



UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL
CURSO DE ENGENHARIA CIVIL

THIAGO VICTOR MEDEIROS DO NASCIMENTO

**MODELAGEM HIDROLÓGICA CHUVA–VAZÃO POR MEIO DE MAPAS AUTO-
ORGANIZÁVEIS PARA A BACIA HIDROGRÁFICA DO ALTO RIO SÃO
FRANCISCO UTILIZANDO DADOS DE PRECIPITAÇÃO TRMM**

JOÃO PESSOA

2020

THIAGO VICTOR MEDEIROS DO NASCIMENTO

**MODELAGEM HIDROLÓGICA CHUVA–VAZÃO POR MEIO DE MAPAS AUTO-
ORGANIZÁVEIS PARA A BACIA HIDROGRÁFICA DO ALTO RIO SÃO
FRANCISCO UTILIZANDO DADOS DE PRECIPITAÇÃO TRMM**

Trabalho de Conclusão de Curso submetido à
Coordenação do Curso de Engenharia Civil da
Universidade Federal da Paraíba, como um dos
requisitos obrigatórios para obtenção do título
de Bacharel em Engenharia Civil.

Orientador: Prof. Dr. Celso Augusto
Guimarães Santos

JOÃO PESSOA

2020

Catálogo na publicação
Seção de Catalogação e Classificação

N244m Nascimento, Thiago Victor Medeiros do.
MODELAGEM HIDROLÓGICA CHUVA-VAZÃO POR MEIO DE MAPAS
AUTO-ORGANIZÁVEIS PARA A BACIA HIDROGRÁFICA DO ALTO RIO
SÃO FRANCISCO UTILIZANDO DADOS DE PRECIPITAÇÃO TRMM /
Thiago Victor Medeiros do Nascimento. - João Pessoa,
2020.
83 f. : il.
Orientação: Celso Augusto Guimarães Santos.
Monografia (Graduação) - UFPB/CT.
1. Modelos chuva-vazão, SOM, TRMM. I. Santos, Celso
Augusto Guimarães. II. Título.


UFPB/BC

FOLHA DE APROVAÇÃO

THIAGO VICTOR MEDEIROS DO NASCIMENTO

**MODELAGEM HIDROLÓGICA CHUVA-AZÃO POR MEIO DE MAPAS AUTO
ORGANIZÁVEIS PARA A BACIA HIDROGRÁFICA DO ALTO RIO SÃO
FRANCISCO UTILIZANDO DADOS DE PRECIPITAÇÃO TRMM**

Trabalho de Conclusão de Curso em 06/03/2020 perante a seguinte Comissão Julgadora:



Celso Augusto Guimarães Santos
Departamento de Engenharia Civil e Ambiental do CT/UFPB

APROVADO



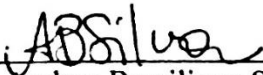
Francisco Jácome Sarmento
Universidade Federal da Paraíba

APROVADO



Aline Flavia Nunes Remigio Antunes
Universidade Federal da Paraíba

APROVADO



Prof. Andrea Brasiliano Silva
Matrícula Siape: 1549557
Coordenadora do Curso de Graduação em Engenharia Civil

Dedico este trabalho a Deus e às pessoas que amo e que me deram suporte para buscar os meus sonhos. Gratidão.

O orgulho leva a pessoa à queda, e a vaidade
faz cair na desgraça.

Provérbios 16:18

AGRADECIMENTOS

Num momento tão especial eu não poderia deixar de agradecer a todas as pessoas que estiveram comigo durante toda a minha caminhada. Foram dias de lutas, mas também de muitas vitórias, e da certeza de que Deus estava comigo em todos os momentos.

Agradeço primeiramente a Deus por tanta bondade e generosidade em minha vida. Sem Ele eu não teria chegado até aqui.

Agradeço a minha mãe Rosilene Medeiros, e a minha avó Severina Medeiros que estão sempre ao meu lado, e são as melhores pessoas que eu poderia ter na minha vida. Ter a certeza de que tenho o privilégio de ter duas pessoas tão especiais na minha vida me deixa com uma felicidade sem limites.

Agradeço ao meu avô Severino Medeiros por ter estado comigo até o dia que Deus permitiu. Minha eterna gratidão.

Ao meu pai, José Pereira por todo o apoio e dedicação direcionadas a mim.

Agradeço a minha tia Maria José por sempre ter sido como mãe.

A minha família, que apesar de não ser perfeita, é uma boa família.

Agradeço a minha namorada Manuela Amaral por ter estado comigo em boa parte da caminhada acadêmica e sempre ter me apoiado.

Ao meu professor, orientador e grande amigo Celso Santos pela oportunidade e confiança.

Aos professores Cristiano, Francisco Sarmiento, Karine e Ana Cláudia por terem sido tão importantes no desenvolvimento da minha paixão por recursos hídricos.

Aos melhores amigos que a universidade poderia me presentear, dos quais não poderia deixar de citar Rafaela, Laíssa, Raul, Yesus, Eduardo, Ana Carolina e Alexandre.

Aos meus amigos Valdir, Arthur, Adelar, Aryellen, Emmanuella, Michael, Matheus, Fagner, Andrielle, Waleska, Yan, Claudemir, Bárbara e tantos outros que posso ter esquecido de citar.

Aos amigos de infância que ainda permanecem e mesmo distantes, sempre estão presentes, como Stênia Rodrigues e tantos outros.

Aos meus amigos da Primeira Igreja Batista de Itabaiana por tanto amor e cuidado.

Aos meus amigos da Juventude Batista do Vale do Paraíba (JUBAVALE).

Aos meus amigos e colegas de trabalho do grupo de orientandos do professor Celso Santos e Richarde Marques e do LEPPAN.

RESUMO

O bom gerenciamento da água é essencial para o desenvolvimento das mais variadas atividades numa sociedade, como o abastecimento humano, hidroeletricidade, indústria, irrigação, dessedentação dos animais, entre outros. Nesse sentido, para a realização de uma boa gestão e controle desse recurso, estudos e pesquisas na área são essenciais, e para as quais é requerida a existência de séries de dados hidrológicos longas e confiáveis (e.g. precipitação, vazão etc.). Entretanto, é comum a inexistência de séries históricas consistentes, como as de chuva e vazão em bacias hidrográficas. Mas, ainda nesse aspecto de monitoramento de variáveis hidrológicas, há uma alternativa para o preenchimento, e até a previsão de dados de vazão afluente em rios e reservatórios onde há a limitação de dados disponíveis, que é a utilização de modelos chuva-vazão. Como uma alternativa para a modelagem chuva-vazão, existe também a utilização de Redes Neurais Artificiais (RNA). Assim, o presente trabalho utiliza dados de precipitação do satélite *Tropical Rainfall Measurement Mission* (TRMM) na modelagem chuva-vazão numa rede RNA chamada de mapas auto-organizáveis (do inglês *Self-Organizing Maps* – SOM), em um reservatório da bacia do Alto São Francisco, onde é feita a modelagem de vazões afluentes naturais no mesmo. São realizadas quatro modelagens com períodos para treinamento e teste dos resultados diferentes, dentro do período total de dados disponíveis 1998-2019, devido a não homogeneidade da série completa, além disso, são calculadas métricas estatísticas nos resultados para uma análise do desempenho dos modelos e análise da influência ou não da homogeneidade dos dados na rede SOM. Os resultados mostraram que a homogeneidade dos dados influenciou na modelagem chuva-vazão de acordo com a representatividade da série completa na etapa de treinamento. Para a etapa de calibração, os modelos apresentaram índices de eficiência máximos, e para a de teste, os coeficientes de NASH foram superiores a 0,80 para três modelagens, assim como o coeficiente de determinação, juntamente com valores menores do que 9% para o viés relativo (RB). Os piores valores foram encontrados para a modelagem que fez uso de dois períodos de comportamento diferente para a etapa de treinamento e teste (modelagem #1). Assim, reforça-se a relevância e possibilidade de utilização de redes SOM na modelagem chuva-vazão a nível mensal, bem como a eficiência de dados TRMM de precipitação como uma alternativa aos dados tradicionais como pluviômetros, por exemplo.

Palavras Chave: Modelos chuva-vazão, SOM, TRMM.

ABSTRACT

The good water management is essential for the development of the most varied activities in a society, such as human supply, hydroelectricity, industry, irrigation, and animal feed, among others. In this sense, for the realization of a good management and control of this resource, studies and research in the area are essential, and for which the existence of long and reliable hydrological data series (e.g. precipitation, flow, etc.) is required. However, it is common to lack consistent historical series, such as rainfall and flow in hydrographic basins. But, still in this aspect of monitoring hydrological variables, there is an alternative for filling, and even forecasting, affluent flow data in rivers and reservoirs where there is a limitation of available data, which is the use of rain-flow models. As an alternative for rain-flow modeling, there is also the use of Artificial Neural Networks (ANN). Thus, the present work uses precipitation data from the Tropical Rainfall Measurement Mission (TRMM) satellite in rain-flow modeling in an RNA network called self-organizing maps, in a reservoir in the Alto basin São Francisco, where the modeling of natural affluent flows is made in it. Four models are performed with periods for training and testing the different results, within the total period of available data 1998-2019, due to inhomogeneity of the complete series, in addition, statistical metrics are calculated in the results for an analysis of the models' performance and analysis of the influence or not of the homogeneity of the data in the SOM network. The results showed that the homogeneity of the data influenced the rainfall-flow modeling according to the representativeness of the complete series in the training stage. For the calibration step, the models showed maximum efficiency indexes, and for the test, the NASH coefficients were greater than 0.80 for three models, as well as the determination coefficient, together with values less than 9% for the relative bias (RB). The worst values were found for the modeling that used two periods of different behavior for the training and testing stage (modeling # 1). Thus, the relevance and possibility of using SOM networks in rain-flow modeling at a monthly level is reinforced, as well as the efficiency of TRMM precipitation data as an alternative to traditional data such as rain gauges, for example.

Keywords: Rainfall-inflow models, SOM, TRMM.

LISTA DE FIGURAS

Figura 1 – Esquema de uma bacia hidrográfica.	21
Figura 2 - O ciclo da água.	22
Figura 3 – Concepção artística do satélite TRMM.....	26
Figura 4 – Concepção de uma rede neural artificial.....	33
Figura 5 – Estrutura de um modelo SOM com detalhe do neurônio vencedor, e seus neurônios vizinhos.....	36
Figura 6 – As distâncias entre os neurônios de um modelo SOM para a determinação da vizinhança.....	39
Figura 7 – Localização da sub-bacia do Alto São Francisco.....	44
Figura 8 – Topografia da região.	45
Figura 9 – Uso e ocupação do solo.....	46
Figura 10 – Precipitação anual média na região (1998-2018) baseado em dados TRMM.	47
Figura 11 – Imagem de satélite do reservatório Três Marias.	48
Figura 12 – Espacialização das quadrículas TRMM na área de estudo.	50
Figura 13 – Série de precipitações médias mensais para a região baseadas nas 91 quadrículas TRMM (1998-2019).	51
Figura 14 – Série de vazões médias afluentes naturais mensais para o reservatório Três Marias (1998–2019).	53
Figura 15 – Arquitetura da rede SOM utilizada.	55
Figura 16 – Série de vazões observadas e calculadas para os períodos de calibração, teste e para todo o período, para a modelagem #1.....	61
Figura 17 – Plotagem dos valores observados e calculados e regressão linear para os períodos de calibração, teste e para todo o período, para a modelagem #1.	62
Figura 18 – Série de vazões observadas e calculadas para os períodos de calibração, teste e para todo o período, para a modelagem #2.....	64
Figura 19 – Plotagem dos valores observados e calculados e regressão linear para o os períodos de calibração, teste e para todo o período, para a modelagem #2.....	65
Figura 20 – Série de vazões observadas e calculadas para os períodos de calibração, teste e para todo o período, para a modelagem #3.....	67
Figura 21 – Plotagem dos valores observados e calculados e regressão linear para o os períodos de calibração, teste e para todo o período, para a modelagem #3.....	68
Figura 22 – Série de vazões observadas e calculadas para os períodos de calibração, teste e para todo o período, para a modelagem #4.....	70
Figura 23 – Plotagem dos valores observados e calculados e regressão linear para o os períodos de calibração, teste e para todo o período, para a modelagem #4.....	71

LISTA DE QUADROS

Quadro 1 - Tipos de produtos do satélite TRMM.	27
--	----

LISTA DE TABELAS

Tabela 1 - Índices estatísticos para as precipitações observadas.....	49
Tabela 2 - Índices estatísticos para as vazões observadas.....	52
Tabela 3 - Índices desempenho para a etapa de calibração.....	54
Tabela 4 - Índices desempenho para as etapas de calibração, teste e para todo o período, para a modelagem #1.....	60
Tabela 5 - Índices desempenho para as etapas de calibração, teste e para todo o período, para a modelagem #2.....	64
Tabela 6 - Índices desempenho para as etapas de calibração, teste e para todo o período, para a modelagem #3.....	67
Tabela 7 - Índices desempenho para as etapas de calibração, teste e para todo o período, para a modelagem #4.....	70
Tabela 8 – Tabela resumo com os índices de desempenho para os dados na etapa de teste das quatro modelagens utilizadas no trabalho.....	73

LISTA DE ABREVIATURAS E SIGLAS

ANA – Agência Nacional de Águas
ANEEL – Agência Nacional de Energia Elétrica
BMU – Best matching unit
CERES – Cloud and Earth Radiant Energy Sensor
FAO – Food and Agricultural Organization
FFNN – Feed-forward
GPM – Global Precipitation Measurement
GSMaP – Global Satellite Mapping of Precipitation project,
IMERG – Integrated Multisatellite Retrievals for the Global Precipitation Measurement
JAXA – Agência de Exploração Aeroespacial do Japão
KSOM – Kohonen self-organizing map
LIS – Lightning Imaging Sensor
MME – Ministério de Minas e Energia
MODHAC – Modelo hidrológico auto calibrável
NASA – Administração Nacional da Aeronáutica e Espaço
ONS – Operador Nacional do Sistema Elétrico
PERSIANN-CDR – Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks – Climate Data Record
PR – Precipitation Radar
RB – Relative bias
RMSE – Raiz quadrada do erro médio
RNA – redes neurais artificiais
SIG – Sistema de Informações Geográficas
SMAP – Soil Moisture Accounting Program
SOM – Self-Organizing Maps
SRTM – Shuttle Radar Topography Mission
SWAT – Soil and Water Assessment Tool
TMI – TRMM Microwave Imager
TMPA – Multi-satellite Precipitation Analysis
TRMM – Tropical Rainfall Measuring Mission
USGS – United States Geological Survey
VIRS – Visible Infrared Radiometer

WMO – World Meteorological Organization

SUMÁRIO

1	INTRODUÇÃO.....	17
1.1	OBJETIVOS	19
1.1.1	Objetivo geral.....	20
1.1.2	Objetivo específicos.....	20
2	REFERENCIAL TEÓRICO	21
2.1	BACIA HIDROGRÁFICA.....	21
2.2	CICLO HIDROLÓGICO.....	22
2.2.1	Precipitação	23
2.2.1.1	Definição.....	23
2.2.1.2	Formas de obtenção de dados	24
2.2.1.3	Dados de Precipitação TRMM	25
2.2.2	Vazão.....	28
2.2.2.1	Definição.....	28
2.2.2.2	Medição de vazão	28
2.2.2.3	Vazões naturais	29
2.3	MODELAGEM HDROLÓGICA	30
2.3.1	Definição	30
2.3.2	Modelos chuva-vazão.....	31
2.4	REDES NEURAS ARTIFICIAIS	31
2.4.1	Definição e estrutura das RNA	31
2.4.2	Considerações sobre o treinamento das RNA	33
2.4.3	Vantagens das RNA	34
2.4.4	Aplicações de RNA na previsão de séries temporais	34
2.5	MAPAS AUTO-ORGANIZÁVEIS	35
2.5.1	Definição	35
2.5.2	Estrutura.....	35
2.5.3	Treinamento	37
2.5.4	Previsão utilizando SOM.....	40
2.5.5	Aplicações do SOM em séries temporais	40
3	MATERIAIS E MÉTODOS	43
3.1	ÁREA DE ESTUDO	43
3.2	DADOS DE PRECIPITAÇÃO.....	48
3.3	DADOS DE VAZÃO	51
3.4	METODOLOGIA PARA A MODELAGEM SOM.....	53
3.4.1	Modelagens SOM utilizadas	53
3.4.2	Preparação dos dados.....	54
3.4.3	Arquitetura da rede SOM.....	55
3.4.4	Calibração do modelo.....	56
3.4.5	Teste do modelo.....	56
3.5	CÁLCULO DOS ÍNDICES DE DESEMPENHO.....	57
3.5.1	Coefficiente de determinação (R^2)	57
3.5.2	Coefficiente de Nash.....	57
3.5.3	Viés Relativo.....	58
3.5.4	Raiz do Erro Médio Padrão	58

4	RESULTADOS E DISCUSSÕES.....	59
4.1	MODELAGEM #1	59
4.2	MODELAGEM #2	62
4.3	MODELAGEM #3	66
4.4	MODELAGEM #4	69
4.5	DISCUSSÕES DOS RESULTADOS	72
5	CONSIDERAÇÕES FINAIS	75
6	REFERÊNCIAS BIBLIOGRÁFICAS.....	77

1 INTRODUÇÃO

A água é a substância mais presente na superfície da terra, e cobre mais de 70% da superfície do globo, todas as formas de vida necessitam de água para sobreviver, e é a única substância na terra naturalmente presente nas formas líquida, sólida e gasosa (COLLISCHONN e DORNELLES, 2015). Assim, percebe-se a importância enorme que esse elemento possui.

Ainda de acordo com Collischonn e Dornelles (2015), no que se refere aos usos múltiplos da água, ela possui como tipos de usos principais, o abastecimento humano, a irrigação, a dessedentação animal, a geração de energia elétrica, a navegação, a diluição de efluentes, a pesca, a recreação e o paisagismo. Além disso, esses usos podem ser classificados como sendo consuntivos, ou não consuntivos, e em relação se é necessário, ou não, a retirada da água do corpo d'água para ser realizado o uso.

Em relação a essas duas classificações, os usos consuntivos são os que alteram substancialmente a quantidade da água disponível para outros usuários (irrigação), e os não consuntivos são os que pouco alteram, a quantidade de água disponível para outros usuários, mas que podem modificar a sua qualidade (hidroelétricas e o uso de água para navegação, por exemplo). Em relação a necessidade de retirada, ou não retirada da água dos corpos hídricos originais para a sua utilização, tem-se respectivamente irrigação e abastecimento, contra navegação e geração de energia, por exemplo, como exemplos onde é necessária a retirada, e onde não o é.

É importante destacar também que a água se movimenta na terra através de um conceito chamado de ciclo hidrológico, e esse movimento é o que permite a variação da água nas suas diversas formas (líquida, gasosa e sólida). De maneira breve, a água de superfícies hídricas como lagos e oceanos evapora, vai para a atmosfera, se condensa na forma de nuvens que acabam por gerar a precipitação, a água retorna para a superfície da terra, e outra parte infiltra, retornando para os corpos hídricos e assim formando esse ciclo contínuo, onde a água está em constante movimento.

Portanto, destaca-se a importância e relevância do estudo e entendimento da água e os seus processos no meio ambiente. Conhecer a precipitação e principalmente seu impacto sobre a bacia é indispensável, visto que a partir disso pode-se fazer o gerenciamento otimizado dos recursos hídricos, conseguindo avaliar a necessidade de irrigação, prever enchentes nos rios,

avaliar o atendimento às demandas para abastecimento público, avaliar as operações de hidroelétricas, organizar projetos de engenharia hidráulica.

Nesse sentido, entender e monitorar as variáveis climáticas como a precipitação e a vazão é essencial. Em linhas gerais, pode-se classificar como tendo um bom monitoramento da precipitação no Brasil e no mundo, sobretudo quando comparado com o monitoramento do fluxo de vazão em rios e reservatórios. Em geral, é comum a ausência de dados consistentes de fluviômetros para a maior parte dos rios no Brasil. Nas bacias desprovidas de registros sistemáticos de cota e descarga, faz-se necessário desenvolver metodologias que possibilitem a transferência das informações hidrológicas existentes em outras bacias, com características semelhantes a bacia que se quer estudar.

Assim, surge a importância e utilidade da modelagem ambiental como uma ferramenta para a previsão e preenchimento de dados de vazão em rios e reservatórios. Segundo Tucci (1998), um modelo hidrológico é classificado como sendo uma ferramenta utilizada para representar processos que ocorrem numa bacia hidrográfica, possibilitando uma análise de previsão de consequências de diferentes ocorrências de valores observados. No caso de a variável desejada ser a vazão de um rio, pode-se imaginar como parâmetros de entrada a variável precipitação, por exemplo. O modelo transforma a chuva precipitada em vazão.

Como exemplos de modelos chuva-vazão pode-se citar SMAP (*Soil Moisture Accounting Program*) descritos por Lopes et al. (1981), *Tank Model* (Sugawara, 1961), MODHAC (Modelo hidrológico auto calibrável) desenvolvido por Lanna Schwarzbach (1989), e o *Soil and Water Assessment Tool – SWAT* (ARNOLD et al., 2012). São modelos hidrológicos que utilizam dados de precipitação como como variáveis de entrada para a modelagem da vazão. Entretanto, a maioria dos modelos faz uso de variadas variáveis ambientais como dados de entrada e.g. evapotranspiração, uso e ocupação do solo, precipitação, temperatura do ar. Alternativas de modelos que utilizem menos variáveis se mostram como uma vantagem, a partir do momento em que passem a mostrar uma boa segurança na modelagem.

Nesse sentido, é relevante citar que mais recentemente, modelos baseados em redes neurais artificiais tem sido amplamente utilizado na análise chuva–vazão (Coulibaly et al., 2001; Jeong e Kim, 2005; Farias et al., 2007; Wu & Chau, 2011; e Santos et al., 2012 a, b). As redes neurais artificiais (RNAs) podem ser definidas como sendo ferramentas matemáticas, inspiradas no entendimento de que as funções neurais biológicas, como por exemplo a memória, são armazenadas em neurônios, e nas conexões entre os mesmo, e desse modo, o processo de

aprendizado consiste no estabelecimento de novas conexões entre os neurônios, e em modificações nas conexões existentes (FREIRE, 2012).

Dentro da área de recursos hídricos, a utilização de RNAs é presente em diversos ramos, sobretudo na previsão de variáveis hidrológicas como chuva, vazão, evapotranspiração, níveis freáticos, salinidade do solo, variáveis que indicam qualidade da água etc. (Gorgij et al., 2008; Farias et al., 2010; Adeloye et al. 2011; Santos et al., 2014; Farias et al., 2015; Bayatvarkeshi et al., 2018; Mirabbasi et al., 2018; Noureddine et al., 2019; Silva et al., 2019; Honorato et al., 2019; Machado et al., 2019; e Santos et al., 2019).

Dentre os tipos de RNA pode-se destacar um em especial que são os Mapas Auto-Organizáveis (do inglês *Self-Organizing Maps* – SOM), também conhecidos como redes auto-organizáveis de Kohonen (*Kohonen self-organizing map* - KSOM), e foram propostas por Kohonen (1982). Eles podem ser definidos como redes não supervisionadas que agrupam dados de entrada em classes de acordo com suas semelhanças por meio de métodos de treinamento competitivo (Kohonen, 1982; Haykin, 1999; Silva et al., 2010).

