

<http://dx.doi.org/10.1590/2318-0331.011615124>

Eficácia da arquitetura MLP em modo closed-loop para simulação de um Sistema Hidrológico

Efficiency of MLP architecture on closed-loop mode for the simulation of a hydrological system

Aline Bernarda Debastiani¹, Ricardo Dal'Agnol da Silva² e Sílvio Luíz Rafaeli Neto¹

¹Universidade do Estado de Santa Catarina, Lages, SC, Brasil

²Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP, Brasil

E-mails: aline.debastiani@gmail.com (ABD), ricds@hotmail.com (RDS), silvio.rafaeli@udesc.br (SLRN)

Recebido: Agosto 06, 2015 - Revisado: Julho 15, 2016 - Aceito: Agosto 09, 2016

RESUMO

Para a elaboração do plano de bacia se faz necessária a realização de estimativas da resposta hidrológica. Sendo assim, o objetivo desse estudo foi de avaliar a simulação do comportamento hidrológico da bacia hidrográfica do Alto Canoas localizada em Santa Catarina, através de redes neurais artificiais *Multi Layer Perceptron* (MLP), bem como de analisar a contribuição das variáveis de entrada para a modelagem. Foram testados doze tratamentos com combinações de variáveis de precipitação, evapotranspiração (ET0) e vazão, além de transformações e deslocamentos temporais dessas, a fim de determinar as variáveis que promovessem o melhor desempenho da modelagem da vazão. A MLP foi treinada em modo *open-loop* utilizando parte das vazões observadas. As vazões foram simuladas em *open-loop* e *closed-loop* para o período de teste, sendo em *closed-loop* utilizado a vazão simulada no passo de tempo anterior como entrada. O algoritmo de aprendizado utilizado foi o de *Levenberg-Marquardt*. O tratamento que apresentou melhor desempenho (Nash e Sutcliffe (NS) = 0,9119, Raiz do Erro Médio Quadrático (RMS) = 14,29 m³/s) empregou a precipitação diária das quatro estações pluviométricas (Urubici, Vila Canoas, Lomba Alta e Anitápolis), precipitação das quatro estações com tempo de resposta de -2 dias, e vazão simulada do dia anterior. Apesar do baixo RMS, a vazão modelada pela MLP foi, em geral, superestimada.

Palavras-chave: Modelo chuva-vazão; Algoritmo de Garson; Bacia hidrográfica do Rio Canoas.

ABSTRACT

Estimates of hydrological responses are needed for the watershed planning. The aim of this study was to evaluate the hydrological behavior simulation of the Upper Canoas basin using artificial neural networks *Multi Layer Perceptron* (MLP) method, as well as to analyze the contribution of the input variables for modeling. It were tested 12 treatments with combinations of variables such as precipitation, evapotranspiration (ET0) and discharge, as well as transformations and temporal displacements of these variables, in order to determine the variables that promoted the better performance on discharge modeling. The MLP was trained in *open-loop* mode using part of the observed discharges. The discharges for the whole series were simulated in *closed-loop*, using the discharge simulated on the previous time step as input. The learning algorithm used was the *Levenberg-Marquardt*. The treatment with the best performance (NS = 0.9119, RMS = 14.29 m³/s) employed the daily precipitation of the four rainfall stations (Urubici, Vila Canoas, Lomba Alta e Anitápolis), precipitation of the four stations with -2 days of response time, and simulated discharge from the previous day. Despite the low RMS, the modeled discharge using MLP was generally overestimated.

Keywords: Rainfall-runoff model; Garson algorithm; Basin of Canoas.



INTRODUÇÃO

Uma etapa importante do planejamento de recursos hídricos é a projeção hidrológica de médio e longo prazo, a partir de projeções climáticas. De acordo com Pooch e Wall (1993), o modelo de um sistema pode ser aplicado para prever comportamentos e auxiliar no planejamento de ações futuras, sintetizar e avaliar alternativas de solução do problema, avaliar as magnitudes das variáveis do sistema e realizar prescrições que tragam efeitos desejados sobre as saídas do sistema. Neste contexto, as projeções do comportamento de sistemas hidrológicos podem ser aplicadas como suporte na etapa de prognóstico da disponibilidade hídrica, durante o processo de construção de um plano de bacia. No caso de Santa Catarina, nenhum dos planos de bacia, propostos ou em construção até o momento, prevê o uso de projeções de médio e longo prazo para a estimativa da disponibilidade hídrica, por meio de modelos científicos dos sistemas hidrológicos. Carecem, portanto, de representatividade e rigor técnico nestas projeções. Algumas razões que podem ser apontadas são a complexidade dos sistemas hidrológicos, o desconhecimento de qual modelo melhor representaria um determinado sistema, a dificuldade em encontrar o melhor modelo e o próprio despreparo técnico das equipes que elaboram estes planos, quanto à modelagem destes sistemas.

Os modelos conceituais de base física procuram representar os sistemas hidrológicos como sistemas caixas-cinzas. Estes modelos simplificam a realidade hidrológica da bacia por meio de armazenamentos e transferências hídricas em diferentes graus de detalhamento conceitual e de discretização das unidades espaciais de resposta hidrológica. Estes modelos tendem a ser complexos quando se procura modelar sistemas hídricos visando os planos de bacias. A necessidade de prognósticos factíveis de disponibilidade hídrica, por vezes, exige que os sistemas hidrológicos sejam discretizados em patamares de complexidade que podem dificultar, ou mesmo inviabilizar, uma solução técnica.

Os modelos de base numérica, por outro lado, abordam o sistema hídrico como caixa-preta. Os componentes do sistema, os seus comportamentos, as suas relações internas e os seus reguladores internos são desconhecidos. Contudo, o produto do comportamento do sistema é determinado a partir de insumos fornecidos ao modelo (RAFAELI NETO, 2000). O sistema é significativamente simplificado, ao ponto de representa-lo por meio de relações numéricas entre entradas e saídas.

As Redes Neurais Artificiais (RNAs) que utilizam o método *Multi Layer Perceptron* (MLP) têm sido cada vez mais estudadas na Hidrologia. As MLPs possuem capacidade de modelar sistemas não-lineares sem a necessidade de realizar suposições sobre a interação das variáveis que compõem tais sistemas (RIAD et al., 2004).

A estrutura da MLP é constituída por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída (FU, 1994). A configuração da MLP em *open-loop* é comumente estudada em simulação de sistemas hidrológicos e outros estudos de modelagem, o que consiste em fornecer dados observados no tempo $t-1$ como entrada para simular a saída no tempo t . No entanto, uma variação a esta abordagem chamada de *closed-loop* consiste em fornecer como entrada algum valor simulado em um tempo $t-1$ (ABRAHART; KNEALE; SEE, 2004). Uma MLP em modo *closed-loop* permite fazer simulações de vazão a longo prazo,

utilizando como entrada a precipitação resultante de modelos climáticos por exemplo, e a vazão simulada pela própria MLP, limitação essa encontrada no modelo em *open-loop* visto que utiliza a vazão observada no tempo $t-1$ (Q_{t-1}) a qual deve ser inserida na MLP a cada passo de tempo, tornando o modelo estático.

A combinação de vetores na entrada da RNA tem se demonstrado um fator relevante no sucesso do modelo para simulação hidrológica. Santos e Pereira Filho (2003) concluíram que o conjunto de vetores composto pelas variáveis de precipitação diária observada no tempo t (P_t) e vazão observada no tempo $t-1$ (Q_{t-1}) produziu coeficiente de determinação (R^2) acima de 0,93, ao passo que o uso de apenas P_t produziu um R^2 máximo de 0,409 na bacia hidrográfica do Rio Tamanduateí (SP). Sousa e Sousa (2010) obtiveram um R^2 de 0,92, também na estimativa de vazão média mensal na bacia hidrográfica do Rio Piancó (PB), ao aplicarem vetores de entrada compostos por combinações de cinco estações pluviométricas relacionadas com a área de estudo.

