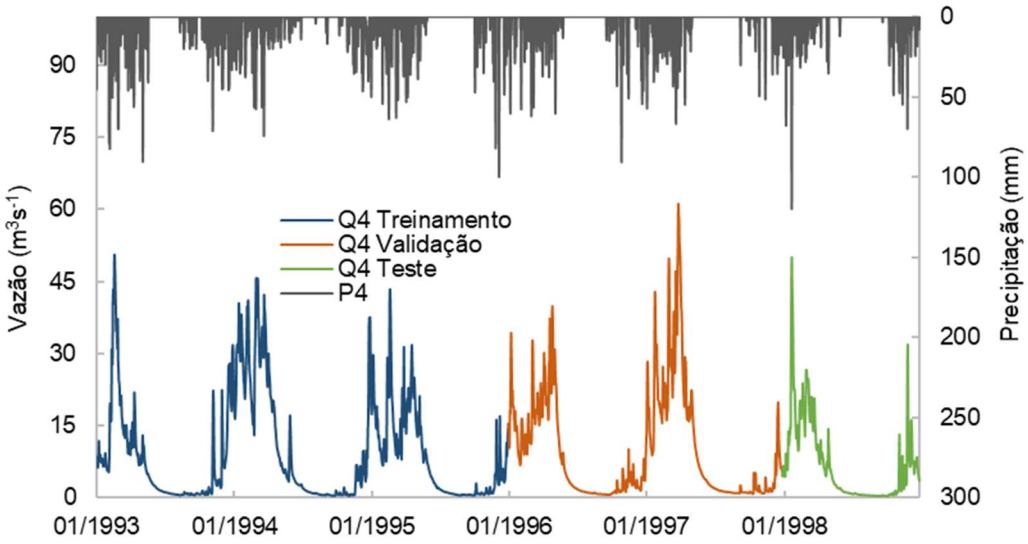

Figura 55: Hidrogramas de vazões da estação Q4 entre 1993-1998

A Tabela 13 mostra a estatística descritiva das variáveis P4 e Q4. Foi observado que no ano de 1998 foram evidenciados os menores índices pluviométricos e fluviométricos. Neste ano, a precipitação anual foi de 1.389 mm, e as vazões médias e de referência Q_{95} foram de 6,84 m^3s^{-1} e 0,46 m^3s^{-1}, respectivamente. Esses índices estão relacionados ao El Niño, ocorrido de magnitude forte em 1997-1998 (INPE, 2021). Entretanto, quando são analisados os maiores índices, é observado divergência entre as variáveis. O ano com o maior índice pluviométrico ocorreu no ano de 1993, com precipitação anual de 2.310 mm, enquanto que o maior índice fluviométrico ocorreu no ano de 1994, com vazões média e de referência Q_{95}, sendo iguais a 11,29 m^3s^{-1} e 0,67 m^3s^{-1}, respectivamente. Isto está relacionado à variabilidade intra-anual da precipitação, pois as precipitações ocorridas nos meses outubro, novembro e dezembro de 1993 influenciam diretamente no hidrograma característico de 1994.

Tabela 13: Estatística descritiva das variáveis P4 e Q4

Variável hidrológica	Estatística	Treinamento			Validação		Teste
		1993	1994	1995	1996	1997	1998
P4	Média[1] (mm/dia)	6,33	4,76	4,92	4,29	4,13	3,81
	Anual (mm/ano)	2.310,00	1.736,60	1.794,40	1.570,40	1.507,40	1.389,40
	Máx (mm/dia)	90,50	74,20	100,00	90,50	66,60	120,00
	Min[2] (mm/dia)	1,70	1,00	1,50	0,80	0,20	0,10
	Coef. de variação (%)	229,31	224,82	263,13	276,12	253,64	293,64
Q4	Média (m³/s)	8,08	11,29	8,19	8,37	10,83	6,84
	Máx (m³/s)	50,32	45,55	43,30	39,84	60,96	50,01
	Min (m³/s)	0,67	0,52	0,63	0,56	0,77	0,42
	Coef. de variação (%)	120,79	114,12	105,22	108,37	121,10	121,29
	$Q_{50\%}$ (m³/s)	3,96	4,45	4,59	3,69	3,80	3,16
	$Q_{95\%}$ (m³/s)	0,75	0,67	0,71	0,63	0,87	0,46

[1]Considerando dias sem chuva
[2]Acima de zero

A Figura 56 mostra as funções de correlação cruzada entre as séries temporais de Q4 e P4, considerando diferentes resoluções temporais. Considerando os anos entre 1993-1995, e um nível de significância de 5%, destacam-se as precipitações diárias defasadas entre 0-30 dias. Considerando os anos entre 1994-1995, e o nível de significância de 5% (α = 0,05), destacam-se as precipitações diárias defasadas entre 0-30 dias. E, considerando apenas o ano de 2012, e um nível de significância de 5%, destacam-se as precipitações diárias defasadas 0-26 dias. Todas as correlações têm coeficiente menores que 0,5, o que as classificam como correlações fracas (PECK et al., 2015). O ano de 2015 apresentou função de correlação cruzada menor que as demais.

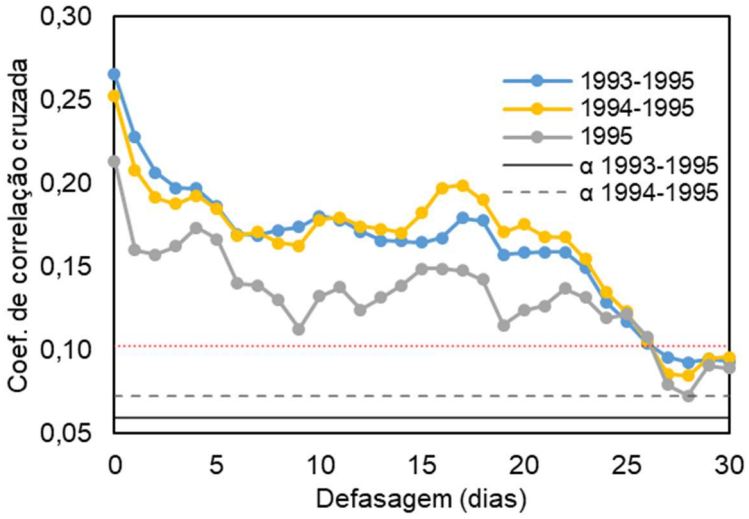

Figura 56: Função de correlação cruzada entre as estações P4 e Q4

As Figuras 57 e 58 mostram as funções de autocorrelação e autocorrelação parcial de Q_4, respectivamente. Considerando os anos entre 1993-1995, e um nível de significância de 5%,destacam-se as vazões diárias de 1-2 dias, 10-11 dias, 15 dias, 20 dias e entre 26-28 dias. Considerando os anos entre 1994-1995, e um nível

de significância de 5%, destacam-se as vazões defasadas entre 1-2 dias, 5 dias, 11 dias, 20 dias, entre 26-28 dias defasados. E, considerando apenas o ano de 1995, e o nível de significância de 5% (α = 0,05), destacam-se as vazões diárias entre 1 dia, 4 dias, 10 dias, 17 e 27 dias defasados. Segundo Peck et al. (2015), a vazão defasada em 1 dia têm forte correlação, entretanto, as demais variáveis são consideradas fracas. A função de autocorrelação referente ao ano de 1995 também apresentou diminuição dos coeficientes.

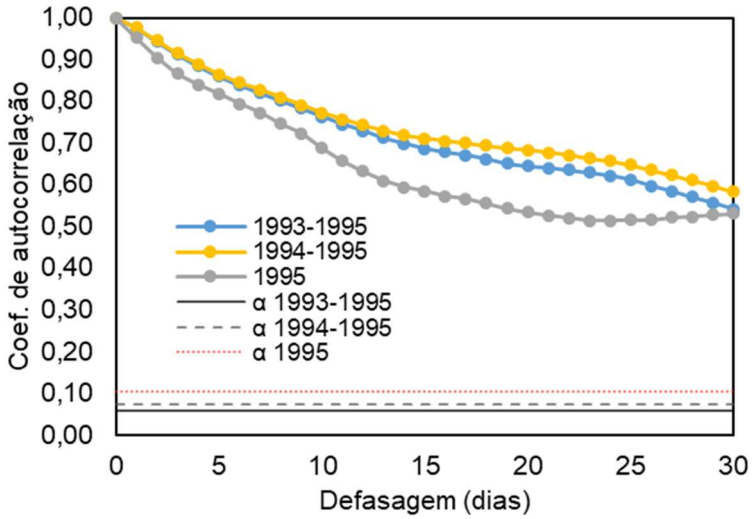

Figura 57: Função de autocorrelação da estação Q4

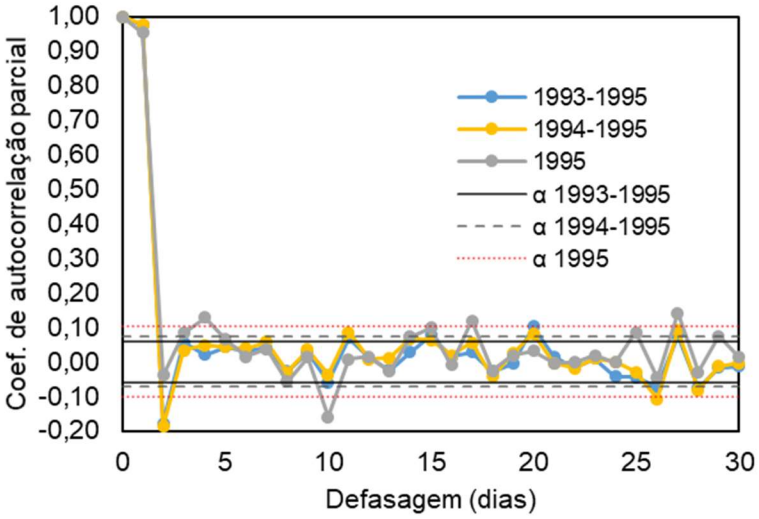

Figura 58: Função de autocorrelação parcial da estação Q4

Com base nas correlações expostas anteriormente, foram escolhidos seis diferentes vetores de entrada para simulação de vazões em Q4, de modo que todas as correlações relevantes fossem inseridas gradativamente. A descrição dos vetores e os critérios de desempenho das simulações constam da Tabela 14. A RNN-NARX que demonstrou melhor resultado foi a com as precipitações defasadas entre 0-11

dias e as vazões defasadas em 1-5 dias ($P4_{0:11}Q4_{1:5}$), e três neurônios na camada oculta. Este modelo alcançou índices de NSE de 0,725, RSR de 0,52, RMSE de 4,34 m^3s^{-1}, PBIAS de -5,87% e MAPE de 64,66%, o que o classifica com bons resultados (MORIASI et al., 2007).

Tabela 14: Desempenhos dos modelos propostos para simulação de Q4

Modelo	NO	NSE	RSR	RMSE	PBIAS	MAPE
$P4_{0:11}Q4_{1}$	3	0,680	0,56	4,69	-9,66	54,97
$P4_{0:11}Q4_{1:5}$	3	0,725	0,52	4,34	-5,87	64,66
$P4_{0:17}Q4_{1:5}$	3	0,501	0,71	5,85	-20,57	60,6
$P4_{0:17}Q4_{1:11}$	3	0,673	0,57	4,74	3,16	84,4
$P4_{0:26}Q4_{1:11}$	2	0,619	0,62	5,11	-5,71	70,62
$P4_{0:27}Q4_{1:27}$	3	0,593	0,64	5,28	-21,29	79,23

NO – Neurônios Ocultos; NSE – Coeficiente de Eficiência de Nash-Sutcliffe; RSR – Raiz do erro quadrático médio padronizado; PBIAS – Viés percentual; MAPE – Erro absoluto percentual médio

As Figuras 59 e 60 mostram as vazões observadas e simuladas pela RNN-NARX $P4_{0:11}Q4_{1}$ em paralelo sob a forma de hidrograma de vazões e gráfico de dispersão, respectivamente. Os maiores erros percentuais são observados nos meses de outubro (MAPE de 174%) e novembro (149%), período em que começa o período chuvoso na região. O maior erro absoluto percentual corresponde à vazão observada com 0,726 m^3s^{-1}, no dia 25/10/1998, que foi simulada com 4,384 m^3s^{-1} (superestimação de 503%). Outro exemplo é a vazão observada no dia 13/11/1998, com 1,453 m^3s^{-1}, que foi simulada com 8,446 m^3s^{-1} (superestimação de 481%). No entanto, os menores erros percentuais estão no mês de setembro (MAPE de 7,34%), período em que são observadas as menores vazões do ano, e o mês de março (MAPE de 12,67%). O menor erro absoluto percentual foi de 0,07%, com a vazão observada de 22,8 m^3s^{-1}, no dia 12/03/1998¸ que foi simulada com 22,8 m^3s^{-1}.

