

Universidade Federal do Paraná

Isadora Chiamulera

**EMPREGO DE REDES NEURAIS ARTIFICIAIS PARA A  
MODELAGEM CHUVA-VAZÃO DA BACIA DO RIO DAS  
MORTES**

CURITIBA  
2014

Isadora Chiamulera

**EMPREGO DE REDES NEURAIS ARTIFICIAIS PARA A  
MODELAGEM CHUVA-VAZÃO DA BACIA DO RIO DAS  
MORTES**

Projeto Final II apresentado ao Curso de Engenharia Ambiental da Universidade Federal do Paraná como parte dos requisitos para a obtenção do título de Bacharel em Engenharia Ambiental.

Orientador: Prof. Dr. Ricardo Carvalho de Almeida

CURITIBA  
2014

EMPREGO DE REDES NEURAIAS ARTIFICIAIS PARA A MODELAGEM  
CHUVA-VAZÃO DA BACIA DO RIO DAS MORTES

Isadora Chiamulera

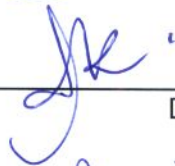
PROJETO FINAL SUBMETIDO AO CORPO DOCENTE DO DEPARTAMENTO DE  
ENGENHARIA AMBIENTAL DA UNIVERSIDADE FEDERAL DO PARANÁ COMO  
PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE  
BACHAREL EM ENGENHARIA AMBIENTAL.

Examinada por:



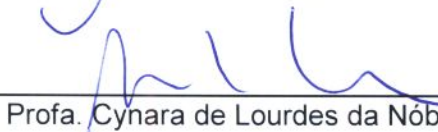
---

Prof. Dr. Ricardo Carvalho de Almeida



---

Dra. Akemi Kan



---

Profa. Cynara de Lourdes da Nóbrega Cunha

CURITIBA, PR - BRASIL  
DEZEMBRO DE 2014

## RESUMO

Este trabalho avaliou o desempenho de um modelo empírico baseado em Redes Neurais Artificiais para a previsão de vazões diárias na Bacia do Rio das Mortes - MG. A metodologia de treinamento e validação da rede é apresentada e utiliza como variáveis de entrada dados pretéritos de precipitação e vazão. O modelo conceitual Sacramento foi calibrado para a mesma bacia durante o estágio acadêmico realizado pela autora no Centro de Hidráulica e Hidrologia Prof. Parigot de Souza (CEHPAR). Uma ferramenta de calibração automática foi utilizada para a calibração dos parâmetros do modelo, o qual tem como variáveis de entrada séries de evapotranspiração e de precipitação. Apesar de consistirem em modelos com características distintas, os resultados obtidos pela Rede Neural Artificial e pelo modelo Sacramento puderam ser comparados através de diversos índices de qualidade. Para o período analisado o modelo de RNA gerou resultados mais acurados. Conclui-se que fatores como o grau de detalhamento requerido, o tempo e os recursos computacionais disponíveis devem ser levados em consideração para a escolha de um modelo hidrológico.

## **ABSTRACT**

This study evaluated the performance of an empirical model based on Artificial Neural Networks (ANNs) for the prediction of daily waterflow at the Rio das Mortes catchment in the state of Minas Gerais, Brazil. The methodology used for training and validation of the network presented uses past precipitation and waterflow data as input variables. The conceptual model Sacramento was calibrated for the same catchment during the academic internship carried out by the author at the Centro de Hidráulica e Hidrologia Prof. Parigot de Souza (CEHPAR). An automatic calibration tool was used for calibrating the model parameters. Precipitation and evapotranspiration data were used as inputs to the model. Even though the ANN-based model and the Sacramento model display distinct features, it was possible to compare the results obtained from both through different statistical parameters. For the time frame analyzed, the ANN-based model produced the most accurate results. It can be concluded that factors such as the level of detail that is required as well as the time and computational resources available must be taken into account when choosing the best suited hydrological model.

## LISTA DE FIGURAS

Figura 1 – Representação do Método dos Polígonos de Thiessen.....	6
Figura 2 – Componentes do Modelo Sacramento. ....	11
Figura 3 - Representação de um neurônio artificial. ....	16
Figura 4 - Representação gráfica de diferentes funções de ativação .....	17
Figura 5 - Modelo de rede neural com arquitetura do tipo <i>multilayer perceptron</i> . ....	18
Figura 6 – Superfície de erro bidimensional. ....	19
Figura 7 – Diagrama esquemático ilustrando o problema de overfitting.....	20
Figura 8 – Erro dos conjuntos de teste e de treinamento em função das épocas de treinamento. ....	21
Figura 9 - Localização da bacia hidrográfica do Rio das Mortes.....	26
Figura 10 – Localização das estações pluviométricas e fluviométricas de interesse na bacia do Rio das Mortes.....	28
Figura 11 – Configuração do modelo Sacramento no software SOBEK. ....	30
Figura 12 - Calibração da bacia do Rio das Mortes no software RRL.....	31
Figura 13 – Discretização em sub-bacias conforme configuração semi-distribuída (acima) e concentrada (abaixo). ....	32
Figura 14 - Coeficiente de Nash-Sutcliffe para a escolha do número de neurônios da camada intermediária.....	40
Figura 15 - Coeficiente de Nash-Sutcliffe para a definição das variáveis de entrada. ....	41
Figura 16 - Coeficiente de Nash-Stucliffe para configurações com a ausência (azul) e presença (vermelho) da variável evapotranspiração. ....	41
Figura 17 - Erros dos conjuntos de treinamento (azul) e de teste (vermelho) ao longo das épocas de treinamento da configuração T_5_38. ....	42
Figura 18 – Diagrama de dispersão da configuração T_5_38. ....	43
Figura 19 - Curvas de vazão observada e prevista pelo modelo de RNA e pelo modelo Sacramento de 2009 a 2010.....	45
Figura 20- Curvas de vazão observada e prevista pelo modelo de RNA e pelo modelo Sacramento de 2011 a 2012.....	46

# SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	1
1.1. TEMA.....	1
1.2. DEFINIÇÃO DO PROBLEMA .....	2
1.3. OBJETIVOS.....	3
<b>2. REVISÃO BIBLIOGRÁFICA</b> .....	3
2.1. PROCESSOS HIDROLÓGICOS .....	3
2.2. MODELOS HIDROLÓGICOS .....	6
2.2.1. Classificação de modelos hidrológicos.....	7
2.2.2. Modelos conceituais.....	8
2.2.3. Modelos empíricos .....	13
2.3. REDES NEURAS ARTIFICIAIS.....	14
2.4. APLICAÇÕES DE REDES NEURAS ARTIFICIAIS EM HIDROLOGIA .....	22
<b>3. METODOLOGIA</b> .....	25
3.1. CARACTERIZAÇÃO DA REGIÃO DE ESTUDO .....	25
3.2. CONFIGURAÇÃO DO MODELO SACRAMENTO .....	26
3.2.1. Seleção e controle de qualidade dos dados de entrada.....	27
3.2.2. Configuração do modelo no software SOBEK .....	29
3.2.3. Calibração e validação do modelo Sacramento .....	30
3.3. CONFIGURAÇÃO DA REDE NEURAL ARTIFICIAL.....	32
3.3.1. Seleção e controle de qualidade dos dados de entrada.....	32
3.3.2. Análise estatística dos dados .....	34
3.3.3. Treinamento e operação da Rede Neural Artificial.....	35
<b>4. RESULTADOS E DISCUSSÃO</b> .....	38
4.1. MODELO SACRAMENTO .....	38
4.2. MODELO DE REDE NEURAL ARTIFICIAL.....	39
4.3. COMPARAÇÃO DOS RESULTADOS DOS MODELOS .....	44
<b>5. CONCLUSÕES</b> .....	47
<b>6. REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	50

# 1. INTRODUÇÃO

## 1.1. TEMA

O Brasil é dotado de uma vasta e densa rede de drenagem, presente em mais de 90% do território nacional, sendo classificado como um país rico em água doce pelas Nações Unidas (REBOUÇAS, 2003).

Atualmente, entre os principais usos da água podem ser citados o abastecimento humano, a geração de energia, a irrigação e a navegação. A disponibilização da água para os usos citados, em termos quantitativos e qualitativos, representa um grande desafio para a sociedade além de uma parcela significativa dos orçamentos nacionais e regionais (TUCCI, 2013). Sendo assim, a água é um bem de importância estratégica e seu estudo vem se tornando cada vez mais necessário.

A hidrologia é a ciência que trata do estudo da água na Terra. Ela compreende a análise de dados coletados para o estabelecimento de suas relações mútuas e a aplicação dos conhecimentos adquiridos para a solução de inúmeros problemas práticos (PINTO *et al.*, 1976). Por se relacionar com outros fenômenos naturais e sociais, essa ciência é extremamente ampla e interdisciplinar.

Ainda que os fenômenos hidrológicos mais comuns, como as chuvas e o escoamento dos rios, possam parecer suficientemente conhecidos, esta ciência ainda requer o aprimoramento e o desenvolvimento de técnicas que possibilitem a utilização racional e o gerenciamento adequado da água de modo a promover o crescimento econômico e o bem estar social de um país (TUCCI, 2013).

Entre essas técnicas está a modelagem hidrológica, que busca reproduzir o comportamento dos sistemas, tanto em seus componentes naturais (transformação da chuva em escoamento superficial) como nos artificiais (usinas, reservatórios, captações, canais e redes) (KRAUSKOPF NETO, 2005).

A modelagem hidrológica é uma atividade bastante complexa, devido à variabilidade espacial e temporal dos processos envolvidos. Esta variabilidade leva, via de regra, à simplificação do modelo e dos dados utilizados, gerando erros nas simulações dos sistemas hidrológicos.



Dado que as previsões hidrológicas são parte de sistemas de tomada de decisão, erros nestes valores representam menor eficácia nas decisões tomadas, o que se traduz em perdas, geralmente econômicas, mas no pior dos casos, em perda de vidas (KRAUSKOPF NETO, 2005).

## 1.2. DEFINIÇÃO DO PROBLEMA

A ocorrência de enchentes gera diversos prejuízos que podem ser de natureza econômica, social ou ambiental. Os prejuízos econômicos estão relacionados às perdas materiais, como a destruição de casas, e à interrupção das atividades econômicas da região. Com relação aos prejuízos sociais, tem-se o risco da transmissão de doenças, do desgaste psicológico por estresse traumático das populações atingidas além da perda de vidas. Por fim, a degradação do meio ambiente também causa danos pelo esgotamento dos recursos naturais e a contaminação da água pela inundação de depósitos de material tóxico, estações de tratamentos, entre outros.

Os modelos hidrológicos de previsão de vazão são ferramentas que auxiliam o gerenciamento dos reservatórios e podem ter extrema importância na prevenção de danos decorrentes de eventos extremos, como por exemplo, através da emissão de alertas para as comunidades ribeirinhas.

O presente trabalho complementa o projeto de Pesquisa e Desenvolvimento (P&D) intitulado “Implantação de Sistema para Acoplamento de Modelos e Informação Telemétrica Visando à Otimização da Operação de Reservatórios em Tempo Real, com Foco no Controle de Cheias”, realizado pelo Centro de Hidráulica e Hidrologia Prof. Parigot de Souza (CEHPAR).

O projeto visa propor a utilização de um sistema inovador que garanta a integração das etapas da operação de reservatórios e especialmente, possibilite a minimização de riscos de enchente em cidades localizadas à jusante das barragens. Para isso, estão sendo realizados estudos de modelagem hidrológica em bacias hidrográficas do Estado de Minas Gerais.

A partir dos resultados obtidos pelo modelo Sacramento, utilizado no

projeto para a previsão de vazão, observou-se a possibilidade da melhoria de resultados através do emprego de modelos baseados em Redes Neurais Artificiais.

O presente trabalho aborda os procedimentos envolvidos no emprego de Redes Neurais Artificiais para a previsão de vazões diárias em uma das bacias estudadas no projeto de P&D.

### 1.3. OBJETIVOS

O objetivo geral deste trabalho é configurar um modelo de Rede Neural Artificial para a previsão de vazão diária da bacia hidrográfica do Rio das Mortes, no Estado de Minas Gerais.

Os objetivos específicos são:

- Fazer a análise de dados hidrometeorológicos que sejam relevantes na previsão da vazão;

- Efetuar o treinamento, a verificação e a validação de um modelo de RNA com arquitetura do tipo MLP (*Multilayer Perceptrons*), através do algoritmo *backpropagation* para geração de uma série de vazão prevista;

- Avaliar o desempenho do modelo de Rede Neural Artificial e do modelo Sacramento de maneira objetiva através da comparação de índices estatísticos.

Salienta-se que a configuração do modelo Sacramento foi efetuada como parte do projeto de Pesquisa e Desenvolvimento citado previamente, durante estágio realizado pela autora no CEHPAR.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1. PROCESSOS HIDROLÓGICOS

O ciclo hidrológico é o fenômeno global de circulação fechada da água entre a atmosfera e a superfície terrestre. Caminhos complexos incluem a passagem da água do seu estado gasoso para os corpos d'água e também para o solo até retornar à atmosfera (TUCCI, 2013).

Para analisarmos esses fenômenos hidrológicos é necessária a definição de um elemento fundamental de análise, a bacia hidrográfica.

Segundo Tucci (2013), a bacia hidrográfica é uma área de captação natural da água da precipitação que faz convergir os escoamentos para um único ponto de saída, seu exutório. A bacia pode ser considerada um sistema físico, com fronteira no divisor de águas, que tem o papel de transformar uma entrada de volume de precipitação (*input*) concentrada no tempo em um escoamento (*output*), de forma mais distribuída no tempo (TUCCI, 2013). O balanço hídrico de uma bacia hidrográfica tenta quantificar essas entradas e saídas de água no sistema (CINTRA *et al.*, 2000).

O ciclo hidrológico pode ser explicado a partir da presença de vapor d'água na atmosfera que, sob determinadas condições meteorológicas condensa-se, possibilitando a precipitação. Entende-se por precipitação a água proveniente do meio atmosférico que atinge a superfície terrestre de diferentes formas como chuva, granizo, orvalho, neblina, neve ou geada (PINTO *et al.*, 1976). Quando as gotículas de água atingem uma dimensão suficiente para que seu peso seja superior às forças que as mantêm em suspensão elas adquirem então uma velocidade de queda superior às componentes verticais ascendentes dos movimentos atmosféricos (TUCCI, 2013).

Antes de chegar à superfície terrestre, parte do volume precipitado fica retido na vegetação. Esse processo conhecido como interceptação interfere no balanço hídrico uma vez que armazena uma parcela da precipitação (ARCOVA *et al.*, 2003).

O volume precipitado que atinge o nível do solo pode ficar armazenado em depressões ou sofrer o processo de infiltração no solo. Essa transferência da água se dá por efeito das forças de gravidade e capilaridade e está diretamente relacionada com as características do solo, como porosidade, umidade e grau de compactação (ARCOVA *et al.*, 2003).

Quando a intensidade da precipitação excede a capacidade de infiltração do solo, a água se movimenta através do escoamento superficial, impulsionada pela ação da gravidade. Esse processo ocorre naturalmente, quando a água é drenada da bacia hidrográfica até que forme os cursos dos rios para finalmente dirigirem-se aos grandes volumes de água constituídos pelos lagos, mares e oceanos (PINTO *et al.*, 1976). A parcela de água infiltrada também é escoada. Entretanto, em camadas inferiores do solo, os

escoamentos subsuperficiais e subterrâneos acontecem de forma mais lenta.

Em todas as etapas do ciclo, a água fica disponível para a evaporação. Esse fenômeno ocorre quando a água muda do estado líquido para vapor. As plantas também contribuem para esse processo através da transpiração, quando a passagem da água para a atmosfera ocorre através dos estômatos localizados nas folhas. A transpiração é um mecanismo vital para a manutenção da vida dos vegetais, pois a partir dela ocorre o resfriamento das folhas e a movimentação de nutrientes na planta (PIMENTEL, 2004). Comumente a evaporação e a transpiração são tratadas como um único fator, a evapotranspiração, resultado da soma das duas grandezas. Uma vez disposta na atmosfera a água se torna disponível para a precipitação, dando início novamente ao ciclo hidrológico.

Um fator primordial para os estudos hidrológicos é a existência de dados representativos que possibilitem uma análise precisa da região de interesse. Esses dados referem-se principalmente a séries históricas de precipitação e vazão, medidas por meio de estações pluviométricas e fluviométricas, respectivamente.

Em estações pluviométricas, aparelhos específicos como pluviômetros e pluviógrafos são utilizados para a quantificação do volume de chuva acumulado durante um determinado período de tempo. Considerando uma bacia hidrográfica, é comum expressar a altura de chuva através da precipitação média  $\bar{p}$ , que leva em conta os dados observados em todas as estações presentes na bacia.

Um dos métodos mais utilizados para o cálculo de  $\bar{p}$  é o método dos Polígonos de Thiessen, que leva em consideração a área de influência associada a cada estação. Essas áreas são obtidas através da análise do mapa da bacia. Primeiramente, as estações são conectadas por linhas retas e, em seguida, são traçadas as mediatrizes de cada linha de conexão, formando os polígonos (Figura 1). Por fim, os dados de precipitação  $P_i$  das  $n$  estações são ponderados através das áreas  $A_i$  correspondentes aos seus polígonos, obtendo-se uma precipitação média, conforme a equação 1 (BRAZ *et al.*, 2007).

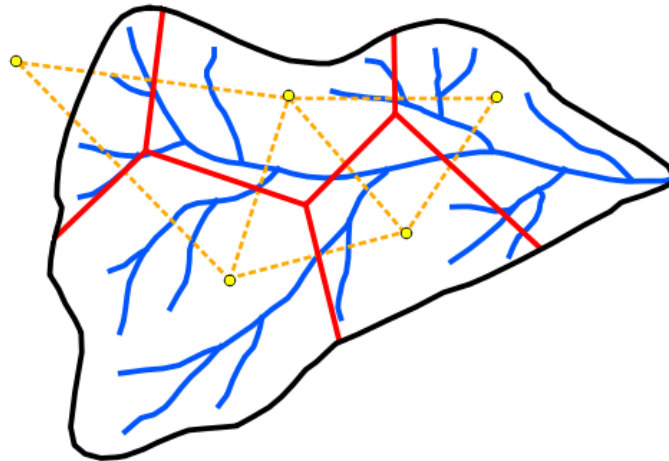


Figura 1 – Representação do Método dos Polígonos de Thiessen.

$$\bar{p} = \frac{\sum_{i=1}^n P_i A_i}{\sum_{i=1}^n A_i} \quad (1)$$

Este método é eficaz em terrenos levemente acidentados, mas pode produzir resultados equivocados em regiões de relevo irregular, onde as barreiras orográficas podem mudar a distribuição das chuvas, e em grandes bacias, onde a precipitação não é homogênea (TUCCI, 2013).

Em estações fluviométricas a medição de vazões é de difícil operação e alto custo, desta forma, na maioria delas são utilizadas escalas linimétricas para a leitura da cota dos rios. De maneira indireta, a vazão gerada pela bacia hidrográfica pode ser obtida pelo nível de água no exutório, através da curva chave, que expressa a relação cota-vazão.

## 2.2. MODELOS HIDROLÓGICOS

O ciclo hidrológico é um processo físico extremamente complexo, afetado por muitos fatores que podem ou não estar relacionados entre si. Entre esses fatores, destacam-se as características do evento de precipitação, as particularidades da bacia hidrográfica e os aspectos geomorfológicos e climáticos da região (JAIN e KUMAR, 2007). Os modelos hidrológicos são uma das ferramentas desenvolvidas pela ciência para melhor entender e representar o comportamento de uma bacia hidrográfica.