Como aplicações, possível realizar com as redes SOM nos dados, o agrupamento de variáveis ambientais, a redução da quantidade de dados de treinamento para identificação do modelo, interpolação não linear e extrapolação (ou seja, previsão), generalização e compactação de informações para transmissão fácil (Kangas & Simula, 1995; Kohonen et al., 1996; Parasuraman et al., 2006; Tananaki et al., 2007; Chang et al., 2010; Adeloye et al., 2011; Rivera et al., 2011; Adeloye e Rustum, 2012, Nourani et al., 2013; Farias et al., 2012; Farias et al., 2013; e Filho e Farias, 2018). Por exemplo, os trabalhos de Farias et al. (2012), Farias et al. (2013); Filho e Farias (2018) aplicaram o SOM em bacias hidrográficas para realizar a previsão de vazões a nível mensal, baseada nas precipitações e vazões do próprio mês, e de um e dois meses anteriores.

Dadas as dificuldades de obtenção de variáveis ambientais para a modelagem ambiental, e a possibilidade de utilização de RNAs como alternativas mais diretas nesses modelos, o presente trabalho possui o objetivo de analisar a aplicabilidade da rede neural artificial SOM na modelagem chuva-vazão a nível mensal, na sub-bacia hidrográfica do Alto Rio São Francisco, inserida na bacia hidrográfica do Rio São Francisco, utilizando como variáveis de entrada para o modelo apenas dados de precipitação provenientes do satélite TRMM.

1.1 OBJETIVOS

1.1.1 Objetivo geral

Aplicar a rede neural artificial do tipo SOM na modelagem chuva-vazão, na sub-bacia do Alto São Francisco, para o cálculo de vazões naturais mensais, utilizando dados de precipitação da *Tropical Rainfall Measuring Mission* (TRMM).

1.1.2 Objetivo específicos

- a) Utilizar dados de precipitação provenientes TRMM na modelagem chuva-vazão a nível mensal numa rede SOM;
- b) Analisar se a homogeneidade de uma série de vazões influenciou na calibração de um modelo chuva-vazão numa rede SOM;
- c) Calibração e teste de um modelo chuva-vazão treinado numa rede SOM para os períodos de respectivamente de 01/01/1998-30/04/2014 (calibração) e 01/05/2014-31/08/2019 (teste); 01/07/2003-31/08/2019 (calibração), 01/01/1998-30/06/2003 (teste); 01/01/1998-31/08/2008 (calibração) e 01/09/2008-30/04/2014 (teste); e 01/05/2014-31/08/2017 (calibração) e 01/09/2017-31/08/2019 (teste);
- d) Verificação do desempenho dos modelos propostos com base em métricas estatísticas.

2 REFERENCIAL TEÓRICO

2.1 BACIA HIDROGRÁFICA

Segundo Rolim (2004), bacia hidrográfica refere-se a uma área de captação natural da água de precipitação que faz convergir os escoamentos para um único ponto de saída, que é chamado de exutório. Ainda de acordo com ele, uma bacia é constituída por um conjunto de superfícies vertentes – terreno sobre o qual escoam a água precipitada – e de uma rede de drenagem formada por cursos d'água que confluem até resultar um leito único no exutório. É interessante notar também, que devido a ser uma concepção física e natural, todos os pontos da superfície terrestre integram uma bacia, ou sub bacia. Além disso, salienta-se que mais de uma sub-bacia pode estar associada a uma bacia principal, e essa análise pode ser feita de acordo com a finalidade de estudo. Assim, uma bacia pode ser dividida em sub bacias associadas a um corpo hídrico principal (HORNBERGER, 1998). Na Figura 1 é mostrado o desenho de uma bacia hidrográfica, com relevo, corpos d'água e exutório.

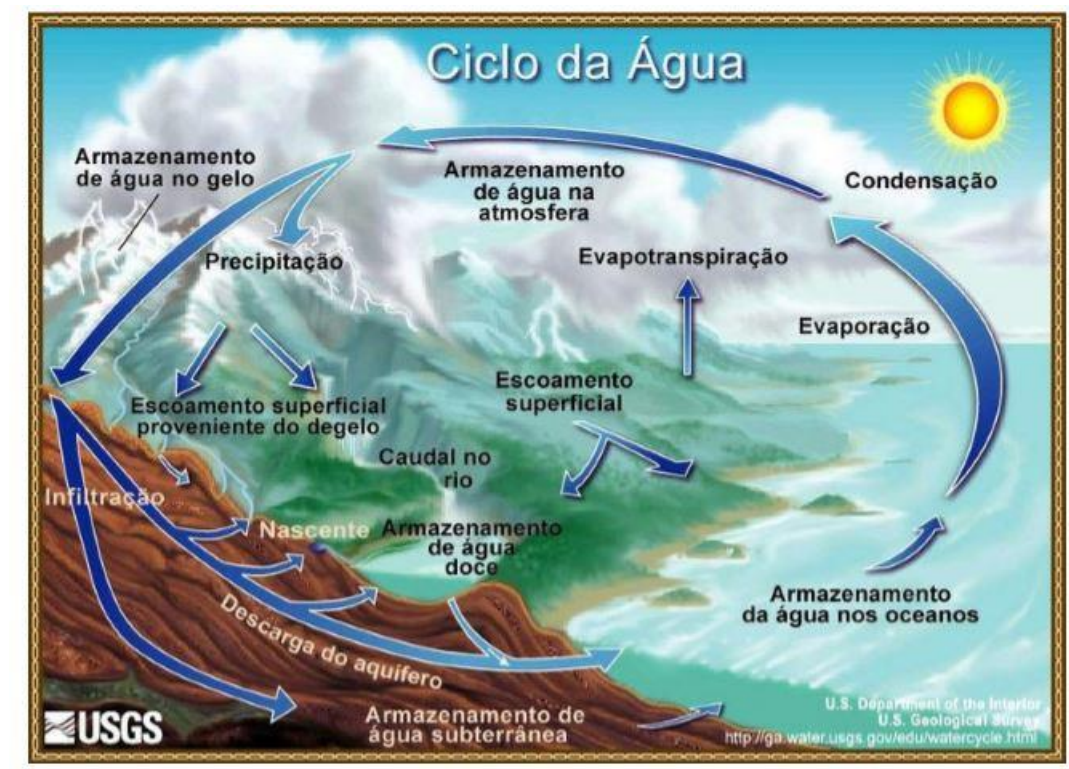
Figura 1 – Esquema de uma bacia hidrográfica.



2.2 CICLO HIDROLÓGICO

De acordo com Tucci (2009), o ciclo hidrológico (conforme mostrado na Figura 2) pode ser definido como um processo fechado, com abrangência global, onde a água circula entre os reservatórios chamados por ele como terrestres (continentes e oceanos) e atmosférico, e que possui como o impulsionador central o sol. Em todo esse processo, a água acaba por se apresentar em seus três estados físicos, i.e. o sólido, o líquido e o gasoso. Ele representa o processo natural de evaporação, condensação, precipitação, detenção e escoamento superficial, infiltração, percolação da água no solo e nos aquíferos, escoamentos fluviais e interações entre esses componentes.

Figura 2 - O ciclo da água.



Fonte: United States Geological Survey (USGS).

Como o próprio nome já sugere, o processo é representado como um ciclo, assim, não possui um início, ou fim, ocorrendo de maneira contínua. Tomando-se a Figura 2 como base ilustrativa, e considerando-se o início do ciclo nos oceanos, ele processa-se com a evaporação

da água deles, e com o conseqüente transporte do mesmo pelo movimento contínuo de massas de ar na atmosfera. Em condições ideais a água é condensada na forma de nuvens.

As nuvens, conseqüentemente, resultam na precipitação da água, que pode ser na forma de neve, granizo e de chuva propriamente dita. A água precipitada retorna para a terra, onde parte gera escoamento superficial, onde a água alimenta rios e reservatórios (lagoas, açudes etc.), e outra infiltra no solo, abastecendo aquíferos que em última análise seguem para os corpos d'água superficiais rios e lagos, e que seguem com destino aos oceanos (governadas por gravidade), e onde se reinicia o ciclo. Parte da água retorna diretamente para a atmosfera, através do processo de evapotranspiração, que representa a perda de água do solo por evaporação e a perda de água de plantas e animais por transpiração.

Assim, toma-se dois dos processos mais importantes no processo do ciclo hidrológico, i.e., a precipitação e o escoamento superficial (vazão de rios, no caso). Ambos são inteiramente ligados, uma vez que geralmente, a precipitação é uma variável de entrada importante responsável pelo escoamento em corpos d'água. A seguir esses dois processos são discutidos mais profundamente.

2.2.1 Precipitação

2.2.1.1 Definição

De acordo com Collischonn (2015), a precipitação consiste no retorno da água à superfície da terra nas formas de chuva, granizo, neve, orvalho neblina ou geada. A chuva representa a forma mais importante de precipitação. Na engenharia, a chuva é a principal causa dos processos hidrológicos, e sua quantificação e entendimento corretos é essencial para o gerenciamento, dos recursos naturais e hídricos, quanto para a execução de obras civis etc.

Dentro de uma bacia hidrográfica, a precipitação é a única forma de entrada de água, e assim, é ela quem fornece subsídios para o gerenciamento de recursos hídricos, como a quantificação do abastecimento de água e irrigação; o controle de desastres ambientais como inundações, o processo de erosão do solo, etc., e é fundamental para o adequado dimensionamento de obras hidráulicas, por exemplo (COLLISCHONN, 2015)

Além disso, destaca-se a importância da precipitação no desenvolvimento de vários estudos relacionados aos recursos hídricos, seja para o monitoramento das secas, para processos de modelagem hidrológica e hidrossedimentológica, para o dimensionamento e gerenciamento de reservatórios ou para análises de tendências etc. O processo de precipitação na forma de chuva pode ocorrer de três formas principais, i.e., chuvas convectivas, orográficas e frontais.

2.2.1.2 Formas de obtenção de dados

De forma geral, as medições de chuva são feitas por postos pluviométricos, por meio de aparelhos como o pluviógrafos e o pluviômetro. O pluviômetro consiste em um aparelho que é dotado de uma superfície de captação horizontal, e de um reservatório para acumular a água recolhida, ligado a essa área de captação. É um aparelho que fornece o total de água acumulado durante um intervalo de tempo definido pela leitura dele, efetuada por um operador. O pluviógrafos assemelha-se ao pluviômetro em sua estrutura, diferenciando-se pela capacidade de gerar medições contínuas da chuva, e possibilitando uma análise temporal mais detalhada, bem como de suas intensidades etc.

Esses aparelhos representam a mais importante e direta forma de obtenção de dados de precipitação, mas devido a fatores geográficos e em função dos custos de operação e manutenção da rede de monitoramento pluviométrico, esse monitoramento tem se tornado cada vez mais desafiador (Terink et al., 2018). No Brasil, por exemplo, essa situação tende a ser ainda mais agravante dada as dimensões continentais do país e a diversidade de climas, características físicas e regiões pluviometricamente homogêneas existentes. O estudo feito por Gadelha et al. (2019), mostrou que a rede de monitoramento de chuva do Brasil conta com aproximadamente 11.820 estações pluviométricas, resultando em uma densidade de cerca de um pluviômetro a cada 720 km², que é inferior a recomendação pela WMO (1994), que é de um aparelho para cada 575 km².

Além disso, estações com séries de dados ininterruptos no Brasil são escassas e as séries geralmente apresentam falhas em sua composição, destacando-se principalmente as regiões Nordeste, da Floresta Amazônica brasileira, e as centrais do Brasil como as áreas que tem a rede mais deficiente do país, resultando em uma caracterização pluviométrica insuficiente (Buarque et al., 2011; Collischonn et al., 2008; Curtarelli et al., 2014; Delahaye et al., 2015).

Nesse sentido, dois métodos que também podem ser utilizados para a obtenção de dados de precipitação são os radares climáticos ou através do monitoramento via satélite.

Em relação as observações por radar, Tang et al. (2015) afirmam que as estimativas de chuva têm grandes incertezas associadas a terrenos complexos, estando esse fato relacionado a erros de sinais eletrônicos diante de ambientes de operação desafiadores. Já a partir do monitoramento via satélite, estimativas de precipitação a nível global podem ser obtidos de maneira sistemática a partir da capacidade das microondas de penetrarem as nuvens e interagirem com as partículas de precipitação (Liao, 2015).

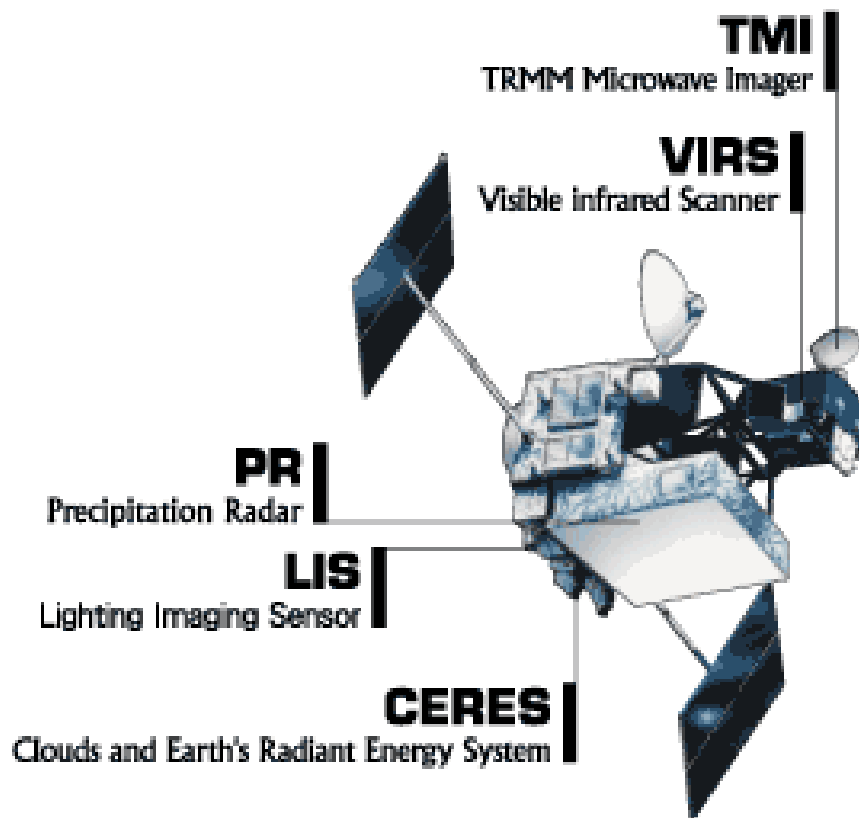
Atualmente, existem vários produtos de satélite de sensoriamento remoto com diferentes resoluções espaço-temporais capazes de gerar estimativas de precipitação de maneira eficaz. Dentre eles, destacam-se o Tropical Rainfall Measuring Mission (TRMM), Multi-satellite Precipitation Analysis (TMPA), o Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) – Climate Data Record (CDR), o Global Satellite Mapping of Precipitation (GSMaP) project, e Integrated Multisatellite Retrievals for the Global Precipitation Measurement (IMERG). Cada um desses satélites tem suas vantagens e desvantagens e a validação de seus produtos vem sendo avaliada por diversos estudos, como Chen et al. (2013), Ning et al. (2015), Sultana et al. (2018), Zeng et al. (2018), Hur et al. (2018) e Gadelha et al. (2019).

2.2.1.3 Dados de Precipitação TRMM

De acordo com a NASA (2020), o TRMM consiste numa missão espacial, lançada em 28 de novembro de 1997, realizada com uma parceria entre a Administração Nacional da Aeronáutica e Espaço (NASA) e a Agência de Exploração Aeroespacial do Japão (JAXA), com o objetivo de medir a precipitação e desenvolver pesquisas climáticas. Este satélite foi destinado a melhorar a compreensão da distribuição e variabilidade da precipitação em regiões tropicais e subtropicais da Terra, abrangendo mais de dois terços da precipitação total mundial. A coleta de dados da TRMM sobre precipitação tropical ajuda a melhorar nosso entendimento sobre clima e clima. Ele fornece importantes informações de precipitação para aumentar a compreensão das interações entre o vapor d'água, nuvens e precipitação, que são fundamentais para a regulação do clima da Terra (NASA, 2020).

É interessante destacar também que apesar da missão TRMM ter tido fim no ano de 2015 a continuidade da coleta de dados foi garantida pela missão conjunta NASA/JAXA de *Global Precipitation Measurement* (GPM) lançada em 2014. Na Figura 3 é mostrada uma concepção artística do satélite TRMM. No Quadro 1 são mostrados simplificados os produtos obtidos tanto do tipo orbitais, quanto em grid, para o satélite TRMM.

Figura 3 – Concepção artística do satélite TRMM.



Fonte: NASA (2020).

O satélite utiliza vários instrumentos para detectar chuvas, incluindo sensores de radar, imagens de microondas e raios. Voa a uma altitude orbital baixa de 240 milhas (400 quilômetros). Em relação aos instrumentos do satélite mostrados na Figura 3, tem-se que:

- a) VIRS (*Visible Infrared Radiometer*) é um radiômetro que opera em cinco diferentes canais (0.63, 1.6, 3.75, 10.8 e 12 μm), e que promove uma alta resolução nas observações das nuvens, captando informações a respeito da cobertura, do tipo e das temperaturas das nuvens;

- b) TMI (*TRMM Microwave Imager*) é um radiômetro de micro-ondas passivo que opera em cinco diferentes frequências (10.65, 19.35, 37.0, 85.5 e 22.235 GHz) e que fornece informações integradas sobre a coluna de precipitação, a nuvem líquida de água, a nuvem de gelo, a intensidade da chuva e os tipos de precipitação;
- c) PR (*Precipitation Radar*) é um radar de escaneamento eletrônico que opera na frequência de 13.8 GHz e que mensura a distribuição da precipitação tanto no oceano quanto na terra de maneira tridimensional;
- d) CERES (*Cloud and Earth Radiant Energy Sensor*) mede a energia no topo da atmosfera bem como estima os níveis de energia dentro da atmosfera e na superfície terrestre;
- e) LIS (*Lightning Imaging Sensor*) é um sensor pequeno, mas altamente sofisticado que detecta e localiza os raios sobre a região tropical do globo terrestre.

Quadro 1 - Tipos de produtos do satélite TRMM.

Produtos orbitais		Produtos em grid	
ID	Nome	ID	Nome
2A21	Precipitation Radar (PR) Surface Cross-Section	3B31	Monthly 5° × 5° Combined Rainfall
2A23	Precipitation Radar (PR) Rain Characteristics	3A46	Monthly 1° × 1° SSM/I Rain
2A25	Precipitation Radar (PR) Rainfall Rate and Profile	3B42	3-hour 0.25° × 0.25° TRMM and Other-GPI Calibration Rainfall
2B31	Combined Rainfall Profile (PR, TMI)	3B43	Monthly 0.25° × 0.25° TRMM and Other Sources Rainfall

Fonte: Adaptado NASA (2017).

2.2.2 Vazão

2.2.2.1 Definição

Vazão, ou caudal, é o volume de água que passa por uma determinada seção de um rio ao longo de uma unidade de tempo. Assim, se o volume é dado em litros, e o tempo é medido em segundos, a vazão pode ser expressa em unidades de litros por segundo (L/s), sendo que no caso de rios, o mais usual a expressão de vazão em metros cúbicos por segundo (m^3/s) (Collischonn et al., 2015). Nos cursos d'água, o escoamento superficial é uma das principais formas pela qual eles recebem água, junto com a precipitação, o escoamento subsuperficial e o subterrâneo.

De acordo com ANA (2013), a formação do escoamento superficial está condicionada a vários fatores, sendo divididos em agroclimáticos (relacionados com o tipo e uso do solo) e fisiográficos (referentes ao relevo e área de drenagem da bacia). As condições de superfície que mais influenciam no escoamento superficial são o tipo de solo, topografia, rede de drenagem, obras hidráulicas presentes na bacia.

Além disso, é interessante salientar que os rios podem ser classificados, segundo o tempo de fluxo de água, em perenes, intermitentes e efêmeros. Rios perenes são os que possuem um fluxo de água durante a maior parte do ano, e onde os canais possuem uma forma definida. Os rios intermitentes são os que apresentam fluxo apenas durante a estação chuvosa, e que possuem canais definidos. Já os rios efêmeros são os que apresentam fluxo apenas durante os períodos de chuva, e não possuem canais bem definidos.

2.2.2.2 Medição de vazão

A vazão de um rio está intimamente relacionada com uma seção específica do mesmo, ou seja, depende do ponto da bacia hidrográfica analisada, uma vez que o rio vai recebendo contribuição da bacia hidrográfica ao longo de todo o seu trajeto. No processo de medição de vazão, as grandezas geométricas da seção a ser analisada são essenciais, como área, perímetro molhado, largura etc., além da velocidade.

De acordo com UFRGS (2009), dentre os principais métodos usuais de medição de vazão, pode-se citar uso de molinete, o método acústico, o método químico, o método com flutuadores,

o uso de dispositivos regulares, e indiretamente pela medição do nível da água (utilização da curva-chave do rio).

O método do molinete hidrométrico consiste em determinar a área da seção de um rio, e medir a velocidade em mais de um ponto da seção, obtendo-se assim a velocidade média vertical da seção, e conseqüentemente calculando-se a vazão. O método acústico consiste na utilização de aparelhos que fazem a análise do eco de pulsos de ultrassom refletidas pelas partículas sólidas em suspensão na massa líquida e pela superfície sólida do fundo. O método químico consiste em injetar uma substância concentrada e medir a concentração em um certo ponto a jusante. A utilização de flutuadores consiste na determinação da velocidade de um objeto flutuante, medindo o tempo que ele percorre um trecho de rio de comprimento conhecido. Dispositivos regulares podem ser utilizados também, como vertedores, calhas *Parshall*, que possuem uma relação entre o nível da água e a vazão que os atravessa conhecida.

Finalmente, o método pelo nível da água é utilizado por meio de uma relação entre o nível de uma seção de rio, e a vazão, que é regida por uma curva-chave, e de acordo com as leituras feitas diariamente, permitem a determinação de vazões regularmente para uma seção. A relação proposta pela curva-chave constitui uma simplificação, sendo desse modo considerada válida apenas para o caso de uma morfologia constante do rio, com uma geometria da seção que não se modifica com o tempo. É o método mais aplicado em postos e estações fluviométricas, e as formas de obtenção do nível da seção vão desde as manuais, realizadas por um operador regularmente, quanto através de dispositivos eletrônicos que realizam as leituras em tempo real.

2.2.2.3 Vazões naturais

Vazões naturais representam a vazão que deveria ocorrer em uma determinada seção do rio, caso não houvesse a influência de ações antrópicas na bacia contribuinte. Dentre essas ações, destacam-se as regularizações de vazões por reservatórios, os desvios de água, as evaporações em reservatórios, a irrigação, a criação animal e abastecimento humano.

O processo de cálculo das vazões naturais consiste numa reconstituição que considera a vazão observada no local, juntamente com suas informações relativas às ações provenientes de ação humana na bacia. No Brasil, o Operador Nacional do Sistema Elétrico – ONS é o órgão

responsável pelo cálculo e monitoramento de vazões naturais afluentes em reservatórios estratégicos do país.

2.3 MODELAGEM HDROLÓGICA

2.3.1 Definição

De acordo com Tucci (1998), um modelo hidrológico consiste numa ferramenta utilizada para representar processos que ocorrem numa bacia hidrográfica. Por exemplo, um modelo pode representar a vazão em um rio, baseado em parâmetros observados de precipitação. Os modelos são utilizados para a gestão de sistemas, e para a possibilidade de realizar prognósticos dele. Entra-se com os parâmetros necessários de entrada observados, e o resultado é o parâmetro desejado a ser previsto.

Os modelos podem ser classificados, de acordo com o tipo de variáveis utilizadas (estocásticos ou determinísticos), no sentido da relação entre as variáveis (empíricos ou baseados em processo), em relação a forma de representação dos dados (discretos ou contínuos), em relação a existência ou não de relações espaciais (pontuais ou distribuídos), e a existência de dependência temporal (estáticos ou dinâmicos).

