

O Desempenho das Redes Neurais Artificiais (RNAs) para Simulação Hidrológica Mensal

Guilherme Garcia de Oliveira*, Olavo Correa Pedrollo*, Nilza Maria dos Reis Castro*
g.g.oliveira10@gmail.com

Recebido: 19/07/13 - revisado: 07/11/13 - aceito: 14/01/14

RESUMO

O objetivo deste estudo foi avaliar o desempenho de modelos hidrológicos baseados em RNAs para a simulação hidrológica da vazão média mensal, utilizando apenas dados meteorológicos e pluviométricos. Este estudo foi realizado na bacia do rio Ijuí, na área de contribuição ao posto fluviométrico Santo Ângelo (5.414 km²), no Rio Grande do Sul, Brasil, com vazão média de 138 m³/s no período entre 1941 e 2005. Os procedimentos de pesquisa adotados foram: i) organização e análise de consistência dos dados; ii) interpolação espacial dos dados meteorológicos e pluviométricos; iii) estimativa da evapotranspiração de referência; iv) extração dos valores mensais de precipitação, evapotranspiração potencial e de vazão; v) modelagem hidrológica com RNAs. A avaliação dos modelos foi efetuada através de análises visuais comparativas dos hidrogramas e das curvas de permanência (observado e simulado), de indicadores quantitativos de desempenho como o coeficiente de Nash-Sutcliffe (NS), e de uma análise de sensibilidade para compreender e interpretar o funcionamento da RNA. Além disto, foi realizada uma comparação com o modelo hidrológico SMAP, modelo conceitual adaptado para simulações mensais, proposto por Lopes et al. (1982). O modelo com RNAs mais adequado para a simulação das vazões apresentou apenas três variáveis de entrada, com 16 pesos sinápticos. Neste modelo, foi obtido um NS igual a 0,904, com RMSE de 37,1 m³/s. Através da análise de sensibilidade, foi observado que a RNA escolhida relacionou corretamente as variáveis de entrada com a saída da rede, respeitando os princípios físicos envolvidos no sistema hidrológico. Além disto, verificou-se que o modelo de RNA simplificado apresentou razoável capacidade de extrapolação. Ao comparar o desempenho das simulações realizadas com a RNA e o modelo SMAP, no período entre 2002 e 2009, foi observada uma grande diferença no NS (RNAs: 0,898; SMAP: 0,672).

Palavras-chave: Bacia hidrográfica do rio Ijuí; modelagem hidrológica; vazão média mensal; análise de sensibilidade.

INTRODUÇÃO

As Redes Neurais Artificiais (RNAs) são modelos matemáticos empíricos, com capacidade para reproduzir o funcionamento dos sistemas nervosos biológicos. Esta habilidade é reforçada através do treinamento e da resposta aos estímulos (variáveis de entrada).

A criação das RNAs está ligada à publicação de McCulloch e Pitts (1943), em que o neurônio artificial foi formulado a partir de observações puramente biológicas. A rede possuía apenas capacidade de representar, embora de forma semelhante aos neurônios biológicos, funções lógicas bastante simplificadas. Somente com o algoritmo retropropagativo, desenvolvido por Rumelhart et al. (1986),

é que as RNAs passaram a ser difundidas e receber investimentos em pesquisa de forma contínua e crescente. O algoritmo retropropagativo é uma generalização, aplicada às redes multicamadas, da Regra Delta, criada por Widrow e Hoff (1960). A atualização dos pesos de um neurônio depende apenas dos erros, das entradas e da derivada do próprio neurônio. Os erros da camada interna de neurônios são calculados a partir dos erros da camada de neurônios de saída.

As entradas de uma RNA funcionam como estímulos e a composição básica de um neurônio artificial é uma função de soma ponderada destas entradas e uma função de ativação, normalmente não linear, que modifica o resultado da soma ponderada. Isto resulta em um sinal correspondente à resposta desejada, à medida que os parâmetros (pesos sinápticos) sejam ajustados adequadamente por meio do treinamento da rede.

Recentemente, diversas pesquisas obtiveram excelentes resultados com a aplicação de RNAs na

*Instituto de Pesquisas Hidráulicas (IPH), Universidade Federal do Rio Grande do Sul (UFRGS).

área de recursos hídricos e hidrologia, principalmente no desenvolvimento de modelos para simulação, previsão e classificação (Jain et al., 2004; Bowden et al., 2005; Jain e Kumar, 2007; Leahy et al., 2008; Oliveira et al., 2013a).

Shamseldin (1997) utilizou RNAs em simulações hidrológicas do tipo chuva-vazão em seis bacias, com áreas de drenagem entre 1.207 e 18.000 km². Em cinco destas, obteve coeficiente Nash-Sutcliffe (NS) superior a 0,7. No estudo de Sajikumar e Thandaveswara (1999), as RNAs foram aplicadas para a simulação hidrológica em duas bacias, rio Lee (Reino Unido) e rio Thuthapuzha (Índia), com 1.419 e 1.030 km², respectivamente, resultando coeficientes NS superiores a 0,8. Anmala et al. (2000) aplicaram as RNAs em simulações hidrológicas mensais em três bacias hidrográficas no Kansas, EUA, com áreas de drenagem entre 518 e 640 km², obtendo desempenho superior a dois outros modelos empíricos, com coeficiente R² superior a 0,9. Hsu et al. (1995) utilizaram a técnica das RNAs para a simulação da vazão diária no rio Leaf, Mississippi, Estados Unidos (EUA), com área de 1.949 km². A técnica apresentou desempenho superior à simulação com o modelo conceitual *Sacramento Soil Moisture Accounting* e o modelo autorregressivo ARMAX. Em outro estudo, Oliveira et al. (2013a) apresenta uma comparação entre as RNAs e do modelo conceitual IPH II na simulação de vazões diárias em quatro sub-bacias do rio Ijuí, no Rio Grande do Sul, Brasil, com áreas entre 19 e 9.426 km². As RNAs superaram o modelo conceitual, com NS variando entre 0,72 e 0,94, enquanto que no modelo IPH II o mesmo índice ficou entre 0,3 e 0,88.

A pesquisa de Campolo et al. (2003) aborda a previsão de níveis em tempo real com RNAs para a bacia do rio Arno (4.000 km²). Foram obtidos baixos erros percentuais (7% a 15%), evidenciando que o modelo de previsão com RNAs é adequado para tal aplicação. No estudo de Rajurkar et al. (2004) as RNAs foram aplicadas na modelagem chuva-vazão com dados diários em períodos de cheia, na Índia, em sete bacias hidrográficas, 1.207 a 26.200 km², obtendo resultados satisfatórios, com coeficiente de NS superior a 0,8 na maioria das previsões. Na mesma temática, Jain e Kumar (2007) aplicaram modelos com RNAs e modelos autorregressivos para a previsão hidrológica mensal em uma bacia do rio Colorado, Estados Unidos, com área de 290.000 km². Foram testadas diferentes configurações de modelos, e em todos os casos, os modelos baseados em RNAs obtiveram resultados melhores que os autorregressivos. Por exemplo, em um modelo de quarta ordem, enquanto a RNA apresentou um

coeficiente de correlação linear (r) de 0,89, o modelo autorregressivo obteve r de 0,81.