As dificuldades da modelagem hidrológica em representar os

fenômenos naturais surgem devido às características não lineares inerentes aos fenômenos hidrológicos, o que justifica a complexidade de se construir um modelo capaz de simular de forma precisa as respostas a diferentes eventos (BEVEN, 1989). Por serem simplificações extremas da realidade, mesmo que esses modelos sejam bons descritores de processos bem definidos, homogêneos e estacionários, não têm a mesma capacidade em descrever os sistemas reais tridimensionais, heterogêneos e não permanentes (BEVEN, 1989).

Dependendo dos objetivos estabelecidos, diferentes modelos podem ser utilizados no gerenciamento dos recursos hídricos. Dentre estes, estão os modelos hidrodinâmicos, os de qualidade da água e os de chuva-vazão.

Neste estudo serão detalhados aspectos relacionados a modelos do tipo chuva-vazão que representam a parte do ciclo hidrológico entre a precipitação e a vazão. Este modelo permite antecipar, por exemplo, a ocorrência de enchentes em uma determinada região ou os impactos da urbanização de uma bacia antes que ela ocorra (TUCCI, 2005). As variáveis de entrada utilizadas pelos modelos chuva-vazão são séries históricas de dados de precipitação, evaporação e da própria série de vazão para que os resultados das simulações possam ser validados.

### 2.2.1. Classificação de modelos hidrológicos

Os modelos hidrológicos podem ser classificados em função de diferentes critérios tais como: o tipo das variáveis utilizadas no modelo (estocásticos e determinísticos) e o comportamento no tempo (discretos ou contínuos) (LUCAS *et. al*, 2009).

Pode-se também classificar a discretização da bacia de acordo com a existência ou não de relações espaciais. De forma geral são adotadas uma das seguintes estruturas (TUCCI, 2005):

- Concentrado: os processos hidrológicos são representados por variáveis concentradas no espaço.
- Distribuído por sub-bacias: o modelo considera a divisão da bacia em sub-bacias. A divisão pode ser realizada com base na disponibilidade de dados, locais de interesse e variabilidade

de parâmetros físicos da bacia.

- Distribuído por módulos: a discretização é realizada por formas geométricas como quadrados, sem relação direta com a forma da bacia.

Ainda, os modelos podem ser diferenciados pelo tipo de relação entre as variáveis de entrada e saída. Os modelos conceituais buscam representar os processos físicos observados nos sistemas naturais enquanto os modelos empíricos ajustam os valores calculados aos dados observados através de funções que não têm nenhuma relação explícita com os processos físicos envolvidos (TUCCI, 2005).

Entre as principais dificuldades encontradas pelo uso dos modelos conceituais estão a grande variabilidade espacial dos processos e a dificuldade da estimativa dos parâmetros (TUCCI, 2005). Em geral, os modelos conceituais necessitam de uma grande quantidade de dados de entrada para calibração e validação (YAPO *et al.*, 1996) e, além disso, demandam um grande esforço computacional (FRAKES e YU, 1999).

Os modelos empíricos, por sua vez, se baseiam em relações estatísticas, dependendo fortemente de uma base de dados confiáveis para a sua aplicação.

### 2.2.2. Modelos conceituais

Os modelos conceituais baseiam-se na análise e no desenvolvimento das relações entre as características físicas relacionadas aos processos hidrológicos (FRAKES e YU, 1999).

Geralmente a capacidade desses modelos em representar a variabilidade espacial e temporal dos processos hidrológicos, faz com que sua aplicação seja mais adaptada às grandes bacias hidrográficas, onde as características do solo e da distribuição das chuvas, por exemplo, são mais heterogêneas (FRAKES e YU, 1999).

Por mais que esses modelos sejam suficientemente confiáveis na predição das principais variáveis hidrológicas, a implementação e a calibração desses modelos normalmente apresentam obstáculos, requerendo ferramentas matemáticas sofisticadas (DUAN *et al.*, 1992) e um bom grau de

entendimento do modelo pelo usuário (ZEALAND *et al.*, 1999).

Entre os diversos modelos hidrológicos conceituais existentes, pode-se citar o modelo Sacramento, desenvolvido por Burnash *et al.* em 1973 no *National Weather Service* (NWS) dos Estados Unidos. Ele é mundialmente utilizado tanto para fins de estudo como para gerar previsões operacionais. Em função do seu amplo tempo de utilização, existe um considerável acervo de conhecimentos sobre todos os seus aspectos, tanto conceituais como operacionais (KRAUSKOPF NETO, 2005).

O modelo utiliza a precipitação e a evapotranspiração da bacia estudada como variáveis de entrada para o processo de determinação da vazão. Ele considera o armazenamento de água no solo através de um conjunto de reservatórios de capacidades determináveis que são conectados por processos, permitindo ao sistema aproximar muitas das condições de umidade do solo que determinam a vazão. O estoque de água aumenta com a precipitação e é reduzido através de percolação vertical, evapotranspiração e/ou drenagem lateral (BURNASH, 1995).

A maneira mais fácil de visualizar os componentes básicos do modelo Sacramento é entendendo a umidade presente no solo (NWS, 2002). Para isso, faz-se necessária a definição de dois tipos de umidade do solo. A água pelicular é a porção de umidade que se adere às partículas do solo por processos de atração molecular, diminuindo sua mobilidade, podendo retornar à atmosfera através da evapotranspiração. Já a água gravitacional representa o volume no estado líquido infiltrado no solo que se movimenta com mais facilidade. Essa água percola suprindo as deficiências de umidade que encontrar no solo (BURNASH, 1995).

Uma das soluções que foi encontrada pelos autores do modelo Sacramento para a representação das características de drenagem e percolação foi a divisão do solo e conseqüentemente das águas peliculares e gravitacionais em duas camadas: a camada superior e a camada inferior (NWS, 2002).

A camada superior representa os processos mais dinâmicos da bacia, de menor tempo de resposta, e é modelada por dois tipos de armazenamento (KRAUSKOPF NETO, 2005):



- Reservatório de água pelicular: nesse reservatório, a água pelicular deixa o reservatório através de evapotranspiração ou como fluxo para o reservatório gravitacional, quando a capacidade de armazenamento for superada;

- Reservatório de água gravitacional: este componente apresenta como entrada o volume proveniente do reservatório de água pelicular e deixa o reservatório através do escoamento subsuperficial e de percolação. Quando toda sua capacidade de armazenamento é utilizada, ocorre o escoamento superficial.

A camada inferior representa o armazenamento subterrâneo e a maior parcela da umidade presente no solo, modelando os fenômenos de variação mais lenta e maior duração no processo hidrológico. Esta camada é composta por dois componentes de armazenamento:

- Reservatório de água pelicular: representa o armazenamento de água pelicular. Este componente é alimentado pelo volume percolado da camada superior e por fluxos provenientes dos reservatórios gravitacionais da camada inferior;

- Reservatórios de água gravitacional: semelhantemente ao reservatório gravitacional da camada superior, este componente é alimentado pelo volume percolado e apresenta como saída o fluxo que gera o escoamento de base. Para representar o escoamento de base com um maior grau de não-linearidade, o volume é distribuído em dois reservatórios:

Reservatório de água gravitacional primária: representa o escoamento de base de maior inércia, de resposta mais lenta e contínua;

Reservatório de água gravitacional suplementar: representa o escoamento de base de resposta mais rápida, mas de menor duração.

Os fluxos de água também podem ocorrer a partir dos reservatórios de água gravitacional para os reservatórios de água pelicular de modo a atender déficits de transpiração. A Figura 2 apresenta de maneira simplificada os

componentes do modelo Sacramento, com a influência de cada escoamento no hidrograma unitário.



Figura 2 – Componentes do Modelo Sacramento.

Para a calibração do modelo devem ser determinados 17 parâmetros relacionados ao tamanho, clima, topografia, geologia e cobertura vegetal da bacia. A Tabela 1 apresenta o detalhamento desses parâmetros (BURNASH, 1995):

Tabela 1 – Descrição dos parâmetros do modelo Sacramento.

	Parâmetro	Unidade	Descrição
Escoamento direto	PCTIM	--	Fração da área impermeável permanente da bacia.
	ADIMP	--	Fração máxima da área impermeável variável da bacia.
	SARVA	--	Fração da bacia coberta por córregos, lagos e vegetação que possa esgotar vazão por evapotranspiração.
Camada superior do solo	UZTWM	mm	Capacidade do reservatório de água pelicular.
	UZFWM	mm	Capacidade do reservatório de água gravitacional.
	UZK	1/dia	Taxa de drenagem do reservatório de água gravitacional.
Percolação	ZPERC	--	Máxima fração de aumento da demanda de percolação mínima.
	REXP	--	Expoente da equação de percolação.
Camada inferior do solo	LZTWM	mm	Capacidade do reservatório de água pelicular.
	LZFPM	mm	Capacidade do reservatório primário de água gravitacional.
	LZFSM	m	Capacidade do reservatório suplementar de água gravitacional.
	LZPK	1/dia	Taxa de drenagem do reservatório primário de água gravitacional.
	LZSK	1/dia	Taxa de drenagem do reservatório suplementar de água gravitacional.
	PFREE	--	Fração da percolação destinada aos reservatórios de água gravitacional.
	RSERV	--	Capacidade de água gravitacional indisponível para transpiração
	SIDE	--	Fração do escoamento de base total que compõe a perda de base.
	SSOUT	m <sup>3</sup> /s/km <sup>2</sup>	Taxa fixa do aporte total ao canal que compões a perda de base.

Alguns dos parâmetros apresentados na Tabela 1 podem ser calibrados através da análise de hidrogramas ou a partir de características fisiográficas da bacia e outros devem ser estimados com base em método de tentativa e erro. Também é possível fazer uso de algoritmos de calibração

automática baseados em técnicas de otimização do erro global (GIBERTONI *et al.*, 2014).

O modelo Sacramento é implementado em uma série de softwares, que facilitam sua aplicação, um deles, é o software SOBEK.

Desenvolvido pelo instituto holandês Deltares, o SOBEK é um software que auxilia no gerenciamento de corpos hídricos. Uma de suas principais aplicações é a simulação de complexos processos hidrológicos em bacias hidrográficas (DELTARES, 2013).

O SOBEK não oferece um módulo de calibração automática. Para a realização desta etapa podem ser utilizadas ferramentas auxiliares. O software *Rainfall Runoff Library* (RRL), desenvolvido pelo Programa de Pesquisa do Comportamento Hidrológico na Austrália (*CRC for catchment hydrology*) permite a calibração automática ou manual dos parâmetros de cinco modelos hidrológicos, entre eles o modelo Sacramento.

O RRL utiliza como dados de entrada séries diárias de precipitação, evapotranspiração e vazão, além da área da bacia, a qual é necessária para a conversão da altura de precipitação em escoamento superficial. Os períodos de calibração e de verificação podem ser especificados manualmente. Para a calibração dos parâmetros são oferecidos oito diferentes métodos de otimização que buscam minimizar os erros entre os valores modelados e observados. Além disso, também são disponibilizadas ferramentas que permitem a visualização do desempenho do método simultaneamente ao ajuste dos parâmetros, o que facilita a calibração do modelo (CRC, 2004).

### 2.2.3. Modelos empíricos

Os modelos empíricos não utilizam formulações físicas para descrever o comportamento hidrológico da região estudada. Ao invés disso, são baseados em relações estatísticas entre a precipitação e o escoamento obtidas a partir de observações.

As principais vantagens do uso de modelos empíricos, também chamados “caixa-preta”, são a baixa demanda computacional e o reduzido tempo necessário para calibração. A capacidade de processamento exigida

para a implementação desses modelos é menor quando comparada a outros modelos, facilitando sua aplicação (FRAKES e YU, 1999).

Esses modelos têm grande acurácia e são de fácil operação, pois envolvem um conhecimento mínimo a respeito dos processos subjacentes à modelagem hidrológica (MINOCHA, 2003).

Entretanto, a simplificação demasiada da distribuição temporal e espacial das características da região pode tornar o modelo empírico ineficaz, dependendo de como for implementado. Isto pode ocorrer em grandes bacias, onde condições de homogeneidade não são observadas.

Os modelos empíricos, como as regressões, geralmente se baseiam em relações lineares, que raramente acontecem na natureza. No entanto, o uso de Redes Neurais Artificiais vem superando essa dificuldade (FRAKES e YU, 1999).

### 2.3. REDES NEURAS ARTIFICIAIS

As Redes Neurais Artificiais (RNAs) são ferramentas computacionais capazes de desenvolver algoritmos matemáticos que modelam relações envolvendo séries temporais complexas em várias áreas do conhecimento (ZANETTI *et al.*, 2008).

RNAs possuem propriedades valiosas que não são encontradas em outros métodos. Primeiramente, elas deduzem soluções empiricamente, a partir de dados de entrada, sem o conhecimento prévio de suas características. As redes neurais são capazes de identificar padrões e podem modificar o seu comportamento em resposta ao ambiente, de forma a ajustar-se para obter uma resposta consistente. As RNAs também têm a propriedade de generalizar cenários a partir de exemplos anteriores e aplicá-los em novos casos. Essa capacidade faz com que as redes abstraíam características essenciais a partir de entradas que possam conter dados irrelevantes, o que acontece frequentemente, uma vez que os dados do mundo real contêm ruídos, são distorcidos e muitas vezes incompletos (ZEALAND *et al.*, 1999).

Outra característica importante é a capacidade de modelar as relações não lineares entre os dados, isto é, as redes neurais podem resolver alguns problemas complexos com mais acurácia do que técnicas lineares o

fazem. Finalmente, as RNAs contêm muitas operações independentes que podem ser executadas simultaneamente, frequentemente tornando a modelagem mais rápida quando comparada a outros métodos numéricos (ZEALAND *et al.*, 1999).

As RNAs fazem uso do que se sabe sobre a funcionalidade das redes biológicas para a resolução de problemas complexos. Os neurônios biológicos têm três principais unidades funcionais: o corpo celular, os dendritos e o axônio. Os dendritos recebem informações, na forma de sinais elétricos, vindas de outros neurônios e as fazem passar pelo corpo celular. Já os axônios recebem os sinais do corpo celular e, através da sinapse, os transmitem para os dendritos dos neurônios vizinhos (BASHEER e HAJMEER, 2000).

Mesmo que os neurônios estejam muito próximos uns aos outros não existe contato físico entre as terminações dos axônios e dendritos. Para que o sinal seja transmitido através desse espaço, a fenda sináptica, o sistema nervoso utiliza substâncias químicas, os neurotransmissores (BASHEER e HAJMEER, 2000).

A chegada dos impulsos nervosos nas terminações dos axônios libera neurotransmissores em quantidades proporcionais à intensidade dos sinais. Os neurotransmissores são difundidos pela fenda sináptica até os dendritos dos neurônios vizinhos. Quando a quantidade de neurotransmissores recebida for maior que certo limiar (dependente de cada neurônio), um novo sinal elétrico será gerado (sinapse excitatória), caso contrário, o sinal não será transmitido (sinapse inibitória).

Pode-se fazer uma analogia entre um neurônio biológico e um neurônio artificial. As conexões entre os nós representam os axônios e os dendritos, os pesos representam o potencial elétrico nas sinapses e a aplicação do limiar se aproxima da atividade no corpo celular. De maneira similar, o neurônio artificial só é ativado quando o conjunto de estímulos recebido representar um valor maior que o seu limiar (BASHEER e HAJMEER, 2000).

No neurônio artificial os sinais de entrada  $x_i$  são ponderados por seus pesos sinápticos  $w_i$  e transmitidos através das  $n$  conexões, como mostra a Figura 3.

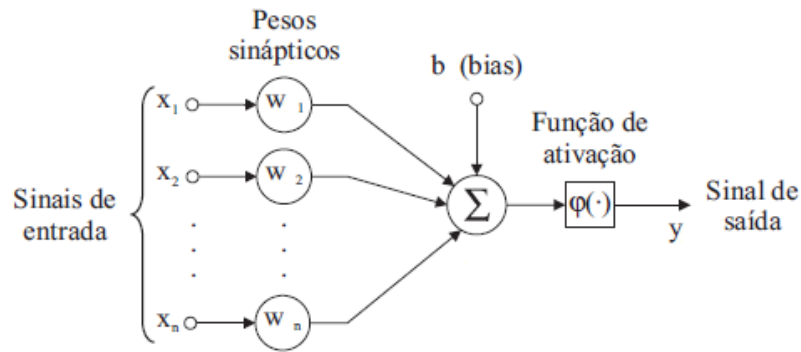


Figura 3 - Representação de um neurônio artificial.

Uma função de ativação  $\varphi$  é aplicada ao somatório  $\xi$  dessa ponderação e ao *bias*, para determinar a atividade do neurônio.

O *bias* do neurônio pode ser considerado como um *input* adicional, cujo valor é sempre unitário e seu peso é igual a  $b$  (BASHEER e HAJMEER, 2000). Ele tem o efeito de aumentar ou diminuir o *input* do neurônio, atuando como um limitante para a ativação do neurônio.

Em termos matemáticos, a atividade em um neurônio pode ser descrita através do seguinte par de equações (Equação 2):

$$\begin{cases} \xi = \sum_{i=1}^n x_i w_i \\ y = \varphi(\xi + b) \end{cases} \quad (2)$$

As funções de ativação são utilizadas para simular a resposta do neurônio ao receber os sinais de entrada, limitando a amplitude do sinal de saída para um valor finito. O sinal de saída  $y$  é produzido se o resultado gerado pela função de ativação ultrapassar o limiar pré-determinado, nesse caso, o neurônio será ativado e as informações serão transmitidas ao neurônio seguinte. Esse comportamento é análogo ao produzido pelas sinapses excitatórias, observadas no neurônio biológico.

Para que valores muito discrepantes em relação à média não tendenciem o ajuste dos pesos, deve ser realizada a normalização dos valores dos pares entrada-saída, normalmente limitados ao intervalo  $[-1,1]$ , dependendo da função de ativação escolhida.

A normalização de variáveis com distribuição assimétrica, tais como a precipitação, não deve se basear no valor médio da série, pois isso levaria a valores incorretos da série normalizada. Para contornar esse problema os valores mínimo e máximo da série devem ser levados em consideração em oposição ao valor médio. Uma das possíveis equações utilizadas nesse caso é mostrada a seguir (Equação 3), onde  $x_n$  corresponde ao valor normalizado da variável original  $x$  e os valores máximo e mínimo da série a ser normalizada são representados por  $x_{máx}$  e  $x_{mín}$ , respectivamente (CARDOSO, 2009).

$$x_n = \frac{2(x - x_{mín})}{x_{máx} - x_{mín}} - 1 \quad (3)$$

Segundo Fiorin *et al.* (2011), as funções de ativação mais comumente utilizadas são a função limiar (ou degrau), a função identidade e as funções sigmóides logística e tangente hiperbólica, respectivamente (Figura 4).