Um modelo estocástico, possui pelo menos uma relação aleatória no comportamento das variáveis. Ou seja, se para uma entrada o modelo fornece apenas uma saída, ele é classificado como determinístico. Modelos empíricos são os que utilizam relações entre as variáveis baseadas em observações, já os baseados em processos, procuram explicar determinado fenômeno baseado na complexidade das relações em geral físicas ou conceituais existentes. Modelos contínuos são os que apresentam variáveis de dados com uma continuidade no tempo, baseado nas suas observações. Modelos concentrados ou pontuais são os que servem apenas para um determinado espaço geográfico (variáveis aleatórias de tempo e espaço), enquanto os distribuídos são mais gerais (variável aleatória como apenas tempo). E finalmente, modelos estáticos e dinâmicos tem a ver com a existência ou não de iterações no processo de modelagem a partir da entrada dos dados.

De acordo com Almeida e Serra (2017), exemplos de aplicação de modelos hidrológicos são na análise de consistência e preenchimento de falhas, na previsão de vazões por modelos

chuva-vazão e vazão-vazão, modelos que simulam a qualidade da água em rios e reservatórios, modelos que determinam a operação ótima em sistemas de reservatórios etc.

2.3.2 Modelos chuva-vazão

Um modelo chuva – vazão é aquele que a partir de dados de precipitação, simula vazões para uma determinada bacia hidrográfica. Exemplos de modelos do tipo comumente utilizados: SMAP (*Soil Moisture Accounting Program*) descritos por Lopes et al. (1981), o modelo *Tank Model* (Sugawara, 1961), o MODHAC (Modelo hidrológico auto calibrável), desenvolvido por Lanna Schwarzbach (1989), desenvolvido no Instituto de Pesquisas Hidráulicas da Universidade Federal do Rio Grande do Sul, e o *Soil and Water Assessment Tool* – SWAT (Arnold et al., 2012).

Mais recentemente, modelos baseados em redes neurais artificiais tem sido amplamente utilizado na análise chuva – vazão, como visto, por exemplo, nos trabalhos de Coulibaly et al. (2001), Jeong & Kim (2005), Farias et al. (2007), Wu & Chau (2011) e Santos et al. (2012a,b).

2.4 REDES NEURAIAS ARTIFICIAIS

2.4.1 Definição e estrutura das RNA

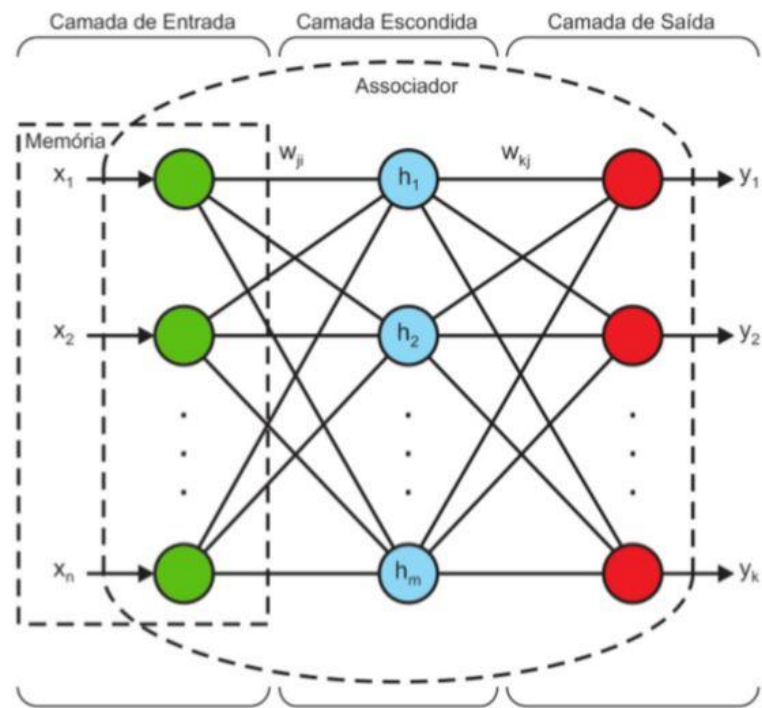
De acordo com Freire (2012), as Redes Neurais Artificiais (RNA) são definidas como ferramentas matemáticas, inspiradas no entendimento de que as funções neurais biológicas, como a memória, são armazenadas em neurônios, e nas conexões entre os mesmo, e desse modo, o processo de aprendizado consiste no estabelecimento de novas conexões entre os neurônios, e em modificações nas conexões existentes.

Os neurônios artificiais, são definidos como uma estrutura lógico-matemática que procuram simular a forma, o comportamento e as funções de um neurônio biológico. Fazendo uma analogia do biológico com o artificial, os dendritos consistem nas entradas, as sinapses são os pesos, os estímulos captados pelas entradas são a função de soma, e o limiar de disparo do neurônio e a função de transferência.

De acordo com Freire (2012), nas RNA, as entradas são os sinais iniciais que são responsáveis pela ativação da RNA, os pesos são os valores de ligação, ajustados pelo processo de formação. As tendências são os vieses, ajustados pelo processo de formação, com o objetivo de aumentar ou diminuir a entrada para a estrutura. O campo induzido são as somas de operações matemáticas realizadas com os pesos e vieses do neurônio artificial. A função de ativação restringe a faixa de amplitude de um neurônio ao valor finito como uma função do campo produzido. E finalmente, a saída é o resultado apresentado pelo neurônio da RNA.

Em relação a estrutura, as RNA podem ser entendida como conjuntos de neurônios artificiais, que são interligadas por canais de comunicação, representados por um correspondente numérico. No caso de uma estrutura ser composta por três camadas, por exemplo, elas representam as entradas, onde os dados são introduzidos na rede, a camada intermediária, onde ocorre o processamento de dados, e finalmente uma camada de saída, onde os resultados do processamento são produzidos. Na prática, as redes podem ocorrer com várias camadas intermediárias. Matematicamente, a transmissão de informações entre as camadas de uma rede neural é feita através de uma função de combinação e uma função de ativação. Por exemplo, se os dados trabalhados tiverem uma relação linear, ela representa o somatório da entrada com seus respectivos pesos. Na Figura 4 é mostrado um exemplo da estrutura de uma rede RNA.

Figura 4 – Concepção de uma rede neural artificial.



Fonte: Freire (2012).

2.4.2 Considerações sobre o treinamento das RNA

Em relação ao processo de treinamento, os mesmos podem ser classificados para as RNA como sendo principalmente supervisionados e não supervisionados. O treinamento supervisionado é o que se deve dispor das amostras e das respectivas saídas desejadas para que os pesos sejam ajustados continuamente durante a aprendizagem. O não supervisionado consiste no tipo em que a saída deve se auto organizar em relação às particularidades do conjunto de amostras, e assim identificar subconjuntos similares.

Além disso, é interessante destacar que o processo de treinamento também é formado por ciclos, que são uma apresentação de todos os pares de entrada e saída do conjunto treinamento no processo de aprendizado da RNA. Nesse sentido, a correção dos pesos num ciclo pode ser executada de dois modos, i.e., o modo padrão, e o modo *batch*. O modo padrão consiste no que a correção dos pesos acontece a cada apresentação a rede de um exemplo do conjunto de treinamento. Assim, dentro de um ciclo, ocorre o número de pares entrada e saída de correções. Já o modelo Batch realiza apenas uma correção no ciclo onde todos os elementos do conjunto de treinamento são apresentados à rede, e as correções são feitas com o erro médio.

2.4.3 Vantagens das RNA

Assim, de acordo com Freire (2012), como diferenciais, pode-se citar que as RNA são capazes de representar relações fortemente não lineares entre variáveis, levando à representação de sistemas complexos com uma precisão aceitável. Além disso, as redes podem extrair modelos e assim generalizar o mapeamento de valores de entrada e saída. E finalmente, elas podem ser retreinadas fazendo com que possam ser facilmente ajustáveis em tempo real. Desse modo, elas são amplamente utilizadas em diversos campos da ciência que se referem a estudos de previsão de séries temporais.

2.4.4 Aplicações de RNA na previsão de séries temporais

Na literatura, pode-se encontrar um vasto registro de aplicações de RNA. Diversos ramos científicos como a medicina, robótica, engenharia, finanças, segurança etc. fazem uso das mesmas, conforme Haykin, (1999) e Demuth et al. (2008). Na área de recursos hídricos, a utilização de RNA é notada sobretudo na previsão de variáveis hidrológicas como chuva, vazão, evapotranspiração, níveis freáticos, variáveis que indicam qualidade a água etc.

Noureddine et al. (2019) fizeram uso RNA na modelagem espacial da salinidade do solo, enquanto que Silva et al. (2019) usaram a RNA para a previsão do uso e ocupação do solo em uma bacia do semiárido Paraibano. Na previsão de sedimentos e de erosão em bacias hidrográficas pode-se destacar os estudos de Farias et al. (2010), Farias et al. (2015). Os estudos de Adeloye (2011) se concentraram na utilização da rede para a previsão de evapotranspiração. Bayatvarkeshi et al. (2018) estudaram a previsão de umidade relativa por meio da técnica. Mirabbasi et al. (2018) fizeram previsões de chuva por meio de RNA. Gorgij et al. (2008) fizeram a previsão de água subterrânea por meio da RNA. No que se refere a previsão de vazões em reservatórios, pode-se citar os estudos de Santos et al. (2014), Honorato et al. (2019), Macedo et al. (2019) e Santos et al. (2019).

2.5 MAPAS AUTO-ORGANIZÁVEIS

2.5.1 Definição

Mapas auto-organizáveis (do inglês *Self-Organizing Maps* – SOM) são redes neurais artificiais não supervisionadas que agrupam dados de entrada em classes de acordo com suas semelhanças por meio de métodos de treinamento competitivo (Kohonen, 1982; Haykin, 1999; Silva et al., 2010). São também conhecidos como redes auto-organizáveis de *Kohonen* (*Kohonen self-organizing map* - *KSOM*), e foram propostas por Kohonen (1982). De acordo com Kohonen et al. (1996), elas são o tipo de rede neural mais utilizado.

De acordo com Back et al. (1998), o principal objetivo do SOM é transformar um padrão de sinal de entrada de dimensão arbitrária em um mapa discreto bidimensional. Para isso, a rede agrupa os dados de entrada padrões de tal maneira que padrões semelhantes são representados pelos mesmos neurônios de saída, ou por um de seus vizinhos. O SOM é formado por uma matriz de unidades organizadas numa grade, e onde cada uma dessas grades são chamadas de neurônios.

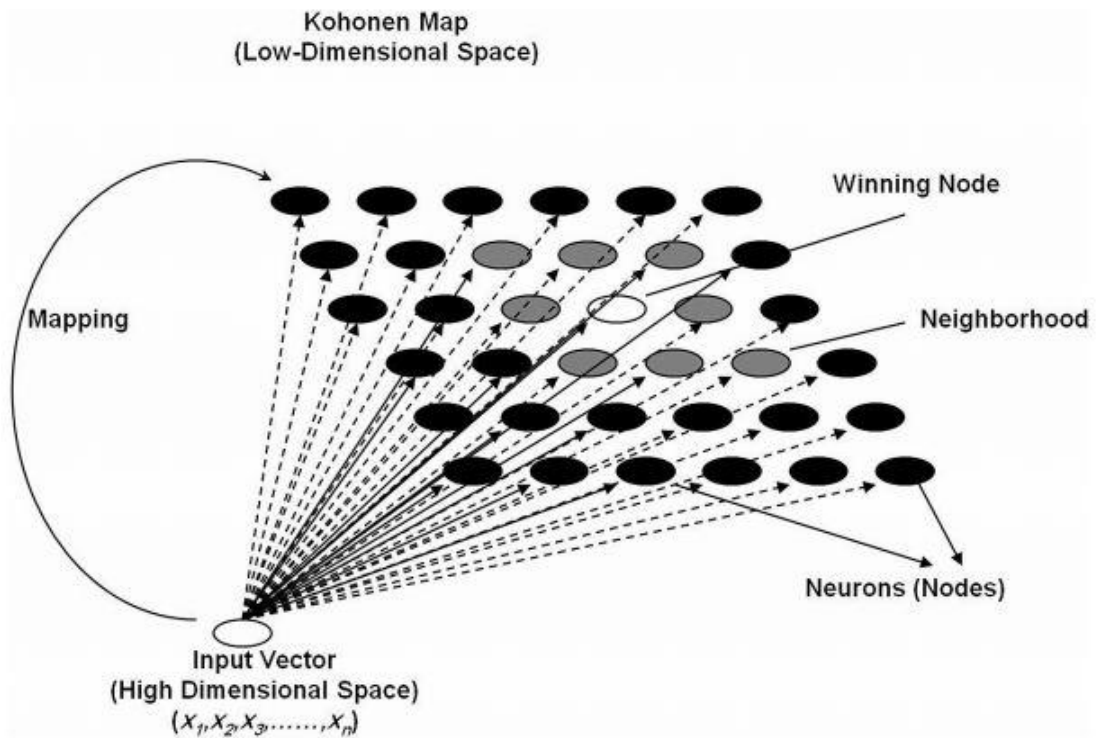
Nesse sentido, segundo Kohonen et al. (1996), o SOM pode ser também entendido como uma ferramenta para reduzir a quantidade de dados agrupando, assim, convertendo uma complexa relação estatística não linear entre dados de várias dimensões, numa relação bidimensional, por exemplo. É importante salientar que o agrupamento e mapeamento realizado pela rede SOM preserva o relacionamento topológico e métrico mais importante dos elementos dos dados originais, o que implica que não são perdidas muitas informações durante o mapeamento dos dados. Nesse sentido, essa capacidade de reduzir um conjunto de dados multidimensionais a uma matriz bidimensional de características, faz com com que possa ser utilizada para fins de análise e de previsão.

2.5.2 Estrutura

A estrutura de uma rede SOM é formada por duas camadas, sendo a camada multidimensional de entrada, e a camada de saída, chamada também de camada competitiva. Esta é geralmente unidimensional ou bidimensional. Na camada de saída os neurônios

competem entre si para ser escolhido o neurônio vencedor. Ou seja, é escolhida a classe mais adequada para um determinado vetor de entrada. Na Figura 5 é mostrado um esquema da arquitetura de uma rede SOM, e nela é possível ver que as duas camadas são interconectadas.

Figura 5 – Estrutura de um modelo SOM com detalhe do neurônio vencedor, e seus neurônios vizinhos.



Fonte: Adeloje et al. (2011).

De acordo com Garcia e Gonzalez (2004), o número de neurônios da rede pode ser definido pela Eq. (1).

$$M = 5\sqrt{N} \quad \text{Eq. (1)}$$

onde N é o número de dados da amostra. Outra orientação de Garcia e Gonzalez (2004), é que com M conhecido, o número de linhas e colunas no SOM pode ser determinado pela Eq. (2):

$$\frac{l_1}{l_2} = \sqrt{\frac{e_1}{e_2}} \quad \text{Eq. (2)}$$

onde l_1 e l_2 são o número de linhas e colunas, respectivamente, e_1 é o maior autovalor do conjunto de dados de treinamento, e e_2 é o segundo maior autovalor.

2.5.3 Treinamento

Assim, de acordo com o proposto por Kohonen (1996), os dados de entrada multidimensionais são primeiro padronizados deduzindo a média e depois dividindo o resultado pelo desvio padrão (normalizados). Em seguida, são dados valores gerados aleatoriamente aos neurônios na camada de saída. Um vetor de entrada padrão é então escolhido aleatoriamente e apresentado a cada um dos neurônios individualmente do SOM para comparação com seus vetores de código, com o objetivo de identificar o vetor de código mais semelhante ao vetor de entrada apresentado. Nessa identificação é utilizada a distância euclidiana, que é definida pela Eq. (3):

$$DI_i = \sqrt{\sum_{j=1}^n m_j (x_j - w_{ij})^2} \quad \text{Eq. (3)}$$

onde x_j é o j -ésimo componente do vetor de entrada x ; n é a dimensão do vetor de entrada x ; i variando de 1 até M , e M sendo o número total de neurônios na camada de saída, e m_j sendo a máscara da função (geralmente atribui-se $m_j = 1$), ou, para fins de exclusão da contribuição do elemento x_j no cálculo da distância euclidiana, muito utilizado quando a variável de entrada possui valores faltando.

Assim, o neurônio cujo vetor se aproxima mais do vetor de dados de entrada (ou seja, para o qual o D_i é mínimo) é escolhido como o neurônio vencedor, ou a melhor unidade correspondente (do inglês *best matching unit* - BMU). Os vetores de código deste sistema do nó vencedor e os dos seus neurônios adjacentes são então ajustados para combinar os dados de entrada por meio da Eq. (4), trazendo assim os vetores do código ainda mais de acordo com o vetor de entrada (Vesanto et al., 2000).

$$w_{ij}(t + 1) = m_i(t) + \alpha(t)h_{bi}(t)[x - m_i(t)] \quad \text{Eq. (4)}$$

onde t é o tempo, $\alpha(t)$ é a taxa de aprendizado y , e $h_{bi}(t)$ é a função da vizinhança entre cada neurônio i e a melhor unidade b correspondente. Dessa maneira, cada nó no mapa desenvolve internamente a capacidade de reconhecer vetores de entrada semelhantes a ele. Esta característica é chamada de auto-organização, porque não há informações externas que são fornecidas para levar a uma classificação (Penn, 2005).

De acordo com Farias et al. (2013), a regra de *Kohonen* força os pesos ligados ao neurônio vencedor e seus vizinhos a se deslocarem na direção do vetor de entrada apresentado à rede, fazendo com que a distância Euclidiana se torne cada vez menor e com que estes neurônios aprendam a classificar vetores similares. Segundo Adeloje et al. (2011), o processo de comparação e ajuste continua até que o número ideal de iterações seja alcançado ou que os critérios de erro especificados sejam atingidos. Tanto a taxa de aprendizado quanto a função de vizinhança afetam a eficácia da aprendizagem do SOM e deve ser escolhida com cuidado. Em particular, a taxa de aprendizado diminui monotonicamente com o aumento do número de iterações, conforme a Eq. (5).

$$\alpha(t) = \alpha_0 \left(\frac{0.005}{\alpha_0} \right)^{\frac{t}{T}} \quad \text{Eq. (4)}$$

onde α_0 é a taxa inicial de aprendizagem e T é o treinamento, forçando assim o vetor de peso convergir. A função de vizinhança é normalmente escolhida como sendo a função gaussiana centrada na unidade vencedora c , conforme a Eq. (6).

$$h_{ci}(t) = e^{\left(-\frac{\|r_c - r_i\|^2}{2\sigma^2(t)} \right)} \quad \text{Eq. (6)}$$

onde r_c e r_i são as posições dos nós c e i na grade SOM e $\sigma(t)$ é o raio da vizinhança. Como o aprendizado taxa $\alpha(t)$, $\sigma(t)$ também diminui monotonicamente à medida que o número de iterações aumenta.

A qualidade da rede SOM treinada é medida pelo total erro médio de quantização e erro topológico total, dado pela Eq. (7).

$$q_e = \frac{1}{N} \sum_{i=1}^N \|X_i - W_c\| \quad \text{Eq. (7)}$$

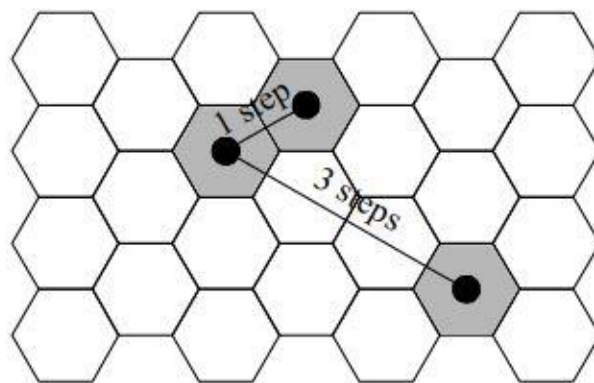
onde q_e é o erro de quantização, X_i é a i -ésima amostra de dados ou vetor, W_c é o vetor protótipo da BMU para X_i e $\| \cdot \|$ denota a distância euclidiana. O erro topológico é dado pela Eq. (8).

$$t_e = \frac{1}{N} \sum_{i=1}^N u(X_i) \quad \text{Eq. (8)}$$

onde u_i é um número inteiro binário tal que seja igual a 1 se o primeiro e o segundo BMUs para X_i , não são sejam unidades adjacentes. Caso contrário, é zero.

Ainda de acordo com Farias et al. (2013), a apresentação de vetores de entrada para a rede também pode ser feito usando todo o conjunto de dados antes de qualquer peso ser atualizado aos neurônios. Essa forma de apresentação é conhecida como modo *batch*. Nesse caso, a busca pelo neurônio vencedor é realizada para cada vetor de entrada e o vetor de pesos é movido para uma posição específica calculada pelo média de vetores de entrada para os quais o neurônio será o vencedor ou vizinho do vencedor. Além disso, ele também afirma que os pesos tendem a se estabilizar após várias iterações do conjunto de dados de entradas.

Figura 6 – As distâncias entre os neurônios de um modelo SOM para a determinação da vizinhança.



Fonte: Farias et al. (2013).

E finalmente, o treinamento da SOM ocorre em duas fases: a *ordering phase* e a *tuning phase* de afinação. Na primeira, o treinamento é limitado por um determinado número de apresentações dos dados e o raio da vizinhança começa com uma dada distância que diminui até um. Esta medida permite que os pesos dos neurônios se auto organizem no espaço de entrada, de forma consistente com as suas respectivas posições. A segunda fase dura o número restante de apresentações para o treinamento definido. Nesta, o raio da vizinhança fica abaixo de um, e assim, só há atualização com o peso do neurônio vencedor. E conforme Beale et al. (2012), durante essa fase é esperado que os pesos se modificarão relativamente uniformemente no espaço de entrada, mantendo a topologia definida na primeira fase.

2.5.4 Previsão utilizando SOM

E assim, uma vez com a rede SOM treinada, é possível utilizar o modelo como uma ferramenta para previsão, ou cálculo de variáveis. Para esse propósito, de acordo com Farias et al. (2013), deve-se considerar no vetor de entrada a ausência da variável a ser fornecida pelo modelo através dos passos:

- a) Calcule as distâncias euclidianas entre os vetores de entrada e pesos anexados à saída neurônios desconsiderando o elemento j a ser fornecido. Isso pode ser feito incluindo uma variável booleana m_j , como mostrado pela Eq. (3) A variável m_j é usada para que inclua ($m_j = 1$) ou exclua ($m_j = 0$) a contribuição de um dado elemento j do vetor de entrada no cálculo das distâncias euclidianas;
- b) Determina-se o neurônio vencedor com base na menor distância euclidiana;
- c) Utiliza-se o peso do neurônio vencedor conectado ao elemento ausente j do vetor de entrada como o valor previsto.

2.5.5 Aplicações do SOM em séries temporais

As redes SOM são amplamente utilizadas na modelagem e análise de séries temporais. e modelagem. Ele pode ser usado em diversas aplicações práticas, como a redução da quantidade de dados de treinamento para identificação do modelo, interpolação não linear e extrapolação (ou seja, previsão), generalização e compactação de informações para transmissão

fácil (Kangas & Simula, 1995; Kohonen et al., 1996; Parasuraman et al., 2006; Tananaki et al., 2007; Chang et al., 2010; Adeloje et al., 2011; Rivera et al., 2011; Adeloje e Rustum, 2012).

Como exemplos da aplicação de SOM no campo da engenharia, e especificamente da hidrologia, García e González (2004), aplicaram o SOM para entender o comportamento das variáveis envolvidas nos processos de funcionamento de uma estação de tratamento de águas residuais. Norvarini et al. (2019) fizeram uso da capacidade de agrupamentos da rede SOM nas etapas de modelagem do gerenciamento ótimo das pressões em redes de abastecimento de água. Voutilainen e Arvola (2017) aplicaram o SOM no agrupamento de 21 anos de dados ambientais complexos para um pequeno lago boreal intocado, e como resultados, como resultados, concluiu-se que ele é um método útil para agrupar as variáveis de um grande conjunto de dados multidimensionais, e assim possibilitar análises mais efetivas.