Entretanto, apesar dos resultados promissores, as RNAs são rotuladas como modelos "caixa preta", pois não explicam os processos físicos inerentes à dinâmica das bacias que originam o escoamento (HSU et al., 1995; JAIN et al., 2004). Ainda assim, deve-se considerar a possibilidade de analisar a sensibilidade do modelo aos dados de entrada, de modo a investigar se os impulsos nas entradas estão provocando o efeito esperado na resposta hidrológica da bacia, utilizando, por exemplo, o método de análise de sensibilidade proposto por Lek et al. (1996). A partir desta técnica, pode-se, portanto, verificar se o modelo hidrológico baseado em RNAs representa de modo adequado os processos físicos hidrológicos. Além disto, recentemente, várias técnicas de análise de RNAs treinadas para a modelagem hidrológica têm sido utilizadas para a extração de informações dos processos físicos relacionados à hidrologia (JAIN e KUMAR, 2009) e para a pesquisa da importância das variáveis explicativas (OLIVEIRA et al., 2011).

Tendo em vista a excelente capacidade de representação de processos hidrológicos com o uso de modelos baseados em RNAs, este trabalho pretende contribuir com a apresentação de um modelo hidrológico que se diferencia primeiramente pela discretização temporal em nível mensal, podendo ser aplicado, por exemplo, em estudos de disponibilidade hídrica sazonal em bacias rurais. Além disto, neste estudo, pretende-se demonstrar como o uso de transformações nos dados de entrada interfere no desempenho da modelagem e na capacidade de generalização da rede para valores que extrapolam os limites observados no passado.

Neste contexto, o objetivo deste estudo é avaliar o desempenho das RNAs para a simulação hidrológica da vazão média mensal, utilizando apenas dados meteorológicos e pluviométricos, na bacia do rio Ijuí, na área de contribuição ao posto fluviométrico Santo Ângelo.

MATERIAIS E MÉTODOS

Área de estudo

A seção relativa ao posto fluviométrico Santo Ângelo, na bacia do rio Ijuí, situa-se no noroeste do Rio Grande do Sul (RS), Brasil. A bacia possui 5.414 km² e está localizada entre as seguintes coordenadas geográficas: latitudes de 27,98°S a 28,74°S e longitudes de 53,21°O e 54,28°O (Fig. 1). Neste

posto fluviométrico, entre 1941 e 2005, a vazão média (\bar{Q}) foi de 138 m³/s, tendo como períodos de estiagem e de cheia os meses de março (\bar{Q} = 72 m³/s) e de outubro (\bar{Q} =211 m³/s), respectivamente.

A escolha da área de estudo se deve, primeiramente, ao fato de estar sendo realizado um monitoramento pelo IPH/UFRGS, desde 1989 em bacias da região, com disponibilidade de um banco de dados hidrológicos superior a 20 anos (CASTRO et al., 1999; CASTRO et al., 2010). Além disso, a região é altamente dependente de atividades agrícolas, podendo sofrer sérios impactos socioeconômicos com as alterações climáticas.

Materiais

Os seguintes materiais foram utilizados no presente estudo: a) séries históricas de precipitações disponibilizadas através do portal HydroWeb, da Agência Nacional de Águas (ANA), no período entre 1961 e 2011, em 77 postos pluviométricos no raio de abrangência de 100 km dos limites da bacia (Fig. 2); b) séries históricas de precipitações disponibilizadas pelo IPH (CASTRO et al., 2000), no período entre 1989 e 2011, em 22 postos pluviométricos (Fig. 2); c) séries históricas de precipitação, temperatura, velocidade do vento, radiação solar, pressão atmosférica e umidade relativa do ar, disponibilizadas através do portal BDMEP (Banco de Dados Meteorológicos para Ensino e Pesquisa) do INMET (Instituto Nacional de Meteorologia), no período entre 1961 e 2011, em cinco postos meteorológicos (Fig. 2); d) série histórica de vazões do posto Santo Ângelo (Cód.: 75230000), localizado nas coordenadas 28,36°S e 54,27°O, disponibilizada no portal HydroWeb, no período entre 1961 e 2011.

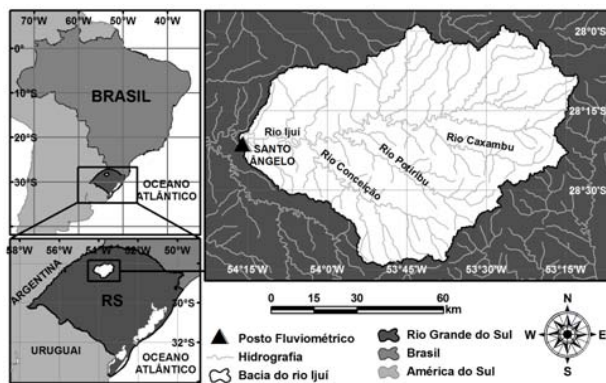


Figura 1 - Localização da bacia do rio Ijuí, seção do posto Santo Ângelo, RS, Brasil

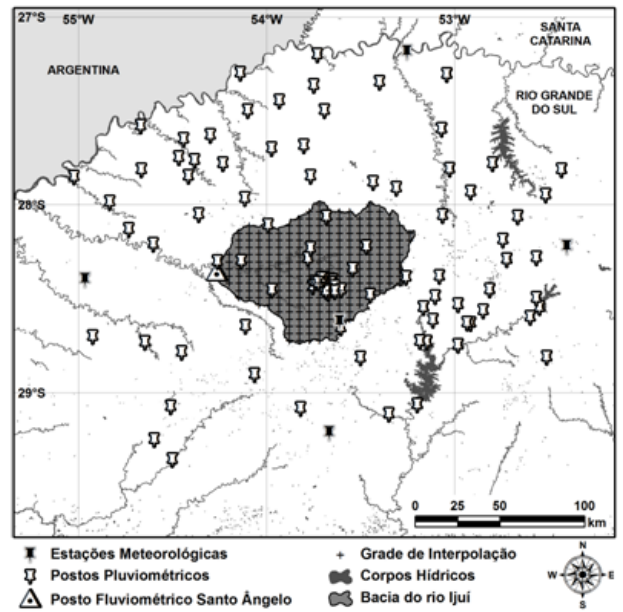


Figura 2 - Localização dos postos pluviométricos e estações meteorológicas utilizadas.

Métodos

A metodologia adotada neste estudo compreendeu o uso de um modelo hidrológico baseado em RNAs, composto por variáveis meteorológicas e pluviométricas transformadas. O treinamento da rede foi realizado através do algoritmo retropropagativo com validação cruzada. A simplificação da RNA, com redução de variáveis de entrada e de neurônios na camada interna, foi realizada através de um algoritmo que verifica o desempenho do modelo após a imposição de pequenas perturbações nos dados de entrada da RNA. A avaliação dos modelos testados foi efetuada através de análises visuais comparativas dos hidrogramas e das curvas de permanência, de indicadores quantitativos de desempenho e de uma análise de sensibilidade para compreender e interpretar o funcionamento da RNA. Além disto, foi realizada uma comparação com o modelo hidrológico SMAP, modelo conceitual adaptado para simulações mensais, proposto por Lopes et al. (1982).

A primeira etapa do trabalho compreendeu a compatibilização das informações das três fontes de séries históricas (HydroWeb, INMET, IPH), de modo que todos os dados pudessem ser lidos e analisados por um programa padrão. Todas as séries brutas foram consistidas para eliminação ou correção de valores incertos ou duvidosos.

Na segunda etapa, os dados diários de precipitação, temperatura do ar, umidade relativa do ar, pressão atmosférica, radiação solar e velocidade do vento foram interpolados pelo método do vizinho natural para toda a área da bacia, em uma grade regular com espaçamento de 5 x 5 km (Fig. 2). Este método de interpolação foi o que obteve os melhores resultados no estudo apresentado por Silva (2011), com séries semelhantes às que foram utilizadas no presente trabalho, também na bacia do rio Ijuí. No estudo citado foram testados também os seguintes métodos: vizinho mais próximo, triangulação linear e o inverso ponderado das distâncias.

