

## DESENVOLVIMENTO DE REDES NEURAIS ARTIFICIAIS PARA ESTIMATIVA DAS VAZÕES DIÁRIAS NA BACIA DO RIO PIRACICABA<sup>5</sup>

**EDUARDO MORGAN ULIANA<sup>1</sup>; DEMETRIUS DAVID DA SILVA<sup>2</sup>; MICHEL CASTRO MOREIRA<sup>3</sup>; DONIZETE DOS REIS PEREIRA<sup>4</sup>; SILVIO BUENO PEREIRA<sup>2</sup> E FREDERICO TERRA DE ALMEIDA<sup>1</sup>**

<sup>1</sup>Universidade Federal de Mato Grosso (UFMT), Instituto de Ciências Agrárias e Ambientais (ICAA), Avenida Alexandre Ferronato, CEP.: 78557-267, Sinop – MT, Brasil, e-mail: morganuliana@ufmt.br, fredterr@gmail.com.

<sup>2</sup>Universidade Federal de Viçosa (UFV), Departamento de Engenharia Agrícola, Avenida Peter Henry Rolfs, CEP.: 36570-900, Viçosa – MG, Brasil, e-mail: demetrius.ufv@gmail.com.

<sup>3</sup>Universidade Federal do Oeste da Bahia, Centro das Ciências Exatas e das Tecnologias, Rua Professor José Seabra de Lemos, CEP.: 47808-021, Barreiras – BA, Brasil, e-mail: michelcm@gmail.com.

<sup>4</sup>Universidade Federal de Viçosa (UFV), Instituto de Ciências Agrárias, Rodovia LMG 818, km 06, Florestal – MG, Brasil, e-mail: donizete.pereira@ufv.br.

5O artigo é referente ao capítulo 3 da tese de doutorado do primeiro autor.

### 1 RESUMO

As Redes Neurais Artificiais (RNAs) são uma alternativa na modelagem hidrológica para a estimativa das vazões dos cursos de água a partir de dados hidrometeorológicos. O objetivo do trabalho foi desenvolver Redes Neurais Artificiais para estimar as vazões diárias na bacia hidrográfica do rio Piracicaba, Minas Gerais. O estudo foi realizado em três seções de monitoramento de vazão da bacia do rio Piracicaba, localizada no Estado de Minas Gerais - Brasil. No desenvolvimento das RNAs foram realizados a coleta e seleção dos dados; a definição da arquitetura da rede; e o treinamento e validação das redes desenvolvidas. A maior parte das RNAs desenvolvidas apresentou coeficiente de Nash-Sutcliffe maior que 0,80 o que permitiu classificar os modelos como bons para a estimativa das vazões. Com base nos resultados, pode-se concluir que as RNAs são adequadas para a estimativa das vazões diárias na bacia do rio Piracicaba e podem ser empregadas na estimativa de eventos extremos e no gerenciamento dos recursos hídricos.

**Palavras-Chave:** modelo empírico inteligência artificial, modelagem hidrológica.

**ULIANA, E. M.; SILVA, D. D.; MOREIRA, M. C.; PEREIRA, D. R.; PEREIRA, S. B.; ALMEIDA, F. T.**

**ARTIFICIAL NEURAL NETWORKS FOR DAILY FLOW ESTIMATES IN THE PIRACICABA RIVER BASIN**

### 2 ABSTRACT

Artificial neural networks (ANNs) have been used alternatively in hydrologic modeling to estimate accurately watercourse flows based on hydrometeorological data. This study developed artificial neural networks to estimate daily flows in Piracicaba river basin, in Minas Gerais state (Brazil). For this, we used three runoff-monitoring sections of the Piracicaba river

basin, with an area of 5,304.0 km<sup>2</sup>, and located in the State of Minas Gerais – Brazil. For designing the ANNs to estimate daily flows, we adopted the following steps: data collection and selection, network architecture definition, training and validation of results. The results showed that ANNs are adequate to estimate daily flows in Piracicaba river basin.

**Keywords:** empirical model, artificial intelligence, hydrologic modeling.

### 3 INTRODUÇÃO

A maior parte dos processos hidrológicos apresenta alto grau de variabilidade espacial e temporal, aumentando a complexidade da modelagem computacional dos processos envolvidos. Na representação do ciclo hidrológico, o empirismo desempenha papel importante em estudos relacionados à modelagem hidrológica. Os modelos empíricos tratam o sistema hidrológico como uma “caixa preta” e tentam encontrar uma relação entre os dados de entrada e os de saída (GOVINDARAJU, 2000a).

Entre os modelos empíricos, as Redes Neurais Artificiais (RNAs) apresentam resultados promissores para a estimativa das vazões de cursos de água, como demonstrado por Aichouri et al. (2015), Elsafi (2014), Meng et al. (2015), Oliveira et al. (2013), Sattari; Apaydin; Ozturk (2012) e Setiono (2015).

As RNAs imitam o funcionamento do cérebro humano, por meio da aquisição de conhecimento em um processo de aprendizagem. A capacidade de extrair a relação entre as entradas e saídas de um processo, sem considerar os processos físicos que estão sendo explicitamente envolvidos, torna a RNA ferramenta capaz de resolver problemas complexos e não lineares, como é o caso da relação entre a chuva e a vazão em bacias hidrográficas, que depende de inúmeros fatores, como umidade inicial do solo, uso e manejo do solo, geomorfologia da bacia hidrográfica, evaporação, infiltração, distribuição e duração da chuva, entre outros (ARAÚJO; DANTAS NETO; SOUZA FILHO, 2015; GOVINDARAJU, 2000b).

O desenvolvimento de uma RNA requer a definição de sua arquitetura, que consiste na determinação dos dados de entrada e saída, do número de camadas intermediárias e de neurônios artificiais, e das funções de transferência (MENG et al., 2015). Segundo Asadi et al. (2013), a arquitetura de uma RNA é variante e deve ser definida de forma que proporcione o menor erro de estimativa dos dados de saída.

Após a definição da arquitetura, a RNA deve passar pela etapa de treinamento, que consiste em ajustar pesos para as conexões da rede com auxílio de um algoritmo de aprendizagem. Por fim, a RNA deve ser validada com dados que não foram utilizados durante o treinamento, de forma a verificar seu desempenho para estimativa dos dados de saída (LUDWIG JÚNIOR; COSTA, 2007).

Considerando que a modelagem das vazões de cursos de água possibilita a previsão de eventos extremos e subsidia o planejamento e a avaliação de estratégias de gerenciamento de recursos hídricos, objetivou-se neste trabalho desenvolver redes neurais artificiais para estimar vazões diárias na bacia hidrográfica do Rio Piracicaba, Minas Gerais.

### 4 MATERIAL E MÉTODOS

#### 4.1 Área de estudo e base de dados

O estudo foi conduzido na Bacia do Rio Piracicaba, localizada no Estado de Minas Gerais - Brasil. O rio Piracicaba é um afluente do rio Doce, e sua área de drenagem abrange 21 municípios mineiros, cuja

economia é baseada, principalmente, em três atividades: mineração, siderurgia e monocultura de eucalipto.