Adeloje et al. (2011) aplicaram SOM para prever a evapotranspiração de referência em certas plantações com base em dados climáticos diários observados em duas bacias climáticas distintas (clima temperado e semiárido), tomando-se como referência a modelagem pela formulação convencional da *Food and Agricultural Organization* – FAO, a *Penman-Monteith*, empregando o conjunto completo de dados meteorológicos nos dois locais. Os resultados indicaram eficácia de utilização da rede para a previsão da evapotranspiração, e ainda por cima, a capacidade de modelagem inclusive na ausência de determinadas variáveis de entrada.

Farias e Santos (2014) e Farias et al. (2015) utilizaram a rede SOM na modelagem vazão-erosão. Farias e Santos (2014) fizeram uma comparação entre a utilização de uma regressão linear múltipla com o SOM na modelagem da erosão, utilizando variáveis de entrada a chuva, a duração da chuva, a vazão do corpo d'água, a máxima e mínima temperatura diária, e o período de dias secos. E assim, e os resultados indicaram uma superioridade do mesmo em frente a regressão linear múltipla, produzindo resultados mais confiáveis, devido a sua característica de detecção e extração de tendências não lineares.

Farias et al. (2015) compararam algumas estruturas SOM com o objetivo de estimar a produção de sedimentos com base em dados de vazão e climatológicos na escala da microbacia. Foram utilizados dados de vazão, altura média da vegetação, duração da chuva, intensidade da chuva, número de dias secos e a quantidade de chuva total como variáveis de entrada, e os resultados indicaram SOM mostrou-se muito prático na aplicação em determinadas bacias, e concluíram que a metodologia também exige dados de longo prazo e recalibração dinâmica com informações atualizadas de possíveis mudanças na bacia.

Na aplicação da modelagem chuva-vazão, tem-se como exemplos os trabalhos propostos por Adeloye e Rustum (2012), Farias et al. (2012), Farias et al. (2013), Nourani et al. (2013) e Filho e Farias (2018). Adeloye e Rustum (2012) fizeram a aplicação do SOM na modelagem chuva-vazão para a reconstrução de vazões em bacias inadequadamente compostas por postos pluviométricos. Os resultados mostraram sucesso na utilização dessa ferramenta para vencer problemas práticos na espacialização de postos pluviométricos em bacias.

Nourani et al. (2013) aplicaram o SOM para a identificação de agrupamentos espacialmente homogêneos de dados de satélite de precipitação e para escolher os dados mais operacionais e eficazes para uma rede neural de *feed-forward* (FFNN) para modelar processo chuva-vazão em escala de tempo diária. O modelo proposto de FFNN, associado ao método de agrupamento SOM, diminuiu a dimensionalidade das variáveis de entrada e, conseqüentemente, a complexidade dos modelos de FFNN. Salienta-se também que nesse trabalho, Nourani et al. (2013) fizeram uso também da transformada *wavelet* (WT) também foi usada para extrair recursos dinâmicos e em várias escalas das séries temporais de escoamento não estacionário e para remover ruídos. Ou seja, foi realizada uma aplicação SOM-FFNN e SOM-WT-FFNN, que mostraram a aplicabilidade do SOM nesse tipo de previsão.

Farias et al. (2012), Farias et al. (2013), e Filho e Farias (2018) aplicaram o SOM em bacias hidrográficas para realizar a previsão de vazões a nível mensal, baseada nas precipitações e vazões do próprio mês, e de um e dois meses anteriores. O modelo demonstrou sucesso na modelagem, bem como foi útil ao demonstrar as correlações entre as variáveis de entrada e os resultados obtidos. Ou seja, nesse sentido, demonstrou-se a forte correlação entre os dados de precipitação, sobretudo o do mês corrente e o anterior na modelagem.

3 MATERIAIS E MÉTODOS

3.1 ÁREA DE ESTUDO

A bacia do rio São Francisco é considerada a terceira mais importante bacia hidrográfica brasileira, e é dividida em quatro sub-bacias, i.e. Superior, Sub-Médio, Médio e Inferior (Santos e Morais 2013; Santos e Silva 2013). O presente estudo se concentra na área da sub-bacia do Alto São Francisco, que é considerada uma área estratégica para o gerenciamento dos recursos hídricos do Brasil (Santos et al., 2019). A sub-bacia está localizada entre as coordenadas 18.125°-20.875° S e 43.875°-46.625° O, conforme é mostrado na Figura 7.

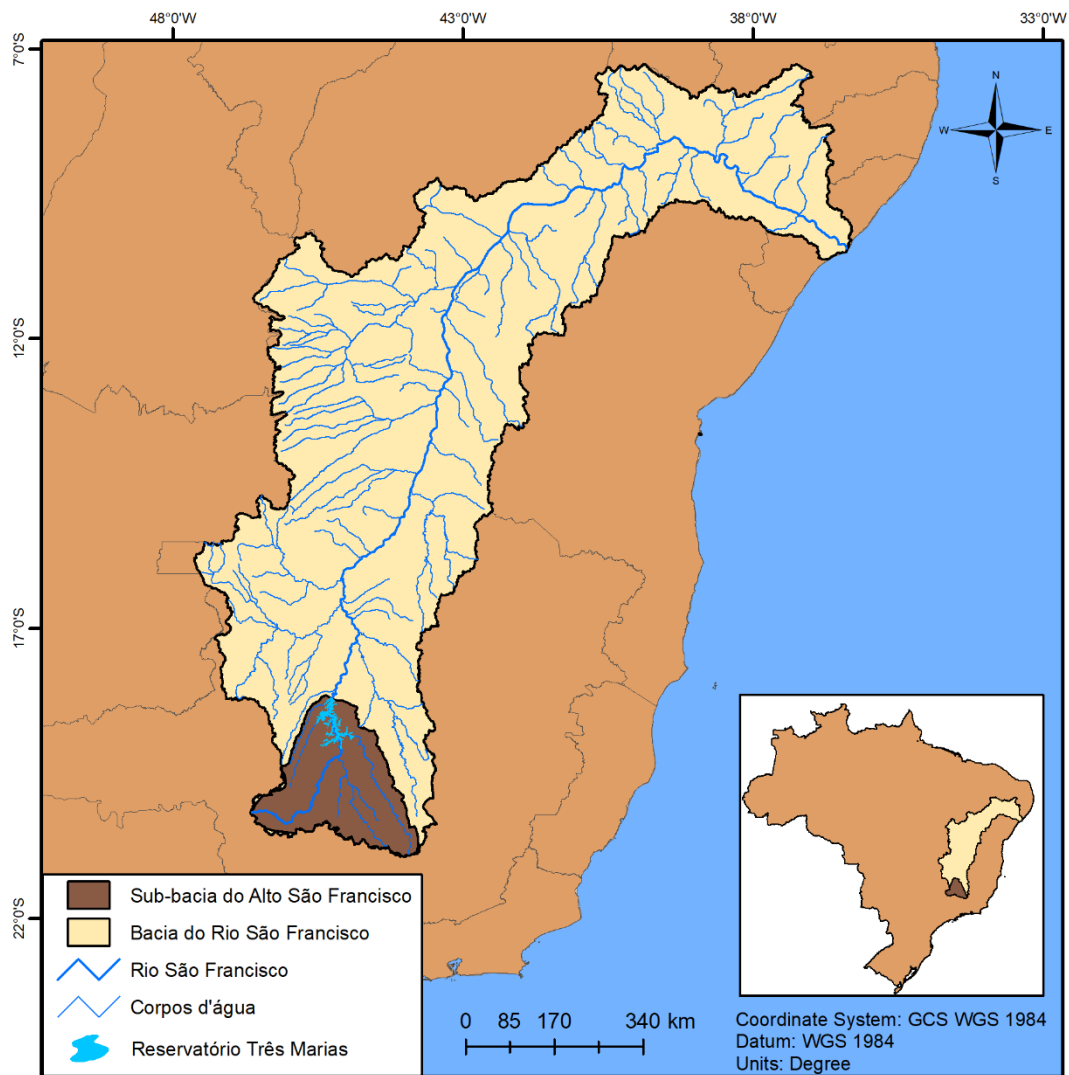
Na Figura 7, é mostrado um mapa com a localização da área de estudo (mostrada num tom de marrom escuro), dentro da bacia hidrográfica do rio São Francisco (num tom da cor creme). Na figura também são mostradas as áreas de drenagem principal da bacia, juntamente com o rio principal destacado (que é o rio São Francisco), e o reservatório Três Marias, que conforme será explicado mais na frente faz parte da área de estudo do presente trabalho. A área foi escolhida para ser estudada devido a sua relevância no cenário hídrico brasileiro, e pela disponibilidade de dados de vazão afluente observada no reservatório trabalhado para ser utilizado nas modelagens. Destaca-se ainda a perenidade do rio de onde as vazões afluentes são observadas.

A sub-bacia do Alto São Francisco possui uma área de cerca de 49.574 km², uma área consideravelmente grande, sendo maior do que países como a Holanda, a Dinamarca, ou a Suíça, por exemplo. Em relação ao relevo da região, destaca-se que ela possui uma topografia ondulada, com elevações que variam de 600 a 1600 m, conforme a Figura 8. A Figura 8 mostra a espacialização das elevações da região, e é representada com um tom que vai desde o verde (cerca de 600 m) até o laranja escuro (1600 m).

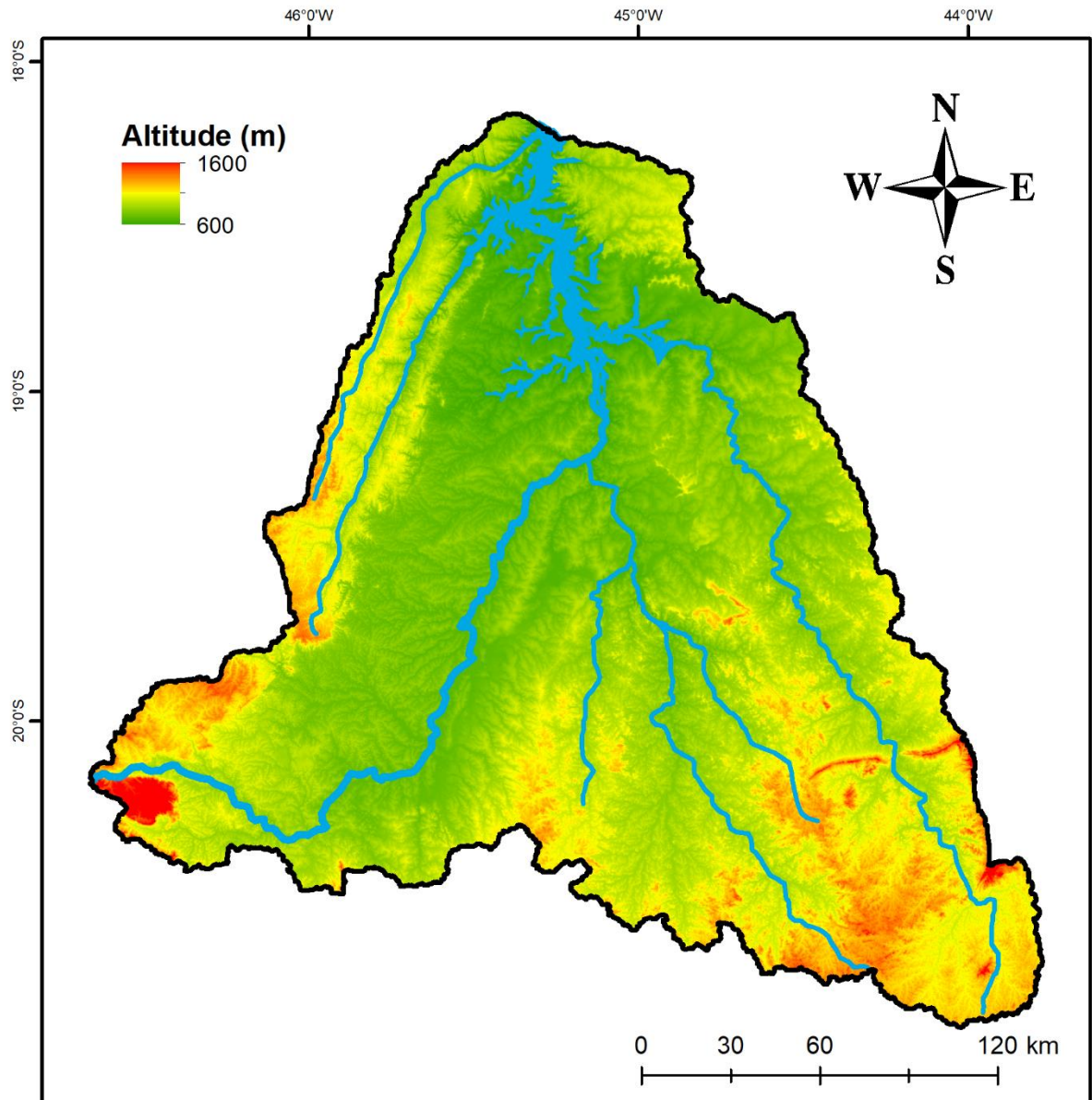
Essa área de estudo é localizada no bioma cerrado, e possui um uso e ocupação do solo da região variando de vegetação savânica, áreas de agricultura, áreas de pastagem, áreas urbanas, terra estéril e até florestas (Silva et al., 2018). Na Figura 9 é mostrado o uso e ocupação do solo da região, e pela característica mostrada percebe-se uma predominância dos tipos floresta e agropecuária na região. Os tons de vermelho representam a área urbanizada, e é notável uma proporção bastante pequena desta em relação a área total da região, destacando que a mesma bacia não pode ser classificada como uma bacia urbana.

O clima da região pode ser classificado como tropical úmido, pela classificação de Köppen, apresentando duas estações distintas bem definidas, sendo um verão chuvoso (outubro-janeiro), e um inverno seco (junho-agosto), com uma média de precipitação anual variando entre 1.100 a 1.600 mm. Na estação chuvosa é a que ocorre cerca de 85% da precipitação anual. Na Figura 10 é mostrada uma espacialização da precipitação média anual na região. A temperatura média anual da região é de cerca de 22°C e a evaporação é de cerca de 1.000 mm por ano (MMA, 2017).

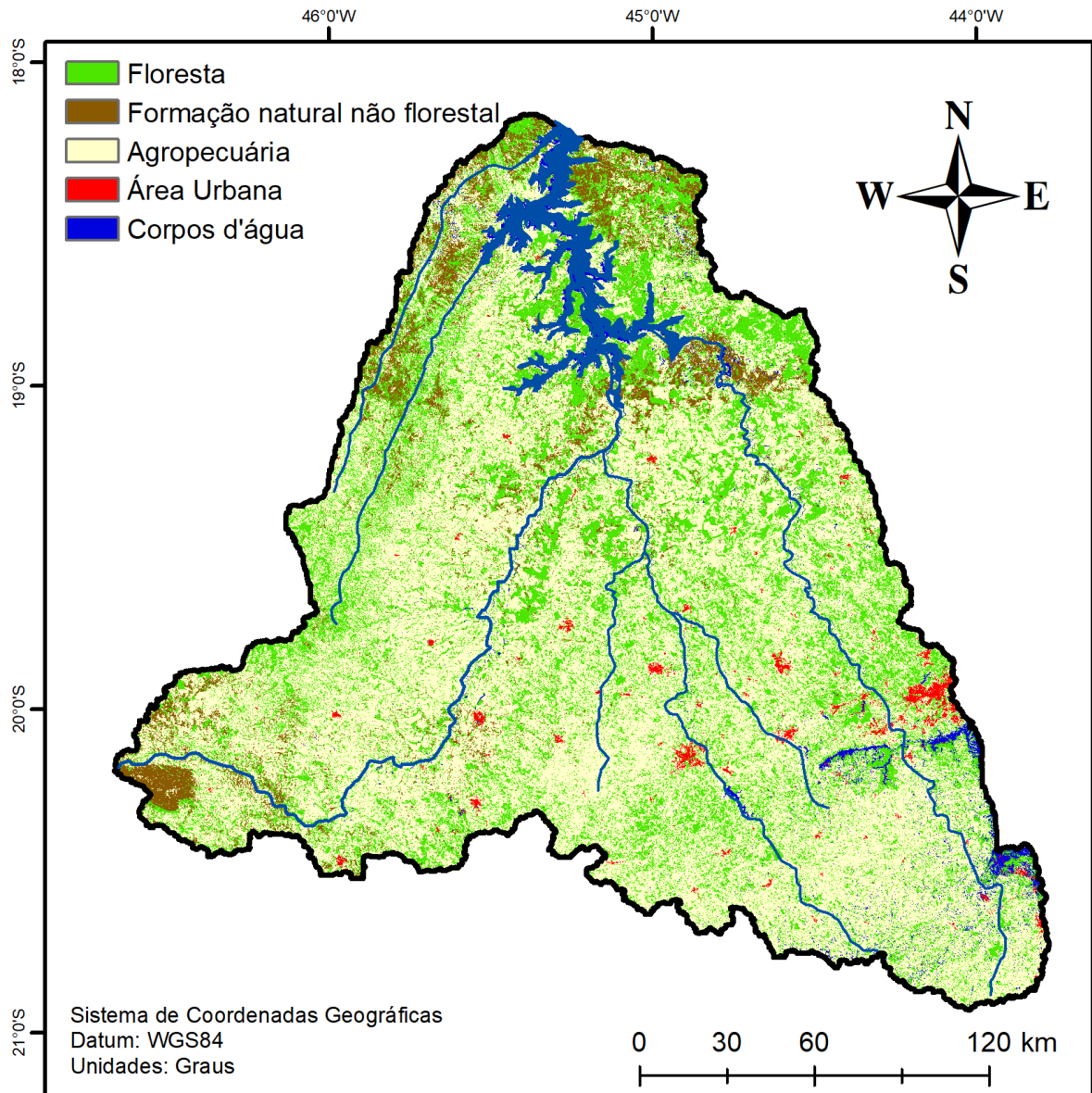
Figura 7 – Localização da sub-bacia do Alto São Francisco.



Fonte: Elaborado pelo autor.

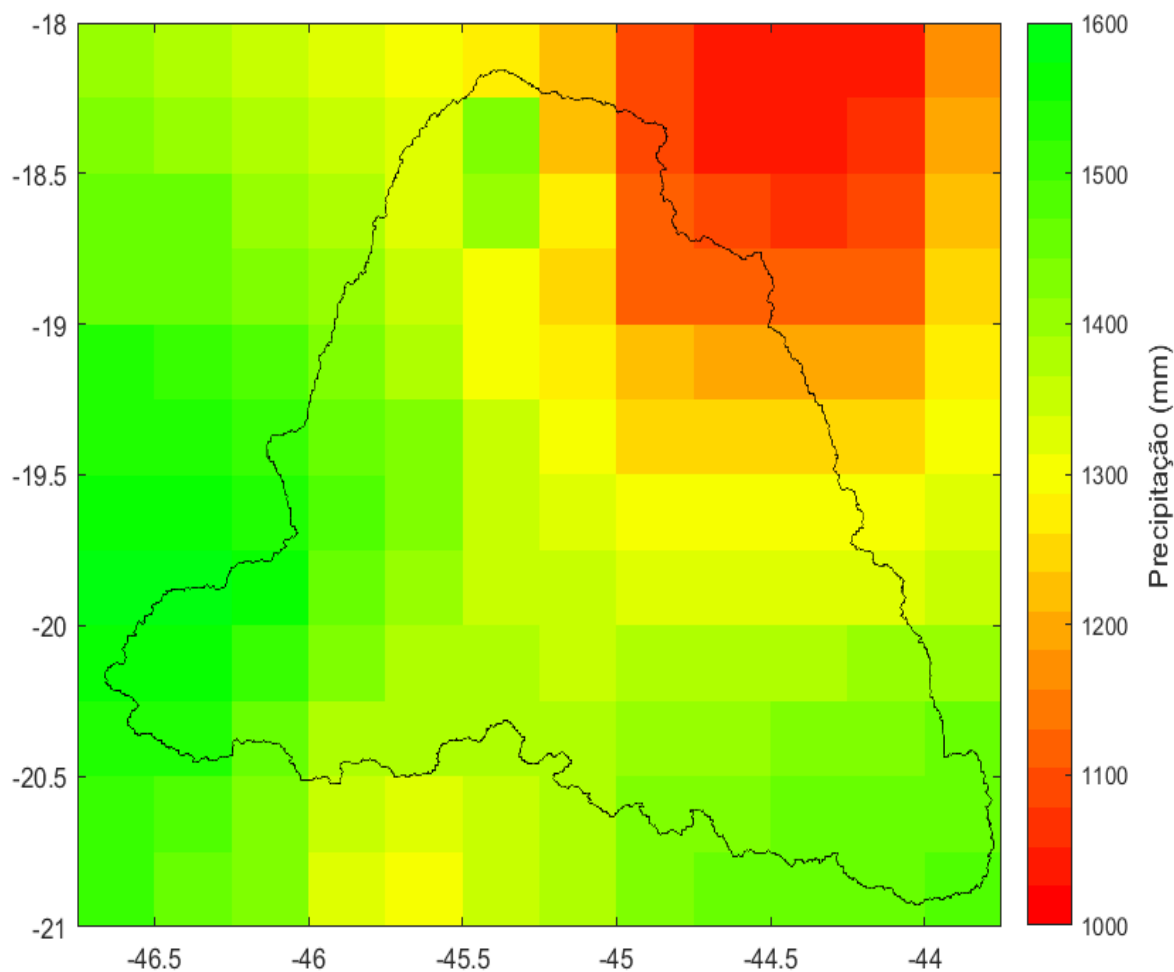
Figura 8 – Topografia da região.

Fonte: Elaborado pelo autor.

Figura 9 – Uso e ocupação do solo.

Fonte: Elaborado pelo autor.

Figura 10 – Precipitação anual média na região (1998-2018) baseado em dados TRMM.



Fonte: Elaborado pelo autor.

O reservatório dos quais as vazões afluentes são estudadas pode ser observado por uma imagem de satélite na Figura 11, e é conhecido como reservatório Três Marias, e corresponde a uma hidroelétrica, que possui uma barragem com comprimento de cerca de 2.700 metros, e uma capacidade de cerca de 21 bilhões de metros cúbicos de água, localizado cerca de 2.221 km acima da foz do rio. Ele é a maior obra regularizadora de vazão da região (CEMIG, 2020). Pela figura fica notória a grande extensão do reservatório, destacando que é uma obra de proporções grandes para a região onde está localizada.

Figura 11 – Imagem de satélite do reservatório Três Marias.



Fonte: Google Earth Pro (2020).

3.2 DADOS DE PRECIPITAÇÃO

Os dados de precipitação para a área de estudo foram obtidos através do satélite TRMM (*Tropical Rainfall Measuring Mission*) da NASA, algoritmo 3B42 (versão 7) que coleta informações sobre o clima das regiões tropicais e subtropicais. Devido ao funcionamento do satélite ter começado no fim do ano de 1997, a nossa série foi extraída a partir do início do ano de 1998 (01 de janeiro de 1998), até 31 de agosto de 2019, o que totalizam 260 meses de análise.

Para a análise, foram utilizados os pontos das quadrículas TRMM contidas dentro da área de estudo. Cada uma das quadrículas é equidistante a cada $0,25^\circ$ (centroides). A grade inicial para a área de estudo era composta por 144 quadrículas, mas apenas as que estavam efetivamente dentro da área da sub-bacia foram utilizadas, restando 91 para a análise dos dados

TRMM, conforme é mostrado na Figura 12. O vetor de entrada para os dados de precipitação consistiu numa matriz com 91 colunas, representando os centroides dos pontos do TRMM, por 260 linhas, onde cada uma representa cada um dos meses de análise.

