



Universidade Federal Da Paraíba

Centro De Tecnologia

**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL E
AMBIENTAL – MESTRADO**

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

Por

Thiago Victor Medeiros do Nascimento

*Dissertação de Mestrado apresentada à Universidade Federal da Paraíba para a obtenção
do grau de Mestre*

João Pessoa – Paraíba

Dezembro de 2021



Universidade Federal Da Paraíba

Centro De Tecnologia

**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL E
AMBIENTAL – MESTRADO**

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

Dissertação submetida ao Programa de Pós-graduação em Engenharia Civil e Ambiental da Universidade Federal da Paraíba, como parte dos requisitos para a obtenção do título de Mestre.

Thiago Victor Medeiros do Nascimento

Orientador: Prof. Dr. Celso Augusto Guimarães Santos

João Pessoa – Paraíba

Dezembro de 2021

Catálogo na publicação
Seção de Catalogação e Classificação

N244m Nascimento, Thiago Victor Medeiros do.

Modelagem hidrológica mensal chuva-razão utilizando dados de precipitação TRMM e mapas auto-organizáveis para a bacia do Alto Rio São Francisco / Thiago Victor Medeiros do Nascimento. - João Pessoa, 2021.

87 f. : il.

Orientação: Celso Augusto Guimarães Santos.
Dissertação (Mestrado) - UFPB/CT.

1. Meio ambiente hidrosférico. 2. Modelos chuva-vazão. 3. Redes Neurais Artificiais - RNA. 4. Self-Organizing Maps - SOM. 5. Tropical Rainfall Measurement Mission - TRMM. I. Santos, Celso Augusto Guimarães. II. Título.

UFPB/BC

CDU 502.51(043)



***“MODELAGEM HIDROLÓGICA MENSAL CHUVA–VAZÃO UTILIZANDO MAPAS
AUTO-ORGANIZÁVEIS E DADOS DE PRECIPITAÇÃO TRMM PARA A BACIA DO ALTO
RIO SÃO FRANCISCO”***

THIAGO VICTOR MEDEIROS DO NASCIMENTO
Dissertação aprovada em 14 de dezembro de 2021

Período Letivo: 2021.2

Prof. Dr. Celso Augusto Guimarães Santos – UFPB
Orientador

Prof. Dr. Richarde Marques da Silva – UFPB
Examinador Interno

Prof. Dr. Camilo Allyson Simões de Farias – UFCG
Examinador Externo

João Pessoa/PB
2021

Dedico este trabalho a Deus, meus pais e a
minha avó. Gratidão.

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

Provérbios 16:18

AGRADECIMENTOS

Meu sincero sentimento de gratidão a todos que de forma direta ou indireta estiveram comigo durante toda essa trajetória. Como costume dizer, 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 graça derramadas em minha vida.

Agradeço a minha mãe Rosilene Medeiros, e a minha avó Severina Medeiros que estiveram e estão sempre ao meu lado, e são as melhores pessoas que eu poderia ter na minha vida.

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.

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

A toda a minha família.

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

Aos professores Cristiano e Francisco Sarmento, por ter sido um dos impulsionadores da minha paixão pelos recursos hídricos.

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

Aos amigos que o Programa de Pós Graduação me proporcionou, em particular Tatiane, Filipe e Judivan.

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 nessas linhas.

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 (JUBA VALE).

Aos meus amigos e colegas de trabalho do grupo de orientandos do professor Celso Santos e Richarde Marques e do LEPPAN, em especial ao parceiro de trabalho Reginaldo Neto.

Enfim, a todos que de maneira direta e indireta contribuíram para que eu alcançasse mais esse degrau na escalada de um sonho. Gratidão.

RESUMO

O bom gerenciamento de recursos hídricos é uma atividade essencial para a realização de uma eficiente gestão e controle na sociedade, e assim, estudos e pesquisas na área são essenciais. Assim, é requerida a existência de séries de dados hidrológicos longas e confiáveis, e.g. precipitação, e vazão. Entretanto, é comum a presença de séries históricas com dados faltosos, como as de chuva e vazão em bacias hidrográficas. Portanto, dentro desse aspecto de monitoramento de variáveis hidrológicas, a modelagem chuva-vazão é 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. E nesse aspecto, como uma alternativa para a modelagem chuva-vazão, existe a utilização de Redes Neurais Artificiais (RNA). O presente trabalho utiliza uma RNA chamada de mapas auto-organizáveis (do inglês *Self-Organizing Maps* – SOM), na modelagem chuva-vazão, utilizando produtos de precipitação do satélite *Tropical Rainfall Measurement Mission* (TRMM), e vazões naturais afluentes de um reservatório da bacia do Alto São Francisco. São trabalhados com três conjuntos de dados, sendo realizadas três modelagens distintas, com períodos para treinamento e teste diferentes, dentro do intervalo de 1998-2019. E por fim, são calculadas métricas estatísticas para os resultados, como objetivo de analisar o desempenho dos modelos hidrológicos, bem como também analisar a influência da homogeneidade dos dados na calibração e teste de uma rede SOM. Os resultados mostraram que a homogeneidade dos dados influenciou na modelagem chuva-vazão. Para a etapa de calibração, os modelos apresentaram altos índices de eficiência, e para a de teste, os resultados variaram a depender do *dataset* e do tipo da modelagem. Os piores valores das métricas encontradas foram para a modelagem que fez uso de 70% dos dados iniciais da série na calibração, e 30% dos finais no teste para o *dataset* #1, que é o não homogêneo. As demais oito modelagens apresentaram métricas relativamente de boas a ótimas, i.e., Nash-Sutcliffe (NSE) e Coeficiente de determinação ($CD > 0,70$), Viés relativo (VR) e a Raiz do erro quadrático médio normalizado (NRMSE) próximos de zero. Portanto, fica clara a possibilidade de utilização de dados de precipitação TRMM e redes SOM na modelagem chuva-vazão a nível mensal.

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

ABSTRACT

The good management of water resources is an essential activity for the achievement of an efficient management of it. Therefore, studies and research in the area are essential, and for which the existence of long and long series of hydrological data is required, e.g., precipitation, and streamflow. However, it is common to have historical series with missing data, such as rainfall and flow in hydrographic basins. Within this aspect of monitoring hydrological variables, rainfall-streamflow modeling is an alternative for filling, and even forecasting, affluent flow data in rivers and reservoirs where there is limited data available. And in this respect, as an alternative for rainfall-streamflow modeling, there is also the use of Artificial Neural Networks (ANN). The present work uses an RNA called self-organizing maps (SOM), in rainfall-streamflow modeling, using precipitation products from the Tropical Rainfall Measurement Mission (TRMM) satellite, affluent natural flows from a reservoir the Alto São Francisco basin. They are worked with three sets of data, being carried through three different models, with periods for training and testing of the different results, within the total period of available data 1998-2019. Finally, statistical metrics are calculated for the results, in order to analyze the performance of the models, as well as to analyze the influence of data homogeneity in the calibration and testing of a SOM network. The results showed that the homogeneity of the data influenced the rainfall-flow modeling. For the calibration step, the models showed relatively high efficiency indexes, and for the test, the results varied depending on the dataset and the type of modeling. The worst values of the metrics found were for the modeling that made use of 70% of the initial data of the series in the calibration, and 30% of the final in the test for dataset # 1, which is the non-homogeneous. The remaining eight models presented relatively good to excellent metrics, i.e., Nash-Sutcliffe (NSE) and Determination coefficient (DC) > 0.70, and (Relative bias) RB and Normalize root-mean-square error (NRMSE) close to zero. Therefore, the possibility of using TRMM precipitation data and SOM networks in rainfall-flow modeling at a monthly level is clear.

Keywords: Rainfall-streamflow models, RNA, SOM, TRMM.

LISTA DE FIGURAS

Figura 1 - Esquema simplificado de uma bacia hidrográfica (Fonte: Brasil Escola)	22
Figura 2 – Esquematização do ciclo da água. (Fonte: <i>United States Geological Survey</i> - USGS).	23
Figura 3 – Concepção artística do satélite TRMM (Fonte: NASA, 2020).	27
Figura 4 - Esquema da concepção de uma rede neural artificial (Fonte: Freire, 2012).	32
Figura 5 - Estrutura de um modelo SOM com detalhe do neurônio vencedor, e seus neurônios vizinhos (Fonte: Adeloyle et al., 2011).	35
Figura 6 - As distâncias entre os neurônios de um modelo SOM para a determinação da vizinhança (Fonte: Farias et al., 2013).	36
Figura 7 - Localização da sub-bacia do Alto São Francisco, com destaque para a sua rede de drenagem e para o reservatório Três Marias.	42
Figura 8 - Topografia da região.	43
Figura 9 - Uso e ocupação do solo.	44
Figura 10 – Interpolação Spline da precipitação anual média na região (1998-2019) baseado em dados TRMM.	45
Figura 11 - Imagem de satélite do reservatório Três Marias (Fonte: Adaptado, Google Earth Pro, 2020).	46
Figura 12 - Espacialização das quadrículas TRMM na área de estudo.	48
Figura 13 - Série de precipitações médias mensais para a região baseadas nas 91 quadrículas TRMM (1998-2019).	49
Figura 14 - Série de vazões médias afluentes naturais mensais para o reservatório Três Marias (1998–2019).	51
Figura 15 - Divisão dos três datasets utilizados na modelagem das vazões do presente trabalho.	52
Figura 16 – Topologia das redes SOM utilizadas para os três conjuntos de dados trabalhados no presente estudo (a) Topologia 9x9, (b) Topologia 8x8 e (c) Topologia 6x6 (Fonte: Adaptado MATLAB).	54
Figura 17 – Mapa de <i>Hits</i> das redes SOM para as etapas de calibração e teste para as três modelagens do <i>dataset #1</i>	58
Figura 18 – Mapa dos componentes do plano da quadrícula 39 para as etapas de calibração dos três modelos desenvolvidos para o <i>dataset #1</i>	59
Figura 19 – Hidrograma de vazões observadas e calculadas para a etapa de calibração das três modelagens realizadas com o <i>dataset #1</i>	60
Figura 20 – Hidrograma de vazões observadas e calculadas para a etapa de teste das três modelagens realizadas com o <i>dataset #1</i>	62
Figura 21 – Mapa de <i>Hits</i> das redes SOM para as etapas de calibração e teste para as três modelagens do <i>dataset #2</i>	64
Figura 22 – Mapa dos componentes do plano da quadrícula 39 para as etapas de calibração dos três modelos desenvolvidos para o <i>dataset #2</i>	65
Figura 23 – Hidrograma de vazões observadas e calculadas para a etapa de calibração das três modelagens realizadas com o <i>dataset #2</i>	66
Figura 24 – Hidrograma de vazões observadas e calculadas para a etapa de teste das três modelagens realizadas com o <i>dataset #1</i>	68
Figura 25 – Mapa de <i>Hits</i> das redes SOM para as etapas de calibração e teste para as três modelagens do <i>dataset #3</i>	69

Figura 26 – Mapa dos componentes do plano da quadrícula 39 para as etapas de calibração dos três modelos desenvolvidos para o <i>dataset</i> #3.....	70
Figura 27 – Hidrograma de vazões observadas e calculadas para a etapa de calibração das três modelagens realizadas com o <i>dataset</i> #3.....	71
Figura 28 – Hidrograma de vazões observadas e calculadas para a etapa de teste das três modelagens realizadas com o <i>dataset</i> #3.....	73

LISTA DE TABELAS

Tabela 1 - Estatísticas descritivas para os dados de vazão e precipitação TRMM (médias ponderadas de Thiessen) mensais no reservatório de Três Marias (1998–2019).	50
Tabela 2 - Informações sobre os conjuntos de dados usados neste trabalho.	53
Tabela 3 – Métricas estatísticas da etapa de calibração e teste para o dataset #1.	59
Tabela 4 – Métricas estatísticas da etapa de calibração e teste para o dataset #2.	65
Tabela 5 – Métricas estatísticas da etapa de calibração e teste para o dataset #3.	71
Tabela 6 - Performance das métricas CD, NSE e VR (Fonte: Adaptado, Moriasi et al., 2015).	74

LISTA DE ABREVIATURAS E SIGLAS

ANA – Agência Nacional de Águas
ANEEL – Agência Natural de Energia Elétrica
BMU – *Best matching unit*
CD – *Coeficiente de Determinação*
CERES – *Cloud and Earth Radiant Energy Sensor*
CHIRPS – *Climate Hazards Group InfraRed Precipitation with Station data*
CMORPH – *Climate Prediction Center Morphing*
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*
MMA – Ministério do Meio Ambiente
MME – Ministério de Minas e Energia
MODHAC – Modelo hidrológico auto calibrável
NASA – Administração Nacional da Aeronáutica e Espaço
NRMSE – *Normalized Root Mean Square Error*
NSE – *Nash-Sutcliffe*
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 – *Root Mean Square Error*
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*

VR – *Viés Relativo*

WMO – *World Meteorological Organization*

WT – *Wavelet*

SUMÁRIO

RESUMO

ABSTRACT

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

1	INTRODUÇÃO.....	17
2	OBJETIVOS.....	21
2.1	OBJETIVO GERAL	21
2.2	OBJETIVO ESPECÍFICOS.....	21
3	REFERENCIAL TEÓRICO	22
3.1	BACIA HIDROGRÁFICA E CICLO HIDROLÓGICO.....	22
3.1.1	Precipitação	24
3.1.2	Produtos de precipitação TRMM.....	26
3.1.3	Vazão.....	28
3.2	MODELAGEM HDROLÓGICA	30
3.3	REDES NEURAIAS ARTIFICIAIS	31
3.3.1	Definição e estrutura	31
3.3.2	Aplicações de RNA na modelagem e previsão de séries temporais.....	33
3.4	MAPAS AUTO-ORGANIZÁVEIS	34
3.4.1	Definição e estrutura das redes SOM	34
3.4.2	Treinamento	35
3.4.3	Cálculo e previsão dos valores das variáveis por meio de redes SOM.....	38
3.4.4	Aplicações do SOM na modelagem e na previsão de séries temporais.....	38
4	CARACTERIZAÇÃO DA ÁREA DE ESTUDO.....	41
5	MATERIAIS E MÉTODOS	47
5.1	DADOS UTILIZADOS.....	47
5.1.1	Dados de precipitação.....	47
5.1.2	Dados de vazão	49
5.2	MODELAGEM SOM.....	51
5.2.1	Modelos e conjuntos de dados utilizados	51
5.2.2	Arquitetura da rede SOM.....	53
5.2.3	Fase de treinamento.....	54
5.2.4	Fase de teste.....	55
5.3	MÉTRICAS ESTATÍSTICAS.....	55
6	RESULTADOS E DISCUSSÕES.....	57
6.1	<i>DATASET #1</i>	57
6.1.1	Componentes do plano e modelagem SOM.....	57
6.1.2	Métricas estatísticas.....	59
6.2	<i>DATASET #2</i>	63
6.2.1	Componentes do plano e modelagem SOM.....	63
6.2.2	Métricas estatísticas.....	65
6.3	<i>DATASET #3</i>	68

6.3.1	Componentes do plano e modelagem SOM.....	68
6.3.2	Métricas estatísticas.....	70
6.4	DISCUSSÕES DOS RESULTADOS	73
7	CONCLUSÕES E RECOMENDAÇÕES.....	76
8	REFERÊNCIAS BIBLIOGRÁFICAS.....	78

1 INTRODUÇÃO

A água é um recurso natural que possui uma importância significativa no desenvolvimento e ciclo das diversas formas de vida da terra. Ela é a substância mais presente na superfície terrestre, e cobre mais de 70% de sua superfície. A água é um elemento essencial na sociedade, apresentando como tipos de usos principais que vão desde 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 (Collischonn & Dornelles, 2015).

É importante destacar que a água se move na atmosfera através de um processo contínuo chamado ciclo hidrológico, e esse movimento é o que permite a variação da água nas suas líquida, gasosa e sólida (Vilella & Matos, 1975). 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; posteriormente 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.

Nesse sentido, a gestão efetiva dos recursos hídricos é fundamental para o desenvolvimento social e econômico de qualquer região. A gestão adequada desses recursos é uma das formas que possibilita a mitigação dos efeitos de diversos fenômenos naturais, como as secas que afetam diretamente o setor de energia (Santos *et al.*, 2019). Assim, compreender as variáveis ambientais que afetam o processo de escoamento e vazões em bacias hidrográficas, em especial, a precipitação, é essencial para otimizar o gerenciamento de recursos hídricos. Esse gerenciamento se otimizado engloba aspectos que vão desde a avaliação da necessidade de irrigação, previsão de enchentes em rios, avaliação do atendimento de demandas de abastecimento público, avaliação de operações hidrelétricas, e organização e elaboração de projetos de engenharia hidráulica.

A precipitação é considerada a principal fonte de entrada de água de uma bacia hidrográfica (Muhammad *et al.*, 2018). Em linhas gerais, pode-se afirmar que há 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 brasileiros. 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.

A modelagem ambiental é uma ferramenta que pode ser utilizada em diversas áreas dentro do âmbito ambiental, e especificamente dos recursos hídricos, e como tal, mostra-se como uma ferramenta para a previsão e preenchimento de dados de vazão em rios e reservatórios. 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 (Tucci, 2009). Ao trabalhar-se com a variável de saída como sendo a vazão, pode-se utilizar dados de vazão anteriores, precipitação, evapotranspiração, e diversas outras variáveis ambientais como sendo as de entrada.

Nesse sentido, a modelagem chuva-vazão é um processo de alta complexidade, não estacionariedade e não linearidade nas escalas espacial e temporal, devido à complexidade hidrológica do sistema, uma vez que estão sujeitos a muitas incertezas inerentes. Dentre essas incertezas destaca-se a precipitação, características, cobertura e uso do solo, interceptação, evapotranspiração e uso da água. Nesse sentido, modelar essa relação é uma atividade importante e complexa. Um modelo de chuva-vazão é baseado em dados de chuva como variáveis de entrada, para simular as vazões em uma seção específica de um rio.

Existem variados modelos hidrológicos que utilizam diversas variáveis ambientais como dados de entrada, e.g., precipitação, evapotranspiração, uso e ocupação do solo, precipitação, temperatura do ar, humidade relativa. Exemplos de modelos de chuva-vazão comumente usados são: Soil Moisture Accounting Program-SMAP (Lopes *et al.*, 1981), the Tank Model (Sugawara, 1961), o self-calibrating hydrological model-MODHAC (Lanna and Schwarzbach, 1989), e o Soil and Water Assessment Tool - SWAT (Arnold *et al.*, 2012). Nesse sentido, o desenvolvimento de modelos simplificados que utilizem menos variáveis ambientais como dados de entrada, e consigam mesmo assim representar bem o comportamento da variável de saída (vazão) mostram-se vantajosos, dada a dificuldade de obtenção de muitas das variáveis hidrológicas supracitadas.

Ainda referente a modelagem hidrológica de maneira geral, mais recentemente diversos modelos baseados em redes neurais artificiais (RNA) estão sendo utilizados. RNA são 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 mesmos.

O processo de aprendizado consiste no estabelecimento de novas conexões entre os neurônios, e em modificações nas conexões existentes (Freire, 2012). RNA já foram utilizados para modelar chuva-vazão (Hsu *et al.*, 1995; Jain *et al.*, 2004; Filho *et al.*, 2018 e Santos *et al.*, 2019), para previsão do fluxo do rio (Campolo *et al.*, 1999; Kisi, 2004; Kumar *et al.*, 2004; Cigizoglu & Kisi, 2005; e Honorato *et al.*, 2018; e Santos *et al.*, 2014), para a previsão de precipitação (Hall *et al.*, 1999; Silverman & Dracup, 2000; Freiwan & Cigizoglu, 2005; Mirabbasi *et al.*, 2018); modelagem de água subterrânea (Coulibaly *et al.* 2001; Nayak *et al.* 2006; Mohanty *et al.* 2010), modelagem de qualidade da água (Singh *et al.* 2009, Gazzaz *et al.* 2012), estimativa de evapotranspiração (Kumar *et al.* 2002, Trajkovic *et al.* 2003, Zanetti *et al.* 2007, Adeloye *et al.* 2011) e transporte de sedimentos (Nagy *et al.*, 2002; Cigizoğlu e Alp, 2006; Melesse *et al.*, 2011; e Farias e Santos, 2014).

Nesse sentido, apresenta-se os Mapas Auto-organizáveis (do inglês *Self-Organizing Maps* – SOM), também conhecidos como redes auto-organizáveis de Kohonen (*Kohonen self-organizing map* - KSOM). Eles são um tipo de Redes Neurais Artificiais (RNA) proposto por Kohonen (1982). As redes SOM são redes neurais não supervisionadas que agrupam os dados de entrada em classes de acordo com suas semelhanças por meio de métodos de aprendizagem competitivos (Kohonen, 1982; Haykin, 1999). Essa rede neural reduz o conjunto de dados multidimensionais a uma matriz de recursos unidimensional ou bidimensional, que pode ser usada para análise de agrupamentos (cluster) e fins de previsão. A rede SOM é composta por neurônios (as pequenas unidades da estrutura), que são organizados de maneira a preservar as relações vizinhas. Para isso, a rede agrupa os dados de entrada padrão de forma que padrões semelhantes sejam representados pelos mesmos neurônios de saída, ou por um de seus vizinhos.

Vários estudos relataram o sucesso da aplicação do SOM em problemas ambientais e de recursos hídricos. García e González (2004) aplicaram a rede para entender o comportamento das variáveis envolvidas nos processos de operação de uma estação de tratamento de águas residuais. Novarini *et al.* (2019) fez uso da capacidade de agrupamento da rede SOM nas fases de modelagem de gerenciamento de pressão ideal em redes de abastecimento de água. Voutilainen e Arvola (2017) a aplicaram para agrupar dados ambientais complexos para um pequeno lago boreal intocado. Mannan *et al.* (2017) utilizaram a RNA para regionalizar as características da precipitação na Índia, incorporando variáveis climáticas. Adeloye *et al.* (2011) fizeram uso do 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). Farias e Santos (2014) e Farias *et al.* (2015) utilizaram

a rede SOM na modelagem de erosão por fluxo. Nourani *et al.* (2013) e Ismail *et al.* (2012) usaram SOM em modelos híbridos com outras RNAs, e onde SOM para identificar aglomerados espacialmente homogêneos de precipitação para serem usados nos modelos chuva-escoamento. E, por fim, Adeloye e Rustum (2012), Farias *et al.* (2012), Farias *et al.* (2013), e Filho e Farias (2018), aplicaram o SOM na modelagem chuva-vazão.

Do ponto de vista das variáveis hidrológicas comumente utilizadas em modelos hidrológicos, é comum a falta de séries temporais de chuvas contínuas de longo prazo em países em desenvolvimento e em áreas extensas e até mesmo inacessíveis. No Brasil, por exemplo, o monitoramento por pluviômetros pode ser classificado como uma tarefa desafiadora devido à extensão territorial do país, às diversidades climáticas e físicas existentes, e à existência de uma rede pluviométrica com diversas falhas em sua composição (Pereira *et al.*, 2013; Curtarelli *et al.*, 2014; Rao *et al.*, 2015).

Além disso, segundo Gadelha *et al.* (2019), a rede brasileira de monitoramento de chuva possui aproximadamente 11.820 estações pluviométricas, resultando em uma densidade de cerca de um pluviômetro por 720 km², inferior ao recomendado pela WMO (1994), que é de um a cada 575 km². Desse modo, o uso de satélites de sensoriamento remoto surge como alternativa viável para monitorar o regime das chuvas sobre as regiões mais remotas e complexas do mundo (Neto, 2020).

Assim, para lidar com esse fato, o presente trabalho utilizou dados de precipitação obtidos da *Tropical Rainfall Measuring Mission* (TRMM), que é uma missão conjunta de duas importantes agências espaciais, a saber, a National Aeronautics and Space Administration (NASA) e a Japan Agência de Exploração Aeroespacial (JAXA). Ao cobrir as regiões tropicais e subtropicais da Terra, o TRMM fornece informações muito necessárias e importantes sobre a precipitação e tem sido útil para vários estudos (por exemplo, Plouffe *et al.* 2015; Teng *et al.* 2016; Santos *et al.*, 2019).

Portanto, o presente trabalho propõe a aplicação da rede neural artificial SOM na modelagem chuva-vazão utilizando apenas dados de precipitação TRMM como variável de entrada. Foi trabalhado em escala mensal, na sub-bacia do Alto Rio São Francisco, considerada uma área estratégica para o gerenciamento dos recursos hídricos do Brasil (Santos *et al.*, 2019).