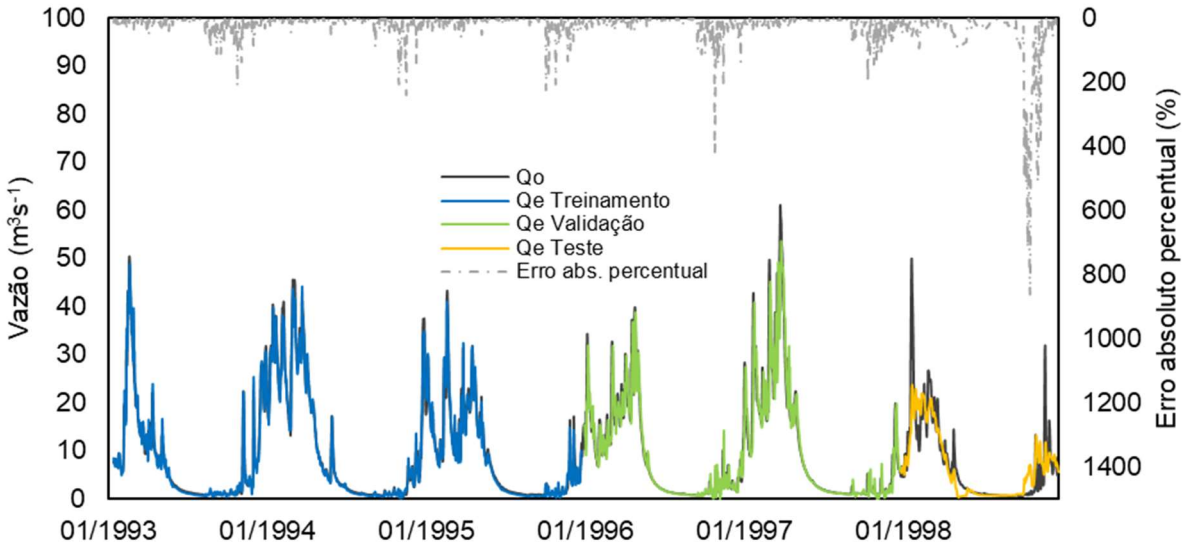

Figura 59: Hidrogramas observados e simulados para a estação Q4

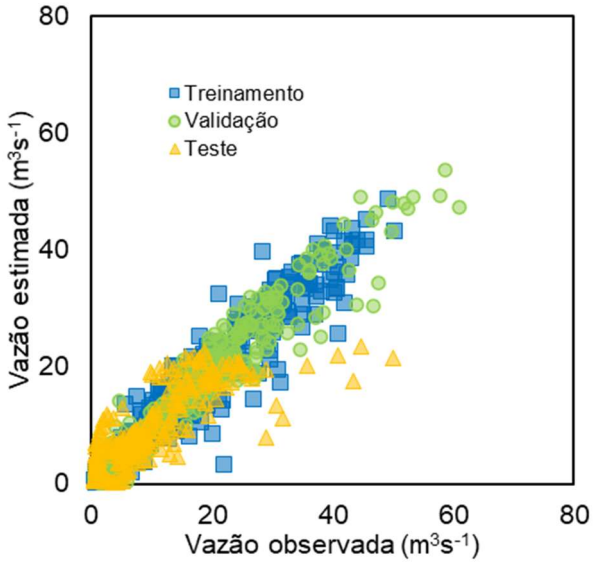

Figura 60: Dispersão entre os dados observados e simulados para a estação Q4

A Figura 61 apresenta os resultados da função de autocorrelação de erro para as simulações de vazões em Q4. É verificado que há presença de autocorrelações significativas nos erros, inseridos no nível de significância de α=5%, com correlação linear do erro defasado de primeira ordem de 0,86. A média e o desvio padrão do erro são iguais a respectivamente -0,40 m^3s^{-1} e 4,33 m^3s^{-1}.

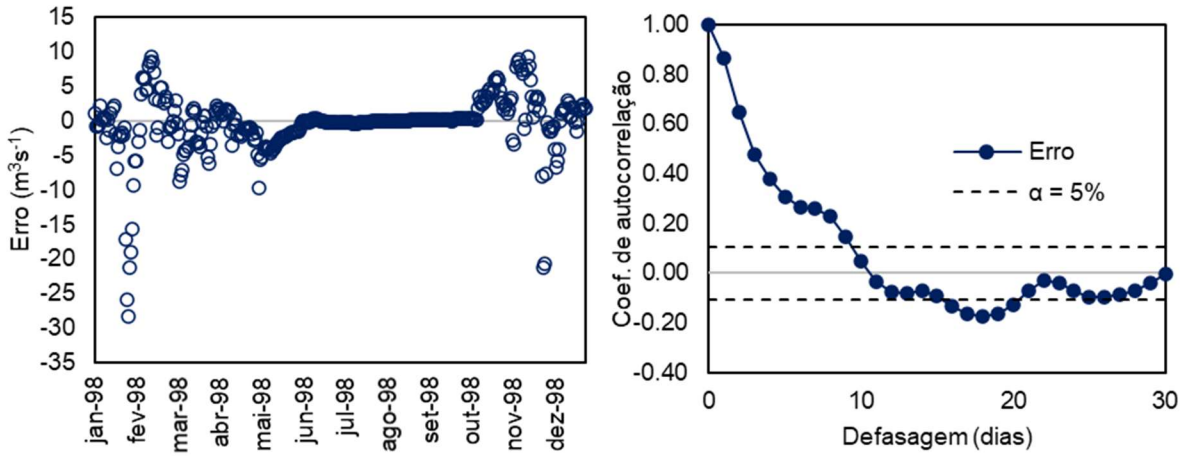

Figura 61: Série temporal e função de autocorrelação do erro em Q4

A Figura 62 mostra as curvas de permanência observadas e simuladas para a estação Q4. Os índices estatísticos mostram simulações muito boas, com índices NSE de 94,30%, RSR de 0,24, RMSE de 1,98 m^3s^{-1}, PBIAS de -6,03%, e MAPE de 17,78%. As vazões de referência Q_{90} e Q_{95}, observadas com 0,50 m^3s^{-1} e 0,46 m^3s^{-1}, respectivamente, foram simuladas com 0,56 m^3s^{-1} e 0,50 m^3s^{-1}.

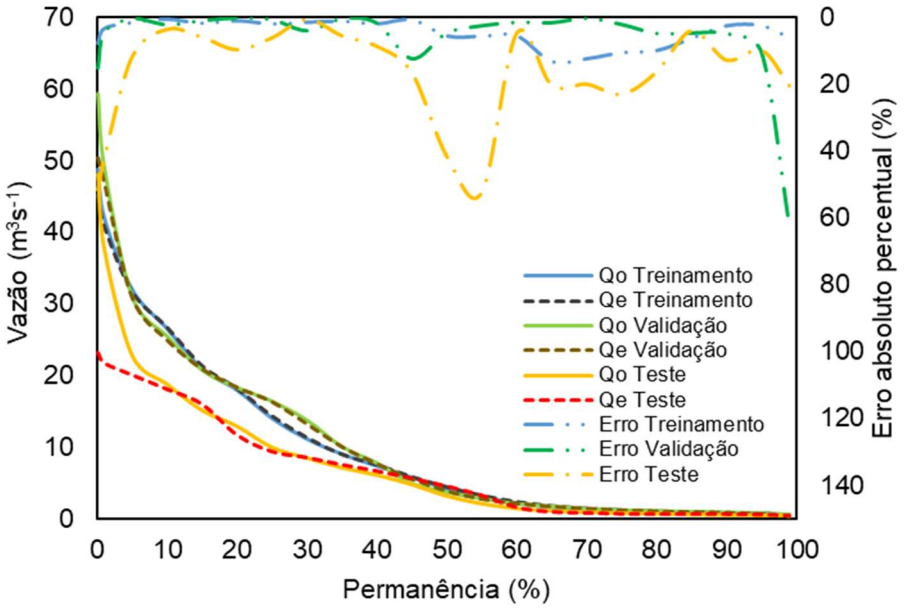

Figura 62: Curvas de permanência observadas e simuladas para a estação Q4

O treinamento supervisionado teve 213 iterações, sendo que a configuração ideal para a RNN-NARX foi encontrada na iteração 40, como mostra a Figura 63. Apesar da diminuição contínua do erro de treinamento, o mesmo comportamento não foi evidenciado nos demais conjuntos de dados. Sendo assim, os RMSE de treinamento, validação e teste foram de 3,96 m^3s^{-1}, 5,79 m^3s^{-1} e 4,04 m^3s^{-1} respectivamente.

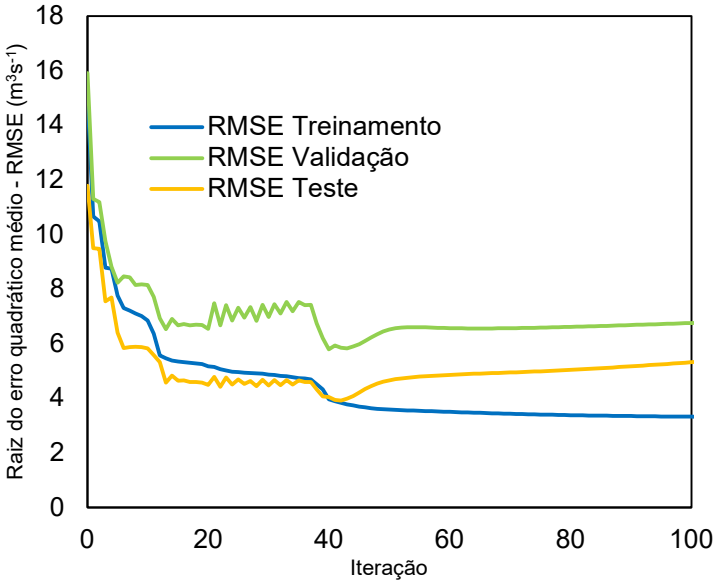

Figura 63: Raiz do erro quadrático médio em função das iterações do modelo para simulação de vazões da estação Q4

A Figura 64 mostra a contribuição relativa das variáveis de entrada para simulação das vazões em Q4. Com percentuais de 14,30% e 11,19% de contribuição relativa, as precipitações defasadas em 11 e 10 dias têm relevantes importâncias para as simulações. A vazão defasada em 1 dia obteve o percentual de 13,67%. Assim, pode-se subentender que o comportamento das vazões simuladas tem relação com precipitações defasadas para descrever a dinâmica da água na bacia, assim como relatado em bacias anteriores.

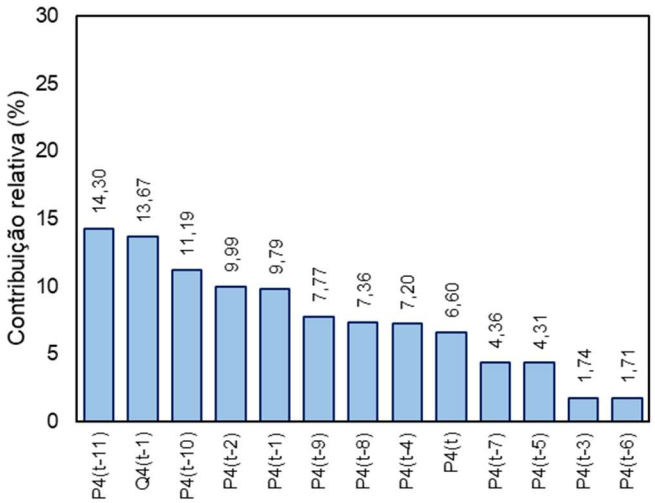

Figura 64: Contribuição relativa dos dados de entrada do modelo para simulação de vazões da estação Q4

A sensibilidade da RNN-NARX à diferentes resoluções temporais é expressa na Tabela 15. Verifica-se que somente a RNN-NARX em paralelo utilizando os anos de 1993-1995 alcançaram desempenho estatístico classificado como muito bom, e a maior parte das proposições atingiram apenas o desempenho bom. A utilização do ano de 1995 para treinamento supervisionado apresenta diminuição do NSE em 11,42%, e aumento do RSR de 0,11, RMSE de 0,88 m^3s^{-1}, da subestimação em 4,07% e MAPE de 12,26%. Deste modo, a utilização de um ano hidrológico é suficiente para a obtenção de resultados satisfatórios. Porém, as RNN-NARX em série-paralelo não apresentam diminuição clara dos desempenhos estatísticos com a diminuição das amostras, o que é considerado uma boa perspectiva.

A Figura 65 mostra os hidrogramas observado e simulado para a estação Q4 no ano de 1998. O hidrograma simulado a partir do ano de 1995 apresenta picos de vazão acentuados no hidrograma nos meses de janeiro, novembro e dezembro, sendo as principais fontes de erros do modelo. A simulação de 1994-1995 apresentou resultados de subestimação principalmente entre os meses de fevereiro e maio.

A Figura 66 mostra a análise sensibilidade do modelo sob a forma de curvas de permanência. Comparando as simulações, nota-se que o período entre 1993-1995 se ajustou melhor, com menores erros percentuais.