O método do vizinho natural (SIBSON, 1981) baseia-se no conceito de área de influência dos pontos amostrais, determinada pelos polígonos de Voronoi. Estes polígonos são obtidos a partir da triangulação de Delaunay. Para cada ponto da grade de interpolação, é calculado o peso de cada ponto amostral em virtude da área de influência. Por fim, o valor diário de cada variável na bacia foi obtido pela média dos valores interpolados em todos os nós da grade regular.

Na terceira etapa, a estimativa da evapotranspiração de referência foi realizada pelo método de Penman-Monteith (PENMAN, 1948; MONTEITH, 1965), que tem sido considerado por diversos autores como o método mais confiável (SEDIYAMA, 1996; ALLEN et al., 1998), sendo adotado como o método padrão pela Organização das Nações Unidas para a Agricultura e Alimentação (FAO). Este método está parametrizado para uma área totalmente coberta com grama de 12 cm de altura, considerando a resistência aerodinâmica da superfície de 70 s.m⁻¹ e albedo de 0,23, em solo sem déficit hídrico.

Na quarta etapa, com base nos valores médios diários de vazão, de precipitação e de evapotranspiração, foram extraídos os valores mensais no período entre 1961 e 2011. Deste modo, foi obtida uma série de dados mensais para cada uma destas variáveis, consistindo na base de dados para a simulação hidrológica com as RNAs.

Na quinta etapa foi realizada a modelagem hidrológica baseada em RNAs. A programação do modelo de RNA foi realizada com o editor de scripts do aplicativo MATLAB R2010a, utilizando, no treinamento, o algoritmo retropropagativo (RUMELHART et al., 1986) com validação cruzada, com uma camada de nós de entrada (variáveis brutas ou transformadas da precipitação e da evapotranspiração), uma camada de neurônios intermediários e outra camada de saída com um único neurônio (vazão simulada).

Conforme o teorema apresentado por Hecht-Nielsen (1987), dada uma relação contínua $r: [0,1]^n \rightarrow R^m$, existe uma rede neural progressiva, com apenas uma camada interna e com $2n+1$ neurônios artificiais nesta camada, capaz de aproximá-la exatamente. Ainda sobre a arquitetura das RNAs, Hornik et al. (1989) enunciou que dada uma relação mensurável $r: R^n \rightarrow R^m$, existe uma RN progressiva capaz de aproximá-la, com qualquer precisão desejada, dispondo apenas de uma camada interna e de um número compatível de neurônios artificiais nesta camada, desde que o treinamento seja adequado.

No presente trabalho, após diversos testes, o número de neurônios na camada intermediária da RNA foi definido como sendo igual ao número de variáveis de entrada. Esta característica foi adotada também com base nos resultados obtidos pela pesquisa de Oliveira et al. (2013b), na qual foram realizadas 83 simulações em diversas bacias do rio Jacuí, RS, Brasil. No estudo, foram testadas quantidades de neurônios na camada interna (entre 1 e $2n+1$, onde n é o número de variáveis de entrada), sendo comparadas através do critério de parcimônia Akaike's Information Criterion (AIC), proposto por Akaike (1974). Cerca de 90% das RNAs selecionadas por este critério apresentaram um número de neurônios na camada intermediária igual ou inferior ao número de variáveis de entrada.

Tendo em vista a utilização do algoritmo retropropagativo para o treinamento da rede, a função de ativação utilizada foi a sigmoide unipolar, sendo esta, não linear, contínua e derivável em todo o seu domínio. A função sigmoide unipolar apresenta saídas no intervalo entre 0 e 1. A derivada é calculada em função da saída, conforme ilustram as Equações 1 e 2. O funcionamento da RNA, considerando apenas uma camada intermediária, é expresso pela Equação 3.

$$a = f(n) = \frac{1}{1 + e^{-n}} \quad (1)$$

$$f'(n) = a(1 - a) \quad (2)$$

$$y = f_s \left(\sum_n w_s f_h \left(\sum_e w_h x + b_h \right) + b_s \right) + e_s \quad (3)$$

Onde: a é a saída da função de ativação; n é a entrada líquida; x e y são as matrizes com as entradas (e) e as saídas (s), respectivamente; w_h , b_h , w_s , b_s são os pesos sinápticos (w) e tendências (b) das

camadas interna (h) e de saída (s); f_h e f_s são as funções de ativação da camada interna e de saída, respectivamente; e_s é o erro esperado na camada de saída.

Quanto ao treinamento da rede, o algoritmo retropropagativo (RUMELHART et al., 1986) é um método de procura dos pesos sinápticos para a minimização dos erros (Equação 4), com o uso da Regra Delta (WIDROW e HOFF, 1960). O uso desta regra para o treinamento da camada intermediária do modelo requer o conhecimento dos erros nesta camada, calculados pela Equação 5, em que os erros da camada anterior (h) dependem somente dos erros e das propriedades da camada posterior (s).

$$w_{k+1} = w_k + (\tau e_k \delta_k P_k) \quad (4)$$

$$e_h = \Sigma(w_s e_s \delta_s) \quad (5)$$

Onde: w_k são os valores dos pesos sinápticos no ciclo k ; τ é a taxa de aprendizagem; e_k são os erros das saídas da camada; δ_k é a derivada das funções de ativação; P_k são as entradas na própria camada, no ciclo k ; e_h é o erro na camada interna; w_s são os pesos da camada posterior; δ_s são as derivadas da função de ativação na camada posterior.

Na inicialização de cada iteração do processo, os pesos sinápticos são gerados aleatoriamente e a taxa de aprendizagem é redefinida para 0,01. A aceleração da convergência se dá pela variação da taxa de aprendizado durante os ciclos de treinamento, de acordo com as soluções apresentadas em Vogl et al. (1988). A técnica consiste na aceitação ou não dos novos pesos sinápticos, de acordo com o desempenho (normalmente, representado pelo erro quadrático), em um ciclo, em função dos erros no ciclo anterior, com aumento da taxa de aprendizado no primeiro caso (a taxa é multiplicada por 1,1), e redução desta no segundo (a taxa é multiplicada por 0,5).

Como o método de treinamento é dependente das condições iniciais, o treinamento pode estacionar em mínimos locais. Tal fenômeno é muito frequente em redes complexas.

Para contornar a incerteza associada à aleatoriedade dos valores iniciais dos parâmetros da RNA para o treinamento, o procedimento foi repetido 30 vezes, de forma iterativa, adotando-se o modelo de melhor desempenho. Este número de itera-

ções foi definido com base no estudo de Dornelles et al. (2013), no qual se pode constatar que, de modo geral, após 30 iterações, o desempenho dos modelos hidrológicos com RNAs manteve-se estável, com melhoramentos pouco significativos.

Em cada iteração, o algoritmo é composto de aplicações sucessivas, com cálculo dos erros nas camadas e da atualização dos pesos sinápticos. Cada aplicação é um ciclo de treinamento, sendo necessários, muitas vezes, milhões de ciclos, dependendo da complexidade do problema.

Neste estudo, foram definidos dois critérios de parada:

- I. por precisão desejada, definida como um erro máximo de 0,001 nas amostras;
- II. por falta de convergência: quando o primeiro critério de parada não é atingido, definiu-se o número de 8.000 ciclos na primeira iteração. Nas iterações seguintes, utilizou-se o número de ciclos necessários para a estabilização dos erros no procedimento anterior, acrescido de 20% (margem de segurança).

Para evitar o superajustamento, adotou-se um método de interrupção antecipada: a técnica de validação cruzada. O superajustamento ocorre quando o treinamento reproduz até mesmo o comportamento individual, incluindo os erros e a aleatoriedade das amostras de treinamento. Tal fenômeno prejudica a capacidade de generalização do modelo.