2 OBJETIVOS

2.1 OBJETIVO GERAL

O objetivo geral do presente trabalho é aplicar a rede neural artificial do tipo SOM no desenvolvimento de modelos chuva-vazão em diferentes tipos de intervalos temporais de dados dentro do período de 22 anos (1998-2019), na área da sub-bacia do Alto São Francisco, para o cálculo de vazões naturais mensais utilizando dados de precipitação TRMM.

2.2 OBJETIVO ESPECÍFICOS

- a) Calibração e teste de um modelo chuva-vazão treinado numa rede SOM para os períodos compreendidos entre 1998-2019, 1998-2014 e 2014-2019, de modo a testar a influência da homogeneidade e tamanho da série temporal na qualidade da modelagem chuva-vazão;
- b) Divisão dos três conjuntos de dados em três modelagens cada (a primeira utilizando 70% dos dados iniciais como calibração, a segunda dividindo os dados de calibração entre os 35% iniciais e os 35% finais, e a terceira com os dados de calibração sendo os 70% finais;
- c) Verificação do desempenho de cada uma das modelagens hidrológicas chuva-vazão desenvolvidas com base em métricas estatísticas.

3 REFERENCIAL TEÓRICO

3.1 BACIA HIDROGRÁFICA E CICLO HIDROLÓGICO

Uma bacia hidrográfica corresponde a uma região definida topograficamente, delimitada por elementos do relevo denominados divisores de água, que é drenada por um curso d'água ou sistema de cursos, por onde todo o escoamento do efluente é descarregado por um único ponto, denominado exutório (Vilella e Matos, 1975).

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, destacando-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. Na Figura 1 é mostrado um desenho esquemático de uma bacia hidrográfica, com destaque para o seu relevo e os corpos d'água seguindo para o exutório.

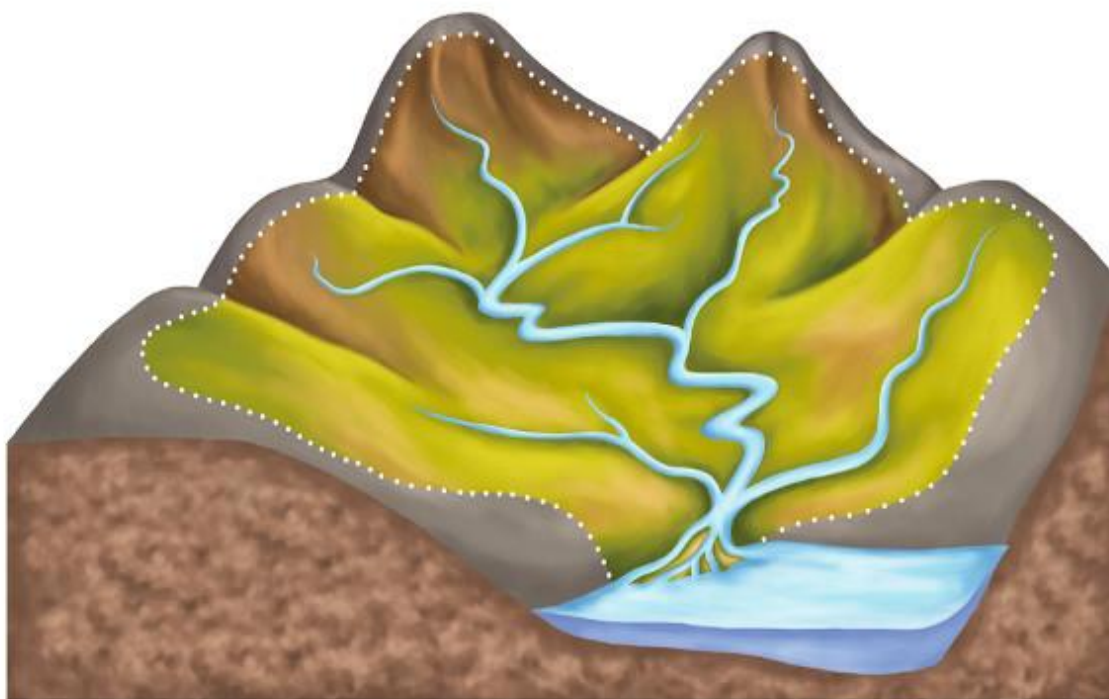


Figura 1 - Esquema simplificado de uma bacia hidrográfica (Fonte: Brasil Escola)

A água se move na atmosfera por meio de um processo contínuo denominado ciclo hidrológico. O ciclo hidrológico consiste num 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 (Tucci, 2009), e em todo esse processo, a água pode se apresentar em cada um dos seus três estados físicos, i.e. o sólido, o líquido e o gasoso.

Na Figura 2 é mostrado uma esquematização do ciclo hidrológico, sendo destacados os seus processos, começando com a evaporação e evapotranspiração, condensação, armazenamento da água na atmosfera e/ou no gelo, precipitação, detenção e escoamento superficial, infiltração, percolação da água no solo e nos aquíferos, escoamentos fluviais e finalmente o armazenamento da água nos oceanos.



Figura 2 – Esquematização do ciclo da água. (Fonte: *United States Geological Survey* - USGS).

Conforme é mostrado na Figura 2, ao considerar como ponto de partida os oceanos, a água é primeiramente evaporada. Em seguida, a água é condensada na forma de nuvens, que consequentemente, resultam na precipitação da água na forma de neve, granito e de chuva.

Assim, a água precipitada retorna para a terra, onde parte gera escoamento superficial, alimentando rios e lagos, e outra infiltra no solo, abastecendo aquíferos, que como destino final acabam seguindo para corpos d'água superficiais, que consequentemente seguem para os oceanos, e assim, reinicia-se o processo cíclico com a evaporação. É interessante destacar que parte da água acaba por retornar diretamente para a atmosfera, por meio do processo de evapotranspiração, i.e., 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 a vazão de rios, como os objetos de estudo do presente trabalho, sendo ambos inteiramente ligados, uma vez que, a precipitação é uma variável de entrada importante no ciclo hidrológico de uma bacia hidrográfica.

3.1.1 Precipitação

A precipitação pode ser definida como o retorno da água à superfície da terra nas formas de chuva, granizo, neve, orvalho neblina ou geada (Collischonn & Dornelles, 2015). 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. Como definição, destaca-se ainda que o processo de precipitação na forma de chuva pode ocorrer de três formas principais, i.e., chuvas convectivas, orográficas e frontais.

Assim, dada a relevância do recurso água e do seu movimento na atmosfera, corrobora-se a importância e relevância do estudo e entendimento dela e dos seus processos no meio ambiente. Portanto, é importante conhecer-se a precipitação os seus mais variados impactos sobre uma bacia hidrográfica, 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 (Collischonn & Dornelles, 2015).

É interessante destacar também que a precipitação apresenta alta variabilidade espaço-temporal, principalmente nas regiões áridas e semiáridas do planeta (Farhangi *et al.*, 2016), o que torna o seu monitoramento desafiador (Neto, 2020). De forma geral, as medições de chuva são feitas por postos pluviométricos (Cao *et al.*, 2018), utilizando aparelhos como o pluviógrafos e o pluviômetro. O pluviômetro é um aparelho dotado de uma superfície de captação horizontal, e de um reservatório, utilizado para a acumulação da água recolhida,

consistindo num aparelho que fornece o total de água acumulado durante um intervalo de tempo. Por outro lado, mesmo que os pluviógrafos se assemelhem ao pluviômetro quanto a sua estrutura, este é capaz de realizar medições contínuas da chuva, o que possibilita numa análise temporal mais detalhada.

Entretanto, mesmo que esses aparelhos representem a mais importante e direta forma de obtenção de dados de precipitação, eles apresentam um caráter pontual, o que torna a caracterização espaço-temporal em zonas extensas e remotas como um fator dependente da quantidade e disponibilidade de aparelhos. Além disso, 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, i.e., no Brasil existem cinco diferentes regimes climáticos (Reboita *et al.*, 2010; Rao *et al.*, 2015) e cerca de 25 diferentes zonas pluviometricamente homogêneas espalhadas por todo o território (Keller Filho *et al.*, 2005),

Em relação a quantidade de postos pluviométricos, a *World Meteorological Organization* WMO (1994) recomenda que a densidade de pluviométrica das regiões sejam de pelo menos um aparelho para cada 575 km², para um bom monitoramento, entretanto, em diversas regiões do mundo isso não é o que ocorre. No Brasil, por exemplo, Gadelha *et al.* (2019), mostrou que a rede de monitoramento de chuva do país conta com aproximadamente 11.820 estações pluviométricas, resultando em uma densidade de cerca de um pluviômetro a cada 720 km².

Além da densidade de estações pluviométricas ser escassa, o Brasil ainda conta com o fato de que séries de dados de precipitação ininterruptos serem 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 (Buarque *et al.*, 2011; Collischonn *et al.*, 2008; Curtarelli *et al.*, 2014; Delahaye *et al.*, 2015).

Assim, dois métodos que vem sendo amplamente também estudados e trabalhados para a obtenção de dados de precipitação são os radares climáticos, e o monitoramento via satélite. Em relação as observações por radar, de acordo com Tang *et al.* (2015), essas estimativas de chuva realizadas pelos mesmos acabam por apresentar grandes incertezas associadas a terrenos complexos, fazendo com que esse fato relacionado a erros de sinais eletrônicos diante de ambientes de operação desafiadores. E por outro lado, a partir do monitoramento via satélite, é

possível realizar estimativas de precipitação a nível global 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, e diferentes algoritmos, capazes de gerar estimativas de precipitação de maneira eficaz, e.g., *TRMM Multisatellite Precipitation Analysis data* (TMPA) (Huffman *et al.*, 2007), *Precipitation Estimation from Remote Sensing Information using Artificial Neural Network* (PERSIANN) (Sorooshian *et al.*, 2000), *Climate Prediction Center Morphing* (CMORPH) (Joyce *et al.*, 2004), *Global Satellite Mapping of Precipitation* (GSMaP) (Kubota *et al.*, 2007), *Integrated Multi-satellite Retrievals for GPM mission* (IMERG) (Hou *et al.*, 2014) e *Climate Hazards Group InfraRed Precipitation with Station data* (CHIRPS) (Funk *et al.*, 2015). 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), Cao *et al.* (2018), e Gadelha *et al.* (2019). Dessa forma, o presente estudo utiliza os produtos de precipitação oriundos do TRMM.

3.1.2 Produtos de precipitação TRMM

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 em regiões tropicais do planeta (Kummerow *et al.*, 1998). 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 o entendimento sobre o clima. Ele fornece importantes informações sobre precipitação, com o objetivo de 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). Dentre os diversos produtos disponíveis, o TRMM Multi-satellite Precipitation Analysis (TMPA) por exemplo, combina os dados de precipitação estimados por medições de sensoriamento remoto com as observações de precipitação de estações pluviométricas e radares, quando disponíveis.

É interessante destacar que o satélite caiu em meados de 2015, após 18 anos de atividade (Curtarelli *et al.*, 2014). Mas apesar do fim da missão, 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.

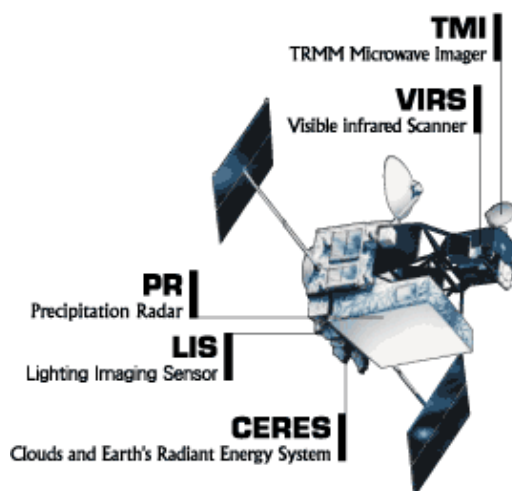


Figura 3 – Concepção artística do satélite TRMM (Fonte: NASA, 2020).

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.

Os dados de precipitação TRMM estão disponíveis em vários níveis relacionados com o grau de processamento aplicado aos dados de entrada. De acordo com a Nasa (2020) eles podem ser classificados em seis níveis. O Nível 0 denota os dados brutos; O nível 1A consiste nos dados que fazem uso de informações auxiliares, desde coeficientes de calibração radiométrica e geométricas, até a parâmetros de georreferenciamento; O Nível 1B consiste nos dados do Nível 1A radiograficamente corrigidos e geolocalizados; O Nível 2 representa os parâmetros geofísicos derivados na mesma resolução e localização dos dados de Nível 1; O Nível 3 denota os parâmetros geofísicos que foram reamostrados espacial ou temporalmente derivados dos níveis 1 e 2; Finalmente, o Nível 4 consiste em resultados de modelos que utilizam dados de nível inferior como entradas, e portanto não podem ser considerados dados brutos.

É interessante notar que as estimativas de precipitação TRMM foram consideradas as melhores dentre os produtos estimados por satélite de sua era (Jiang *et al.*, 2017), e isso se dá ao fato de que desde o seu lançamento, a missão TRMM passou por vários avanços, o que garantiu o aumento na acurácia dos produtos de precipitação. De acordo com os diversos estudos da acurácia do produto TRMM, na maioria dos casos, é apontado que os dados estimados pelo satélite TRMM são capazes de captar o comportamento espaço-temporal da precipitação em regiões com diferentes características, e.g., Collischon *et al.*, 2008; Plouffe *et al.*, 2015; Soares *et al.*, 2016; Teng *et al.*, 2016; Rozante *et al.*, 2018; Le *et al.*, 2020; Peng *et al.*, 2020, Neto, 2020).

3.1.3 Vazão

Vazão pode ser definida como o volume de água que passa por uma determinada seção de um rio ao longo de uma unidade de tempo, ou seja, ao trabalhar com a unidade de volume sendo dado em litros, e o tempo 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 (m³/s) (Collischonn & Dornelles, 2015). O escoamento superficial é uma das principais formas pela qual os cursos d'água recebem água, junto com a precipitação, o escoamento subsuperficial e o subterrâneo.

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) ANA (2013). 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, os rios podem ser classificados, segundo o tempo de fluxo de água, em perenes, intermitentes e efêmeros. Os corpos d'água classificados como 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 intermitentes são os que apresentam fluxo apenas durante a estação chuvosa, e que possuem canais definidos. E finalmente os efêmeros são os que apresentam fluxo apenas durante os períodos de chuva, e não possuem canais bem definidos.

Dentro do âmbito das vazões, as 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, i.e., 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 espalhados por todo o país.

É interessante destacar que a vazão que passa em um corpo d'água está intimamente relacionada com uma seção específica dele, 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 e velocidade.

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) (Collischon & Dornelles, 2015; UFRGS, 2009). O método do molinete consiste na determinação da área da seção de um rio, e posterior medição da velocidade em mais de um ponto da seção, obtendo-se assim a velocidade média vertical da seção, e em seguida, 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, e assim, proceder com o cálculo da vazão. O método químico consiste na injeção de uma

substância concentrada numa determinada seção do rio, e em seguida medir a concentração em um certo ponto a jusante, calculando-se assim a vazão. 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. Outros 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 (UFRGS, 2009). E por fim, o método pelo nível da água é um método que utiliza uma relação entre o nível de uma seção de rio, e a vazão, que é regida por uma curva-chave pré estabelecida, e de acordo com as leituras feitas diariamente, é possível realizar a determinação de vazões regularmente para essa seção. É importante salientar que a relação proposta por meio da curva-chave consiste numa simplificação, sendo, portanto, considerada válida apenas para o caso de ocorrência de uma morfologia constante do rio, com uma geometria da seção que não se modifica com o tempo. Pode ser considerado como sendo o método mais aplicado em postos e estações fluviométricas.

3.2 MODELAGEM HDROLÓGICA

Um modelo hidrológico consiste numa ferramenta utilizada para representar processos que ocorrem numa bacia hidrográfica (Tucci, 2009), ou seja, um modelo consiste numa ferramenta capaz de calcular coerentemente a vazão em uma determinada seção de um rio, baseado em parâmetros observados de precipitação, evapotranspiração e tipo e uso do solo daquela determinada região. 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).

Modelos determinísticos são aqueles que para uma mesma entrada, o modelo sempre fornecerá a mesma saída. Os estocásticos, consistem em modelos que possuem pelo menos uma relação aleatória no comportamento das variáveis, ou seja, para uma mesma entrada, a cada modelagem, os resultados fornecidos poderão variar, a depender da aleatoriedade da modelagem. Modelos empíricos são os modelos que fazem uso de relações entre as variáveis baseadas em observações reais, já os baseados em processos, procuram explicar determinado fenômeno baseado na complexidade das relações em geral físicas ou conceituais existentes. Modelos contínuos são caracterizados por serem aqueles 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, enquanto os distribuídos são mais gerais. 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.

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 (Almeida e Serra, 2017). Nesse sentido, um modelo chuva-vazão é aquele que a partir de dados de chuva, simula vazões para uma determinada bacia hidrográfica, e.g., *Soil Moisture Accounting Program* (SMAP) (Lopes *et al.*, 1981), *Tank Model* (Sugawara, 1961), o Modelo hidrológico auto calibrável (MODHAC) (Lanna Schwarzbach, 1989) e o *Soil and Water Assessment Tool* (SWAT) (Arnold *et al.*, 2012). Além disso, mais recentemente, destaca-se a ampla utilização de modelos baseados em redes neurais artificiais na análise chuva – vazão, e.g., Coulibaly *et al.* (2001), Jeong & Kim (2005), Farias *et al.* (2007), Wu & Chau (2011) e Santos *et al.* (2012a,b).

3.3 REDES NEURAIS ARTIFICIAIS

3.3.1 Definição e estrutura

Redes neurais artificiais são ferramentas computacionais desenvolvidas inspiradas no funcionamento do cérebro humano e no seu sistema nervoso, que apresenta uma grande capacidade para realizar tarefas complexas, atribuído ao processamento paralelo e distribuído natureza dos neurônios biológicos. Nesse sentido, a RNA imita tais estruturas onde o cálculo é processado através de unidades simples, chamadas neurônios artificiais, que estão interconectados para formar uma rede (Santos & Silva, 2014).

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 viés 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.

Quanto a estrutura, as RNAs podem ser entendidas como sendo conjuntos de neurônios artificiais interligadas por canais de comunicação, representados por um correspondente numérico. No caso de uma estrutura ser composta por três camadas, e.g., 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 fornecidos. 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, a RNA representa o somatório da entrada da rede com seus respectivos pesos. Na Figura 4 é mostrado um esquema da concepção de uma rede neural artificial.

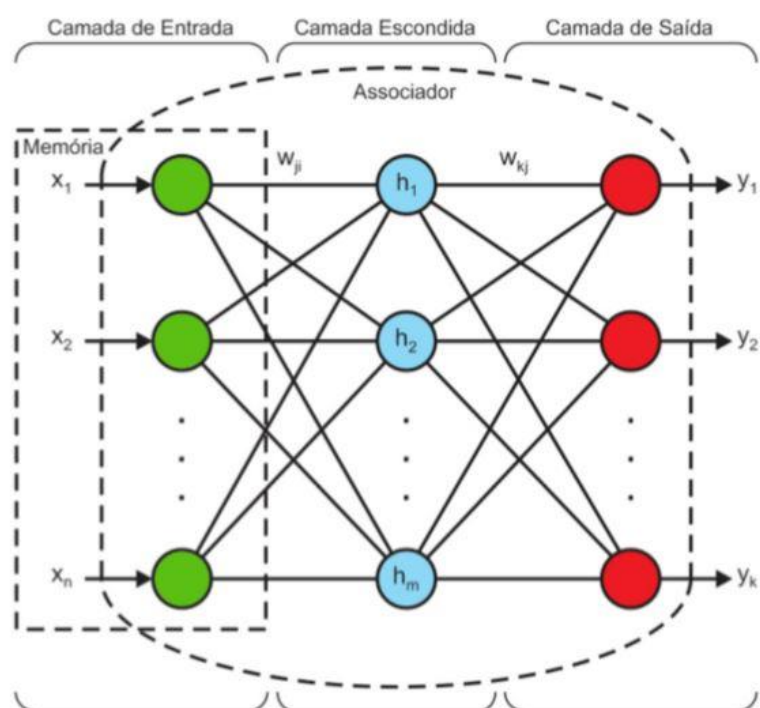


Figura 4 - Esquema da concepção de uma rede neural artificial (Fonte: Freire, 2012).

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.

3.3.2 Aplicações de RNA na modelagem e previsão de séries temporais

De acordo com Freire (2012), as RNAs são capazes de representar relações fortemente não lineares entre variáveis, levando à representação de sistemas complexos com uma precisão aceitável. Além disso, destaca-se que as redes podem extrair modelos e assim generalizar o mapeamento de valores de entrada e saída, além de que há a possibilidade de que as mesmas possam 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. Na literatura, diversos ramos científicos como a medicina, robótica, engenharia, finanças, segurança etc. fazem uso das RNAs, conforme Haykin, (1999) e Demuth *et al.* (2008).

Na área de recursos hídricos, a utilização de RNA é notada em diversas aplicações, e.g., desenvolvimento de modelos chuva-vazão (Hsu *et al.*, 1995; Jain *et al.*, 2004; Filho *et al.*, 2018 e Santos *et al.*, 2019), para a previsão da vazão em reservatórios (Campolo *et al.*, 1999; Kisi, 2004; Kumar *et al.*, 2004; Cigizoglu & Kisi, 2005; Santos & Silva, 2014; e Honorato *et al.*, 2018; e Santos *et al.*, 2019), para a previsão de precipitação (Hall *et al.*, 1999; Silverman & Dracup, 2000; Freiwan & Cigizoglu, 2005; Mirabbasi *et al.*, 2018); modelagem de água subterrânea (Coulibaly *et al.* 2001; Nayak *et al.* 2006; Mohanty *et al.* 2010), modelagem de qualidade de água (Singh *et al.* 2009, Gazzaz *et al.* 2012), estimativa de evapotranspiração (Kumar *et al.* 2002, Trajkovic *et al.* 2003, Zanetti *et al.* 2007, Adeloye *et al.* 2011), transporte de sedimentos (Nagy *et al.*, 2002; Cigizoğlu e Alp, 2006; Melesse *et al.*, 2011; e Farias e Santos, 2014) e modelagem da salinidade do solo (Nouredine *et al.*, 2019).

3.4 MAPAS AUTO-ORGANIZÁVEIS

3.4.1 Definição e estrutura das redes SOM

Redes SOM são redes neurais não supervisionadas que por meio de métodos de aprendizagem competitivos, agrupam os dados de entrada em classes de acordo com suas semelhanças (Kohonen, 1982). Essa rede neural reduz o conjunto de dados multidimensionais a uma matriz de recursos unidimensional ou bidimensional, que pode ser usada para análise de agrupamentos e fins de modelagem. 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 que possa ser utilizada para fins de análise e de previsão.

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. A rede SOM é composta por neurônios (as pequenas unidades da estrutura), que são organizados de maneira a preservar as relações vizinhas. Para isso, a rede agrupa os dados de entrada padrão de forma que padrões semelhantes sejam representados pelos mesmos neurônios de saída, ou por um de seus vizinhos.

A estrutura de uma rede SOM pode ser dividida em duas camadas, sendo a primeira a camada multidimensional de entrada, e a segunda a camada de saída, chamada também de camada competitiva. Esta é geralmente unidimensional ou bidimensional. Na Figura 5 é mostrado um esquema da arquitetura de uma rede SOM bidimensional.

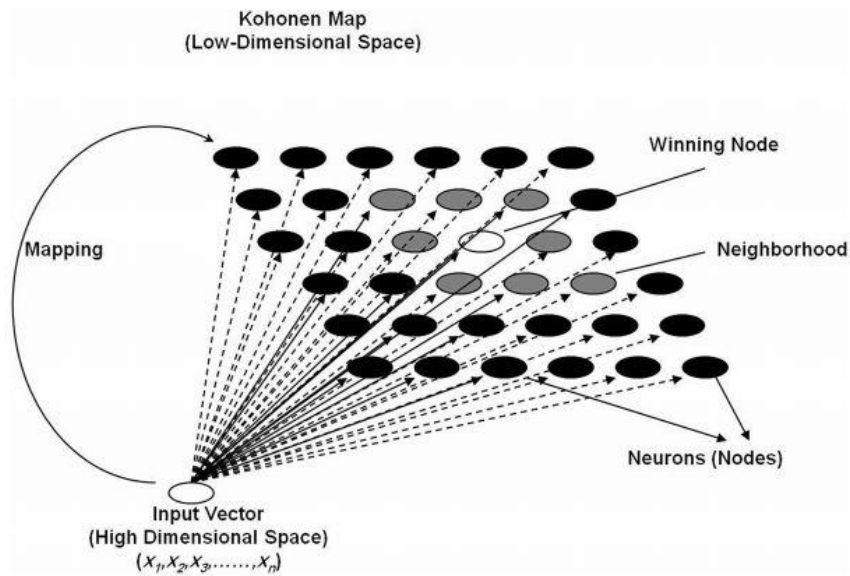


Figura 5 - Estrutura de um modelo SOM com detalhe do neurônio vencedor, e seus neurônios vizinhos (Fonte: Adeloze et al., 2011).