Tabela 15: Sensibilidade do modelo à diferentes resoluções temporais para simulações de Q4

Período	NO	NSE	RSR	RMSE	PBIAS	MAPE
1993-1995	3	0,725	0,52	4,34	-5,87	64,66
1994-1995	4	0,628	0,61	5,05	-11,47	58,12
1995	4	0,709	0,54	4,47	-0,25	80,05

NO – Neurônios Ocultos; NSE – Coeficiente de Eficiência de Nash-Sutcliffe; RSR – Raiz do erro quadrático médio padronizado; PBias – Viés percentual; MAPE – Erro absoluto percentual médio

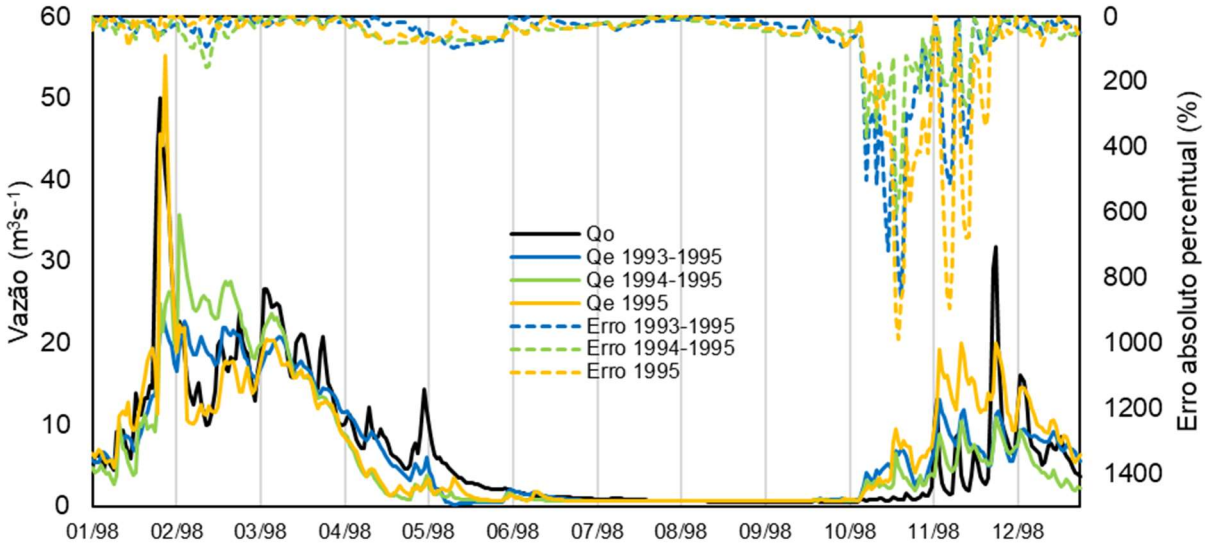

Figura 65: Hidrogramas observado e simulados para Q4, utilizando diferentes intervalos temporais

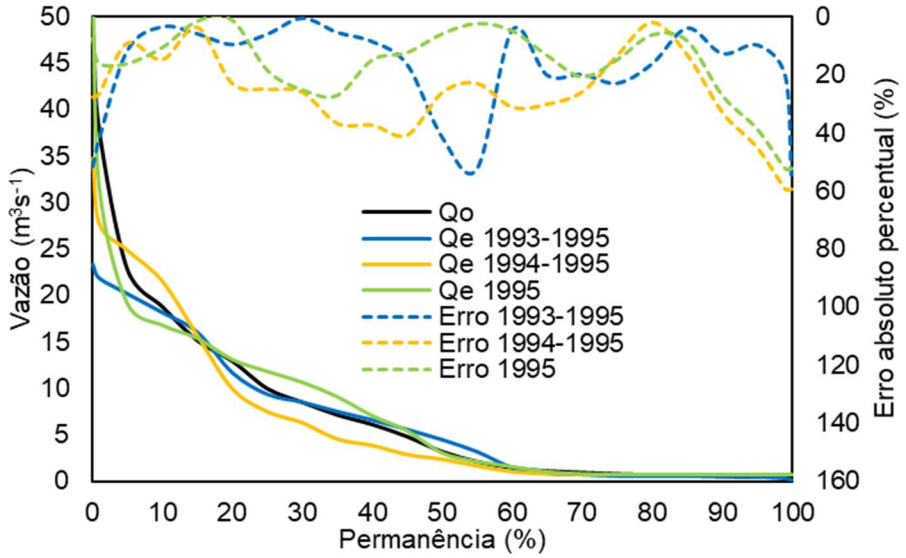

Figura 66: Curvas de permanência observada e simuladas para Q4, utilizando diferentes intervalos temporais

4.5 RIO BRAÇO NORTE

A estação fluviométrica Base do Cachimbo (Q5) (Cód. 17345000) está situada no rio Braço Norte, município de Novo Progresso, sul do Estado do Pará. A área de contribuição referente a esta estação é de aproximadamente 465 km^2 (ANA, 2021). Além disso, 42% da área de drenagem estão inseridos na unidade de conservação federal, de proteção integral, denominada Reserva Biológica Nascentes da Serra do Cachimbo (Figura 67) (SILVA et al., 2021).

A Figura 67 mostra a caracterização da bacia de Q5. Verifica-se que as altitudes mínima, máxima e média da bacia são de 455 m, 692 m, 541 m, respectivamente. As declividades média e máxima da bacia são, respectivamente, de 5,24% e 55,34%, ou seja, a declividade da área é classificada como suave ondulada (EMBRAPA, 1999; USGS, 2014). A classe de solo predominante na bacia de Q5 é o Neossolo Quartzarênico Órtico (IBGE, 2021). Segundo Sartori et al. (2005), este tipo de solo apresenta alta taxa de infiltração hídrica, e desfaz-se com facilidade, o que o torna muito susceptível a erosão hídrica. Além disso, o uso da terra é marcado pela formação natural florestal (18,33%) e natural não-florestal (80,88%) (SILVA et al., 2021), sendo a última definida por vegetação de campina e campinarana, muito comuns na região (LLERAS e KIRKBRIDE JR, 1978).

Há 132 km de distância, na direção norte a partir da estação Q5, está localizada a estação pluviométrica KM 947 BR-163 (P5) (Cód. 855000). O clima local é caracterizado como Am, conforme a classificação de Köppen-Geiger, ou seja, é um clima tropical úmido, com altos índices pluviométricos anuais e por apresentar pequeno período seco (ALVARES et al., 2013). O período chuvoso da região é iniciado em outubro e estende-se até o mês de abril, e o período de estiagem estende-se desde o mês de maio até setembro, como mostra a normal climatológica exposta na Figura 68.

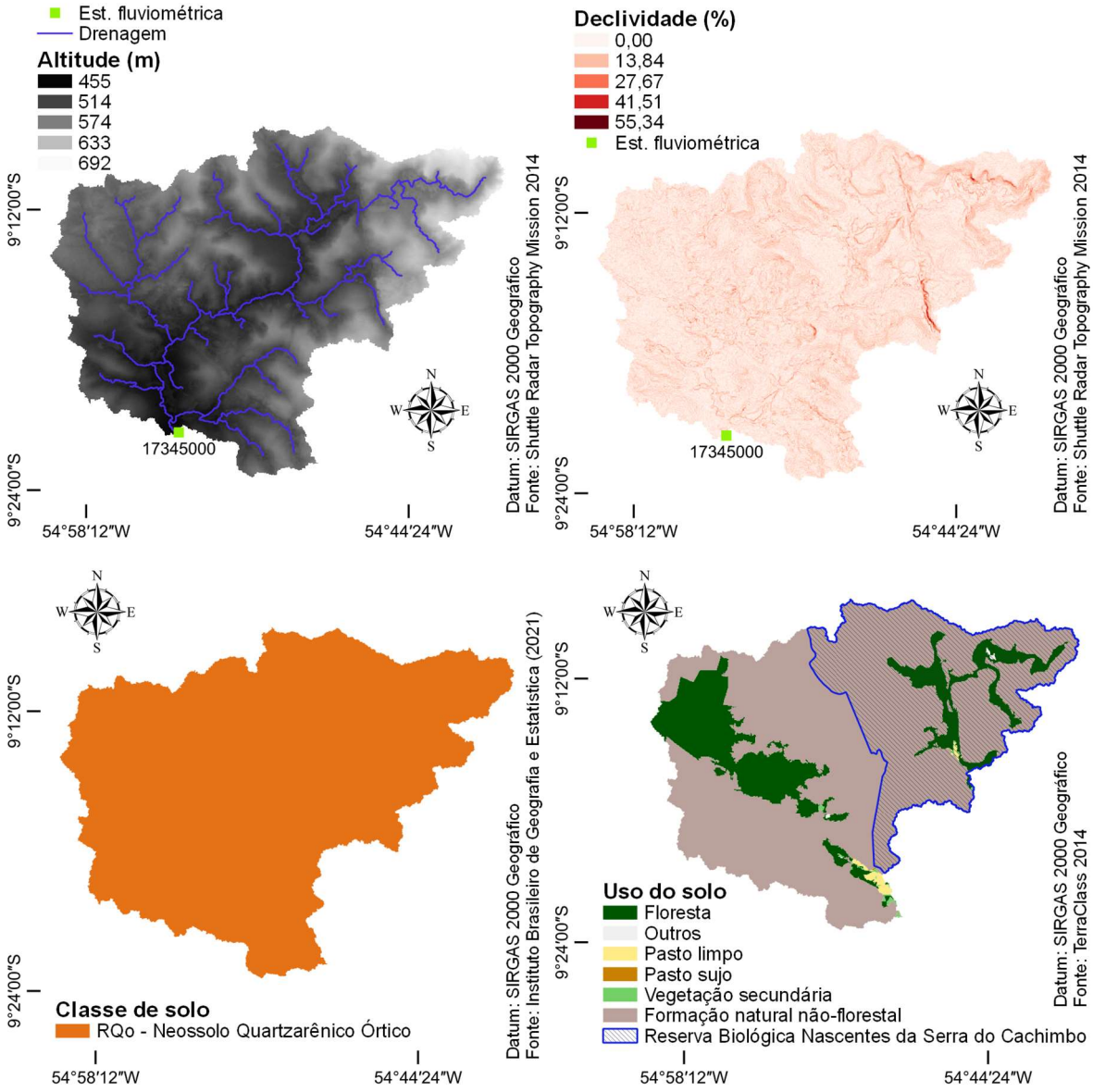

Figura 67: Caracterização da bacia da estação Q5

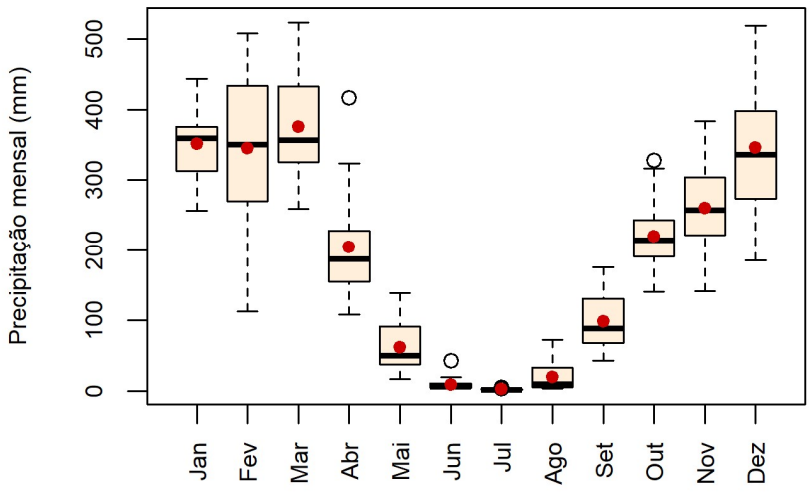

Figura 68: Normal climatológica na bacia de Q5 entre 1991-2020

Fonte: Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS, 2020)

Com o objetivo de realizar as simulações de vazões, as séries temporais Q5 e P5 foram divididas conforme descrito a seguir. Os anos de 2015-2016 (731 dias) foram atribuídos ao treinamento das RNN-NARX e análise exploratória. O ano de 2017 (365 dias) foi utilizado para validação cruzada das RNN-NARX. E, o ano de 2018 (365 dias) foi destinado ao teste estatístico dos modelos. A divisão das séries temporais é exibida na Figura 69.

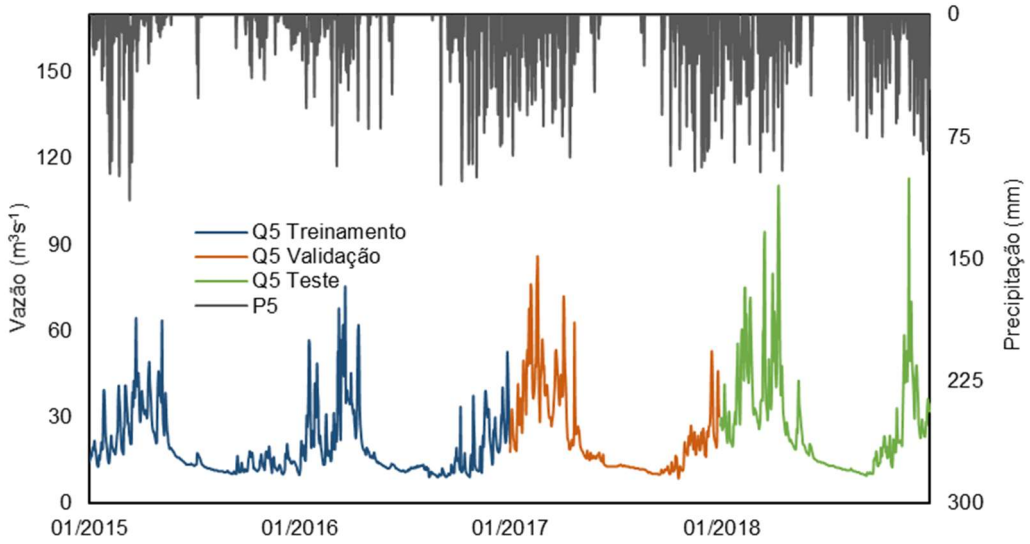

Figura 69: Hidrogramas de vazões da estação Q5 entre 2015-2018

A Tabela 16 mostra a descrição estatística das variáveis utilizadas para simulação das vazões de Q5. Verifica-se que o ano de 2015 apresentou os menores índices hidrológicos, com precipitação anual de 1.970 mm e vazão média igual a 19 m^3s^{-1}, enquanto que o ano de 2018 apresentou os maiores índices, com precipitação anual de 4.197 mm e vazão média de 27 m^3s^{-1}. Esta variabilidade pode estar associada a fenômenos meteorológicos, pois, os anos hidrológicos de 2015-2016 apresentaram fortes índices de anomalia climática, caracterizando-os com o fenômeno El Niño. Por outro lado, os anos de 2017-2018 foram alvos de anomalias positivas, caracterizando-os com o fenômeno La Niña (INPE, 2021).