O uso de dados progressos de uma variável parece melhorar o desempenho da RNA na modelagem de sistemas hidrológicos. Riad et al. (2004) estudando a bacia hidrográfica de Ourika no Marrocos (503 km²), obtiveram R^2 significativo de 0,917 ao comporem os vetores de entrada com as variáveis de precipitação e escoamento superficial dos 7 dias precedentes ao dia t , a partir de séries históricas de sete anos. Oliveira et al. (2013) avaliaram o uso da RNA para previsão de vazão diária em diferentes proporções e características de área controlada na região hidrográfica do Guaíba (RS) (967 a 71.401 km²). Os autores utilizaram dados de vazão média diária de 25 postos fluviométricos, os quais foram combinados em 83 vetores de entrada da RNA. Avaliaram ainda o desempenho destes vetores incluindo-se o valor da vazão no tempo $t-1$ (Q_{t-1}). Os autores afirmam que a inclusão da Q_{t-1} elevou o coeficiente de Nash e Sutcliffe (NS) de 0,75 para 0,94. Oliveira, Pedrollo e Castro (2014) avaliaram o desempenho da simulação da vazão média mensal da bacia do Rio Ijuí (RS) (5.414 km²) por RNA. Utilizaram dados meteorológicos e pluviométricos de 1941 a 2005. A melhor composição de vetores foi formada por três variáveis de entrada (balanço médio de água nos tempos t e $t-1$; média móvel dos valores passados; filtro do balanço de água) com NS de 0,904 e Raiz do Erro Médio Quadrático (RMS) de 37,1 m³/s. Com a RNA treinada, esta foi aplicada para simulações no período de 2002 a 2009 e produziu um NS de 0,89, enquanto outro modelo hidrológico, *Soil Moisture Accounting Procedure* (SMAP), alcançou 0,67.

Estes estudos mostram que as RNAs são eficazes para simular eventos hidrológicos, especialmente para a previsão de vazões. Portanto, a inclusão de observações realizadas nas épocas antecedentes ao evento em análise mostra-se um recurso importante para o sucesso da simulação. Isso porque a resposta hidrológica em uma bacia hidrográfica, após episódios de precipitação, depende de processos hidrológicos complexos (interceptação, infiltração, percolação, armazenamentos, etc.), de características morfométricas (declividades, comprimento da drenagem, solos, etc.) que ajudam a definir o tempo de concentração e a “memória” do sistema hidrológico representado pela bacia

Apesar da variedade de estudos que procuram determinar a melhor composição dos vetores de entrada, há uma carência de determinação do efeito que cada vetor produz na simulação.

Esta análise possibilita identificar não apenas a melhor combinação, mas também qual vetor se mostra mais relevante no conjunto de entrada utilizado. O algoritmo de Garson (1991) é comumente empregado com este propósito.

Este trabalho teve por objetivo analisar a eficácia de uma MLP em configurações *open* e *closed-loop* para modelagem de um sistema hidrológico e avaliar a influência da utilização de variáveis hidroclimáticas e suas variações temporais para a modelagem hidrológica.

MATERIAIS E MÉTODOS

Área de estudo

A área de estudo escolhida é uma sub-bacia da bacia hidrográfica do Rio Canoas, denominada bacia do Alto Canoas, em cujo exutório se encontra a estação fluviométrica Rio Bonito (71300000), operada pela Agência Nacional de Águas (ANA). Esta bacia, localizada no Estado de Santa Catarina (SC), foi selecionada por sua importância hidrológica para a bacia do Rio Canoas. Nela se encontra a nascente do principal rio formador do Rio Uruguai junto a Serra do Corvo Branco, no município de Urubici. Está bem preservada em termos ambientais e é considerada uma importante zona de recarga do Aquífero Guarani (Figura 1).

A bacia do Alto Canoas apresenta uma área de cerca de 1980 km² cujo território abrange os municípios de Urubici, Rio Rufino e Bom Retiro, em SC. Nesta bacia encontram-se as estações pluviométricas de Urubici (P1) (02849021) e Vila Canoas (P2) (02749031), cujas áreas de influências, calculadas pelo método dos polígonos de Thiessen, correspondem respectivamente a 47% e 34% da área total da bacia. As estações de Lomba Alta (P3) (02749007), com 16%, e Anitápolis (P4) (02749065), com aproximadamente 3%, também foram consideradas.

Variáveis das camadas de entrada e saída

O conjunto de dados consistiu nas variáveis diárias de precipitação (P), vazão (Q) e evapotranspiração de referência (ET_0), no período de 01/01/1996 a 31/12/1999 (Figura 2). Este período foi selecionado em virtude da disponibilidade ininterrupta dos dados diários. Os dados diários de precipitação e de vazão foram obtidos no portal HidroWeb da ANA (<http://hidroweb.ana.gov.br/>).

A ET_0 foi calculada pelo método de Penman-Monteith-Padrão (ALLEN et al., 1998) a partir das séries diárias de temperatura, umidade relativa, radiação solar e velocidade do vento, todos estes obtidos da Estação Climatológica de Lages, localizada a 84°NO e cerca de 75 km do ponto central da bacia e com altitude de 937 m, pertencente à Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina (EPAGRI) e Instituto Nacional de Meteorologia (INMET). Esta estação foi