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 & 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, sendo e_1 e e_2 o maior, e o segundo maior autovalor do conjunto de dados de treinamento.

3.4.2 Treinamento

Inicialmente, os dados de entrada são normalizados (ou seja, o valor médio é zero e a variância é um). Esta etapa é necessária para permitir que todos os recursos se movam aproximadamente nas mesmas faixas e, possibilidade de que possam ser tratados pelo SOM da mesma forma. Em seguida, os dados são treinados de forma interativa na rede SOM. Existem valores de peso (w) atribuídos a cada neurônio, e o vetor de entrada n -dimensional (x) tem a distância entre seus valores, e os pesos para cada neurônio são calculados (Na Figura 6 é mostrado um esquema da distância entre de um neurônio para um modelo hexagonal). Cada

neurônio tem o número n -dimensional de pesos, um para cada variável de entrada. O critério mais usado é a distância euclidiana, como é mostrado na Eq. (1).

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

onde x_j é o j -ésimo componente do vetor de entrada x ; n é a dimensão do vetor de entrada x ; i varia de 1 a M , sendo M o número total de neurônios na camada de saída, e m_j sendo a máscara da função (geralmente $m_j = 1$), ou, com o propósito de excluir a contribuição do elemento x_j no cálculo euclidiano distância ($m_j = 0$), amplamente utilizado quando a variável de entrada tem valores ausentes.

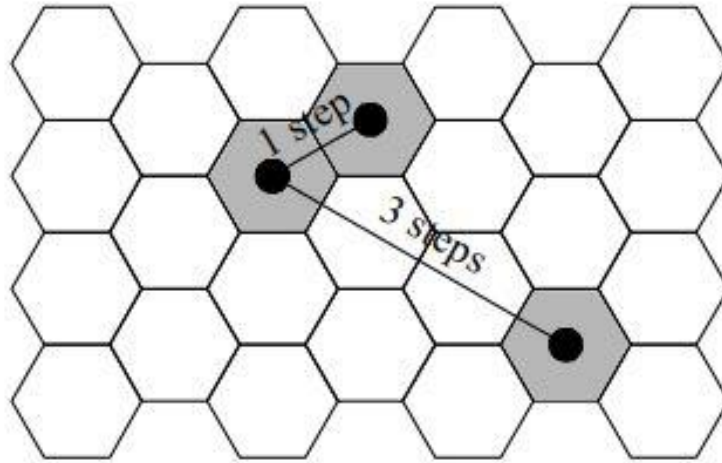


Figura 6 - As distâncias entre os neurônios de um modelo SOM para a determinação da vizinhança (Fonte: Farias et al., 2013).

Neste estudo, o tipo de treinamento escolhido é denominado modo *batch*, e consiste no fato de a apresentação dos vetores de entrada para a rede ser realizada utilizando todo o conjunto de dados antes da atualização de qualquer peso. Nesse sentido, a busca pelo neurônio vencedor é feita para cada vetor de entrada, e então o vetor de peso é movido para a posição média de todos os vetores de entrada para os quais é um vencedor (*best unit matches* - BMU) ou vizinho de um vencedor. Após todos os vetores de entrada serem calculados a distância euclidiana, o neurônio vencedor e o vizinho são selecionados.

Nesse método, os pesos tendem a se estabilizar após várias apresentações do conjunto de dados. É importante observar que este treinamento de rede não é supervisionado, pois não há saídas desejadas. Nesta fase de treinamento, o neurônio de saída cujo vetor de peso mais se

aproxima do vetor de dados de entrada (ou seja, para o qual a distância euclidiana é o valor mais baixo) é selecionado como o neurônio vencedor. Depois disso, é possível usar a função de uma vizinhança para atualizar todos os neurônios com base no vencedor, ou, como é usado neste estudo após cada iteração, os pesos conectados a cada vencedor de neurônio e neurônios vizinhos em um determinado raio de vizinhança são então atualizados pela regra de *Kohonen* (Beale *et al.* 2012) mostrada na Eq. (2).

$$w_{ij}(t + 1) = w_i(t) + \alpha(t)h_{bi}(t)[x - w_i(t)] \quad \text{Eq. (2)}$$

onde t representa o número de interações, $\alpha(t)$ é a taxa de aprendizado, $h_{bi}(t)$ é a função da vizinhança entre cada neurônio i e o BMU b , x é o valor médio das variáveis associadas a cada neurônio, e i corresponde ao BMU.

Tanto a taxa de aprendizagem quanto a função de vizinhança afetam a eficácia da aprendizagem do SOM e devem ser escolhidas com cuidado. Em particular, a taxa de aprendizado diminui monotonicamente com o aumento do número de iterações, conforme a Eq. (3), e uma função de vizinhança, que geralmente é escolhida como a função Gaussiana centrada no BMU, é mostrada na Eq. (4).

$$\alpha(t) = \alpha_0 \left(\frac{0.005}{\alpha_0} \right)^{\frac{t}{T}} \quad \text{Eq. (3)}$$

onde α_0 é a taxa inicial de aprendizado e T é o treinamento, forçando a convergência do vetor peso.

$$h_{ci}(t) = e^{\left(-\frac{\|r_b - r_i\|^2}{2\sigma^2(t)} \right)} \quad \text{Eq. (4)}$$

onde r_b e r_i são as posições dos nós b e i na grade SOM (respectivamente, a posição do BMU e dos outros neurônios) e $\sigma(t)$ é o raio da vizinhança. À medida que a taxa de aprendizagem $\alpha(t)$, $\sigma(t)$ também diminui monotonicamente à medida que o número de iterações aumenta.

O treinamento da rede usando o modo batch geralmente ocorre em duas fases: fase de pedido e fase de ajuste. Na primeira fase, o treinamento é limitado por várias apresentações de todo o conjunto de dados, e o raio da vizinhança começa com uma determinada distância que diminui até o valor da unidade. Essa medida permite que os pesos dos neurônios se organizem no espaço de entrada de forma consistente com as posições dos neurônios na grade dimensional. A fase de ajuste usa o número restante de apresentações estabelecidas para treinamento. Nesse

estágio, o raio da vizinhança está abaixo da unidade, o que significa que há apenas uma atualização do peso do neurônio vencedor. Durante esta fase, espera-se que os pesos sejam modificados de forma relativamente uniforme no espaço de entrada, mantendo a topologia definida na fase de pedido (Beale *et al.* 2012).

3.4.3 Cálculo e previsão dos valores das variáveis por meio de redes SOM

Finalmente, pode-se resumir que uma vez treinada a rede SOM, é possível utilizar o modelo como ferramenta de previsão ou cálculo de variáveis. Para tanto, (i) são calculadas as distâncias euclidianas entre os vetores de entrada e os pesos atribuídos aos neurônios de saída desconsiderando o elemento j a ser fornecido. Isso pode ser feito incluindo uma variável booleana m_j , conforme mostrado pela Eq. (1) A variável m_j é usada para incluir ($m_j = 1$) ou excluir ($m_j = 0$) a contribuição de um dado elemento j do vetor de entrada para o cálculo das distâncias euclidianas. Para o cálculo da variável pela rede SOM, de acordo com Farias *et al.* (2013), deve-se considerar no vetor de entrada $m_j = 0$; (ii) O neurônio vencedor é determinado com base na menor distância euclidiana; e (iii) o peso do neurônio vencedor conectado ao elemento j ausente do vetor de entrada é usado como o valor previsto.

3.4.4 Aplicações do SOM na modelagem e na previsão de séries temporais

Mesmo originalmente sendo aplicadas no agrupamento de variáveis, as redes SOM são também amplamente utilizadas na modelagem e análise de séries temporais. Em termos práticos, as redes SOM podem ser utilizadas para o agrupamento de variáveis, para a redução da quantidade de dados de treinamento num modelo, na interpolação não linear e extrapolação (ou seja, previsão), na 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; Adeloey *et al.*, 2011; Adeloey & Rustum, 2012).

Há diversas aplicações de redes SOM no campo da hidrologia. Na questão de agrupamento de variáveis (*clusterização*), García & González (2004), aplicaram o SOM para entender o comportamento das variáveis envolvidas nos processos de funcionamento de uma estação de tratamento de águas residuais, fazendo uso do agrupamento de variáveis. O estudo de Norvarini *et al.* (2019) fez uso da capacidade de agrupamentos da rede SOM nas etapas de modelagem do gerenciamento ótimo das pressões em redes de abastecimento de água. E o trabalho de Voutilainen e Arvola (2017) aplicou a rede SOM no agrupamento de 21 anos de dados ambientais complexos para um pequeno lago boreal intocado, e 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.

Em termos de modelagem, especificamente no campo da evapotranspiração, destacam-se a aplicação de redes SOM no trabalho de Adelo *et al.* (2011), que realizou o cálculo da evapotranspiração de referência em certas plantações com base em dados climáticos diários observados em duas bacias climáticas distintas (clima temperado e semiárido). Foi tomada como referência os valores de evapotranspiração de referência calculados pela equação de *Penman-Monteith*, empregando o conjunto completo de dados meteorológicos nos dois locais. Os resultados indicaram que além de eficaz para o cálculo da evapotranspiração, a rede SOM mostrou uma grande capacidade de modelagem inclusive na ausência de determinadas variáveis de entrada.

No que diz respeito a modelagem vazão-erosão, destacam-se os trabalhos de Farias & Santos (2014) e Farias *et al.* (2015) na utilização da rede SOM. Farias e Santos (2014) fizeram uma comparação entre a utilização de uma regressão linear múltipla com o SOM na modelagem da erosão, utilizando diversas variáveis de entrada, i.e., a chuva, a duração da chuva, a vazão do corpo d'água, a máxima e mínima temperatura diária, e o período de dias secos. Os resultados mostraram a superioridade das redes SOM frente a regressão linear múltipla, produzindo resultados mais confiáveis, devido a sua característica de detecção e extração de tendências não lineares. No trabalho de Farias *et al.* (2015) foram feitas algumas comparações entre 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 como resultados encontrou-se que as redes SOM se mostraram muito práticas na aplicação em determinadas bacias.

Na aplicação da modelagem chuva-vazão, tem-se como exemplos os trabalhos desenvolvidos por Adelo & Rustum (2012), Farias *et al.* (2012), Farias *et al.* (2013), Nourani *et al.* (2013) e Filho & Farias (2018). Adelo & Rustum (2012) fizeram a aplicação de redes SOM na modelagem chuva-vazão com o objetivo de reconstruir vazões em bacias inadequadamente cobertas por postos pluviométricos, e como resultados se encontrou o sucesso na utilização dessa ferramenta para vencer problemas práticos na espacialização de postos pluviométricos em bacias. Nourani *et al.* (2013) aplicaram as redes 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 posterior aplicação dos mesmos como

variáveis de entrada em uma rede neural de *feed-forward* (FFNN) para modelar processo chuva-vazão em escala de tempo diária, fazendo uso também da transformada *wavelet* (WT) para extração de recursos dinâmicos e em várias escalas das séries temporais de vazão não estacionária e para remover ruídos. O modelo composto (SOM-WT-FFNN) conseguiu diminuir a dimensionalidade das variáveis de entrada através da aplicação SOM, e conseqüentemente, a complexidade e o tempo de processamento da rede FFNN. Farias *et al.* (2012), Farias *et al.* (2013), e Filho & 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. Os modelos tiveram sucesso na eficiência das modelagens, bem como se mostraram capazes de demonstrar as correlações entre as variáveis de entrada e os resultados obtidos.

4 CARACTERIZAÇÃO DA ÁREA DE ESTUDO

A área de estudo deste trabalho é a sub-bacia do Alto São Francisco (Figura 7), que faz parte da bacia do rio São Francisco. A área foi escolhida para o estudo 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, e pelo fato de o rio de onde as vazões afluentes são observadas ser perene.

A sub-bacia está localizada entre as coordenadas 18.125°S e 20.875°S, e 43.875°O e 46.625°O, conforme é mostrado na Figura 7. Na Figura 7 a sub-bacia do Alto São Francisco é destacada dentro da bacia hidrográfica do rio São Francisco, sendo mostradas também as áreas de drenagem principal da bacia, juntamente com o rio principal destacado, i.e., rio São Francisco, e o reservatório Três Maria, sendo localizado neste reservatório o local de medição das vazões naturais utilizadas no presente trabalho. A sub-bacia do Alto São Francisco possui uma área de cerca de 49.574 km², que em termos relativos apresenta-se como maior do que países como a Suíça (41.285 km²), Holanda (41.543km²), a Dinamarca (42.933 km²).

A Figura 8 mostra o relevo da região baseado no produto Shuttle Radar Topography Mission (SRTM) com resolução de 30 m. A topografia da região é considerada ondulada, com elevações que variando de 600 a 1600 m, conforme mostrado na Figura 8. A Figura 9 mostra o uso e ocupação do solo na região de estudo com base no produto da base de dados do MapBiomas (<https://mapbiomas.org/en>). A 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), conforme mostrado na Figura 9.

De acordo com a classificação de *Köppen*, o clima da região é tropical úmido, sendo um verão chuvoso (outubro-janeiro), e um inverno seco (junho-agosto), i.e., duas estações distintas, com uma média de precipitação anual variando entre 1.100 a 1.600 mm. Cerca de 85% da precipitação anual ocorre durante a estação chuvosa (Silva *et al.*, 2018). Na Figura 10 é mostrada uma espacialização, por meio de uma interpolação *Spline*, da precipitação média anual na região. E finalmente, de acordo com o MMA (2017) a temperatura média anual da região é de cerca de 22°C e a evaporação é de cerca de 1.000 mm por ano.

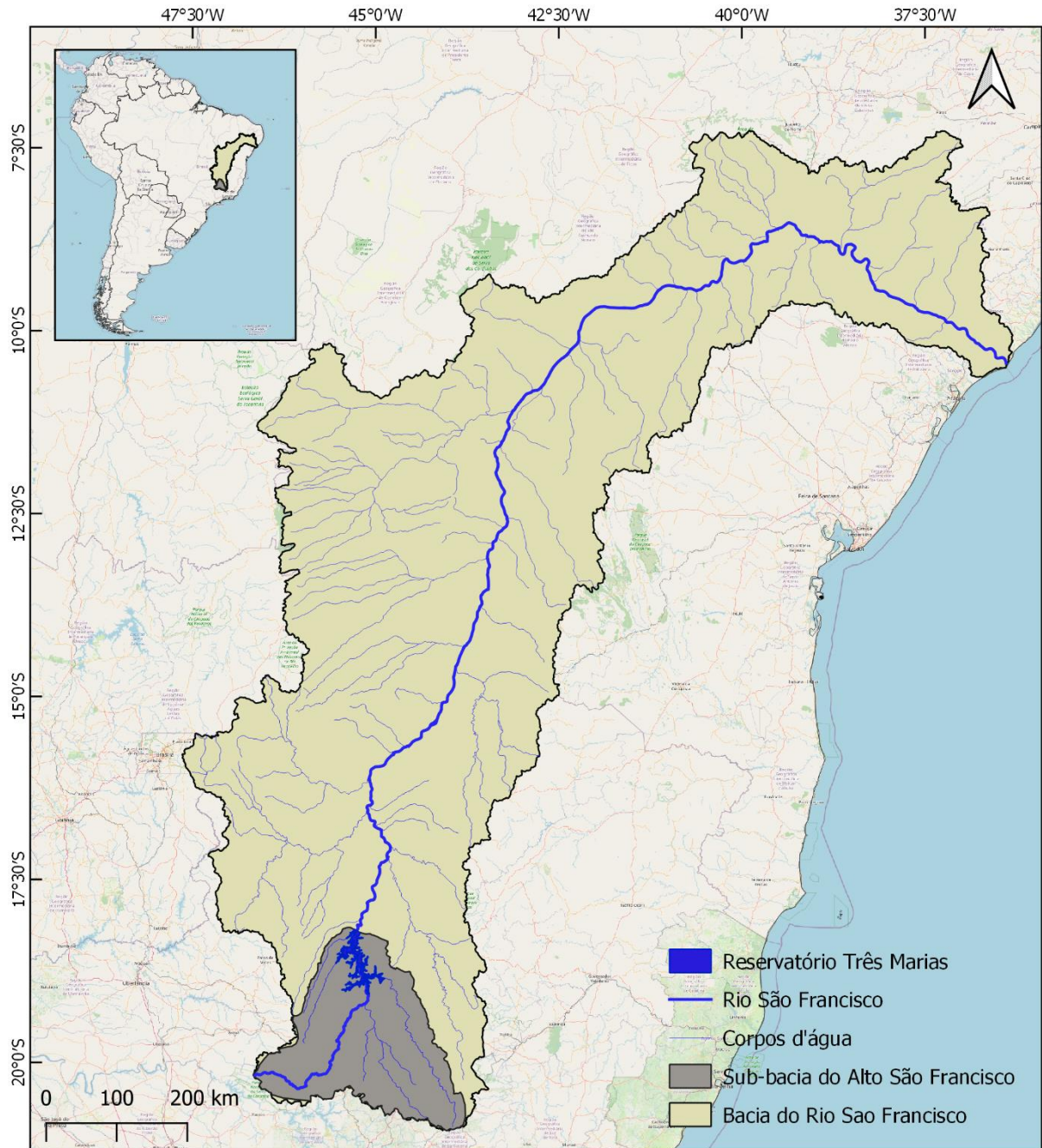


Figura 7 - Localização da sub-bacia do Alto São Francisco, com destaque para a sua rede de drenagem e para o reservatório Três Marias.

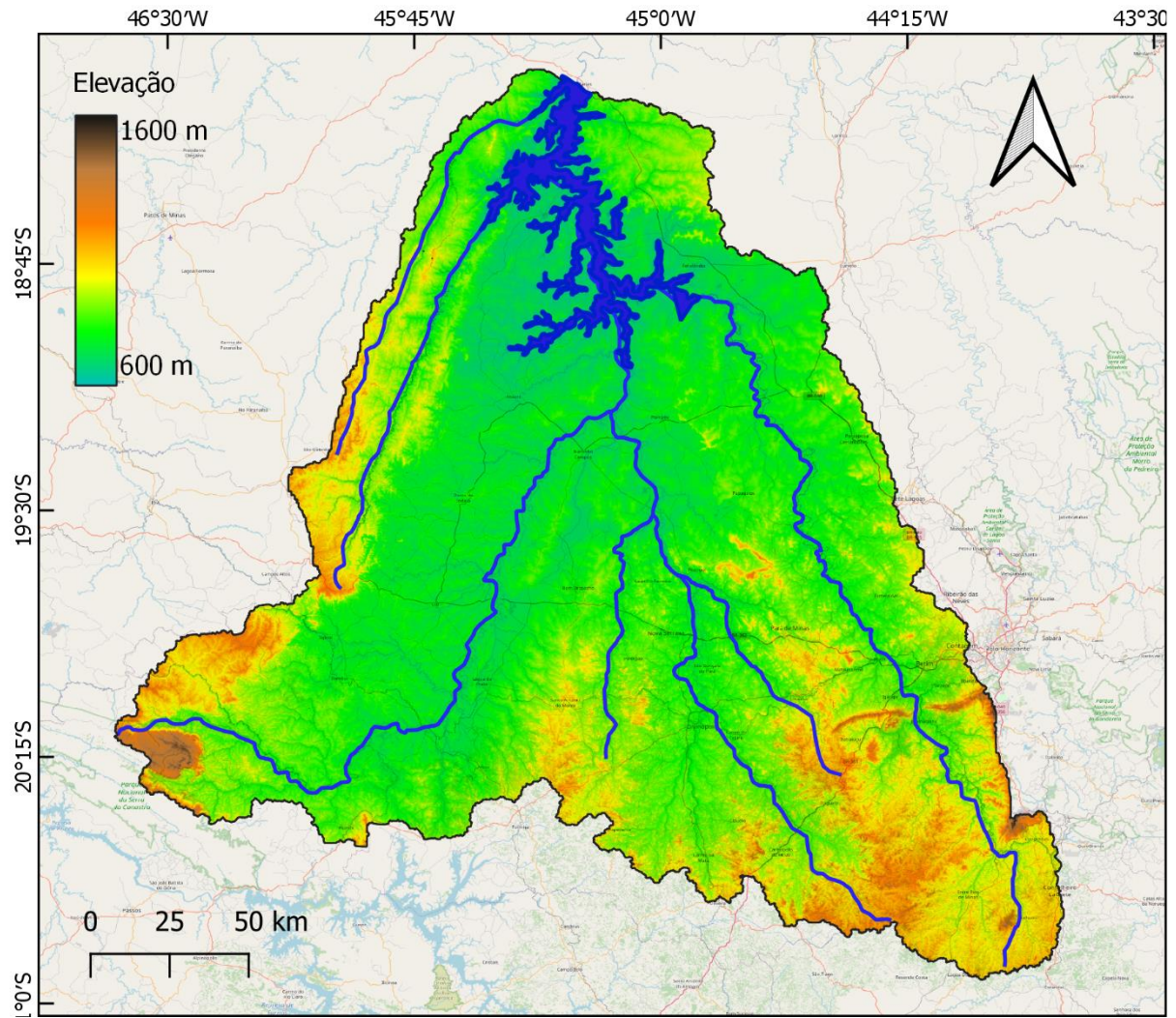


Figura 8 - Topografia da região.

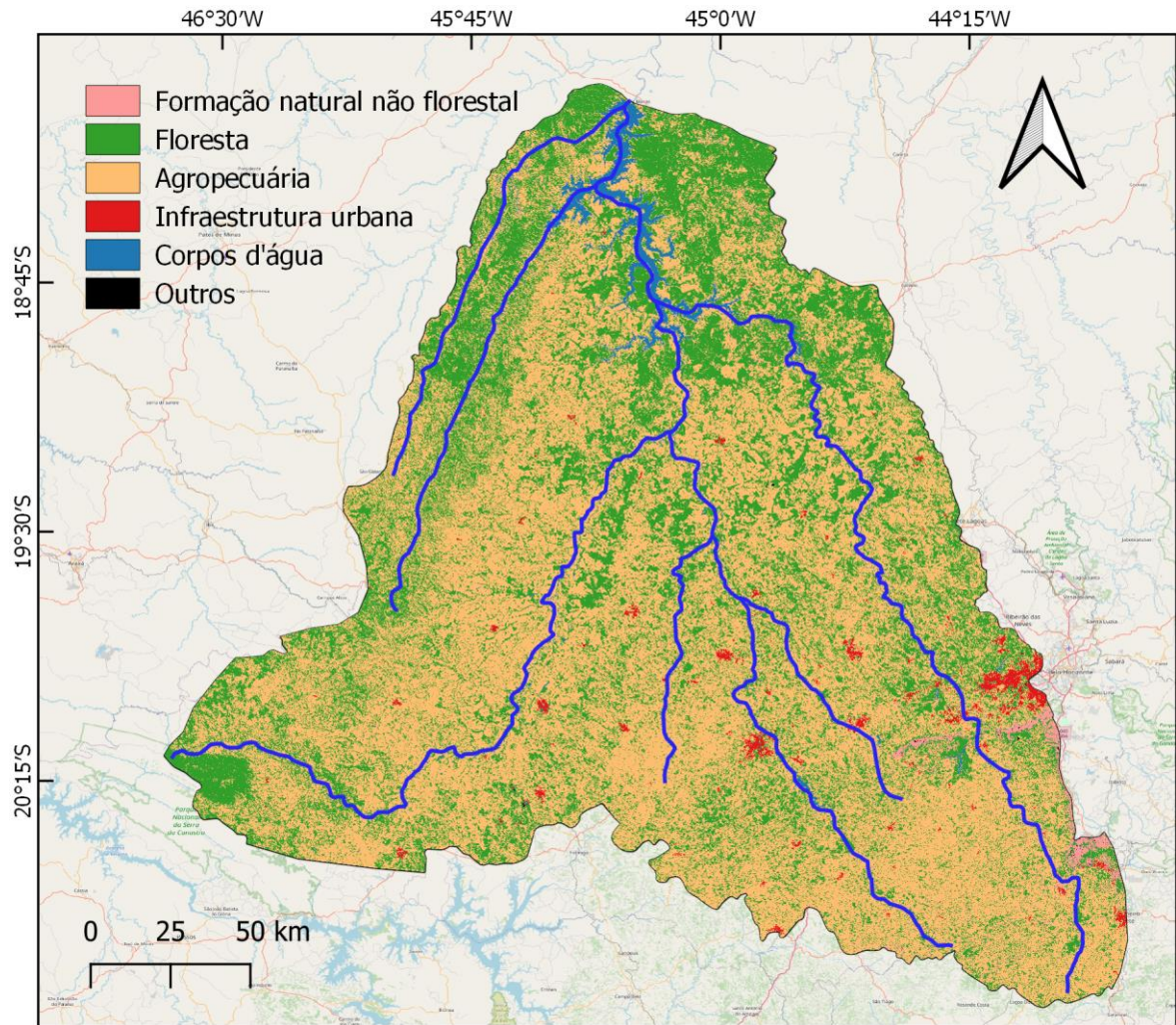


Figura 9 - Uso e ocupação do solo.