Tabela 16: Estatística descritiva das variáveis P5 e Q5

Variável hidrológica	Estatística	Treinamento		Validação	Teste
		2015	2016	2017	2018
P5	Média[1] (mm/dia)	5,40	8,85	9,57	11,50
	Anual (mm/ano)	1.970,90	3.237,70	3.491,40	4.197,10
	Máx (mm/dia)	114,10	104,60	96,30	96,80
	Min[2] (mm/dia)	0,10	0,50	0,90	11,10
	Coeficiente de variação (%)	285,04	226,38	219,21	185,14
Q5	Média (m³/s)	19,01	19,84	23,25	27,36
	Máx (m³/s)	64,54	75,33	85,68	112,89
	Min (m³/s)	9,76	9,06	8,52	9,48
	Coeficiente de variação (%)	0,52	0,59	0,63	0,65
	$Q_{50\%}$ (m³/s)	15,33	15,33	17,85	21,93
	$Q_{95\%}$ (m³/s)	10,64	9,62	10,19	10,45

[1]Considerando dias sem chuva
[2]Acima de zero

A Figura 70 mostra as funções de correlação cruzada entre P5 e Q5 em diferentes intervalos de séries temporais. Adotando os anos de 2015 e 2016, e nível de significância de 5%, destacam-se as precipitações diárias defasadas entre 0-11 dias. E, adotando apenas o ano de 2016, e o nível de significância de 5% (α = 0,05), destacam-se as precipitações diárias defasadas 0-3 dias. Todos os coeficientes de correlação são considerados fracos, uma vez que o maior coeficiente é de aproximadamente 15%. Esta baixa correlação linear pode estar vinculada a distância entre a bacia e a estação pluviométrica.

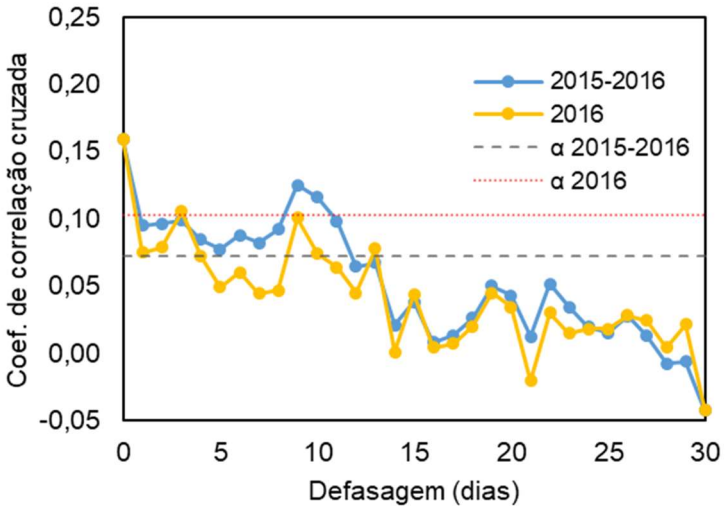

Figura 70: Função de correlação cruzada entre as estações P5 e Q5

As Figuras 71 e 72 mostram as funções de autocorrelação e autocorrelação parcial da estação Base do Cachimbo, respectivamente. Considerando os anos de 201 e 2016, e um nível de significância de 5%, destacam-se as vazões defasadas em 1 dia, 4 dias, 6 dias, 9 dias, 13 dias, 19 dias e 26 dias. E, considerando o ano de 2016, e o nível de significância de 5% (α = 0,05), destacam-se as vazões diárias de 1 dia, 4 dias, 9 dias, 13 dias e 19 dias. Somente a vazão defasada em 1 dia tem

correlação forte (PECK et al., 2015). A função de autocorrelação de 2015-2016 apresentou maiores coeficientes de correlação linear.

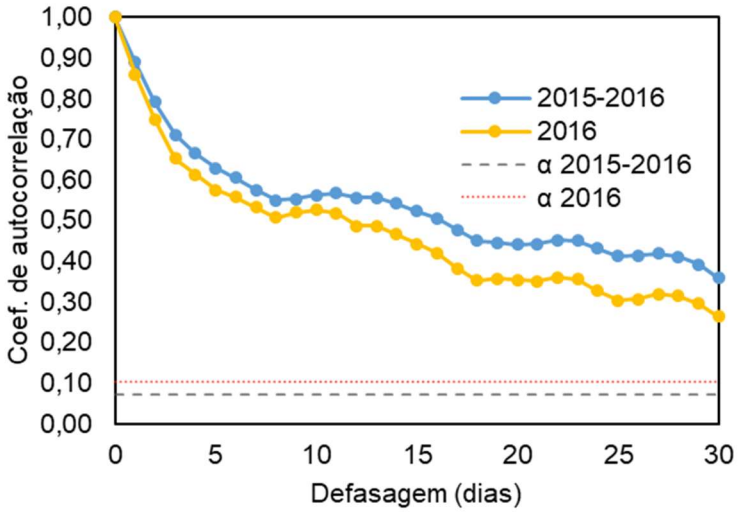

Figura 71: Função de autocorrelação da estação Q5

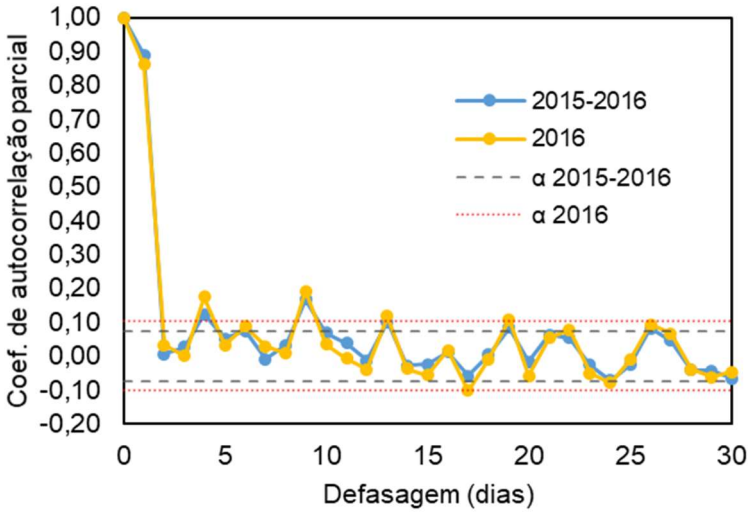

Figura 72: Função de autocorrelação parcial da estação Q5

Diante disso, cinco configurações de dados de entrada foram escolhidas para realizar as simulações de Q5. A Tabela 17 mostra tais vetores de entrada, além dos respectivos índices estatísticos. Verifica-se que todos os modelos apresentaram resultados insatisfatórios para a simulação de hidrogramas. Os modelos com precipitações defasadas entre 0-11 dias e vazões defasadas entre 1-4 dias ($P_{0:11}$ $Q_{1:4}$) apresentaram os melhores resultados. Os critérios de desempenho obtidos foi de NSE de 0,443, RSR de 0,75, PBIAS de 4,31%, RMSE de 13,21 m^3s^{-1} e MAPE de 32,75%.

Tabela 17: Desempenhos dos modelos propostos para simulação de Q5

Modelo	NO	NSE	RSR	RMSE	PBIAS	MAPE
$P5_{0:3}Q5_{1}$	4	0,378	0,79	13,96	-1,79	35,26
$P5_{0:11}Q5_{1:4}$	4	0,443	0,75	13,21	4,31	32,75
$P5_{0:11}Q5_{1:9}$	4	0,359	0,8	14,17	-8,01	27,61
$P5_{0:13}Q5_{1:13}$	4	0,271	0,85	15,11	-12,48	27,94
$P5_{0:19}Q5_{1:19}$	4	0,359	0,8	14,16	-9,23	27,48

NO – Neurônios Ocultos; NSE – Coeficiente de Eficiência de Nash-Sutcliffe; RSR – Raiz do erro quadrático médio padronizado; PBIAS – Viés percentual; MAPE – Erro absoluto percentual médio

A Figura 73 mostra os hidrogramas observados e simulados, e a Figura 74 mostra a dispersão entre as vazões observadas e simuladas para Q_5. É evidente que o modelo apresenta dificuldades na simulação de vazões acima dos 50 m^3s^{-1}. Os meses com os menores erros percentuais foram janeiro (MAPE de 19,60%), fevereiro (MAPE de 21,48%) e março (MAPE de 17,99%). A maior vazão da série histórica ocorreu no dia 25/11/2018, sendo igual a 112,89 m^3s^{-1} e foi simulada com valor de 26,08 m^3s^{-1}, ocasionando subestimação de 76,90%. Além disso, a vazão de 110,41 m^3s^{-1}, observada em 12/04/2018, foi simulada com 47,11 m^3s^{-1}, acarretando subestimação de 57,33%. O maior erro percentual de simulação é evidenciado no dia 19/09/2018, onde a vazão observada é de 10,49 m^3s^{-1} e a vazão simulada é de 23,97 m^3s^{-1}, gerando erro absoluto percentual de 128,53%. Outro exemplo é a vazão do dia 16/09/2018, observada com 10,19 m^3s^{-1} e simulada com 22,28 m^3s^{-1}, com erro de 118,55%. Por outro lado, os menores erros são apresentados em valores próximos à média de vazões observada em 2018 (27,36 m^3s^{-1}), como, por exemplo, a ocorrida em 21/05/2018, observada com 27,78 m^3s^{-1} e simulada com 28,26 m^3s^{-1}, gerando erro absoluto percentual de 1,73%.

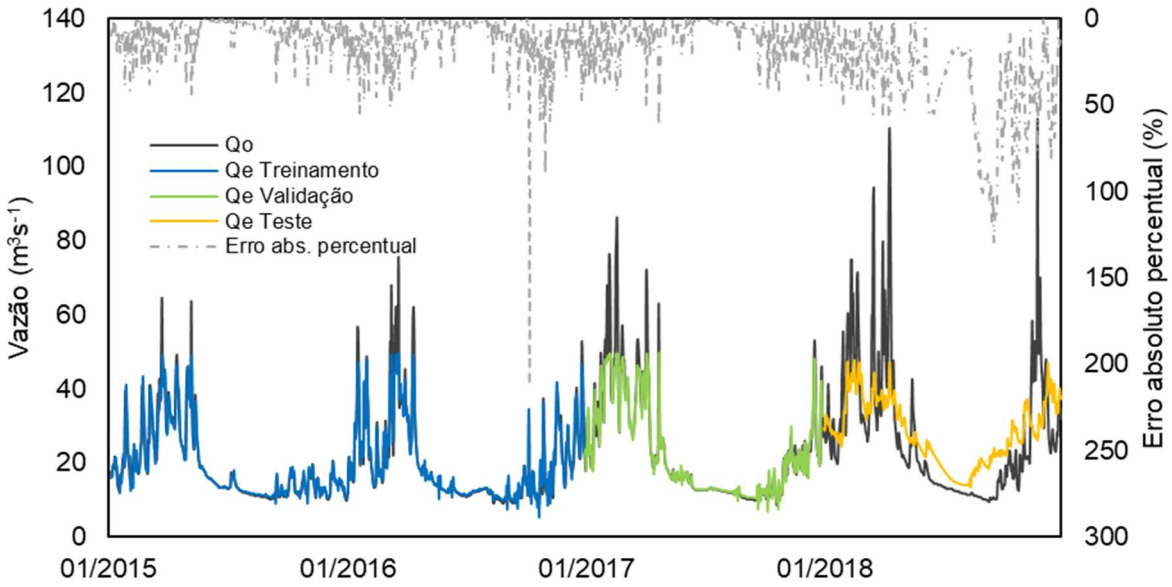

Figura 73: Hidrogramas observados e simulados para a estação Q5

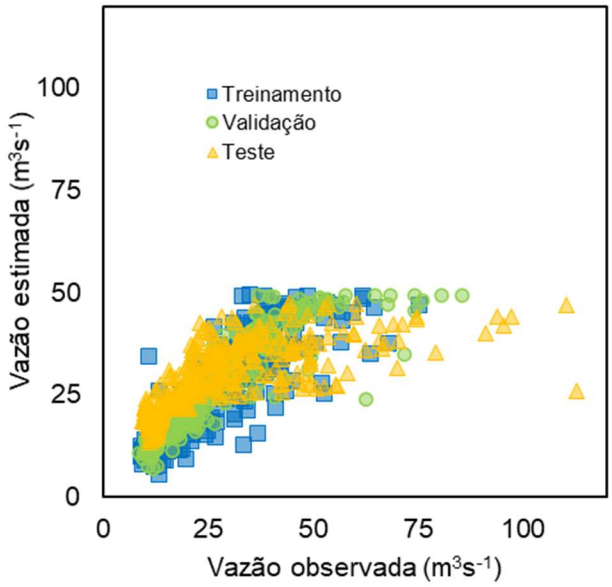

Figura 74: Dispersão entre os dados observados e simulados para a estação Q5

A Figura 75 apresenta os resultados da função de autocorrelação de erro para as simulações de vazões em Q5. É verificado que há presença de autocorrelações significativas nos erros, inseridos no nível de significância de α=5%, com correlação linear do erro defasado de primeira ordem de 0,81. Os erros apresentam média de 0,90 e desvio padrão de 13,12 m^3s^{-1}.