A técnica de validação cruzada utiliza, para o treinamento, além da série na qual será aplicado o algoritmo retropropagativo, uma série adicional (série de validação) para determinar o ponto ideal de parada durante o treinamento, para que este não perca a sua capacidade de generalização, evitando, desta forma, o superajustamento. Neste processo, os erros do treinamento seguem diminuindo, enquanto os erros da série de validação, a partir de um determinado ciclo, tornam a crescer. Este é o instante em que o limiar é atingido, a partir do qual a capacidade de generalização fica comprometida, havendo, portanto, a interrupção do treinamento (HECHT-NIELSEN, 1990).

O conjunto inicial de variáveis de entrada para a simulação da vazão média mensal no tempo t , $Q(t)$, foi composto por: a) precipitação nos tempos t e $t-1$, $P(t)$ e $P(t-1)$, respectivamente; b) evapotranspiração nos tempos t e $t-1$, $E(t)$ e $E(t-1)$, respectiva-

mente; c) precipitação e evapotranspiração média nos meses t e $t-1$, \bar{P} e \bar{E} , respectivamente; c) balanço de água (saldo), isto é, a diferença entre a precipitação e a evapotranspiração, considerando o apenas o tempo t , $S(t)$; d) saldo médio de água nos tempos t e $t-1$, \bar{S} ; e) valores transformados da precipitação e do saldo de água, respectivamente, fP e fS , pela aplicação de filtros de decaimento exponencial (Equações 6 e 7).

$$fP_t = (1 - \alpha) \cdot fP_{t-1} + \alpha \cdot P_t \quad (6)$$

$$fS_t = (1 - \beta) \cdot fS_{t-1} + \beta \cdot S_t \quad (7)$$

Onde: fP e fS são valores transformados pela aplicação dos filtros de decaimento exponencial às precipitações e ao saldo de água, respectivamente; P é a precipitação; S é o saldo de água - diferença entre precipitação e evapotranspiração; α e β são coeficientes que variam entre 0 e 1, que dependem de calibração.

A calibração destes coeficientes foi feita com base na correlação linear (r) obtida entre as variáveis filtradas e a vazão observada, sendo selecionados os valores que resultaram no maior coeficiente r para o conjunto de dados. Assim, os valores obtidos para α e β foram, respectivamente, 0,52 e 0,41.

As duas entradas resultantes da aplicação do filtro com decaimento exponencial (fP e fS) apresentam um comportamento muito similar às vazões observadas na bacia, com elevado coeficiente r (superior a 0,87). Tal constatação já havia sido apresentada em um estudo anterior, na mesma área de estudo (OLIVEIRA et al., 2011).

A Figura 3 ilustra o conjunto total de dados utilizados. Já a Tabela 1 apresenta a divisão das amostras em três séries de dados (treinamento, validação e verificação). Como os dados de entrada foram normalizados no processo de escalonamento que antecede as simulações, a divisão das amostras nas três séries foi realizada de modo a preservar os valores médios, os extremos e os valores de dispersão dos dados, garantindo que os três conjuntos de amostras possuam características semelhantes.

Ainda na quinta etapa, referente à modelagem hidrológica, foi executado um algoritmo para a simplificação da RNA, seguindo a metodologia detalhada em Oliveira et al. (2013a). Este método estima a importância da contribuição das variáveis de entrada no modelo e realiza a redução da dimensionalidade da RNA através da exclusão das variáveis me-

nos importantes. O método pode ser subdividido nos seguintes passos:

- a) realização de uma simulação com a rede inicial, composta por todas as variáveis de entrada inicialmente selecionadas (10 entradas), com verificação do desempenho através de algum indicador. Nesta etapa, o indicador utilizado para a verificação do desempenho foi o RMSE (*Root Mean Square Deviation*);
- b) imposição de pequenas perturbações (20 séries) nos valores de cada uma das variáveis de entrada, seguida de uma nova simulação e medida de desempenho. A utilização de 20 séries de perturbações, com diferentes magnitudes, em cada variável de entrada, mostrou-se necessária para estabilizar os resultados e não comprometer a análise;
- c) cálculo do Índice de Contribuição (CI), de acordo com a Equação 3, para cada variável de entrada. Este índice consiste na média das diferenças de RMSE, obtidas entre a simulação com dados observados e cada uma das 20 séries de dados perturbados;

$$CI_x = \frac{\sum |RMSE_o - RMSE_i|}{p} \quad (8)$$

Onde: CI_x é o Índice de Contribuição da variável de entrada x ; $RMSE_o$ é o valor do RMSE do modelo inicial, com dados originais; $RMSE_i$ é o valor de RMSE do modelo alterado, com perturbação dos registros da variável x , com valores de i variando de 1 a p ; p é o número de séries de perturbações.

- d) cálculo do Índice de Contribuição Relativa (RCI) de cada variável, que consiste na proporção (%) do CI de cada variável no somatório dos CIs das 10 variáveis;
- e) ordenamento das variáveis de entrada pelo RCI e exclusão das variáveis menos importantes, cujo RCI foi inferior a 5%;
- f) repetição das cinco atividades anteriores até que nenhuma variável de entrada possua RCI inferior a 5%, quando atinge-se um modelo simplificado, somente com as variáveis que realmente influenciam na saída da RNA.

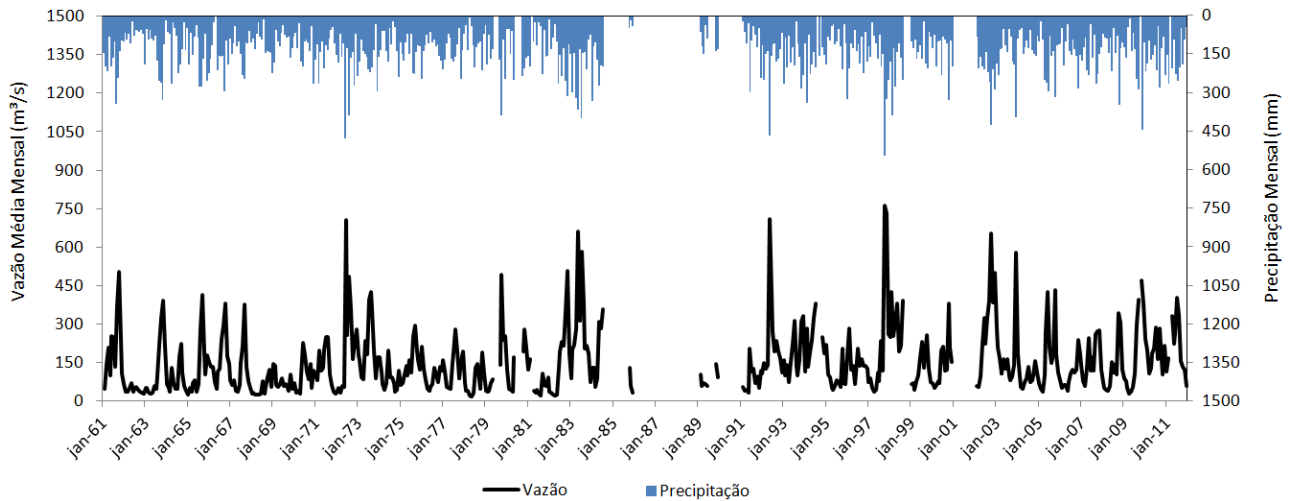


Figura 3 - Hidrograma na bacia do rio Ijuí, seção do posto Santo Ângelo, no período entre 1961 e 2011.

Tabela 1 - Séries de dados para o treinamento, a validação e a verificação do modelo hidrológico com RNAs.