Os dados de precipitação foram multiplicados pela área em relação ao percentual de uma quadrícula completa, de acordo com a área que ela ocupava em relação a sub-bacia. Ou seja, quadrículas como a de número 54, ou 55 tiveram a precipitação observada multiplicada por um, antes de utilizá-la na modelagem, enquanto quadrículas como a 7 e a 28 tiveram um valor bem próximo de 0 ao multiplicar a precipitação da própria quadrícula. Assim, utilizou-se dados de precipitação mais coerentes com a área de estudo.

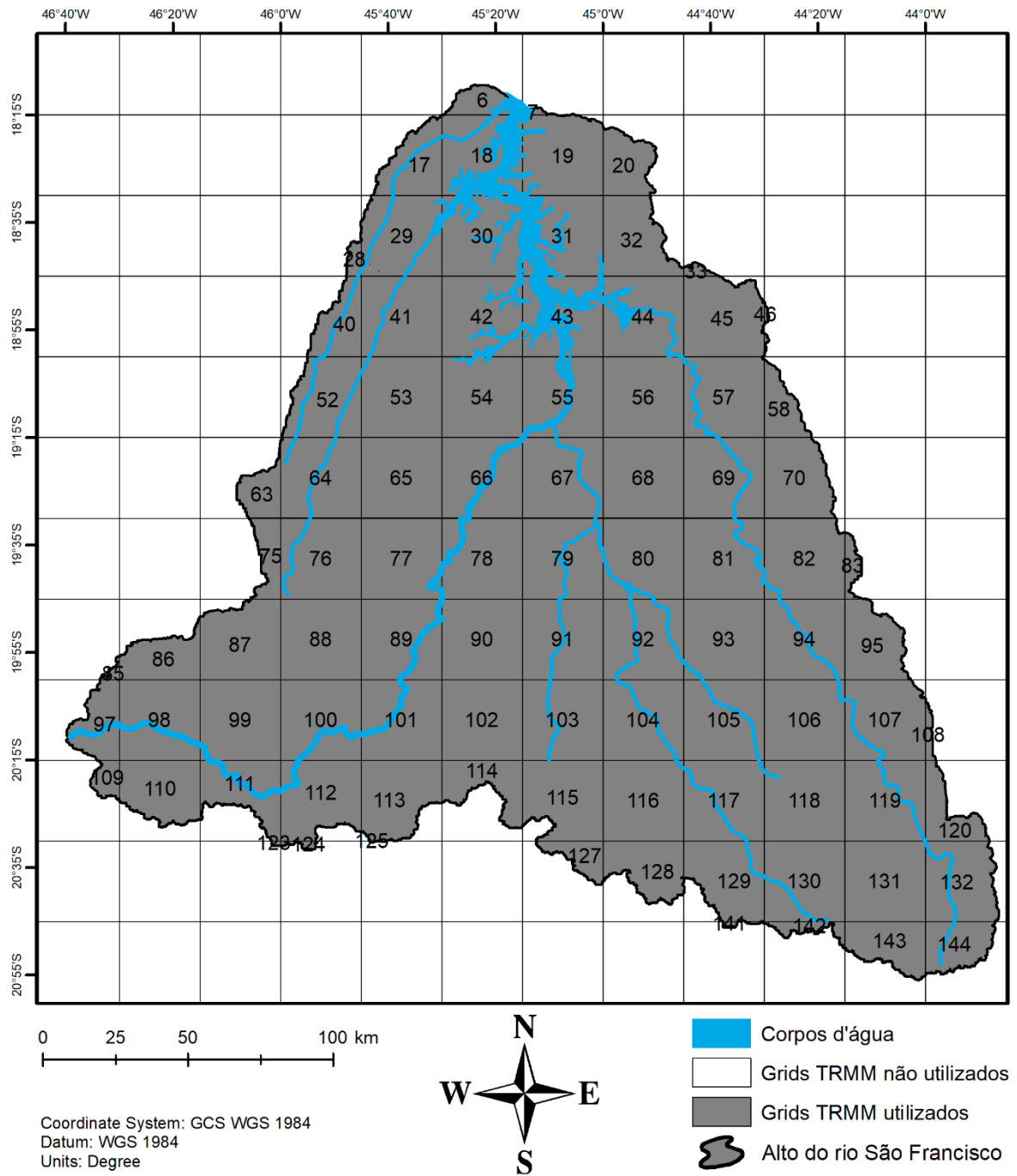
Na Figura 13 é mostrada a série de precipitações mensais para a área de estudo, e na Tabela 1 os índices estatísticos principais para as variáveis (média, soma, desvio padrão etc.).

Tabela 1 - Índices estatísticos para as precipitações observadas.

Índices estatísticos	Valor
Média (mm)	115,49
Mediana (mm)	73,65
Máximo (mm)	437,46
Desvio padrão (mm)	111,87
Variância (mm) ²	12.515,58

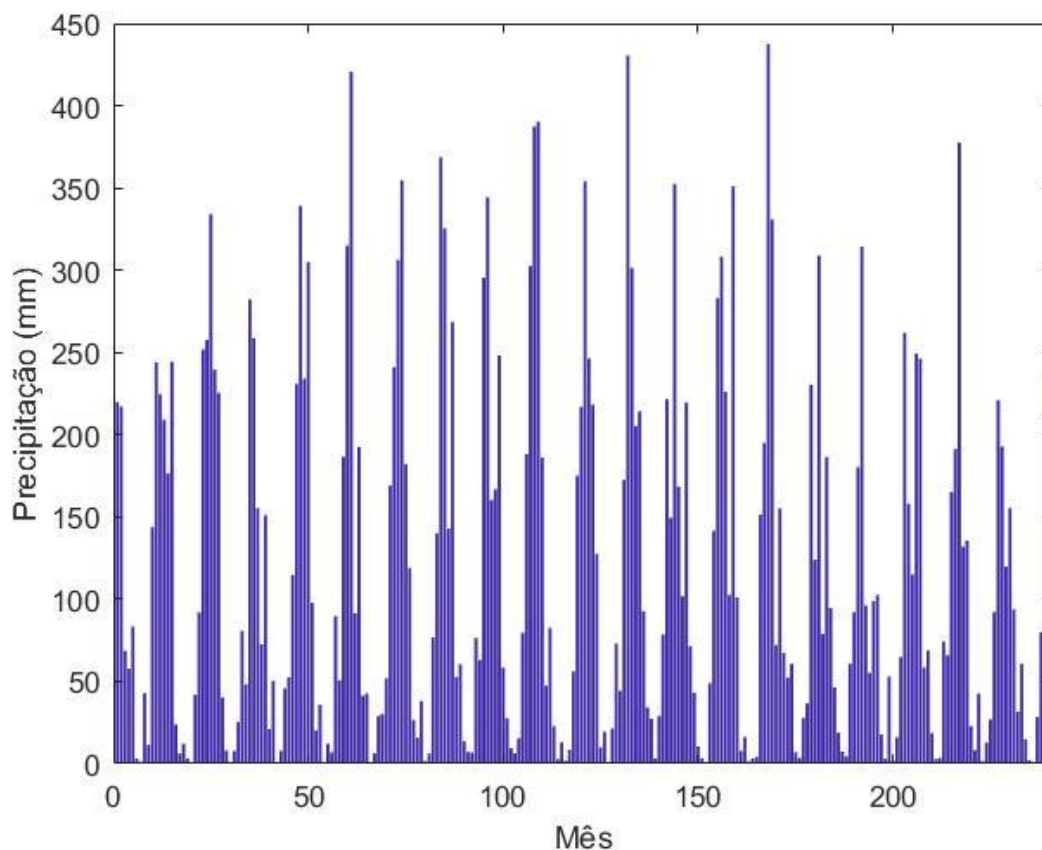
Fonte: Elaborado pelo autor.

Figura 12 – Espacialização das quadrículas TRMM na área de estudo.



Fonte: Elaborado pelo autor.

Figura 13 – Série de precipitações médias mensais para a região baseadas nas 91 quadrículas TRMM (1998-2019).



Fonte: Elaborado pelo autor.

3.3 DADOS DE VAZÃO

Os dados de vazão utilizados no trabalho foram obtidos do operador Nacional do Sistema Elétrico – ONS (2020), e referem-se à série histórica de vazões naturais médias afluentes à hidroelétrica de Três Marias. O período analisado foi compreendido de 01 de janeiro de 1998 a 31 de agosto de 2019. Os dados foram obtidos a nível diário para a região, e foram acumulados a nível mensal para proceder-se a modelagem e análises no presente estudo.

A ONS é o órgão brasileiro responsável pela previsão e geração de cenários de vazões naturais médias diárias, semanais e mensais para todos os locais de aproveitamentos hidroelétricos do país. De acordo com o ONS (2020), o cálculo das mesmas é feito no âmbito

do projeto de revisão das séries de vazões naturais, coordenado pelo ONS, e que contou com o apoio da Agência Nacional de Energia Elétrica – ANEEL, a Agência Nacional de Águas – ANA, do Ministério de Minas e Energia – MME, e dos agentes de geração responsáveis pelos aproveitamentos dessas bacias.

O vetor de entrada com os dados de vazões consistiu num vetor com 260 dados, representando cada um dos meses dos anos de análise. Na Figura 14 é mostrado o Hidrograma com as vazões trabalhadas, e é interessante notar a inconsistência (não homogeneidade) da série de vazões. De acordo com a aplicação do teste de Pettitt (Pettitt, 1979) percebeu-se uma quebra na homogeneidade dos dados a partir do mês 184 (abril de 2013).

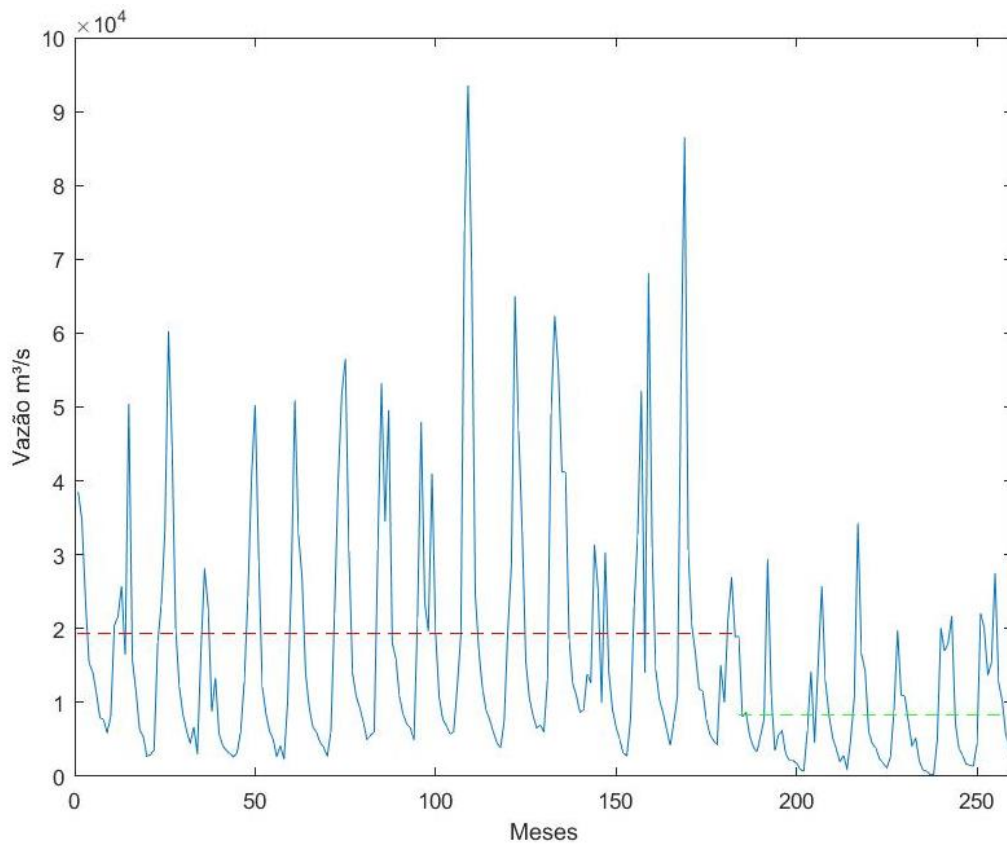
Na Figura 14, essa mudança no comportamento da vazão é mostrada pelas duas linhas médias que são mostradas, onde ficam claros os dois períodos homogêneos que os dados possuem. Essa não homogeneidade foi estudada dentro da modelagem SOM chuva-vazão, de acordo com a escolha dos períodos utilizados para as quatro modelagens. Na Tabela 2 são mostradas as variáveis estatísticas dos dados (média, máximo e mínimo, desvio padrão etc.) para os períodos dos meses 1-184, 185-260 e 1-260, analisados de acordo com a quebra da homogeneidade percebida no teste aplicado.

Tabela 2 - Índices estatísticos para as vazões observadas.

Índices estatísticos/Meses	1–184	185–260	1–260
Média (m ³ /s)	19.361,0	8.206,5	16.551,5
Mediana (m ³ /s)	12.345,5	5.307,5	10.50,5
Soma (m ³ /s)	3.562.430,0	623.700,0	3.972.364,0
Máximo (m ³ /s)	93.501,0	34.252,0	93.501,0
Mínimo (m ³ /s)	2.331,0	204,0	204,0
Desvio padrão (m ³ /s)	17.811,7	7.644,3	16.769,9
Variância (m ³ /s) ²	317.259.228,5	58.435.271,8	281.231.260,6

Fonte: Elaborado pelo autor.

Figura 14 – Série de vazões médias afluentes naturais mensais para o reservatório Três Marias (1998-2019).



Fonte: Elaborado pelo autor.

3.4 METODOLOGIA PARA A MODELAGEM SOM

3.4.1 Modelagens SOM utilizadas

Devido a característica de não homogeneidade dos dados de vazão durante o todo período de 1998-2019, foram feitas quatro modelagens SOM diferentes de acordo com as escalas de tempo utilizadas. Para cada série temporal trabalhada fez-se uso de cerca de 70% dos dados para a calibração do modelo, e aproximadamente 30% para a etapa de teste de cada uma das quatro modelagens. (a) A primeira foi feita com uma calibração para o período de 01/ 1998-04/2014 (182 meses), com teste no período de 05/2014-08/2019 (78 meses); (b) a segunda foi feita com uma calibração para o período de 07/2003-08/2019 (182 meses), com um teste para o período de 01/1998-06/2003 (78 meses); (c) a terceira com uma calibração para o período de

01/1998-08/2007 (128 meses), com um teste para o período de 09/2007-04/2014 (56 meses); (d) e a quarta com a calibração para o período de 05/2014-08/2017 (52 meses), com um teste para o período de 09/2017-08/2019 (24 meses).

Na Tabela 3 são mostrados de forma resumida as informações de cada modelagem, com a quantidade de meses de precipitação utilizados como variáveis de entrada, período de calibração e de teste. O objetivo das modelagens foi analisar a utilização do SOM para diferentes tipos de série temporal. Apenas a modelagem #4 fez uso de cinco meses de precipitação como variáveis de entrada do modelo, e isso se deu devido ao ligeiro aumento de qualidade observado dos índices de desempenho dele ao aumentar o número de meses, o que não foi percebido nas outras modelagens.

Tabela 3 - Índices desempenho para a etapa de calibração.

Modelagem SOM	Dados de entrada	Período de calibração	Período de teste
#1	$P_{t-3}, P_{t-2}, P_{t-1}$ e P_t	01/01/1998– 30/04/2014	01/05/2014– 31/08/2019
#2	$P_{t-3}, P_{t-2}, P_{t-1}$ e P_t	01/07/2003– 31/08/2019	01/01/1998– 30/06/2003
#3	$P_{t-3}, P_{t-2}, P_{t-1}$ e P_t	01/01/1998– 31/08/2008	01/09/2008– 30/04/2014
#4	$P_{t-4}, P_{t-3}, P_{t-2}, P_{t-1}$ e P_t	01/05/2014– 31/08/2017	01/09/2017– 31/08/2019

Fonte: Elaborado pelo autor.

3.4.2 Preparação dos dados

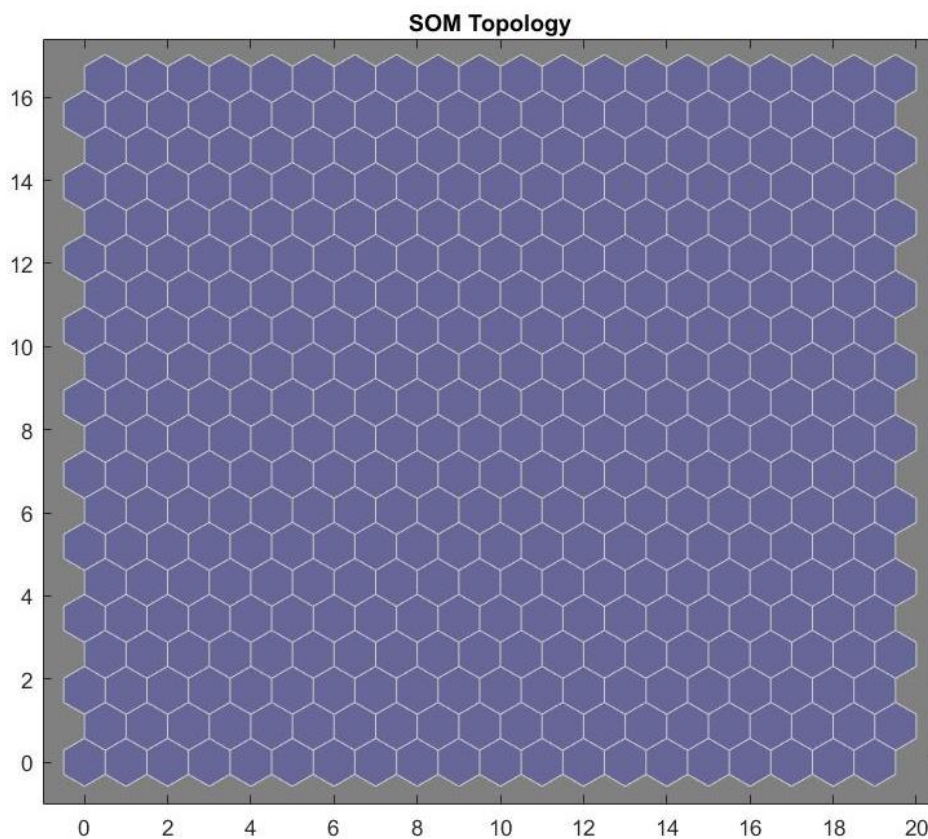
Antes de inserir no modelo SOM para cada modelagem, os dados de precipitação e vazão foram organizados numa mesma matriz, o número de linhas referente ao número de meses trabalhados. A modelagem foi feita fazendo uso da precipitação no mês t , $t-1$ e $t-2$ como variáveis de entrada para os modelos #1, #2 e #3, e t , $t-1$, $t-2$ e $t-3$ para a modelagem #4, com ambas tendo a vazão no mês t como variável de saída. Todos os dados foram manipulados por

meio do *software* MATLAB. Como a rede SOM trata-se de uma RNA de aprendizado não supervisionado, ou seja, toda a matriz deve ser fornecida sem os alvos (*targets*) serem apresentados de forma separada, a matriz de entrada para a etapa de calibração é uma matriz completa, com os dados de vazão (variável a ser calculada) também junta dos dados de precipitação.

3.4.3 Arquitetura da rede SOM

Para todas as modelagens realizadas, a arquitetura da rede SOM trabalhada foi bidimensional 21x21, conforme é mostrado na Figura 15. Notou-se que seguir a orientação dada por Garcia e Gonzalez (2004), apresentada pela Eq. (1) não foi satisfatória para as modelagens, chegando-se à relação de 21x21 através de várias tentativas.

Figura 15 – Arquitetura da rede SOM utilizada.



Fonte: Adaptado MATLAB.

3.4.4 Calibração do modelo

Para a calibração do modelo foi utilizada a função *selforgmap* do software MATLAB. Nesse sentido, inicialmente, com as matrizes de dados de entrada devidamente preparadas, e a arquitetura da rede definida, definiu-se o número de épocas para a calibração no valor de 300. Sabe-se que o modelo chega rapidamente numa estabilização, de acordo com o referencial teórico levado como base.

Finalmente, os dados de calibração são normalizados, e o modelo é calibrado dessa forma. Após a devida modelagem, identifica-se qual o neurônio vencedor da calibração e a sua vizinhança, e baseado nisso, utiliza-se o peso do neurônio vencedor conectado ao elemento ausente do vetor de entrada como o valor previsto (conforme a seção 2.5.5). realiza-se a previsão das variáveis através do modelo SOM. Após a previsão, os valores obtidos são devidamente analisados quanto ao desempenho do modelo através dos índices de desempenho trabalhados, como coeficiente de correlação, coeficiente de Nash-Sutcliffe, etc.

Como trata-se de um modelo estocástico, ou seja, que leva em conta variáveis aleatórias na sua calibração, a cada vez que a rotina computacional é executada através do software MATLAB, os valores da previsão dos dados, mesmo com entradas idênticas, tendem a ser diferentes, mesmo que dada uma determinada arquitetura da rede, e número de épocas de execução, esperar-se um comportamento mais homogêneo nos resultados. Assim, utilizou-se a modelagem tanto com os melhores resultados dos índices de desempenho para a previsão dos dados de calibração no modelo.

3.4.5 Teste do modelo

Finalmente, após a calibração do modelo ser executada e o com os melhores índices de desempenho ser escolhido como o final, realiza-se a previsão dos dados de vazão através do modelo. A previsão segue os passos descritos para o que foi realizado com os dados de calibração (conforme a seção 2.5.5), logo após os dados obtidos pela previsão do modelo são verificados quanto ao seu desempenho por meio das métricas estatísticas aplicadas.

3.5 CÁLCULO DOS ÍNDICES DE DESEMPENHO

As métricas estatísticas utilizados para a análise de desempenho do modelo SOM foram (a) o coeficiente de determinação (R^2); (b) o coeficiente de Nash-Sutcliffe; (c) o viés relativo (do inglês *Relative Bias* – RB); e (d) o erro médio padrão RMSE (do inglês *Root Mean Square Error* - RMSE).

3.5.1 Coeficiente de determinação (R^2)

O coeficiente de determinação calcula a variabilidade de um determinado número de previsões ao redor do seu valor verdadeiro, ou seja, o índice de correlação determina a qualidade de uma relação linear entre duas variáveis, ou seja, se o modelo foi capaz de explicar bem os dados observados, ou não. É um índice que varia de 0 a 1, onde 1 corresponde o ajuste perfeito. Ele é dado pela Eq. (9).

$$R^2 = \frac{(\sum_{i=1}^n (y_i - y_m) - (x_i - x_m))^2}{\sqrt{\sum_{i=1}^n (y_i - y_m)^2 (x_i - x_m)^2}} \quad \text{Eq. (9)}$$

onde x_i representa s valores de vazões observadas e x_m a média observada dos dados, y_i representa os valores de vazão calculadas pelo modelo, e y_m é a de vazão média calculada pelo modelo.

3.5.2 Coeficiente de Nash

O coeficiente de eficiência de Nash-Sutcliffe (Nash e Sutcliffe, 1970) é um índice que varia de $-\infty$ a 1, onde 1 corresponde ao perfeito ajuste entre os dados observados e estimados, e os valores menores do que 0 correspondem a situações em que a média dos dados observados é uma predição melhor do que as estimativas dadas pelo modelo empregado. Esse índice é dado pela Eq. (10).

$$Nash = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - x_m)^2} \quad \text{Eq. (10)}$$

onde x_i o de vazões observadas e x_m a média observada dos dados e y_i representa os valores de vazão calculadas pelo modelo.