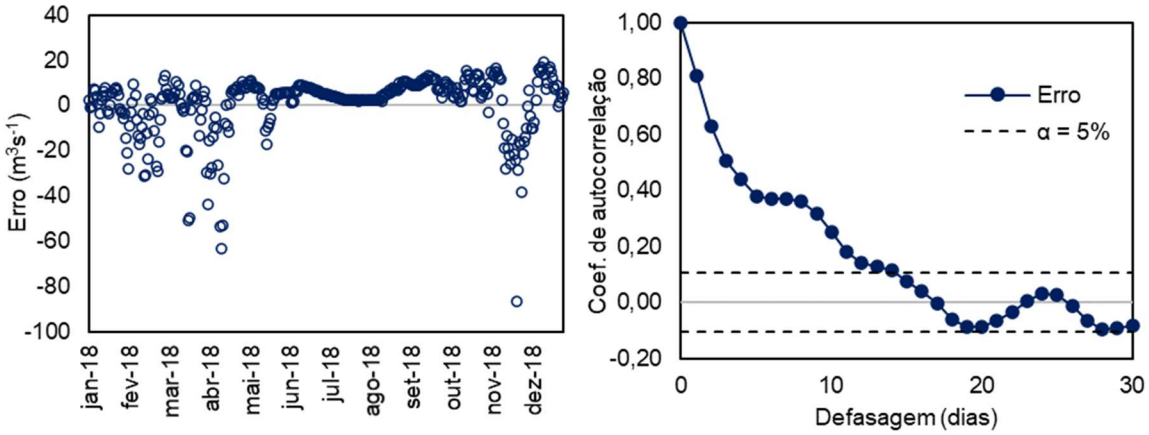

Figura 75: Série temporal e função de autocorrelação do erro em Q5

Notou-se que as curvas de permanência apresentaram melhor ajuste com precipitações defasadas entre 0-11 dias e vazões defasadas entre 1-9 dias ($P5_{0:11}Q5_{1:9}$). A Figura 76 mostra tais simulações. Assim, a RNN-NARX série-paralelo obteve índices NSE de 0,780, RSR de 0,47, PBIAS de -8,01%, RMSE de 8,29 m^3s^{-1} e MAPE de 8,71%. Mesmo assim, ainda é verificado que os erros mais discrepantes ainda estão em vazões acima dos 50 m^3s^{-1}, acima das vazões de referência Q_{20}. As vazões de referência Q_{90} e Q_{95}, observadas com 11,40 m^3s^{-1} e 10,49 m^3s^{-1}, respectivamente, foram simuladas com 10,61 m^3s^{-1} e 10,58 m^3s^{-1}.

Além do aspecto de discretização diária de vazões em pequenas bacias, a variabilidade espacial entre chuva monitorada na estação meteorológica e a efetivamente ocorrida na bacia de Q5 pode ser outro motivo destes resultados, uma vez que há 132 km separando-as. Em outras palavras, a precipitação ocorrida no interior da bacia, que geram os efeitos do ciclo hidrológico, podem não ser as mesmas observadas no ponto de monitoramento. Além disso, o melhor ajuste das vazões mínimas pode estar associado ao padrão sazonal encontrado nas vazões defasadas, ou seja, ter maior dependência da autorregressão.

A redução do RMSE em função das interações do treinamento supervisionado do modelo $P_{0:11}Q_{1:9}$ é exibida na Figura 77. Embora, o RMSE de treinamento tenha diminuído continuamente, os RMSEs de validação e teste não acompanham tal comportamento, em que estes dois últimos conjuntos alcançaram o menor erro na iteração 15. Logo, utilizar os parâmetros pesos e biases, além desta iteração, é um erro, pois configuraria *overffiting*. Sendo assim, a configuração de RNN-NARX escolhida para simulação é a da iteração 15.

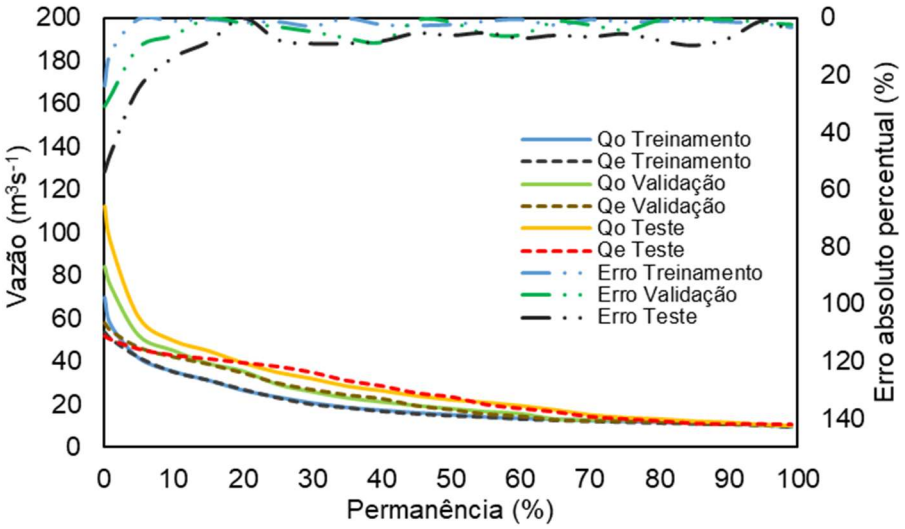

Figura 76: Curvas de permanência observadas e simuladas para a estação Q5

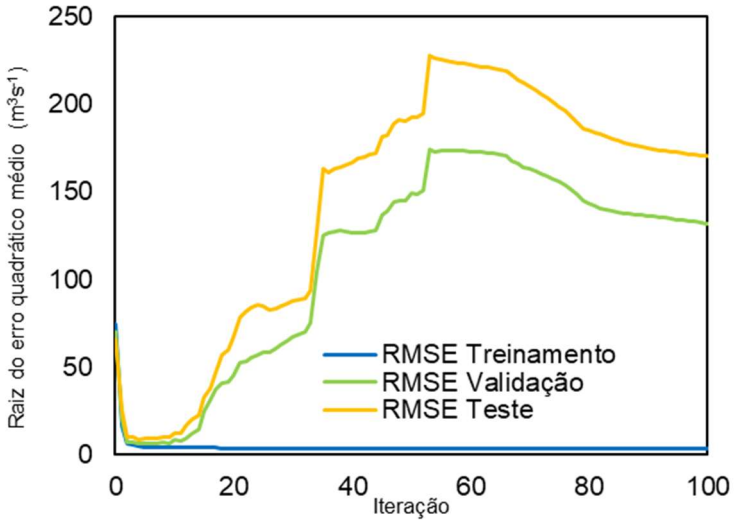

Figura 77: Raiz do erro quadrático médio em função das iterações do modelo para simulação de vazões da estação Q5

A Figura 78 mostra a contribuição relativa das variáveis para simulação das vazões em Q_5. A dependência autorregressiva é notável, pois 67,28% da contribuição relativa é proveniente das vazões defasadas. A maior contribuição é realizada pela vazão defasada $Q_{5(t-1)}$, com 21,93%. A maior contribuição de uma variável de precipitação é a $P_{5(t-3)}$, com 5,52%. Este resultado reforça a ideia de que a distância de 132 km entre a bacia e a estação meteorológica pode estar dificultando simulações mais precisas, principalmente de picos de vazões.

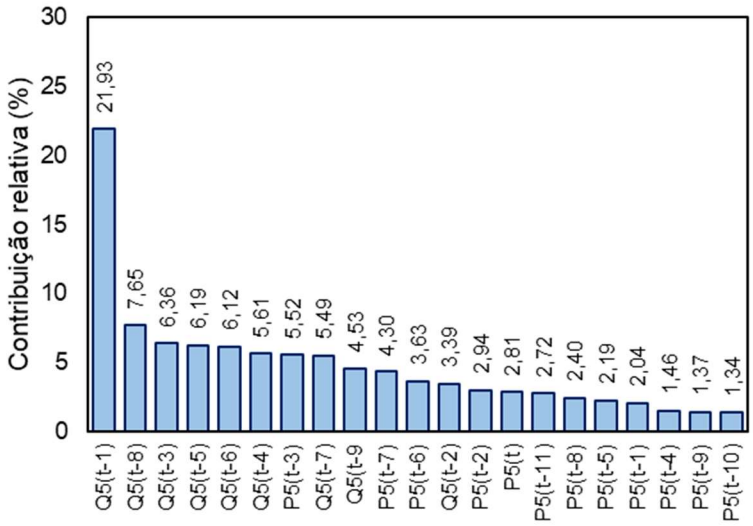

Figura 78: Contribuição relativa dos dados de entrada do modelo para simulação de vazões da estação Q5

A análise de sensibilidade do modelo para simulação de Q5 é apresentada na Tabela 18. É possível notar que não há diferenças significativas nos índices de desempenho ao realizar o treinamento supervisionado com diferentes anos hidrológicos. Ainda assim, todos os modelos são classificados como insatisfatórios para simulação dos hidrogramas (Figura 79). Utilizando o ano de 2016, a curva de permanência (Figura 80) obteve índice NSE de 0,728, RSR de 0,52, PBIAS de -6,45%, RMSE de 9,23 m^3s^{-1} e MAPE de 14,34%, porém, as vazões acima de Q_{20} ainda não são simuladas de forma satisfatória.

Tabela 18: Sensibilidade do modelo à diferentes resoluções temporais para simulações de Q5

Período	NO	NSE	RSR	RMSE	PBIAS	MAPE
2015-2016	4	0,359	0,8	14,17	-8,01	27,61
2016	4	0,324	0,82	14,55	-6,45	29,84

NO – Neurônios Ocultos; NSE – Coeficiente de Eficiência de Nash-Sutcliffe; RSR – Raiz do erro quadrático médio padronizado; PBias – Viés percentual; MAPE – Erro absoluto percentual médio

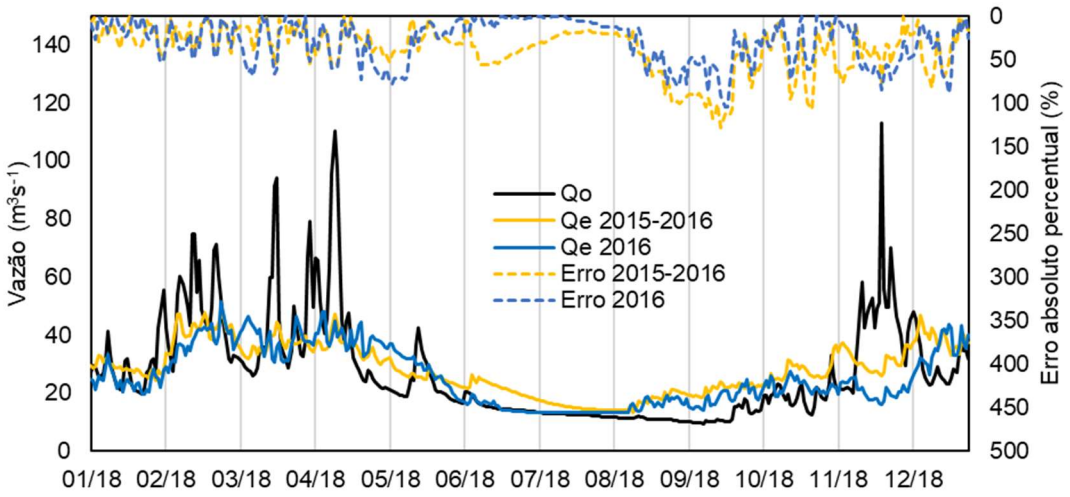

Figura 79: Hidrogramas observado e simulados para Q5, utilizando diferentes intervalos temporais

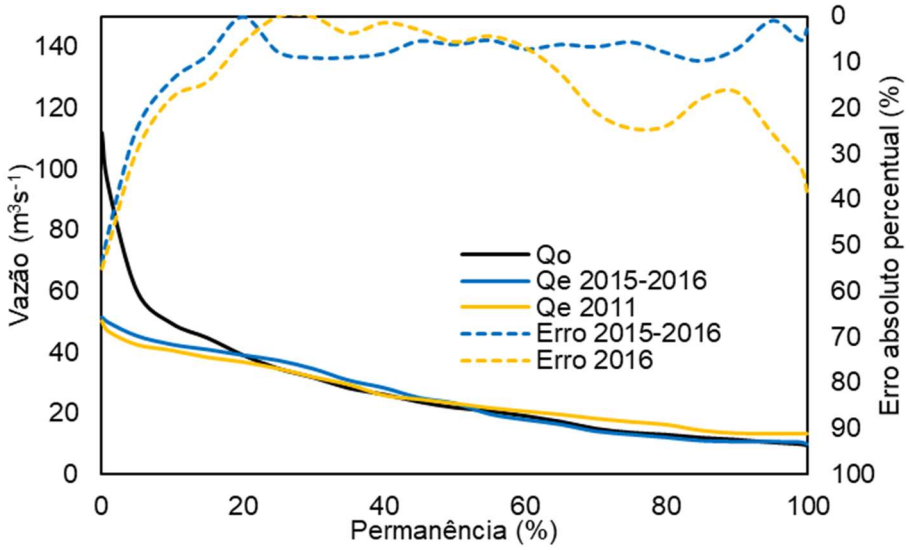

Figura 80: Curvas de permanência observada e simuladas para Q5, utilizando diferentes intervalos temporais