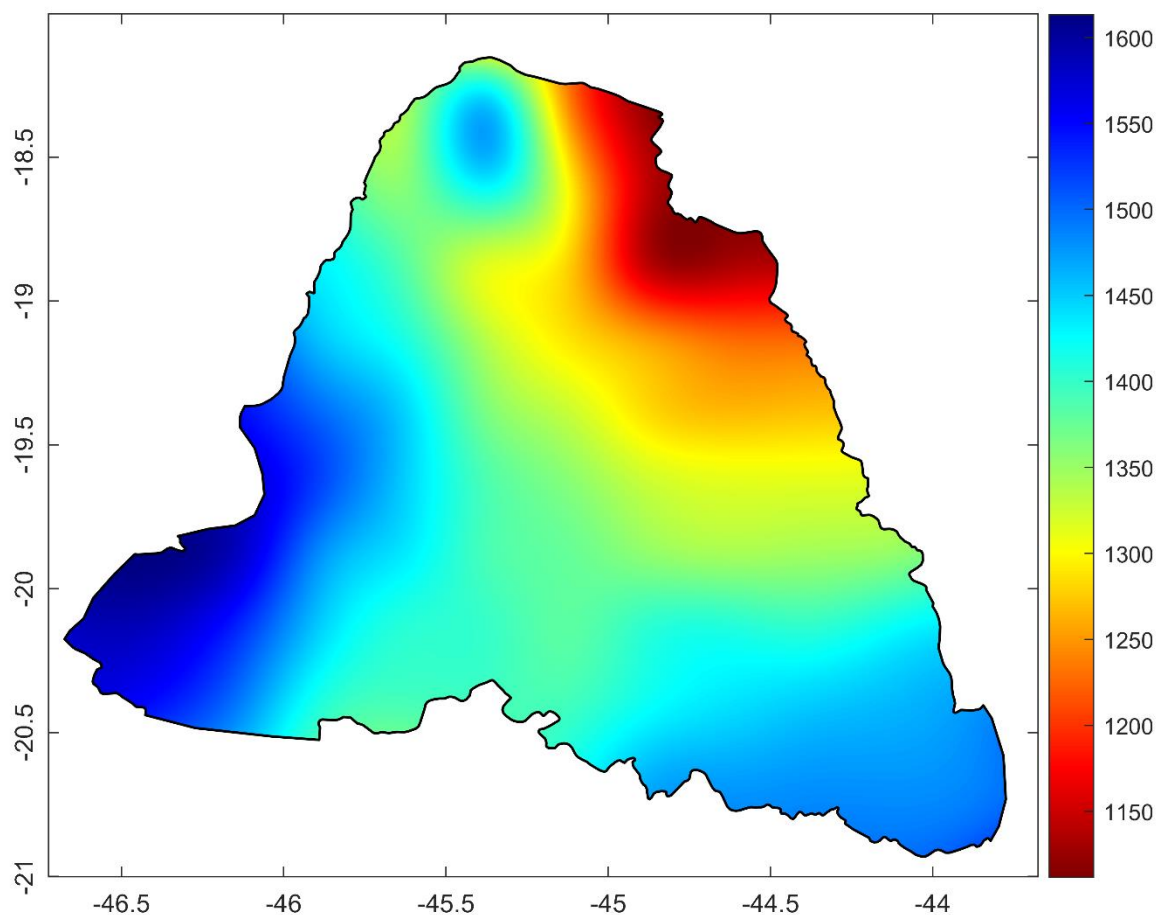


Figura 10 – Interpolação *Spline* da precipitação anual média na região (1998-2019) baseado em dados TRMM.

A Figura 11 mostra o reservatório Três Marias, onde é localizado o ponto de medição das vazões afluentes naturais utilizadas no presente estudo. Ele 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, sendo a maior obra regularizadora de vazão da região (CEMIG, 2020).

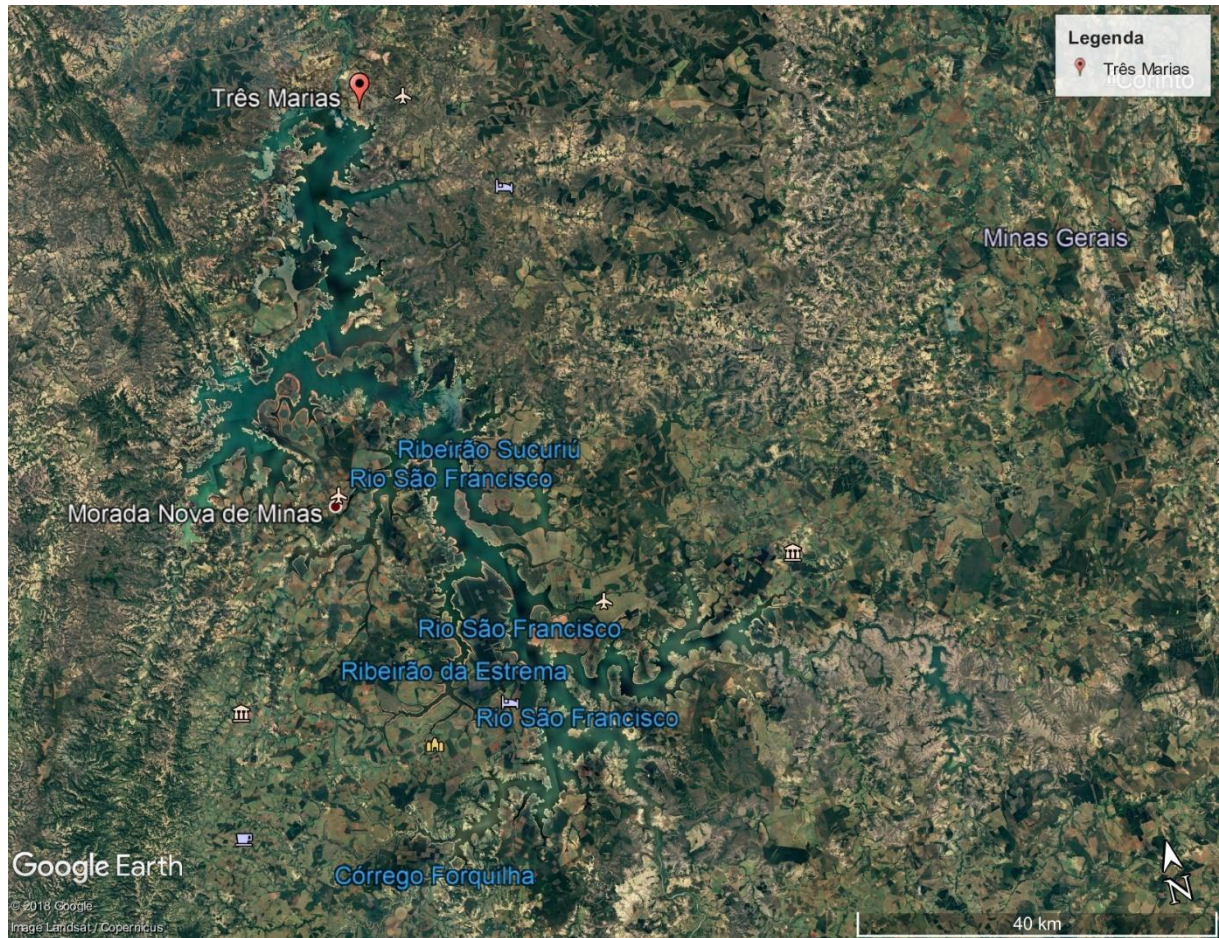


Figura 11 - Imagem de satélite do reservatório Três Marias (Fonte: Adaptado, Google Earth Pro, 2020).

5 MATERIAIS E MÉTODOS

5.1 DADOS UTILIZADOS

5.1.1 Dados de precipitação

No presente trabalho foram utilizados dados de precipitação mensal do satélite TRMM (*Tropical Rainfall Measuring Mission*) da NASA, algoritmo 3B42 (versão 7), que é uma missão conjunta entre a agência espacial americana (NASA) e a agência espacial japonesa (JAXA) (Huffman *et al.*, 2007). Ele 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 série utilizada no presente trabalho variou de janeiro de 1998 até dezembro de 2019, o que totalizam 264 meses de análise.

Foram utilizados os pontos das quadrículas TRMM (equidistantes a cada 0,25°) contidas dentro da área da sub-bacia do Alto Rio São Francisco, conforme mostrado na Figura 12. Mesmo com a grade inicial para a área de estudo composta por 144 quadrículas, para o presente trabalho foram utilizadas apenas as quadrículas efetivamente dentro da área de estudo, totalizando em 91 pontos para análise. Os dados de precipitação foram multiplicados pelo percentual do tamanho de uma quadrícula completa que a sua quadrícula ocupava, ou seja, as quadrículas de número 42 e 43 tiveram a precipitação observada multiplicada por 100%, antes de utilizá-la na modelagem, enquanto quadrículas como as de número 7 e 28 tiveram o seu valor de precipitação observada multiplicado por um valor próximo a 0%.

Assim, o vetor de entrada para o modelo no que tange a precipitação consistiu numa matriz com 91 colunas, representando os centroides dos pontos do TRMM, por 264 linhas, onde cada uma representa cada um dos meses de análise (conforme mostrado na Figura 13). Na Figura 13 é mostrada a série de precipitações mensais para a área de estudo, e na Tabela 1 as medidas de variabilidade estatística principais para as variáveis (média, soma, desvio padrão etc.).

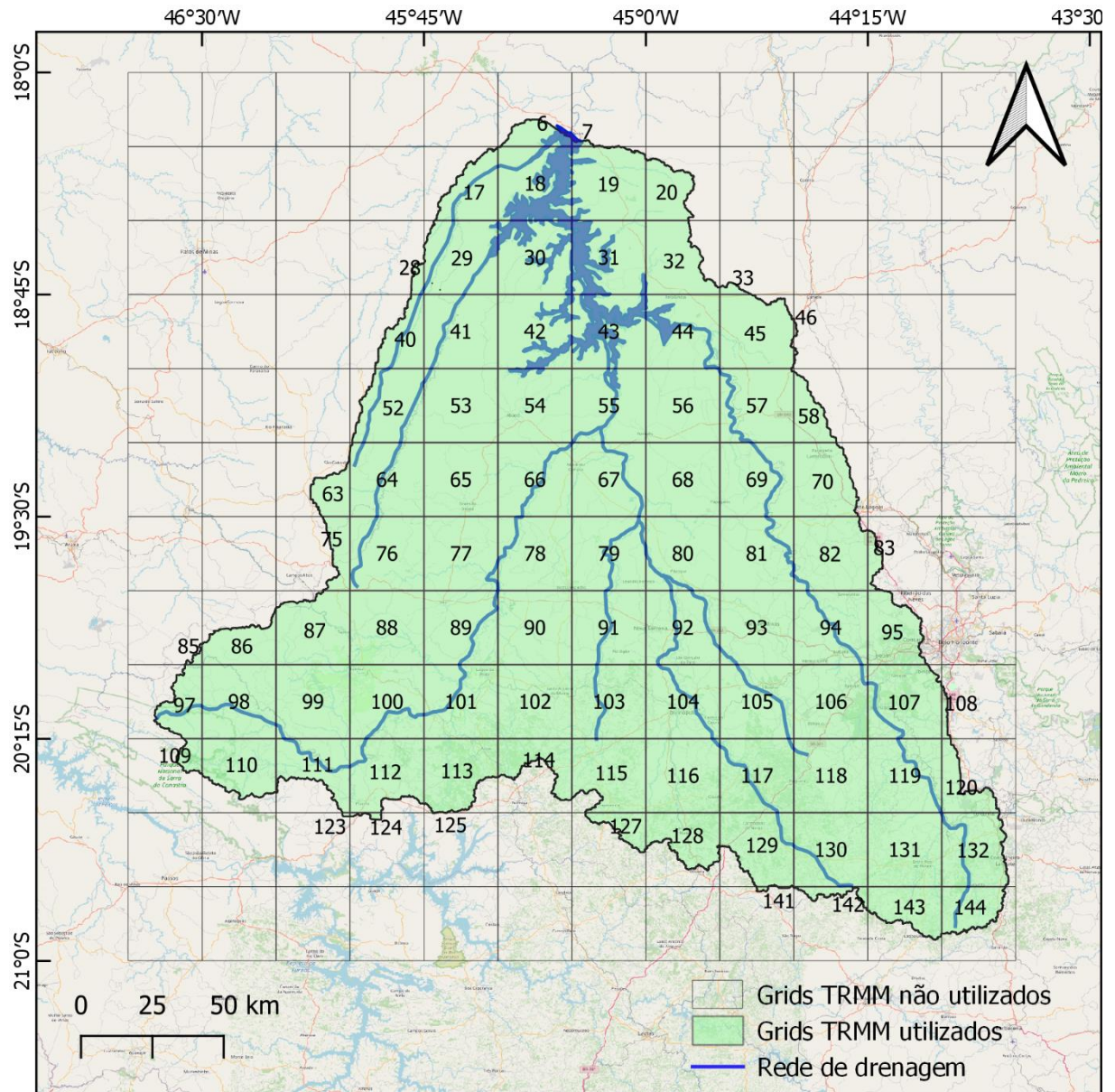


Figura 12 - Espacialização das quadrículas TRMM na área de estudo.

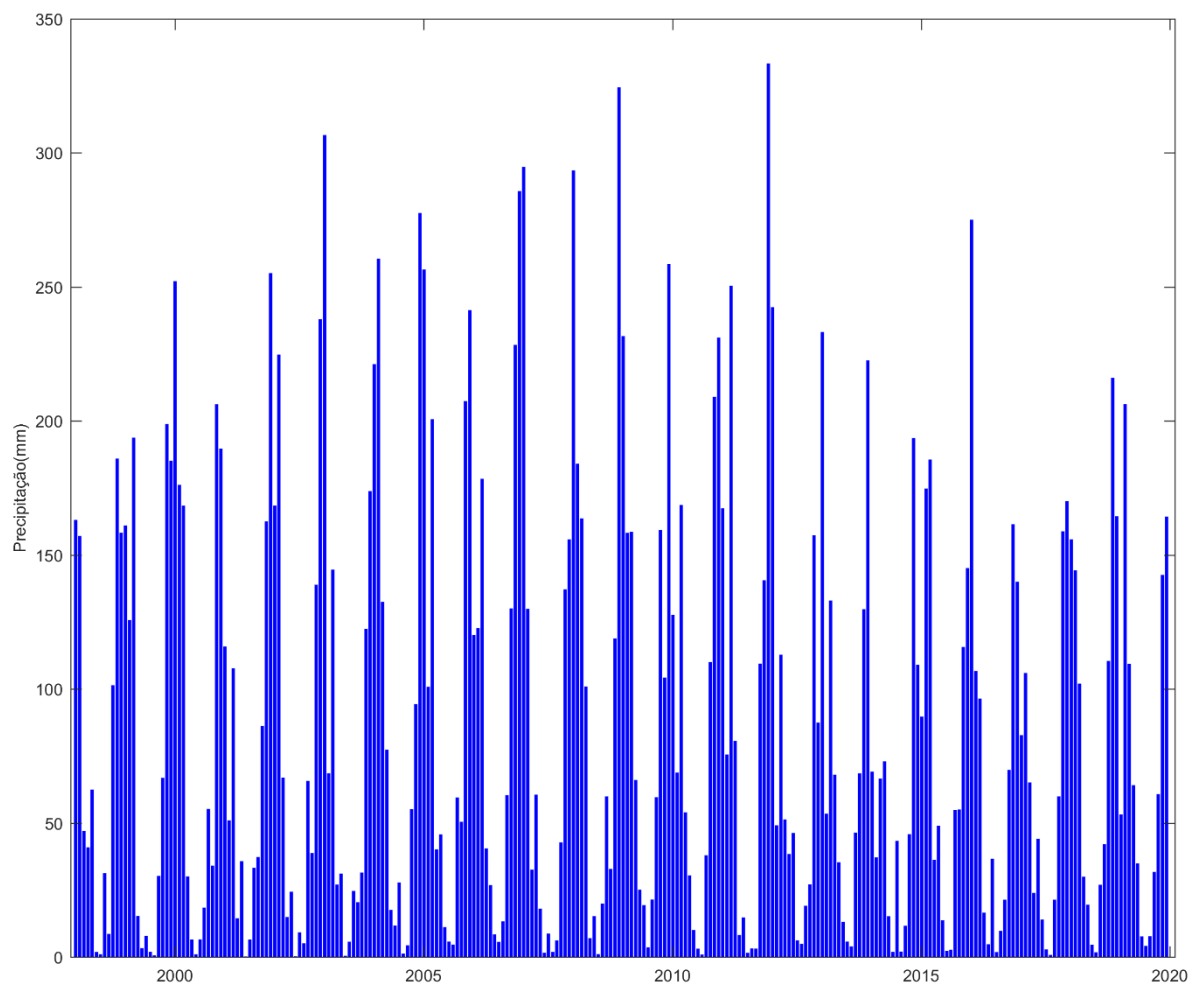


Figura 13 - Série de precipitações médias mensais para a região baseadas nas 91 quadrículas TRMM (1998-2019).

5.1.2 Dados de vazão

Para o presente trabalho foram utilizados dados de vazões naturais médias mensais afluentes à hidroelétrica de Três Marias, obtidos através do Operador Nacional do Sistema Elétrico – ONS (2020). 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, e 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 Natural 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 (ONS, 2020).

O período de estudo foi o de janeiro de 1998 dezembro de 2019. O vetor de entrada com os dados de vazões consistiu num conjunto de 264 dados, representando cada um dos meses analisados. Na Figura 14 é mostrado o Hidrograma com as vazões trabalhadas, e ao proceder-

se as análises dos dados, foi identificada uma inconsistência (não homogeneidade) da série de vazões para o período total trabalhado, i.e., 1998 a 2019. Foi aplicado o teste de homogeneidade de *Pettitt* (Pettitt, 1979), e identificou-se que a ruptura da na homogeneidade dos dados para o primeiro período consistente se deu a partir de abril de 2013, i.e., mês 184 (Figura 14). Portanto, essa não homogeneidade foi estudada dentro da modelagem SOM chuva-vazão, de acordo com a escolha três conjuntos de dados trabalhados (*datasets*). O entendimento do fator causador da mudança no padrão hidrológico na região não faz parte do escopo do presente trabalho, portanto não foi profundamente explorada.

Na Tabela 1 são mostrados as estatísticas descritivas, i.e., média aritmética, mediana, soma, máximo, mínimo, desvio padrão, variância e coeficiente de variação, para os três *datasets* de vazão utilizados no trabalho. Mais uma vez percebe-se a considerável variação entre esses descritores estatísticos para os conjuntos de dados, ressaltando-se a média aritmética da vazão, por exemplo, percebe-se uma diminuição de cerca de 42,19 %, i.e., de 638,4 m³/s para o primeiro *dataset* (janeiro de 1998 a abril de 2013), para cerca de 269,4 m³/s no segundo (maio de 2013 a dezembro de 2019).

Tabela 1 - Estatísticas descritivas para os dados de vazão e precipitação TRMM (médias ponderadas de Thiessen) mensais no reservatório de Três Marias (1998–2019).

Métrica	Vazão			Precipitação TRMM		
	01/1998-04/2013	05/2013-12/2019	01/1998-12/2019	01/1998-04/2013	05/2013-12/2019	01/1998-12/2019
Média aritmética	638.4 m ³ /s	269.4 m ³ /s	526.6 m ³ /s	90.4 mm	71.4 mm	84.6 mm
Mediana	410.6 m ³ /s	174.0 m ³ /s	338.7 m ³ /s	114.2 mm	87.6 mm	105.8 mm
Soma	117464.8 m ³ /s	21550.8 m ³ /s	139015.6 m ³ /s	8223.8 mm	6500.8 mm	7701.7 mm
Máximo	3016.2 m ³ /s	1104.9 m ³ /s	3016.2 m ³ /s	138.0 mm	113.3 mm	130.5 mm
Mínimo	75.2 m ³ /s	6.8 m ³ /s	6.8 m ³ /s	0.0 mm	0.0 mm	0.0 mm
Desvio padrão	588.8 m ³ /s	250.9 m ³ /s	537.6 m ³ /s	44.2 mm	35.1 mm	41.4 mm
Variância	346726.3 (m ³ /s) ²	62956.6 (m ³ /s) ²	289037.8 (m ³ /s) ²	1951.9 mm ²	1229.8 mm ²	1713.2 mm ²
Coeficiente de variação	0.92	0.93	1.02	0.49	0.49	0.49

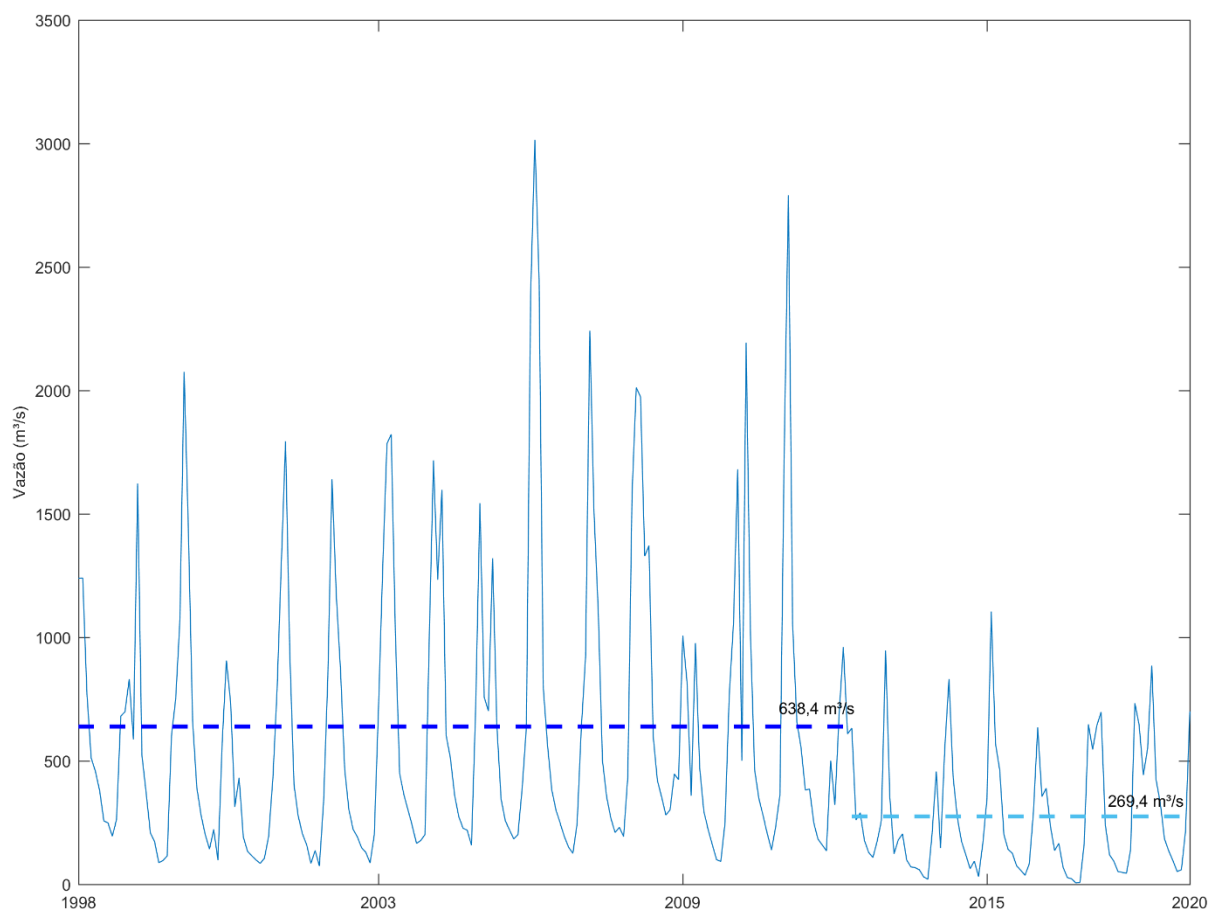


Figura 14 - Série de vazões médias afluentes naturais mensais para o reservatório Três Marias (1998–2019).