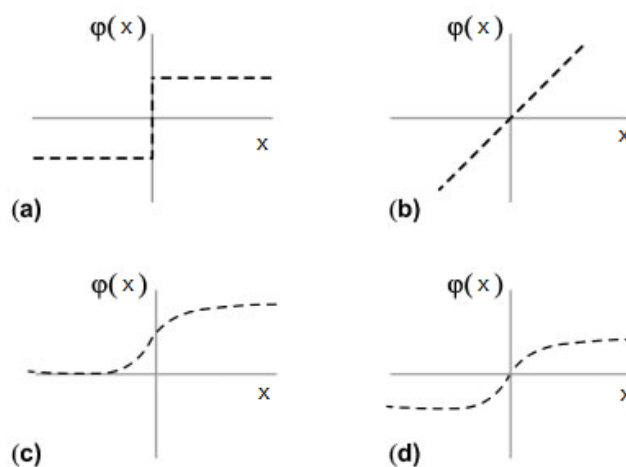


Figura 4 - Representação gráfica de diferentes funções de ativação: (a) função degrau; (b) função linear; (c) função logística; (d) função tangente hiperbólica.

O modo de organização dos neurônios caracteriza a arquitetura da rede. Essa estrutura determina o fluxo das informações no modelo tendo um impacto significativo na relação entre os *inputs* e *outputs* (HSIEH e TANG, 1998).

Em redes com arquitetura do tipo *feedforward*, a propagação da



informação é realizada somente em uma direção, da entrada para a saída, uma vez que os neurônios de uma mesma camada não são conectados entre si (MAIER *et al.*, 2010).

Um das formas mais populares desse tipo de arquitetura são as redes *Multilayer Perceptrons* (MLP). Elas utilizam três ou mais camadas de neurônios artificiais, como apresentado na Figura 5.

A adição de camadas intermediárias entre os neurônios de entrada e de saída aumenta a capacidade da rede em resolver problemas não lineares (BASHEER e HAJMEER, 2000).

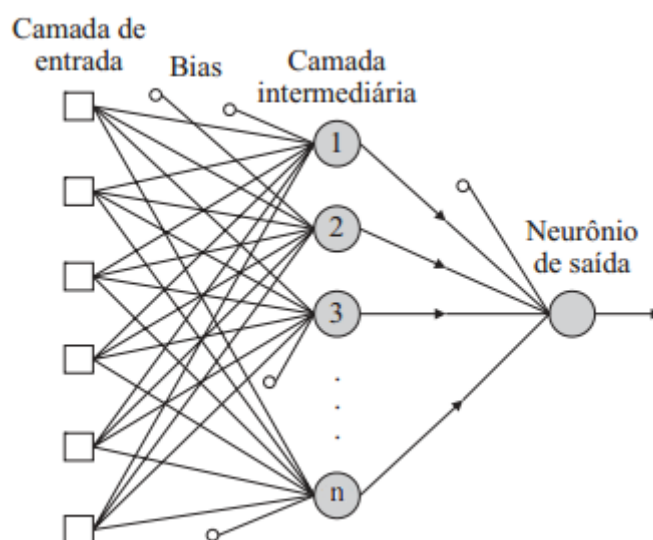


Figura 5 - Modelo de rede neural com arquitetura do tipo *multilayer perceptron*.

Em sistemas artificiais, o aprendizado é visto como o processo de atualização da representação do sistema em resposta a estímulos externos. O objetivo do treinamento é encontrar um conjunto de parâmetros, ou pesos, que possibilitem a melhor representação da relação entre os *inputs* e *outputs*.

O treinamento é realizado de forma iterativa enquanto a rede é apresentada a diferentes exemplos. A cada iteração os pesos são ajustados, visando a minimização dos erros entre os dados simulados e observados. Para que haja um melhor desempenho da rede, os dados de entrada devem ser reapresentados à rede de maneira cíclica, cada ciclo corresponde a uma época de treinamento.

No algoritmo de aprendizagem *backpropagation* os erros são retropropagados da camada de saída até a camada de entrada. Os pesos de cada neurônio são alterados em proporção ao erro gerado por eles (GARDNER e DORLING, 1998).

Cada combinação de pesos resulta em um erro, o conjunto de erros formado por todas as combinações de pesos pode ser imaginado na forma de um plano multidimensional irregular, com mínimos locais e um mínimo global. A Figura 6 apresenta uma superfície de erros bidimensional com a variação do erro em função dos pesos 1 e 2. Um falso ajuste ótimo de pesos pode ser encontrado caso a rede convirja para um mínimo local, ao invés do mínimo global (GARDNER e DORLING, 1998).

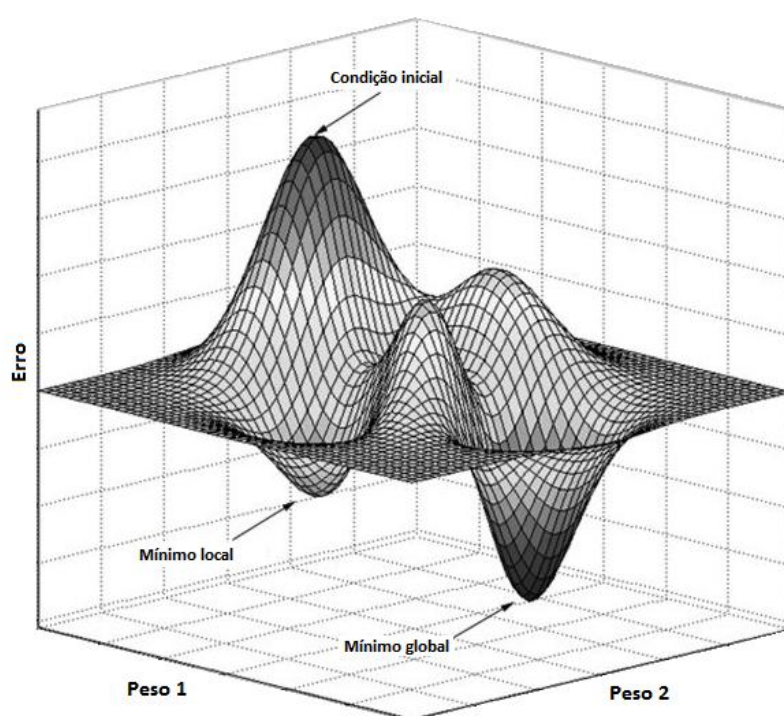


Figura 6 – Superfície de erro bidimensional.

Outra dificuldade comumente encontrada ao longo do treinamento é o *overfitting*. Esse comportamento é causado devido ao excesso de treinamento, o qual pode levar a rede a “memorizar” o problema, tornando-se demasiadamente especializada e perdendo a capacidade de generalização (HSIEH e TANG, 1998).

A Figura 7 ilustra o problema do *overfitting*. Nesta figura, a curva sólida tem um ajuste perfeito aos dados usados para o treinamento

(quadrados), porém não se ajusta adequadamente quando apresentada a novos dados (pontos). A linha pontilhada indica uma rede que não passou por treinamento excessivo, com um bom ajuste para ambos os conjuntos de dados.

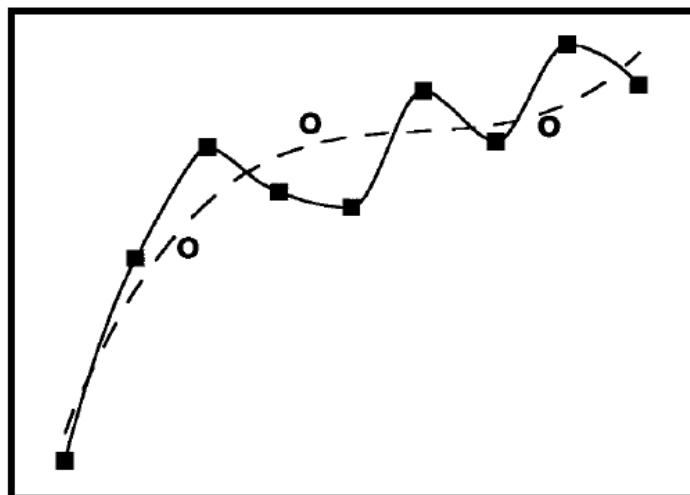


Figura 7 – Diagrama esquemático ilustrando o problema de overfitting.

Uma importante etapa do desenvolvimento de um modelo de RNA é a seleção dos dados de entrada. Durante esta fase deve-se prezar pela inclusão de dados que realmente contribuam com novas informações ao modelo, uma vez que a adição de dados de entrada redundantes aumenta a presença de mínimos locais e as chances de *overfitting*. Uma análise de correlação entre dados de entrada e saída pode ser realizada de modo a encontrar o melhor conjunto de *inputs* (MAIER *et al.*, 2010).

Uma vez definidos, os dados de entrada devem ser então divididos em três conjuntos que serão utilizados para treinamento, teste e validação.

O primeiro conjunto é utilizado para estimar os valores dos pesos sinápticos. Neste conjunto deve-se encontrar a maior quantidade de dados, contendo informações sobre todos os padrões a serem modelados.

O conjunto reservado para a realização do teste é utilizado para se determinar quando o treinamento deve ser interrompido, evitando o *overfitting*. Durante o teste, a rede é apresentada ao conjunto de teste e paralelamente ao conjunto de treinamento sendo que cada conjunto apresenta pares entrada-saída diferentes, mas com padrões semelhantes (BASHEER e HAJMEER, 2000).

A cada época de treinamento os valores dos pesos sinápticos são aplicados aos dois conjuntos de dados e o erro referente a cada um deles é calculado. O erro do conjunto de treinamento tende a diminuir indefinidamente, obtendo-se uma rede cada vez mais especializada. Já o erro do conjunto teste tende a diminuir até um ponto ótimo, quando passa a aumentar novamente (Figura 8). O número de épocas referente a esse ponto indica o momento em que o treinamento deve ser interrompido, resultando nos valores de pesos adotados definitivamente.

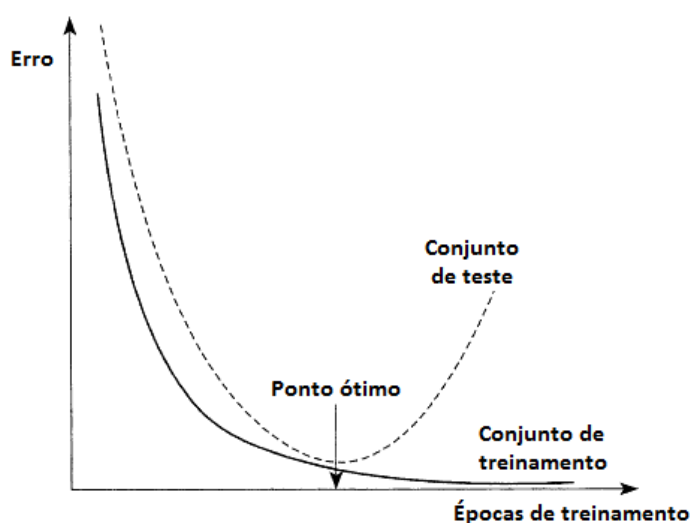


Figura 8 – Erro dos conjuntos de teste e de treinamento em função das épocas de treinamento.

Com a combinação de parâmetros e as épocas de treinamento estabelecidas, pode-se realizar a validação. Os pares entrada-saída ainda não utilizados para treinamento e teste serão então apresentados à rede. Em seguida, os resultados modelados devem passar pelo processo de desnormalização, inverso ao obtido através da Equação 3, para que possam ser comparados com os dados reais.

Por fim, os resultados simulados e observados devem ser comparados com o objetivo de verificar a eficácia do modelo. Na bibliografia analisada, os índices estatísticos mais comumente utilizados para a comparação das vazões são NASH, RMSE, MAE e DV. Esses índices são apresentados nas equações 4, 5, 6 e 7, onde  $Q_{calc}$  e  $Q_{obs}$  indicam respectivamente os valores de vazão previstos e observados,  $\bar{Q}$  é a média

das vazões observadas e  $n$  é o número de observações consideradas:

- NASH – Coeficiente de Nash-Sutcliffe: o seu valor varia de  $-\infty$  a 1, onde 1 indica um ajuste perfeito, valores entre 0 e 1 representam diferentes níveis de desempenho e valores negativos caracterizam a invalidação do modelo (MORIASI *et al.*, 2007).

$$NASH = 1 - \frac{\sum(Q_{calc} - Q_{obs})^2}{\sum(Q_{obs} - \bar{Q})^2} \quad (4)$$

- RMSE – Raiz do Erro Médio Quadrático (Root Mean Square Error): oferece uma representação da distribuição dos erros. Quanto maior for o erro, maior será a sua influência no cálculo do índice. O valor ótimo desta medida estatística é zero (CHAI e DRAXLER, 2014).

$$RMSE = \sqrt{\frac{1}{n} \sum (Q_{calc} - Q_{obs})^2} \quad (5)$$

- MAE – Erro Médio Absoluto (Mean Absolute Error): indica a magnitude do erro médio, porém, dando o mesmo peso aos erros. Valores mais próximos a zero indicam uma maior conformidade dos resultados em relação à realidade (CHAI e DRAXLER, 2014).

$$MAE = \frac{1}{n} \sum |Q_{calc} - Q_{obs}| \quad (6)$$

- DV – Desvio de Vazões: diferença percentual de volume entre as vazões observadas e previstas, sendo que valores positivos e negativos representam, respectivamente, a subestimação e a superestimação das vazões (CARDOSO, 2009).

$$DV = \frac{\sum Q_{obs} - \sum Q_{calc}}{\sum Q_{obs}} \quad (7)$$

#### 2.4. APLICAÇÕES DE REDES NEURAI ARTIFICIAIS EM HIDROLOGIA

As redes neurais são amplamente utilizadas para a modelagem hidrológica, por reproduzirem comportamentos não-lineares, compatíveis com

os fenômenos hidrológicos. Os modelos conceituais, por sua vez, são tradicionalmente aplicados por apresentarem uma base teórica mais bem consolidada. Na literatura existem diversos exemplos da aplicação desses dois tipos de modelos para a previsão de séries temporais.

Maier *et al.*(2010) realizaram uma revisão de 210 artigos publicados entre 1999 e 2007 que faziam o uso de redes neurais artificiais. A maioria desses artigos tinha como foco a previsão de vazões em rios. No que diz respeito ao método utilizado para a seleção dos dados de entrada, 60 estudos realizaram uma abordagem analítica linear, como por exemplo, a correlação. Em relação a arquitetura da rede, a estrutura *multilayer perceptrons*, tradicionalmente empregada em estudos hidrológicos, foi a mais popular, adotada em 178 artigos. Na escolha do algoritmo de treinamento, os métodos determinísticos, locais e de primeira ordem, tais como o algoritmo *backpropagation*, foram utilizados mais frequentemente, totalizando 103 aplicações. Por fim, para a avaliação do desempenho do modelo, os critérios foram escolhidos de maneira mais homogênea, o que segundo os autores é uma boa prática, pois diferentes critérios captam diferentes características do desempenho dos modelos. Ainda assim, as métricas baseadas no quadrado dos erros, como o coeficiente de Nash-Sutcliffe e o RMSE, foram as mais amplamente utilizadas, somando 170 casos.

Cigizoglu (2003) realizou uma comparação entre o uso de redes neurais artificiais e de modelos estatísticos convencionais para a previsão, extrapolação e generalização dos dados de vazão diária em quatro rios na Turquia. Regressões multilíneas e não lineares foram utilizadas como métodos estatísticos. No experimento de extrapolação foi verificada a capacidade dos modelos em prever valores encontrados fora do intervalo de calibração. O experimento de generalização envolvia a habilidade da rede que, uma vez treinada, foi apresentada a diferentes séries temporais para a previsão de séries de vazão. Os resultados dos experimentos mostram uma notável superioridade das RNAs em relação aos métodos estatísticos nas três aplicações estudadas.

Uma comparação entre modelos hidrológicos foi realizada no Rio Raccoon (EUA) por Tokar e Markus (2000). Em seu estudo, os autores utilizaram o modelo Sacramento e quatro diferentes modelos de redes

neurais artificiais para a previsão de vazões diárias. As redes foram treinadas através do algoritmo *backpropagation* e a função tangente hiperbólica foi utilizada como função de ativação. A variação entre os modelos de RNA consistia nas diferentes quantidades de neurônios na camada intermediária e nas diferentes variáveis de entrada utilizadas. O desempenho dos modelos foi comparado considerando três diferentes cenários: chuvoso, seco e intermediário. No cenário intermediário, o modelo Sacramento obteve resultados ligeiramente melhores do que as redes neurais. No entanto, nos cenários chuvoso e seco, um dos modelos de RNA obteve o melhor desempenho, mostrando sua alta capacidade em prever eventos extremos. Este resultado pode ser explicado pela adição de  $Q_{-1}$  (vazão observada no dia anterior à previsão) como dado de entrada no modelo. Isto demonstra a importância desta variável como *input*, uma vez que sua utilização trouxe informações que não estavam contidas nas outras variáveis, melhorando a capacidade de previsão do modelo.

Bravo *et al.* (2008) fizeram o uso de redes neurais artificiais para estimar a vazão diária de afluência ao reservatório de Furnas em Minas Gerais. A bacia estudada foi dividida em cinco sub-bacias, entre elas a bacia drenada pela estação fluviométrica Ibituruna, localidade que será estudada no presente trabalho, conforme apresentado na definição da região de estudo. Segundo os autores, as variáveis de entrada relevantes para a previsão por redes neurais foram a precipitação na bacia contribuinte, a vazão nos exutórios das sub-bacias contribuintes e a vazão no local de previsão. Os resultados obtidos pelo uso de redes neurais apresentaram melhores índices de desempenho quando comparados aos obtidos com o modelo matemático atualmente utilizado pelo ONS (Operador Nacional do Sistema Elétrico).

Cabral (2013) avaliou a capacidade de previsão de vazões diárias de uma rede neural artificial, quando submetida a dados pretéritos de vazão e precipitação, para a bacia do Alto Iguaçu. A análise estatística mostrou que a correlação entre a vazão a ser prevista  $Q$  e as vazões anteriores decresce paulatinamente com o passar dos dias. O mesmo fenômeno não se verificou com a variável precipitação  $P$ , pois neste caso a maior correlação entre a

vazão  $Q$  e a precipitação foi observada em  $P_{-9}$ .

O trabalho realizado por Essenfelder (2009) utilizou RNAs para a previsão de curto prazo de vazões diárias no rio Tibagi (PR). Diferentes conjuntos de dados de entrada foram apresentados à rede, que teve seu desempenho avaliado para horizontes de previsão de 1 a 7 dias. A variação nos conjuntos consistia na natureza e na quantidade dos dados. Os dados de precipitação e vazão foram obtidos em posições remotas e utilizados de forma isolada ou simultânea, com informações pretéritas de 10, 20, ou 30 dias. Previsões com maior acurácia foram geradas para horizontes menores. As previsões de 1 e 2 dias foram consideradas adequadas, as previsões de 3 e 4 dias foram consideradas aceitáveis e as previsões de 5, 6 e 7 dias tiveram resultados insatisfatórios. De maneira geral, os experimentos que consideraram informações de precipitação e vazão conjuntamente apresentaram resultados mais satisfatórios. No entanto, para o horizonte de previsão de 1 dia, as informações de precipitação foram consideradas irrelevantes. Na maioria dos casos, o treinamento da rede mostrou-se mais eficaz quando foram consideradas informações pretéritas de pelo menos 20 dias. Contudo, para o horizonte de previsão de 1 dia, os experimentos que consideravam apenas informações pretéritas de 10 dias tiveram melhor desempenho. A adição de outras informações, neste caso, não agregou informações significativas.

### **3. METODOLOGIA**

#### **3.1. CARACTERIZAÇÃO DA REGIÃO DE ESTUDO**

A bacia selecionada para o estudo é a bacia do Rio das Mortes, inserida na mesorregião do Campo das Vertentes, localizada no sul do Estado de Minas Gerais. A bacia do Rio das Mortes pertence à sub-bacia 61, segundo a codificação adotada pela ANA (Agência Nacional de Águas), e pertence à bacia do Rio Grande.