3.5.3 Viés Relativo

O RB é um índice estatístico que informa se os dados previstos pelo modelo estão sendo subestimados ou superestimados em relação aos dados observados. Quando o valor é negativo, ocorre subestimação dos valores obtidos pela modelagem, e quando ocorre um valor positivo ocorre o oposto. Quanto mais próximo de 0, mais similares são as variáveis observadas e estimadas. Ele é dado pela Eq. (11).

$$RB = \frac{y_m - x_m}{x_m} \quad \text{Eq. (11)}$$

onde x_m a média observada dos dados e y_m a média dos valores de vazão calculados pelo modelo.

3.5.4 Raiz do Erro Médio Padrão

O RMSE é a raiz quadrada do *Mean Square Error* (MSE), e representa a raiz do erro médio quadrático da diferença entre o valor estimado e o valor observado. O seu valor ideal é de 0. Ele é dado pela Eq. (12).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - x_i)^2}{n}} \quad \text{Eq. (12)}$$

onde x_i são os valores observados, e y_i os valores estimados pelo modelo.

4 RESULTADOS E DISCUSSÕES

4.1 MODELAGEM #1

Para a etapa de calibração do modelo #1, foi feito uso dos dados mensais de precipitação e vazões para o período de 01/01/1998 - 01/04/2014. A série de dados do período de calibração tanto observados, quanto calculados de vazões é mostrada na Figura 16a. Os dados de vazão observadas são mostrados com linhas azuis, e os de vazão prevista são em linhas vermelhas. Para a etapa de teste do modelo, fez-se uso dos dados mensais de precipitação e vazões para o período de 01/05/2004–31/08/2019. A série de dados do período de teste tanto observados, quanto calculados de vazões é mostrada na Figura 16b, e nota-se através do gráfico que o ajuste não foi muito bom para a série de dados. E por fim, fez-se a análise dos resultados totais para todo o período trabalhado (01/01/1998–31/08/2019), conforme mostrado na Figura 16c.

Na Tabela 4 são mostrados os índices de desempenho utilizados no presente trabalho (R^2 , NASH, RB e RMS) para as etapas de calibração, teste e para o período de calibração e teste juntos. Para a etapa de calibração, os resultados são os melhores possíveis para os quatro parâmetros, mostrando um ajuste perfeito na etapa de treinamento para as vazões estimadas e observadas. Para a etapa de teste, conforme mostrado na Tabela 4, o resultado do coeficiente de correlação para o modelo não foi de todo modo baixo (0,6147). O gráfico mostrado na Figura 17b mostra a relação linear entre os resultados de vazão observada e prevista, e é possível perceber visualmente (Figura 16b) que há uma relação linear encontrada pelos dados.

Entretanto, quando se trabalha com dados observados e calculados, o ideal do modelo é que não só consiga-se um ajuste do coeficiente de correlação próximo a 1, mas quando o gráfico de valores observados e calculados se comporte próximo a uma linha $Y=T$, o que conduz a valores de vazão calculadas correspondendo exatamente ao valor observado. Ter um valor próximo a 1, implica que há uma relação linear entre os resultados plotados, o que pode levar a dados calculados que não se modelam exatamente ao que é observado, mas que possuem uma correlação que leva a linearidade, como é o caso do modelo. No presente estudo, os dados observados e calculados têm uma razoável relação linear, entretanto, um mal comportamento próximo a linha $Y=T$, o que é um indicativo de uma má modelagem.

Ainda em relação a etapa de teste, para o coeficiente de NASH, os resultados mostraram um mal ajuste entre os valores da vazão observada e a prevista (0,2221). Este ajuste pode ser

observado na Figura 17b, onde os valores calculados e observados não são bem ajustados. Entretanto, é relevante salientar que em geral, o modelo acompanhou os picos e vales da série temporal, mesmo sem fornecer uma semelhança entre os dois valores. Para o viés relativo (RB), como deveria ser esperado pelo ajuste de NASH, o valor foi alto 61,47 %, onde o sinal positivo do valor encontrado está associado com uma superestimação para os valores da modelagem em geral, com valores calculados maiores, do que os observados em boa parte da série. Finalmente, para o valor do RMSE, que foi de 6.781,93 m³/s, observa-se um valor pequeno em relação a ordem de grandeza das vazões no modelo.

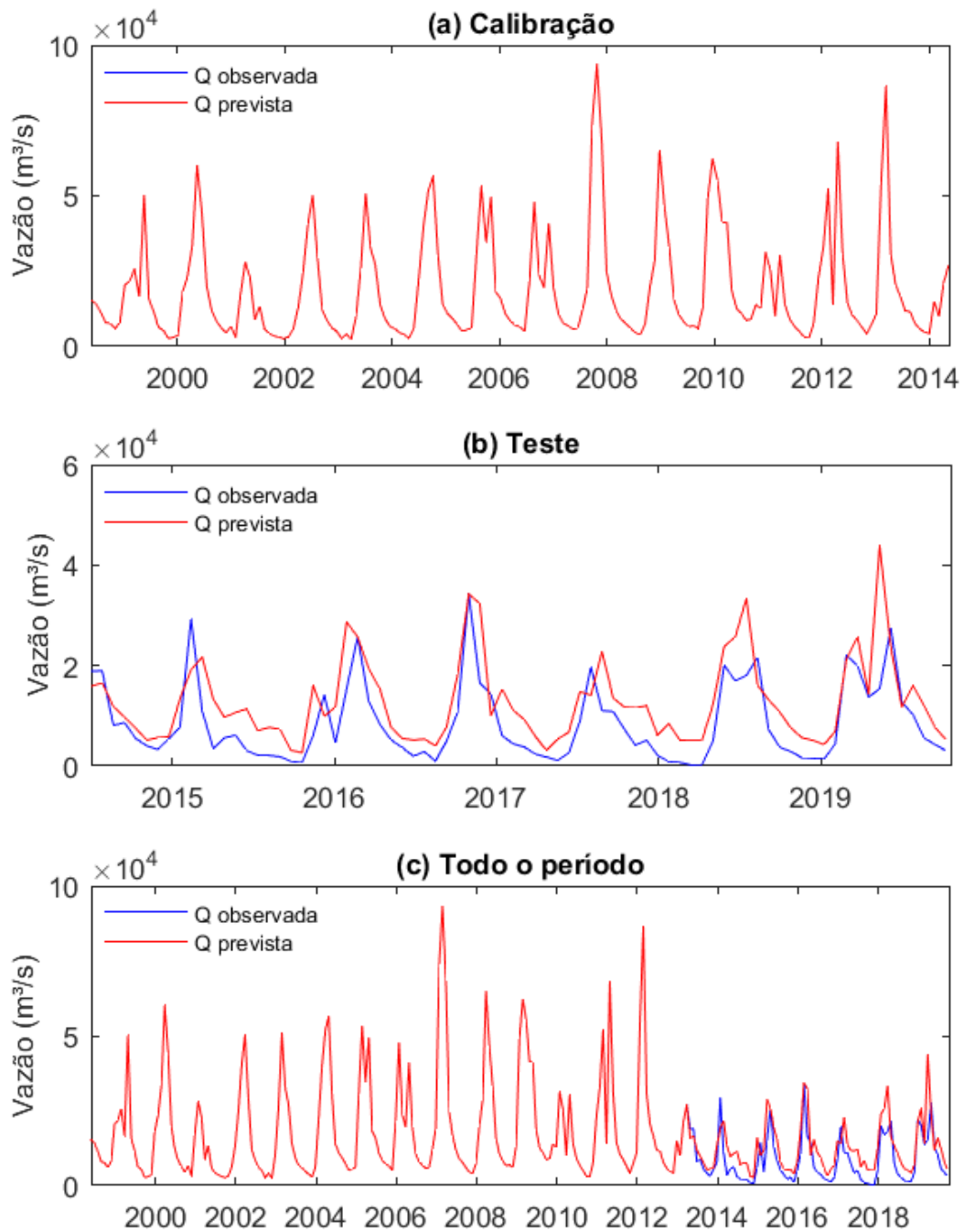
E por fim, para a análise da série completa, o resultado do R² para o modelo foi 0,9538. O gráfico mostrado na Figura 17c mostra a relação linear entre os resultados de vazão observada e prevista. Para o coeficiente de NASH, os resultados mostraram um bom ajuste entre os valores da vazão observada e a prevista, com um valor igual a 0,9544. O valor dele é alto, e deve-se levar em conta de que o mesmo ocorre devido a etapa de calibração que apresentou NASH 1. Para o viés relativo (RB), o valor foi de 8,16 %, onde o sinal positivo do valor encontrado está associado com uma superestimação para os valores da modelagem em geral, com valores calculados menores, do que os observados nos casos de haver discrepância entre ambos. E por fim, para o valor do RMSE, que foi de 3.736,20 m³/s, observa-se um valor pequeno em relação a ordem de grandeza no modelo, fazendo da variável uma melhores índices de desempenho do modelo.

Tabela 4 - Índices desempenho para as etapas de calibração, teste e para todo o período, para a modelagem #1.

Índice estatístico	Calibração	Teste	Todo o período
R ²	1	0,6147	0,9538
NASH	1	0,2221	0,9544
RB	0%	61,47%	8,16%
RMSE	0	6.781,93	3.736,20

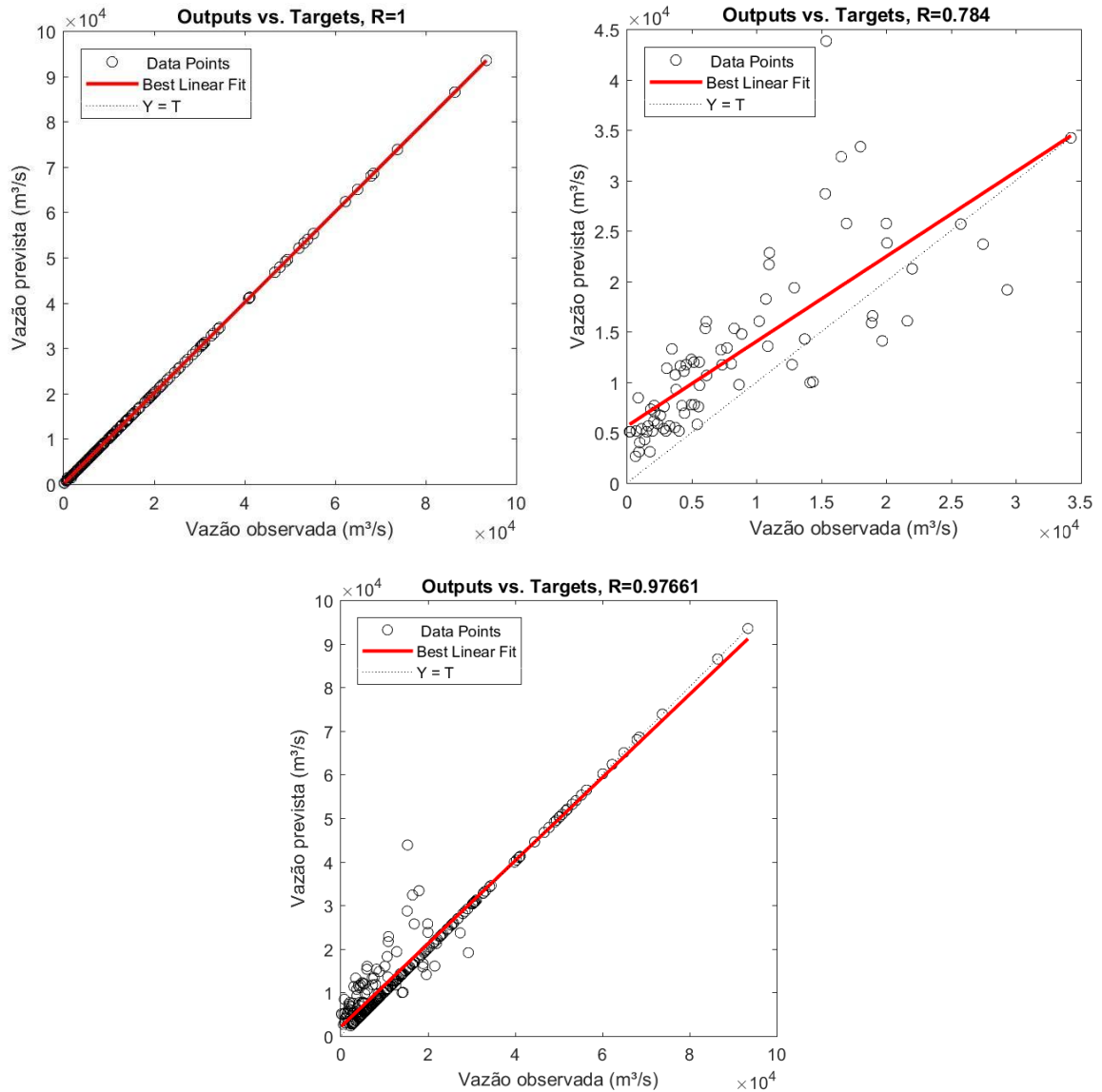
Fonte: Elaborado pelo autor.

Figura 16 – Série de vazões observadas e calculadas para os períodos de calibração, teste e para todo o período, para a modelagem #1.



Fonte: Elaborado pelo autor.

Figura 17 – Plotagem dos valores observados e calculados e regressão linear para os períodos de calibração, teste e para todo o período, para a modelagem #1.



Fonte: Elaborado pelo autor.

4.2 MODELAGEM #2

Para a etapa de calibração do modelo #2, foi feito uso dos dados mensais de precipitação e vazões para o período de 01/07/2003-31/08/2019. A série de dados do período de calibração tanto observados, quanto calculados de vazões é mostrada na Figura 18a. A etapa de teste deste modelo fez uso dos dados mensais de precipitação e vazões para o período de 01/01/1998-

30/06/2003, conforme mostrado na Figura 18b, e, conforme a figura, apresentou bom ajuste entre os dados observados e calculados na etapa de teste. E por fim, fez-se a análise dos resultados totais para todo o período trabalhado (01/01/1998-31/08/2019), conforme a Figura 18c.

Na Tabela 5, são mostrados os índices de desempenho utilizados no presente trabalho (R^2 , NASH, RB e RMSE) para as etapas de calibração, teste e para o período de calibração e teste juntos. Para a etapa de calibração do modelo os resultados dos índices de desempenho foram ótimos, mostrando o perfeito ajuste entre os dados observados e calculados na etapa de treinamento do modelo.

Para a etapa de teste do modelo, o resultado do R^2 foi alto (0,8704). O gráfico mostrado na Figura 19a mostra a relação linear entre os resultados de vazão observada e prevista, e é possível perceber visualmente o perfeito ajuste entre os resultados. O valor do coeficiente de NASH foi 0,8650, o que representa um bom ajuste entre os valores da vazão observada e a prevista. Este ótimo ajuste pode ser visualizado na Figura 18b, onde os valores calculados e observados estão bem próximos, com os picos e vales acompanhados entre ambas as séries. O RB foi de -2,48%, onde o sinal negativo do valor encontrado está associado com uma subestimação para os valores da modelagem em geral, com valores calculados menores, do que os observados nos casos de haver discrepância entre ambos. E por fim, para o valor do RMSE, que foi de 5.375,02 m³/s, observa-se um valor pequeno em relação a ordem de grandeza do modelo.

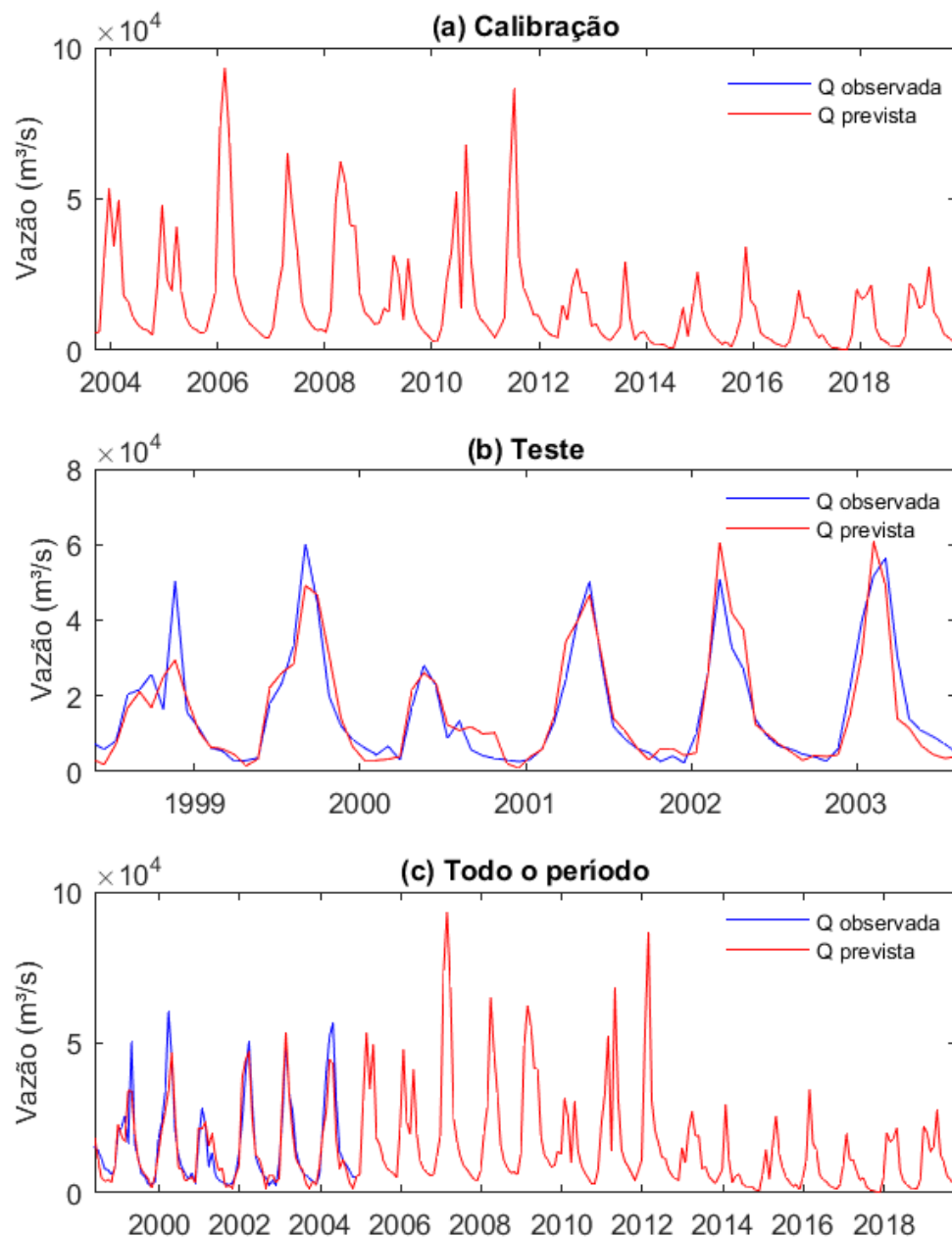
E por fim, para a análise de todo o período junto, o resultado do R^2 para o modelo foi 0,9673. O gráfico mostrado na Figura 19c mostra a relação linear entre os resultados de vazão observada e prevista. O coeficiente de NASH foi alto (0,9576). Para o RB, o valor encontrado foi de -0,76%, onde o sinal negativo do valor encontrado está associado com uma subestimação para os valores da modelagem em geral, com valores calculados menores, do que os observados nos casos de haver discrepância entre ambos. E por fim, para o valor do RMS, que foi de 3.354,08 m³/s, observa-se um valor pequeno em relação a ordem de grandeza das vazões do modelo.

Tabela 5 - Índices desempenho para as etapas de calibração, teste e para todo o período, para a modelagem #2.

Índice estatístico	Calibração	Teste	Todo o período
R ²	1	0,8704	0,9673
NASH	1	0,8650	0,9576
RB	0%	-2,48%	-0,76%
RMS	0	5.375,02	3.354,08

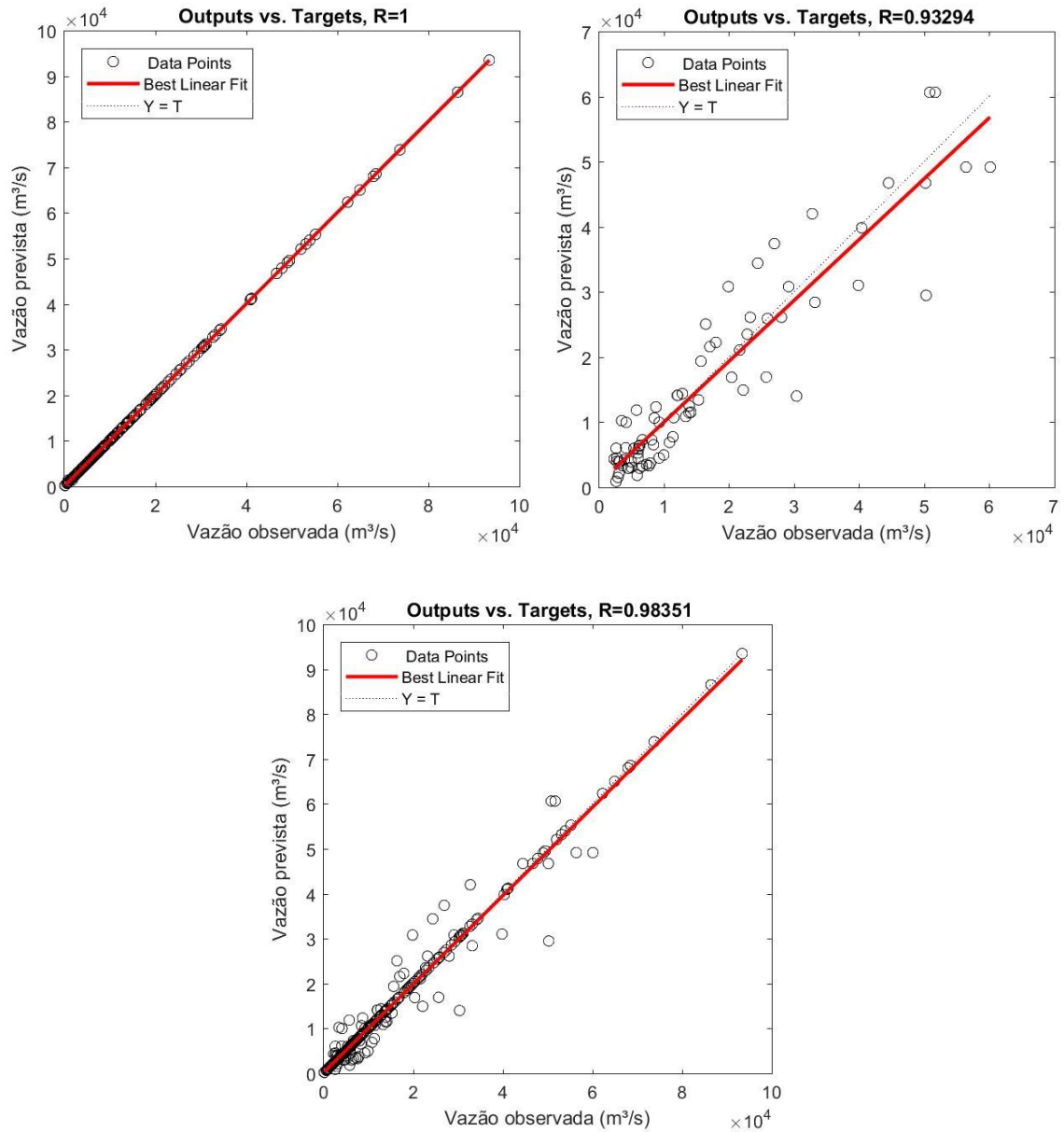
Fonte: Elaborado pelo autor.