5.2 MODELAGEM SOM

5.2.1 Modelos e conjuntos de dados utilizados

Três conjuntos de dados (*datasets*) foram usados e três modelos foram criados para cada um dos três conjuntos (Figura 15). Os *datasets* correspondem a três períodos de estudo diferentes escolhidos devido à falta de homogeneidade da série completa de vazões naturais utilizada. O *dataset* #1 foi escolhido como compreendendo a série temporal de janeiro de 1998 até dezembro 2019; o *dataset* #2 como sendo janeiro de 1998 até abril de 2013; e o *dataset* #3 como maio de 2013 até dezembro de 2019. Cada um dos três conjuntos de dados foi subdividido em dois subconjuntos com 70% para a etapa de calibração e 30% para a etapa de teste. No entanto, às vezes os dados de entrada podem não ter um padrão estacionário, o que pode dificultar a calibração de um modelo proposto ou sua aplicação futura. Assim, a divisão da série temporal de vazões para as etapas de calibração e teste foi feita de três maneiras diferentes, i.e., (i) para a primeira modelagem, o conjunto de dados para a calibração compreendeu os primeiros

70% dos dados da série temporal (D1A, D2A e D3A), (ii) para a segunda modelagem, o conjunto de dados de calibração compreendeu os últimos 70% dos dados da série temporal (D1B, D2B e D3B) e (iii) para a terceira modelagem, o conjunto de dados de calibração compreendeu os primeiros e os últimos 35% dos dados da série temporal (D1C, D2C e D3C). Na Figura 15, é mostrado um esquema do processo de divisão de dados para modelagem. Por fim, a Tabela 2 resume as informações sobre cada conjunto de dados e modelos usados neste estudo. É importante destacar que cada variável de precipitação representa 91 quadrículas TRMM.

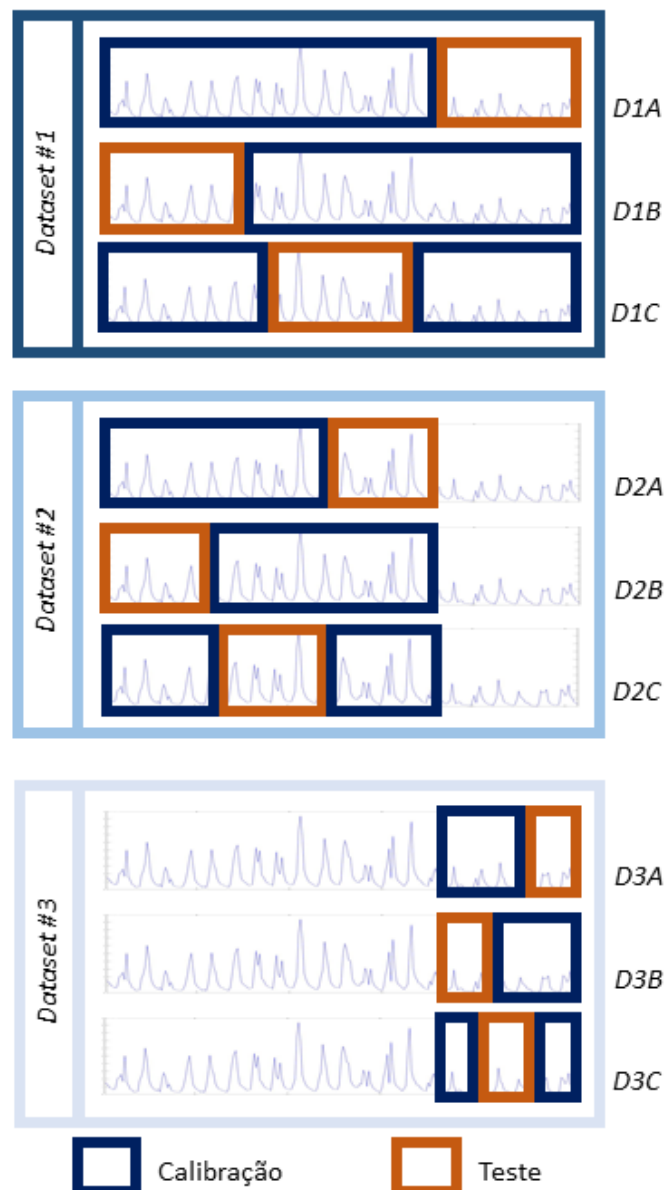


Figura 15 - Divisão dos três *datasets* utilizados na modelagem das vazões do presente trabalho.

Tabela 2 - Informações sobre os conjuntos de dados usados neste trabalho.

<i>Dataset</i>	Variáveis de entrada	Modelo	Período de calibração	Período de teste
#1	$CH_{t-5}, CH_{t-4}, CH_{t-3}, CH_{t-2}, CH_{t-1}$ e CH_t	D1A	01/1998-05/2014	06/2014-12/2019
		D1B	01/2005-12/2019	01/1998-12/2004
		D1C	01/1998-11/2005 & 7/2012-12/2019	12/2005-6/2012
#2	$CH_{t-5}, CH_{t-4}, CH_{t-3}, CH_{t-2}, CH_{t-1}$ e CH_t	D2A	01/1998-08/2007	09/2007-04/2013
		D2B	02/2002-04/2013	01/1998-01/2002
		D2C	01/1998-06/2003 & 03/2008-04/2013	07/2003-02/2008
#3	$CH_{t-5}, CH_{t-4}, CH_{t-3}, CH_{t-2}, CH_{t-1}$ e CH_t	D3A	05/2013-12/2017	01/2018-12/2019
		D3B	06/2015-12/2019	05/2013-05/2015
		D3C	05/2013-12/2015 & 02/2018-12/2019	01/2016-01/2018

5.2.2 Arquitetura da rede SOM

Neste estudo, o número de neurônios M para esta camada SOM foi determinado com base em García & Gonzáles (2004), onde o número de neurônios na camada é calculado com a Eq. (5).

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

onde N é o número total de amostras usadas na fase de calibração.

Foram escolhidas estruturas quadradas para as grades do SOM, e, nesse sentido, havia três conjuntos de dados, ou seja, três vetores de entrada diferentes para a fase de calibração, portanto, o número ótimo de neurônios M foi ~ 67 para o *dataset #1*, ~ 56 para o *dataset #2*, e ~ 36 para o *dataset #3*. Assim, as grades escolhidas visaram o menor número quadrado imediatamente maior do que o número M calculado, ou seja, foram escolhidas grades quadradas respectivamente de 9×9 , 8×8 e 6×6 , com uma estrutura hexagonal.

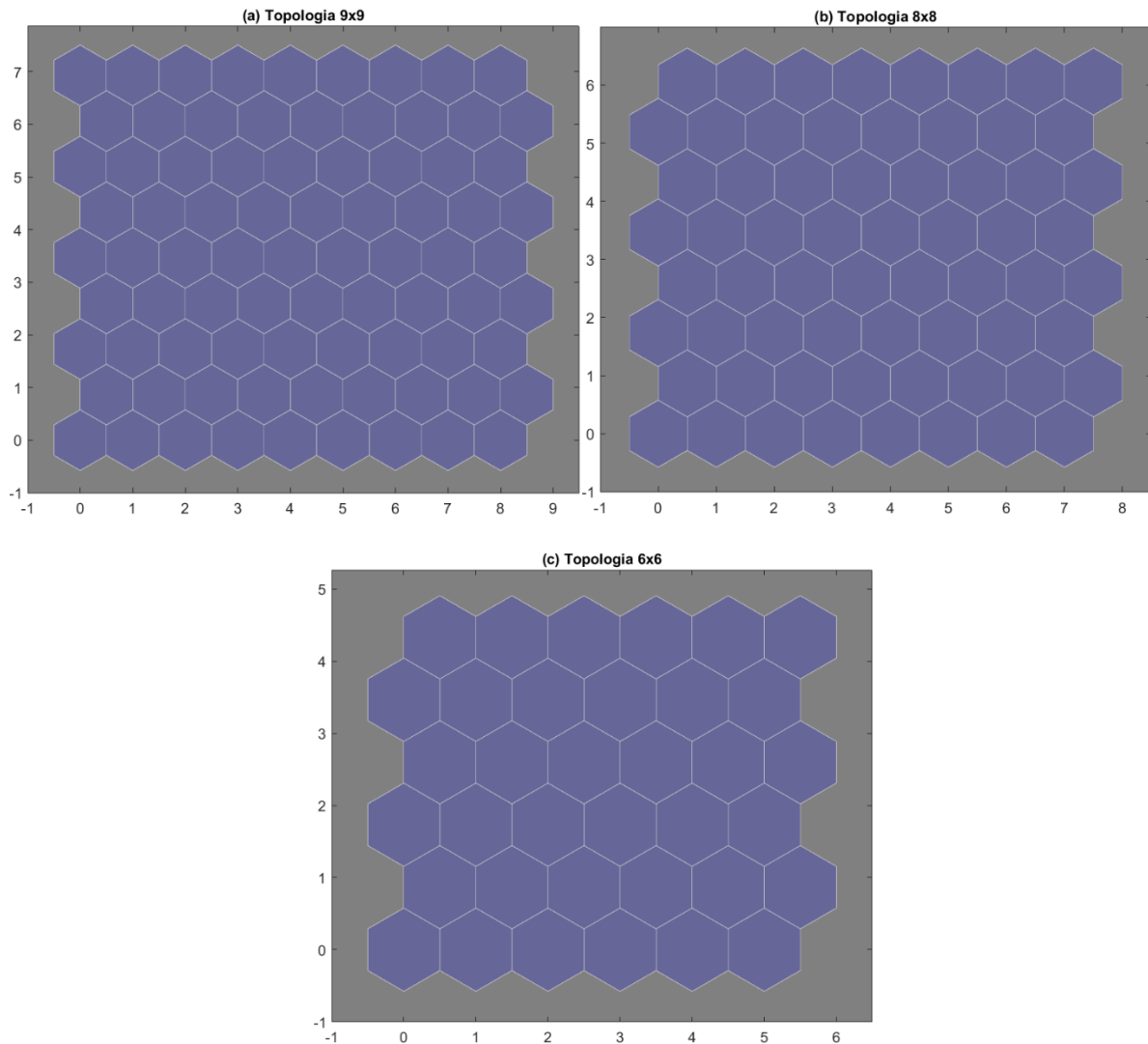


Figura 16 – Topologia das redes SOM utilizadas para os três conjuntos de dados trabalhados no presente estudo (a) Topologia 9x9, (b) Topologia 8x8 e (c) Topologia 6x6 (Fonte: Adaptado MATLAB).

5.2.3 Fase de treinamento

O processo de escalonamento consistiu em normalizar as entradas para que tivessem média zero e desvio padrão unitário (Beale *et al.*, 2012). A fase de treinamento do modelo utilizou o modo batch e, com o objetivo de garantir um aprendizado consistente para o modelo, os conjuntos de dados foram apresentados à estrutura do SOM 3.000 vezes. Para a fase de ordenação, foram utilizadas 300 apresentações do conjunto de dados e um raio de vizinhança inicial igual a três etapas. A fase de ajuste durou as 2700 apresentações restantes. Os modelos SOM foram calculados no *MathWorks' MATLAB R2017a* usando a *Neural Network Toolbox* (Beale *et al.* 2012). Especificamente, foi utilizada a função *selforgmap*.

Após a previsão, os valores obtidos são devidamente analisados quanto ao desempenho do modelo através das métricas estatísticas utilizadas no trabalho. Como trata-se de um modelo estocástico, i.e., que leva em conta variáveis aleatórias na sua calibração, é fato de que a cada vez que a rotina computacional é executada através do software MATLAB, os valores da previsão dos dados na etapa de calibração, mesmo com entradas idênticas, tendem a apresentar diferentes modelagens, assim, utilizou-se as métricas estatísticas para a escolha da modelagem que apresentou os melhores valores das métricas, e só a partir dessa etapa proceder-se com o teste dos modelos.

5.2.4 Fase de teste

Assim, após a fase de treinamento ser realizada, e os modelos com as melhores métricas estatísticas da fase de calibração serem escolhidos, procede-se com o teste dos mesmos, realizando-se o cálculo da vazão por meio dos dados de precipitação de entrada, e em seguida, os resultados do teste são analisados quanto ao desempenho, por meio das métricas utilizadas no trabalho.

5.3 MÉTRICAS ESTATÍSTICAS

Foram utilizadas quatro métricas estatísticas para a análise de desempenho dos modelos SOM no presente trabalho, i.e., (a) coeficiente de determinação (CD); (b) o coeficiente de eficiência de *Nash-Sutcliffe* (NSE; Nash & Sutcliffe, 1970); (c) Viés relativo (VR) e (d) a Raiz do erro quadrático médio normalizado (do inglês *Normalized Root Mean Square Error* - NRMSE). CD, NSE, RMSE e VR são métricas estatísticas amplamente utilizadas em análises hidrológicas (Moriassi *et al.*, 2015). Desta forma, neste estudo, optou-se pelo NRMSE, que é a versão normalizada do RMSE, e pode ajudar a comparar os resultados entre diferentes modelos, neste caso, modelos com diferentes ordens de magnitude (Viljanen *et al.*, 2018). As Equações 6-9 descrevem as métricas utilizadas.

$$CD = \left[\frac{\sum_{i=1}^n (O_i - \bar{O})(C_i - \bar{C})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^n (C_i - \bar{C})^2}} \right]^2 \quad \text{Eq. (6)}$$

$$NSE = 1 - \frac{\sum_{i=1}^n (O_i - C_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad \text{Eq. (7)}$$

$$VR = \frac{\sum_{i=1}^n (O_i - C_i)}{\sum_1^n O_i} \times 100 \quad \text{Eq. (8)}$$

$$NRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - C_i)^2}}{\bar{O}} \quad \text{Eq. (9)}$$

onde O_i são os valores observados, C_i são os valores gerados pelos modelos SOM, $\bar{O}$ é a média dos valores observados, $\bar{C}$ é a média dos valores gerados pelos modelos SOM e n é o tamanho da amostra.

O CD avalia o desempenho do modelo, calculando a variabilidade de um determinado número de previsões ao redor do seu valor verdadeiro, determinando a qualidade de uma relação linear entre duas variáveis, os valores variam de 0 a 1, onde o valor ideal é 1. O NSE avalia o ajuste entre os dados observados e estimados e o intervalo varia de $-\infty$ a 1, e o valor ideal é 1. O VR determina o quão bem o modelo simula as magnitudes médias para os dados calculados, em outras, a faixa varia de $-\infty$ a ∞ , e o valor ideal é 0. Valores positivos indicam que o modelo geralmente está superestimando os dados medidos, e negativos os valores indicam o contrário (o modelo está subestimando os valores). NRMSE é a versão normalizada do erro quadrático médio, em que valores menores de NRMSE indicariam maior precisão, ou seja, o valor ideal é 0.

6 RESULTADOS E DISCUSSÕES

6.1 DATASET #1

6.1.1 Componentes do plano e modelagem SOM

Na Figura 17 é mostrada a rede SOM (mapa de *Hits*) desenvolvida para cada uma das modelagens *dataset #1*, tanto para a etapa de calibração, como o seu resultado para a etapa de teste. Percebe-se uma esperada equivalência entre as redes, com destaque para o neurônio vencedor de cada uma, e para a vizinhança. Percebe-se uma adequação da rede ao número de variáveis, observada pelo preenchimento das variáveis em torno de todos os neurônios utilizados, não deixando neurônios inutilizados na rede. Com os mapas de *Hits* é possível realizar um agrupamento das variáveis de entrada do modelo, caso esse fosse o objetivo do trabalho.

Além disso, as correlações entre as variáveis podem ser identificadas por meio de gradientes de cores em cada componente do plano. Duas variáveis com gradientes paralelos mostram uma correlação direta, enquanto os gradientes antiparalelos mostram uma correlação inversa (Garcia & González, 2004). Na Figura 18 são apresentados os mapas de componentes do plano para as variáveis utilizadas nos modelos para a etapa de calibração da quadrícula 39. Essa quadrícula foi escolhida por ser a que apresentou a maior correlação entre chuva e vazão para o período estudado. Além disso, de acordo com a Figura 18 é possível notar uma semelhança entre os dados de chuva do mês atual, com a vazão do próprio mês para as modelagens D1A e D1C, com menor semelhança, mas ainda existente, para o mês anterior a vazão, evidenciando uma alta correlação entre as variáveis da chuva do mês com o valor da vazão.

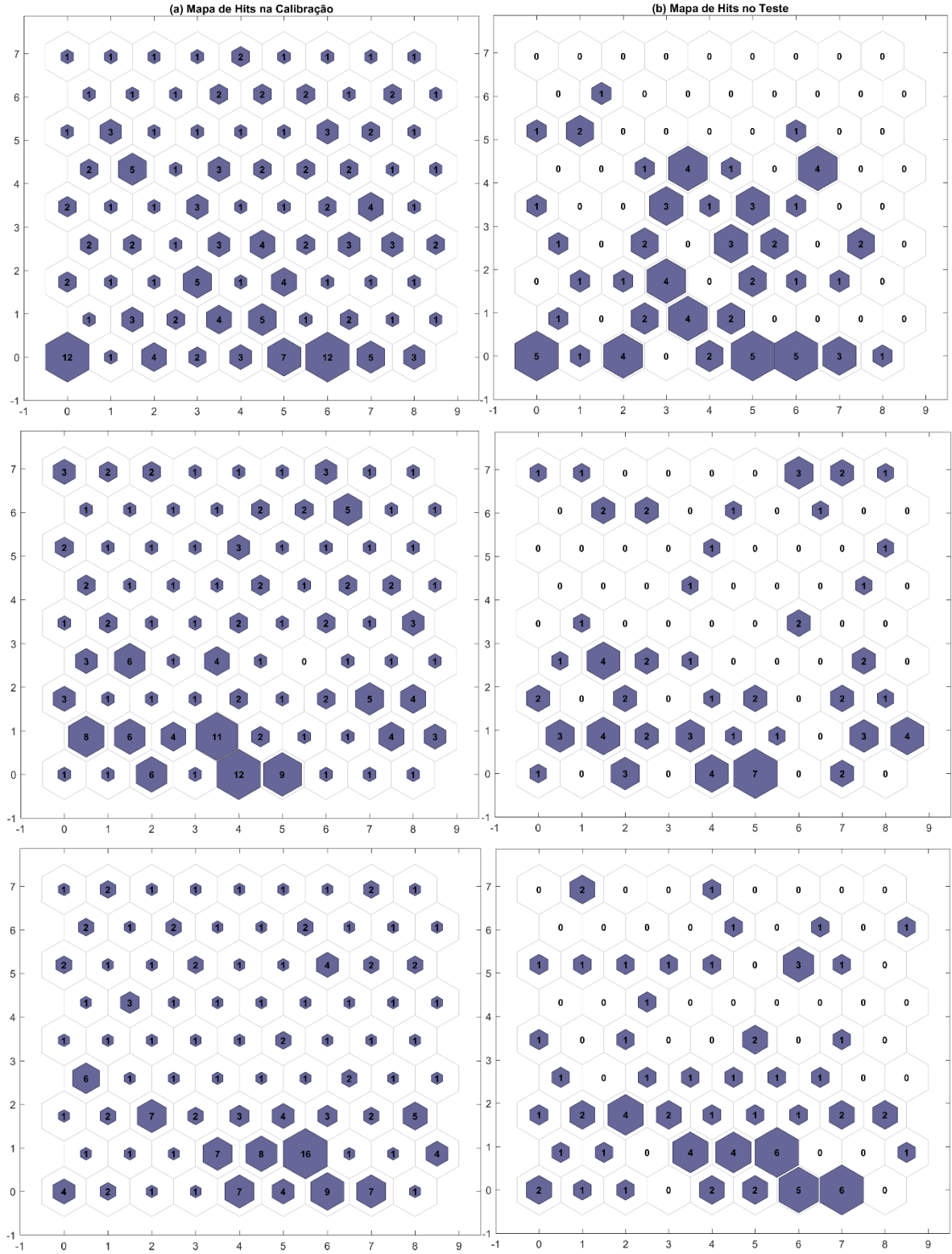


Figura 17 – Mapa de *Hits* das redes SOM para as etapas de calibração e teste para as três modelagens do *dataset #1*.

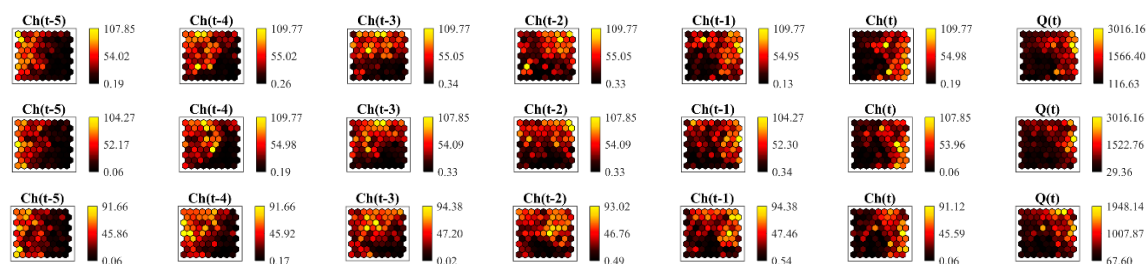


Figura 18 – Mapa dos componentes do plano da quadrícula 39 para as etapas de calibração dos três modelos desenvolvidos para o *dataset #1*.

6.1.2 Métricas estatísticas

Nesta seção são apresentados os resultados das modelagens para o primeiro conjunto de dados, i.e., *dataset #1*, que varia de janeiro de 1998 até dezembro de 2019, para os três modelos utilizados, i.e., D1A, D1B e D1C. O desempenho dos modelos chuva-vazão do SOM utilizados no presente trabalho foi avaliado utilizando as métricas CD, NSE, VR e NRMSE, e os resultados da aplicação das mesmas, tanto para a etapa de calibração, quanto para a de teste, estão mostrados na Tabela 3.

Para as etapas de calibração dos três modelos trabalhados, perceberam-se valores relativamente altos para as métricas CD, NSE e VR, e mesmo o NRMSE não sendo tão próximo de 0 (valor ideal), ele pode ser considerado baixo para os três modelos. Os valores de CD e NSE encontrados foram próximos a 0,94, como percebe-se na Tabela 3 para os três tipos de modelo, mostrando uma proximidade muito grande do valor ideal, i.e., 1. Para o VR, os valores foram bem próximos a 0 para os três modelos. E finalmente, para o NRMSE, os valores estavam entre 0,217 (primeiro modelo), 0,277 (segundo modelo) e 0,233 (terceiro modelo). Na Figura 19 é feita uma comparação gráfica entre os valores e vazão observados e os calculados pelo modelo na etapa de calibração para as três modelagens trabalhadas, percebendo-se visualmente uma boa compatibilidade entre os resultados.

Tabela 3 – Métricas estatísticas da etapa de calibração e teste para o *dataset #1*.