A região analisada tem uma área de drenagem de 6.155 km<sup>2</sup> e compreende o Rio das Mortes desde a sua nascente, no município de Barbacena, até a estação fluviométrica presente na cidade de Ibituruna. O



trecho de rio estudado apresenta uma extensão de aproximadamente 180 km e tem como principal afluente o Rio Elvas.

A classificação de Koppen indica que o clima da região pertence às classes Cwa e Cwb, considerado como temperado úmido com inverno seco e verão moderadamente quente (SÁ JÚNIOR, 2009).

Segundo o ATLAS (2013), a média anual da vazão diária registrada na estação de Ibituruna é de 113,7 m<sup>3</sup>/s e a precipitação anual na bacia varia de 1285 a 1420 mm.

A localização da bacia do Rio das Mortes está representada na Figura 9.

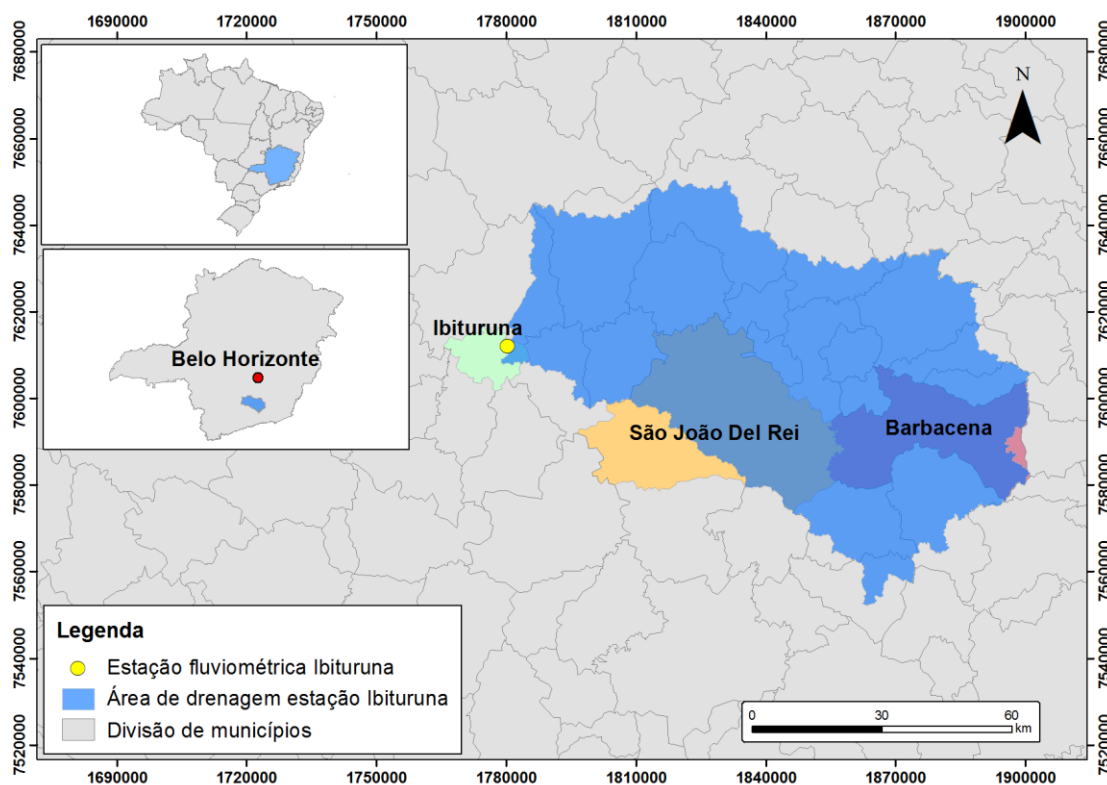


Figura 9 - Localização da bacia hidrográfica do Rio das Mortes.

### 3.2. CONFIGURAÇÃO DO MODELO SACRAMENTO

Cabe ressaltar que a calibração do modelo Sacramento da Bacia do Rio das Mortes, realizada pelo CEHPAR e descrita abaixo, é apenas uma parte do Projeto de P&D citado anteriormente. Esse projeto envolve a modelagem chuva-vazão em outras bacias hidrográficas, e ainda, a utilização de outros modelos hidrológicos para a implantação do sistema integrado de

operação de reservatórios.

### 3.2.1. Seleção e controle de qualidade dos dados de entrada

A seleção das estações pluviométricas e fluviométricas utilizadas para a calibração do modelo chuva-vazão considerou estações convencionais com períodos de dados adequados. A Tabela 2 apresenta o nome, o código, as coordenadas geográficas e o período de observação das estações selecionadas para a aquisição dos dados. Todas estas estações apresentavam um intervalo de um dia entre as medições.

Tabela 2 - Estações selecionadas para a aquisição de dados.

<b>Tipo</b>	<b>Código</b>	<b>Nome</b>	<b>Latitude</b>	<b>Longitude</b>	<b>Período de observação</b>
Pluviométrica	2143006	Barroso	-21:11:14	-43:58:48	01/05/1941 - 01/02/2013
	2143005	Campolide	-21:16:44	-43:49:12	01/05/1941 - 01/04/2013
	2043018	Carandaí	-20:57:21	-43:48:03	01/07/1941 - 01/04/2013
	2143008	Ibertioga	-21:25:38	-43:57:34	01/05/1941 - 01/02/2013
	2144023	Ibituruna	-21:08:35	-44:44:16	01/02/1967 - 01/02/2013
	2144009	Porto do Elvas	-21:09:54	-44:08:08	01/06/1941 - 01/02/2013
	2144002	Porto Tiradentes	-21:07:21	-44:13:57	01/01/1941 - 01/03/2013
	2044038	Resende Costa	-20:55:25	-44:14:18	01/08/1975 - 01/05/2013
	2044050	São Tiago	-20:53:50	-44:30:01	01/02/1984 - 01/05/2013
	2143009	Usina Barbacena	-21:16:29	-43:52:43	01/08/1941 - 01/05/2013
	2144020	Usina São João Del Rei	-21:03:20	-44:12:32	01/01/1961 - 01/02/2013
	2143007	Vargem do Engenho	-21:11:34	-43:36:50	01/05/1941 - 01/05/2013
	2144024	Vila Rio das Mortes	-21:11:22	-44:19:44	01/06/1967 - 01/02/2013
	2143062	Santos Dumont	-21:25:52	-43:33:11	01/05/1989 - 01/04/2014
	Fluviométrica	61135000	Ibituruna	-21:08:33	-44:44:23
61122000		Vila Rio das Mortes	-21:11:44	-44:19:43	01/07/1940 - 01/05/2010
61115000		Usina São João Del Rei	-21:03:09	-44:12:40	01/09/1939 - 01/05/2010
61107000		Porto Tiradentes	-21:07:20	-44:14:00	01/10/1934 - 01/12/2002

A Figura 10 mostra a localização das estações pluviométricas e fluviométricas selecionadas.

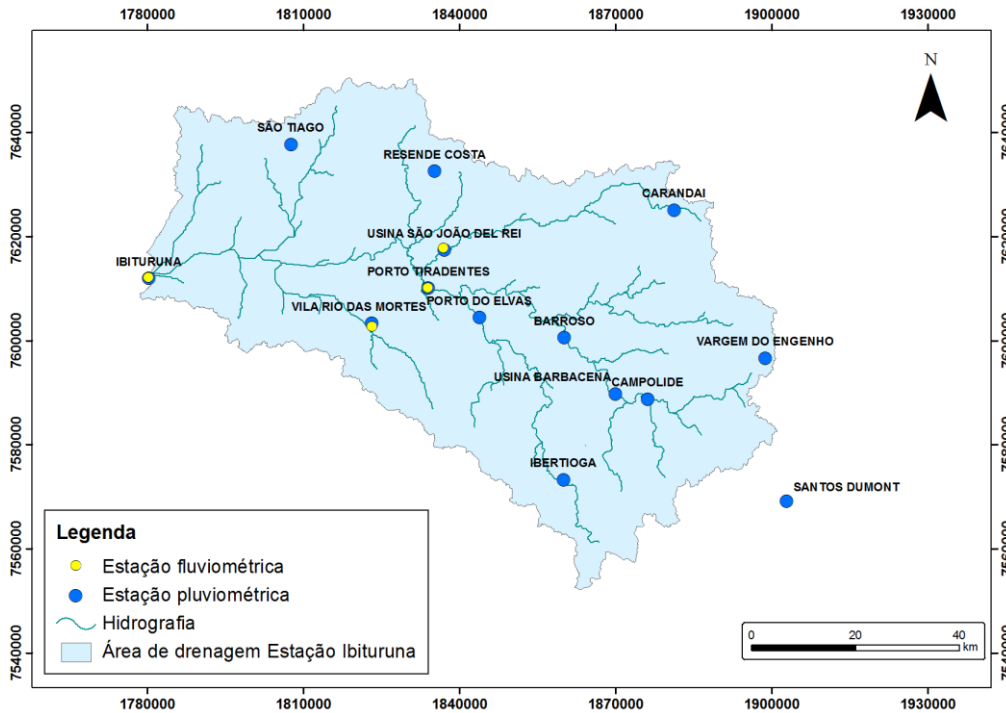


Figura 10 – Localização das estações pluviométricas e fluviométricas de interesse na bacia do Rio das Mortes.

A consistência dos dados pluviométricos se baseou nas curvas duplo cumulativas. Dessa maneira, os períodos ou estações de dados duvidosos foram descartados. A precipitação média em cada uma das sub-bacias consideradas na discretização foi obtida pelo método dos Polígonos de Thiessen. Entre os métodos utilizados na consistência dos dados fluviométricos estão a plotagem dos hidrogramas simultâneos e a avaliação das curvas de descarga (LACTEC, 2014).

A série de dados de evapotranspiração EVT foi estimada pelo método de Thornthwaite (THORNTHWAITE, 1948), expresso pelas equações 8, 9 e 10:

$$EVT = 16 \left( 10 \cdot \frac{T_j}{i} \right)^a \quad (8)$$

$$i = \sum_{j=1}^{12} \left( \frac{T_j}{5} \right)^{1,514} \quad (9)$$

$$a = 0,000000675 \cdot i^3 - 0,0000771 \cdot i^2 + 0,0179 \cdot i + 0,49239 \quad (10)$$

Onde:

$i$  = índice térmico anual ou índice de calor anual

$T_j$  = temperatura média mensal do mês  $j$

Os dados mensais de temperatura foram obtidos do site do INMET. Assim, o resultado da evapotranspiração consiste em médias mensais de longo período. Para obter os valores na escala diária foi utilizado um modelo de desagregação com uma função interpoladora definida pela série de Fourier (KAVISKI *et al.*, 2005).

Os dados do SIG (Sistema de Informação Geográfica) utilizados na esquematização do modelo foram disponibilizados pela Companhia Energética de Minas Gerais (CEMIG).

### 3.2.2. Configuração do modelo no software SOBEK

A esquematização da bacia hidrográfica dentro do software SOBEK requer o pré-processamento de dados do Modelo Digital de Elevação (MDE). Esta etapa foi realizada com o auxílio de ferramentas do SIG e teve como objetivo o delineamento dos cursos d'água, a delimitação e o cálculo das áreas das microbacias e a obtenção dos centróides das mesmas. O modelo, baseado nessas informações, é formado através de nós e *links* de propagação.

Os nós executam diferentes funções dependendo do seu tipo, sendo que os dois mais utilizados são os nós do tipo "Flow-connection" que têm a função de unir dois ou mais *links* de propagação, e nós do tipo "Sacramento", que são localizados nos centróides das microbacias e utilizados para a entrada dos parâmetros do modelo Sacramento. Com o mapa da hidrografia da bacia como base, foi possível criar uma rede com nós e *links* com um comportamento semelhante ao da rede fluvial na realidade. A Figura 11 mostra a esquematização da bacia do Rio das Mortes no SOBEK.

A etapa de discretização da bacia hidrográfica delimitou microbacias com o objetivo de alcançar uma adequada representação da chuva média da bacia do Rio das Mortes, visto que esta será a configuração final a ser adotada pelo projeto de P&D. A precipitação média nas microbacias foi obtida através da média ponderada da precipitação, calculada pelo Método de

Thiessen, pela área de drenagem das referidas microbacias.

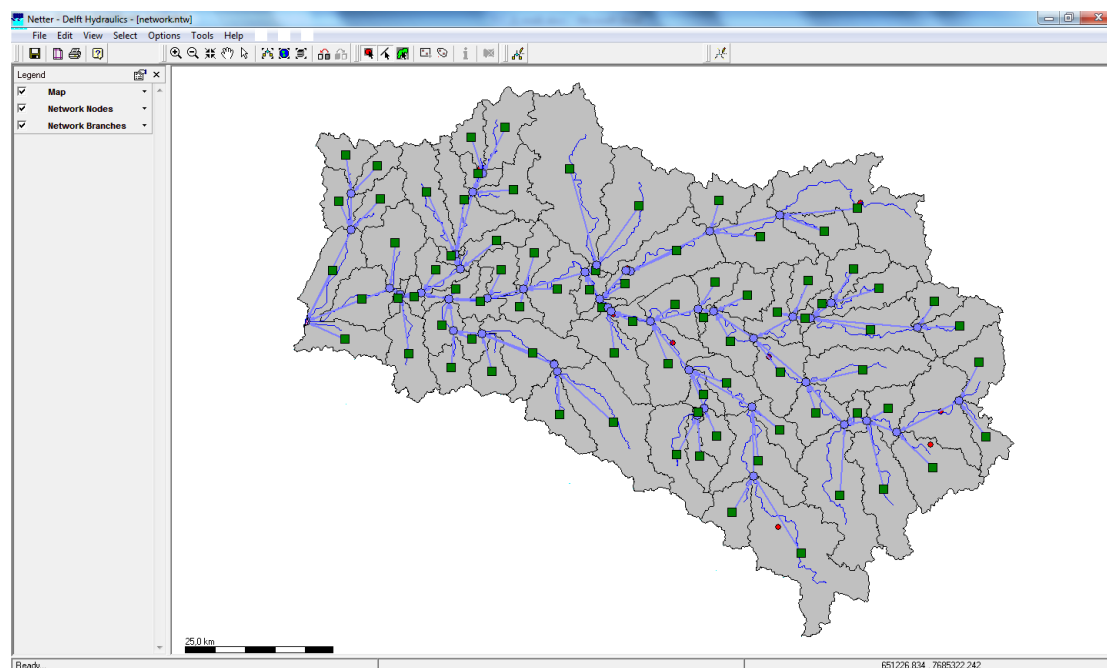


Figura 11 – Configuração do modelo Sacramento no software SOBEK.

### 3.2.3. Calibração e validação do modelo Sacramento

A estimativa dos parâmetros do modelo Sacramento foi realizada com base nos dados históricos. O período para a calibração foi de 01/01/2004 a 31/12/2008 e o de validação foi de 01/01/2009 a 31/12/2012.

Para a calibração do modelo hidrológico da bacia estudada foram utilizados os otimizadores automáticos do pacote RRL. Foram ainda testados todos os métodos de otimização. Entretanto, aquele que demonstrou melhor desempenho foi o do algoritmo genético.

A função objetivo utilizada foi o critério de Nash-Sutcliffe. Além dele, outros índices estatísticos foram utilizados para verificação da eficácia do modelo.

A Figura 12 mostra a interface do programa RRL e os parâmetros encontrados na calibração da bacia do Rio das Mortes.

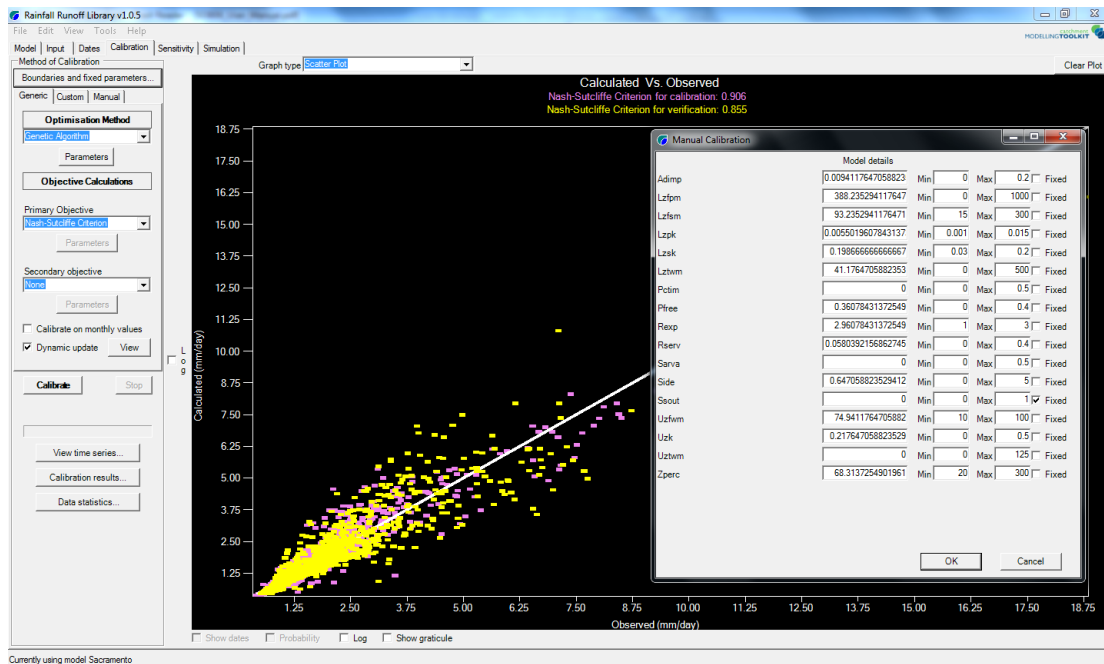


Figura 12 - Calibração da bacia do Rio das Mortes no software RRL.

Ressalta-se que a calibração automática dos parâmetros do modelo Sacramento é dependente de seus valores iniciais, bem como de seu intervalo de variação. Estes valores iniciais podem variar de acordo com eventos selecionados e também com bacia estudada. Foi considerada a faixa de variação dos parâmetros obtida na literatura (ANDERSON *et al.*, 2006; KOREN *et al.*, 2000).

A Figura 13 mostra a representação das duas abordagens de discretização analisadas. Uma mais refinada abrangendo bacias menores (semi-distribuída), e outra considerando a bacia sem divisões (concentrada). Na abordagem semi-distribuída a calibração dos parâmetros do modelo Sacramento no software RRL foi realizada separadamente para cada uma das sub-bacias. Desta forma, espera-se uma melhor representação da variabilidade espacial do sistema. Apesar da calibração das bacias ter considerado a aplicação de modelos concentrados no RRL, dentro do SOBEK estes parâmetros foram aplicados considerando a divisão da bacia em microbacias, pois esta será a configuração final a ser adotada pelo projeto.