4.6 SÍNTESE DOS RESULTADOS

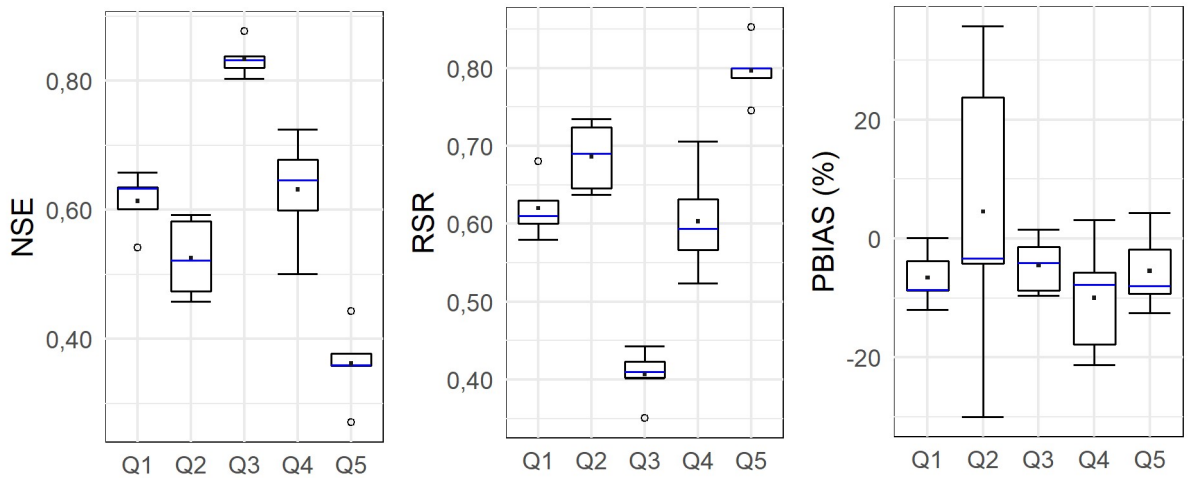

Figura 81: Boxplots dos índices de critérios obtidos

Os resultados demonstram que a capacidade de autocorrelação das vazões diárias, uma característica estocástica muitas vezes atribuída à retenção de água no interior da bacia hidrográfica, é fator preponderante para as simulações via RNN-NARX (como ocorre em Q3). Além disso, com auxílio das precipitações defasadas, as simulações descreveram bem, principalmente, a recessão dos hidrogramas. Por esta última variável se tratar de uma variável exógena, passível de mudança de comportamento ao longo do tempo, as incertezas na simulação de longos períodos é reduzida, uma vez que a principal forçante de água no sistema é a chuva. As pequenas bacias hidrográficas aqui analisadas apresentam características semelhantes, como a declividade variando entre o plano e o suave ondulado, tem solos com alto potencial de infiltração. Além disso, a análise de sensibilidade mostrou que dois anos hidrológicos é suficiente para o treinamento supervisionado das RNN-NARX.

Para simulação de vazões de cheias, a discretização diária das vazões em pequenas bacias e a variabilidade espacial da precipitação, principalmente em anos de anomalias meteorológicas, devem ser melhor analisadas. O principal exemplo disto, é a bacia de Q5, na qual as vazões podem ter sido simuladas majoritariamente pela autorregressão, conforme a interpretação da CR de Garson.

5 CONCLUSÃO

A análise exploratória baseada nas funções de correlação cruzada e autocorrelação parcial mostraram potenciais vetores de entrada para as RNN-NARX. No entanto, esta abordagem ainda é subjetiva para identificação de sistemas

dinâmicos não lineares, uma vez que, a relação entre as variáveis não é totalmente compreendida por correlações lineares. Este fato é facilmente interpretado quando há a comparação entre os vetores de entrada propostos e a contribuição relativa de Garson de cada modelo.

As simulações de vazões dos modelos hidrológicos baseados em RNN-NARX, orientados às pequenas bacias amazônicas, variou entre o insatisfatório e o muito bom. É observada, também, a tendência de subestimar vazões. A contribuição relativa de Garson reforça o argumento de que as vazões diárias de pequenas bacias amazônicas podem ser simuladas, principalmente, a partir de vazões e precipitações defasadas, pois auxiliam a descrever o comportamento dos escoamentos subsuperficial e subterrâneo.

As principais fontes de incertezas dos modelos são provenientes de variabilidade espacial da precipitação, principalmente em anos de anomalias meteorológicas, e de discretização, que resulta na dificuldade em descrever os picos de vazões em pequenas bacias. No entanto, os resultados mais expressivos para simulação de 365 dias são provenientes dos potenciais autorregressivos de cada bacia, que é uma característica estocástica, muitas vezes atribuída à capacidade de estocagem de água no solo ou regularização de vazões por intervenção humana.

A análise de sensibilidade dos modelos RNN-NARX frente a diferentes intervalos de treinamento mostrou que, de modo geral, são alcançados resultados satisfatórios para quatro das cinco bacias hidrográficas (Marambaia, Fazenda Craveiro, Arraial e Próximo à Colinas do Tocantins), a partir da implementação de 2 anos para o treinamento supervisionado das RNN-NARX. No contexto prático, os modelos aqui discutidos são aplicáveis, principalmente, para simulação de vazões mínimas. Assim, podem ser aplicados para avaliação de disponibilidade hídrica.

Sugere-se, como trabalhos futuros, a implementação de técnicas de pré-tratamento dos dados de entrada, tal como a Transformada de Wavelet Discreta. Além disso, a utilização de algoritmos genéticos podem melhorar a escolha dos vetores de entrada ideais, ou até mesmo melhorar a determinação de parâmetros pesos e bias ideais, na etapa de treinamento.

REFERÊNCIAS

ANA – **Agência Nacional de Águas**. Sistema de Informações Hidrológicas - HidroWeb. Disponível em: http://hidroweb.ana.gov.br/. Acesso em março de 2021.

ANDRÉASSIAN, V.; PERRIN, C.; MICHEL, C. Impact of imperfect potential evapotranspiration knowledge on the efficiency and parameters of watershed models. **Journal of Hydrology**, v. 286, n. 1-4, p. 19-35, 2004.

ARNOLD, J. G.; MORIASI, D. N.; GASSMAN, P. W.; ABBASPOUR, K. C.; WHITE, M. J.; SRINIVASAN, R.; SANTHI, C.; HARMEL, R.D.; VAN GRIENSVEN, A.; KANNAN, N.; VAN LIEW, M.W.; JHA, M. K. SWAT: Model use, calibration, and validation. **Transactions of the ASABE**, v. 55, n. 4, p. 1491-1508, 2012.

BANIHABIB, M. E.; BANDARI, R.; PERALTA, R. C. Auto-regressive neural-network models for long lead-time forecasting of daily flow. **Water resources management**, v. 33, n. 1, p. 159-172, 2019.

BAYER, D. M.; CASTRO, N. M. R.; BAYER, F. M. Modelagem e previsão de vazões médias mensais do rio Potiribu utilizando modelos de séries temporais. **Revista Brasileira de Recursos Hídricos**, v. 17, n. 2, p. 229-239, 2012.

BEALE, M. H.; HAGAN, M. T.; DEMUTH, H. B. Deep Learning Toolbox User's Guide. **The Mathworks Inc. Herborn, MA, USA**, 2016.

BEHRANGI, A.; KHAKBAZ, B.; JAW, T. C.; AGHAKOUCHAK, A.; HSU, K.; SOROOSHIAN, S. Hydrologic evaluation of satellite precipitation products over a mid-size basin. **Journal of Hydrology**, v. 397, n. 3-4, p. 225-237, 2011.

BEVEN, K. **Rainfall-runoff modelling: the primer**. Second edition. UK: Wiley-Blackwell. 457 p., 2012.

BEVEN, K.; WOOD, E. F.; SIVAPALAN, M. On hydrological heterogeneity—catchment morphology and catchment response. **Journal of hydrology**, v. 100, n. 1-3, p. 353-375, 1988.

BLANCO, C. J. C.; SANTOS, S. S. M.; QUINTAS, M. C.; VINAGRE, M. V. A.; MESQUITA, A. L. A. Contribution to hydrological modelling of small Amazonian catchments: application of rainfall–runoff models to simulate flow duration curves. **Hydrological Sciences Journal**, v. 58, n. 7, p. 1423-1433, 2013.

BLANCO, C. J.; SECRETAN, Y.; FAVRE, A. C.; SLIVITZKY, M. Modèle pluie–débit pour la simulation de courbes de débits classés sur des petits bassins non jaugés de l'Amazonie. **Canadian Journal of Civil Engineering**, v. 32, n. 5, p. 803-811, 2005.

BLANCO, C. J.C.; SECRETAN, Y.; FAVRE, A. Análise, aplicação e transposição de um modelo chuva-vazão para simulação de curvas de permanência de pequenas bacias da Amazônia. **Revista Brasileira de Recursos Hídricos**, v. 12, n. 1, p. 205-216, 2007.

BLÖSCHL, G.; GRAYSON, R. B.; SIVAPALAN, M. On the representative elementary area (REA) concept and its utility for distributed rainfall-runoff modelling. **Hydrological Processes**, v. 9, n. 3-4, p. 313-330, 1995.

BUTTLE, J. M. Fundamentals of small catchment hydrology. In: **Isotope tracers in catchment hydrology**. Elsevier, 1998. p. 1-49.

CARIELLO, B.; BLANCO, C.; GOMIDE, I.; BARBOSA, A.; SANTOS, D. Application and Calibration of the Modified Universal Soil Loss Equation to estimate the sediment yield in a small Amazon catchment. **Sylwan Journal**, v. 347, 2014.

CHANDWANI, V.; VYAS, S. K.; AGRAWAL, V.; SHARMA, G. Soft computing approach for rainfall-runoff modelling: a review. **Aquatic Procedia**, v. 4, p. 1054-1061, 2015.

CHANG, F. J.; CHANG, L. C.; HUANG, C. W.; KAO, I. F. Prediction of monthly regional groundwater levels through hybrid soft-computing techniques. **Journal of Hydrology**, v. 541, p. 965-976, 2016.

CHEN, S.; BILLINGS, S. A.; GRANT, P. M. Non-linear system identification using neural networks. **International journal of control**, v. 51, n. 6, p. 1191-1214, 1990.

CLARK, M. P.; KAVETSKI, D.; FENICIA, F. Pursuing the method of multiple working hypotheses for hydrological modeling. **Water Resources Research**, v. 47, n. 9, 2011.

COLLISCHONN, B.; COLLISCHONN, W.; TUCCI, C. E. M. Daily hydrological modeling in the Amazon basin using TRMM rainfall estimates. **Journal of Hydrology**, v. 360, n. 1-4, p. 207-216, 2008.

CORDEIRO, D. G.; RAMOS, M. R.; ULHMANN, A.; CUSTODIO, D. P.; FREIRE, T. M.; DOS SANTOS CAMPOS, L. Análise do desempenho do modelo de Kostiakov na determinação da velocidade básica de infiltração em plintossolo pétrico concrecionário. **Revista Brasileira de Agropecuária Sustentável**, v. 11, n. 1, p. 135-142, 2021.

COULTHARD, T. J.; LEWIN, J.; MACKLIN, M. G. Modelling differential catchment response to environmental change. **Geomorphology**, v. 69, n. 1-4, p. 222-241, 2005.

DANIELL, T. M. Neural networks. Applications in hydrology and water resources engineering. In: **National Conference Publication- Institute of Engineers**. Australia. 1991.

DEBASTIANI, A. B.; SILVA, R. D.; RAFAELI NETO, S. L. Efficiency of MLP architecture on closed-loop mode for the simulation of a hydrological system. **Revista Brasileira de Recursos Hídricos**, v. 21, n. 4, p. 821-831, 2016.

DEVIA, G. K.; GANASRI, B. P.; DWARAKISH, G. S. A review on hydrological models. **Aquatic Procedia**, v. 4, p. 1001-1007, 2015.

DIACONESCU, Eugen. The use of NARX neural networks to predict chaotic time series. **Wseas Transactions on computer research**, v. 3, n. 3, p. 182-191, 2008.

EMBRAPA - **Empresa Brasileira de Pesquisa Agropecuária**. Centro Nacional de Pesquisa de Solos. Sistema brasileiro de classificação de solos. Rio de Janeiro, 1999. 412p.

ESPINOZA, J. C.; MARENGO, J. A.; SCHONGART, J.; JIMENEZ, J. C. The new historical flood of 2021 in the Amazon River compared to major floods of the 21st century: Atmospheric features in the context of the intensification of floods. **Weather and Climate Extremes**, p. 100406, 2021.

ESTRANY, J.; GARCIA, C.; BATALLA, R. J. Hydrological response of a small Mediterranean agricultural catchment. **Journal of Hydrology**, v. 380, n. 1-2, p. 180-190, 2010.

FEKETE, B. M.; VÖRÖSMARTY, C. J.; GRABS, W. High-resolution fields of global runoff combining observed river discharge and simulated water balances. **Global Biogeochemical Cycles,** v. 16, n. 3, p. 15-1-15-10, 2002.

FINNOFF, W.; HERGERT, F.; ZIMMERMANN, H. G. Improving model selection by nonconvergent methods. **Neural Networks**, v. 6, n. 6, p. 771-783, 1993.

GARSON, D. G. Interpreting neural network connection weights. **Artificial Intelligence Expert**, v. 6, p. 47–51, 1991.

GEIGER, R. **Landolt-Börnstein–Zahlenwerte und Funktionen aus Physik, Chemie, Astronomie, Geophysik und Technik, alte Serie Vol. 3**. Ch. Klassifikation der Klimate nach W. Köppen. Springer, Berlin, p. 603-607, 1954.

GHOSE, D.; DAS, U.; ROY, P. Modeling response of runoff and evapotranspiration for predicting water table depth in arid region using dynamic recurrent neural network. **Groundwater for Sustainable Development**, v. 6, p. 263-269, 2018.

GRAYSON, R. B.; MOORE, I. D.; MCMAHON, T. A. Physically based hydrologic modeling: 2. Is the concept realistic? **Water resources research**, v. 28, n. 10, p. 2659-2666, 1992.

GUZMAN, S. M.; PAZ, J. O.; TAGERT; M. L. M.; MERCER, A. E. Evaluation of seasonally classified inputs for the prediction of daily groundwater levels: NARX networks vs support vector machines. **Environmental Modeling & Assessment**, v. 24, n. 2, p. 223-234, 2019.