Métrica/Modelo	D1A		D1B		D1C	
	Calibração	Teste	Calibração	Teste	Calibração	Teste
CD	0,946	0,126	0,934	0,811	0,946	0,721
NSE	0,946	0,126	0,934	0,811	0,945	0,721
VR	-0,001	0,507	-0,002	0,043	-0,011	-0,234
NRMSE	0,217	0,870	0,277	0,402	0,233	0,484

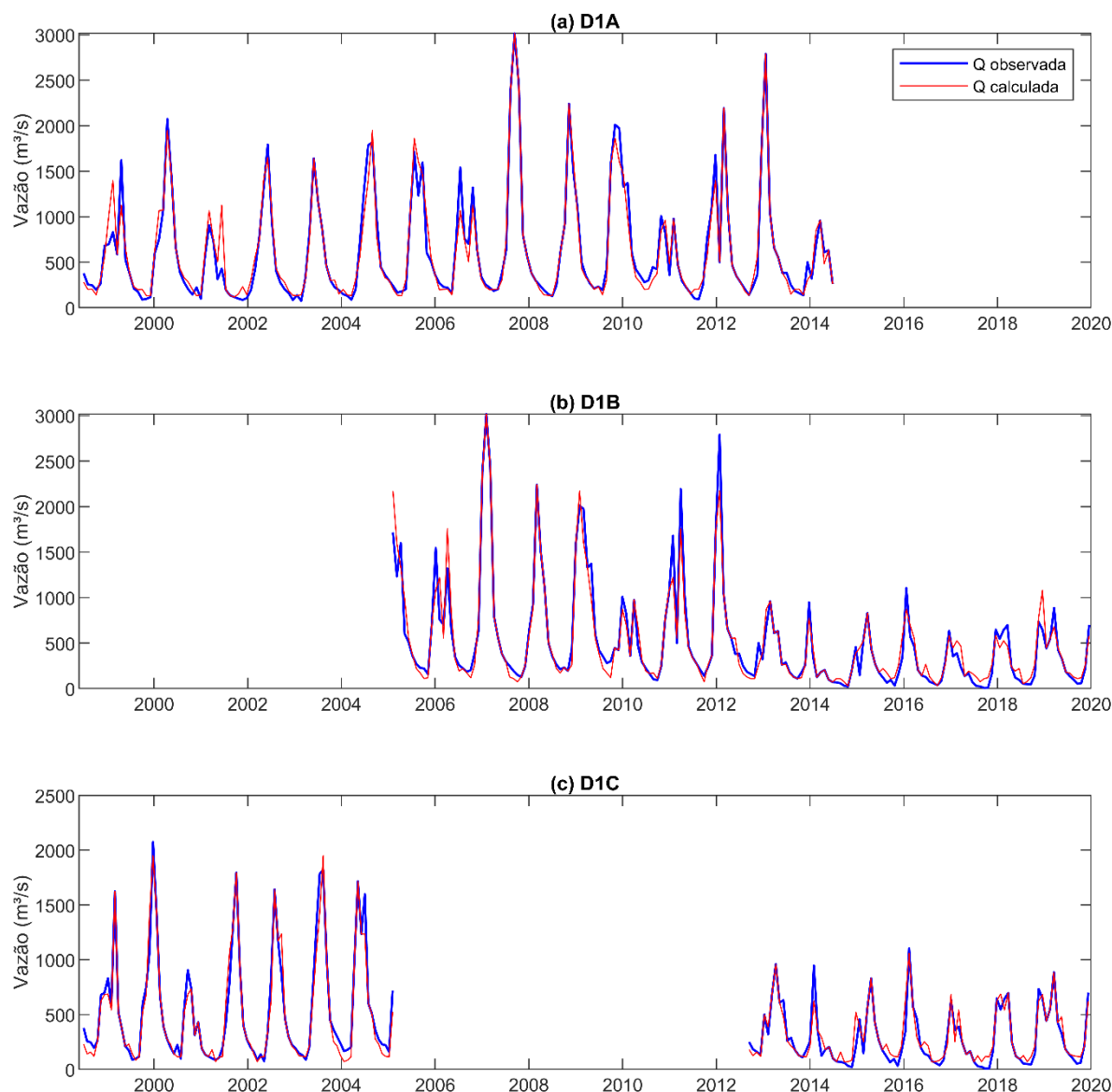


Figura 19 – Hidrograma de vazões observadas e calculadas para a etapa de calibração das três modelagens realizadas com o *dataset #1*.

Para a etapa de teste percebeu-se uma variação maior nos resultados obtidos. As métricas foram calculadas para os três modelos trabalhados, conforme a etapa de calibração. Para o modelo D1A os valores das métricas encontradas foram os piores dentre o *dataset #1*. O valor do CD encontrado foi de 0,126, o que está muito distante de 1 (valor ideal). Isso foi visto para o NSE, que também apresentou valor de 0,126, sendo também muito distante de 1. O valor de VR encontrado foi de 0,507, o que significa que o modelo superestimou os valores de vazão em cerca de 50,7%, em relação aos valores observados. Ao calcular o NRMSE percebe-se um valor extremamente alto de 0,870, muito longe do valor ideal de 0, mostrando 87% de erro pelo NRMSE. Na Figura 20 é mostrada uma comparação das séries observadas e calculadas pelo

modelo, e percebe-se visualmente que mesmo com métricas consideradas ruins, em termos práticos, a modelagem conseguiu acompanhar bem o padrão das vazões observadas, com picos e vales concentrados em regiões coerentes, mostrando que mesmo não sendo uma boa modelagem, o modelo já se mostrou suficientemente capaz de reproduzir, mesmo que em geral superestimado, o padrão da série. Além disso, conforme será discutido na seção de discussões, infere-se como uma das possíveis causas da modelagem pobre, o fato de a série utilizada não ser homogênea, sendo que coincidentemente os dados utilizados na etapa de calibração fazem parte de um comportamento hidrológico diferente dos dados utilizados na etapa de teste.

Para o modelo D1B perceberam-se valores melhores para as métricas estatísticas empregadas na avaliação do modelo na etapa de teste. O valor do CD e do NSE encontrados foi de 0,811, representando um valor alto, e que implica numa boa adequação do modelo em calcular os valores de vazão trabalhados. O valor do VR foi de 0,043, o que implica que o modelo está superestimando os valores de vazão observados em cerca de 4,3%, o que é um valor relativamente próximo de 0% (valor ideal). E finalmente encontrou-se um valor do NRMSE de 0,402, o que implica num valor quase mediano em relação ao valor ideal que seria de 0, mas que pode ser interpretado pelo fato de que mesmo sendo um bom modelo, o mesmo não calcula os valores de vazão exatamente iguais aos observados, mas de acordo com o VR de cerca de 4,7% a mais, o que faz com que a métrica apresente esse valor.

Observando a Figura 20, observa-se uma boa adequação entre as duas séries de vazões, que mesmo não estando exatamente sobrepostas, apresentam valores muito próximos, e que em termos práticos, do ponto de vista de obras hidráulicas, uma vez que a série calculada já acompanha as ordens de grandeza da observada.

Por fim, para o modelo D1C percebeu-se uma alta performance das métricas para o *dataset #1* na etapa de teste. O valor do CD e do NSE encontrados foi de 0,721, o que denota uma boa eficiência no modelo desenvolvido no cálculo das vazões. O valor do VR foi de -0.234, que é maior que o valor encontrado para o modelo D1B, mostrando que o modelo está subestimando os valores de vazão ao comparar com as vazões observadas. Finalmente, o valor de NRMSE encontrado foi de 0,484, o segundo menor dentre os três modelos do *dataset #1*, e fica válido o mesmo comentário feito para a situação da modelagem anterior, onde mesmo não sendo um valor tão próximo de 0 para NRMSE (o que seria o ideal), recorre-se ao fato de que é uma característica da métrica em mostrar as diferenças, mesmo que não tão robustas entre os valores calculados e observados de uma variável, e além disso, conforme percebe-se na Figura

20, é observada uma boa concordância entre as duas séries de dados de vazão, i.e., valores observados e calculados.

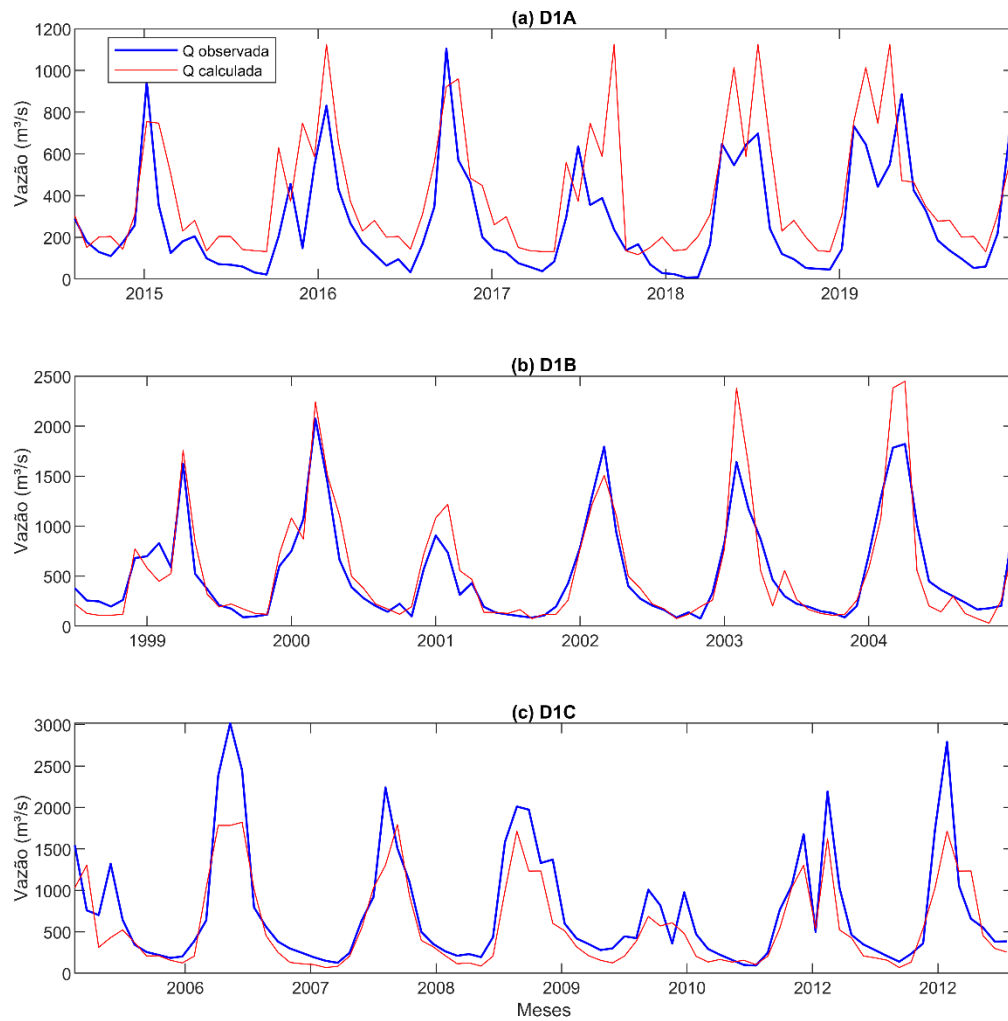


Figura 20 – Hidrograma de vazões observadas e calculadas para a etapa de teste das três modelagens realizadas com o *dataset #1*.

6.2 DATASET #2

6.2.1 Componentes do plano e modelagem SOM

De forma semelhante ao que foi apresentado para o *dataset #1*, na Figura 21 é mostrada a rede SOM que foi desenvolvida para cada uma das modelagens realizadas no *dataset #2*, sendo mostrado também tanto a etapa de calibração, como o seu resultado para a etapa de teste. Percebe-se de forma semelhante ao que foi feito anteriormente uma esperada equivalência entre as redes, com destaque para o neurônio vencedor de cada uma, e para a vizinhança. Destaca-se também a adequação da rede ao número de variáveis, observada pelo preenchimento das variáveis em torno de todos os neurônios utilizados, não deixando neurônios inutilizados na rede.

Além disso, as correlações entre as variáveis são representadas na Figura 22 são apresentados os mapas de componentes do plano para as variáveis utilizadas nos modelos para a etapa de calibração. Para cada um dos três modelos foram selecionados os mapas de componentes para a grade TRMM com a maior correlação entre os dados de chuva, e assim, percebe-se de forma geral como semelhanças entre os dados de chuva usados no modelo e a vazão modelada. De acordo com a Figura 22 é possível notar uma semelhança entre os dados de vazão e o de chuva do mês anterior, para os modelos D2A, D2B e D2C, além da semelhança com o mês atual de chuva, percebe-se também com o mês anterior, evidenciando uma alta correlação entre as variáveis do mês de chuva do mês atual e anterior, e de vazão.

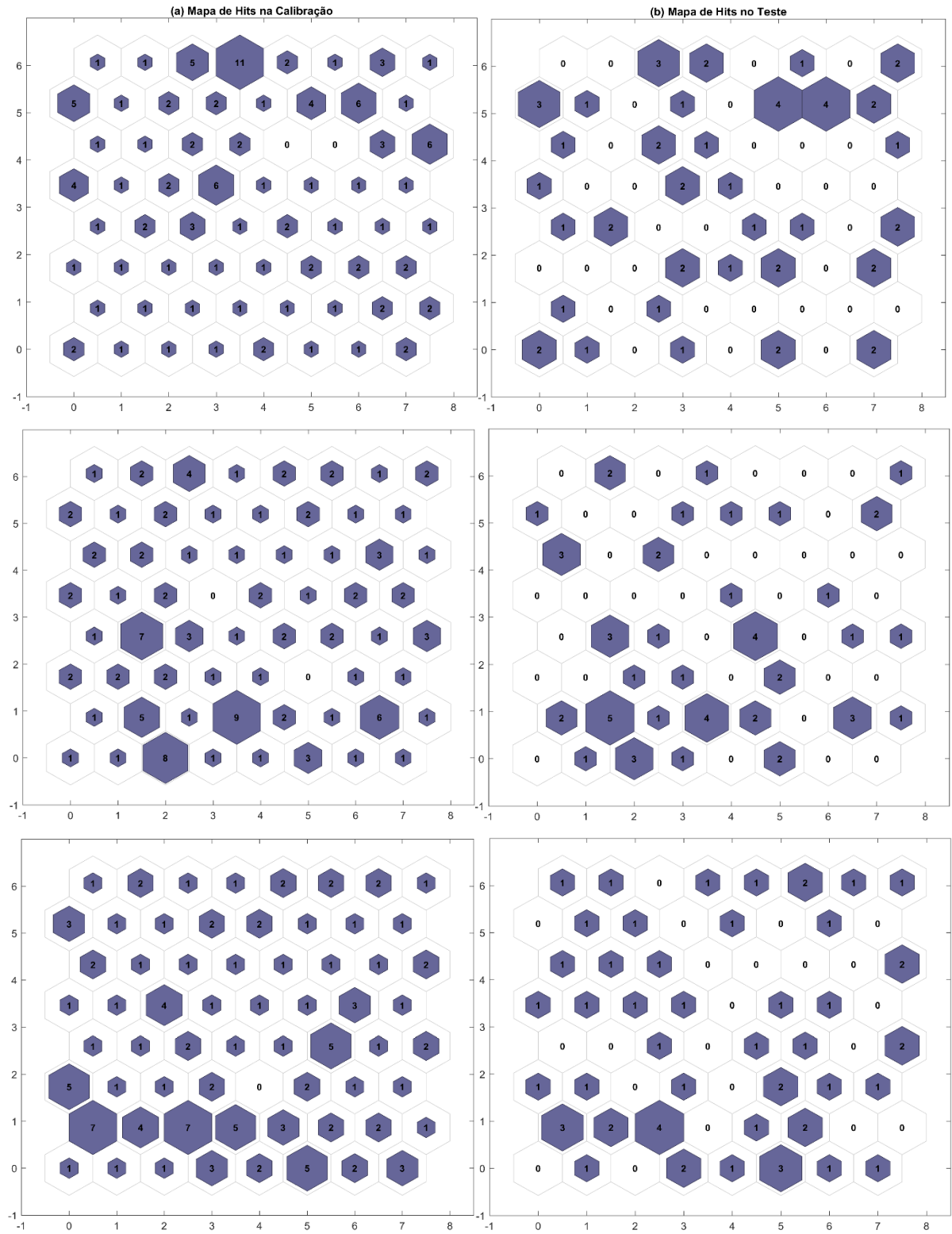


Figura 21 – Mapa de Hits das redes SOM para as etapas de calibração e teste para as três modelagens do *dataset #2*.

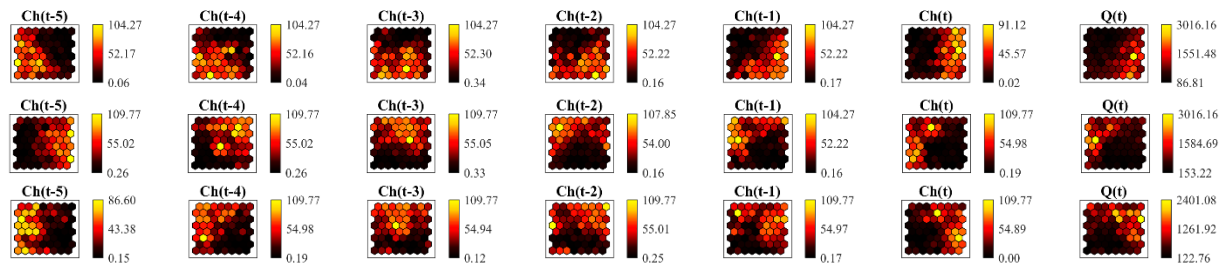


Figura 22 – Mapa dos componentes do plano da quadrícula 39 para as etapas de calibração dos três modelos desenvolvidos para o *dataset #2*.

6.2.2 Métricas estatísticas

Nesta seção são apresentados os resultados das modelagens para o segundo conjunto de dados, i.e., *dataset #2*, que varia de janeiro de 1998 até abril de 2013, para os três modelos utilizados, i.e., D2A, D2B e D2C. Na Tabela 4 são mostrados os resultados das métricas CD, NSE, VR e NRMSE utilizadas, tanto para a etapa de calibração, quanto para a de teste, como foi feito para o *dataset #1*.

Para as etapas de calibração dos três modelos trabalhados, perceberam-se valores relativamente altos para as métricas CD, NSE e VR, e mesmo o NRMSE não sendo tão próximo de 0 (valor ideal), ele pode ser considerado baixo para os três modelos. Os valores de CD e NSE ficaram próximos a 0,90, como percebe-se na Tabela 4 para os três tipos de modelo, mostrando uma proximidade muito grande do valor ideal, i.e., 1. Para o VR, os valores foram bem próximos a 0 para os três modelos, sendo inclusive de valores negativos, o que mostra uma baixa, mas existente, característica de subestimação dos dados na calibração. E finalmente, para o NRMSE, os valores giraram entre 0,256 (D2A), 0,265 (D2B) e 0,171 (D2C). Na Figura 23 é feita uma comparação gráfica entre os valores e vazão observados e os calculados pelo modelo na etapa de calibração para as três modelagens trabalhadas, percebendo-se visualmente uma boa compatibilidade entre os resultados, como foi notado para a etapa de calibração do *dataset #1*.

Tabela 4 – Métricas estatísticas da etapa de calibração e teste para o *dataset #2*.

Métrica/Modelo	D2A		D2B		D2C	
	Calibração	Teste	Calibração	Teste	Calibração	Teste
CD	0,930	0,689	0,916	0,749	0,964	0,805
NSE	0,929	0,689	0,916	0,749	0,964	0,805
VR	-0,015	-0,118	0,004	0,185	-0,005	-0,06
NRMSE	0,256	0,492	0,265	0,469	0,171	0,417

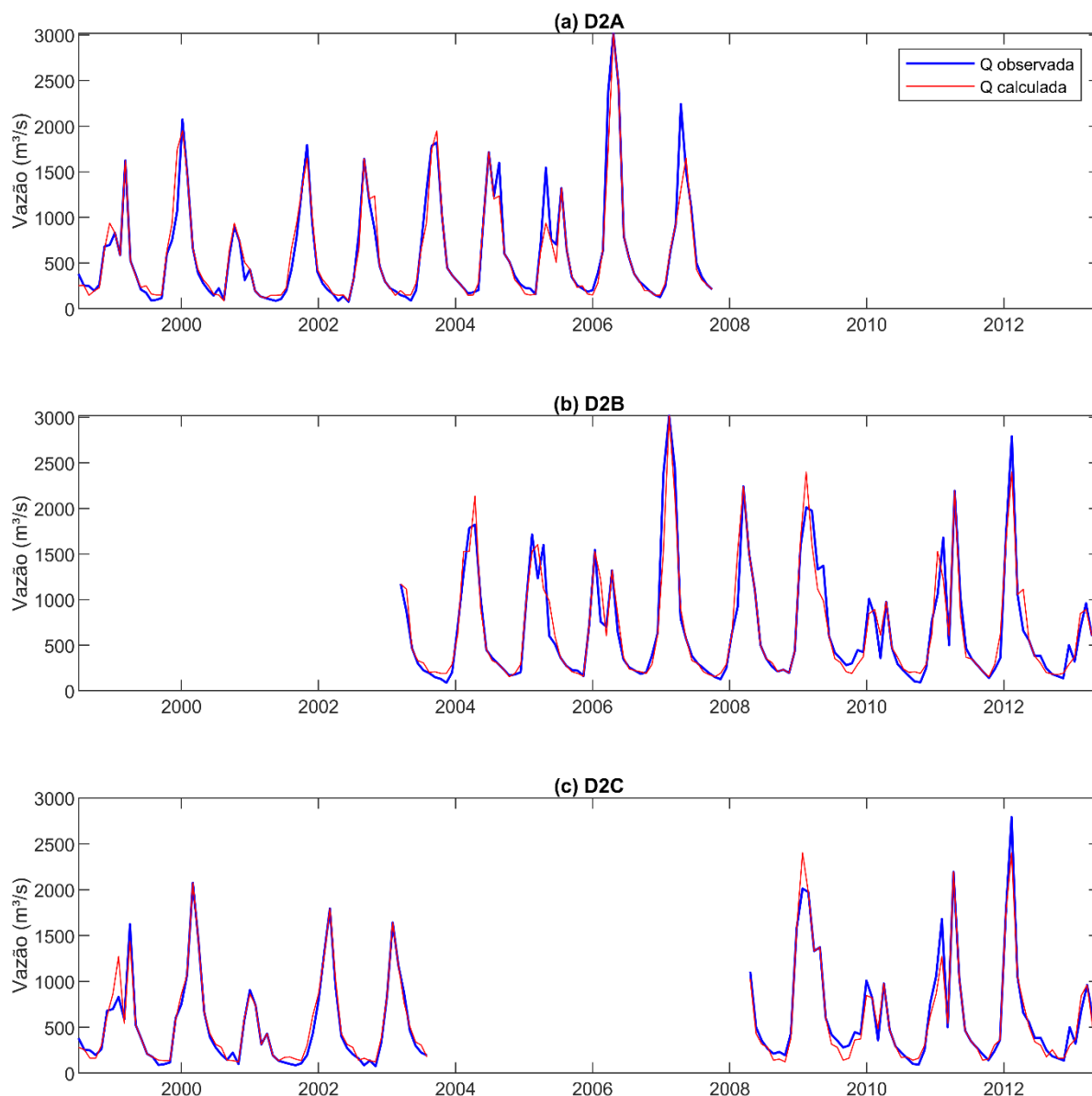


Figura 23 – Hidrograma de vazões observadas e calculadas para a etapa de calibração das três modelagens realizadas com o *dataset* #2.

Para a etapa de teste percebeu-se uma variação maior nos resultados obtidos. As métricas foram calculadas para os três modelos trabalhados, conforme a etapa de calibração. Para o modelo D2A, os valores das métricas encontradas foram os mais baixos encontrados para o *dataset* #2. O valor do CD e NSE encontrados foi de 0,689, o que se mostra um valor aceitável para as métricas (cujo valor ideal seria 1). O valor de VR encontrado foi de -0,118, o que significa que o modelo tendeu a subestimar (assim como na etapa de treino) os valores de vazão em cerca de 11,8% em relação aos valores observados. O valor do NRMSE encontrado foi de 0,492, o que deixa o modelo com um valor mediano para o parâmetro, que tem valor ideal 0. Na Figura 24 é mostrada uma comparação das séries observadas e calculadas pelo

modelo, e percebe-se visualmente o comportamento aceitável da modelo em calcular a série de vazões. Esse comportamento aceitável para o modelo, mesmo utilizando como valores de treino os 70% iniciais, mostra-se diferente do encontrado para o *dataset #1*, que apresentou os piores valores de métrica para esse tipo de modelagem.

Em parte, isso se deve ao fato de que diferentemente do *dataset #1*, o *dataset #2* possui um comportamento homogêneo dentro de seu intervalo temporal. Mesmo com esse do NRMSE, observando-se a Figura 24 percebe-se uma boa adequação entre as duas séries de vazões, que mesmo não estando exatamente sobrepostas, apresentam valores muito próximos, e que em termos práticos, do ponto de vista de obras hidráulicas, como discutido na seção anterior para o *dataset #1*.

Para o modelo D2B perceberam-se valores melhores para as métricas estatísticas empregadas na avaliação do modelo na etapa de teste. O valor do CD e do NSE encontrados foi de 0,749, representando um valor alto, e que implica numa boa adequação do modelo em calcular os valores de vazão trabalhados. O valor do VR foi de 0,182, o que implica que o modelo está superestimando os valores de vazão observados em cerca de 18,2%, o que não é relativamente muito alto, ao comparar com 0% (valor ideal). E finalmente, para este segundo modelo, encontrou-se um valor do NRMSE de 0,469, o que implica num valor quase mediano em relação ao valor ideal que seria de 0. Assim como observado para a modelagem anterior, percebe-se uma boa adequação entre as duas séries de vazões, que mesmo não estando exatamente sobrepostas, apresentam valores muito próximos.