Para treinamento e validação das Redes Neurais Artificiais (RNAs), foram utilizados dados de estações pluviométricas, fluviométricas e meteorológicas localizadas

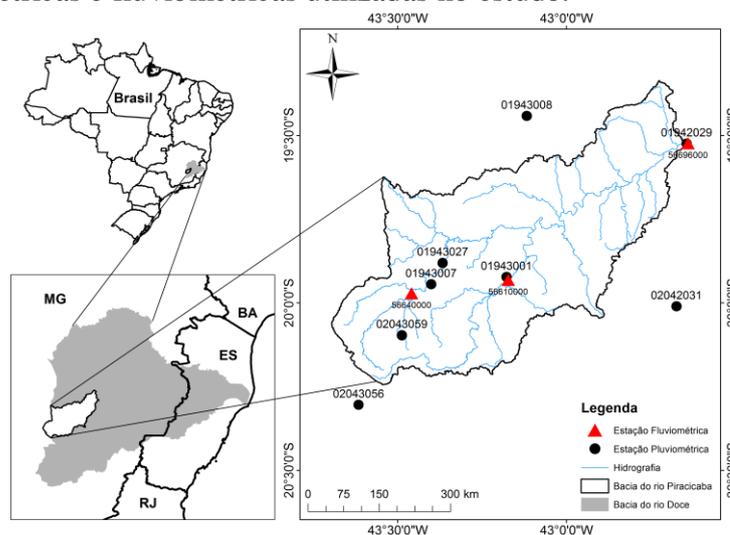
na bacia hidrográfica e em seu entorno (Tabela 1 e Figura 1). A utilização de estações do entorno foi decorrente da baixa quantidade de estações (pluviométrica e meteorológica) localizadas no interior da área de drenagem com série de dados suficiente para realização da modelagem.

**Tabela 1.** Estações pluviométricas, fluviométricas e meteorológicas utilizadas no estudo

Código	Nome	Tipo	Latitude (°)	Longitude (°)	Altitude (m)	AD (km <sup>2</sup> )
01942029	Mário de Carvalho	P	-19,525	-42,644	232	-
01943008	Santa Maria do Itabira	P	-19,442	-43,118	538	-
01943027	Usina Peti	P	-19,881	-43,367	1.110	-
02042031	Fazenda Cach. D'Antas	P	-20,011	-42,674	280	-
02043059	Colégio Caraça	P	-20,097	-43,488	1.300	-
02043056	Fazenda Água Limpa	P	-20,305	-43,616	965	-
01943007	Santa Bárbara	P	-19,945	-43,401	748	-
01943001	Rio Piracicaba	P	-19,923	-43,178	623	-
56640000	Carrapato	F	-19,972	-43,459	755	427,5
56610000	Rio Piracicaba	F	-19,932	-43,173	748	1.164,2
56696000	Mário de Carvalho	F	-19,524	-42,640	232	5.304,0
2042024	Viçosa	M	-20,766	-42,866	712	-

AD: área de drenagem; P: estação pluviométrica; F: estação fluviométrica; e M: estação meteorológica.

**Figura 1.** Bacia hidrográfica do rio Piracicaba, com destaque para a localização das estações pluviométricas e fluviométricas utilizadas no estudo.



Foram utilizados dados diários de vazões correspondentes aos períodos de 01/09/1990 a 31/12/1993, 01/07/2002 a

31/12/2006 e 01/09/1990 a 31/12/1992, registrados, respectivamente, nas estações fluviométricas 56696000 (Mário de

Carvalho), 56610000 (Rio Piracicaba) e 56640000 (Carrapato), assim como dados pluviométricos e climáticos (Tabela 1). Esses períodos foram selecionados por não apresentarem falhas nos dados hidrometeorológicos das estações listadas na Tabela 1.

As RNAs não requerem a utilização de séries históricas contínuas como requisitado por alguns modelos hidrológicos conceituais. Portanto, a utilização de períodos não sequenciais como realizada neste estudo não interfere nos resultados. Destaca-se novamente que a utilização dos períodos apresentados foi condicionada pela inexistência de falhas nos dados de chuva, vazão e temperatura do ar.

Para seleção das variáveis de entrada das RNAs desenvolvidas, foi utilizada a análise de componentes principais (ACP), que permitiu avaliar a importância de cada variável e promover a eliminação daquelas que contribuíram pouco, em termos de variação, no grupo analisado. Na ACP, foram analisadas as seguintes variáveis hidrometeorológicas: precipitação; evapotranspiração de referência; temperatura máxima; temperatura mínima; e vazões registradas nas estações fluviométricas. A ACP permitiu propor as entradas das RNAs para estimativa da vazão diária ( $Q_t$ ), como apresentado na Tabela 2.

**Tabela 2.** Variáveis de entrada utilizadas para o desenvolvimento das Redes Neurais Artificiais

Entrada	Entrada	Entrada	Entrada
$P_t$	$Q_{mt,2}$	$Q_{t-1}; Q_{t-2}$	$Q_{mt,4}; P_t$
$E_t$	$Q_{mt,3}$	$P_t; E_t$	$Q_{t-1}; Q_{t-2}; Q_{t-3}$
$Q_{t-1}$	$Q_{mt,4}$	$Q_{mt,2}; Q_{mt,3}$	$Q_{mt,2}; Q_{mt,3}; Q_{mt,4}$
$Q_{t-2}$	$Q_{mt,5}$	$Q_{t-1}; P_t$	$Q_{t-1}; Q_{t-2}; P_t$
$Q_{t-3}$	$Q_{mt,6}$	$Q_{mt,2}; P_t$	$Q_{mt,2}; Q_{mt,3}; P_t$
$Q_{t-4}$	$Q_{mt,7}$	$Q_{mt,3}; P_t$	$P_t; E_t; Q_{t-1}$

$P_t$  representa a precipitação média (mm) na bacia, obtida pelo método de Thiessen, como descrito por Macêdo et al. (2013);  $E_t$ , a evapotranspiração de referência (mm), obtida pelo método de Hargreaves e Samani, conforme Borges Júnior et al. (2012);  $Q_{t-n}$ , a vazão diária defasada em  $n$  (1, 2, 3, ...,  $n$ ) dias ( $m^3 s^{-1}$ ); e  $Q_{mt,n}$ , a vazão média de  $n$  (2, 3, ...,  $n$ ) dias anteriores ( $m^3 s^{-1}$ ).

#### 4.2 Arquitetura e treinamento das redes

No desenvolvimento das redes, foram realizados a coleta e seleção dos dados de entrada e saída; a definição da arquitetura da rede; e o treinamento e validação das redes desenvolvidas.

Os dados de entrada foram organizados aleatoriamente e normalizados,

$$pn = \frac{2(p-\min p)}{(\max p-\min p)} - 1 \quad (1)$$

Em que  $pn$  é o valor normalizado (variando entre -1 e 1),  $p$  é o valor da variável, e  $\min p$

utilizando a Equação 1. Segundo Silva; Spatti; Flauzino (2010), a normalização garante que cada dado de entrada receba a mesma atenção durante o treinamento da RNA, implicando melhor eficiência numérica diante das operações matemáticas realizadas pela rede.

e  $\max p$  são, respectivamente, o menor e o maior valor da variável na série em estudo.

Após a normalização, os dados foram separados em duas categorias: dados de treinamento, que foram utilizados para treinamento da rede; e dados de validação, que foram utilizados para verificar o desempenho da rede. Neste estudo, 85% dos dados da série foram utilizados para treinamento e 15% para validação. Este

critério de divisão da série foi utilizado também por Rezaei et al. (2015) e Uysal, Sorman e Sensoy (2016) para estimativa de vazões com RNAs.