GYASI-AGYEI, Y.; WILLGOOSE, G.; DE TROCH, F. P. Effects of vertical resolution and map scale of digital elevation models on geomorphological parameters used in hydrology. **Hydrological Processes**, v. 9, n. 3-4, p. 363-382, 1995.3

HAYKIN, S. **Redes neurais: princípios e prática**. Bookman Editora, 2001.

HEBB, D. O. The organization of behavior; a neuropsychological theory. **A Wiley Book in Clinical Psychology**, v. 62, p. 78, 1949.

HEWLETT, J. D.; HIBBERT, A. R. Factors affecting the response of small watersheds to precipitation in humid areas. **Forest hydrology**, v. 1, p. 275-290, 1967.

HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. **Neural networks**, v. 2, n. 5, p. 359-366, 1989.

HUGGINS, L. F.; MONKE, E. J. A mathematical model for simulating the hydrologic response of a watershed. **Water Resources Research**, v. 4, n. 3, p. 529-539, 1968.

IBGE – **Instituto Brasileiro De Geografia e Estatística**. Vetores geoespaciais da pedologia brasileira na escala 1:250.000. Disponível em: https://www.ibge.gov.br/geociencias/informacoes-ambientais/pedologia/10871-pedologia.html?=&t=downloads. Acesso em março de 2021.

INPE – **Instituto Nacional de Pesquisas Espaciais**. Condições atuais dos fenômenos El Niño Oscilação Sul (ENOS). Disponível em: http://enos.cptec.inpe.br/. Acesso em dezembro de 2021.

INPE – **Instituto Nacional de Pesquisas Espaciais**. Mapeamento do uso e cobertura da terra no Cerrado, Projeto TerraClass Cerrado. 2018. Disponível em: https://www.terraclass.gov.br/geoportal-cerrado/. Acesso em dezembro de 2021.

INPE – **Instituto Nacional de Pesquisas Espaciais**. Mapeamento do uso e cobertura da terra na Amazônia Legal, Projeto TerraClass Amazônia Legal. 2014. Disponível em: https://www.terraclass.gov.br/geoportal-aml/. Acesso em dezembro de 2021.

ISHIHARA, J. H.; FERNANDES, L.; DUARTE, A. A. A. M.; LOUREIRO, G. E. Avaliação do monitoramento pluviométrico da Amazônia Legal. **Engenharia Ambiental**, v. 10, n. 3, p. 132-144, 2013.

IZADY, A.; DAVARY, K.; ALIZADEH, A.; MOGHADDAM NIA, A.; ZIAEI, A. N.; HASHEMINIA, S. M. Application of NN-ARX model to predict groundwater levels in the Neishaboor Plain, Iran. **Water Resources Management**, v. 27, n. 14, p. 4773-4794, 2013.

KOPPEN, W. Das geographische system der klimat. **Handbuch der klimatologie**, p. 46, 1936.

KOVÁCS, Z. L. **Redes neurais artificiais. 3° edição**. São Paulo: Editora Livraria da Física, 2002.

LEE, W.; RESDI, T. A. Simultaneous hydrological prediction at multiple gauging stations using the NARX network for Kemaman catchment, Terengganu, Malaysia. **Hydrological Sciences Journal**, v. 61, n. 16, p. 2930-2945, 2016.

LIN, T.; HORNE, B. G.; GILES, C. L. How embedded memory in recurrent neural network architectures helps learning long-term temporal dependencies. **Neural Networks**, v. 11, n. 5, p. 861-868, 1998.

LINSLEY, Ray K.; CRAWFORD, Norman H. Computation of a synthetic streamflow record on a digital computer. **International Association of Scientific Hydrology**, v. 5, p. 526-538, 1960.

LIPTON, Z. C.; BERKOWITZ, J.; ELKAN, C. A critical review of recurrent neural networks for sequence learning. **arXiv preprint arXiv: 1506.00019**, 2015.

MAIER, H. R.; DANDY, G. C. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. **Environmental modelling & software**, v. 15, n. 1, p. 101-124, 2000.

MAIER, H. R.; JAIN, A.; DANDY, G. C.; SUDHEER, K. P. Methods used for the development of neural networks for the prediction of water resource variables in river

systems: Current status and future directions. **Environmental modelling & software**, v. 25, n. 8, p. 891-909, 2010.

MASSMANN, C.; WAGENER, T.; HOLZMANN, H. A new approach to visualizing time-varying sensitivity indices for environmental model diagnostics across evaluation time-scales. **Environmental Modelling & Software**, v. 51, p. 190-194, 2014.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, v. 5, n. 4, p. 115-133, 1943.

MELSEN, L.; TEULING, A.; TORFS, P.; ZAPPA, M.; MIZUKAMI, N.; CLARK, M.; UIJLENHOET, R. Representation of spatial and temporal variability in large-domain hydrological models: case study for a mesoscale pre-Alpine basin. **Hydrology and Earth System Sciences**, v. 20, n. 6, p. 2207-2226, 2016.

MENDONÇA, L M; GOMIDE, I.S.; SOUSA, J.V.; BLANCO, C.J.C. Modelagem chuva-vazão via redes neurais artificiais para simulação de vazões de uma bacia hidrográfica da Amazônia. **Revista de Gestão de Água da América Latina**, v. 18, n. 2021, 2021.

MENEZES JR, J. M. P.; BARRETO, G. A. Long-term time series prediction with the NARX network: An empirical evaluation. **Neurocomputing**, v. 71, n. 16-18, p. 3335-3343, 2008.

MINSKY, M.; PAPERT, S. **Perceptrons: An Introduction to Computational Geometry**. Cambridge. MIT Press, 1969.

MIRUS, B. B.; LOAGUE, K. How runoff begins (and ends): Characterizing hydrologic response at the catchment scale. **Water Resources Research**, v. 49, n. 5, p. 2987-3006, 2013.

MORAES, B. C. D.; COSTA, J. M. N. D.; COSTA, A. C. L. D.; COSTA, M. H. Variação espacial e temporal da precipitação no estado do Pará. **Acta Amazônica**, v. 35, p. 207-214, 2005.

MORIASI, D. N.; ARNOLD, J. G.; VAN LIEW, M. W.; BINGNER, R. L., HARMEL, R. D.; VEITH, T. L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. **Transactions of the ASABE**, v. 50, n. 3, p. 885-900, 2007.

MOSAVI, Amir; OZTURK, Pinar; CHAU, Kwok-wing. Flood prediction using machine learning models: Literature review. **Water**, v. 10, n. 11, p. 1536, 2018.

MULVANEY, T. J. On the use of self-registering rain and flood gauges in making observations of the relations of rainfall and flood discharges in a given catchment. **Proceedings of the institution of Civil Engineers of Ireland**, v. 4, p. 19-31, 1851.

NASH, J. E.; SUTCLIFFE, J. V. River flow forecasting through conceptual models: part I - A discussion of principles. **Journal of hydrology**, v. 10, n. 3, p. 282-290, 1970.

LLERAS, E.; KIRKBRIDE JR, J. H. Alguns aspectos da vegetação da Serra do Cachimbo. **Acta Amazônica**, v. 8, n. 1, p. 51-65, 1978.

OYEBODE, O.; STRETCH, D. Neural network modeling of hydrological systems: A review of implementation techniques. **Natural Resource Modeling**, v. 32, n. 1, p. e12189, 2019.

PAIVA, R. C. D.; COLLISCHONN, W.; BONET, M. P.; GONÇALVES, L. G. G. On the sources of hydrological prediction uncertainty in the Amazon**. Hydrology and Earth System Sciences**, v. 16, n. 9, p. 3127-3137, 2012.

PECHLIVANIDIS, I. G.; JACKSON, B. M.; MCINTYRE, N. R.; WHEATER, H. S. Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. **Global NEST Journal**, v. 13, n. 3, p. 193-214, 2011.

PECK, R.; OLSEN, C.; DEVORE, J. L. **Introduction to statistics and data analysis**. Cengage Learning, 2015.

PEEL, M. C.; MCMAHON, T. A. Historical development of rainfall-runoff modeling. **Wiley Interdisciplinary Reviews: Water**, v. 7, n. 5, p. e1471, 2020.

PRAHARAJ, S.; MOHANTY, P. K.; SAHOO, B. C. Quantification of Error in Estimation of Reference Crop Evapotranspiration by Class A Pan Evaporimeter and Its Correction. In: **Hydrologic Modeling**. Springer, Singapore, 2018. p. 85-96.

PRECHELT, Lutz. Automatic early stopping using cross validation: quantifying the criteria. **Neural Networks**, v. 11, n. 4, p. 761-767, 1998.

QUINTAS, M. C.; BLANCO, C. J. C.; MESQUITA, A L. A. A non-linear rainfall-runoff model applied to Amazon small catchments with limited data to simulate the flow duration curves. **International Journal of Hydrology Science and Technology**, v. 1, n. 1-2, p. 19-36, 2011.

REZAEIAN ZADEH, M.; AMIN, S.; KHALILI, D.; SINGH, V. P. Daily outflow prediction by multi-layer perceptron with logistic sigmoid and tangent sigmoid activation functions. **Water resources management**, v. 24, n. 11, p. 2673-2688, 2010.

RIBEIRO NETO, A.; SILVA, R. C. V.; COLLISCHONN, W.; TUCCI, C. E. Simulação na bacia Amazônica com dados limitados: rio Madeira**. Revista Brasileira de Recursos Hídricos**, v. 13, n. 3, p. 47-58, 2008.

ROGERS, W. F. New concept in hydrograph analysis. **Water Resources Research**, v. 8, n. 4, p. 973-981, 1972.

ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. D. M.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711-728, 2013.

ROLDÃO, A. F.; FERREIRA, V. O. **Climatologia do Estado do Tocantins – Brasil**. Caderno de Geografia. v.29, n. 59, 2019. DOI:10.5752/p.2318-2962.2019v29n59p1161.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, v. 65, n. 6, p. 386, 1958.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, v. 323, n. 6088, p. 533-536, 1986.

SAHOO, S; JHA, M. K. Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment. **Hydrogeology Journal**, v. 21, n. 8, p. 1865-1887, 2013.

SANTANA, L. R.; BLANCO, C. J. C. Contribution to the classification of small catchments according to the drainage area. **International Journal of River Basin Management**, p. 1-12, 2020.

SARTORI, A.; LOMBARDI NETO, F.; GENOVEZ, A. M. Classificação hidrológica de solos brasileiros para a estimativa da chuva excedente com o método do Serviço de Conservação do Solo dos Estados Unidos Parte 1: Classificação. **Revista Brasileira de Recursos Hídricos**, v. 10, n. 4, p. 05-18, 2005.

SAVENIJE, H. H. G. Equifinality, a blessing in disguise? **Hydrological processes**, v. 15, n. 14, p. 2835-2838, 2001.

SHEN, H. Y.; CHANG, L. C. Online multistep-ahead inundation depth forecasts by recurrent NARX networks. **Hydrology and Earth System Sciences**, v. 17, n. 3, p. 935-945, 2013.

SHERMAN, Jack. Crystal Energies of Ionic Compounds and Thermochemical Applications. **Chemical Reviews**, v. 11, n. 1, p. 93-170, 1932.

SINGH, J.; KNAPP, H. V.; ARNOLD, J. G.; DEMISSIE, M. Hydrological modeling of the Iroquois river watershed using HSPF and SWAT 1. **JAWRA Journal of the American Water Resources Association**, v. 41, n. 2, p. 343-360, 2005.

SIT, M., DEMIRAY, B.; XIANG, Z.; EWING, G.; SERMET, Y.; DEMIR, I. A comprehensive review of deep learning applications in hydrology and water resources. **Water Science and Technology**, v. 82, n. 12, p. 2635-2670, 2020.

SIT, M., DEMIRAY, B.; XIANG, Z.; EWING, G.; SERMET, Y.; DEMIR, I. A comprehensive review of deep learning applications in hydrology and water resources. **Water Science and Technology**, v. 82, n. 12, p. 2635-2670, 2020.

SITTERSON, J.; KNIGHTES, C.; PARMAR, R.; WOLFE, K.; AVANT, B.; MUCHE, M. **An overview of rainfall-runoff model types**. In International Congress on Environmental Modelling and Software. 2018.

SOARES, P. A.; PINHEIRO, A.; SOARES, K. H.; ZUCCO, E. Estimativa da disponibilidade hídrica em pequenas bacias hidrográficas com escassez de dados fluviométricos. **Revista de estudos ambientais**, v. 12, n. 1, p. 29-38, 2010.

SOLA, J.; SEVILLA, J. Importance of input data normalization for the application of neural networks to complex industrial problems. **IEEE Transactions on nuclear science**, v. 44, n. 3, p. 1464-1468, 1997.

SONG, X. M.; ZHAN, C. S.; XIA, J.; ZHANG, Y. Y. Methodology and application of parameter uncertainty quantification in watershed hydrological models. **China Water Power Press, Beijing**, 2014.

SONG, X.; ZHANG, J.; ZHAN, C.; XUAN, Y.; YE, M.; XU, C. Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications. **Journal of hydrology**, v. 523, p. 739-757, 2015.

SORJAMAA, A.; HAO, J.; REYHANI; N.; JI, Y.; LENDASSE, A. Methodology for long-term prediction of time series. **Neurocomputing**, v. 70, n. 16-18, p. 2861-2869, 2007.

STATHAKIS, D. How many hidden layers and nodes? **International Journal of Remote Sensing**, v. 30, n. 8, p. 2133-2147, 2009.

TUCCI, C. E. M.; CLARKE, R. T. Impacto das mudanças da cobertura vegetal no escoamento: revisão. **Revista Brasileira de Recursos Hídricos**, v. 2, n. 1, p. 135-152, 1997.