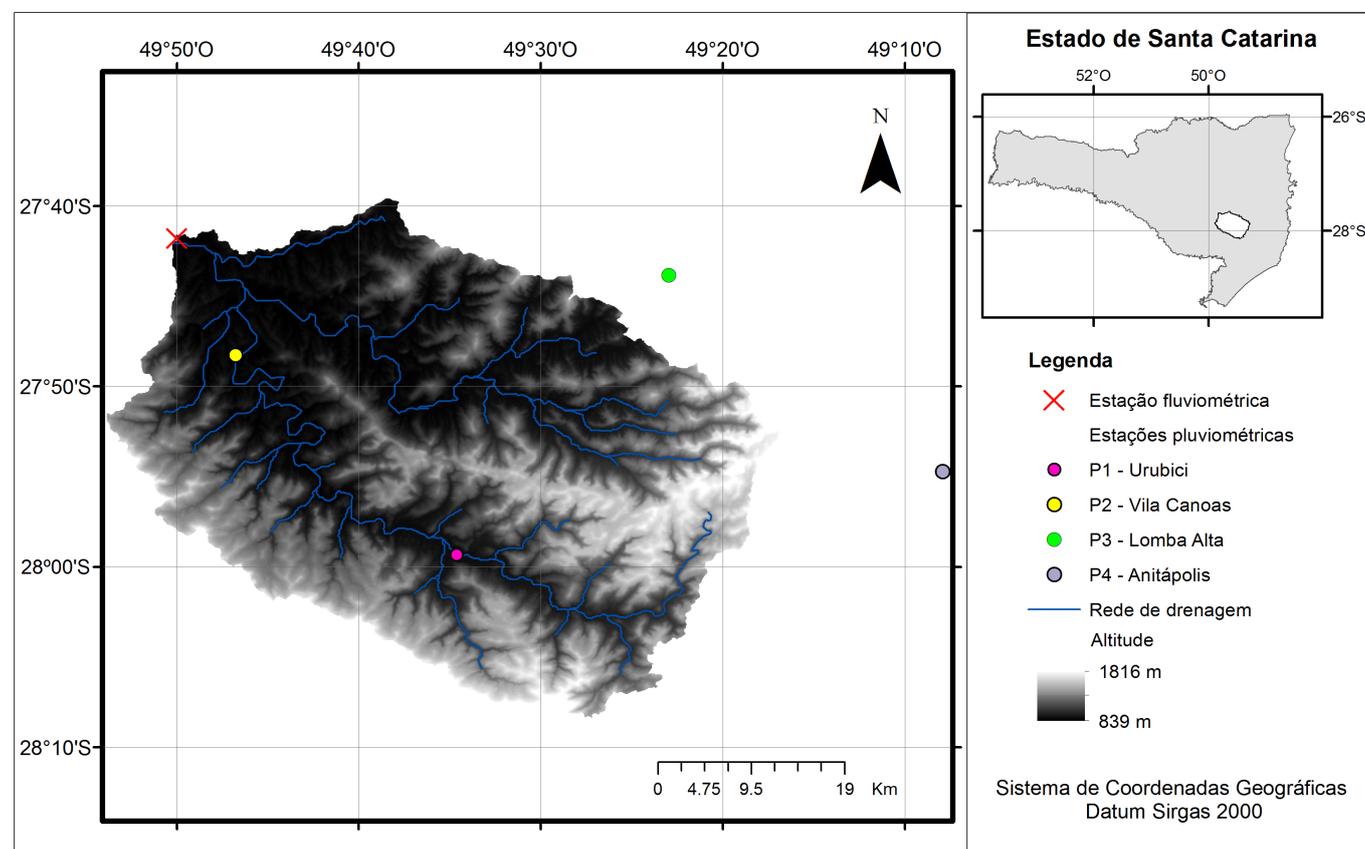


Figura 1. Modelo Digital do Terreno (MDT) referente à localização da bacia hidrográfica do Alto Canoas, estação Rio Bonito e estações pluviométricas utilizadas no estudo.

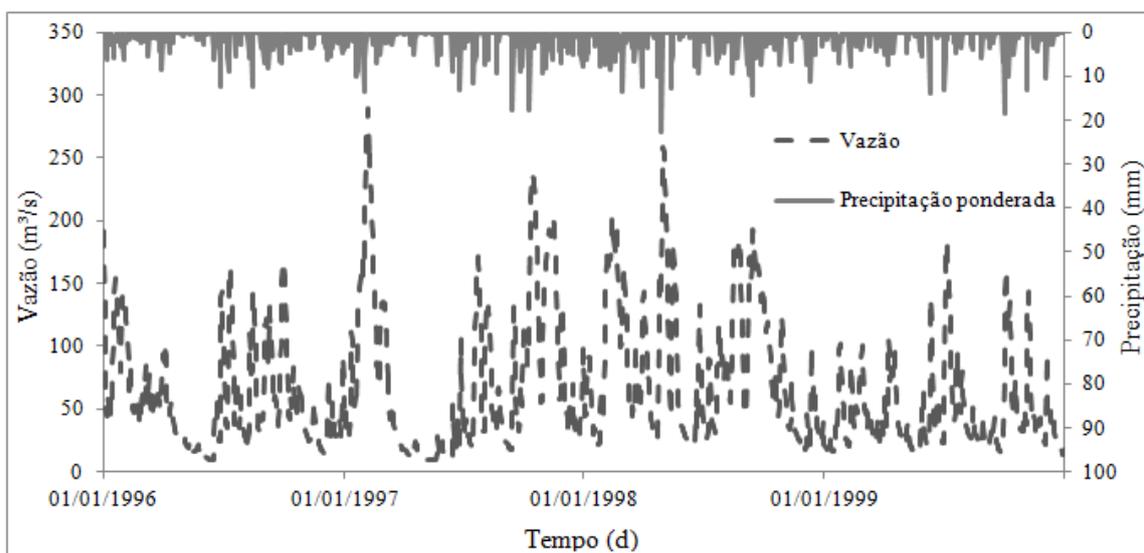


Figura 2. Série histórica de vazão, precipitação e evapotranspiração (ET0) compreendendo os anos de 1996, 1997, 1998 e 1999 utilizado para este estudo.

escolhida por ser a estação mais próxima da área de estudo, contar com uma série contínua de dados e ainda por situar-se próxima à latitude do ponto central da bacia e sua altitude encontrar-se dentro da faixa de altitudes da bacia do Alto Canoas.

Foram definidos doze tratamentos para compor os vetores de entrada e assim avaliar a influência das variáveis sobre a modelagem: precipitação das estações dentro da bacia (P1 e P2) e fora da bacia (P3 e P4), deslocamento temporal de precipitação e vazão, ET0, precipitação ponderada pelo peso da área de contribuição definido pelos polígonos de Thiessen. Os valores de deslocamento temporal das variáveis foram suportados pela análise exploratória.

A vazão média (Q) diária no exutório da bacia foi utilizada como vetor de saída.

Extração de informação

Foi realizada uma análise exploratória com o objetivo de identificar os processos-chaves ligados às variáveis de entrada que influenciam as vazões no exutório da bacia estudada. As variáveis utilizadas na MLP foram analisadas através do coeficiente de correlação de Pearson (R) (Equação 1) e pela função de correlação cruzada (CC) (Equação 2).

$$R = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\sum (x_i - \bar{x})^2\right)\left(\sum (y_i - \bar{y})^2\right)}} \quad (1)$$

$$CC_{y_1 y_2} = \frac{C_{y_1 y_2}^{(k)}}{s_{y_1} s_{y_2}} \quad (2)$$

em que: x_i e y_i : variáveis analisadas; \bar{x} e \bar{y} : valor médio das variáveis analisadas; x : valor das observações de entrada; k : defasagens no

tempo (0, -1, -2, -3, ...); S_{y_1} e S_{y_2} : desvio padrão da amostra da série; C_{y_1} e C_{y_2} : covariância da amostra da série.

A análise de correlação cruzada é a autocovariância padronizada que avalia a relação temporal entre duas variáveis, com defasagens no tempo. Assim sendo, aplicou-se a CC entre as variáveis e a vazão para uma defasagem de 10 dias e considerando intervalos de confiança de 95% em torno da média.

Estrutura da simulação da resposta hidrológica em open-loop e closed-loop

O período de estudo (01/01/1996 a 31/12/1999) foi dividido em três subperíodos para constituir os conjuntos amostrais de treinamento (50% dos dados), validação (25% dos dados) e teste (25% dos dados), estes conjuntos de dados seguiram a divisão da ordem temporal, tendo em vista que é uma condicionante para a arquitetura em *closed-loop*.

O método MLP foi aplicado pela *Neural Network Toolbox* no *software* MATLAB (2014) e sua arquitetura constituíram-se de três camadas (Figura 3): (1) a camada de entrada, constituída pelas combinações de vetores (X), em que cada vetor corresponde a um neurônio da MLP; (2) a camada oculta, constituída por um número variável de neurônios (U), conforme o conjunto de vetores de entrada, sendo esta variação definida por testes de amplitude de neurônios; (3) a camada de saída, com um único neurônio (Y), correspondente ao vetor Q_i .

O objetivo do treinamento foi determinar os valores dos pesos sinápticos através do algoritmo de aprendizado de *Levenberg-Marquardt*. Esse foi escolhido devido a sua rapidez e convergência estável, sendo basicamente uma integração dos métodos clássicos de *Error Back Propagation* e *Gauss-Newton* (HAGAN; MENHAJ, 1994). Utilizou-se função de ativação tangente hiperbólica sigmoideal (*tansig*), a qual comprime a resposta em um intervalo conhecido, de -1 a 1. A taxa de aprendizado foi

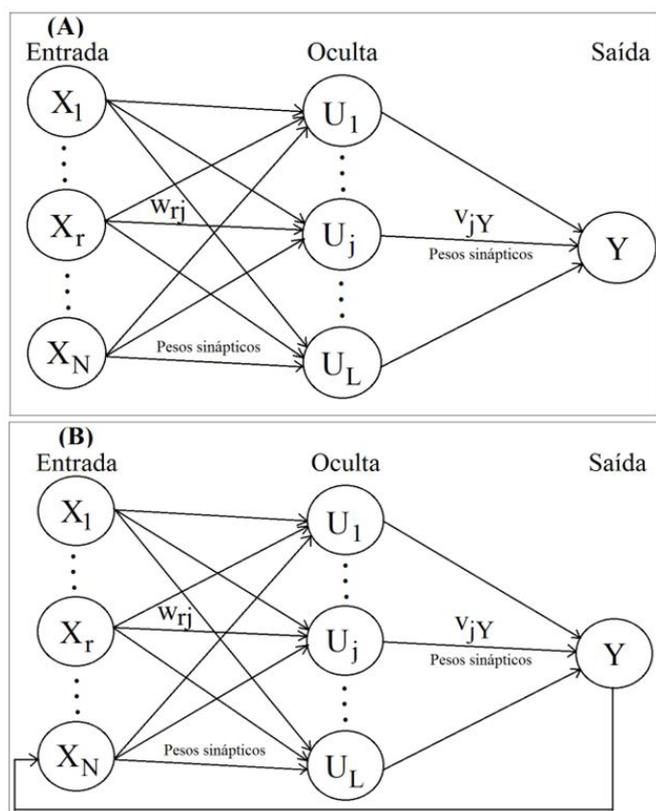


Figura 3. Arquitetura da Multi Layer Perceptron (MLP) com configuração (A) *open-loop*; e (B) *closed-loop*.