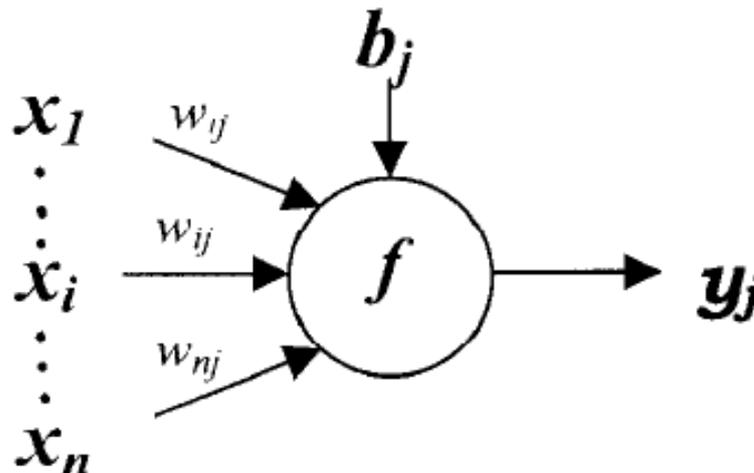
A unidade de processamento das RNAs é o neurônio matemático (Figura 2), sendo seu funcionamento definido na Equação 2.

$$y_j = f\left(\sum_{i=1}^n w_{ij}x_i + b_j\right) \quad (2)$$

Em que  $y_j$  é o valor da resposta do neurônio (valores das variáveis de saída);  $F$ , a função de ativação ou transferência;  $W_{ij}$ , os pesos sinápticos;  $X_i$ , os valores das variáveis de

entrada, que são processados no neurônio; e  $b_j$  (bias), o valor de compensação do neurônio matemático.

**Figura 2.** Esquema de um neurônio artificial.

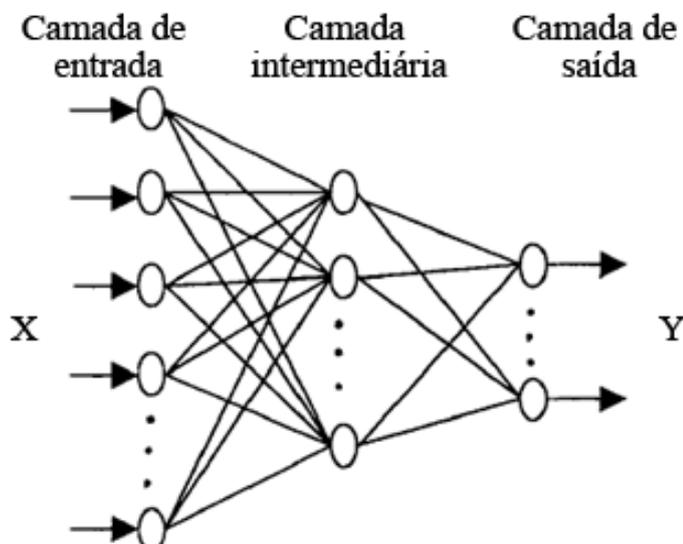


Fonte: Govindaraju (2000a).

Como paradigma neural foi empregado o Perceptron de múltiplas camadas (Figura 3), o qual, segundo Araújo; Dantas Neto; Souza Filho (2015), é muito utilizado para a modelagem de fenômenos

em engenharia. As RNAs foram do tipo retroalimentadas, possuindo uma camada de entrada, duas camadas intermediárias e uma camada de saída.

**Figura 3.** Representação de arquitetura de uma RNA do tipo Perceptron de múltiplas camadas.



Fonte: Adaptado de Govindaraju (2000a).

Como funções de transferência, foram adotadas a tangente hiperbólica nas camadas intermediárias e a tangente linear na camada de saída das RNAs. Foi priorizada a função tangente hiperbólica nas camadas intermediárias, em decorrência de a sua característica de antissimetria melhorar o processo de convergência da rede durante o treinamento, como constatado por Silva; Spatti; Flauzino (2010).

O número de neurônios em cada camada intermediária foi definido empiricamente, durante o processo de treinamento da rede, de forma a obter o menor erro de estimativa das vazões diárias.

Para evitar a memorização da rede (*overfitting*), caracterizada pela redução de sua capacidade de generalização, o número máximo de neurônios artificiais em cada camada intermediária foi limitado a 10. Esse número foi definido com base nos valores estabelecidos por Noori e Kalin (2016), ao utilizarem o modelo para estimativa de vazões.

No treinamento da rede, foi utilizado o algoritmo *backpropagation* (SILVA; SPATTI; FLAUZINO, 2010), conhecido também como algoritmo de retropropagação do erro, incorporado ao algoritmo de otimização de Levenberg-Marquardt, como

proposto por Asadi et al. (2013), de forma a minimizar o esforço computacional e o problema de convergência muito lenta, decorrente da utilização de forma isolada do algoritmo *backpropagation*.

Durante o treinamento, foi estabelecido um limite máximo de épocas (*epochs*), como critério adicional de parada do algoritmo, quando a precisão especificada se tornou inalcançável.

A quantidade de épocas é definida como o número de tentativas de ajuste dos pesos ( $w_{ij}$ ) e bias ( $b_j$ ) (Equação 2) da RNA, de forma a garantir acurácia nas estimativas da variável de saída (GOVINDARAJU, 2000a).

Neste trabalho foi utilizado o software Matlab para definir a arquitetura e realizar o treinamento das RNAs, conforme procedimentos apresentados.

### 4.3 Avaliação do desempenho das RNAs

Para validação do resultado das RNAs, foram utilizadas as seguintes medidas estatísticas: o erro absoluto médio (MAE); a raiz do erro quadrático médio (RMSE); o viés (MBE); o teste t pareado a 5% de significância; o índice de concordância de Willmott (d); e o índice de eficiência de

Nash-Sutcliffe ( $E_{NS}$ ), respectivamente obtidos por:

$$MAE = \frac{1}{N} \sum_{i=1}^N |O_i - P_i| \quad (3)$$

$$RMSE = \left[ \frac{1}{N} \sum_{i=1}^N (O_i - P_i)^2 \right]^{0,5} \quad (4)$$

$$Viés = \frac{1}{N} \sum_{i=1}^N (O_i - P_i) \quad (5)$$

$$t = \sqrt{\frac{(N-1) viés^2}{RMSE^2 - viés^2}} \quad (6)$$

$$d = 1 - \left[ \frac{\sum_{i=1}^N (P_i - O_i)^2}{\sum_{i=1}^N (|P_i - O| + |O_i - O|)^2} \right] \quad (7)$$

$$E_{NS} = 1 - \frac{\sum_{i=1}^N (O_i - P_i)^2}{\sum_{i=1}^N (O_i - O)^2} \quad (8)$$

Em que  $P_i$  é a vazão estimada ( $m^3 s^{-1}$ ),  $O_i$  é a vazão observada ( $m^3 s^{-1}$ ),  $O$  é a média das

vazões observadas ( $m^3 s^{-1}$ ) e  $N$  é o número de valores da amostra.

O erro absoluto médio (MAE) e a raiz do erro quadrático médio (RMSE) medem a magnitude dos erros em um conjunto de estimativas, nas unidades da variável de interesse. Valores de MAE e RMSE iguais a zero indicam ajuste perfeito do modelo à série de dados observados.