SILVA, M. G.; BELTRÃO, N. E. S.; MORALES, G. P. Avaliação e mapeamento dos serviços ecossistêmicos ofertados pela Reserva Biológica Nascentes da Serra do Cachimbo, Pará, Brasil. **Geosul**, v. 36, n. 78, p. 516-536, 2021.

USEPA. United States Environmental Protection Agency. **Stream Corridor Restoration – principles, processes and practices**. EUA: National Technical Info Svc, 1998.

USGS – **United States Geological Survey**. Satellite images, aerial photographs, and cartographic products. 2014. Disponível em: https://earthexplorer.usgs.gov/. Acesso em maio de 2021.

VON SPERLING, V. Afinal, quanta água temos no planeta? **Revista Brasileira de Recursos Hídricos**, v. 11, n. 4, p. 189-199, 2006.

WANG, Jielong; CHEN, Yi. Using NARX neural network to forecast droughts and floods over Yangtze River Basin. **Natural Hazards**, v. 110, n. 1, p. 225-246, 2022.

WIDROW, B.; HOFF, M. E. **Adaptive switching circuits**. Stanford Univ Ca Stanford Electronics Labs, 1960.

WILAMOWSKI, B. M.; CHEN, Y.; MALINOWSKI, A. Efficient algorithm for training neural networks with one hidden layer. In: **IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No. 99CH36339)**. IEEE, 1999. p. 1725-1728.

WOOD, E. F.; SIVAPALAN, M.; BEVEN, K.; BAND, L. Effects of spatial variability and scale with implications to hydrologic modeling. **Journal of hydrology**, v. 102, n. 1-4, p. 29-47, 1988.

WOOLHISER, David A. Hydrologic and watershed modeling–state of the art. **Transactions of the ASAE**, v. 16, n. 3, p. 553-559, 1973.

WUNSCH, A.; LIESCH, T.; BRODA, S. Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX). **Journal of Hydrology**, v. 567, p. 743-758, 2018.

WUNSCH, A.; LIESCH, T.; BRODA, S. Groundwater level forecasting with artificial neural networks: a comparison of long short-term memory (LSTM), convolutional neural networks (CNNs), and non-linear autoregressive networks with exogenous input (NARX). **Hydrology and Earth System Sciences**, v. 25, n. 3, p. 1671-1687, 2021.

WYTHOFF, Barry J. Backpropagation neural networks: a tutorial. **Chemometrics and Intelligent Laboratory Systems**, v. 18, n. 2, p. 115-155, 1993.

XU, L.; LIU, S. Study of short-term water quality prediction model based on wavelet neural network. **Mathematical and Computer Modelling**, v. 58, n. 3-4, p. 807-813, 2013.

YADAV, N.; YADAV, A.; KUMAR, M. **An introduction to neural network methods for differential equations**. Berlin: Springer, 2015.

YAZID, M. H. A.; SATRIA, M. H.; TALIB, S.; AZMAN, N. Artificial neural network parameter tuning framework for heart disease classification. In: **2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)**. IEEE, 2018. p. 674-679.

YONABA, H.; ANCTIL, F.; FORTIN, V. Comparing sigmoid transfer functions for neural network multistep ahead streamflow forecasting. **Journal of hydrologic engineering**, v. 15, n. 4, p. 275-283, 2010.

ZHANG, J.; MORRIS, A. J. A sequential learning approach for single hidden layer neural networks. **Neural networks**, v. 11, n. 1, p. 65-80, 1998.

APÊNDICE

APÊNDICE A – Algoritmo da RNN-NARX

```
%Universidade Federal do Pará - UFPA
%Programa de Pós-Graduação em Eng. Civil - PPGEC
%Área de concentração em Recursos Hídricos e Saneamento Ambiental
%RNN-NARX para simulação de hidrogramas de pequenas bacias

%Carregar conjuntos de dados
%Marambaia Fazenda_Craveiro Arraial Proximo_Colinas_Tocantins Base_do_Cachimbo

load Dataset Proximo_Colinas_Tocantins;
Dados = Proximo_Colinas_Tocantins;

clear Proximo_Colinas_Tocantins
% Informações fundamentais da arquitetura
A_Treinamento = 'trainlm';
Mode = 'open';
Def_P = 0:11;
Def_Q = 1:5;
NO = 8;

% Definição do particionamento dos conjuntos de dados
Treinamento_inicial = '01/01/1993';
Treinamento_final = '31/12/1995';
Validacao_inicial = '01/01/1996';
Validacao_final = '31/12/1997';
Teste_inicial = '01/01/1998';
Teste_final = '31/12/1998';


%Conjuntos de treinamento e validação cruzada
Treinamento = Dados((Dados.Data >= Treinamento_inicial) & (Dados.Data <= Teste_final), :);
P_treinamento = Treinamento.P;
Q_treinamento = Treinamento.Q;

clear Treinamento

%Conjunto de teste
Max_atraso = max([Def_P,Def_Q]);
Atraso_Teste = datetime(Teste_inicial, 'InputFormat', 'dd/MM/yyyy', 'Format','preserveinput')-
days(max(Max_atraso));
Teste = Dados((Dados.Data >= Atraso_Teste) & (Dados.Data <= Teste_final), :);
P_teste = Teste.P;
Q_teste = Teste.Q;

clear Teste Atraso_Teste Dados

% Matriz padrão de RNN
X_tr = tonndata(P_treinamento,false,false);
T_tr = tonndata(Q_treinamento,false,false);

X_te = tonndata(P_teste,false,false);
T_te = tonndata(Q_teste,false,false);

clear P_treinamento P_teste Q_treinamento Q_teste

% Construção da RNN-NARX
net = narxnet(Def_P, Def_Q, NO, Mode,A_Treinamento);
net.inputs{1}.processFcns = {'mapminmax'};
```

```
if isequal(Mode, 'open')
net.inputs{2}.processFcns = {'mapminmax'};
else
net.outputs{2}.processFcns = {'mapminmax'};
end

clear A_Treinamento

net.layers{1}.transferFcn = 'tansig';
net.layers{2}.transferFcn = 'purelin';
[x,xi,ai,Q] = preparets(net,X_tr,{},T_tr);

clear X_tr T_tr

net.trainParam.epochs = 1000;
net.trainParam.goal = 0.005;
net.trainParam.max_fail = 200;


d1 = datetime(Treinamento_inicial, 'InputFormat', 'dd/MM/yyyy', 'Format','preserveinput');
d2 = datetime(Treinamento_final, 'InputFormat', 'dd/MM/yyyy', 'Format','preserveinput');
d3 = datetime(Validacao_inicial, 'InputFormat', 'dd/MM/yyyy', 'Format','preserveinput');
d4 = datetime(Validacao_final, 'InputFormat', 'dd/MM/yyyy', 'Format','preserveinput');
d5 = datetime(Teste_inicial, 'InputFormat', 'dd/MM/yyyy', 'Format','preserveinput');
d6 = datetime(Teste_final, 'InputFormat', 'dd/MM/yyyy', 'Format','preserveinput');
Trein = caldays(between(d1,(d2+1),'days'));
Valid = caldays(between(d3,(d4+1),'days'));
Test = caldays(between(d5,(d6+1),'days'));
Total = caldays(between(d1,(d6+1),'days'));

clear Teste_final Teste_inicial Validacao_inicial Validacao_final Treinamento_inicial Treinamento_final

net.divideFcn = 'divideblock';
net.divideMode = 'time';
net.divideParam.trainRatio = Trein/Total;
net.divideParam.valRatio = Valid/Total;
net.divideParam.testRatio = Test/Total;

clear d1 d2 d3 d4 d5 d6

net.performFcn = 'mse';
net.performParam.regularization = 0.1;

[net,Relatorio] = train(net,x,Q,xi,ai);

%Treinamento da RNN-NARX
Qe = net(x,xi,ai);
Qe = cell2mat(Qe');
Q = cell2mat(Q');
clear x xi ai

Qe_treinamento = Qe(1:Trein- Max_atraso);
Q_treinamento = Q(1:Trein - Max_atraso);

Qe_validacao = Qe(Trein+1-Max_atraso:Trein-Max_atraso+Valid);
Q_validacao = Q(Trein+1-Max_atraso:Trein-Max_atraso+Valid);

Qe_teste = Qe(Trein+Valid-Max_atraso+1:end);
Q_teste = Q(Trein+Valid-Max_atraso+1:end);
```

```
clear Max_atraso Trein Valid Test Total Q Qe

%Convertendo Open-loop em Closed-loop
netc = closeloop(net);
[x,xi,ai,t] = preparets(netc,X_te,{},T_te);

Qe_teste_closed = netc(x,xi,ai);
Qe_teste_closed = cell2mat(Qe_teste_closed');

clear t X_te T_te ai xi x

% Contribuição Relativa de Garson (1991)
Iw = net.IW;
Lw = net.Lw;

if isequal(Mode, 'open')
   Pesos_P = cell2mat(Iw(1));
   Pesos_Q = cell2mat(Iw(1,2));
   Pesos_P_Q = abs(horzcat(Pesos_P, Pesos_Q));
   Pesos_Neuronios = abs(cell2mat(Lw(2,1)));
   Mult_Pesos = Pesos_P_Q .* Pesos_Neuronios';

   Soma_Pesos = sum(Mult_Pesos);
   Soma_total = sum(Soma_Pesos);
   CR = (Soma_Pesos .* 100) ./ Soma_total;

   clear Pesos_P Pesos_Q Pesos_P_Q Pesos_Neuronios Mult_Pesos Soma_Pesos Soma_total Iw
Lw Mode Def_P Def_Q
else
   Pesos_P = cell2mat(Iw(1));
   Pesos_Q = cell2mat(Lw(1,2));
   Pesos_P_Q = abs(horzcat(Pesos_P, Pesos_Q));
   Pesos_Neuronios = abs(cell2mat(Lw(2,1)));

   Mult_Pesos = Pesos_P_Q .* Pesos_Neuronios';
   Soma_Pesos = sum(Mult_Pesos);
   Soma_total = sum(Soma_Pesos);

   CR = (Soma_Pesos .* 100) ./ Soma_total;

   clear Pesos_P Pesos_Q Pesos_P_Q Pesos_Neuronios Mult_Pesos Soma_Pesos Soma_total Iw
Lw Mode Def_P Def_Q
end

% Avaliação estatística
[NSE, RSR, RMSE, PBias, MAPE]  = Desempenho(Q_treinamento, Qe_treinamento);
A = [NO, NSE, RSR, RMSE, PBias, MAPE];

[NSE, RSR, RMSE, PBias, MAPE]  = Desempenho(Q_validacao, Qe_validacao);
B = [NO, NSE, RSR, RMSE, PBias, MAPE];

[NSE, RSR, RMSE, PBias, MAPE]  = Desempenho(Q_teste, Qe_teste);
C = [NO, NSE, RSR, RMSE, PBias, MAPE];

[NSE, RSR, RMSE, PBias, MAPE]  = Desempenho(Q_teste, Qe_teste_closed);
D = [NO, NSE, RSR, RMSE, PBias, MAPE];

Criterios_Desempenho_OpenLoop = vertcat(A,B,C);
Criterios_Desempenho_ClosedLoop = vertcat(A,B,D);
```

```
clear A B C D NSE RMSE MAPE PBias RSR

%Curvas de permanência
a = 1:size(Qe_teste);
b = 1:size(Qe_treinamento);
c = 1:size(Qe_validacao);

Perm_Teste = (a./((max(a))+1));
Perm_Trein = (b./((max(b))+1));
Perm_Val = (c./((max(c))+1));

Perm_Teste_obs = sort(Q_teste, 'descend');
Perm_Teste_est_open = sort(Qe_teste, 'descend');
Perm_Teste_est_closed = sort(Qe_teste_closed, 'descend');

Perm_Trein_Obs = sort(Q_treinamento, 'descend');
Perm_Trein_Est = sort(Qe_treinamento, 'descend');

Perm_Val_Obs = sort(Q_validacao, 'descend');
Perm_Val_Est = sort(Qe_validacao, 'descend');

clear a b c

%Avaliação estatística curvas de permanencia

[NSE, RSR, RMSE, PBias, MAPE]  = Desempenho(Perm_Trein_Obs, Perm_Trein_Est);
A = [NO, NSE, RSR, RMSE, PBias, MAPE];

[NSE, RSR, RMSE, PBias, MAPE]  = Desempenho(Perm_Val_Obs, Perm_Val_Est);
B = [NO, NSE, RSR, RMSE, PBias, MAPE];

[NSE, RSR, RMSE, PBias, MAPE]  = Desempenho(Perm_Teste_obs, Perm_Teste_est_open);
C = [NO, NSE, RSR, RMSE, PBias, MAPE];

[NSE, RSR, RMSE, PBias, MAPE]  = Desempenho(Perm_Teste_obs, Perm_Teste_est_closed);
D = [NO, NSE, RSR, RMSE, PBias, MAPE];

Curvas_Permanencia_OpenLoop = vertcat(A,B,C);
Curvas_Permanencia_ClosedLoop = vertcat(A,B,D);

clear A B C D NSE RMSE MAPE PBias RSR NO

%Funções dos critérios de desempenho;
function [NSE, RSR, RMSE, PBias, MAPE] = Desempenho(y,ye)
NSE = 1-((sum((y-ye).^2))./(sum((y-mean(y)).^2)));
RMSE = sqrt(sum((y-ye).^2)./(length(y)));
PBias = 100.*((sum(ye-y))./(sum(y)));
MAPE = 100.*(sum(abs(ye-y)./y))/length(y);
RSR = RMSE./std(y);
end
```