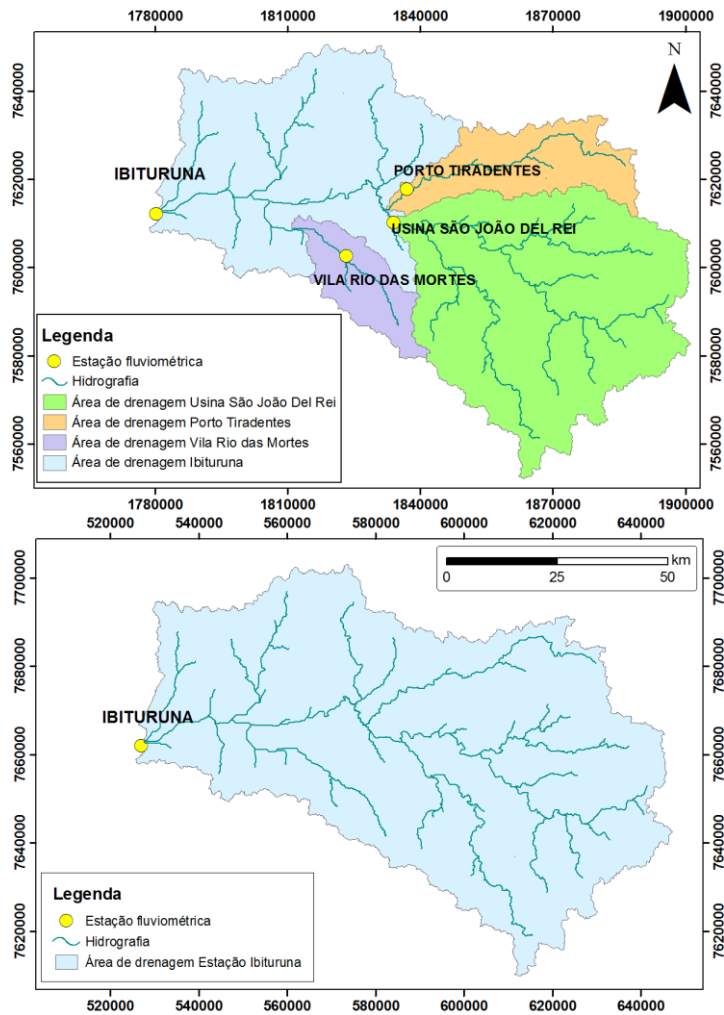


Figura 13 – Discretização em sub-bacias conforme configuração semi-distribuída (acima) e concentrada (abaixo).

Após serem definidos como mais adequados pela ferramenta RRL, os parâmetros foram inseridos no SOBEK para a operação do modelo. As séries de precipitação e evapotranspiração foram introduzidas no software como variáveis de entrada para a geração da série de vazão. O modelo leva em torno de 30 minutos para realizar a etapa de operação.

### 3.3. CONFIGURAÇÃO DA REDE NEURAL ARTIFICIAL

#### 3.3.1. Seleção e controle de qualidade dos dados de entrada

Inicialmente foi realizada uma análise preliminar das estações pluviométricas presentes na sub-bacia que tem como exutório a estação fluviométrica Ibituruna. Foram selecionadas as mesmas estações

pluviométricas utilizadas para a calibração do modelo Sacramento (Tabela 2), com exceção da estação Santos Dumont, não utilizada neste caso por se localizar fora da bacia hidrográfica estudada. Considerando as estações fluviométricas, a única estação utilizada pelo modelo de RNA foi a de Ibituruna. As séries históricas de precipitação e vazão foram obtidas através do portal HIDROWEB da Agência Nacional de Águas (ANA).

Uma análise preliminar dos dados foi feita através de inspeção visual das séries históricas e dos histogramas de cada estação. Os dados obtidos foram consistidos pela ANA até o ano de 1999 e 2005, respectivamente para as estações pluviométricas e fluviométrica. Foi realizada uma análise de consistência para a utilização dos dados do período não consistido.

Primeiramente foram selecionados os maiores valores registrados no período consistido considerando cada mês separadamente. Esses valores foram considerados como limite superior para o período não consistido. Em seguida, verificou-se a incidência de valores duvidosos, acima dos limites de cada mês. Foram encontrados em média cinco valores acima do limite em cada estação. A quantidade de valores duvidosos foi considerada muito pequena em relação ao total de dados da amostra, sendo assim, esses valores foram excluídos das séries históricas sem que houvesse comprometimento dos dados.

O período de estudo foi definido através da seleção da maior série histórica que contasse com dados de todas as estações. O período selecionado foi de 20/02/1984 a 31/12/2012, no qual foram constatadas falhas. Para o treinamento e operação da rede foram utilizados, respectivamente, 10.574 e 1.492 pares entrada-saída.

A precipitação média da bacia foi obtida através do método dos Polígonos de Thiessen. Somente foram considerados os dias que não apresentaram falhas em nenhuma das estações pluviométricas. Por se tratar de uma série histórica muito longa, a não utilização destes dados não foi considerado um problema.

As séries de vazões obtidas através da ANA e da CEMIG foram confrontadas, por meio de inspeção visual percebeu-se que a série da ANA tinha dados incoerentes em seu período final. Para contornar esse problema, a série da ANA foi substituída pela série da CEMIG a partir de 18/03/2004.



### 3.3.2. Análise estatística dos dados

A seleção dos dados de entrada para o treinamento da RNA é uma etapa importante e que define o desempenho da rede.

Primeiramente foram calculados os índices de correlação linear entre a vazão a ser prevista,  $Q$ , e as vazões de dias anteriores, assim como entre  $Q$  e os dados de precipitação de cada estação isoladamente e da precipitação média. A Tabela 3 apresenta a correlação entre os dados.

Tabela 3 - Correlação linear entre a vazão a ser prevista e as vazões de dias anteriores, as precipitações dos dias anteriores e a precipitação média.

Vazão		Q-1	Q-2	Q-3	Q-4	Q-5	Q-6	Q-7	Q-8	Q-9	Q-10	
	Ibituruna		0,9462	0,8736	0,8107	0,7588	0,7168	0,6830	0,6566	0,6390	0,6277	0,6189
Precipitação	P	P-1	P-2	P-3	P-4	P-5	P-6	P-7	P-8	P-9	P-10	
	Barroso	0,2889	0,4018	0,3683	0,3118	0,3010	0,2593	0,2238	0,2027	0,1891	0,1735	0,1967
	Campolide	0,3000	0,3832	0,3869	0,3646	0,3210	0,2951	0,2776	0,2350	0,2143	0,2153	0,2416
	Carandaí	0,2607	0,3395	0,3420	0,3034	0,2749	0,2626	0,2376	0,2127	0,2188	0,1904	0,1903
	Ibertioga	0,2840	0,3924	0,3748	0,3149	0,2833	0,2599	0,2426	0,1994	0,1824	0,1815	0,2214
	Ibituruna	0,3434	0,3898	0,3579	0,3454	0,3052	0,2655	0,2582	0,2320	0,2048	0,2037	0,2365
	Porto do Elvas	0,2775	0,3963	0,3577	0,3214	0,2797	0,2391	0,2191	0,2096	0,1893	0,1795	0,1954
	Porto Tiradentes	0,3643	0,4603	0,4056	0,3425	0,3174	0,2920	0,2625	0,2510	0,2542	0,2674	0,2860
	Resende Costa	0,1592	0,3202	0,4036	0,3371	0,3043	0,2682	0,2412	0,2136	0,1896	0,2001	0,2002
	São Tiago	0,3179	0,4036	0,3633	0,3083	0,2849	0,2755	0,2367	0,2322	0,2179	0,1891	0,2367
	Usina Barbacena	0,2750	0,3952	0,3754	0,3390	0,2972	0,2664	0,2482	0,1990	0,1798	0,1847	0,2028
	Usina São João Del Rei	0,3652	0,4440	0,3723	0,3378	0,2889	0,2723	0,2470	0,2187	0,1975	0,1928	0,2129
	Vargem do Engenho	0,2315	0,3342	0,3250	0,2857	0,2569	0,2225	0,2122	0,1838	0,1960	0,1983	0,1907
	Vila Rio das Mortes	0,3028	0,4159	0,3586	0,3218	0,3010	0,2609	0,2783	0,2554	0,2413	0,2475	0,2512
	Precipitação Média	0,3979	0,5383	0,5075	0,4472	0,4048	0,3649	0,3388	0,3040	0,2847	0,2762	0,3051

Percebe-se que a correlação é maior para  $Q_{-1}$  e  $P_{-1}$ , diminuindo com o passar dos dias. Isto pode ser explicado pelo fato dos processos hidrológicos apresentarem “memória”, isto é, seus fenômenos estão diretamente relacionados com eventos que ocorreram nos dias anteriores. A análise da tabela demonstra que a correlação entre a vazão  $Q$  e a vazão dos dias anteriores é maior do que a correlação entre a vazão  $Q$  e a precipitação. Ainda, observa-se que em muitos casos a vazão  $Q$  tem uma menor

correlação com a precipitação do dia ( $P$ ) do que com a precipitação dos dias anteriores ( $P_{-1}, P_{-2}, P_{-3}, \dots$ ). Além disso, a precipitação média teve maior correlação com a vazão do que as precipitações quando analisadas separadamente. Nota-se também que as estações que apresentam maior correlação com o dia  $P$  são as que se encontram geograficamente mais próximas ao exutório.

### 3.3.3. Treinamento e operação da Rede Neural Artificial

Para a realização do treinamento e da operação da RNA foram utilizados três programas em linguagem de programação FORTRAN, de autoria do Prof. Dr. Ricardo Carvalho de Almeida. Para a utilização dos programas foi necessário estruturar os dados de entrada em arquivos compatíveis. Foram utilizados somente os dias que não apresentaram falhas em nenhum dos dados de entrada e de saída. Em caso de falha em qualquer um dos dados de vazão ou precipitação do dia a ser calculado ou dos dias anteriores o par entrada-saída não era utilizado.

No primeiro programa os dados de entrada são normalizados entre -1 e 1, segundo a Equação 3. Em seguida, um algoritmo aleatório é utilizado para separação dos dados em grupos de treinamento, teste e validação, seguindo as porcentagens de 84%, 15% e 1%, respectivamente. Os dados utilizados para o treinamento compreendem o período inteiro de 1984 a 2012.

Os dados gerados na etapa anterior foram inseridos no segundo programa para treinamento da rede. Esse programa consiste numa rede MLP com três camadas cujo algoritmo de aprendizagem é o *backpropagation*. O conjunto ótimo de pesos sinápticos encontrado durante o treinamento é salvo pelo próprio programa e registrado em um arquivo tipo texto. Em outro arquivo de saída são compilados os índices calculados para avaliação da qualidade de previsão da rede, entre eles o coeficiente de Nash-Sutcliffe, o Desvio de Vazões e o RMSE para os conjuntos de validação e de teste.

Para cada uma das configurações testadas, os procedimentos acima descritos para a separação de dados e o treinamento da rede foram realizados cinco vezes. Foi então calculada a média dos índices de avaliação gerados pelo programa, os resultados obtidos estão disponíveis no Apêndice

A. A média do coeficiente de Nash-Sutcliffe foi adotada como referência para a escolha das melhores configurações.

O treinamento da RNA foi realizado em etapas. Inicialmente, foi definida a topologia da rede, isto é, o número de neurônios na camada intermediária, em seguida, buscou-se a determinação das variáveis de entrada e por fim, a quantidade de épocas de treinamento.

Os testes realizados na primeira etapa estão representados na Tabela 4, onde o caracter asterisco (\*) indica o número de neurônios utilizados na camada intermediária. Foram efetuados testes com cinco, dez e quinze neurônios intermediários, todos com 1000 épocas de treinamento. Nesta etapa enfatizou-se a definição da topologia, mas também foram avaliadas as variáveis de entrada. As configurações testadas continham dados de vazão e precipitação, isoladamente.

Tabela 4 - Testes para a definição da topologia da RNA.

T_*_01	Q-1
T_*_02	Q-1, Q-2
T_*_03	Q-1, Q-2, Q-3
T_*_04	Q-1, Q-2, Q-3, Q-4
T_*_05	Q-1, Q-2, Q-3, Q-4, Q-5
T_*_06	Q-1, Q-2, Q-3, Q-4, Q-5, Q-6
T_*_07	Q-1, Q-2, Q-3, Q-4, Q-5, Q-6, Q-7
T_*_08	Q-1, Q-2, Q-3, Q-4, Q-5, Q-6, Q-7, Q-8
T_*_09	Q-1, Q-2, Q-3, Q-4, Q-5, Q-6, Q-7, Q-8, Q-9
T_*_10	Q-1, Q-2, Q-3, Q-4, Q-5, Q-6, Q-7, Q-8, Q-9, Q-10
T_*_11	P
T_*_12	P,P-1
T_*_13	P,P-1,P-2
T_*_14	P,P-1,P-2,P-3
T_*_15	P,P-1,P-2,P-3,P-4
T_*_16	P-1
T_*_17	P-1,P-2
T_*_18	P-1,P-2,P-3
T_*_19	P-1,P-2,P-3,P-4

A determinação dos dados de vazão e precipitação relevantes para a previsão da vazão em uma bacia hidrográfica não segue uma regra fixa. Sendo assim, diferentes configurações devem ser testadas em busca do

conjunto de dados mais representativo em cada caso. A Tabela 5 mostra as configurações testadas na segunda etapa do treinamento com diferentes combinações contendo simultaneamente, dados de vazão e precipitação. Foi selecionada a topologia com cinco neurônios na camada intermediária, a justificativa desta escolha será apresentada posteriormente.

Tabela 5 - Testes para a definição das variáveis de entrada.

T_5_20	Q-1,P	T_5_38	Q-1,Q-2,Q-3,P,P-1,P-2,P-3,P-4
T_5_21	Q-1,Q-2,P	T_5_39	Q-1,Q-2,Q-3,Q-4,P,P-1,P-2,P-3,P-4
T_5_22	Q-1,Q-2,Q-3,P	T_5_40	Q-1,P-1
T_5_23	Q-1,Q-2,Q-3,Q-4,P	T_5_41	Q-1,Q-2,P-1
T_5_24	Q-1,P,P-1	T_5_42	Q-1,Q-2,Q-3,P-1
T_5_25	Q-1,Q-2,P,P-1	T_5_43	Q-1,Q-2,Q-3,Q-4,P-1
T_5_26	Q-1,Q-2,Q-3,P,P-1	T_5_44	Q-1,P-1,P-2
T_5_27	Q-1,Q-2,Q-3,Q-4,P,P-1	T_5_45	Q-1,Q-2,P-1,P-2
T_5_28	Q-1,P,P-1,P-2	T_5_46	Q-1,Q-2,Q-3,P-1,P-2
T_5_29	Q-1,Q-2,P,P-1,P-2	T_5_47	Q-1,Q-2,Q-3,Q-4,P-1,P-2
T_5_30	Q-1,Q-2,Q-3,P,P-1,P-2	T_5_48	Q-1,P-1,P-2,P-3
T_5_31	Q-1,Q-2,Q-3,Q-4,P,P-1,P-2	T_5_49	Q-1,Q-2,P-1,P-2,P-3
T_5_32	Q-1,P,P-1,P-2,P-3	T_5_50	Q-1,Q-2,Q-3,P-1,P-2,P-3
T_5_33	Q-1,Q-2,P,P-1,P-2,P-3	T_5_51	Q-1,Q-2,Q-3,Q-4,P-1,P-2,P-3
T_5_34	Q-1,Q-2,Q-3,P,P-1,P-2,P-3	T_5_52	Q-1,P-1,P-2,P-3,P-4
T_5_35	Q-1,Q-2,Q-3,Q-4,P,P-1,P-2,P-3	T_5_53	Q-1,Q-2,P-1,P-2,P-3,P-4
T_5_36	Q-1,P,P-1,P-2,P-3,P-4	T_5_54	Q-1,Q-2,Q-3,P-1,P-2,P-3,P-4
T_5_37	Q-1,Q-2,P,P-1,P-2,P-3,P-4	T_5_55	Q-1,Q-2,Q-3,Q-4,P-1,P-2,P-3,P-4

Como o modelo Sacramento utiliza dados de evapotranspiração para a previsão de vazões, questionou-se qual seria o comportamento da rede com a inserção dessa variável como entrada. Para a solução desta incerteza, a série de evapotranspiração foi adicionada às oito configurações que apresentaram os maiores índices do coeficiente de Nash-Sutcliffe. A evapotranspiração estimada do dia foi utilizada como um neurônio adicional na camada de entrada.

A última etapa consistiu na definição da quantidade de épocas de treinamento, buscando a diminuição dos erros e evitando o *overfitting*. Foram realizados testes com até 100.000 épocas de treinamento para oito as configurações selecionadas. O programa leva em média 10 segundos para efetuar 1.000 iterações.

Após a realização de todas essas etapas, as oito melhores configurações passaram então para a fase de operação da rede. O procedimento realizado na etapa de treinamento foi repetido para essas configurações, dessa vez, realizado somente com dados do período de 1984 a 2008. Assim, a rede não seria corrompida por ter sido treinada com dados do período de operação, sendo este definido de 2009 a 2012 para posterior comparação com o modelo Sacramento. No entanto, este período apresentava uma grande falha na série de precipitação. A série foi então preenchida com os valores utilizados pelo CEHPAR. Finalmente, foi executado o terceiro programa, para operação da rede. Um novo arquivo de entrada foi preparado, somente com dados para o período de operação. O conjunto de sinapses determinado durante o treinamento foi utilizado pelo programa para a previsão da série de vazão. Os índices de RMSE, MAE, NASH e Desvio de Vazões foram calculados para a comparação das séries prevista e observada.

## **4. RESULTADOS E DISCUSSÃO**

### **4.1. MODELO SACRAMENTO**

A Tabela 6 mostra os valores do coeficiente de Nash-Sutcliffe para todas as sub-bacias e os parâmetros calibrados através do software RRL e aplicados ao software SOBEK.

No software RRL a configuração concentrada obteve o maior coeficiente de Nash-Sutcliffe, somado a isto, observou-se o resultado pouco satisfatório na sub-bacia Vila Rio das Mortes. Mesmo assim, a configuração semi-distribuída foi adotada, pois quando operada no software SOBEK, esta configuração obteve melhor desempenho se comparada com o modelo concentrado. A divergência de resultados entre dois softwares utilizados pode ser explicada pela utilização de um conjunto de dados independente para a validação, neste caso, os melhores resultados foram obtidos através do modelo semi-distribuído adotado.

Tabela 6 - Valores calibrados no RRL para as sub-bacias nas configuração semi-distribuída e concentrada com seus respectivos coeficientes de Nash-Sutcliffe.

Parâmetro	Faixa de variação	Ibituruna	Usina São João Del Rei	Porto Tiradentes	Vila Rio das Mortes	Ibituruna
		Semi-distribuído			Concentrado	
Localização	-	Incremental	Cabeceira	Cabeceira	Cabeceira	
ADIMP	0 - 0,2	0,01	0,01	0,005	0	0,01
LZFBM	0 - 1000	388	388	635	996	505
LZFSM	15 - 300	93	236	300	54	251
LZPK	0,001 - 0,015	0,005	0,0056	0,004	0,0015	0,004
LZSK	0,03 - 0,2	0,2	0,2	0,2	0,06	0,2
LZTWM	0 - 500	41	250	250	500	223
PCTIM	0 - 0,5	0	0,023	0,016	0,03	0
PFREE	0 - 0,4	0,36	0,37	0,4	0,36	0,39
REXP	1 - 3	3	2,8	2,4	1	2,6
RSERV	0 - 0,4	0,06	0,27	0,3	0,3	0,06
SARVA	0 - 0,5	0	0,014	0,04	0,01	0,01
SIDE	0 - 5	0,65	0,25	0,4	0	0,53
SSOUT	0 - 1	0	0,001	0,001	0,001	0,001
UZFBM	10 - 100	75	52	31	39	90
UZK	0 - 0,5	0,2	0,42	0,064	0,1	0,29
UZTWM	0 - 125	0,1	16	0,1	69	0
ZPERC	20 - 300	68	58	47	20	212
<b>NASH da calibração</b>		<b>0,906</b>	<b>0,827</b>	<b>0,901</b>	<b>0,769</b>	<b>0,927</b>
<b>NASH da validação</b>			<b>0,895</b>			<b>0,878</b>

Segundo Burnash (1995), a conveniência da utilização de otimizadores automáticos pode converter poderosas ferramentas conceituais em sistemas de caixa-preta distorcidos. Os otimizadores tem a capacidade de produzir combinações razoáveis de parâmetros, no entanto, eles podem ser conceitualmente inválidos. Assim, destaca-se a importância da restrição da modificação dos parâmetros para uma faixa que seja compatível com as bases conceituais.