definida como 0,01. Os vetores de entrada foram padronizados para uma mesma escala, média igual a 0 e desvio padrão igual a 1, a fim de melhorar a convergência (FU, 1994). Uma relação inicial de 10 vezes mais equações de treinamento (Equação 3) do que de pesos (Equação 4) foi assumida com o objetivo de determinar o número de neurônios da camada oculta (HEATH, 2010).

$$n_Eq_Treino = n_Amostra_Treino * n_Classes \quad (3)$$

$$n_Pesos = ((n_Atrib + 1) * n_Neur) + ((n_Neur + 1) * n_classes) \quad (4)$$

$$n_Neur = \frac{(-1) + \frac{n_Eq_Treino}{(n_Atrib + n_classes + 1)}}{razão} \quad (5)$$

em que: *n_Amostra_Treino*: número de amostras de treino; *n_classes*: número de classes de saída; *n_Atrib*: número de vetores de entrada; *n_Neur*: número de neurônios ótimo na camada oculta (Equação 5); *razão*: relação de 10 vezes.

A partir deste valor, testou-se uma variação de ± 5 neurônios na camada oculta, a fim de buscar a arquitetura que proporcionasse a melhor solução para o problema. Também foram testadas dez inicializações aleatórias dos pesos sinápticos.

A configuração *open-loop* (Figura 3A) foi utilizada na etapa de treinamento, na qual os vetores de entrada consistiram sempre em dados observados. Durante o processo de treinamento, o desempenho de aprendizado da rede foi monitorado pela validação cruzada utilizando-se o subconjunto de validação, a fim

de determinar o momento em que não houvesse mais ganho de performance e interrupção do treinamento.

A partir da rede já treinada realizou-se a simulação da vazão para o período de teste com configuração em *open-loop* e *closed-loop* (Figura 3B). Em *closed-loop* os valores de $Q_{(t-1)}$ são os valores que a própria MLP calculou, ou seja, os valores simulados. O processo começa na data inicial do período da simulação, adotando-se a primeira vazão pelo valor observado, e prossegue com as demais vazões pelos valores simulados. Dessa forma, na medida em que se avança o passo de tempo de um dia, as vazões $Q_{(t-1)}$ são simuladas em série encadeadas, de modo que a vazão para o dia da simulação $Q_{(t)}$ depende da vazão simulada do dia anterior ($Q_{(t-1)}$).

Eficácia da MLP e contribuição relativa dos vetores de entrada na MLP

A eficácia da MLP foi medida sobre os resíduos das vazões simuladas em relação às vazões observadas. As estatísticas utilizadas para avaliar a MLP em *open-loop* e *closed-loop* foram o índice de Nash e Sutcliffe (1970) (NS), o Coeficiente de Determinação (R^2) e a Raiz do Erro Médio Quadrático (RMS), sendo realizado o *ranking* com essas estatísticas para determinar os melhores tratamentos.

As estatísticas para a fase de treinamento foram calculadas correspondendo aos dados dos primeiros 2 anos (1996 e 1997) e para teste o último ano (1999), sendo que para o teste as vazões foram simuladas pela arquitetura em *open-loop* e *closed-loop*.

A análise de resíduos consistiu na verificação da diferença da vazão observada pela simulada (resíduo) em níveis mínimos, médios e máximos, com objetivo de analisar em quais classes de vazão ocorreu subestimativa ou superestimativa pelo modelo. A delimitação dos níveis baseou-se em Xu (2001) considerando a média dos valores observados, pelas seguintes regras: (1) mínima: vazões menores que $0,75 * (\text{média observada})$; (2) média: vazões entre $0,75 * (\text{média observada})$ e $1,25 * (\text{média observada})$; e (3) máxima: vazões maiores que $1,25 * (\text{média observada})$.

A Contribuição Relativa (CR) dos vetores de entrada (GARSON, 1991) e a análise de resíduos (XU, 2001; BIEGER; HORMANN; FOHRER, 2012) foram realizadas sobre os tratamentos de entrada que apresentaram os melhores desempenhos no teste em *closed-loop* por se tratar da metodologia em destaque neste estudo.

O cálculo do índice percentual da CR consiste em ponderar as variações dos valores absolutos dos pesos entre os vetores da camada de entrada e da camada de saída (Equação 6), a fim de determinar a relevância relativa de cada vetor da camada de entrada.

$$CR_{ik} = \frac{\sum_{j=1}^L \left(\frac{w_{ij}}{N} - v_{jk} \right)}{\sum_{i=1}^N \left(\sum_{j=1}^L \left(\frac{w_{ij}}{N} - v_{jk} \right) \right)} \quad (6)$$

em que: CR_{ik} : representa a influência do vetor de entrada i sobre o vetor de resposta k ; $\sum_{r=1}^N w_{rj}$: é a soma dos pesos de conexão entre a camada de entrada i e o neurônio j ; N : corresponde ao total de vetores de entrada; L : corresponde ao total de neurônios da camada oculta; v_{jk} : peso da conexão entre neurônio da camada oculta j e o vetor da saída k .

RESULTADOS E DISCUSSÃO

Análise exploratória

A matriz de correlação linear (Tabela 1) entre as séries das precipitações registradas nas estações pluviométricas apresentou correlações positivas significativas (0,54 a 0,8), o que serve de indicador do nível de variabilidade espacial das chuvas na área de estudo.

As fracas correlações da $ET0$ com as séries de precipitações registradas podem indicar que a $ET0$ não deva ser controlada pela chuva nessa área de estudo ou que não haja forte ciclo de retroalimentação entre essas duas variáveis. Além disso, a correlação

linear entre a $ET0$ e vazão é fraca, acredita-se que por este fato, a retirada de água do solo pela $ET0$ não deva causar muita influência na vazão. As baixas correlações entre as precipitações e a vazão indicam que as vazões mensuradas no exutório da bacia dependem de outros fatores, tais como, as chuvas antecedentes que modificam o teor de umidade do solo, escoamento subterrâneo que predomina em longos períodos sem precipitação, entre outros processos do ciclo hidrológico.

A correlação linear entre a vazão no tempo t (Q_t) com a vazão antecedente, com defasagem de 1 dia ($Q_{t-1} = 0,96$) e de 2 dias ($Q_{t-2} = 0,88$) corrobora com a correlação cruzada (Figura 4F), em que dias próximos a correlação é maior e ao se distanciar do tempo t , a correlação da vazão diminui. Corrobora com a correlação cruzada no sentido de que quanto mais distante na escala de tempo a correlação diminui, como esperado.

As estações que apresentaram maiores correlações cruzadas com a vazão foram as estações $P1$ e $P2$, situadas no interior da bacia, próximas ao canal principal (Figura 4A-D). Além disso, é possível observar um sutil decréscimo da máxima correlação da estação $P1$ até $P4$ (0,37 a 0,32), o que pode estar relacionado com a distância de cada estação ao centro da bacia ou, ainda, ao percentual da área de contribuição das estações, que também decresce no mesmo sentido.