O MAE e o RMSE foram analisados ao mesmo tempo, com a finalidade de diagnosticar a variação dos erros nas séries de vazões estimadas. O valor de RMSE será sempre superior ou igual ao MAE. Quanto maior a diferença entre eles, maior também a variância dos erros individuais na amostra. No entanto, se apresentarem valores iguais, isso significa que todos os erros possuem a mesma ordem de grandeza.

O viés (MBE) foi utilizado para verificar a existência de subestimativa (valor

positivo) ou superestimativa (valor negativo) pelo modelo.

O teste  $t$  pareado a  $\alpha$  5% de significância foi empregado com a finalidade de confirmar se a diferença ( $\Delta$ ) entre os dados observados e simulados foi estatisticamente nula, sendo esta a condição esperada.

As hipóteses do teste  $t$  foram:  $H_0: \Delta = 0$ ;  $H_1: \Delta \neq 0$ . Quando o valor em módulo de  $t$  (Equação 6) apresentou valor superior ao quantil da distribuição  $t$ , de Student, com  $v = n-1$  grau de liberdade  $\left(t_{\frac{\alpha}{2}, v}\right)$ , então se rejeitou a hipótese nula ( $H_0$ ) bilateral, ou seja, existe diferença entre os valores de vazões observadas e simuladas pelo modelo, o que não é desejável.

O valor do índice de concordância de Willmott varia entre 0 e 1, e valores próximos da unidade indicam concordância, ou seja, melhor desempenho do modelo na estimativa.

Para avaliar o desempenho dos modelos com relação ao índice de eficiência de Nash-Sutcliffe ( $E_{NS}$ ), foi adotada a classificação seguida por Pereira et al. (2016). De acordo com a classificação:  $E_{NS} = 1$  significa ajuste perfeito dos dados preditos pelo modelo;  $E_{NS} > 0,75$  indica que o modelo é adequado e bom;  $0,36 < E_{NS} < 0,75$  indica que o modelo é considerado satisfatório; e  $E_{NS} < 0,36$  indica que o modelo não é satisfatório.

## 5 RESULTADOS E DISCUSSÃO

### 5.1 Estação Mário de Carvalho (56696000)

Na Tabela 3 são apresentados os dados referentes à arquitetura e ao treinamento das Redes Neurais Artificiais (RNAs) desenvolvidas para a estimativa das vazões diárias na seção de controle Mário de Carvalho (56696000), bem como as medidas quantitativas do desempenho dessas RNAs.

Observa-se pelos valores da Tabela 3, que as RNAs que apresentaram bons resultados para estimativa das vazões diárias possuem como variáveis de entrada as vazões do curso de água defasadas no tempo, a precipitação média da bacia hidrográfica e a evapotranspiração de referência.

O número máximo de neurônios nas camadas intermediárias, estabelecido como

igual a 10, só foi empregado na primeira camada da RNA, com dado de entrada  $Q_{mt,2}$ ;  $P_t$  (Tabela 3). Esse resultado demonstra que a maior parte das RNAs desenvolvidas, foram capazes de extrair conhecimento com um número menor de neurônios que o número máximo estabelecido para as camadas intermediárias.

Os dados apresentados na Tabela 3 confirmam que os modelos de RNAs apresentaram concordância e são bons para a estimativa das vazões diárias, visto que o índice de concordância de Willmott ( $d$ ) apresentou valores próximos a unidade, e o índice de eficiência de Nash-Sutcliffe ( $E_{NS}$ ) valores maiores que 0,75.

Os valores do viés indicaram subestimativa na maior parte das configurações das RNAs. Com relação ao teste  $t$  pareado a 5% de significância, percebe-se que as RNAs apresentaram valores de  $t$  não significativo, ou seja, pode-se afirmar com 95% de probabilidade que a diferença entre os valores observados e estimados com esses modelos é estatisticamente nula, indicando que as estimativas são confiáveis.

Os valores do erro absoluto médio (MAE) e da raiz do erro quadrático médio (RMSE) das RNAs variaram entre 12,92 e 17,67  $m^3 s^{-1}$  e entre 31,57 e 42,84  $m^3 s^{-1}$ , respectivamente. Além disso, o distanciamento entre os valores de MAE e RMSE permite afirmar que a variância dos erros individuais é pequena (Tabela 3).

**Tabela 3.** Variáveis de entrada, quantidade de épocas, número de neurônios das camadas intermediárias ( $N_1$  e  $N_2$ ) e medidas quantitativas do desempenho das melhores Redes Neurais Artificiais (RNAs) desenvolvidas para a estimativa da vazão diária ( $m^3 s^{-1}$ ) na seção de monitoramento Mário de Carvalho

Entrada	$Q_{mt,2}$	$Q_{mt,2}; Q_{mt,3}$	$Q_{mt,2}; Q_{mt,3}; Q_{mt,4}$	$P_t; E_t; Q_{t,1}$	$Q_{mt,2}; P_t$	$Q_{mt,3}; P_t$
Épocas	400	200	300	200	300	200
$N_1$	6	3	2	4	10	9
$N_2$	4	6	5	9	2	7
$\bar{O}$	107,05	107,05	107,05	107,05	107,05	107,05
$\bar{P}$	106,71	105,76	107,52	105,59	107,04	107,33
MAE	17,67	16,12	16,77	12,92	16,11	15,82
RMSE	42,84	39,15	39,28	31,57	39,14	39,04
Viés	0,34	1,29	-0,46	1,46	0,02	-0,28
d	0,94	0,96	0,96	0,97	0,96	0,96
$E_{NS}$	0,81	0,84	0,84	0,9	0,84	0,84
t	0,27 <sup>(ns)</sup>	1,10 <sup>(ns)</sup>	0,39 <sup>(ns)</sup>	1,54 <sup>(ns)</sup>	0,01 <sup>(ns)</sup>	0,24 <sup>(ns)</sup>

$N_1$  e  $N_2$  são o número de neurônios artificiais na primeira e na segunda camada intermediária, respectivamente;  $\bar{O}$  e  $\bar{P}$  são as vazões médias ( $m^3 s^{-1}$ ) observadas e estimadas, respectivamente; MAE é o erro absoluto médio ( $m^3 s^{-1}$ ); RMSE é a raiz do erro quadrático médio ( $m^3 s^{-1}$ ); d é o índice de concordância de Willmott;  $E_{NS}$  é o índice de eficiência de Nash-Sutcliffe; t é o valor da estatística de teste do teste t pareado; (ns) não significativo a 5% de significância;  $Q_{t-n}$  é a vazão diária defasada em n (1, 2, 3, ..., n) dias ( $m^3 s^{-1}$ );  $P_t$  é a precipitação média da bacia no tempo t (mm) obtida pelo método de Thiessen;  $E_t$  é a evapotranspiração de referência no tempo t (mm) obtida pelo método de Hargreaves e Samani; e  $Q_{mt,n}$  é a vazão média de n (2, 3, ..., n) dias anteriores ( $m^3 s^{-1}$ ).

Verifica-se, ainda, que a RNA com as variáveis de entrada  $P_t$ ;  $E_t$ ; e  $Q_{t,1}$  alcançou desempenho superior para a estimativa das vazões diárias, em comparação com os outros modelos de RNAs. Esse resultado foi obtido com a utilização de 200 épocas de treinamento e com quatro e nove neurônios artificiais na primeira e na segunda camada intermediária, respectivamente (Tabela 3).