Figura 18 – Série de vazões observadas e calculadas para os períodos de calibração, teste e para todo o período, para a modelagem #2.



Fonte: Elaborado pelo autor.

Figura 19 – Plotagem dos valores observados e calculados e regressão linear para o os períodos de calibração, teste e para todo o período, para a modelagem #2.



Fonte: Elaborado pelo autor.

4.3 MODELAGEM #3

Para a etapa de calibração do modelo 3, o período utilizado foi de 01/01/1998-31/08/2008. A série de dados do período de calibração tanto observados, quanto calculados de vazões é mostrada na Figura 20a. Para o teste foi utilizado o período de 01/09/2008-30/04/2014, e os seus valores de vazões são mostrados na Figura 20b. E por fim, foi realizada a análise dos resultados totais para todo o período trabalhado (01/01/1998-30/04/2014), com as vazões mostradas na Figura 20c.

Para a calibração do modelo, os resultados dos índices de desempenho são mostrados na Tabela 6. Os resultados são os melhores possíveis para os quatro parâmetros, mostrando um ajuste perfeito na etapa de treinamento para as vazões estimadas e observadas.

Para a etapa de teste, as métricas estatísticas também são mostrados na Tabela 6. O resultado do R^2 para o modelo foi alto (0,8431), e o gráfico mostrado na Figura 21b representa a relação linear entre os resultados de vazão observada e prevista, e é possível perceber visualmente o ótimo ajuste entre os resultados. O coeficiente de NASH mostra um ajuste relativamente bom entre os valores da vazão observada e a prevista, com um valor igual a 0,8204. O RB apresentou o valor de -9,00 %, um valor relativamente alto, e onde o sinal negativo está associado com uma subestimação para os valores da modelagem em geral, com valores calculados menores, do que os observados nos casos de haver discrepância entre ambos. Finalmente, para o valor do RMSE, que foi de 7.753,42 m³/s, o observa-se um valor pequeno em relação a ordem de grandeza adotada no modelo, o que leva a um entendimento de um bom modelo, com uma raiz de erros médios padrões baixos.

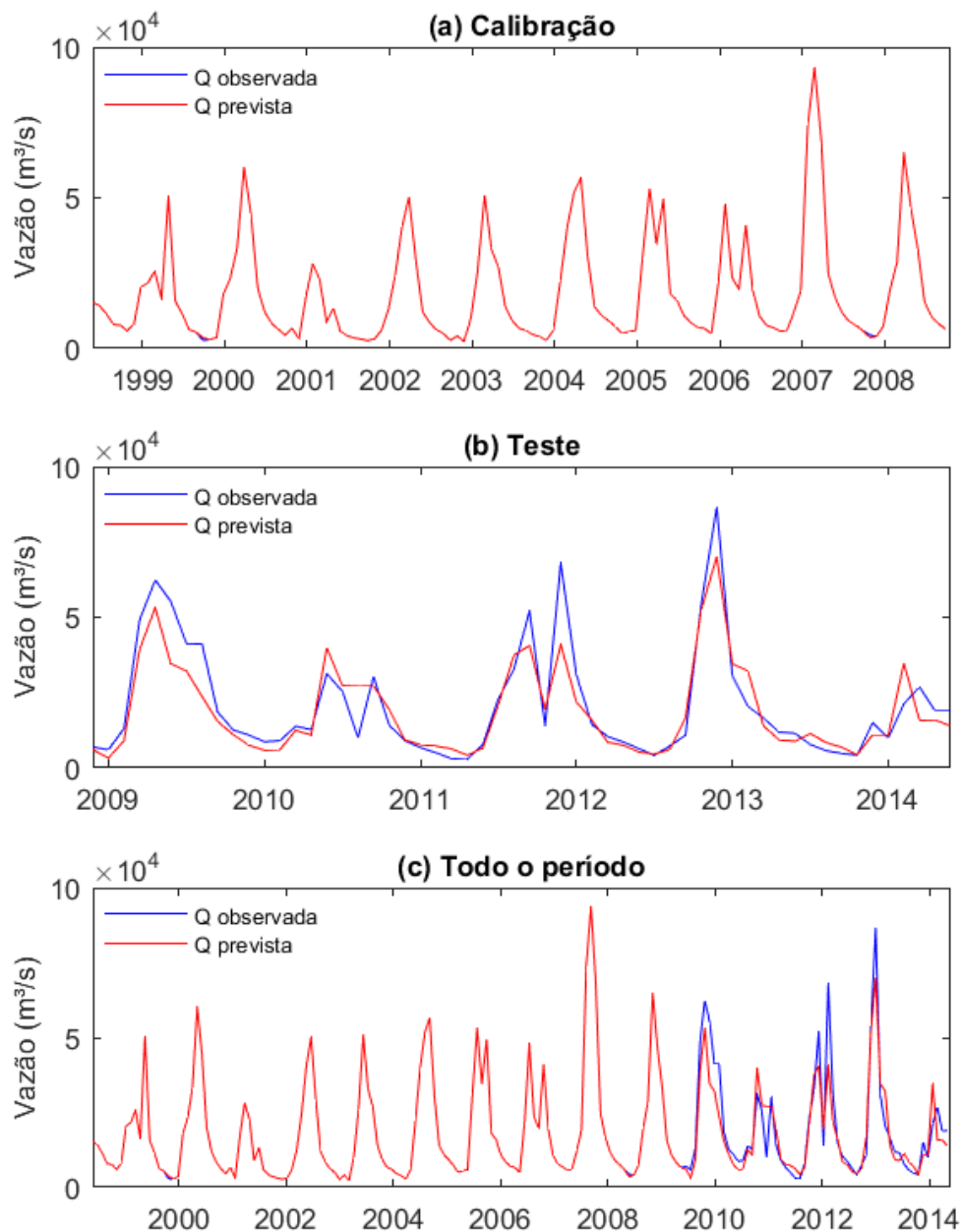
Finalmente, para a análise conjunta de todo o período de estudo, conforme mostrado na Tabela 6, o R^2 encontrado para o modelo foi alto (0,9434), e no gráfico da Figura 21c é mostrada a relação linear entre os resultados de vazão observada e prevista. Para o coeficiente de NASH, os resultados mostraram um bom ajuste entre os valores da vazão observada e a prevista, com um valor igual a 0,9414. O RB foi de -3,00%, representando uma subestimação para os valores da modelagem em geral, com valores calculados menores, do que os observados nos casos de haver discrepância entre ambos. E por fim, para o valor do RMSE, que foi de 4.313,68 m³/s, o observa-se um valor pequeno em relação a ordem de grandeza das vazões observadas no período de estudo.

Tabela 6 - Índices desempenho para as etapas de calibração, teste e para todo o período, para a modelagem #3.

Índice estatístico	Calibração	Teste	Todo o período
R ²	1	0,8431	0,9434
NASH	1	0,8204	0,9414
RB	-0%	-9,00%	-3,00%
RMSE	0	7.753,42	4.313,68

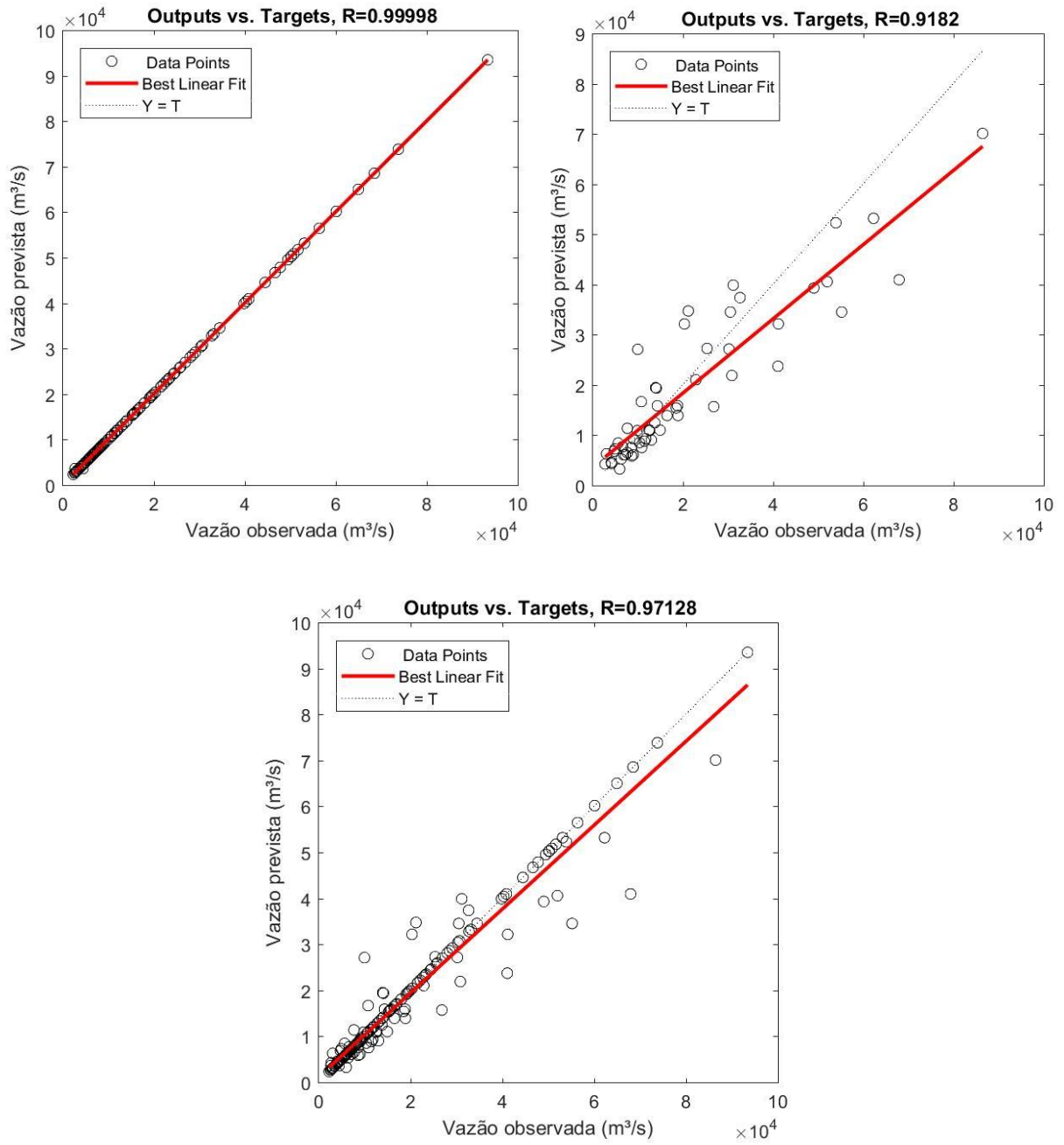
Fonte: Elaborado pelo autor.

Figura 20 – Série de vazões observadas e calculadas para os períodos de calibração, teste e para todo o período, para a modelagem #3.



Fonte: Elaborado pelo autor.

Figura 21 – Plotagem dos valores observados e calculados e regressão linear para o os períodos de calibração, teste e para todo o período, para a modelagem #3.



Fonte: Elaborado pelo autor.

4.4 MODELAGEM #4

Para a calibração da modelagem #4, o período utilizado foi o de 01/05/2014–31/08/2017. A série de dados do período de calibração tanto observados, quanto calculados de vazões é mostrada na Figura 22a. Para a etapa de teste do modelo, o período trabalhado foi entre 01/09/2017–31/08/2019, e os valores observados e estimados estão mostrados na Figura 22b. E por fim, conforme mostrado na Figura 22c, realizou-se a análise dos resultados totais para todo o período trabalhado (01/05/2014–31/08/2019).

Para a calibração do modelo, os resultados dos índices de desempenho são mostrados na Tabela 7, e são os melhores possíveis para os quatro parâmetros, mostrando um ajuste perfeito na etapa de treinamento para as vazões estimadas e observadas, conforme foi observado nas outras três modelagens SOM realizadas no trabalho.

Para a etapa de teste, os resultados dos índices de desempenho do modelo são mostrados na Tabela 7, e conforme a mesma, o valor do R^2 para o modelo foi de 0,9374, um valor alto. O gráfico mostrado na Figura 23b mostra a relação linear entre os resultados de vazão observada e prevista, e é possível perceber visualmente o ótimo ajuste entre os resultados. Para o coeficiente de NASH, os resultados mostraram um bom ajuste entre os valores da vazão observada e a prevista, com um valor igual a 0,9370. O RB apresentou um valor de $-1,37\%$, representando uma subestimação para os valores da modelagem em geral, com valores calculados menores, do que os observados. E por fim, o valor do RMSE, que foi de 2.055,66 m^3/s , observa-se um valor pequeno em relação a ordem de grandeza dos dados do modelo.

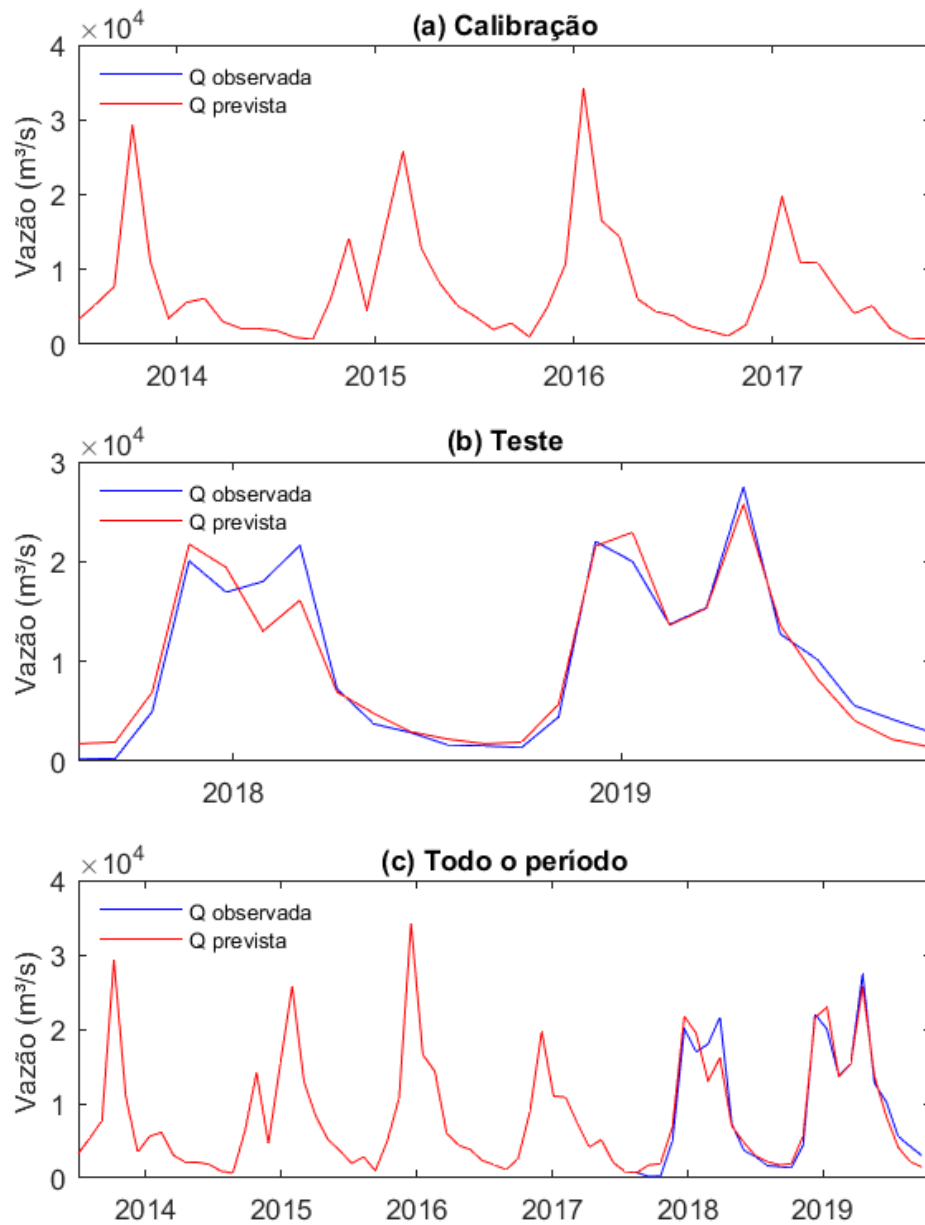
E finalmente, analisou-se a série completa, e os resultados dos índices de desempenho estão na Tabela 7. A série de dados completa, tanto de dados calculados quanto observados é mostrada na Figura 22c. O resultado do R^2 para o modelo, considerando toda a série de dados de teste e calibração juntos, foi alto (0,9768). O gráfico mostrado na Figura 23c mostra a relação linear entre os resultados de vazão observada e prevista, e é possível perceber visualmente o ótimo ajuste entre os resultados. O coeficiente de NASH foi de 0,9767. O RB, como deveria ser esperado pelo ajuste ótimo, apresentou o valor de $-0,55\%$, E por fim, para o valor do RMSE, que foi de 1.186,84 m^3/s , observa-se um valor pequeno em relação a ordem de grandeza de vazões do modelo.

Tabela 7 - Índices desempenho para as etapas de calibração, teste e para todo o período, para a modelagem #4.

Índice estatístico	Calibração	Teste	Todo o período
R ²	1	0,9374	0,9768
NASH	1	0,9370	0,9767
RB	0%	-1,37%	-0,55%
RMSE	0	2.055,66	1.186,84

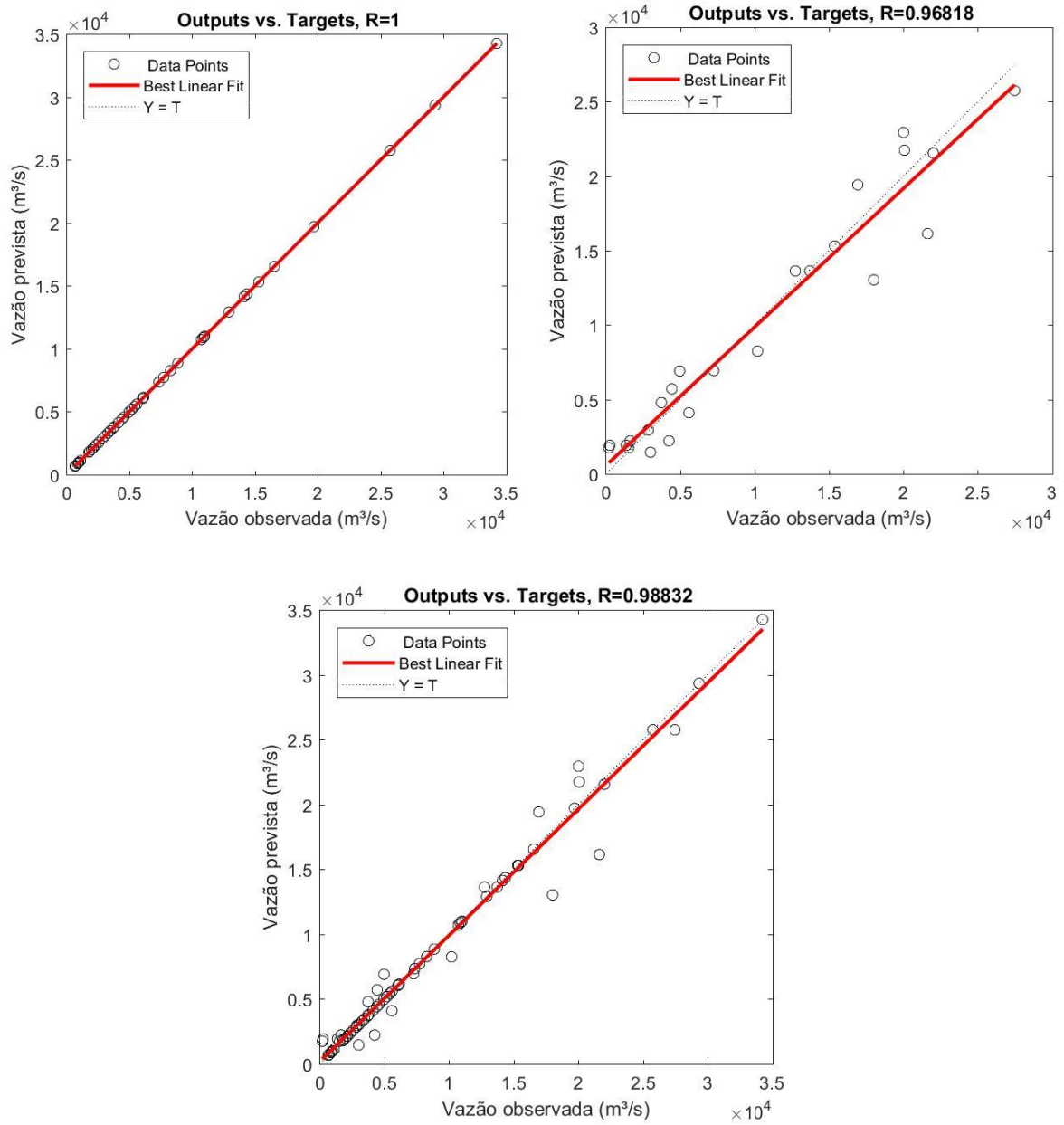
Fonte: Elaborado pelo autor.

Figura 22 – Série de vazões observadas e calculadas para os períodos de calibração, teste e para todo o período, para a modelagem #4.



Fonte: Elaborado pelo autor.

Figura 23 – Plotagem dos valores observados e calculados e regressão linear para os períodos de calibração, teste e para todo o período, para a modelagem #4.



Fonte: Elaborado pelo autor.

4.5 DISCUSSÕES DOS RESULTADOS

De acordo com as quatro modelagens feitas com a rede SOM nos dados utilizados, percebe-se uma variação significativa dos índices de desempenho de cada uma (Tabela 8). Inicialmente, pode-se interpretar que a rede SOM não foi boa de ser utilizada no caso de a série de dados a serem calculados não ser homogênea, e dentro dessa não homogeneidade, os dados utilizados para teste apresentarem o comportamento diferente dos de calibração, como é o caso da modelagem #1. Na mesma, a série geral apresenta uma mudança no comportamento temporal das vazões, e como a calibração dos dados foi realizada majoritariamente com a série apresentando o primeiro tipo de comportamento (Figura 16), quando a etapa de teste fez uso da série temporal com o comportamento temporal diferente, a modelagem não foi boa, conforme foi visto nos índices de desempenho.

Ao analisar a modelagem #2, continuou-se trabalhando com a série não homogênea, entretanto, inverteu-se a posição dos dados de calibração, colocando-os no final da série temporal, e assim tendo como dados de treinamento representantes dos dois tipos de comportamento hidrológico da vazão afluente no reservatório. Para essa modelagem, os dados calculados na etapa de teste tiveram um bom desempenho, conforme são vistos nos seus índices calculados (Tabela 8). Portanto, percebe-se que o problema na rede SOM ao trabalhar com dados não homogêneos se concentrou na variedade de comportamentos hidrológicos dados para os dados na etapa de calibração. Como a modelagem #2 apresenta dois comportamentos diferentes de vazão na etapa de calibração, o resultado do modelo ao ser aplicado na etapa de teste foi melhor.

Portanto, em suma, ainda em relação a problemática da homogeneidade, é possível de que se a série trabalhada fosse com um período maior, a variação de comportamentos hidrológicos fosse naturalmente maior, e assim, se observaria uma melhor modelagem naturalmente para os dados de vazão da modelagem #1.