Série	Nº de Amostras	% de Amostras	Períodos (mês/ano)
Treinamento	264	51,4%	02/1966 - 06/1974
			02/1978 - 02/1998
Validação	122	23,7%	09/1963 - 01/1966
			07/1974 - 01/1978 09/2007 - 12/2011
Verificação	128	24,9%	02/1961 - 08/1963
			03/1998 - 08/2007
Total	514	100,0%	01/1961 - 12/2011

Após a simplificação do modelo, com a obtenção de uma arquitetura de RNA mais eficiente, a sexta etapa consistiu na avaliação do modelo obtido, considerando os seguintes itens:

- análise visual comparativa dos hidrogramas (observado e simulado);
- análise visual comparativa das curvas de permanência das vazões mensais, observadas e simuladas;
- análise quantitativa através de indicadores de desempenho, como o coeficiente de Nash-Sutcliffe (NS), o RMSE e o erro relativo absoluto (ER);
- análise de sensibilidade do modelo, através de uma adaptação da metodologia apresentada por Lek et al. (1996): nesta adaptação verificou-se se o efeito causado na vazão de saída, ao alterar os valores de cada variável de entrada (fixando as demais variáveis em seus valores médios), é condizente com os processos físicos conhecidos. Neste estudo,

os valores adotados para a análise de sensibilidade de cada variável de entrada variaram de acordo com os dados apresentados na Tabela 2. Os valores adotados extrapolam os limites dos valores observados de cada variável, permitindo assim obter uma avaliação do comportamento da RNA em situações mais extremas do que as registradas na série histórica. Para a extrapolação dos limites foram observadas as condições climáticas previstas para o futuro, com base no modelo Eta (BLACK, 1994), no período entre 2011 e 2041.

Na sétima etapa foi realizada uma comparação do desempenho da RNA com relação ao desempenho do modelo SMAP, apresentado por Lopes et al. (1982). Trata-se de um modelo conceitual concentrado, do tipo chuva-vazão, que utiliza os parâmetros do modelo proposto pelo Soil Conservation Service - modelo SCS (SCS, 1975) - para a separação do escoamento. A principal diferença entre o

Tabela 2 - Intervalo de valores observados na série histórica e adotados para a análise de sensibilidade de cada variável.

Variáveis de Entrada	Valores Observados (mm)			Limites Adotados (mm)	
	Média	Mínimo	Máximo	Mínimo	Máximo
Precipitação: $P(t)$ e $P(t-1)$	143,8	6,7	546,9	0,0	630,0
Precipitação média (2 meses): \bar{P}	143,8	17,6	436,7	0,0	510,0
Evapotranspiração: $E(t)$ e $E(t-1)$	109,0	35,1	208,4	6,0	240,0
Evapotranspiração média (2 meses): \bar{E}	109,0	39,6	208,3	9,0	240,0
Balanco de água: $S(t)$	35,3	-189,4	456,5	-300,0	570,0
Balanco médio de água (2 meses): \bar{S}	69,7	-356,0	663,5	-540,0	840,0
filtro da precipitação: fP	144,2	45,8	357,4	0,0	390,0
filtro do balanço de água: fS	35,2	-109,5	246,8	-180,0	300,0

SMAP e o modelo SCS é que, no primeiro, é possível a simulação de séries contínuas de vazões, não apenas cheias de projeto. O modelo SMAP, na versão mensal, utiliza apenas dois reservatórios (subsuperficial e subterrâneo).

A escolha do modelo SMAP para a comparação se deve aos bons resultados obtidos com a sua aplicação em estudos recentes como em Block et al. (2009) e Kwon et al. (2012). Além disso, outro fator que influenciou na escolha do modelo hidrológico está relacionado à sua simplicidade, com apenas quatro parâmetros: capacidade de saturação do solo, coeficiente de escoamento, coeficiente de recarga e coeficiente de recessão do escoamento de base. Estes parâmetros devem ser calibrados com o uso de alguma função objetivo, como por exemplo, o RMSE.

RESULTADOS E DISCUSSÕES

Neste item são apresentados os principais resultados obtidos quanto ao desempenho do modelo baseado nas RNAs, à análise de sensibilidade e à comparação entre o modelo de RNAs e o modelo SMAP.

De acordo com o que foi descrito na metodologia do trabalho, o modelo hidrológico inicial apresentou 10 variáveis de entrada, combinadas com o mesmo número de neurônios na camada intermediária, o que resultou em 121 pesos sinápticos no modelo (ligações entre as variáveis e os neurônios da camada intermediária e de saída). Através do algoritmo de simplificação da RNA foram excluídas sete variáveis de entrada, resultando em um modelo hidrológico menos complexo, com apenas três entradas (\bar{S} , fP e fS) e 16 pesos sinápticos.

Estas variáveis remanescentes caracterizam-se principalmente por cada uma representar mais do que apenas um processo hidrológico que ocorre na bacia, enquanto que as entradas brutas de precipitação e de evapotranspiração (P e E) explicam apenas uma parcela reduzida do ciclo hidrológico. Variáveis como o saldo médio de água (\bar{S}), o filtro de precipitação (fP) e o filtro do saldo de água (fS) constituem-se em uma conjugação de processos hidrológicos, representando a disponibilidade e o armazenamento de água no solo e nos rios da bacia. Estas variáveis, por já conterem as informações das entradas brutas, além dos outros processos representados, reduzem as possíveis contribuições adicionais destas entradas, as quais tornam-se irrelevantes para a reprodução das vazões resultantes.

A Tabela 3 apresenta o Índice de Contribuição Relativa destas três variáveis (\bar{S} , fP e fS) no modelo inicial (10 entradas) e no modelo final. A variável mais importante durante a simulação foi a fS , com contribuição variando de 33% no modelo inicial (10 entradas) a 40% no modelo mais simplificado (três entradas). É importante destacar a elevada importância destas três variáveis desde o primeiro modelo, quando as três somavam mais de 70% da contribuição (RCI) de todas as variáveis juntas.

Ao comparar o desempenho das duas RNAs, com 10 entradas e com três entradas, através dos hidrogramas obtidos (Fig. 4 e 5), considerando a série total (1961-2011), pode-se observar que o desempenho dos dois modelos foi muito satisfatório, com pequena vantagem para o modelo inicial, com 10 variáveis de entrada. Ambos os modelos representaram de modo adequado os períodos de estiagem e de cheia, com suave tendência de subestimação das vazões mais altas.

Tabela 3 - Índice de Contribuição Relativa (RCI) das três variáveis mais importantes no modelo com RNAs.

Variável	RCI no modelo inicial	RCI no modelo final
\bar{S} (balanço médio de água nos tempos t e $t-1$)	23%	31%
fP (filtro com decaimento exponencial - precipitação)	15%	29%
fS (filtro com decaimento exponencial - balanço de água)	33%	40%
Somatório destas variáveis	71%	100%

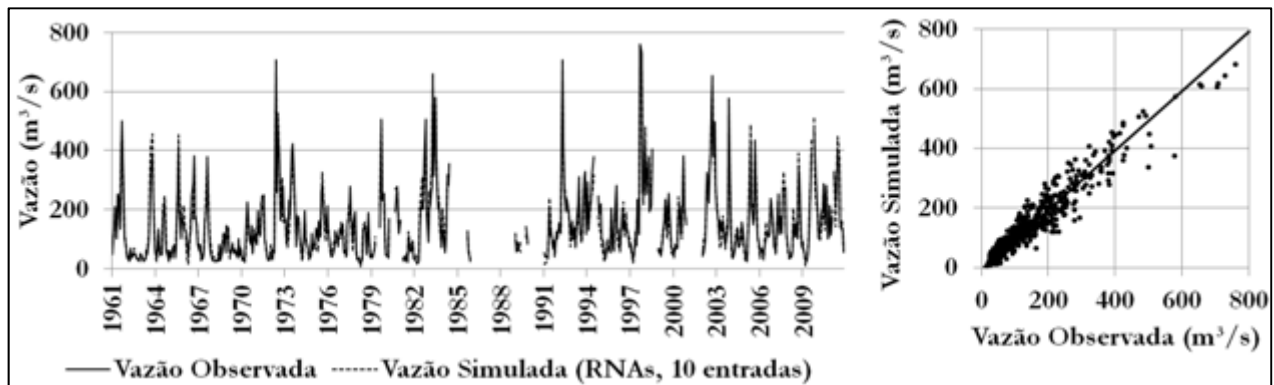


Figura 4 - Hidrogramas: observado e simulado pela RNA inicial (10 entradas), 1961 a 2011, bacia do rio Ijuí, posto Santo Ângelo.