Todos os modelos de RNAs que apresentaram resultados adequados para a estimativa das vazões na Seção Mário de Carvalho requerem como dado de entrada as vazões de dias anteriores, como pode ser observado na Tabela 3.

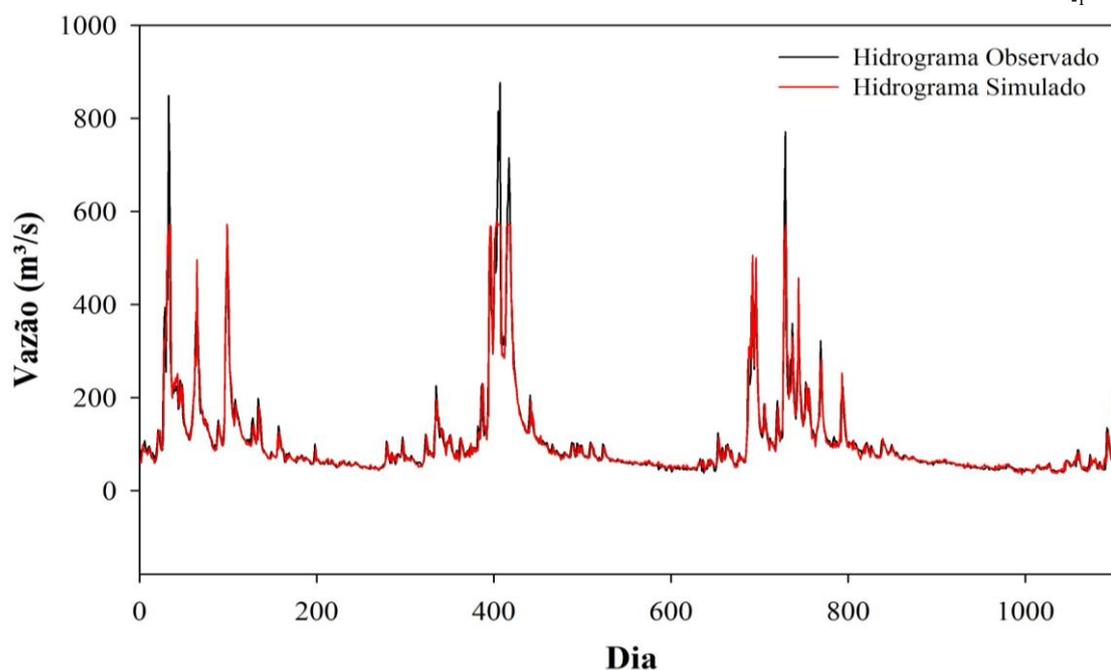
Oliveira et al. (2013) enfatizaram que o bom desempenho de um modelo de RNA para a estimativa de vazões está condicionado à consideração das vazões de dias anteriores como dado de entrada, corroborando os resultados obtidos para a seção Mário de Carvalho. Segundo esses autores, esse resultado se deve ao fato de as

vazões dos dias anteriores representarem o efeito do armazenamento.

O hidrograma das vazões observadas e estimadas com a RNA com as variáveis de entrada  $P_t$ ;  $E_t$ ;  $Q_{t,1}$  (Tabela 3), que apresentaram melhor desempenho, pode ser visualizado na Figura 4. Pela análise do hidrograma verifica-se, de maneira geral, uma boa concordância entre os valores estimados pelo modelo e aqueles observados, porém se observa que a RNA apresentou dificuldade em simular valores máximos. Esse comportamento também foi verificado por Mehr et al. (2015), Tayyab et al. (2016) e Tongal (2013), ao utilizarem RNAs para estimativa de vazões em cursos de água.

Segundo Mehr et al. (2015), a subestimativa das vazões de pico pelas RNAs pode estar relacionada à baixa ocorrência desses eventos utilizados no treinamento da RNA, ao método de normalização dos dados e, como sugere a pesquisa de Tayyab et al. (2016), ao tipo da RNA utilizada.

**Figura 4.** Hidrograma das vazões diárias observadas na seção de controle Mário de Carvalho e estimadas com Rede Neural Artificial com as variáveis de entrada  $P_t$ ;  $E_t$ ;  $Q_{t-1}$ .



## 5.2 Estação Rio Piracicaba (56610000)

Na Tabela 4 estão apresentados os dados referentes à arquitetura, treinamento e desempenho das RNAs com os melhores resultados para a estimativa das vazões diárias na seção de controle Rio Piracicaba (56610000). Verifica-se na tabela que os melhores resultados foram obtidos com a utilização das vazões do curso de água defasadas no tempo e da precipitação média da bacia hidrográfica como dados de entrada.

Os dados apresentados na Tabela 4 confirmam que, na seção de controle Rio Piracicaba, os modelos de RNAs possuem concordância e são bons para a estimativa das vazões, visto que o índice de Willmott ( $d$ ) apresentou valores próximos a unidade e o

índice de eficiência de Nash-Sutcliffe ( $E_{NS}$ ) valores maiores que 0,75.

Os valores do viés indicaram subestimativa das vazões pelas RNAs (Tabela 4). Com relação ao teste  $t$  pareado a 5% de significância, percebe-se que as RNAs apresentaram valor de  $t$  não significativo, ou seja, pode-se afirmar com 95% de probabilidade que a diferença entre os valores observados e estimados com esses modelos é estatisticamente nula (Tabela 4).

Os valores do erro absoluto médio (MAE) e da raiz do erro quadrático médio (RMSE) das RNAs variaram entre 4,42 e 5,66  $m^3 s^{-1}$  e entre 10,39 e 12,90  $m^3 s^{-1}$ , respectivamente. Além disso, o distanciamento entre os valores de MAE e RMSE permite afirmar que a variância dos erros individuais é pequena (Tabela 4).

**Tabela 4.** Variáveis de entrada, quantidade de épocas, número de neurônios das camadas intermediárias ( $N_1$  e  $N_2$ ) e medidas quantitativas do desempenho das melhores Redes Neurais Artificiais (RNAs) desenvolvidas para a estimativa da vazão diária ( $m^3 s^{-1}$ ) na seção de monitoramento Rio Piracicaba

Entrada	$Q_{t,1}; Q_{t,2}$	$Q_{t,1}; P_t$	$Q_{mt,2}; P_t$
Épocas	250	500	500
$N_1$	6	10	8
$N_2$	8	6	4
$\bar{O}$	26,01	26,01	26,01
$\bar{P}$	25,85	25,49	25,64
MAE	5,66	4,47	4,42
RMSE	12,9	12,03	10,39
Viés	0,16	0,52	0,37
d	0,94	0,95	0,96
$E_{NS}$	0,78	0,81	0,86
t	0,49 <sup>(ns)</sup>	1,69 <sup>(ns)</sup>	1,40 <sup>(ns)</sup>

$N_1$  e  $N_2$  são o número de neurônios artificiais na primeira e na segunda camada intermediária, respectivamente;  $\bar{O}$  e  $\bar{P}$  são as vazões médias ( $m^3 s^{-1}$ ) observadas e estimadas, respectivamente; MAE é o erro absoluto médio ( $m^3 s^{-1}$ ); RMSE é a raiz do erro quadrático médio ( $m^3 s^{-1}$ ); d é o índice de concordância de Willmott;  $E_{NS}$  é o índice de eficiência de Nash-Sutcliffe; t é o valor da estatística de teste do teste t pareado; (ns) não significativo a 5% de significância;  $Q_{t-n}$  é a vazão diária defasada em n (1, 2, 3, ..., n) dias ( $m^3 s^{-1}$ );  $P_t$  é a precipitação média da bacia no tempo t (mm) obtida pelo método de Thiessen; e  $Q_{mt,n}$  é a vazão média de n (2, 3, ..., n) dias anteriores ( $m^3 s^{-1}$ ).