Para a modelagem #3, que fez uso de uma série de dados homogênea, foram percebidos bons valores nos índices de desempenho, mostrando que ocorreu uma boa modelagem. E finalmente, a modelagem que apresentou os melhores resultados para os índices de desempenho da etapa de teste foi a modelagem #4. É interessante notar que ela fez uso de uma série temporal relativamente curta (apenas seis anos), porém de comportamento homogêneo, analisando a série como um todo (calibração e teste). Assim, destaca-se a não obrigatoriedade de uma série longa

ao trabalhar com dados homogêneos, destacando-se a possibilidade de modelagem em situações de não abundância de dados observados. Além disso, é interessante destacar que para esta modelagem, foram utilizados dados de 4 meses de precipitação (P_{t-3} , P_{t-2} , P_{t-1} e P_t), uma vez que foi com eles encontrou-se melhores valores para os índices de desempenho para a modelagem da etapa de teste. Esse melhor ajuste com mais meses de precipitação é provável de ser notado devido ao fato de trabalhar-se com uma série temporal pequena.

Tabela 8 – Tabela resumo com os índices de desempenho para os dados na etapa de teste das quatro modelagens utilizadas no trabalho.

Índice estatístico	Modelagem			
	#1	#2	#3	#4
R ²	0,6147	0,8704	0,8431	0,9374
NASH	0,2221	0,8650	0,8204	0,9370
RB	61,47%	-2,48%	-9,00%	-1,37%
RMSE	6.781,93	5.375,02	7.753,42	2.055,66

Fonte: Elaborado pelo autor.

De modo geral, destaca-se que os resultados obtidos para as modelagens através de uma rede SOM foram bons em relação a literatura para os índices de desempenho, conforme é possível ser observado ao comparar-se trabalhos como o de Farias et al. (2012) e Farias et al. (2013), que se concentraram em modelos chuva-vazão com a rede parecidos, e os valores de referência dados por Moriasi et al. (2007).

Para o caso de Farias et al. (2012), a melhor calibração encontrada para o modelo chuva-vazão apresentou dados para os índices de desempenho utilizados (R², NASH e RB) de respectivamente 0,96, 0,89 e -27 % para a etapa de teste. Esses valores são de índice de desempenho são relativamente próximos dos resultados da modelagem #3 do presente trabalho, , que depois da modelagem #1, foi a que apresentou os piores índices de desempenho.

Além disso, é interessante notar que o modelo do presente trabalho contou apenas com dados de precipitação de dois e um meses anteriores, e o do próprio mês para a modelagem, enquanto o trabalho de Farias et al. (2012) fez uso também da variável de vazão do mês anterior para a previsão. A utilização apenas de chuva permite que o modelo consiga preencher séries com falhas contínuas nos valores de vazão, o que o modelo do outro trabalho não poderia, uma vez que precisaria da variável vazão anterior. Finalmente, deve-se destacar também que ele fez

uso dos valores de precipitação oriundos de postos pluviométricos, e não de satélite, como foi o caso do presente estudo.

No trabalho de Farias et al. (2013), a análise foi feita mais uma vez com dados de postos pluviométricos para a região de estudo, e mesmo não utilizando a variável de vazão do mês anterior em sua melhor modelagem, igual ao presente trabalho, ele contou com valores de R^2 , NASH e RB respectivamente, 0,85, 0,81 e -20%, sendo os parâmetros relativamente próximos ao ser comparado com a modelagem #3 do presente trabalho.

Continuando a discussão dos resultados do presente trabalho para os índices de desempenho, o trabalho de Moriasi et al. (2007) fez um estudo onde apresentou escalas que iam desde o insatisfatório, até o muito bom para a qualidade dos índices. De acordo com ele, para o NASH, valores acima de 0,75 são considerados muito bons, para a escala de tempo mensal e dados de vazão. Assim, ao comparar os valores observados para o presente trabalho, observa-se a classificação de muito boa para as modelagens do presente estudo. Ele também fez também uma classificação para o RB, onde apresentou os valores menores do que $\pm 10\%$ como muito bons, o que mais uma vez é aplicado para o presente trabalho. Portanto, os valores mostrados por Moriasi et al. (2007) tanto para NASH quanto para o RB, classificam classificando os modelos #2, #3 e #4 propostos pela rede SOM para o atual trabalho como sendo muito bons.

5 CONSIDERAÇÕES FINAIS

Assim, diante do estudo realizado é demonstrada a possibilidade de utilização de redes SOM na modelagem chuva-vazão, utilizando dados de precipitação provenientes de satélite TRMM a nível mensal de forma eficiente e prática. Das quatro modelagens realizadas no presente trabalho, três mostraram índices de desempenho muito bons nas etapas de calibração e teste, e apenas o modelo proposto #1 ter resultados insatisfatórios devido a questão da não homogeneidade dos dados aplicados.

A mudança de comportamento na série e com uma conseqüente não homogeneidade nos dados no período completo disponível, aliado com a escolha de um período não totalmente representativo para a série na etapa de treinamento fez com que o modelo #1 não fosse eficiente. Caso se trabalhasse com uma série de dados maior, é provável que se possuísse uma variedade maior no comportamento dos dados, o que transpassaria naturalmente essa limitação encontrada. Nesse sentido, ressalta-se a importância de que na etapa de treinamento das redes SOM, sejam utilizados dados o mais representativos possível para todo o período de trabalho, conforme é o caso da modelagem #2, que conseguiu transpassar as limitações do modelo #1, ao inverter-se o período de treinamento para o fim da série temporal, o que permitiu ter a representação dos dois comportamentos hidrológicos na etapa de calibração do modelo, e assim, ter um modelo mais eficiente na previsão dos dados.

Com a modelagem #3, percebeu-se um rápido ajuste da rede SOM, uma vez que a série de dados utilizada foi de todo homogênea, e para a #4, além de reforçar o percebido na #3, destaca-se a eficiência da rede SOM na modelagem de uma série com uma quantidade tão limitada de dados (cerca de 48 meses de dados para calibração, e de 24 para o teste), o que é uma vantagem significativa em relação a modelos tradicionais de chuva-vazão, que em tese, necessitam de longas séries temporais para atingir uma eficiência significativa nas modelagens.

Portanto, a modelagem chuva-vazão através de redes SOM, além de eficientes, mostraram-se totalmente práticas e aplicáveis, sendo bastante úteis no gerenciamento de recursos hídricos e no planejamento de um recurso tão importante para o desenvolvimento humano nos tempos atuais. A modelagem mostrou-se como uma ferramenta de preenchimento de dados de vazões em bacias não tão bem monitoradas, possibilitando estudos e projetos mais bem elaborados na área. Além disso, a utilização de dados de satélite como uma alternativa prática e eficiente para as variáveis de precipitação, dada as limitações que dados provenientes

de formas tradicionais, como de pluviômetros e pluviógrafos mostram em muitos casos, como o caso de bacias não aparelhadas e localizadas em áreas remotas, por exemplo.

6 REFERÊNCIAS BIBLIOGRÁFICAS

ADELOYE A. J.; RUSTUM, R.; KARIYAMA, D. (2011). “Kohonen self-organizing map estimator for the reference crop evapotranspiration”. *Water Resources Research* 47, W08523, pp. 1-19.

ADELOYE, A. J. & RUSTUM, R. (2011). SOM and rainfall-runoff modelling in inadequately gauged basins. *Hydrology Research*, 43.5, 2012.

AGÊNCIA NACIONAL DE ÁGUAS (ANA). Princípios de hidrologia ambiental. Guarulhos, São Paulo, 2017. Disponível em: < <https://capacitacao.ana.gov.br/conhecerh/handle/ana/240>> Acesso em 31 de dezembro de 2019.

ARNOLD, J.G. et al. (2012). SWAT: model use, calibration, and validation. *Transactions of the ASABE*, St. Joseph, v.55, n.4, p.1491-1508.

BACK, B., K. SERE, e HANNA, V. (1998), Managing complexity in large database using self organising map, *Accounting Manag. Inf. Technol.*, 8, 191–210.

BAYATVARKESHI, M.; MOHAMMADI, K. ; KISI, O. ; FASIHI, R. (2018). A new wavelet conjunction approach for estimation of relative humidity: wavelet principal component analysis combined with ANN. *NEURAL COMPUTING & APPLICATIONS*, v. 31, p. 1.

BUARQUE, D. C.; DE PAIVA, R. C. D.; CLARKE, R. T.; MENDES, C. A. B. (2011). A comparison of Amazon rainfall characteristics derived from TRMM, CMORPH and the Brazilian national rain gauge network. *Journal of Geophysical Research: Atmospheres*, 116(19), 1–12. doi: 10.1029/2011JD016060

CHANG, F. J.; CHANG, L. C.; KAO, H. S.; WU, G. W. (2010). Assessing the effort of meteorological variables for evaporation estimation by self-organizing map neural network. *Journal of Hydrology* 384, 118–129.

CHEN, Y; EBERT, E. E.; WALSH, K. J. E.; DAVIDSON, N. E. (2013). Evaluation of TRMM 3B42 precipitation estimates of tropical cyclone rainfall using PACRAIN data. *Journal of Geophysical Research Atmospheres*, 118(XX), 2184–2196. doi: 10.1002/jgrd.50250

COLLISCHONN, B.; COLLISCHONN, W.; TUCCI, C. E. M. (2008). Daily hydrological modeling in the Amazon basin using TRMM rainfall estimates. *Journal of Hydrology*, 360(1), 207–216. doi: 10.1016/j.jhydrol.2008.07.032.

COLLISCHONN, W.; DORNELLES, F. (2015). Hidrologia para engenharias e ciências ambientais. Associação Brasileira de Recursos Hídricos – ABRH. 2ª Edição. Porto Alegre, 350p.

COULIBALY, P.; ANCTIL, F.; BOBÉE, B. (2001). Multivariate reservoir inflow forecasting using temporal neural networks, *J. Hydrol. Engng. ASCE*, 6, 367–376

CURTARELLI, M. P.; RENNÓ, C. D.; ALCÂNTARA, E. H. (2014). Evaluation of the tropical rainfall measuring mission 3B43 product over an inland area in Brazil and the effects of satellite boost on rainfall estimates. *Journal of Applied Remote Sensing*, 8(1), 1–14. doi: 10.1117/1.JRS.8.083589

DELAHAYE, F.; KIRSTETTER, P. E.; DUBREUIL, V.; MACHADO, L. A. T.; VILA, D. A.; CLARK, R. (2015). A consistent gauge database for daily rainfall analysis over the legal Brazilian Amazon. *Journal of Hydrology*, 527(XX), 292–304. doi: 10.1016/j.jhydrol.2015.04.012

FARIAS, C. A. S.; BEZERRA, U. A. ; SILVA FILHO, J. A. (2015). Runoff-erosion modeling at micro-watershed scale: a comparison of self-organizing maps structures. *Geoenvironmental Disasters*, v. 2, p. 1-8.

FARIAS, C. A. S.; CARNEIRO, T. C. e LOURENÇO, A. M. G. (2012). Mapas auto-organizáveis para modelagem chuva-vazão. *Anais do XI Simpósio de Recursos Hídricos do Nordeste*, p. 1-14.

FARIAS, C. A. S.; SANTOS, C. A. G.; LOURENÇO, A. M. G. e CARNEIRO, T. C. (2013). Kohonen Neural Networks for rainfall-runoff modeling: Case study of Piancó river basin. *Journal of Urban and Environmental Engineering (UFPB)*, v. 7, p. 176-182.

FARIAS, C.A.S. ; ALVES, F. M.; SANTOS, C. A. G. ; SUZUKI, K. (2010). An ANN-based approach to modelling sediment yield: a case study in a semi-arid area of Brazil. *IAHS-AISH Publication*, v. 337, p. 316-321.

FARIAS, C. A. S.; KADOTA, A.; CELESTE, A. B.; SUZUKI, K. (2007) RNN based inflow forecasting applied to reservoir operation via implicit stochastic optimization, *IAHS Publ.* 313, 452–462.

FARIAS, C. A. S.; SANTOS, C. A. G. (2014). The use of Kohonen neural networks for runoff-erosion modeling. *Journal of Soils and Sediments (Print)* v. 14, p. 1242-1250.

FREIRE, P. K. M. M.; O uso da transformada Wavelet na previsão de vazão. 2012. 130 f. Dissertação (Mestrado em Engenharia Urbana e Ambiental) – Programa de Pós-Graduação em Engenharia Urbana e Ambiental, Universidade Federal da Paraíba, João Pessoa, 2012.

GADELHA, A. N.; HUGO, V.; COELHO, R.; XAVIER, A. C.; ROMERO, L.; MELO, D. C. D.; XUAN, Y.; HU, G. J.; PETERSEN, W. A.; ALMEIDA, N. (2019). Grid box-level evaluation of IMERG over Brazil at various space and time scales 218, 231–244. <https://doi.org/10.1016/j.atmosres.2018.12.001>

GARCÍA, H. L.; GONZÁLEZ, I. M. (2004). “Self-organizing map and clustering for wastewater treatment monitoring”. *Engineering Applications of Artificial Intelligence*, v. 17, n. 3, pp. 215–225.

GORGIJ, A. D.; KISI, O.; MOGHADDAM, A. A.(2017). Groundwater budget forecasting, using hybrid wavelet-ANN-GP modelling: a case study of Azarshahr Plain, East Azerbaijan, Iran. *Nordic Hydrology (Cessou em 2007. Cont. ISSN 1998-9563 Hydrology Research (Print))*, v. 48, p. 455-467.

HAYKIN, S. (1999). *Neural Networks: a Comprehensive Foundation*. Prentice Hall, Upper Saddle River, USA, 842 p.

HONORATO, A. G. S. M.; SILVA, G. B. L.; GUIMARÃES SANTOS, C. A. (2019). Monthly streamflow forecasting using neuro-wavelet techniques and input analysis. *Hydrological Sciences Journal*, v. 63, p. 1-16.

HORNBERGER, G. M. (1998). *Elements of Physical Hydrology*. The John Hopkins University Press, 302 pp.

HUR, J.; RANGHAVAN, S. V.; NGUYEN, N. S.; LIONG, S. Y. (2018). Are satellite products good proxies for gauge precipitation over Singapore? *Theoretical and Applied Climatology*, 132(3-4), 921–932. doi: 10.1007/s00704-017-2132-7

JEONG, D.; KIM, Y. O. (2005) Rainfall-runoff models using artificial neural networks for ensemble streamflow prediction, *Hydrol. Process.*, 19, 3819–3835.

KANGAS, J.; SIMULA, O. (1995). Process monitoring and visualization using self organizing map. In: *Neural Networks for Chemical Engineers*, Ch. 14 (A. B. Bulsari, ed.). Elsevier Science, Dordrecht, The Netherlands.

- KOHONEN, T. (1982). “Self-organized formation of topologically correct feature maps”. *Biological Cybernetics* 43, pp. 59-69
- KOHONEN, T.; SIMULA, O.; VISA, A. (1996), Engineering applications of the self-organizing map, *Proc. IEEE*, 84(10), 1358–1384.
- LANNA, A. E.; SCHWARZBACH, M. (1989). MODHAC - Modelo Hidrológico Auto-Calibrável. Recursos Hídricos, Publicação 21. Pós-Graduação em Recursos Hídricos e Saneamento, Universidade Federal do Rio Grande do Sul.
- LIAO, R. W.; ZHANG D. B.; SHEN, Y. (2015). Validation of six satellite-derived rainfall estimates over China. *Meteorological Monthly*, 41(XX), 970–979.
- LOPES, J. E. J; BRAGA JR. B. P. F.; CONEJO, J. G. L. (1981). “Simulação hidrológica: Aplicações de um modelo simplificado”. In: *Anais do III Simpósio Brasileiro de Recursos Hídricos*, v.2, p.42-62. Fortaleza-CE.
- MIRABBASI, R.; KISI, O.; SANIKHANI, H.; GAJBHIYE MESHRAM, S. (2018). Monthly long-term rainfall estimation in Central India using M5Tree, MARS, LSSVR, ANN and GEP models. *Neural Computing & Applications*, v. 31, p. 1.
- MMA, 2017. Programa de revitalização da bacia hidrográfica do Rio São Francisco. Disponível em: www.mma.gov.br. Acesso em 31 de dezembro de 2019.
- MORIASI, D. N.; ARNOLD, J. G.; VAN LIEW, M. W.; BINGER, R. L.; HARMEL, R. D.; VEITH, T. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the American Society of Agricultural & Biological Engineers*, St. Joseph, v. 50, n. 3, p. 885-900.
- NASA. National Aeronautics and Space Administration. Disponível em: <https://pmm.nasa.gov/TRMM>. Acesso em: 30 de dezembro de 2019.
- NASH, J. E.; SUTCLIFFE, J. V. (1970). River flow forecasting through conceptual models I: a discussion of principles. *Journal of Hydrology*, v. 10, n.1, p. 282-290.
- NING, S. W.; SONG, F.; UDMALE, P.; JIN, J. L.; THAPA, B. R.; ISHIDAIRA, H. (2017). Error analysis and evaluation of the latest GSGMap and IMERG precipitation products over Eastern China. *Advances in Meteorology*, XX(XX), XX–XX. doi: 10.1155/2017/1803492

- NOURANI, V.; BAGHANAM, A. H.; ADAMOWSKI, J.; GEBREMICHAEL, M. (2013). Using self-organizing maps and wavelet transforms for space–time pre-processing of satellite precipitation and runoff data in neural network based rainfall–runoff modeling. *Journal of Hydrology*, v. 476, p. 228-243.
- NOUREDDINE, K.; MOHAMMED, A.; SANTOS, C. A. G.; ABDELKADER, D.; ABDELHAMID, B.; NASCIMENTO, T. V. M. (2019). Spatial modeling of soil salinity using multiple linear regression, ordinary kriging and artificial neural network methods in the Lower Cheliff plain, Algeria. *Journal of Urban And Environmental Engineering (UFPB)*, v. 13, p. 34-41.
- NOVARINI, B.; BRENTAN, B. M.; MEIRELLES, G. e JUNIOR, E. L. (2019). Optimal pressure management in water distribution networks through district metered area creation based on machine learning. *Brazilian Journal of Water Resources*, v.24, e37, <https://doi.org/10.1590/2318-0331.241920180165>
- ONS, Operador Nacional do Sistema Elétrico, 2019. Disponível em:<www.ons.org.br>.
- PARASURAMAN, K. A.; ELSHORBAGY, A.; CAREY, S. K. (2006). Spiking modular neural networks: a neural network modelling approach for hydrological processes. *Water Resource Research* 42, W05412.
- SILVA, L. P.; XAVIER, A. P. C.; SILVA, R. M.; SANTOS, C. A. G. (2019). Modeling land cover change based on an artificial neural network for a semiarid river basin in northeastern Brazil. *Global Ecology and Conservation*, v. 20, p. e00811.
- PETTITT, A. N. A Non-Parametric Approach to the Change-Point Problem. *Applied Statistics*, v. 28, n. 2, p. 126-135, 1979.
- PRADO, T. B. G. (2005). Evolução do uso das terras e produção de sedimentos na bacia hidrográfica do rio Jundiaí-Mirim. Dissertação (Mestrado em Agricultura Tropical e Subtropical) – Instituto Agrônômico de Campinas. 72 pp.
- RIVIERA, D.; LILLO, M.; UVO, C. B.; BILLIB, M.; ARUMI, J. L. (2011). Forecasting monthly precipitation in central Chile: a self-organising map approach using filtered sea surface temperature. *Theoretical and Applied Climatology* 107 (1–2), 1–13.

- SANTOS, C. A. G.; MORAIS, B. S. (2013). Identification of precipitation zones within São Francisco River basin (Brazil) by global wavelet power spectra. *Hydrological Sciences Journal*, 58(4), 789–796. <https://doi.org/10.1080/02626667.2013.778412>
- SANTOS, C. A. G.; SILVA, G. B. L. (2013). Daily streamflow forecasting using a wavelet transform and artificial neural network hybrid models. *Hydrological Sciences Journal*, 59(2), 312–324. <https://doi.org/10.1080/02626667.2013.800944>
- SANTOS, C. A. G.; FREIRE, P. K. M. M.; SILVA, G. B. L.; SILVA, R. M. (2014). Discrete wavelet transform coupled with ANN for daily discharge forecasting into Três Marias reservoir. *Proceedings of the International Association of Hydrological Sciences*, v. 364, p. 100-105.
- SANTOS, C. A. G.; FREIRE, P. K. M. M.; MISHRA, S. K. (2012a) Cuckoo search via Lévy flights for optimization of a physically-based runoff-erosion model. *J. Urban Environ. Engng.* 6(2), 123-131.
- SANTOS, C. A. G.; FREIRE, P. K. M. M.; ARRUDA, P. M. (2012b) Application of a simulated annealing optimization to a physically-based erosion model. *Water Sci. and Technol.* 66(10), 2099–2108. doi: 10.2166/wst.2012.426
- SANTOS, C. A. G.; FREIRE, P. K. M. M.; SILVA, R. M.; AKRAMI, S. A. (2019). Hybrid Wavelet Neural Network Approach for Daily Inflow Forecasting Using Tropical Rainfall Measuring Mission Data. *Journal of Hydrologic Engineering*, v. 24, p. 04018062, 2019.
- FILHO, J. S.; FARIAS, C. (2018). Stochastic modeling of monthly river flows by Self-Organizing Maps. *Journal of Urban and Environmental Engineering (UFPB)*, v. 12, p. 219-230.
- SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. (2010). *Redes Neurais Artificiais para engenharia e ciências aplicadas*, Artliber.
- SILVA, R. M., DANTAS, J. C., BELTRÃO, J. A.; SANTOS, C. A. G. (2018) Hydrological simulation in a tropical humid basin in the Cerrado biome using the SWAT model. *Hydrology Research*, 49(3). <https://doi.org/10.2166/nh.2018.222>.
- SUGAWARA, M. (1961). Automatic calibration of the Tank-Model, *Hydrological science, bulletin*, v. 24, n. 3. p.375, 1961
- SULTANA, R.; NASROLLAHI, N. (2018). Evaluation of remote sensing precipitation estimates over Saudi Arabia. *Journal of Arid Environments*, 151(XX), 90–103. doi: 10.1016/j.jaridenv.2017.11.002

- TANANAKI, C.; THRASYVOULOU, A.; GIRAUDEL, J. L.; MONTURY, M. (2007). Determination of volatile characteristics of Greek and Turkish pine honey samples and their classification by using Kohonen self organizing maps. *Food Chemistry* 101, 1687–1693.
- TANG, G. Q.; LI, Z.; XUE, X. W.; HU, Q.F.; YONG, B.; HONG, Y. (2015). A Study of substitutability of TRMM remote sensing precipitation for gauge-based observation in Ganjiang River Basin. *Advances in Water Science*, 26(3), 340–346. doi: 10.14042/j.cnki.32.1309.2015.03.005
- TERINK, W.; LEIJNSE, H.; EERTWEGH, G.; UIJLENHOET, R. (2018). Spatial resolutions in areal rainfall estimation and their impact on hydrological simulations of a lowland catchment. *J. Hydrol.* 563, 319–335. <https://doi.org/10.1016/j.jhydrol.2018.05.045>
- TUCCI, C. E. M. (2009). *Hidrologia: ciência e aplicação*. Editora da Universidade Federal do Rio Grande do Sul – ABRH. 4ª Edição. Porto Alegre, 943 p
- VOUTILAINEN, A.; ARVOLA, L. M. J. (2017). SOM clustering of 21-year data of a small pristine boreal lake, *Knowledge and Management of Aquatic Ecosystems* , vol. 2017 , no. 418 . <https://doi.org/10.1051/kmae/2017027>
- WMO. (1994). *Guide to Hydrological Practices: Data Acquisition and Processing, Analysis, Forecasting and Other Applications*, WMO 168. Geneva: World Meteorological Organization.
- WU, C. L.; CHAU, K. W. (2011). Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis. *J. Hydrol.* 399(3-4), 394–409.
- ZENG, Q. L.; WANG, Y. Q.; CHEN, L. F.; WANG, Z. F.; ZHU, H.; LI, B. (2018). Inter-comparison and evaluation of remote sensing precipitation products over China from 2005 to 2013. *Remote Sensing*, 10(2), 168–XX. doi: 10.3390/rs10020168