A maior correlação cruzada da precipitação com a vazão ocorre em -2 dias, indicando que pode estar relacionado com o tempo de concentração dessa bacia, visto que o tempo de

Tabela 1. Matriz de correlação linear de Pearson (R) entre as variáveis utilizadas para compor os vetores de entrada da MLP.

Variáveis	$P1$	$P2$	$P3$	$P4$	$ET0$	Q_t	$Q_{(t-1)}$	$Q_{(t-2)}$
$P1$	1,00	0,79	0,80	0,60	0,07	0,13	0,03	0,03
$P2$	0,79	1,00	0,73	0,54	0,13	0,12	0,02	0,01
$P3$	0,80	0,73	1,00	0,61	0,08	0,13	0,03	0,02
$P4$	0,60	0,54	0,61	1,00	0,03	0,13	0,05	0,04
$ET0$	0,07	0,13	0,08	0,03	1,00	0,01	0,03	0,03
Q_t	0,13	0,12	0,13	0,13	0,01	1,00	0,96	0,88
$Q_{(t-1)}$	0,03	0,02	0,03	0,05	0,03	0,96	1,00	0,96
$Q_{(t-2)}$	0,03	0,01	0,02	0,04	0,03	0,88	0,96	1,00

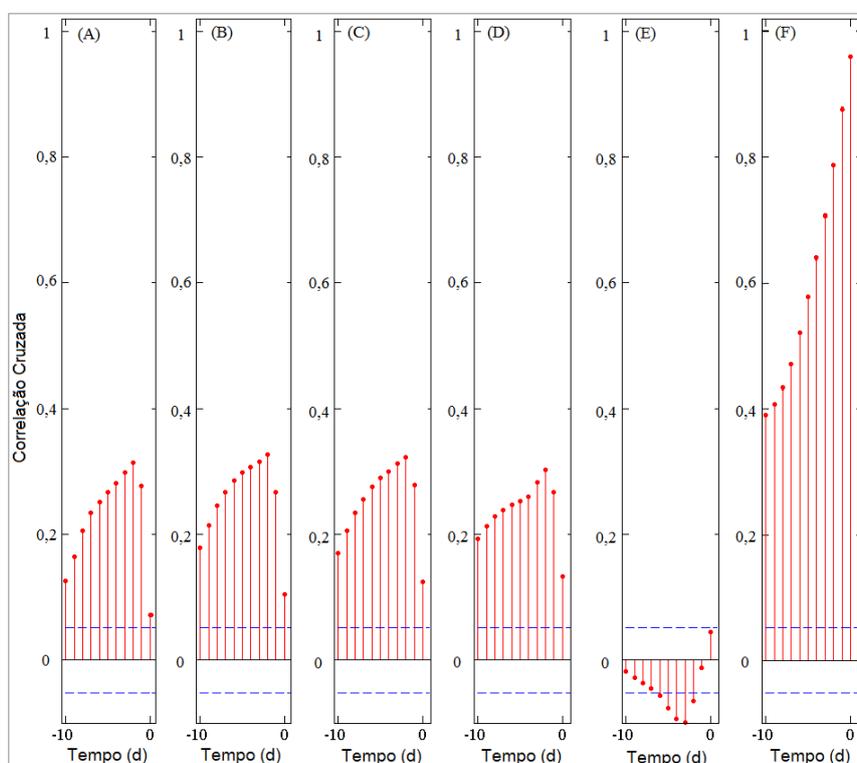


Figura 4. Correlação cruzada (CC) com intervalo de confiança de 95% (linha tracejada) entre a vazão e; (A) precipitação da estação de Urubici ($P1$); (B) precipitação da estação de Vila Canoas ($P2$); (C) precipitação da estação de Lomba Alta ($P3$); (D) precipitação da estação de Anitápolis ($P4$); (E) evapotranspiração ($ET0$); (F) vazão com defasagem de 1 dia (Q_{t-1}).

concentração representa o tempo necessário para que toda a área da bacia contribua para o escoamento superficial no exutório.

Observou-se correlações não significativas, e muito fracas, entre as variáveis média diária de ET0 e vazão média diária (-0,13 até 0,5). Contudo, Sá (2014) e Antunes (2015) demonstraram que a ET0 desta estação foi útil na calibração dos modelos TOPMODEL e SWAT para esta mesma bacia.

Eficácia da MLP em open-loop e closed-loop

A eficácia da MLP foi testada sobre 12 tratamentos, variando-se a composição do conjunto de vetores de entrada (Tabela 2). Os vetores foram definidos a partir dos indicadores de correlações obtidos na análise exploratória, onde esses foram constituídos pela precipitação pontual das estações (P1, P2, P3, P4), precipitação com atraso de dois dias (P1_(t-2), P2_(t-2), P3_(t-2), P4_(t-2) por apresentar maior correlação com a vazão), precipitação de cada estação multiplicada pelo peso correspondente a área de contribuição (PP1, PP2, PP3, PP4), precipitação ponderada de toda a bacia (PP) e vazão com atraso de 1 dia (Q_(t-1), maior correlação com a vazão), sendo este último vetor o condicionante para o funcionamento da arquitetura closed-loop.

A vazão com tempo de resposta de -1 dia (Q_(t-1)) esteve presente em todos os tratamentos, com exceção do tratamento 12. Pode-se verificar que sem a adição da Q_(t-1) nesse tratamento o NS foi de 0,582, R² de 0,009 e RMSE de 35,9%, demonstrando o baixo desempenho da estimativa da vazão diária sem a adição da Q_(t-1), corroborando com os resultados encontrados por Santos e Pereira Filho (2003), Riad et al. (2004) e Oliveira et al. (2013).

A eficácia da MLP em open-loop foi em torno de NS = 0,90, com exceção dos tratamentos 11 e 12, enquanto que em closed-loop, observou-se uma diversidade de comportamentos. Os melhores índices de eficácia em closed-loop foram observados nos tratamentos

10 (NS = 0,919, RMS = 14,29 m³/s) e tratamento 8 (NS = 0,922, RMS = 15,65 m³/s).

Já era esperado um bom desempenho da MLP em open-loop, visto que a previsão realizada era de apenas 1 dia adiante (+1d), sendo que os tratamentos sempre utilizaram como entrada a vazão observada do dia anterior. Por outro lado, esperavam-se resultados inferiores em closed-loop, pois o modelo simulou toda a série de vazões utilizando como entrada, a cada passo de tempo, a vazão simulada no passo de tempo anterior (-1d).

Houve melhoria sensível de eficácia a partir de 3 vetores de precipitação, conforme se constata nos tratamentos de 3 a 5 (NS de 0,810 a 0,908). Considerando-se a correlação similar entre os vetores de precipitação e o vetor de vazão (R de 0,12 a 0,13) e a forte correlação entre si (R de 0,54 a 0,8), acredita-se que o ganho na eficácia tenha ocorrido devido à distribuição espacial relativamente equitativa das estações pluviométricas na região da bacia, estando localizadas cerca de 20 a 40 km uma da outra. Este fato provavelmente garantiu a cada estação uma porção de variabilidade intrínseca a sua localização espacial, que colaborou consideravelmente para a modelagem. Um exemplo disso é o vetor P4 que, apesar de apresentar a correlação mais fraca com as demais estações (R = 0,54), causou uma melhoria no desempenho com sua adição nos tratamentos 2 para 4 e 3 para 5.

A tentativa de ponderar a precipitação de cada estação pelas áreas de influências das estações pluviométricas (tratamento 7) não apresentou resultado satisfatório (NS = 0,273, R² = 0,250 e RMSE = 50,1%), o que pode ser explicado pela correlação entre os vetores de precipitação e a distância espacial das estações pluviométricas. Desempenho um pouco superior foi encontrado ao utilizar a precipitação ponderada total no tempo t e com atraso de 1 e 2 dias (tratamento 11) que resultou em NS = 0,394, R² = 0,747 e RMSE = 35,3, porém este ainda inferior a outros tratamentos. Portanto, não há justificativa para utilização das precipitações

Tabela 2. Índices de eficácia da simulação da vazão média diária pela MLP em open e closed-loop para o sub-conjunto de teste sobre os tratamentos.