Nesta seção de monitoramento, a RNA com as variáveis de entrada  $Q_{mt,2}; P_t$  foi a que alcançou desempenho superior para a estimativa das vazões diárias. Destaca-se novamente que, para o bom desempenho das RNAs, foi preciso considerar como dado de entrada as vazões registradas nos dias anteriores, semelhante ao que ocorreu na seção Mário de Carvalho.

A RNA com melhor desempenho ( $Q_{mt,2}; P_t$ ) foi obtida com 500 épocas de treinamento e com oito e quatro neurônios artificiais na primeira e na segunda camada intermediária, respectivamente (Tabela 4). Verifica-se nesta Seção de Monitoramento, comparativamente à de Mário de Carvalho (Tabelas 3 e 4), que o número de épocas necessário para o treinamento foi maior na seção Rio Piracicaba e a primeira camada

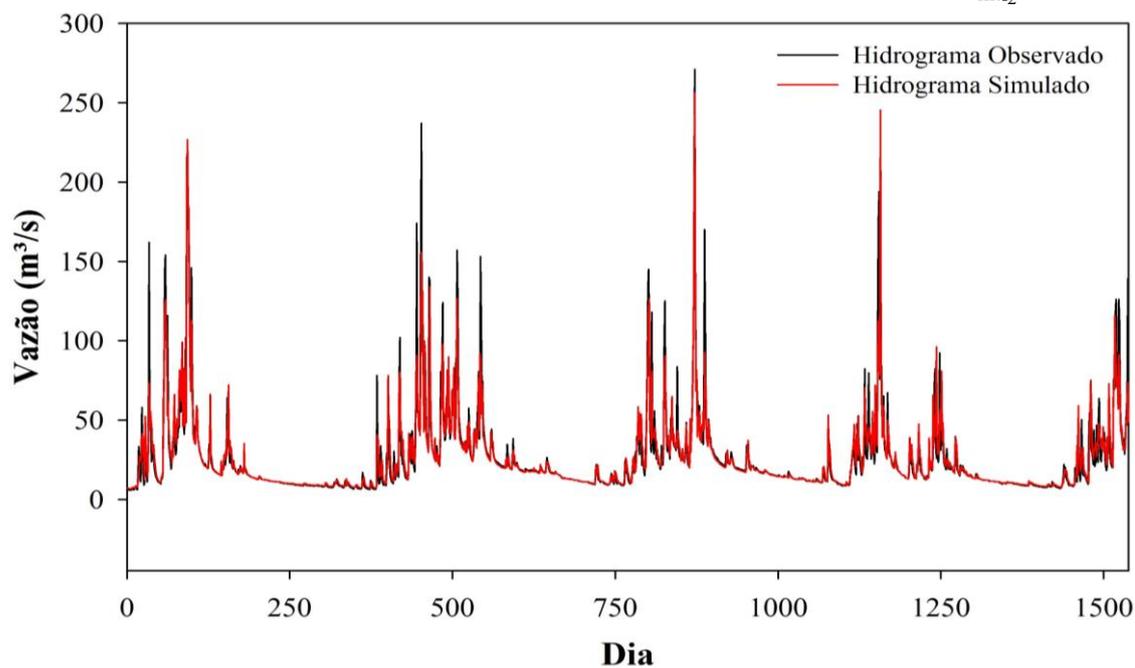
intermediária passou a possuir a maior quantidade de neurônios.

O hidrograma das vazões observadas e estimadas com a RNA com as variáveis de entrada  $Q_{mt,2}; P_t$  pode ser observado na Figura 5.

Pela análise da Figura 5, percebe-se, de maneira geral, uma boa concordância entre os valores estimados pelo modelo e aqueles observados, porém verifica-se certa dificuldade em simular alguns valores do pico de vazão.

Na seção Mário de Carvalho, a rede desenvolvida também não conseguiu simular adequadamente todas as vazões de cheias, e a discussão apresentada anteriormente é também pertinente para a seção Rio Piracicaba.

**Figura 5.** Hidrograma das vazões diárias observadas na seção de controle Rio Piracicaba e estimadas com Rede Neural Artificial com as variáveis de entrada  $Q_{mt,2}$ ;  $P_t$ .



### 5.3 Estação Carrapato (56640000)

Na Tabela 5 podem ser observados os dados referentes à arquitetura e ao treinamento das RNAs desenvolvidas para a

estimativa das vazões diárias na seção de controle Carrapato (56640000), bem como são apresentadas as medidas quantitativas do desempenho dessas RNAs.

**Tabela 5.** Variáveis de entrada, quantidade de épocas, número de neurônios das camadas intermediárias ( $N_1$  e  $N_2$ ) e medidas quantitativas do desempenho das melhores Redes Neurais Artificiais (RNAs) desenvolvidas para a estimativa da vazão diária ( $m^3 s^{-1}$ ) na seção de monitoramento Carrapato

Entrada	$Q_{mt,3}; P_t$	$Q_{mt,4}; P_t$	$Q_{t,1}; Q_{t,2}; Q_{t,3}$	$Q_{mt,2}; Q_{mt,3}; Q_{mt,4}$
Épocas	400	150	250	200
$N_1$	6	4	4	7
$N_2$	4	9	5	7
$\bar{O}$	18,43	18,43	18,43	18,43
$\bar{P}$	17,71	17,62	18,22	17,69
MAE	4,6	4,8	4,62	4,96
RMSE	13,3	13,5	13,47	13,56
Viés	0,73	0,81	0,21	0,74
d	0,89	0,88	0,9	0,89
$E_{NS}$	0,68	0,67	0,68	0,67
t	1,49 <sup>(ns)</sup>	1,65 <sup>(ns)</sup>	0,43 <sup>(ns)</sup>	1,49 <sup>(ns)</sup>

$N_1$  e  $N_2$  são o número de neurônios artificiais na primeira e na segunda camada intermediária, respectivamente;  $\bar{O}$  e  $\bar{P}$  são as vazões médias ( $m^3 s^{-1}$ ) observadas e estimadas, respectivamente; MAE é o erro absoluto médio ( $m^3 s^{-1}$ ); RMSE é a raiz do erro quadrático médio ( $m^3 s^{-1}$ ); d é o índice de concordância de Willmott;  $E_{NS}$  é o índice de eficiência de Nash-Sutcliffe; t é o valor da estatística de teste do teste t pareado; (ns) não significativo a 5% de significância;  $Q_{t,n}$  é a vazão diária defasada em n (1, 2, 3, ..., n) dias ( $m^3 s^{-1}$ );  $P_t$  é a precipitação média da bacia no tempo t (mm) obtida pelo método de Thiessen; e  $Q_{mt,n}$  é a vazão média de n (2, 3, ..., n) dias anteriores ( $m^3 s^{-1}$ ).