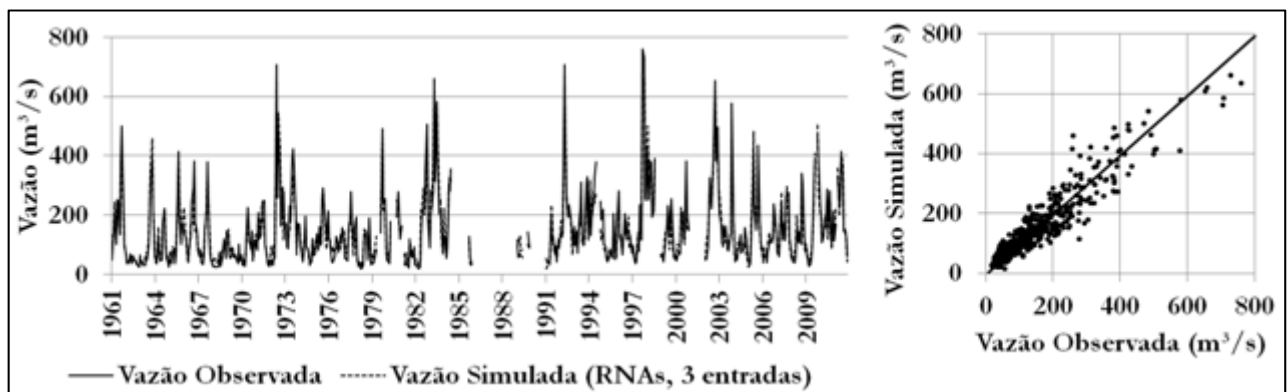


Figura 5 - Hidrogramas: observado e simulado pela RNA final (3 entradas), 1961 a 2011, bacia do rio Ijuí, posto Santo Ângelo.

A diferença de desempenho entre os modelos se torna ainda menor, praticamente imperceptível, quando considerada somente a série de verificação (Fig. 6). Pode-se verificar que a dispersão dos pontos nos dois gráficos da Figura 6 é muito peque-

na, indicando que os valores simulados são semelhantes aos observados. O coeficiente NS do modelo simplificado (RNA final: três entradas) foi de 0,904, enquanto que no modelo inicial (RNA: 10 entradas) foi de 0,907. Quanto ao RMSE, a RNA final apresen-

tou um valor de 37,1 m³/s, enquanto que a RNA inicial obteve RMSE igual a 36,4 m³/s. O erro relativo absoluto (ER) dos modelos foi de 21,7% (RNA com três entradas) e 19,1% (RNA com 10 entradas).

As curvas de permanência obtidas através da simulação (Fig. 7) ilustram também os bons resultados da modelagem com as RNAs. Pode-se observar que o ajuste da curva de permanência entre a Q₅ e a Q₆₀ é excelente nos dois modelos. No intervalo de vazões mais baixas (por exemplo, Q₈₀ ou Q₉₀), o modelo inicial apresentou resultados levemente mais consistentes, tendendo à subestimativa destas, enquanto que no modelo final, a tendência é de superestimativa das vazões neste intervalo.

Na estimativa da Q₅₀, cujo valor observado foi de 116,5 m³/s, os modelos atingiram ótimo desempenho, com valores de 114,6 m³/s e de 114,4 m³/s, nas RNAs com 3 e 10 variáveis de entrada, respectivamente. Já na estimativa da Q₉₀ (valor observado igual a 38,8 m³/s, o modelo com 10 entradas se aproximou mais, com valor de 38,4 m³/s, enquanto que o modelo mais simplificado obteve a vazão de 47,3 m³/s. Na estimativa da Q₉₅, cujo valor observado foi de 32 m³/s, os modelos apresentaram valores de 29,5 m³/s e de 38,9 m³/s, nas RNAs com 10 e 3 variáveis de entrada, respectivamente.

Ainda que, de modo geral, o desempenho do modelo inicial tenha sido levemente superior ao do modelo simplificado, é importante destacar que o modelo final possui aproximadamente 10% da dimensionalidade da RNA inicial, indicando que o mesmo é mais parcimonioso. A simplicidade do modelo resultante facilita a interpretação do funcionamento da rede, o que deve ser considerada como uma grande vantagem da aplicação do modelo com menos variáveis de entrada. Além disso, ao realizar a análise de sensibilidade, pode-se verificar que os processos físicos inerentes às variáveis de entrada foram representados mais adequadamente a partir da RNA com apenas três entradas, o que concede, *a posteriori*, uma certa base conceitual ao modelo resultante. No modelo final, a vazão apresentou um aumento contínuo, à medida que os valores de \bar{S} , fP e fS aumentaram, inclusive nos intervalos extremos, para os quais a RNA não foi treinada (Fig. 8). Este comportamento condiz com o esperado, pois, tanto as variáveis resultantes da aplicação dos filtros às precipitações quanto o saldo de água representam a quantidade de água disponível na bacia. Fisicamente, quanto maior a disponibilidade de água, maior é a vazão.

A Figura 8 ilustra a tendência de aumento da vazão simulada em função do aumento dos valores de \bar{S} , fP e fS. Considerando os limites de -540 e

840 mm para a variável de balanço de água (\bar{S}), fixando-se as demais variáveis com o valor médio observado, as vazões variam de 84 a 407 m³/s. Já para a variável fP (filtro de precipitação), para valores de 0 a 390 mm, as vazões oscilam entre 64 e 332 m³/s. Por fim, para a variável fS (filtro do balanço de água), as vazões também apresentaram valores crescentes, de 37 a 348 m³/s.

O mesmo não ocorre com o modelo inicial, quando pode-se observar que algumas variáveis não foram bem interpretadas na RNA (Fig. 9). Um bom exemplo é a variável P(t-1), a qual deveria influenciar, de forma contínua e positivamente, no aumento das vazões, tendência esta que não é observada no gráfico.

Podem-se mencionar também as variáveis de evapotranspiração no tempo t, E(t), e de evapotranspiração média nos tempos t e t-1, \bar{E} , que deveriam provocar a redução da vazão à medida que seus valores aumentam, uma vez que representam a perda de água para a atmosfera. Entretanto, este processo físico foi interpretado de forma muito equivocada pelo modelo de RNAs inicial, como pode ser observado claramente na Figura 9

Isto indica que, embora os resultados tenham sido um pouco mais precisos, a capacidade de generalização da RNA inicial não é tão satisfatória quanto a da RNA simplificada, sendo praticamente impossível interpretar/compreender o funcionamento do modelo com 10 entradas.

Esta constatação deve ser considerada na escolha do modelo hidrológico, ainda mais quando pretende-se simular a vazão em cenários de mudanças climáticas. Neste caso, é imperativa a necessidade de se obter um modelo capaz de proporcionar a adequada interpretação do processo físico que envolve cada variável de entrada.