#### 4.2. MODELO DE REDE NEURAL ARTIFICIAL

Os gráficos apresentados a seguir são resultado das etapas descritas no item 3.3.3 e foram ampliados para permitir uma melhor visualização.

A Figura 14 apresenta o resultado do coeficiente de Nash-Sutcliffe para as diferentes configurações indicadas na Tabela 4. As barras vermelhas, azuis e verdes representam, respectivamente, as configurações com 5, 10 e 15 neurônios da camada intermediária. No Apêndice B encontram-se os

gráficos referentes a outros coeficientes.

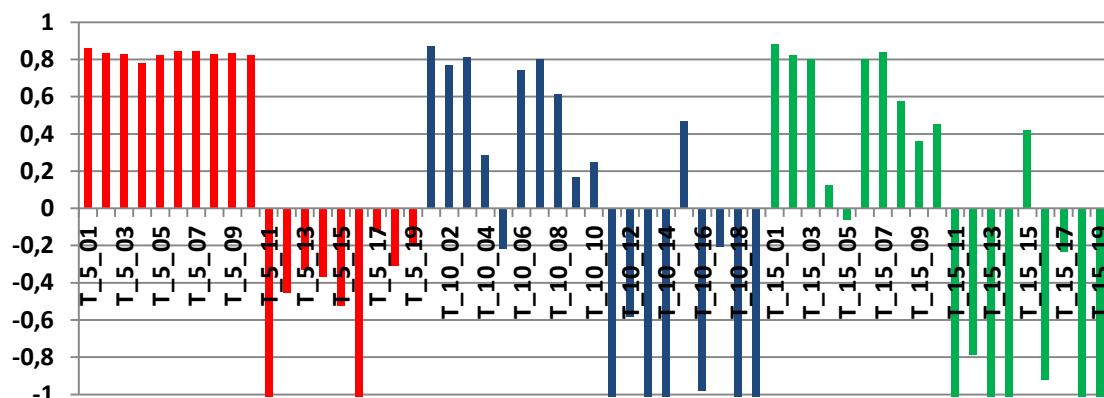


Figura 14 - Coeficiente de Nash-Sutcliffe para a escolha do número de neurônios da camada intermediária.

Através da análise do gráfico da Figura 14, percebe-se que as configurações que continham cinco neurônios na camada intermediária obtiveram um desempenho melhor em relação ao coeficiente de Nash-Sutcliffe. Além disso, as configurações que apresentavam apenas a vazão como variável de entrada alcançaram melhores resultados, quando comparadas com aquelas que tiveram apenas a precipitação como variável de entrada. De acordo com a Tabela 3 isto era previsível, pois a vazão pretérita apresenta maior relevância na modulação das vazões presente e futura.

A Figura 15 apresenta o coeficiente de Nash-Sutcliffe para as diferentes configurações apresentadas na Tabela 5. As barras vermelhas representam as oito configurações de dados de entrada para a RNA que obtiveram os melhores índices. No Apêndice C são apresentados gráficos referentes aos outros coeficientes de avaliação de qualidade de previsão da rede.

Pode-se notar uma significativa melhora dos resultados com a inserção de dados de entrada contendo vazão e precipitação simultaneamente. Por mais que os dados de precipitação tenham apresentado uma menor correlação, eles contém informações importantes para o treinamento da RNA que não são contempladas pela série de vazão.

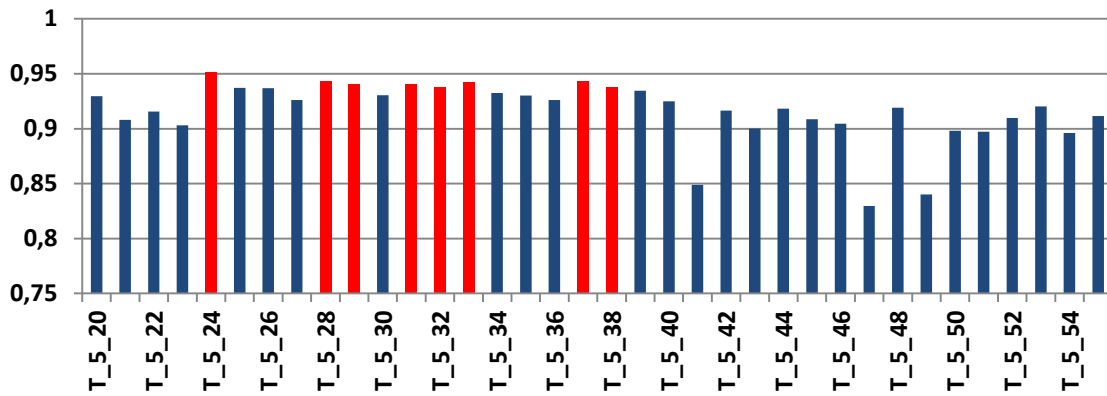


Figura 15 - Coeficiente de Nash-Sutcliffe para a definição das variáveis de entrada.

A Figura 16 mostra o resultado da inserção da variável evapotranspiração. As barras em azul representam as configurações originais e as em vermelho representam as mesmas configurações, porém com a inserção da variável evapotranspiração. Os gráficos referentes aos outros coeficientes de avaliação de qualidade de previsão da rede são apresentados no Apêndice D.

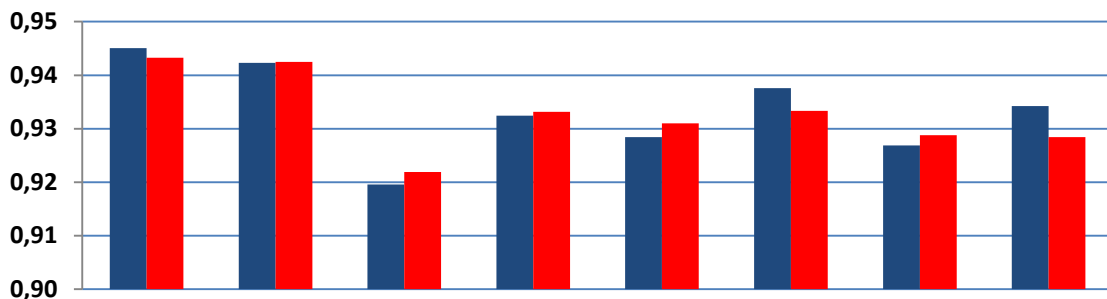


Figura 16 - Coeficiente de Nash-Sutcliffe para configurações com a ausência (azul) e presença (vermelho) da variável evapotranspiração.

A análise da figura acima denota que a inserção da evapotranspiração não interfere significativamente nos resultados do modelo de RNA, aumentando ou diminuindo o coeficiente de referência nas mesmas proporções. Pode-se atribuir esse fato à má qualidade desses dados, uma vez que a série de evapotranspiração foi obtida por meio de cálculos indiretos, e não através de medição realizada em campo.

A Figura 17 mostra as curvas de erro para os conjuntos de



treinamento (azul) e de teste (vermelho) da configuração T\_5\_38. Por meio de análise visual nota-se que com 30.000 épocas de treinamento os erros se encontram estáveis. Isso significa que o aumento do número de épocas não melhora o desempenho da rede. Não foi verificada a ocorrência de *overfitting*. Para diminuir o custo computacional, foram estabelecidas 30.000 épocas de treinamento.

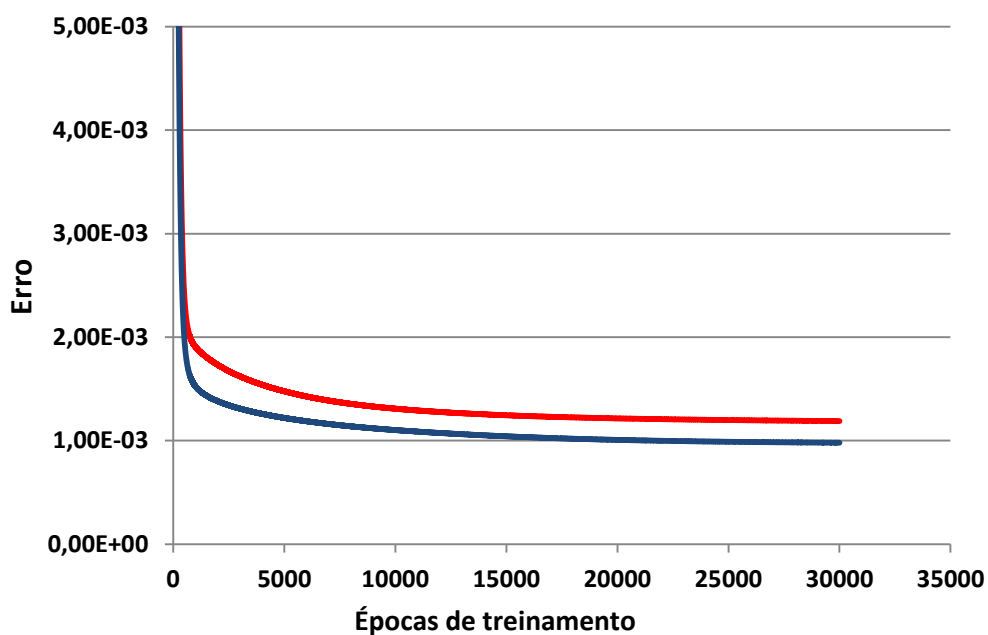


Figura 17 - Erros dos conjuntos de treinamento (azul) e de teste (vermelho) ao longo das épocas de treinamento da configuração T\_5\_38.

A Figura 7 apresenta o resultado dos índices calculados na etapa de operação da rede para as oito melhores configurações. Nota-se que todas estas configurações continham a variável *P* como entrada, ela foi considerada relevante pelo modelo de RNA mesmo apresentando menor correlação do que a precipitação de dias anteriores.

Tabela 7 - Resultado dos índices NASH, RMSE, MAE e Desvio de Vazões.

	NASH	RMSE (m³/s)	MAE (m³/s)	Desvio de vazões (%)
T_5_24	0,961	23,831	11,952	0,017
T_5_28	0,962	23,312	11,800	0,016
T_5_29	0,962	23,497	11,788	0,016
T_5_31	0,960	23,763	11,940	0,018
T_5_32	0,961	23,709	11,056	0,014
T_5_33	0,962	23,369	11,397	0,015
T_5_37	0,960	23,833	11,094	0,016
T_5_38	0,962	23,316	11,203	0,014

Entre as configurações de RNA testadas, destaca-se a configuração T\_5\_38, que além de apresentar o maior coeficiente de Nash-Sutcliffe, está entre as configurações com menor RMSE, MAE e Desvio de Vazões. Nesta configuração as oito variáveis de entrada utilizadas tiveram um efeito positivo no resultado. No entanto, a configuração T\_5\_24 obteve um desempenho equivalente com apenas três variáveis de entrada, o que permite concluir que não existe uma relação direta entre número de variáveis e o resultado esperado.

A Figura 18 apresenta o diagrama de dispersão dos dados do conjunto de validação para a configuração T\_5\_38.

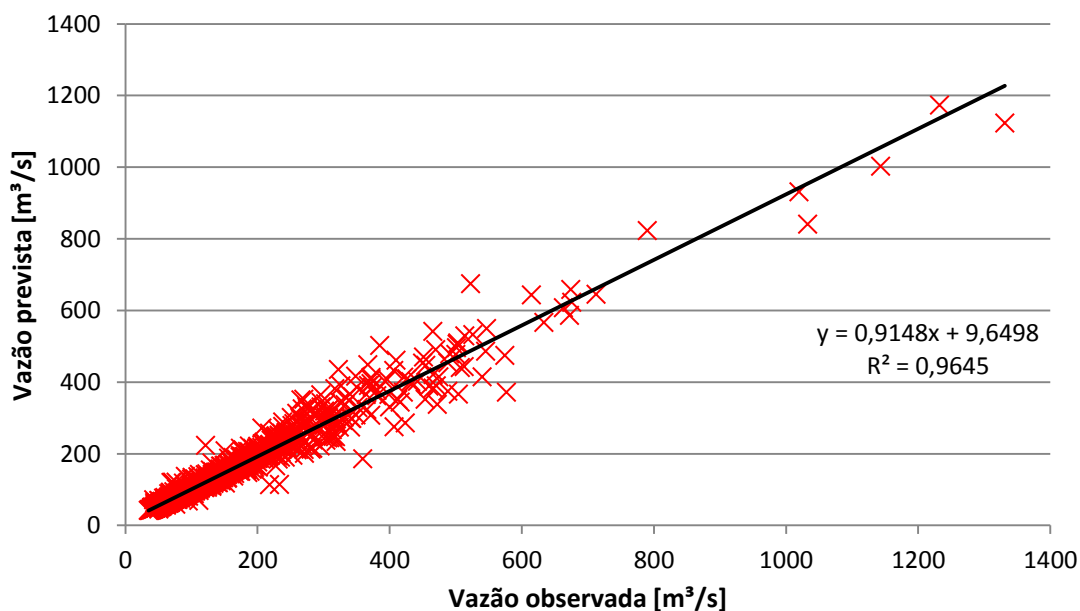


Figura 18 – Diagrama de dispersão da configuração T\_5\_38.

Nota-se que o modelo de RNA tem um bom ajuste à série observada, porém, em casos de cheia, a rede tem uma tendência a subestimar as vazões previstas. Esse resultado é natural uma vez que a rede tende a se ajustar melhor a valores de maior ocorrência. Sendo assim, os poucos eventos de cheia observados não foram suficientes para que a rede obtivesse um bom ajuste a este tipo de evento.

#### 4.3. COMPARAÇÃO DOS RESULTADOS DOS MODELOS

A Tabela 8 compara os resultados dos modelos estudados para o período de 01/01/2009 a 31/12/2012. O modelo de RNA obteve melhores resultados relativos aos índices de Nash-Sutcliffe, RMSE e MAE, o que demonstra um ajuste mais acurado entre o hidrograma previsto pela rede e o observado. O único índice em que o modelo Sacramento superou o modelo de RNA foi o Desvio de Vazões. Este índice avalia o balanço hídrico geral sem considerar a dinâmica do hidrograma. Isto significa dizer que os erros obtidos pelo modelo Sacramento, subestimando ou superestimando as vazões, se compensam de maneira quase equivalente.

Tabela 8 - Índices de avaliação da qualidade para a melhor configuração de RNA e para o modelo Sacramento.

	<b>NASH</b>	<b>RMSE (m<sup>3</sup>/s)</b>	<b>MAE (m<sup>3</sup>/s)</b>	<b>Desvio de vazões (%)</b>
<b>T_5_38</b>	0,962	23,316	11,203	0,014
<b>Sacramento</b>	0,895	39,675	22,632	-0,003

A Figura 20 e a Figura 20 mostra as curvas de vazão observada e prevista pelo modelo de RNA e pelo modelo Sacramento. Observa-se que, em geral a rede se ajusta melhor à série observada. Nos períodos de seca dos anos 2009 e 2012, o modelo Sacramento subestimou a vazão enquanto que os picos de cheia do ano de 2011 foram superestimados. Vale ressaltar que a calibração do modelo Sacramento foi realizada de modo a prover uma previsão mais acurada dos eventos de cheia, por ser o objetivo do projeto de P&D.

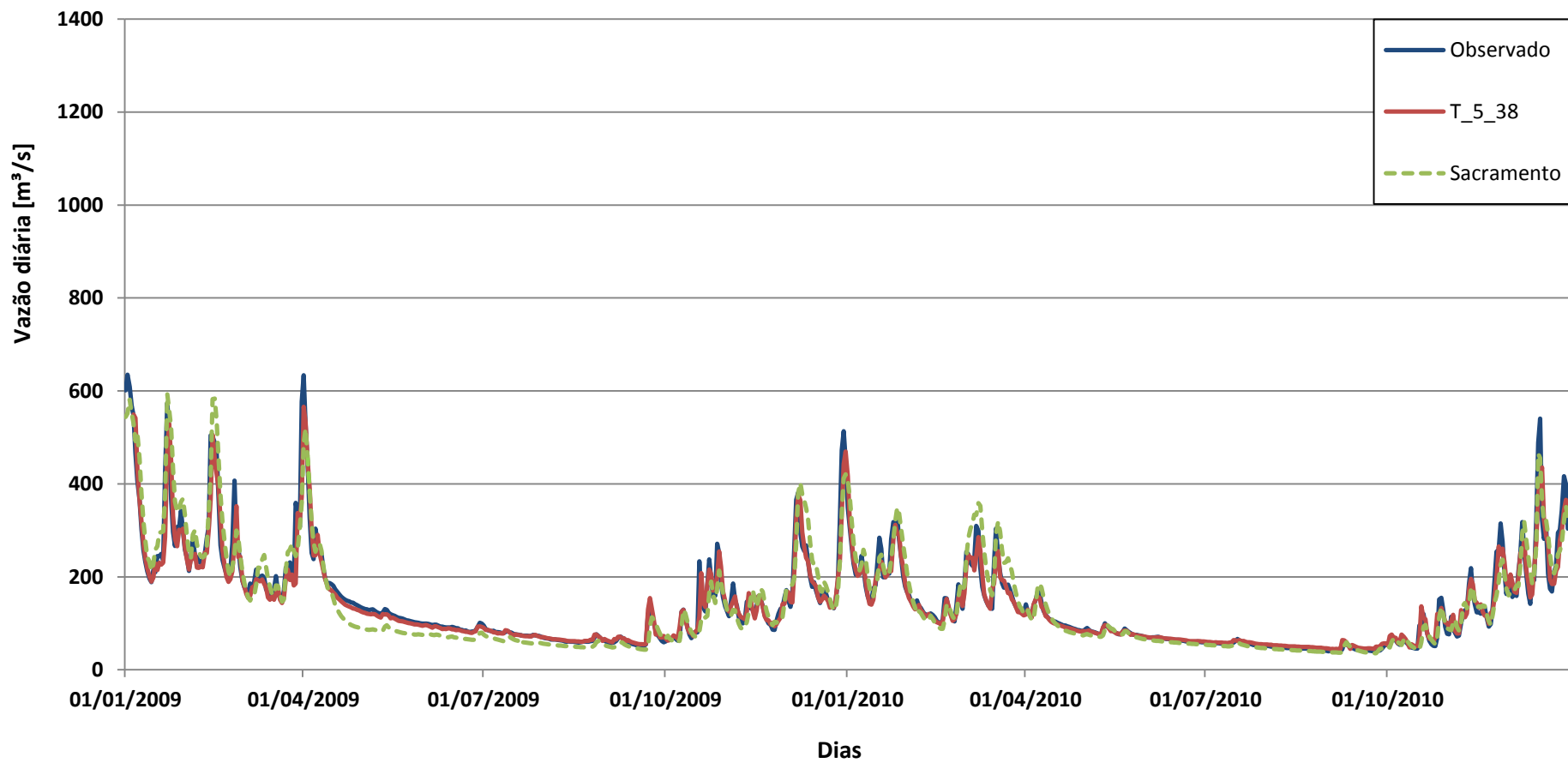


Figura 19 - Curvas de vazão observada e prevista pelo modelo de RNA e pelo modelo Sacramento de 2009 a 2010.