Observa-se na Tabela 5 que as RNAs que apresentaram bons resultados para a estimativa das vazões diárias possuem como variáveis de entrada as vazões do curso de água defasadas no tempo e a precipitação média da bacia hidrográfica, assim como para a seção Rio Piracicaba.

Os modelos de RNAs desenvolvidos apresentam concordância e são satisfatórios para a estimativa das vazões, visto que o índice de Willmott (d) exibiu valores próximos a 0,9 e o índice de eficiência de Nash-Sutcliffe ( $E_{NS}$ ), valores variando entre 0,67 e 0,68 (Tabela 5).

Os valores do erro absoluto médio (MAE) e da raiz do erro quadrático médio (RMSE) das RNAs variaram entre 4,60 e 4,96  $m^3 s^{-1}$  e entre 13,30 e 13,56  $m^3 s^{-1}$ , respectivamente (Tabela 5).

A RNA com melhor desempenho foi a que teve como variáveis de entrada a vazão média dos três dias anteriores e a precipitação ( $Q_{mt,3}; P_t$ ). Observa-se na Tabela 5 que para o desenvolvimento dessa RNA foram

necessárias 400 épocas de treinamento, além de seis e quatro neurônios artificiais na primeira e na segunda camada intermediária, respectivamente. O número de neurônios, requeridos nas camadas intermediárias dessa RNA, ficou menor que o máximo estabelecido nesse trabalho (10 neurônios artificiais).

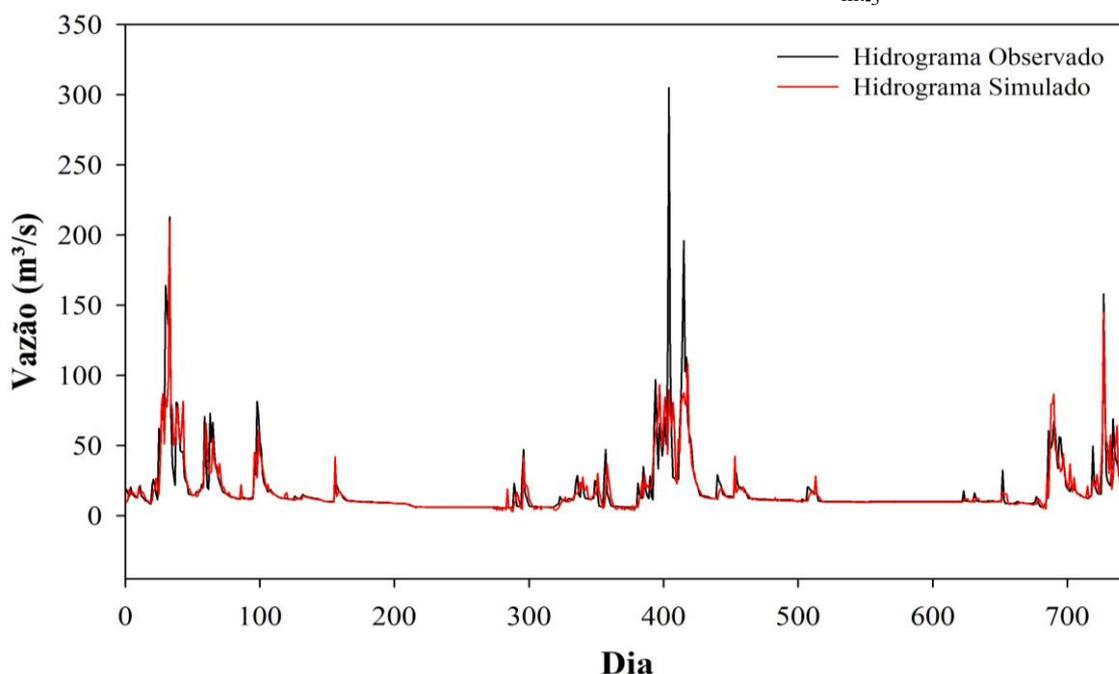
Ao comparar a quantidade de épocas e neurônios artificiais nas camadas intermediárias da seção de monitoramento Carrapato com a obtida para as outras duas seções, verificou-se que a da seção Carrapato apresentou valores próximos aos da seção Rio Piracicaba. Essas duas seções possuem áreas de drenagem menores que a de Mário de Carvalho, permitindo sugerir que, quanto menor a área de drenagem, maior também a quantidade de épocas de treinamento e de neurônios artificiais na primeira camada intermediária.

O hidrograma das vazões observadas e estimadas com a RNA que apresentou o melhor desempenho para predição das vazões pode ser visualizado na Figura 6. Essa

RNA requer como dados de entrada a vazão média dos três dias anteriores ( $Q_{mt,3}$ ) e a

precipitação média ( $P_t$ ), como pode ser observado na Tabela 5.

**Figura 6.** Hidrograma das vazões diárias observadas na seção de controle Carrapato e estimadas com Rede Neural Artificial com as variáveis de entrada  $Q_{mt,3}$ ;  $P_t$ .



A análise do hidrograma da Figura 6 permite constatar que o modelo apresentou maior dificuldade para simular vazões de cheias, porém se verificou, em geral, boa concordância entre os valores estimados pelo modelo e aqueles observados. As causas dessa maior dificuldade para a estimativa de alguns valores de pico são as mesmas que foram apresentadas para as seções de monitoramento Mário de Carvalho e Rio Piracicaba, mas percebe-se, nesse caso, uma dificuldade ainda maior, a qual provavelmente decorre da menor área de drenagem associada à seção Carrapato.

Ao analisar os valores apresentados nas Tabelas 3, 4 e 5, verifica-se que as RNAs tiveram melhor desempenho para estimar as vazões na Seção Mário de Carvalho, com área de drenagem de 5.304,0 km<sup>2</sup>, seguida da Estação Rio Piracicaba, com área de 1.164,2 km<sup>2</sup>; e, por fim, o pior desempenho para determinação das vazões na Seção de Monitoramento Carrapato, com área de

drenagem de 427,5 km<sup>2</sup>. Desse modo, percebe-se que o desempenho do modelo foi melhor à medida que a área de drenagem aumentou.

A escala de tempo de simulação, a alta variabilidade espacial e temporal da chuva e o rápido deslocamento do escoamento superficial em bacias hidrográficas menores podem ter ocasionado o decréscimo do desempenho dos modelos de RNAs nas bacias com menor área de drenagem.

Segundo Viola et al. (2009), quando a simulação de vazões é executada nas escalas de tempo diária ou mensal, principalmente em pequenas e médias bacias hidrográficas tropicais, torna-se mais difícil obter resultados com boa exatidão, devido à alta variabilidade espacial e temporal da precipitação, bem como ao rápido deslocamento do escoamento superficial.

As RNAs desenvolvidas para as seções de monitoramento de vazão da bacia

do rio Piracicaba apresentaram desempenho adequado e semelhante aos obtidos por Aichouri et al. (2015), Elsafi (2014), Mehr et al. (2015), Okkan et al. (2012), Setiono (2015), Tayyab et al. (2016) e Tongal et al. (2013), quando da utilização de RNAs com o mesmo objetivo, em bacias hidrográficas de diferentes países.