Assim, com base nos resultados e discussões apresentados, aceita-se que o modelo simplificado de RNAs, com apenas três variáveis, é o mais adequado para a simulação hidrológica na bacia do rio Ijuí, para aplicações que envolvam apenas dados relacionados à precipitação e à evapotranspiração.

Uma vez escolhida a melhor configuração de rede (hidrológicamente consistente), foi realizada uma comparação dos resultados obtidos pela RNA com o modelo hidrológico SMAP. A Figura 10 apresenta uma comparação entre os dois modelos através de um gráfico de dispersão (Q observada x Q simulada), considerando apenas o período de verificação, de fevereiro de 2002 a setembro de 2009.

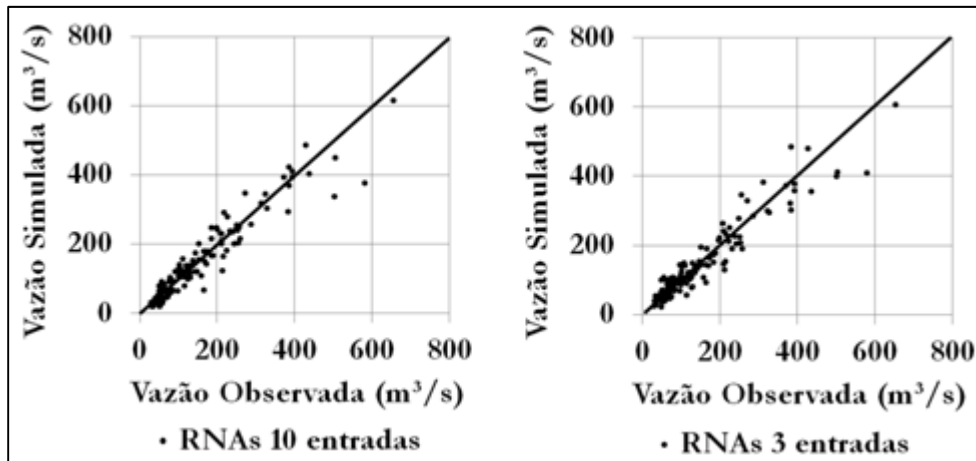


Figura 6 - Gráfico de dispersão entre as vazões calculadas e observadas, considerando o período de verificação (1961-1963; 1998-2007) – bacia do rio Ijuí, posto Santo Ângelo.

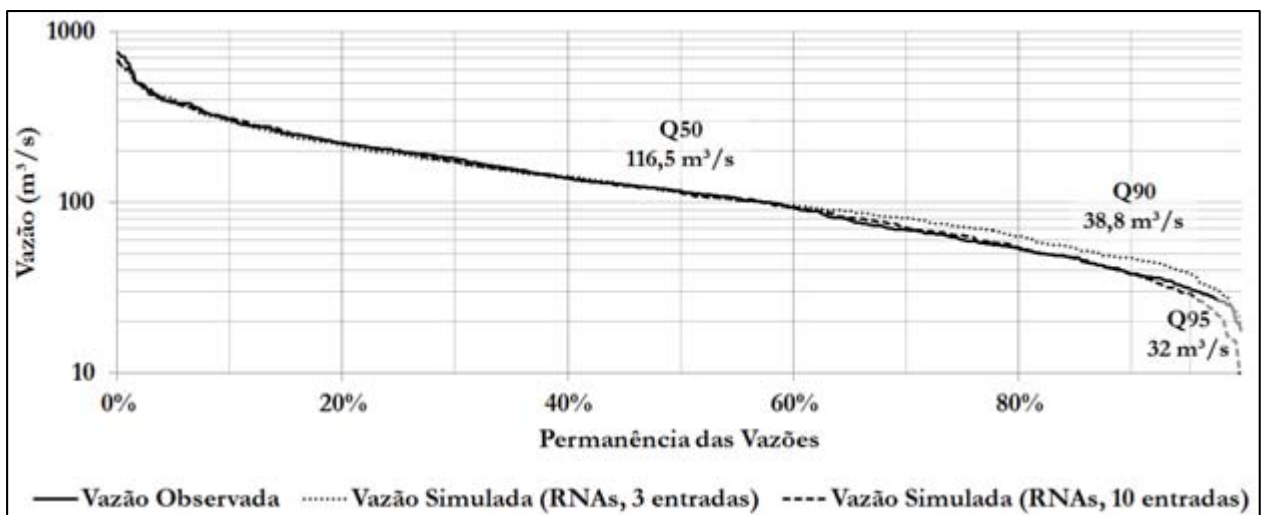


Figura 7 - Curvas de permanência das vazões no rio Ijuí, posto Santo Ângelo (1961-2011): observada e simuladas com os modelos de RNAs (com 3 e 10 entradas).

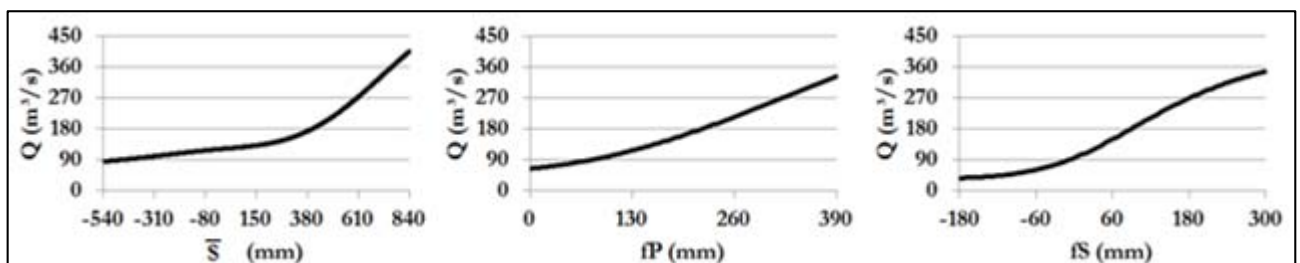


Figura 8 - Sensibilidade da vazão no rio Ijuí às variáveis: balanço de água (\bar{S}), filtro de precipitação (fP) e filtro de balanço de água (fS) - Modelo de RNAs simplificado, com três entradas.