N	Tratamento	Arquitetura MLP	Open-loop			Closed-loop		
			NS	R ²	RMS m ³ /s (%)	NS	R ²	RMS m ³ /s (%)
1	P1, Q _(t-1)	2/18/1	0,912	0,871*	12,25 (7,2)	0,515	0,515*	71,37 (42,2)
2	P1, P2, Q _(t-1)	3/14/1	0,913	0,872*	11,72 (6,9)	0,223	0,350*	34,73 (20,5)
3	P1, P2, P3, Q _(t-1)	4/12/1	0,905	0,859*	12,16 (7,2)	0,908	0,824*	23,05 (13,6)
4	P1, P2, P4, Q _(t-1)	4/11/1	0,910	0,867*	12,11 (7,1)	0,862	0,840*	17,91 (10,5)
5	P1, P2, P3, P4, Q _(t-1)	5/10/1	0,907	0,866*	11,91 (7,1)	0,810	0,848*	16,68 (9,9)
6	P1, P2, P3, P4, ET0, Q _(t-1)	6/9/1	0,907	0,853*	11,8 (6,9)	-1,10	0,785*	166,99 (98)
7	PP1, PP2, PP3, PP4, Q _(t-1)	5/10/1	0,907	0,864*	11,91 (7,1)	0,273	0,250*	84,68 (50,1)
8	P1 _(t-2) , P2 _(t-2) , P3 _(t-2) , P4 _(t-2) , Q _(t-1)	5/10/1	0,915	0,851*	12,41 (7,5)	0,922	0,833*	15,65 (9,2)
9	P1 _{(t-2)AC} , P2 _{(t-2)AC} , P3 _{(t-2)AC} , P4 _{(t-2)AC} , Q _(t-1)	5/10/1	0,894	0,823*	13,56 (8,1)	0,340	0,351*	28,24 (16,7)
10	P1, P2, P3, P4, P1 _(t-2) , P2 _(t-2) , P3 _(t-2) , P4 _(t-2) , Q _(t-1)	9/6/1	0,921	0,878*	12,0 (7,1)	0,919	0,868*	14,29 (8,4)
11	PP _t , PP _{t-1} , PP _{t-2} , Q _(t-1)	3/7/1	0,235	0,555	109,17 (64,6)	0,394	0,747*	99,25 (35,3)
12	P1, P2, P3, P4, ET0	5/11/1	0,582	0,009	60,73 (35,9)	-	-	-

* p-valor < 0,05. Sendo: Arquitetura: número de neurônios na camada de entrada/ camada oculta / camada de saída; ET0: evapotranspiração de referência; Q_(t-1): vazão observada (treinamento) ou simulada (teste) para o dia anterior t-1; PP: precipitação ponderada, precipitação da respectiva estação multiplicada pelo seu peso; PP_t: somatório da precipitação ponderada total da bacia; P_(t-2): precipitação registrada a 2 dias anteriores; P_{(t-2)AC}: precipitação acumulada nos dois dias anteriores. As precipitações P1, P2, P3, P4 são as precipitações registradas nas respectivas estações de Urubici, Vila Canoas, Lomba Alta e Anitápolis.

ponderadas, pois conforme observado nos tratamentos 1 a 5, a MLP foi eficaz com a precipitação pontual.

A utilização das precipitações com defasagem de 2 dias no tratamento 8 (NS = 0,922, $R^2 = 0,833$ e RMSE = 9,2%) apresentou resultados pouco superiores ao do tratamento 5 (NS = 0,810, $R^2 = 0,848$ e RMSE = 9,9%), que não leva em consideração este tempo. Seria esperado que, devido à ocorrência da correlação máxima entre as variáveis de precipitação e vazão em -2 dias, houvesse uma diferença maior entre esses tratamentos. O tratamento 10 foi o que apresentou o melhor desempenho no *ranking* das estatísticas de avaliação (NS = 0,919, $R^2 = 0,868$ e RMSE = 8,4%) ao associar os tratamentos 5 e 8.

O acúmulo da precipitação a -2 dias, no tratamento 9, apresentou resultados insatisfatórios no período de teste (NS = 0,340, $R^2 = 0,351$ e RMSE = 16,7%), sendo indicado o uso da precipitação com defasagem de 2 dias sem acumular.

A adição da variável *ET0* no tratamento 5, gerando o tratamento 6, acarretou em insuficiência completa a eficácia da MLP (NS = -1,10, $R^2 = 0,785$, RMSE = 98%), o que pode ser explicado pela fraca correlação da *ET0* com a vazão (Figura 4E).

Na Figura 5 é representado o comportamento da vazão observada e simulada pela MLP nos tratamentos 5, 8 e 10 para o período de teste (1999). Observa-se que o tratamento 5 obteve bom ajuste das vazões de base e superestimou as vazões máximas, enquanto que o tratamento 8 obteve melhor ajuste nas vazões máximas e superestimou as vazões mínimas. Conforme esperado, o tratamento 10 obteve um resultado intermediário entre 5 e 8, apresentando valores razoáveis tanto para vazões mínimas, quanto para vazões máximas.

Resultado similar foi obtido por Sá (2014) para esta mesma bacia e série histórica de dados. O autor também observou que

o modelo TOPMODEL superestimou as vazões básicas, apesar de ter alcançado NS de 0,84.

Análise de resíduos

As classes de vazões médias e máximas apresentaram distribuição próxima da normal e com mediana negativa (indicador de superestimativa), com exceção do tratamento 8 que nas vazões máximas, apresentou mediana positiva (indicador de subestimativa). Já a classe de vazões mínimas apresentou distribuição diferente da normal, com valores negativos e próximos de zero. Em geral, o tratamento 10 apresentou valores negativos (indicando superestimativas), porém mais próximos de zero do que o restante dos tratamentos (Figura 6).

Assim como verificado por Sá (2014), a superestimativa das vazões de recessão foi inevitável. Visto que o modelo TOPMODEL utilizado pelo autor, baseado em processos, e o modelo baseado em dados do presente trabalho, apontaram para a superestimativa da vazão de recessão, é provável que exista algum processo não previsto pelos modelos, ocorrendo nessa bacia que promova a retirada de água, como a existência de uma ou mais estações de captação de água para uso urbano ou para irrigação na agricultura.

Contribuição relativa dos vetores de entrada

Analisando a CR em *closed-loop*, observa-se que o vetor da estação de Urubici (*P1*) correspondeu por cerca de 50% da eficácia da MLP (Figura 7). As demais estações parecem dividir entre si as contribuições relativas, cujo conjunto é necessário para que a MLP atinja alto índice de eficácia. Destas, as estações *P3* e *P4* parecem