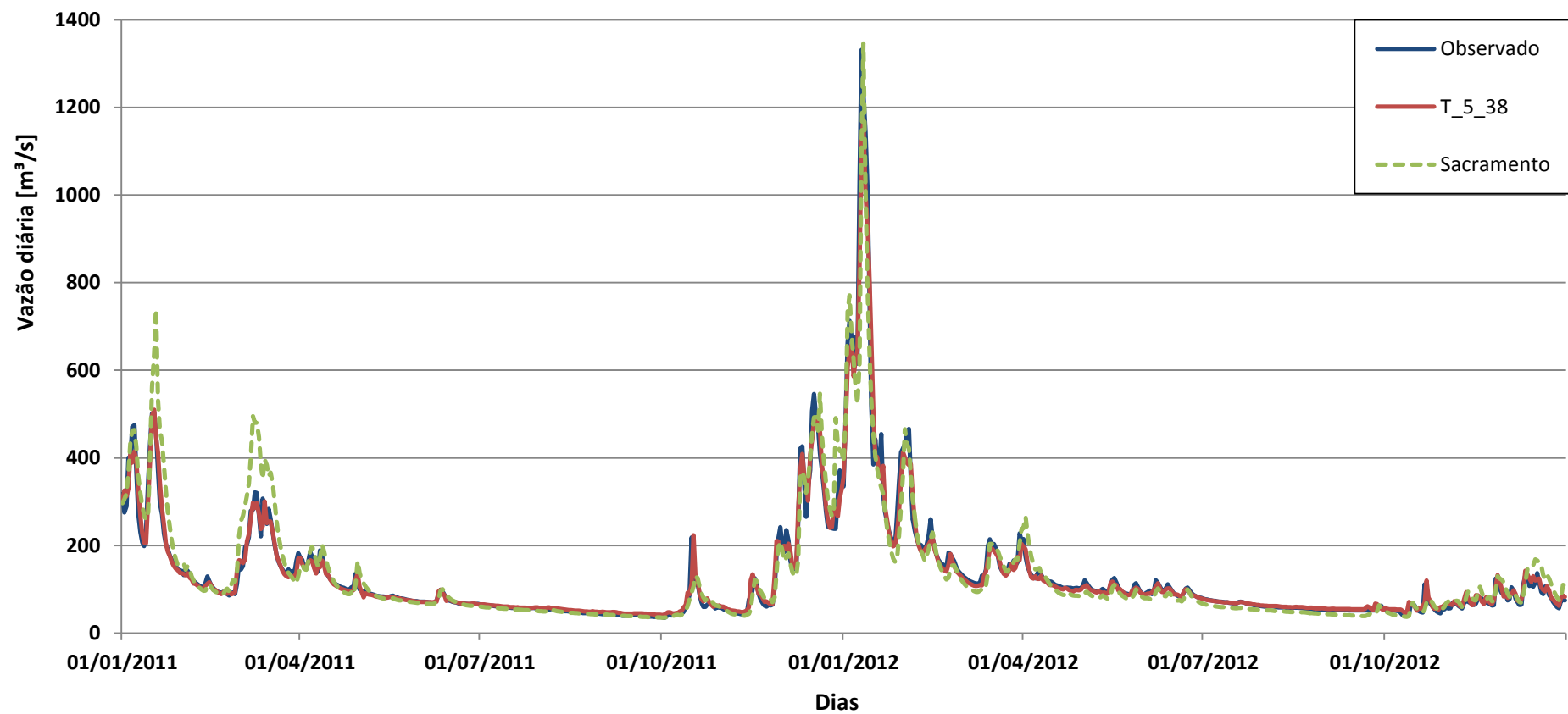


Figura 20- Curvas de vazão observada e prevista pelo modelo de RNA e pelo modelo Sacramento de 2011 a 2012.

## 5. CONCLUSÕES

Este trabalho teve como objetivo configurar um modelo empírico baseado em RNA para a previsão de vazão diária na bacia do Rio das Mortes. Nesta mesma bacia os parâmetros do modelo conceitual Sacramento haviam sido previamente determinados durante estágio acadêmico realizado pela autora em 2014 no CEHPAR. Para a calibração dos parâmetros do modelo Sacramento foi utilizada uma ferramenta auxiliar de otimização, onde foram testadas duas discretizações, uma semi-distribuída e outra concentrada. A abordagem semi-distribuída teve um melhor desempenho, pois considerou a variabilidade espacial das características da bacia.

Para a implementação do modelo de RNA foi necessário avaliar a confiabilidade das séries de dados entrada disponíveis, para tanto foi realizada uma análise de consistência dos dados. Numa segunda análise verificou-se a correlação linear das variáveis de entrada, orientando a fase de treinamento. O treinamento do modelo de RNA foi realizado em etapas, inicialmente para a definição da configuração da rede, em seguida para a seleção das variáveis de entrada e por fim para a quantificação das épocas de treinamento. Também foi testada a efetividade da inserção da variável evapotranspiração no modelo de RNA, que não gerou melhoria nos resultados. Isso pode estar relacionado com a incerteza gerada na quantificação dessa variável devido às limitações nas equações utilizadas e nas observações realizadas (TUCCI, 2005).

O resultado da análise de correlação obtido por Cabral (2013) indicou que correlação entre a vazão a ser prevista  $Q$  e a precipitação aumentou diariamente até atingir seu máximo em  $P_{-9}$ . Esta situação não foi verificada no presente trabalho, uma vez que a maior correlação foi observada entre  $Q$  e  $P_{-1}$ , decrescendo até  $P_{-10}$ . Nos dois casos, os modelos de RNA tiveram um melhor desempenho quando as variáveis de maior correlação foram consideradas, destacando a importância desta análise para a definição das variáveis de entrada.

Tokar e Markus (2000) destacaram a importância da introdução da variável  $Q_{-1}$  para a previsão de vazões. Neste mesmo sentido, os resultados

aqui apresentados mostram que além desta variável, a adição de vazões de dias anteriores contribuiu um melhor desempenho do modelo de RNA.

Pode-se questionar a validade da comparação de dois modelos conceitualmente divergentes e com diferentes variáveis de entrada. Apesar de ambos utilizarem como entrada séries de precipitação, os modelos Sacramento e de RNA utilizam adicionalmente séries de evapotranspiração e vazão, respectivamente. Ainda, a série de vazão foi utilizada de maneira indireta na calibração dos parâmetros do modelo Sacramento por otimizadores automáticos. Ressalta-se que estão sendo comparadas duas ferramentas que tem o mesmo objetivo, mas que utilizam as informações disponíveis de diferentes maneiras. Do ponto de vista da avaliação da eficácia da modelagem chuva-vazão em escala diária, é válido analisar os resultados para uma comparação objetiva das metodologias. No caso de previsão de vazão com horizonte de um dia, o modelo de RNA teve maior acurácia de resultados do que o modelo Sacramento, tendo em vista que obteve o maior coeficiente de Nash-Sutcliffe e os menores erros expressos pelos índices RMSE e MAE.

Frakes e Yu (1999) realizaram um estudo no qual compararam um modelo de RNA e o modelo conceitual GIS-HMS. Segundo os autores esses dois tipos de modelos são comparáveis por meio dos índices de avaliação de desempenho. No entanto, sua conclusão se baseia no fato de que não há um modelo melhor que o outro irrestritamente e sim que a seleção de um modelo deve ser baseada na necessidade. Os usuários do modelo devem se perguntar qual o nível de detalhamento necessário, qual a escala espacial da bacia de interesse e o quanto rapidamente a informação é requerida para tomadas de decisão ou planejamento. Se um modelo de alta resolução é essencial, então um modelo conceitual é preferível. O preço por essa informação é o investimento de mais tempo e recursos computacionais. No entanto, se a necessidade se baseia na previsão de vazão em diferentes cenários visando a obtenção de resultados em um curto período de tempo, então, modelos empíricos como as RNAs podem ser a melhor solução.

Recomenda-se em estudos futuros a utilização de uma Rede Neural Artificial recorrente ao invés da rede do tipo *feedforward*. Neste caso, a saída de alguns neurônios serviria como entrada para outros neurônios,

possibilitando que a rede utilizasse sua própria vazão prevista como entrada para as previsões seguintes. Esse procedimento dispensaria a utilização da série de vazão como variável de entrada do modelo de RNA o qual se aproximaria do Sacramento. Assim, uma comparação mais equilibrada entre os dois modelos seria possível.



## 6. REFERÊNCIAS BIBLIOGRÁFICAS

ANDERSON, R. M., KOREN, V. I., REED, S. M., 2006, "Using SSURGO data to improve Sacramento model a priori parameter estimates", **Journal of Hydrology**, v. 320(1), pp. 103-116.

ATLAS DIGITAL DAS ÁGUAS DE MINAS. Disponível em: <[http://www.atlasdasaguas.ufv.br/atlas\\_digital\\_das\\_aguas\\_de\\_minas\\_gerais.html](http://www.atlasdasaguas.ufv.br/atlas_digital_das_aguas_de_minas_gerais.html)>. Acesso em: junho de 2013.

ARCOVA, F.C.S., CICCO, V., ROCHA, P.A.B., 2003, "Precipitação efetiva e interceptação das chuvas por floresta de Mata Atlântica em uma microbacia experimental em Cunha – São Paulo", **Revista Árvore**, v. 27, n.2, pp. 257-262.

BASHEER, I.A., HAJMEER, M., 2000, "Artificial neural networks: fundamentals, computing, design, and application", **Journal of Microbiological Methods**, v. 43, pp. 3-31.

BEVEN, K., 1989, "Changing ideas in hydrology - The case of physically-based models", **Journal of Hydrology**, v.105, pp.157-172.

BRAVO, J.M., PEDROLLO, O.C., COLLISCHONN, W., PAZ, A.R., 2008, "Previsões de Curto Prazo de Vazão Afluente ao Reservatório de Furnas Utilizando Redes Neurais Artificiais", **Revista Brasileira de Recursos Hídricos**, v. 13, n. 2, pp.77-88.

BRAZ, R.L., RIBEIRO, C.A.D., FERREIRA, D.S., CECÍLIO, R.A., 2007, "Uso de Séries Históricas e Técnicas de SIG no Estudo da Distribuição Temporal e Espacial da Pluviosidade na Bacia Barra Seca Localizada ao Norte do Estado do Espírito Santo", **VII Encontro Latino Americano de Pós-Graduação**, Universidade do Vale do Paraíba, pp. 54-56.

BURNASH, R.J.C., 1995, "The NWS River Forecast System – Catchment Modeling". In: SINGH , V.P.(Ed.), **Computer Models of Watershed Hydrology**, Louisiana, Estados Unidos da América.

CABRAL, I.G., 2013, **Previsões de vazão diária para a bacia do Alto Iguaçu com o emprego de Redes Neurais Artificiais**. Projeto Final, Universidade Federal do Paraná, Curitiba, Paraná, Brasil.

CARDOSO, M.R., 2009, **Modelagem chuva-vazão para períodos mensais utilizando Redes Neurais Artificiais**. Projeto Final, Universidade Federal do Paraná, Curitiba, Paraná, Brasil.

CHAI, T., DRAXLER, R.R., 2014, “Root mean square error (RMSE) or mean absolute error (MAE)?”, **Geoscientific Model Development Discussions**, v. 7, pp. 1525-1534.

CIGIZOGLU, H.K., 2003, “Estimation, forecasting and extrapolation of river flows by artificial neural networks”, **Hydrological Sciences Journal**, v. 48, n. 3, pp. 349-361.

CINTRA, F.L.D., LIBARDI, P.L., SAAD, A.M., 2000, “Balanço hídrico no solo para porta-enxertos de citros em ecossistema de tabuleiro costeiro”, **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.4, n.1, pp. 23-28.

CRC FOR CATCHMENT HYDROLOGY, 2004, **Rainfall Runoff Library – User Guide**. Disponível em: <  
<http://www.toolkit.net.au/Tools/DownloadDocumentation.aspx?id=1000094>>. Acesso em: maio de 2014.

DELTARES, 2013, **SOBEK User Manual**. Disponível em: <  
[http://content.oss.deltares.nl/delft3d/manuals/SOBEK\\_User\\_Manual.pdf](http://content.oss.deltares.nl/delft3d/manuals/SOBEK_User_Manual.pdf)>. Acesso em: maio de 2014.

DUAN, Q., SOROOSHIAN, S., GUPTA, V.K., 1992, “Effective and efficient global optimization for conceptual rainfall–runoff models”, **Water Resources Research**, v. 28, n. 4, pp. 1015–1031.

ESSENFELDER, A.H., 2009, **Previsão de Curto Prazo da Vazão de um Rio Utilizando Redes Neurais Artificiais**. Projeto Final, Universidade Federal do Paraná, Curitiba, Paraná, Brasil.

FIORIN, D.V., MARTINS, F.R., SCHUCH, N.J., PEREIRA, E.B., 2011, "Aplicações de redes neurais e neurais e previsões de disponibilidade de recursos energéticos solares", **Revista Brasileira de Ensino de Física**, v. 33, n. 1, pp. 1309-1329.

FRAKES, B., YU, Z., 1999, "An Evaluation of two hydrologic models for climate change cenarios", **Journal of the American Water Resources Association**, v. 35, n. 6, pp. 1351-1363.

GARDNER, M.W., DORLING, S.R., 1998, "Artificial Neural Networks (The Multilayer Perceptron) - A Review of Applications in the Atmospheric Sciences", **Atmospheric Environment**, v. 32, n. 14/15, pp. 2624-2636.

GIBERTONI, R.F.C., PINHO, J.L.S., MUHLENHOFF, A.P., KAN, A., SCHWANEBERG, D., REIS, A.A., 2014, "Sacramento Model Calibration Using Different Methodologies – Application to the Itabapoana-MG River Basin", **6<sup>th</sup> International Conference on Flood Management**, São Paulo, Brasil.

HIDROWEB. Disponível em: <<http://hidroweb.ana.gov.br>>. Acesso em: abril 2014.

HSIEH, W.W., TANG, B., 1998, "Applying Neural Network Models to Prediction and Data Analysis in Meteorology and Oceanography", **Bulletin of the American Meteorological Society**, v. 79, n. 9, pp. 1855-1870.

JAIN, A., KUMAR, A.M., 2007, "Hybrid neural network models for hydrologic time series forecasting", **Applied Soft Computing**, v. 7, pp. 585-592.

KAVISKI, E.; CUNHA, L. M.; GIBERTONI, R. F. C., 2005, "Método de desagregação utilizando séries de Fourier", **Simpósio Nacional de Recursos Hídricos**, Maceió, Brasil.

KOREN, V.I., SMITH, M., WANG, D., ZHANG, Z., 2000, "Use of Soil Property Data in the Derivation of Conceptual Rainfall-Runoff Model Parameters", **American Meteorological Society 15th Conference on Hydrology**, Long Beach, Estados Unidos da América.

KRAUSKOPF NETO, R.K., 2005, **Atualização de Modelos Chuva-Vazão-Propagação com Estimadores de Estado**. Dissertação de mestrado, Universidade Federal do Paraná, Curitiba, Paraná, Brasil.

LACTEC – INSTITUTO DE TECNOLOGIA PARA O DESENVOLVIMENTO, 2014, **Implantação de Sistema para Acoplamento de Modelos e Informação Telemétrica Visando à Otimização da Operação de Reservatórios em Tempo Real, com Foco no Controle de Cheias - Relatório Técnico Nº 8 – Calibração Hidrológica para as Bacias Hidrográficas do Grupo 2**, Curitiba, Brasil.

LUCAS, E.W.M., SOUSA, F.A.S., SILVA, F.D.S.S., LUCIO, P.S., 2009, “Modelagem hidrológica determinística e estocástica aplicada à região hidrográfica do Xingu”, **Revista Brasileira de Meteorologia**, v. 24, n. 3, pp. 308-322.

MAIER, H.R., JAIN, A., DANDY, G.C., SUDHEER, K.P., 2010, “Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions”, **Environmental Modelling & Software**, v. 25, pp. 891-909.

MINOCHA, V.K., 2003, “Discussion of Comparative Analysis of Event-based Rainfall-runoff Modeling Techniques – Deterministic, Statistical and Artificial Neural Networks” by Ash Jain and S.K.V. Prasad Indurthy”, **Journal of Hydrologic Engineering**, v. 9, pp. 551-553.

MORIASI, D.N., ARNOLD, J.G., VAN LIEW, M.W., BINGNER, R.L., HARMEL, R.D., VEITH, T.L., 2007, “Model Evaluation Guidelines For Systematic Quantification Of Accuracy In Watershed Simulations”, **American Society of Agricultural and Biological Engineers**, v. 50, n. 3, pp.885-900.

NWS (NATIONAL WEATHER SERVICE), 2002, “Conceptualization of the Sacramento Soil Moisture Accounting Model”. Disponível em: <[http://www.nws.noaa.gov/oh/hrl/nwsrfs/users\\_manual/part2/\\_pdf/23sacsma.pdf](http://www.nws.noaa.gov/oh/hrl/nwsrfs/users_manual/part2/_pdf/23sacsma.pdf)>. Acesso em: outubro de 2014.

PIMENTEL, C., 2004, **A relação da planta com a água**. 1 ed. Seropédica, Edur, 2004.

PINTO, N.L.S., HOLTZ, A.C.T., MARTINS, J.A., GOMIDE, F.L.S., 1976, **Hidrologia Básica**. 1 ed. São Paulo, Edgard Blücher.

REBOUÇAS, A.C., 2003, “Água no Brasil: abundância, desperdício e escassez”, **Bahia Análise & Dados**, v. 13, n. especial, pp. 341-345.

SÁ JÚNIOR, A., 2009, **Aplicação da Classificação Köppen para o Zoneamento Climático do Estado de Minas Gerais**. Dissertação de mestrado, Universidade Federal de Lavras, Lavras, Minas Gerais, Brasil.

TOKAR, A.S., MARKUS, M., 2000, “Precipitation-Runoff Modeling Using Artificial Neural Networks and Conceptual Models”, **Journal Of Hydrologic Engineering**, v. 5, pp. 156-161.

THORNTHWAITTE, C. W., 1948, “An approach towards a rational classification of climate”, **Geographical Review**, v. 38, pp. 55-94.

TUCCI, C.E.M., 2005, **Modelos Hidrológicos**. 2 ed. Porto Alegre, UFRGS.

TUCCI, C.E.M., (org.), 2013, **Hidrologia: ciência e aplicação**. 4 ed. Porto Alegre, UFRGS/ABRH.

ZANETTI, C.M., SOUSA, E.F., CARVALHO, D.F., SALASSIER, B., 2008, “Estimação da evapotranspiração de referência no Estado do Rio de Janeiro usando redes neurais artificiais”, **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 12, n. 2, pp.174–180.

ZEALAND, C.M., BURN, D.H., SIMONOVIC, S.P., 1999, “Short term streamflow forecasting using artificial neural networks”, **Journal of Hydrology**, v. 214, pp. 32-48.

YAPO, P., GUPTA, V.K., SOROOSHIAN, S., 1996, “Calibration of conceptual rainfall–runoff models: Sensitivity to calibration data”, **Journal of Hydrology**, v. 181, pp. 23-48.