Por fim, para o modelo D2C, percebeu-se uma boa performance par a maioria das métricas para o *dataset #2* na etapa de teste, que é mesmo comportamento observado para o *dataset #1*. Conforme percebe-se na Figura 24, é observada uma boa concordância entre as duas séries de dados de vazão, i.e., valores observados e calculados.

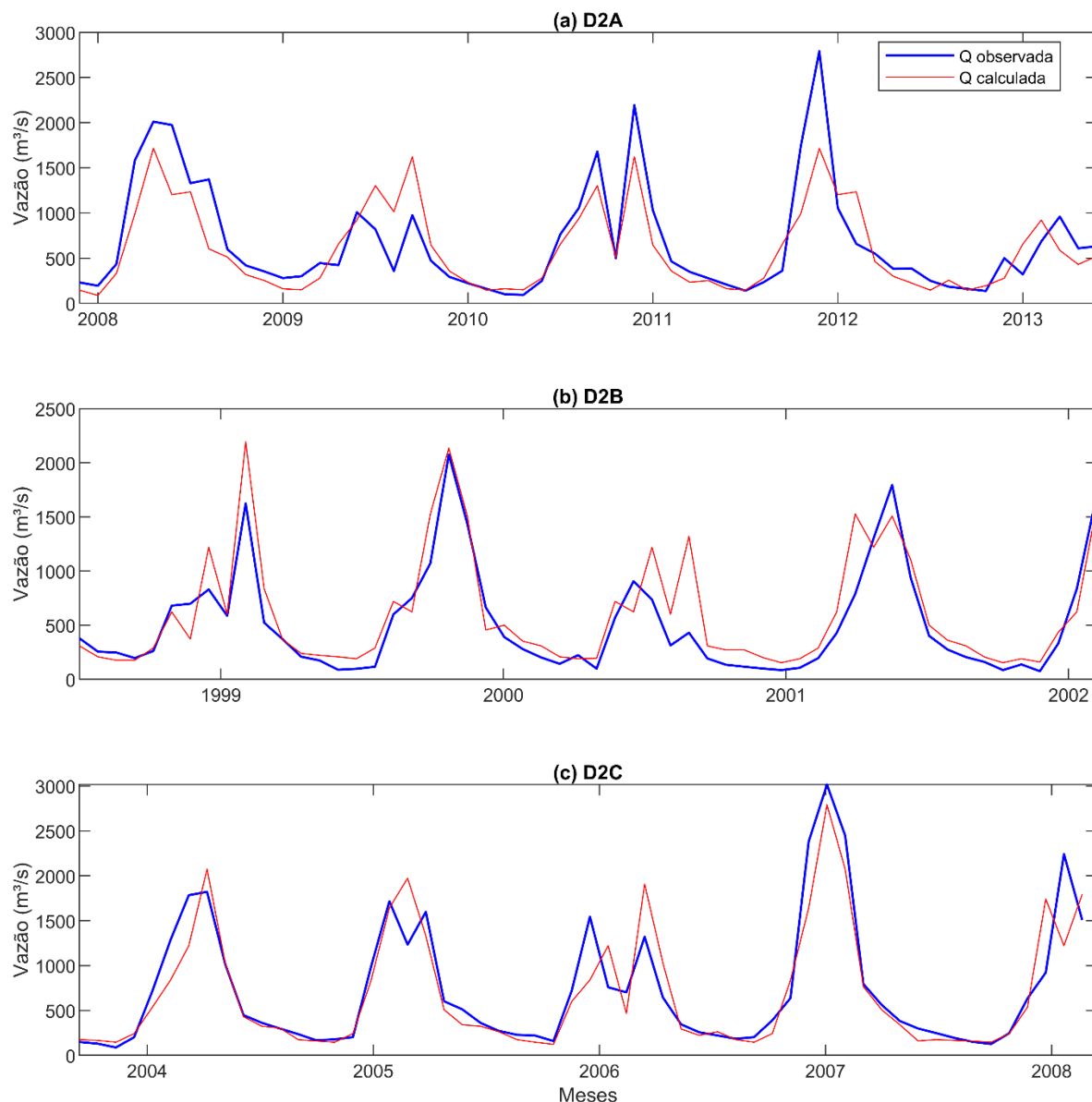


Figura 24 – Hidrograma de vazões observadas e calculadas para a etapa de teste das três modelagens realizadas com o *dataset #1*.

6.3 DATASET #3

6.3.1 Componentes do plano e modelagem SOM

E para o *dataset #3*, na Figura 25 é mostrada a rede SOM desenvolvida para cada uma das modelagens, tanto para a etapa de calibração, como o seu resultado para a etapa de teste. Percebe-se uma esperada equivalência entre as redes, com destaque para o neurônio vencedor de cada uma, e para a sua vizinhança. É interessante frisar também, assim como foi para os outros conjuntos de dados uma adequação da rede ao número de variáveis, observada pelo

preenchimento das variáveis em torno de todos os neurônios utilizados, não deixando neurônios inutilizados na rede.

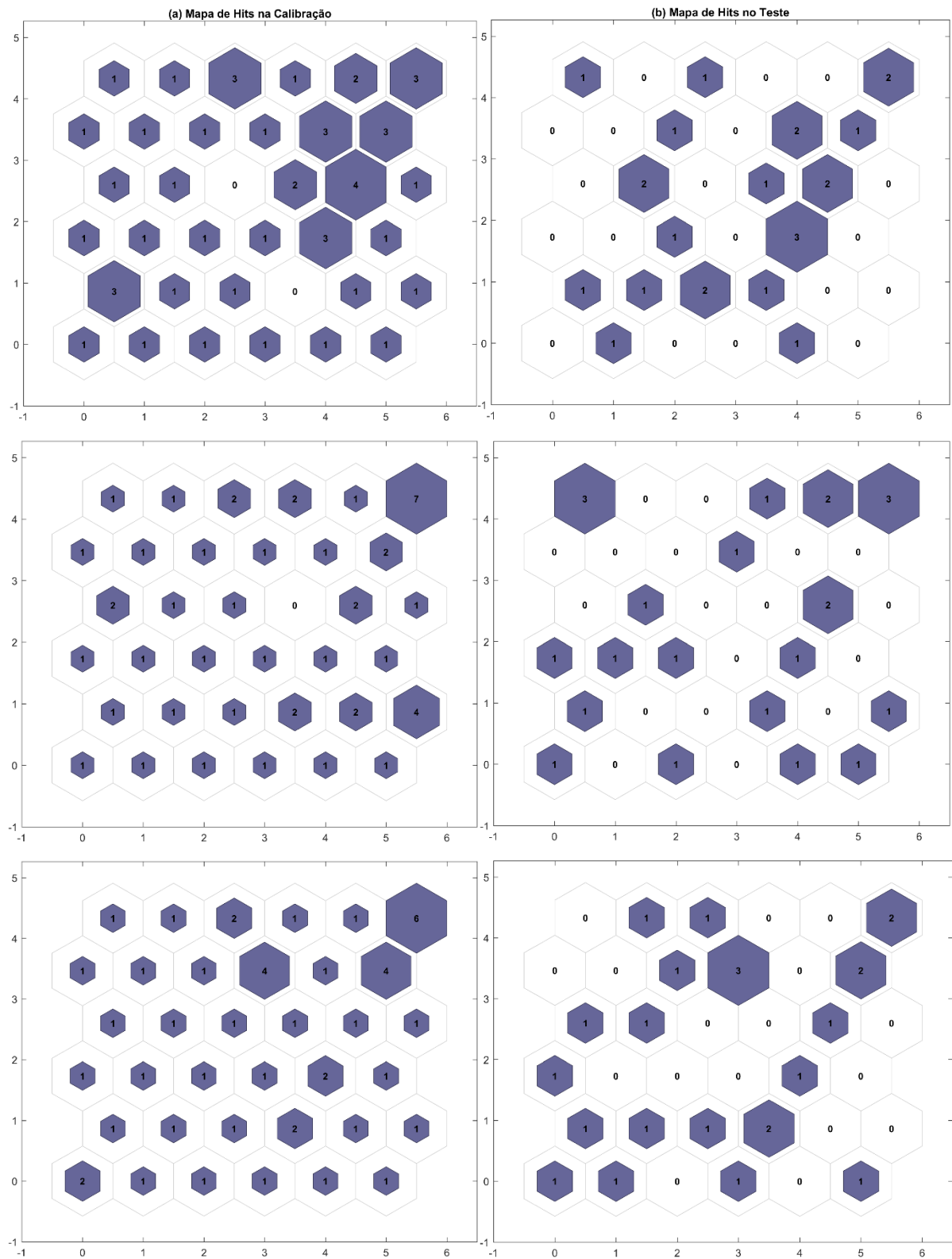


Figura 25 – Mapa de Hits das redes SOM para as etapas de calibração e teste para as três modelagens do *dataset* #3.

E dando continuidade, na figura 26 são apresentados os mapas de componentes do plano para as variáveis utilizadas nos modelos para a etapa de calibração. Para cada um dos três modelos foram selecionados os mapas de componentes para a grade TRMM com a maior correlação entre os dados de chuva, e assim, percebe-se de forma geral como semelhanças entre os dados de chuva usados no modelo e a vazão modelada. E partindo da Figura 26, como foi visto para os outros dois conjuntos de dados, é possível notar uma semelhança entre os dados de chuva do mês atual com a vazão apurada no próprio mês, e de forma menor, com o do mês anterior, evidenciando uma alta correlação entre as variáveis do mês de chuva e de vazão.

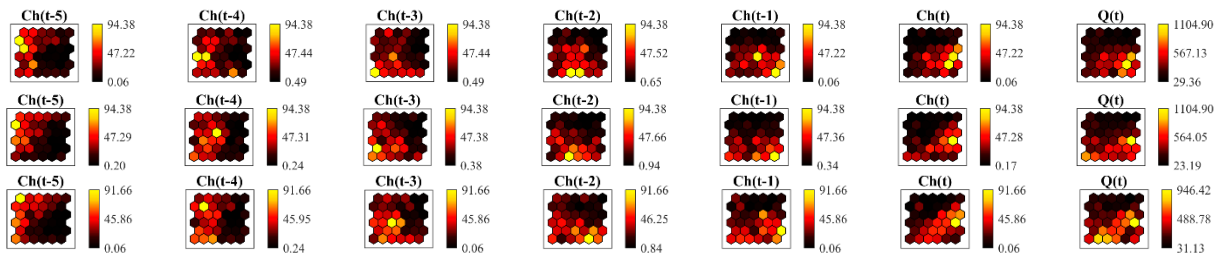


Figura 26 – Mapa dos componentes do plano da quadrícula 39 para as etapas de calibração dos três modelos desenvolvidos para o *dataset #3*.

6.3.2 Métricas estatísticas

Nesta seção são apresentados os resultados das modelagens para o terceiro conjunto de dados trabalhado, i.e., *dataset #3*, que varia de maio de 2013 até dezembro de 2019, para os três modelos utilizados, i.e., D3A, D2B e D3C. Na Tabela 5 são mostrados os resultados das métricas CD, NSE, VR e NRMSE utilizadas, tanto para a etapa de calibração, quanto para a de teste, como foi feito para os outros dois conjuntos de dados trabalhado.

Para as etapas de calibração dos três modelos trabalhados, perceberam-se valores relativamente altos para as métricas CD, NSE e VR, e mesmo o NRMSE, apresentou valores mais próximos de 0 (valor ideal) do que as calibrações dos outros três conjuntos de dados. As métricas da etapa de calibração das três modelagens do *dataset #3* apresentou os melhores valores, quando comparados com o *dataset #1* e *#2*. Os valores de CD e NSE foram acima de 0,96 para as três modelagens, chegando a 0,99 para o terceiro modelo, como percebe-se na Tabela 3 para os três tipos de modelo, mostrando uma proximidade muito grande do valor ideal, i.e., 1. Para o VR, os valores foram bem próximos a 0 para os três modelos (valor ideal). E finalmente, para o NRMSE, os valores foram relativamente baixos, chegando a 0,090 para a terceira modelagem. Na Figura 27 é feita uma comparação gráfica entre os valores e vazão observados e os calculados pelo modelo na etapa de calibração para as três modelagens

trabalhadas, percebendo-se visualmente uma boa compatibilidade entre os resultados, como foi notado para a etapa de calibração do *dataset #1* e *#2*.

Tabela 5 – Métricas estatísticas da etapa de calibração e teste para o *dataset #3*.

Métrica/Modelo	D3A		D3B		D3C	
	Calibração	Teste	Calibração	Teste	Calibração	Teste
CD	0,962	0,821	0,966	0,757	0,990	0,720
NSE	0,961	0,821	0,966	0,757	0,990	0,720
VR	0,003	-0,165	0,000	0,082	0,004	-0,047
NRMSE	0,196	0,335	0,169	0,472	0,090	0,561

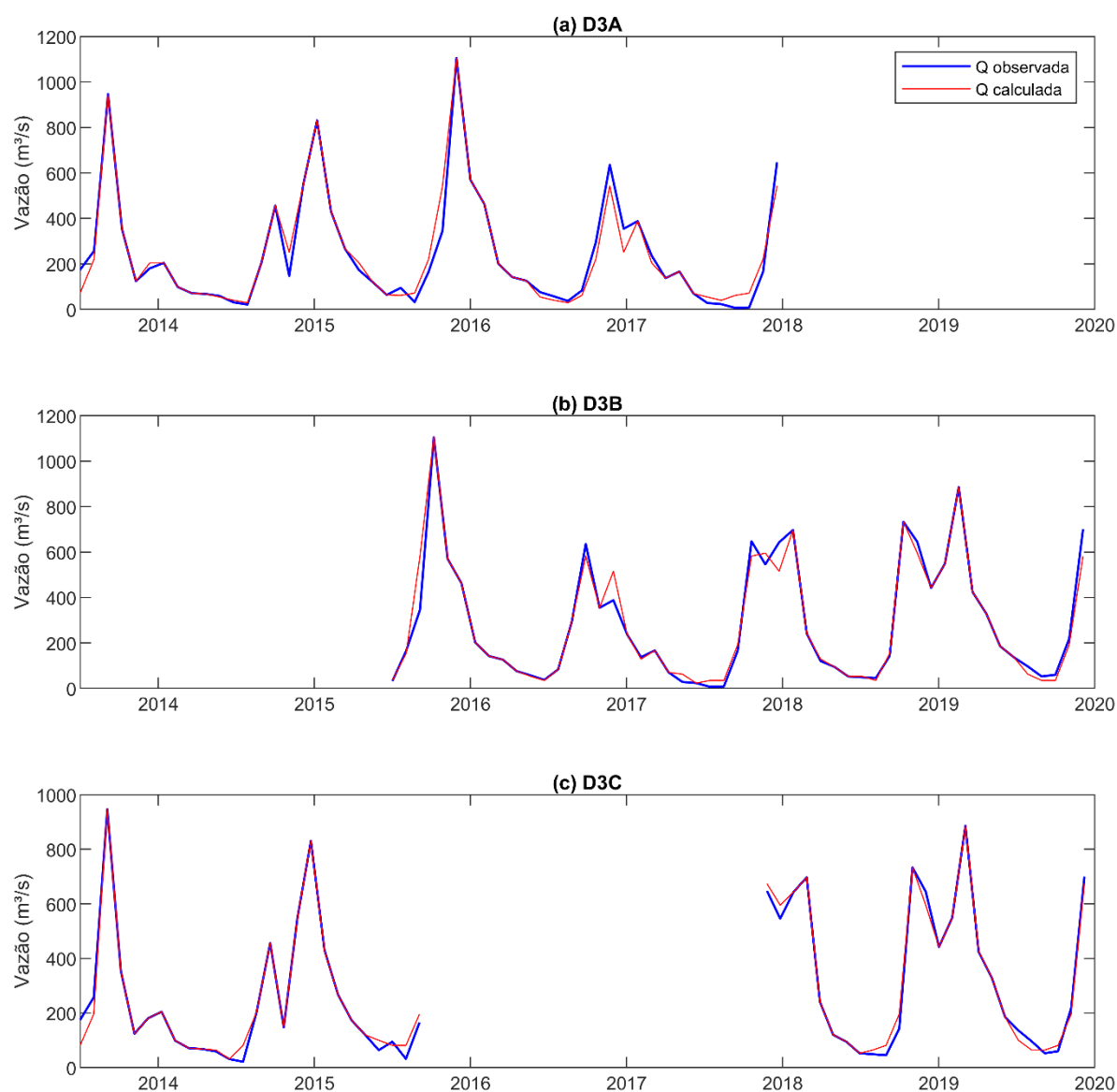


Figura 27 – Hidrograma de vazões observadas e calculadas para a etapa de calibração das três modelagens realizadas com o *dataset #3*.

Para a etapa de teste percebeu-se uma variação maior nos resultados obtidos, como foi notado para os outros dois conjuntos de dados discutidos anteriormente também. As métricas foram calculadas para os três modelos trabalhados, conforme a etapa de calibração. Para o primeiro modelo (D3A), o valor do CD e NSE encontrados foi de 0,821, o que se mostra um valor relativamente alto (cujo valor ideal seria 1). O valor de VR encontrado foi de -0,165, o que significa que o modelo tendeu a subestimar os valores de vazão em cerca de 16,5% em relação aos valores observados. O valor do NRMSE encontrado foi de 0,335, o que é um valor relativamente não alto.

Na Figura 28 é mostrada uma comparação das séries observadas e calculadas pelo modelo, e percebe-se visualmente um bom comportamento do modelo em calcular a série de vazões. Esse comportamento aceitável para o modelo, mesmo utilizando como valores de treino os 70% iniciais, mostra-se diferente do encontrado para o *dataset #1*, que apresentou os piores valores de métrica para esse tipo de modelagem, e a explicação tange os mesmos aspectos discutidos para o *dataset #2*, de que isso se deve ao fato de que diferentemente do *dataset #1*, os *dataset #2* e *#3* possuem um comportamento homogêneo dentro de sua série temporal. Percebe-se uma boa adequação entre as duas séries de vazões, que mesmo não estando exatamente sobrepostas, apresentam valores muito próximos, e que em termos práticos, do ponto de vista de obras hidráulicas, já se mostra, em grande parte, suficiente, uma vez que em diversos estudos trabalha-se com margem de erros, e o importante é o valor médio encontrado.

Para o segundo modelo (D3B), perceberam-se valores menores do que o anterior, mas de toda forma aceitáveis para as métricas estatísticas empregadas na avaliação do modelo na etapa de teste. O valor do CD e do NSE encontrados foi de 0,757, representando um valor alto, e que implica numa boa adequação do modelo em calcular os valores de vazão trabalhados. O valor do VR foi de 0,082, o que implica que o modelo está superestimando os valores de vazão observados em cerca de 8,2%, o que não é relativamente muito alto, ao comparar com 0%, que é o valor ideal. E finalmente, para este segundo modelo, encontrou-se um valor do NRMSE de 0,472, o que implica num valor quase mediano em relação ao valor ideal que seria de 0.

Por fim, para o terceiro modelo (D3C), percebeu-se boa performance para todas as métricas utilizadas no *dataset #3* na etapa de teste, que é mesmo comportamento observado para os *dataset #1* e *dataset #2*. Conforme percebe-se na Figura 28, é observada uma boa concordância entre as duas séries de dados de vazão, i.e., valores observados e calculados.

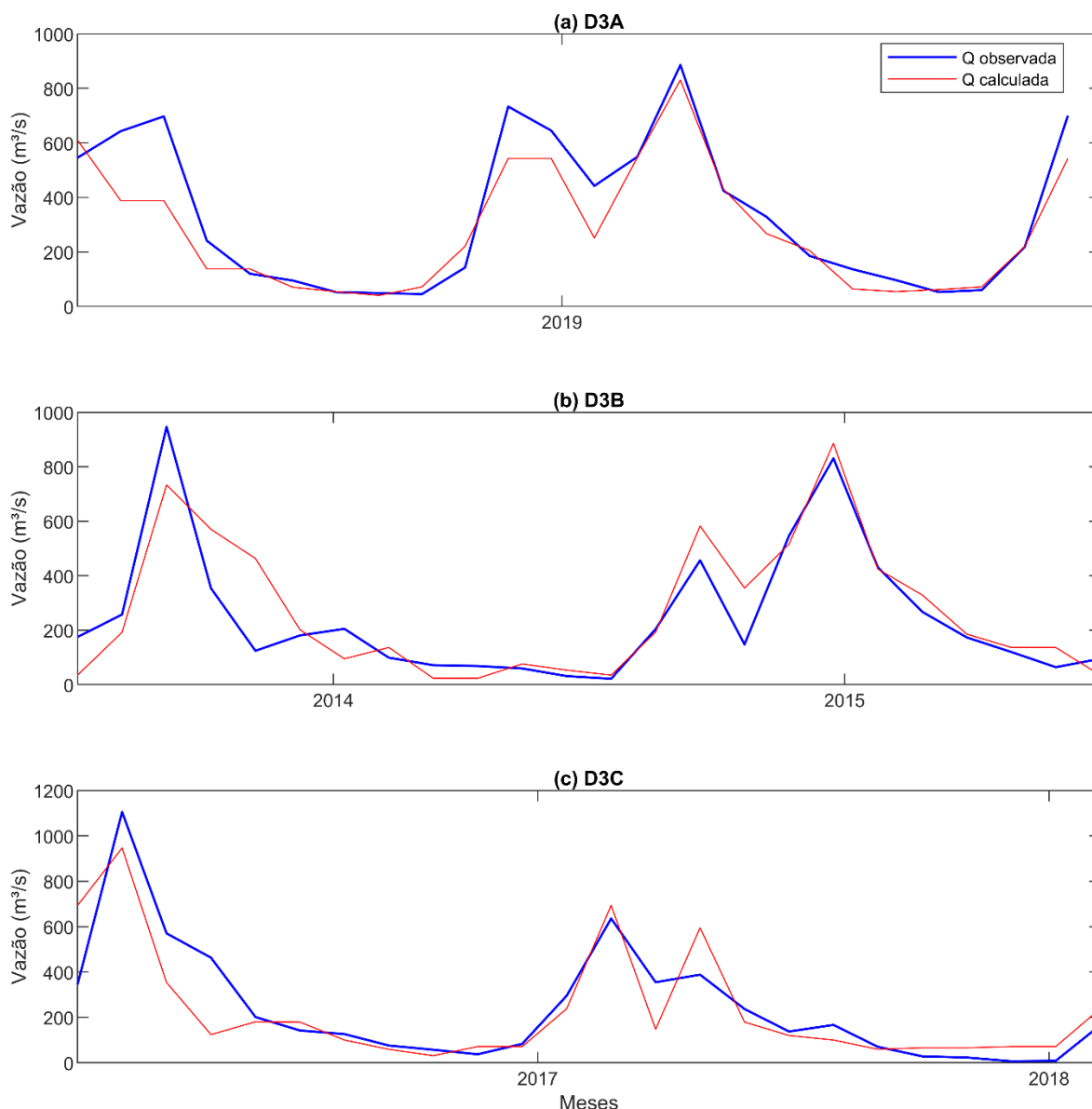


Figura 28 – Hidrograma de vazões observadas e calculadas para a etapa de teste das três modelagens realizadas com o *dataset* #3.

6.4 DISCUSSÕES DOS RESULTADOS

Moriasi *et al.* (2015) fizeram uma classificação para os valores das métricas CD, NSE e VR, que são amplamente utilizadas na área de recursos hídricos (Tabela 6). Na classificação, considerando a escala mensal, os valores de CD são considerados “*muito bom*” acima de 0,85, “*bom*” entre 0,75 e 0,85, “*satisfatório*” entre 0,60 e 0,75 e “*não satisfatório*” aqueles abaixo de 0,60. Para o NSE, os valores são considerados “*muito bom*” por estarem acima de 0,80, “*bom*” entre 0,70 e 0,80, “*satisfatório*” entre 0,50 e 0,70 e “*não satisfatório*” aqueles abaixo de 0,50. E por fim, para o VR, os valores são considerados “*muito bom*” quando abaixo de + -

5%, “*bom*” entre 5 e 10%, “*satisfatório*” entre 10 e 15% e “*não satisfatório*” aqueles maiores que 15%. Assim, ao comparar com os resultados do presente trabalho, percebe-se que em geral as métricas dos modelos do presente trabalho apresentam resultados concentrados entre as faixas de “*muito bom*” a “*bom*”. Dentro os modelos trabalhados, apenas o D1A apresentou todas as três métricas (CD, NSE e VR) com valores classificados como “*não satisfatório*”.