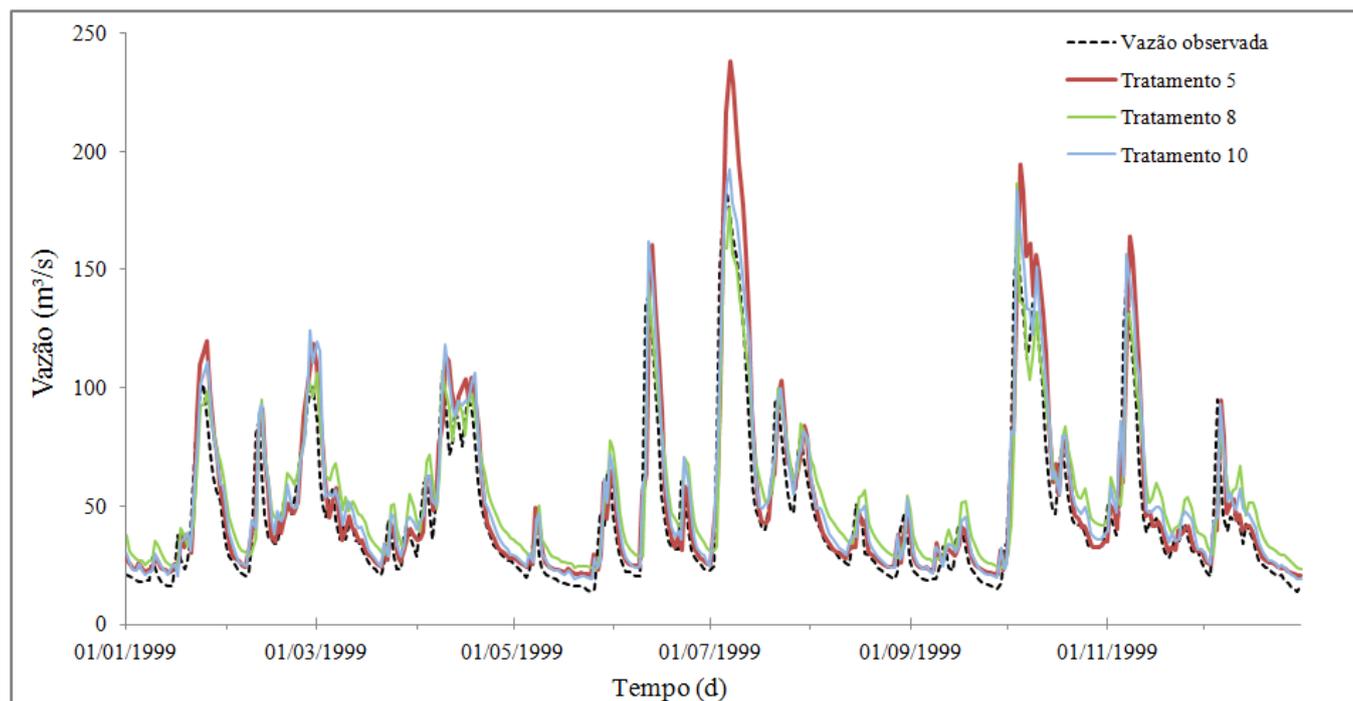


Figura 5. Hidrograma da vazão observada e simulada pela MLP em *closed-loop* nos tratamentos 5, 8 e 10, para o período de teste (1999).

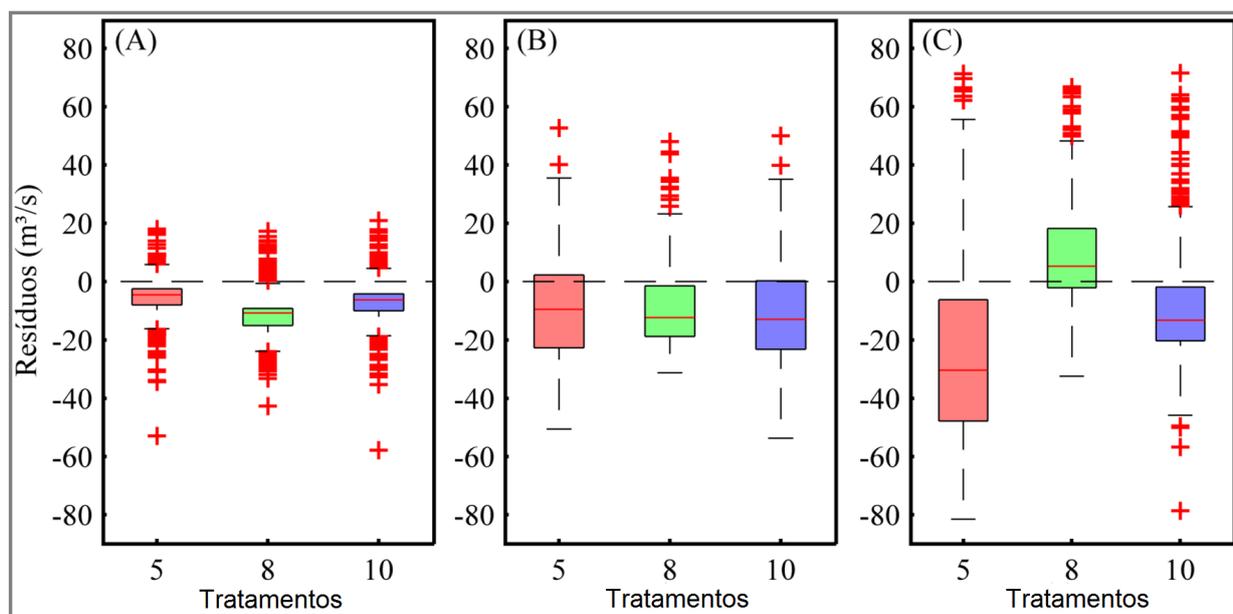


Figura 6. Boxplots dos resíduos das (A) vazões mínimas (n = 770); (B) vazões médias (n = 276) e (C) vazões máximas (n = 414).

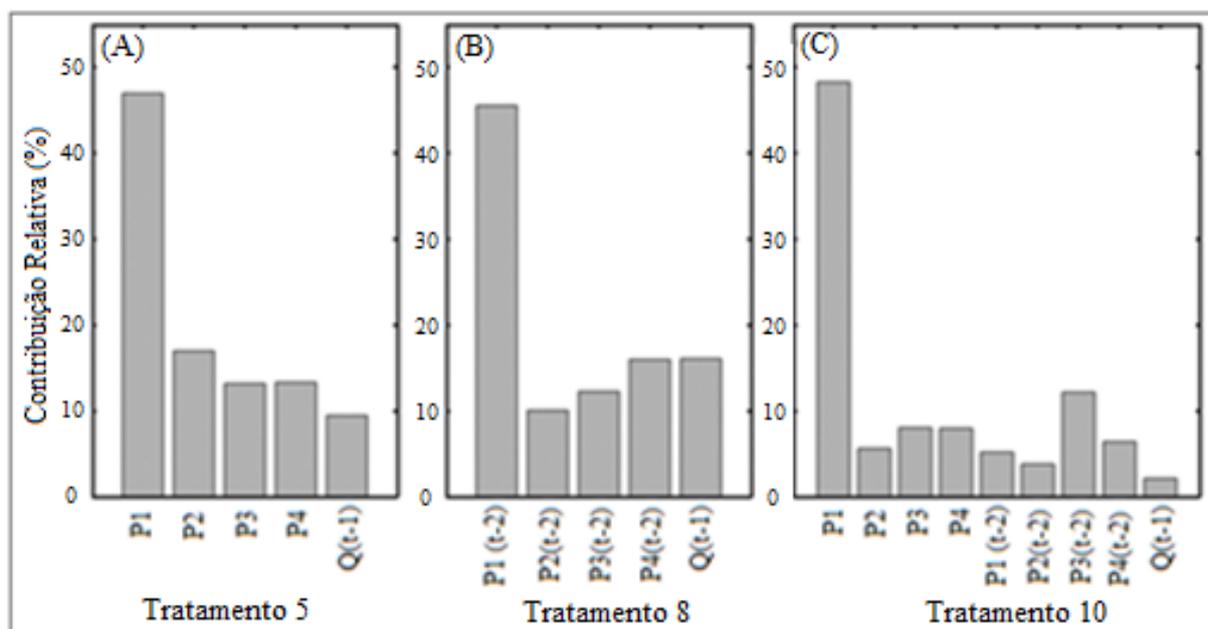


Figura 7. Análise da contribuição relativa (CR) de cada variável para os tratamentos 5 (A); 8 (B) e 10 (C).

oferecer uma melhoria equitativa. A estação *P2* demonstrou ser relevante em relação às demais, porém quando usado o vetor com dados do dia anterior ($-1d$), sua importância relativa diminuiu.

A vazão com tempo de resposta de -1 dia ($Q_{(t-1)}$) apresentou CR inferior ou similar às estações pluviométricas *P2* à *P4*. Com isso, entende-se que essa variável não foi decisiva na estimativa da vazão em *closed-loop*. É possível que a vazão com tempo de resposta de -1 dia ($Q_{(t-1)}$) tenha servido de base para ajuste da magnitude dos valores, enquanto que as variáveis de precipitação foram responsáveis pelas variações diárias. Provavelmente, este seja o motivo da modelagem em *closed-loop* ter obtido êxito, por não depender fortemente das vazões estimadas anteriormente.