## APÊNDICE A

Configuração	Média RMSE validação (m <sup>3</sup> /s)	Média RMSE teste (m <sup>3</sup> /s)	Média NASH	Média desvio de vazões (%)
T_5_01	24,226	24,393	0,862	0,324
T_5_02	27,720	29,259	0,835	1,015
T_5_03	28,409	27,607	0,828	0,265
T_5_04	32,606	28,159	0,780	-0,204
T_5_05	29,042	27,728	0,821	1,241
T_5_06	31,016	29,933	0,844	-0,153
T_5_07	31,057	30,525	0,845	0,163
T_5_08	26,913	28,624	0,828	-0,216
T_5_09	26,889	32,358	0,832	0,319
T_5_10	29,396	28,188	0,824	0,065
T_5_11	98,970	98,945	-16,443	-0,079
T_5_12	69,457	70,373	-0,453	0,553
T_5_13	67,199	70,384	-0,328	0,144
T_5_14	65,916	69,988	-0,367	-1,191
T_5_15	69,166	68,252	-0,523	-0,715
T_5_16	96,283	104,969	-8,928	-2,410
T_5_17	66,335	65,473	-0,123	0,490
T_5_18	67,369	63,972	-0,309	-0,380
T_5_19	64,867	64,118	-0,189	-0,260
T_10_01	22,450	25,398	0,872	-0,176
T_10_02	31,260	28,613	0,769	0,451
T_10_03	26,508	32,661	0,815	0,328
T_10_04	58,039	49,614	0,288	0,965
T_10_05	69,031	50,884	-0,215	3,416
T_10_06	33,967	33,185	0,740	-0,139
T_10_07	28,936	28,580	0,802	-0,573
T_10_08	38,473	35,370	0,611	-0,124
T_10_09	50,612	52,344	0,167	-1,443
T_10_10	50,726	41,243	0,249	1,973
T_10_11	80,301	81,444	-2,873	-0,160
T_10_12	69,345	67,215	-0,582	0,490
T_10_13	103,426	108,799	-7,930	4,423
T_10_14	93,760	103,094	-4,204	1,768
T_10_15	59,712	49,293	0,470	1,510
T_10_16	75,154	74,865	-0,980	-0,347
T_10_17	68,999	66,019	-0,205	1,413
T_10_18	103,826	106,502	-5,229	5,075
T_10_19	98,610	92,840	-5,216	2,771
T_15_01	24,115	22,867	0,884	-0,196
T_15_02	28,669	28,609	0,824	-0,695
T_15_03	30,281	31,498	0,801	-0,153

Configuração	Média RMSE validação (m <sup>3</sup> /s)	Média RMSE teste (m <sup>3</sup> /s)	Média NASH	Média desvio de vazões (%)
T_15_04	59,876	46,617	0,125	1,896
T_15_05	54,751	50,415	-0,062	0,918
T_15_06	29,449	29,513	0,802	-0,406
T_15_07	28,756	30,899	0,838	-0,981
T_15_08	37,738	30,694	0,578	0,626
T_15_09	50,198	49,529	0,364	-0,087
T_15_10	47,883	40,137	0,453	-0,279
T_15_11	79,402	82,677	-2,522	0,254
T_15_12	73,921	66,232	-0,788	0,226
T_15_13	106,588	104,240	-3,894	-1,537
T_15_14	95,330	94,100	-3,824	-1,713
T_15_15	62,538	51,910	0,418	1,249
T_15_16	75,509	74,362	-0,923	0,170
T_15_17	67,141	65,572	-0,233	0,238
T_15_18	106,369	108,494	-3,990	2,157
T_15_19	88,221	92,883	-3,218	-2,185
T_5_20	22,534	22,063	0,930	0,745
T_5_21	26,999	26,048	0,908	-0,394
T_5_22	24,960	24,311	0,916	-0,353
T_5_23	25,655	25,903	0,903	-0,277
T_5_24	19,796	21,272	0,952	-0,279
T_5_25	22,168	22,419	0,937	0,135
T_5_26	22,391	22,391	0,937	-0,168
T_5_27	24,284	24,937	0,926	-0,120
T_5_28	19,989	21,373	0,943	0,000
T_5_29	20,846	22,713	0,941	-0,587
T_5_30	21,991	23,209	0,930	-0,511
T_5_31	21,180	22,279	0,941	-0,364
T_5_32	20,655	21,352	0,938	-0,231
T_5_33	21,431	21,518	0,942	0,032
T_5_34	22,963	22,812	0,933	0,474
T_5_35	22,901	21,108	0,930	-0,041
T_5_36	23,613	22,880	0,926	-0,142
T_5_37	21,243	21,857	0,943	-0,050
T_5_38	22,158	21,498	0,938	0,062
T_5_39	21,944	20,937	0,935	-0,088
T_5_40	25,834	27,021	0,925	-0,244
T_5_41	32,330	29,970	0,849	0,943
T_5_42	26,969	28,112	0,917	0,022
T_5_43	27,909	25,575	0,900	0,752
T_5_44	24,902	26,041	0,918	0,098
T_5_45	25,328	25,782	0,909	-0,497
T_5_46	26,606	28,146	0,905	-0,486

<b>Configuração</b>	<b>Média RMSE validação (m<sup>3</sup>/s)</b>	<b>Média RMSE teste (m<sup>3</sup>/s)</b>	<b>Média NASH</b>	<b>Média desvio de vazões (%)</b>
T_5_47	34,076	32,891	0,830	0,229
T_5_48	25,036	24,533	0,919	0,324
T_5_49	33,566	32,417	0,840	0,183
T_5_50	26,542	25,673	0,898	-0,181
T_5_51	28,897	27,035	0,897	-0,657
T_5_52	25,781	25,217	0,910	-0,360
T_5_53	23,661	24,348	0,920	-0,333
T_5_54	28,304	25,604	0,896	-0,298
T_5_55	25,396	25,334	0,912	-0,405
T_5_56	21,329	19,355	0,943	0,291
T_5_57	20,892	20,026	0,943	0,345
T_5_58	22,632	22,856	0,922	0,352
T_5_59	22,833	21,684	0,933	-0,094
T_5_60	22,813	22,754	0,931	0,349
T_5_61	22,701	22,911	0,933	0,215
T_5_62	22,034	20,950	0,929	-0,549
T_5_63	22,830	21,404	0,928	0,228



## APÊNDICE B

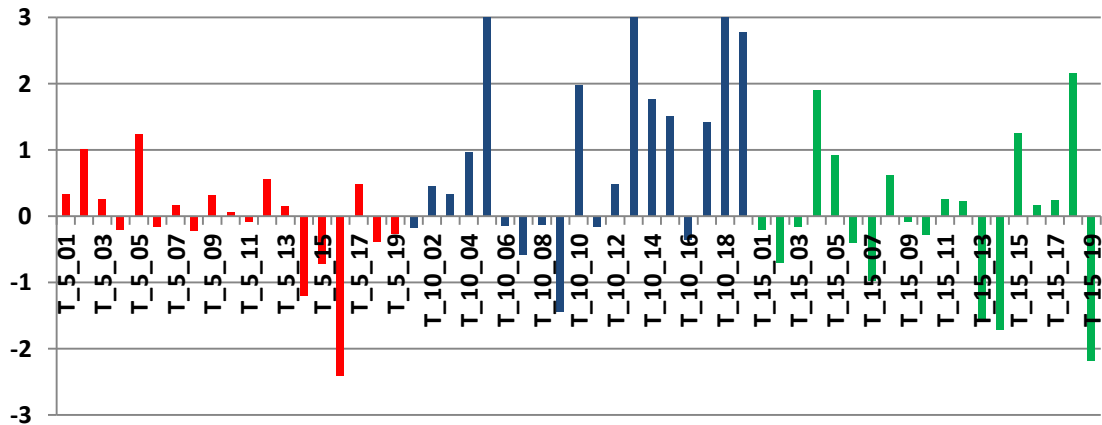


Figura B.1 - Desvio de vazões para a escolha do número de neurônios da camada intermediária.

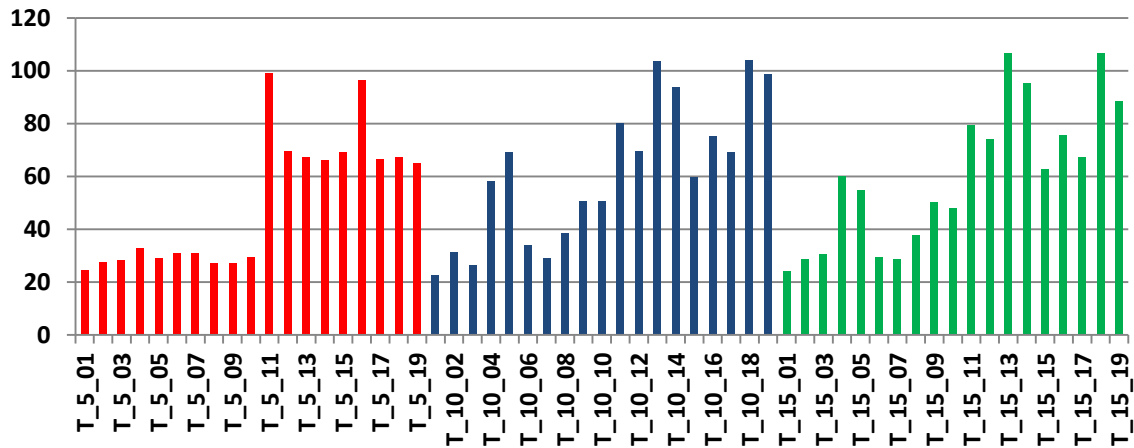


Figura B.2 – RMSE da validação para a escolha do número de neurônios da camada intermediária.

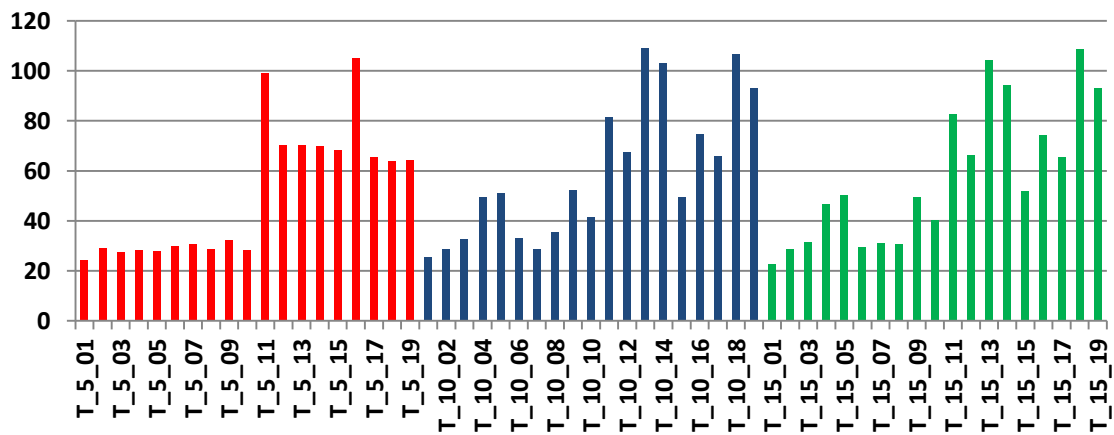


Figura B.3 – RMSE do teste para a escolha do número de neurônios da camada intermediária.

## APÊNDICE C

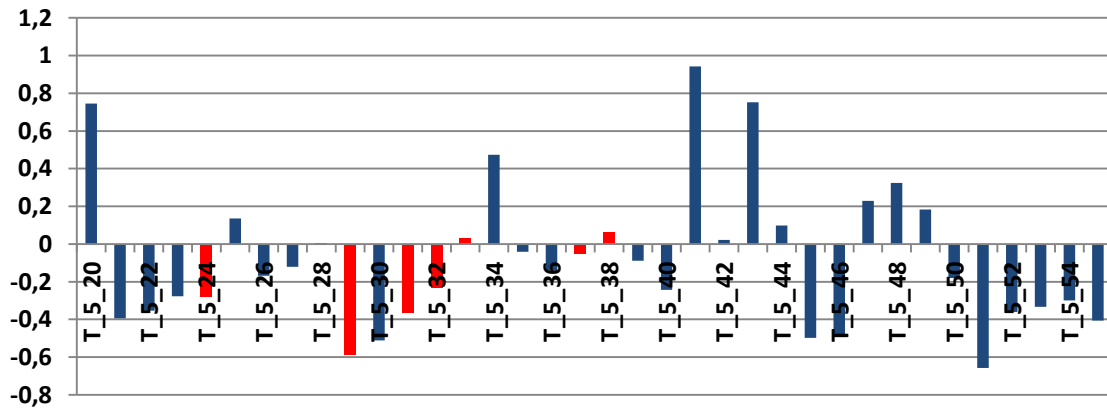


Figura C.1 - Desvio de vazões para a definição das variáveis de entrada.

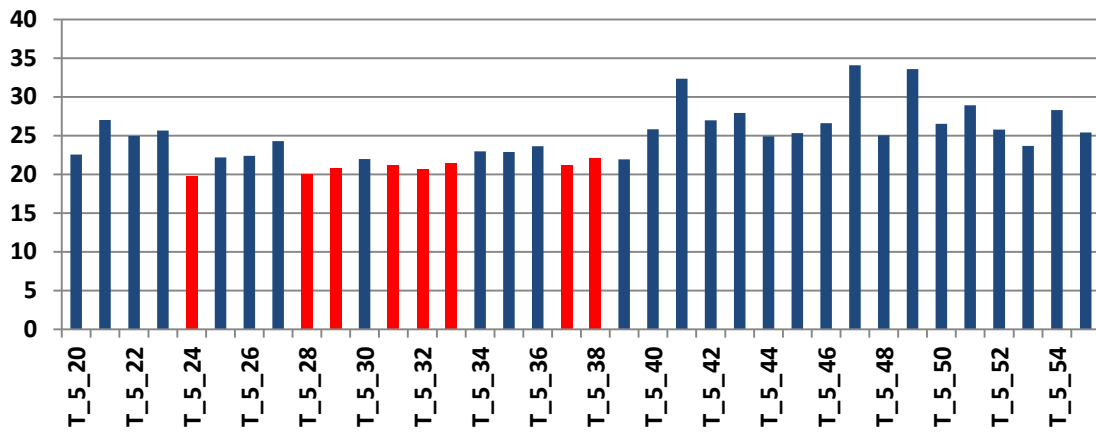


Figura C.2 - RMSE da validação para a definição das variáveis de entrada.

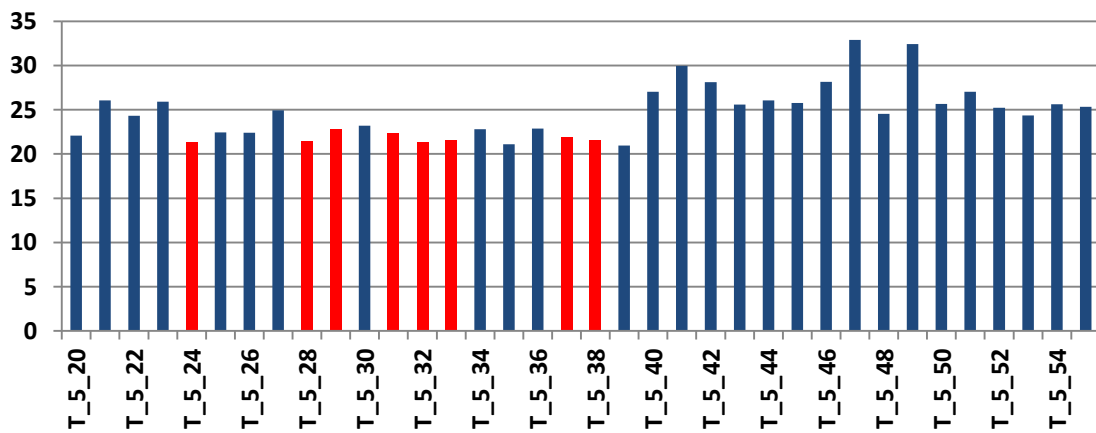


Figura C.3 - RMSE do teste para a definição das variáveis de entrada.

## APÊNDICE D

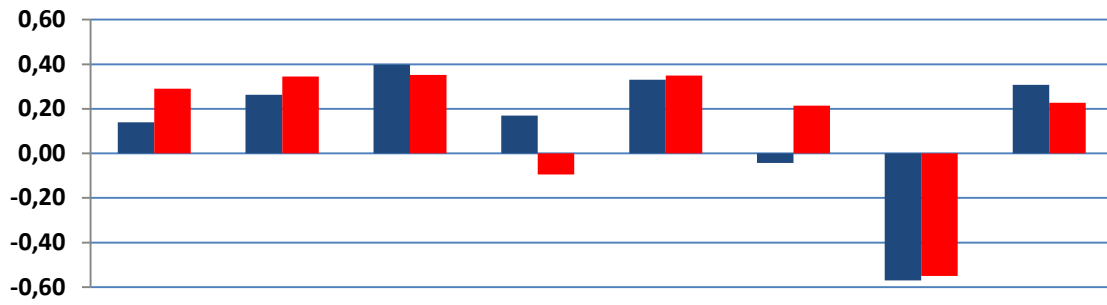


Figura D.1 – Desvio de vazões para configurações com a ausência (azul) e presença (vermelho) da variável evapotranspiração.

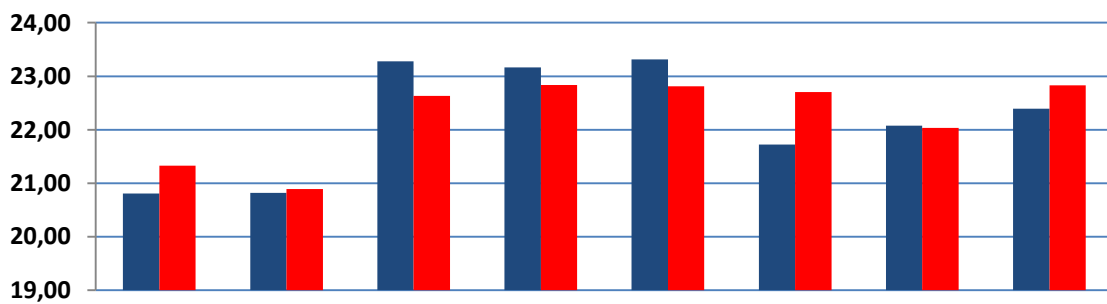


Figura D.2 – RMSE da validação para configurações com a ausência (azul) e presença (vermelho) da variável evapotranspiração.

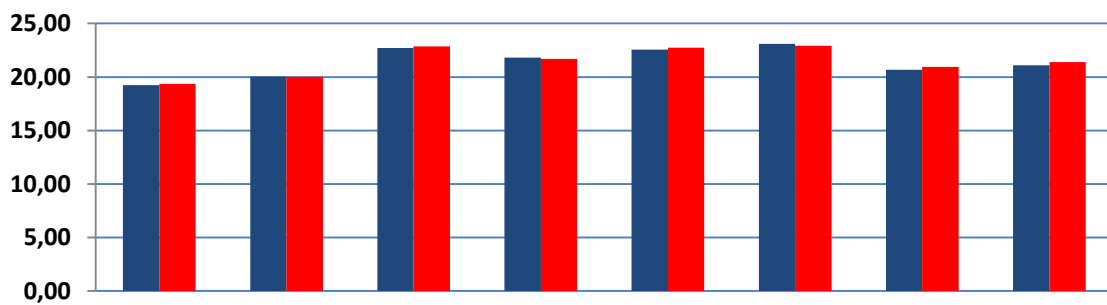


Figura D.3 – RMSE do teste para configurações com a ausência (azul) e presença (vermelho) da variável evapotranspiração.