Os resultados apresentados indicaram que o uso da precipitação média como dado de entrada é importante, mas não suficiente para a estimativa das vazões diárias, sendo necessários os dados de vazões defasadas no tempo e de evapotranspiração de referência, em alguns casos. Esses resultados são corroborados pelos obtidos por Sattari; Apaydin; Ozturk (2012), que concluíram que a precipitação média da bacia hidrográfica não é suficiente para a estimativa das vazões, sendo necessários dados de temperatura e vazões defasadas.

Destaca-se, entretanto, que no caso das três seções da bacia do rio Piracicaba a utilização dos dados de temperatura não proporcionou resultados satisfatórios para a estimativa das vazões diárias.

A partir da análise dos resultados deste estudo, percebe-se que as RNAs possuem potencial para serem utilizadas na previsão de vazão de curto prazo e no preenchimento de falhas de séries históricas.

## 6 CONCLUSÕES

Com base nos resultados, pode-se concluir que:

- As Redes Neurais Artificiais (RNAs) são adequadas para estimativa das vazões diárias;
- As variáveis de entrada das Redes Neurais Artificiais (RNAs) que proporcionam os melhores resultados para estimativa das vazões diárias são precipitação, evapotranspiração de referência e vazões do próprio curso de água defasadas no tempo; e
- O desempenho das RNAs para estimativa das vazões diárias está condicionado à consideração da vazão do curso de água defasada no tempo como variável de entrada;
- As RNAs apresentam melhor desempenho para a estimativa de vazões em bacias hidrográficas com maior área de drenagem; e
- As RNAs desenvolvidas são adequadas para a estimativa das vazões diárias da bacia do rio Piracicaba.

## 7 AGRADECIMENTOS

Os autores agradecem a Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio a pesquisa.

## 8 REFERÊNCIAS

- AICHOURI, I.; HANI, A.; BOUGHERIRA, N.; DJABRI, L.; CHAFFAI, H.; LALLAHEM, S. River flow model using artificial neural networks. **Procedia Engineering**, Oxford, v. 74, n. 1, p. 1007-1014, 2015.
- ARAÚJO, C. B. C. D.; DANTAS NETO, S. A. ; SOUZA FILHO, F. D. A. Streamflow forecasting for the dam Orós/CE from hydrometeorological data using perceptrons. **Revista Brasileira de Meteorologia**, São José dos Campos, v. 30, p. 37-46, 2015.
- ASADI, S.; SHAHRABI, J.; ABBASZADEH, P.; TABANMEHR, S. A new hybrid artificial neural networks for rainfall-runoff process modeling. **Neurocomputing**, Amsterdam, v. 121, p. 470-480, 2013.

BORGES JÚNIOR, J. C. F.; ANJOS, R. J.; SILVA, T. J. A.; LIMA, J. R. S.; ANDRADE, C. L. T. Métodos de estimativa da evapotranspiração de referência diária para a microrregião de Garanhuns, PE. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v. 16, n. 4, p. 380-390, 2012.

ELSAFI, S. H. Artificial Neural Networks (ANNs) for flood forecasting at Dongola Station in the river Nile, Sudan. **Alexandria Engineering Journal**, Alexandria, v. 53, n. 3, p. 655-662, 2014.

GOVINDARAJU, R. S. Artificial neural networks in hydrology I: preliminary concepts. **Journal of Hydrologic Engineering**, Reston, v. 5, n. 2, p. 115-123, 2000a.

GOVINDARAJU, R. S. Artificial neural networks in hydrology II: hydrologic applications. **Journal of Hydrologic Engineering**, Reston, v. 5, n. 2, p. 124-137, 2000b.

LUDWIG JÚNIOR, O.; COSTA, E. M. M. **Redes neurais: fundamentos e aplicações com programas em C**. Rio de Janeiro: Editora Ciência Moderna, 2007.

MACÊDO, M. N. C.; DIAS, H. C. T.; COELHO, F. M. G.; ARAÚJO, E. A.; SOUZA, M. L. H.; SILVA, E. Precipitação pluviométrica e vazão da bacia hidrográfica do Riozinho do Rôla, Amazônia Ocidental. **Revista Ambiente & Água**, Taubaté, v. 8, n. 1, p. 206-221, 2013.

MEHR, A. D.; KAHYA, E.; SAHIN, A.; NAZEMOSADAT, M. J. Successive-station monthly streamflow prediction using different artificial neural network algorithms. **International Journal of Environmental Science and Technology**, Tehran, v. 12, n. 7, p. 2191-2200, 2015.

MENG, X.; YIN, M.; NING, L.; LIU, D.; XUE, X. A threshold artificial neural network model for improving runoff prediction in a karst watershed. **Environmental Earth Sciences**, Berlin, v. 74, n. 6, p. 5039-5048, 2015.

NOORI, N.; KALIN, L. Coupling SWAT and ANN models for enhanced daily streamflow prediction. **Journal of Hydrology**, Amsterdam, v. 533, n. 1, p. 141-151, 2016.

OKKAN, U. Wavelet neural network model for reservoir inflow prediction. **Scientia Iranica**, Tehran, v. 19, n. 6, p. 1445-1455, 2012.

OLIVEIRA, G. G.; PEDROLLO, O. C.; CASTRO, N. M. R.; BRAVO, J. M. Simulações hidrológicas com diferentes proporções de área controlada na bacia hidrográfica. **Revista Brasileira de Recursos Hídricos**, Porto Alegre, v. 18, n. 3, p. 193-204, 2013.

PEREIRA, D. R.; ULIANA, E. M.; MARTINEZ, M. A.; SILVA, D. D. Desempenho de um Modelo Hidrológico Concentrado e de um Semidistribuído na Predição de Vazões Diárias. **Irriga**, Botucatu, v. 21, n. 2, p. 409-424, 2016.

REZAEI, M.; MOTLAQ, A. A. A.; MAHMOUEI, A. R.; MOUSAVI, S. H. River Flow Forecasting using artificial neural network (Shoor Ghaen). **Ciência e Natura**, Santa Maria, v. 37, p. 207-215, 2015.

SATTARI, M. T.; APAYDIN, H.; OZTURK, F. Flow estimations for the Sohu Stream using artificial neural networks. **Environmental Earth Sciences**, Berlin, v. 66, n. 7, p. 2031-2045, 2012.

SETIONO, R. H. Analysis of rainfall-runoff neuron input model with artificial neural network for simulation for availability of discharge at Bah Bolon Watershed. **Procedia Engineering**, Oxford, v. 125, n. 1, p. 150-157, 2015.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais: para engenharia e ciências aplicadas**. São Paulo: Editora Artliber, 2010.

TAYYAB, M.; ZHOU, J.; ZENG, X.; ADNAN, R. Discharge forecasting by applying artificial neural networks at the Jinsha river Basin, China. **European Scientific Journal**, Kocani, v. 12, n. 9, p. 108-127, 2016.

TONGAL, H. Nonlinear forecasting of stream flows using a chaotic approach and artificial neural networks. **Earth Sciences Research Journal**, Bogotá, v. 17, n. 2, p. 119-126, 2013.

UYSAL, G.; SORMAN, A. A.; SENSOY, A. Streamflow forecasting using different neural network models with satellite data for a snow dominated region in Turkey. **Procedia Engineering**, Oxford, v. 154, p. 1185-1192, 2016.

VIOLA, M. R.; MELLO, C. R.; ACERBI JUNIOR, F. W. Modelagem hidrológica na bacia hidrográfica do Rio Aiuruoca, MG. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v. 13, n. 5, p. 581-590, 2009.