Estas constatações evidenciam a importância das estações situadas na cabeceira da bacia do Alto Canoas. As chuvas nesta região tendem a ter influência orográfica, devido à Serra Geral e relevo com vertentes íngremes. A estação *P2* é a única que está fora deste contexto e próxima ao exutório.

CONCLUSÃO

Este trabalho avaliou a estimativa da vazão diária na estação fluviométrica Rio Bonito, na bacia do Alto Canoas-SC, através de modelagem com redes neurais artificiais MLP. Foram testadas combinações com as variáveis de precipitação, vazão e *ET0* para a

estimativa da vazão, bem como transformações e deslocamentos temporais das variáveis.

A previsão da série histórica de vazão em *closed-loop* a partir de uma MLP treinada em *open-loop* foi eficaz. A melhor estimativa da vazão foi obtida pela utilização dos vetores $P1, P2, P3, P4, P1_{(t-2)}, P2_{(t-2)}, P3_{(t-2)}, P4_{(t-2)}$ e $Q_{(t-1)}$ (NS = 0,919, $R^2 = 0,868$ e RMS = 8,4%). A combinação das variáveis de precipitação sem e com tempo de resposta de -2 dias apresentou melhor desempenho que a utilização destas separadamente. Este tratamento apresentou não só melhores índices estatísticos, mas também melhor estimativa das vazões mínimas e máximas.

A utilização dos dados de todas as estações pluviométricas mostrou-se mais interessante do que sua utilização separada devido a sua distribuição espacial esparsa que confere uma variabilidade diferenciada a cada uma das estações e é captada pela MLP. A precipitação ponderada não se mostrou eficaz, bem como a utilização da *ETO*.

O vetor que mais contribuiu para a eficácia da MLP foi a precipitação observada na estação de Urubici (*P1*). O vetor da vazão estimada (Q_{t-1}) teve pouca contribuição relativa, mas manteve-se indispensável na MLP com configuração em *closed-loop*.

AGRADECIMENTOS

À Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (FAPESC) pela concessão de bolsa de estudos a primeira autora.

REFERÊNCIAS

- ABRAHART, R.; KNEALE, P. E.; SEE, L. M. *Neural networks for hydrological modeling*. London: Taylor & Francis, 2004. 316 p.
- ALLEN, R. G.; PEREIRA, L. S.; RAES, D.; SMITH, M. *Crop evapotranspiration: guidelines for computing crop water requirements*. Rome: FAO, 1998. 300 p. (FAO Irrigation and Drainage Paper, 56).
- ANTUNES, T. A. *Modelagem hidrológica da bacia hidrográfica do Alto Canoas através do modelo SWAT*. 2015. 130 f. Dissertação (Mestrado em Engenharia Florestal) – Programa de Pós-graduação em Engenharia Florestal, Universidade do Estado de Santa Catarina, Lages, 2015.
- BIEGER, K.; HORMANN, G.; FOHRER, N. Using residual analysis, auto-and cross-correlations to identify key processes for the calibration of the SWAT model in a data scarce region. *Advanced Geoscience*, v. 31, p. 23-30, 2012. <http://dx.doi.org/10.5194/adgeo-31-23-2012>.
- FU, L. *Neural networks in computer intelligence*. New York: McGraw-Hill, 1994.
- GARSON, G. D. Interpreting neural-network connection weights. *Artif. Intell. Expert*, v. 6, n. 4, p. 47-51, 1991.
- HAGAN, M. T.; MENHAJ, M. Training feed-forward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, v. 5, n. 6, p. 989-993, 1994. PMID:18267874. <http://dx.doi.org/10.1109/72.329697>.
- HEATH, G. E. *Training, testing and validating data set in Neural Network*. 2010. Disponível em: http://www.mathworks.com/matlabcentral/newsreader/view_thread/295781#917734. Acesso em: 04 jul. 2013.
- MATLAB. *Version 8.3.0.532*. Natick: The MathWorks Inc., 2014.
- NASH, J. E.; SUTCLIFFE, J. V. River flow forecasting through conceptual models part 1: a discussion of principles. *Journal of Hydrology (Amsterdam)*, v. 10, n. 3, p. 282-290, 1970. [http://dx.doi.org/10.1016/0022-1694\(70\)90255-6](http://dx.doi.org/10.1016/0022-1694(70)90255-6).
- OLIVEIRA, G. G.; PEDROLLO, O. C.; CASTRO, N. M. O desempenho das redes neurais artificiais (RNAs) para simulação hidrológica mensal. *RBRH: Revista Brasileira de Recursos Hídricos*, v. 19, n. 2, p. 251-265, 2014. <http://dx.doi.org/10.21168/rbrh.v19n2.p251-265>.
- OLIVEIRA, G. G.; PEDROLLO, O. C.; CASTRO, N. M. R.; BRAVO, J. M. Simulações hidrológicas com diferentes proporções de área controlada na bacia hidrográfica. *RBRH: Revista Brasileira de Recursos Hídricos*, v. 18, n. 3, p. 193-204, 2013. <http://dx.doi.org/10.21168/rbrh.v18n3.p193-204>.
- POOCH, U. W.; WALL, J. A. *Discrete event simulation: a practical approach*. Boca Raton: CRC Press, 1993. 412 p.
- RAFAELI NETO, S. L. *Um modelo conceitual de sistemas de apoio à decisão espacial para gestão de desastres por inundações*. 2000. 208 f. Tese (Doutorado em Engenharia) - Programa de Pós Graduação em Engenharia, Universidade de São Paulo, São Paulo, 2000.
- RIAD, S.; MANIA, J.; BOUCHAOU, L.; NAJJAR, Y. Rainfall-runoff model using an artificial neural network approach. *Mathematical and Computer Modelling*, v. 40, n. 7-8, p. 839-846, 2004. <http://dx.doi.org/10.1016/j.mcm.2004.10.012>.
- SÁ, E. A. S. *Estudo do modelo TOPMODEL na bacia hidrográfica do Alto Canoas-SC*. 2014. 103 f. Dissertação (Mestrado em Engenharia Florestal) - Programa de Pós-graduação em Engenharia Florestal, Universidade do Estado de Santa Catarina, Lages, 2014.
- SANTOS, C. C.; PEREIRA FILHO, A. J. Modelagem hidrológica urbana por meio de redes neurais artificiais: uma aplicação para a bacia do Rio Tamanduateí em São Paulo, SP. *Revista Brasileira de Meteorologia*, v. 18, n. 2, p. 149-159, dez. 2003. Disponível em: http://www.rbmet.org.br/port/revista/revista_artigo.php?id_artigo=31. Acesso em: 09 nov. 2014.
- SOUSA, W. S.; SOUSA, F. A. S. Rede neural artificial aplicada à previsão de vazão da bacia hidrográfica do Rio Piancó. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v. 14, n. 2, p. 173-180, 2010. <http://dx.doi.org/10.1590/S1415-43662010000200008>.
- XU, C. Y. Statistical analysis of parameters and residuals of a conceptual water balance model: methodology and case study. *Water Resources Management*, v. 15, n. 2, p. 75-92, 2001. <http://dx.doi.org/10.1023/A:1012559608269>.

Contribuição dos autores

Aline Bernarda Debastiani: Contribuiu com pesquisa bibliográfica, planejamento da pesquisa, análise dos dados, processamento e simulações com a MLP, escrita e formatações do artigo.

Dal'Agnol da Silva: Contribuiu com planejamento da pesquisa, adaptação das rotinas de processamento da MLP, escrita e revisão do artigo.

Sílvio Luís Rafaeli Neto: Professor orientador, contribuiu na concepção do trabalho, análise dos processamentos e revisão do artigo.