Tabela 6 - Performance das métricas CD, NSE e VR (Fonte: Adaptado, Moriasi *et al.*, 2015).

Metrica/Performace	Muito bom	Bom	Satisfatório	Não satisfatório
CD	$CD > 0.85$	$0.75 \leq CD \leq 0.85$	$0.60 \leq CD \leq 0.75$	$CD \leq 0.60$
NSE	$NSE > 0.80$	$0.70 \leq NSE \leq 0.80$	$0.50 \leq NSE \leq 0.70$	$NSE \leq 0.50$
VR	$VR < \pm 5$	$\pm 5 \leq VR \leq \pm 10$	$\pm 10 \leq VR \leq \pm 15$	$VR \geq \pm 15$

É interessante destacar que ao partir para uma análise do comportamento do modelo D1A, infere-se que como já foi notado nas seções anteriores, o mal comportamento da modelagem se deveu ao fato de que além de trabalhar-se com uma série não homogênea em sua totalidade, fazer uso de uma etapa de calibração e de teste onde os dois comportamentos hidrológicos são totalmente distintos. O *dataset #1* possui uma mudança no comportamento hidrológico da vazão a partir do mês de abril de 2013, conforme é notado pelo teste de *Pettitt*, discutido na seção de materiais utilizados, e conseqüentemente, ao utilizar-se 70% dos primeiros dados como treino, acabou-se por não utilizar dados do segundo regime hidrológico na calibragem, fazendo o modelo falho na previsão de um regime novo, e não antes ensinado na calibração.

Ao inverter-se a ordem dos dados utilizados na calibração, escolhendo-se 70% finais da série completa de vazões do *dataset #1*, ou até mesmo utilizando os 70% dos dados divididos em início e fim da série para calibração, percebe-se uma imediata melhora nos valores das métricas estatísticas calculadas para os modelos, deixando claro que as redes SOM se mostraram capazes de trabalhar com séries não homogêneas para a modelagem chuva-vazão, contanto que obviamente, na etapa de treino, se apresente todos os comportamentos hidrológicos que a série completa venha a apresentar.

Além disso, para os três subconjuntos usados, todos os que usaram na modelagem os dados de calibração divididos pela metade no início da série e a outra metade no final, ou seja, os modelos que usaram 30% do meio da série de dados como dados de teste (D1C, D2C e D3C), apresentaram valores na faixa de classificação de “*muito bom*” segundo Moriasi *et al.* (2015),

além de métricas com melhores resultados em relação aos outros dois modelos aplicados para cada subconjunto. As métricas encontradas mostram altos valores de CD e NSE, bem como baixos valores de VR e NRMSE, mostrando uma boa concordância entre a série calculada e a prevista. Esse comportamento deve ser destacado como uma alternativa ao tipo de modelagem no cálculo e preenchimento de séries com dados faltantes, o que é bastante comum em análises hidrológicas em regiões hidrológicas que não possuem uma instrumentação totalmente eficiente.

7 CONCLUSÕES E RECOMENDAÇÕES

Diante do estudo realizado é demonstrada a possibilidade de utilização de produtos de precipitação TRMM na modelagem chuva-vazão através de redes SOM a nível mensal de forma eficiente e prática. Do total de modelagens realizadas no presente trabalho, apenas uma, i.e., a que fez uso dos 70% dos dados iniciais do *dataset #1*, apresentou métricas estatísticas com valores considerados insatisfatórios pela literatura, e a causa pode ser atribuída a escolha de dados não representativos para a etapa de treino da série, que coincidentemente ocorreu devido a característica de não homogeneidade da série completa, que apresenta dois períodos hidrológicos distintos dentro do conjunto total.

O estudo demonstrou também a possibilidade de se trabalhar com séries não homogêneas na modelagem chuva-vazão pelas redes SOM, como foi observado para as duas demais modelagens feitas com o *dataset #1*, que ao fornecer dados mais representativos para etapa de treinamento da rede, os resultados na etapa de teste foram bem mais satisfatórios. É interessante notar que trabalhou-se com uma série relativamente pequena do ponto de visto de análises hidrológicas (apenas cerca de 21 anos de dados), e caso se trabalhasse com uma série de dados maior, é provável que se possuísse uma variabilidade maior no comportamento dos dados, o que transpassaria naturalmente essa limitação encontrada. Assim, a utilização de dados o mais representativos possível do conjunto total de dados é o ideal a se proceder.

O trabalho também demonstrou a possibilidade de uma boa modelagem chuva-vazão a nível mensal, fazendo uso de produtos de precipitação TRMM, mesmo ao se trabalhar com séries temporais curtas, como foi o caso do *dataset #3*, que possui apenas sete anos de dados, porém apresentou ótimos valores de métricas estatísticas, tanto para a calibração, quanto para o teste, nos três tipos de modelagens empregadas.

Além disso, destaca-se o fato de que de maneira geral foram encontradas boas modelagens para os três conjuntos de dados que empregaram os dados de calibração divididos com 35% no início e 35% no final, ou seja, o teste foi feito para os 30% valores centrais da série temporal (DC e NSE maior do que 0.95 para calibração, e maior do que 0.70 para teste). Como discutido anteriormente, esse fato deve ser apresentado como uma alternativa no cálculo e preenchimento de séries com dados faltantes, o que é um dos problemas mais recorrentes no estudo de séries temporais na hidrologia, sobretudo em regiões hidrológicas que não possuem uma instrumentação totalmente eficiente. Além disso, por tratar-se de um preenchimento de falhas ou até mesmo cálculo para o mês atual da vazão, a quantidade de meses anteriores

utilizada não atrapalha o processo, uma vez que são dados já disponíveis de precipitação, e por tratar-se de produtos de satélite a sua disponibilidade não mostra-se um problema, como o é na maioria das situações para produtos de pluviômetros.

Portanto, destaca-se que a modelagem chuva-vazão através de redes SOM se mostrou eficiente e prática, podendo ser útil no gerenciamento de recursos hídricos, que é um recurso tão importante para o desenvolvimento humano em todo o mundo. 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 nessas áreas. Além disso, a utilização de dados do satélite TRMM surgem 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 muitas situações. Por exemplo, no caso de bacias não aparelhadas e localizadas em áreas remotas, por exemplo. Por fim, recomenda-se o desenvolvimento de novos estudos aplicando a metodologia em outras regiões hidrológicas distintas, a fim de validar com novos estudos a metodologia de utilização de dados de satélite, na modelagem chuva-vazão através de redes SOM, tanto a nível mensal, quanto diário.

8 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. doi: 10.1029/2011WR010690
- ADELOYE, A. J. & RUSTUM, R. (2011). SOM and rainfall-runoff modelling in inadequately gauged basins. *Hydrology Research*, 43.5, 2012. doi: 10.2166/nh.2012.017
- 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 maio de 2020.
- 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. doi: 10.1007/s00521-018-3916-0
- BEALE, M.; HAGAN, M.; DEMUTH, H. (2012). *Neural Network toolbox 7.0.3: User's Guide*. Natick: The MathWorks Inc.
- 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
- CAMPOLO, M.; ANDREUSSI, P.; SOLDATI, A. (1999). River flood forecasting with a neural network model. *Water Resources Research*, 35(4), 1191–1197. doi: 10.1029/1998WR900086
- CAO Y., ZHANG W. E WANG W. (2018). Evaluation of TRMM 3B43 data over the Yangtze River Delta of China. *Scientific Reports*, 8(1), 5290–5301. doi: 10.1038/s41598-018-23603-z

- 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. doi: 10.1016/j.jhydrol.2010.01.016
- 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, 2184–2196. doi: 10.1002/jgrd.50250
- CIĞIZOĞLU, H.K.; ALP, M. (2006). Generalized regression neural network in modelling river sediment yield. *Advances in Engineering Software*, 37 (2), 63–68. doi: 10.1016/j.advengsoft.2005.05.002
- CIGIZOGLU, H. K.; KISI, Ö. (2005). Flow prediction by three back propagation techniques using k-fold partitioning of neural network training data. *Nordic Hydrology*, 36(1), 1–16. doi: 10.2166/nh.2005.0005
- 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.; ARAVENA, R.; BOBDE, B. (2001). Artificial neural network modeling of water table depth fluctuations. *Water Resources Research*, 37 (4), 885–896. doi: 10.1029/2000WR900368
- 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, 292–304. doi: 10.1016/j.jhydrol.2015.04.012

- DEMUTH, H.; BEALE, M. (2005). Neural network toolbox for use with Matlab. The MathWorks, 469 Inc., Natick, USA.
- FARHANGI, M., KHOLGHI, M.; CHAVOSHIAN, S. A. (2016). Rainfall Trend Analysis of Hydrological Subbasins in Western Iran. *Journal of Irrigation and Drainage Engineering*, 142(10), 1–11. doi: 10.1061/(ASCE)IR.1943-4774.0001040
- 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. doi: 10.1186/s40677-015-0022-9
- 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.; 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*, v. 14, p. 1242-1250. doi: 10.1007/s11368-013-0841-9
- 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*, v. 7, p. 176-182. doi: 10.4090/juee.2013.v7n1.176-182
- 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.
- FILHO, J. S.; FARIAS, C. (2018). Stochastic modeling of monthly river flows by Self-Organizing Maps. *Journal of Urban and Environmental Engineering*, v. 12, p. 219-230. doi: 10.4090/juee.2018.v12n2.219-230
- FREIRE, P. K. M. M. O uso da transformada Wavelet na previsão de vazão. 130 f. Dissertação (Mestrado em Engenharia Urbana e Ambiental) – Programa de Pós-Graduação em Engenharia Urbana e Ambiental, Universidade Federal da Paraíba, João Pessoa.

- FREIWAN, M.; CIGIZOGLU, H.K. (2005). Prediction of total monthly rainfall in Jordan using feed forward backpropagation method. *Fresenius Environmental Bulletin*. 14(2), 142–151.
- FUNK, C.; PETERSON, P.; LANDSFELD, M.; PEDREROS, D.; VERDIN, J.; SHUKLA, S.; MICHAELSEN, J. (2015). The climate hazards infrared precipitation with stations – A new environmental record for monitoring extremes. *Scientific Data*, 2, 150066. doi: 10.1038/sdata.2015.66
- 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. doi: 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. doi: 10.1016/j.engappai.2004.03.004
- GAZZAZ, N.M.; YUSOFF, M. K.; ARIS, A. Z.; HUAHIR, H.; RAMLI, M. F. (2012). Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. *Marine Pollution Bulletin*, 64 (11), 2409–2420. doi: 10.1016/j.marpolbul.2012.08.005
- 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*, ISSN 1998-9563 *Hydrology Research*, v. 48, p. 455-467.
- Hall, T.; Brooks, H.E.; Doswell III, C.A. (1999). Precipitation forecasting using a neural network. *Weather and Forecasting*, 14(3), 338–345. doi: 10.1175/1520-0434(1999)014<0338:PFUANN>2.0.CO;2
- 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. doi: 10.1080/02626667.2018.1552788
- HORNBERGER, G. M. (1998). *Elements of Physical Hydrology*. The John Hopkins University Press, 302 pp.

- HOU, A.Y.; KAKAR, R. K.; NEECK, S.; AZARBARZIN, A. A.; KUMMEROW, C. D.; KOJIMA, M.; IGUCHI, T. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, 95(5), 701–722. doi: 10.1175/bams-d-13-00164.1
- HSU, K.; GUPTA, H.V.; SOROOSHIAN, S. (1995). Artificial neural network modelling of the rainfall runoff process. *Water Resources Research*, 31(10), 2517–2530. doi: 10.1029/95WR01955
- HUFFMAN, G. J., BOLVIN, D. T.; NELKIN, E. J.; WOLFF, D. B.; ADLER, R. F.; GU, G.; STOCKER, E. F. (2007). The TRMM multi-satellite precipitation analysis (TMPA): Quasi-global, multiyear, combined sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8(1), 38–55. doi:10.1175/JHM560.1
- 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
- ISMAIL, S.; SHABRI, A.; SAMSUDIN, R. (2012). A hybrid model of self organizing maps and least square support vector machine for river flow forecasting. *Hydrological earth system science*, v. 16, p. 4417-4433. doi: 10.5194/hess-16-4417-2012
- JAIN, A.; SUDHEER, K.P.; SRINIVASULU, S. (2004). Identification of physical processes inherent in artificial neural network rainfall runoff models. *Hydrological Processes*, 18(3), 571–581. doi: 10.1002/hyp.5502
- JEONG, D.; KIM, Y. O. (2005) Rainfall-runoff models using artificial neural networks for ensemble streamflow prediction, *Hydrol. Process.*, 19, 3819–3835.
- Jiang, S.; Ren, L.; Zhou, M.; Young, B.; Zhang, Y; Ma, M. (2017). Drought monitoring and reliability evaluation of the latest TMPA precipitation data in the Weihe River Basin, Northwest China. *Journal of Arid Land*, 9(2), 256–269. doi: 10.1007/s40333-017-0007-5
- 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.
- KELLER FILHO, T.; ASSAD, E. D.; LIMA, P. R. S. R. (2005). Regiões pluviometricamente homogêneas no Brasil. *Pesquisa Agropecuária Brasileira*, 40(4), 311–322. doi: 10.1590/S0100-204X2005000400001

- KISI, Ö. (2004). River flow modeling using artificial neural networks. *J. Hydrologic Engineering*, 9(1), 60–63. doi: 10.1061/(ASCE)1084-0699 (2004) 9:1(60)
- 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.
- KUMAR, D. N.; RAJU, K. S.; SATHISH, T. (2004) River flow forecasting using recurrent neural network. *Water Resources Management*, 18(2).
- KUMAR, M.; RAGHUWANSHI, N. S.; SINGH, R.; WALLENDER, W. W.; PRUITT, W. O. (2002). Estimating evapotranspiration using artificial neural network. *Journal of Irrigation and Drainage Engineering*, 128 (4), 224–233. doi:10.1061/(ASCE)0733-9437(2002)128:4(224)
- KUMMEROW, C.; BARNES, W.; KOZU, T.; SHIUE, J.; SIMPSON, J. (1998). The tropical rainfall measuring mission (TRMM) sensor package. *Journal of Atmospheric and Oceanic Technology*, 15(3), 809–817. doi: 10.1175/1520-0426(1998)015<0809:TTRMMT>2.0.CO;2
- 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.*
- LE, M. H.; LAKSHMI, V.; BOLTEN, J.; BUI, D. D. (2020). Adequacy of Satellite-derived Precipitation Estimate for Hydrological Modeling in Vietnam Basins. *Journal of Hydrology*, 586, 124820. doi: 10.1016/j.jhydrol.2020.124820
- LIAO, R. W.; ZHANG D. B.; SHEN, Y. (2015). Validation of six satellite-derived rainfall estimates over China. *Meteorological Monthly*, 41, 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.
- 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(14), 1–14. doi: 10.1117/1.JRS.8.083589

- MANNAN, A.; CHAUDHARY, S.; DHANYA, C. T.; SWAMY, A. K. (2018). Regionalization of rainfall characteristics in India incorporating climatic variables and using self-organizing maps. *ISH Journal of Hydraulic Engineering*, v. 24, n.2, p. 147-156. doi: 10.1080/09715010.2017.1400409
- MELESSE, A. M.; AHMAD, S.; MCCLAIN, M. E.; WANG, X.; LIM, Y. H. (2011). Suspended sediment load prediction of river systems: an artificial neural network approach. *Agricultural Water Management*, 98 (5), 855–866. doi: 10.1016/j.agwat.2010.12.012
- 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. doi: 10.1007/s00521-018-3519-9
- 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.
- MOHANTY, S.; JHA, M. K.; KUMAR, A.; SUDHEER, K. P. (2010). Artificial Neural Network modeling for groundwater level forecasting in a river island of eastern India. *Water Resources Management*, 24 (9), 1845–1865. doi:10.1007/s11269-009-9527-x
- MORIASI, D. N.; GITAU, M.W.; PAI, N.; DAGGUPATI, P. (2015). Hydrologic and water quality models: performance measures and evaluation criteria. *Transactions of the ASABE (American Society of Agricultural and Biological Engineers)* 58(6):1763-1785. doi: 10.13031/trans.58.10715
- 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. doi: 10.13031/2013.23153
- MUHAMMAD, W., YANG, H., LEI, H., MUHAMMAD, A., YANG, D. (2018). Improving the regional applicability of satellite precipitation products by ensemble algorithm. *Remote Sens.* 10, 577. doi: 10.3390/rs10040577
- NAGY, H. M.; WATANABE, K. A. N. D.; HIRANO, M. (2002). Prediction of sediment load concentration in rivers using artificial neural network model. *Journal of Hydraulic Engineering*, 128 (6), 588–595. doi: 10.1061/(ASCE)0733-9429(2002)128:6(588)

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. doi: 10.1016/0022-1694(70)90255-6

NAYAK, P. C.; RAO, Y. R. S.; SUDHEER, K. P. (2006). Groundwater level forecasting in a shallow aquifer using Artificial Neural Network approach. *Water Resources Management*, 20 (1), 77 90. doi: 10.1007/s11269-006-4007-z

NETO, R. M. B. (2020) Avaliação da seca e de suas características sobre o estado da Paraíba utilizando dados TRMM e de pluviômetros (1998–2017). 190 f. Dissertação (Mestrado em Recursos Hídricos) – Programa de Pós-Graduação em Engenharia Urbana e Ambiental, Universidade Federal da Paraíba, João Pessoa.

NING, S. W.; SONG, F.; UDMALE, P.; JIN, J. L.; THAPA, B. R.; ISHIDAIRA, H. (2017). Error analysis and evaluation of the latest GSDMap and IMERG precipitation products over Eastern China. *Advances in Meteorology*, Vol. 2017, Article ID 1803492. 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. doi: 10.1016/j.jhydrol.2012.10.054

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*, v. 13, p. 34-41. doi: 10.4090/juee.2019.v13n1.034041

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, doi: 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.
- PENG, F.; ZHAO, S.; CHEN, C.; CONG, D.; WANG, Y.; OUYANG, H. (2020). Evaluation and comparison of the precipitation detection ability of multiple satellite products in a typical agriculture area of China. *Atmospheric Research*, 236, 104814. doi: 10.1016/j.atmosres.2019.104814
- PEREIRA, G.; SILVA, M. E. S.; MORAES, E. C.; CARDOZO, F. S. (2013). Avaliação dos Dados de Precipitação Estimados pelo Satélite TRMM para o Brasil. *Revista Brasileira de Recursos Hídricos*, 18(3), 139–148. doi: 10.21168/rbrh.v18n3.p139-148
- PETTITT, A. N. A Non-Parametric Approach to the Change-Point Problem. *Applied Statistics*, v. 28, n. 2, p. 126-135, 1979.
- PLOUFFE, C. C. F.; ROBERTSON, C.; CHANDRAPALA, L. (2015). Comparing interpolation techniques for monthly rainfall mapping using multiple evaluation criteria and auxiliary data sources: A case study of Sri Lanka. *Environ. Modell. Software* 67: 57–71. doi: 10.1016/j.envsoft.2015.01.011.
- 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 Agronômico de Campinas. 72 pp.
- RAO, V. B.; FRANCHITO, S. H.; SANTO, C. M. E.; GAN, M. A. (2015). An update on the rainfall characteristics of Brazil: seasonal variations and trends in 1979–2011. *International Journal of Climatology*, 36(1), 291–302. doi: 10.1002/joc.4345
- REBOITA, M.S.; GAN, M. A.; ROCHA, R. P.; AMBRIZZI, T. (2010). Regimes de precipitação na América do Sul: Uma revisão bibliográfica. *Revista Brasileira de Meteorologia*, 25(2), 185– 204. doi: 10.1590/S0102-77862010000200004
- RIVIERA, D.; LILLO, M.; UVO, C. B.; BILLIB, M.; ARUMI, J. L. (2011). Forecasting monthly precipitation in central Chile: a self-organizing map approach using filtered sea surface temperature. *Theoretical and Applied Climatology* 107 (1–2), 1–13. doi: 10.1007/s00704-011-0453-5

- ROZANTE, J. R.; VILA, D. A.; CHIQUELTO, J. B.; FERNANDES, A. A.; ALVIM, D. S. (2018). Evaluation of TRMM/GPM Blended Daily Products over Brazil. *Remote Sensing*, 10(6), 882–898. doi: 10.3390/rs10060882
- 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. doi: 10.5194/piahs-364-100-2014, 2014
- 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.; MISHRA, S. K. (2012a) Cuckoo search via Lévy flights for optimization of a physically-based runoff-erosion model. *Journal of Urban Environmental Engineering*, 6(2), 123–131. doi: 10.4090/juee.2012.v6n2.%25p
- 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.
- 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. doi: 10.1080/02626667.2013.778412
- SANTOS, C. A. G.; SILVA, G. B. L. (2014). Daily streamflow forecasting using a wavelet transform and artificial neural network hybrid models. *Hydrological Sciences Journal*, 59(2), 312–324. doi: 10.1080/02626667.2013.800944
- SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. (2010). *Redes Neurais Artificiais para engenharia e ciências aplicadas*, Artliber.
- 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. doi: 10.1016/j.gecco.2019.e00811
- 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). doi: 10.2166/nh.2018.222.

- SILVERMAN, D.; DRACUP, J.A. (2000). Artificial neural networks and long-range precipitation prediction in California. *J. Applied Meteorology*, 39(1), 57–66, 2000. doi: 10.1175/1520 0450(2000)039<0057:ANNALR>2.0.CO;2
- Singh, K. P.; ANKITA, B.; AMRITA. M; GUNJA, J. (2009). Artificial neural network modeling of the river water quality – a case study. *Ecological Modelling*, 220 (6), 888–895. doi: 10.1016/j.ecolmodel.2009.01.004
- SOARES, A. S. D.; DA PAZ, A. R; PICCILLI, D. G. A. (2016). Avaliação das estimativas de chuva do satélite TRMM no Estado da Paraíba. *Revista Brasileira de Recursos Hídricos*, 21(2), 288-299. doi: 10.21168/rbrh.v21n2.p288-299
- SOROOSHIAN, S.; HSU, K. L.; GAO, X.; GUPTA, H. V.; IMAM, B.; DAN, B. (2000). Evaluation of PERSIANN system satellite–based estimates of tropical rainfall. *Bulletin of the American Meteorological Society*, 81(9), 2035–2046. doi: 10.1175/15200477(2000)081<2035:EOPSSE>2.3.CO;2
- 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, 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. doi: 10.1016/j.foodchem.2006.04.042
- 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
- TENG, H.; ROSSEL, R. A. V.; SHI, Z.; BEHRENS, T.; CHAPPELL, A.; E. BUI. (2016). “Assimilating satellite imagery and visible–near infrared spectroscopy to model and map soil loss by water erosion in Australia.” *Environ. Modell. Software* 77: 156–167. doi: 10.1016/j.envsoft.2015.11.024.

- 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. doi: 10.1016/j.jhydrol.2018.05.045
- TRAJKOVIC, S.; TODOROVIC, B.; STANKOVIC, M. (2003). Forecasting of reference evapotranspiration by artificial neural networks. *Journal of Irrigation and Drainage Engineering*, 129 (6), 454–457. doi:10.1061/(ASCE)0733-9437(2003)129:6(454)
- 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
- VILELLA, S.M.; MATTOS, A. (1975). *Hidrologia Aplicada*, McGraw-Hill, São Paulo – SP.
- VILJANEN, N.; HONKAVAARA, E.; NÄSI, R.; HAKALA, T.; NIEMELÄINEN, O.; KAIIVOSOJA, J. (2018) A Novel Machine Learning Method for Estimating Biomass of Grass Swards Using a Photogrammetric Canopy Height Model, Images and Vegetation Indices Captured by a Drone. *Agriculture*. 8. 70. doi:10.3390/agriculture8050070
- 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. 418(36). doi: 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.
- ZANETTI, S. S.; SOUSA, E. F.; OLIVEIRA, V. P. S.; ALMEIDA, F. T.; BERNARDO, S. (2007). Estimating evapotranspiration using artificial neural network and minimum climatological data. *Journal of Irrigation and Drainage Engineering*, 133 (2), 83–89. doi: 10.1061/(ASCE)0733-9437(2007)133:2(83)
- ZENG, Q. L.; WANG, Y. Q.; CHEN, L. F.; WANG, Z. F.; ZHU, H.; LI, B. (2018). Intercomparison and evaluation of remote sensing precipitation products over China from 2005 to 2013. *Remote Sensing*, 10(2), 168. doi: 10.3390/rs10020168