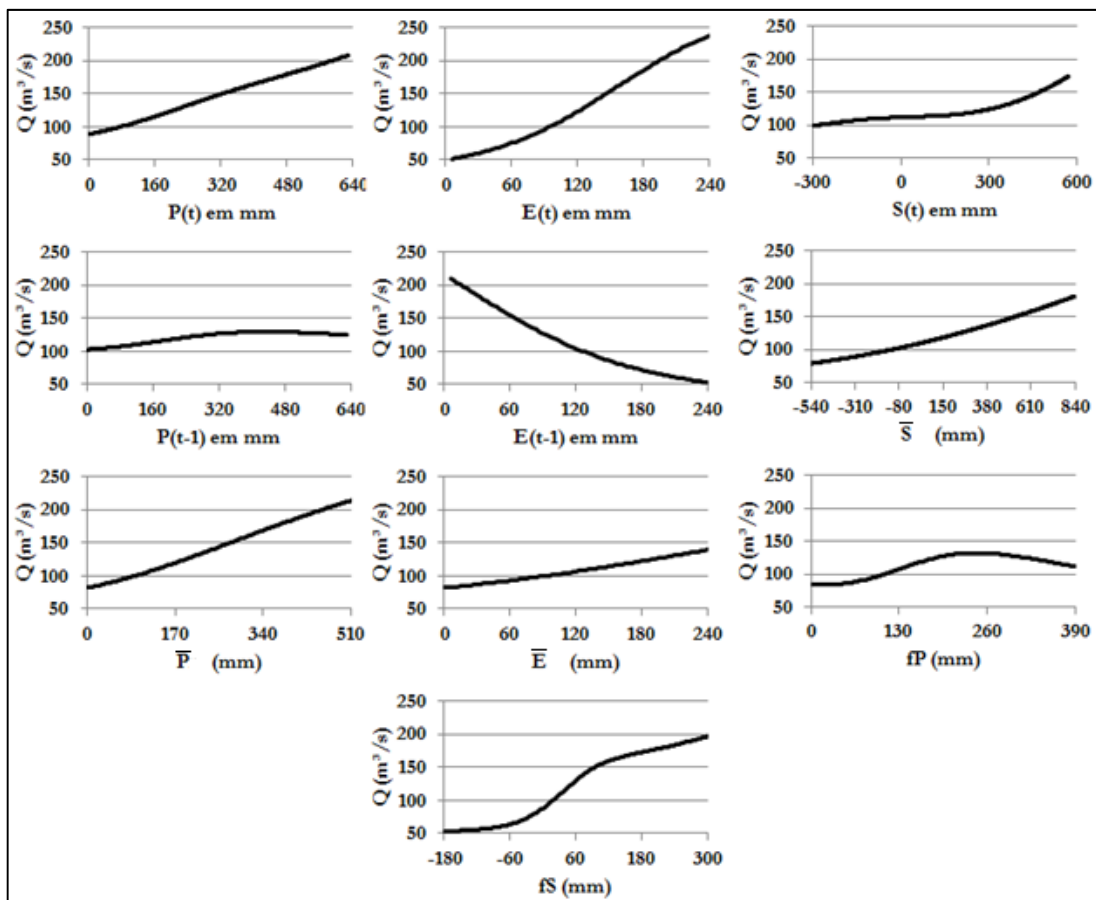


Figura 9 - Sensibilidade da vazão no rio Ijuí às variáveis do modelo de RNAs inicial (com 10 entradas).

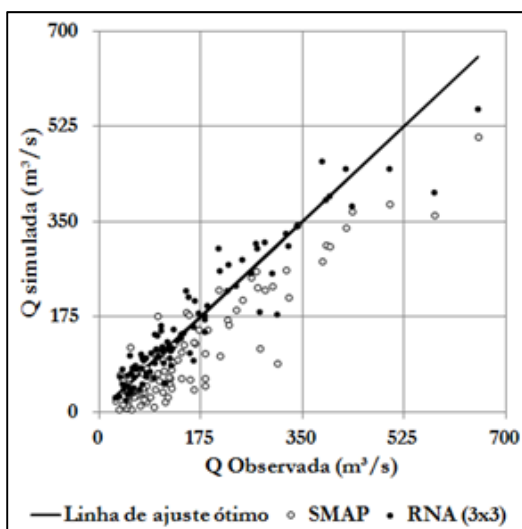


Figura 10 - Gráfico de dispersão entre as vazões calculadas pelo modelo SMAP e pelo modelo de RNAs, e as vazões observadas no posto Santo Ângelo (fev/2002 – set/2009).

Fica evidente que a simulação utilizando o modelo com RNAs apresenta resultados muito mais consistentes em relação ao SMAP, para todos os níveis de vazão (períodos de estiagem e de cheia). O SMAP, de modo geral, apresentou uma significativa tendência de subestimação das vazões.

Ao comparar dois indicadores de desempenho dos modelos, considerando o período de verificação (2002-2009), a diferença fica ainda mais evidente: a) o coeficiente NS da simulação com o modelo SMAP foi de 0,672, enquanto que com as RNAs o NS foi de 0,898; b) o RMSE foi significativamente superior com a aplicação do modelo SMAP, com valor de $55,7 m^3/s$, enquanto que com as RNAs o valor foi de $38,2 m^3/s$.

Diante dos resultados observados na avaliação dos modelos hidrológicos com RNAs, e considerando a comparação com o modelo SMAP, pode-se afirmar que a rede neural com apenas três variáveis de entrada representou satisfatoriamente o comportamento hidrológico do rio Ijuí, e constitui-se em

uma excelente alternativa para a modelagem das vazões médias mensais. Mesmo tratando-se de um modelo hidrológico de origem empírica, sem embasamento físico preconcebido, as RNAs interpretaram adequadamente os processos hidrológicos associados às precipitações e à evapotranspiração na bacia.

Isto ocorreu porque o modelo resultante tornou-se bastante sofisticado, compreendendo também transformações das variáveis de entrada e transformações de escala. Como exemplo, pode-se mencionar a inclusão de variáveis com o filtro de decaimento exponencial, aplicado às precipitações e ao saldo de água, que gera um significativo incremento no desempenho do modelo hidrológico. Tais variáveis representam a condição antecedente de umidade do solo na bacia hidrográfica, integrando satisfatoriamente mais processos hidrológicos do que as entradas de precipitação e de evapotranspiração nos tempos anteriores. Deste modo, a RNA tornou-se o núcleo deste modelo mais amplo, que acabou incorporando alguma base conceitual, como resultado da pesquisa para a sua definição final.

Além disto, outra contribuição relevante deste trabalho se refere à possibilidade da aplicação deste tipo de modelo, em intervalo de tempo mensal, em estudos referentes à disponibilidade hídrica no futuro, tendo em vista a razoável capacidade de extrapolação em relação aos dados observados no passado, conforme observado na análise de sensibilidade. Tal modelo pode ser aplicado em estudos hidrológicos relacionados à variabilidade e mudança climática, temáticas intensamente discutidas na atualidade por cientistas da área ambiental e de recursos hídricos.

CONCLUSÕES

Considerando o conjunto de resultados e das discussões apresentadas neste trabalho, conclui-se que a metodologia proposta para a modelagem das vazões mensais, na bacia hidrográfica do rio Ijuí, com base em RNAs, apresentou resultados muito consistentes. Esta consistência é demonstrada não apenas pelos indicadores quantitativos de desempenho, mas também pelas análises de hidrogramas e de curvas de permanência, além das análises de sensibilidade dos modelos, que proporcionaram uma interpretação do funcionamento das RNAs. Além disso, a avaliação ainda compreendeu uma breve comparação com o modelo determinístico SMAP.

O modelo identificado como o mais adequado para a simulação das vazões na bacia do rio

Ijuí, considerando os indicadores de eficiência, a parcimônia e a representação dos processos hidrológicos, foi o obtido no final do procedimento de simplificação da RNA, com apenas três das dez variáveis de entrada selecionadas previamente. Foram elas: o balanço médio de água nos tempos t e $t-1$ (\bar{S}); o filtro da precipitação (FP); o filtro do balanço de água (FS). Neste modelo, o número de pesos sinápticos foi reduzido a apenas 16.

Utilizando este modelo, foi obtido um coeficiente NS igual a 0,904, com RMSE de 37,1 m³/s e erro relativo absoluto de 21,7%. Embora este não tenha sido o melhor desempenho (quantitativo) observado durante todas as tentativas de simulação, esta configuração se mostrou consistente e a mais eficiente, principalmente, por sua simplicidade (fácil interpretação e mais controle por parte do pesquisador) e por representar adequadamente os processos físicos envolvidos entre as variáveis de entrada e a vazão de saída, como foi observado na análise de sensibilidade do modelo. Além disto, o modelo de RNA selecionado apresentou uma razoável capacidade de extrapolação, tendo-se em vista que foram testados valores que excediam os limites utilizados no treinamento da rede.

Por fim, ao comparar o desempenho das simulações realizadas com a RNA simplificada e o modelo SMAP, ficou ainda comprovada a qualidade do modelo baseado em RNAs. A diferença de NS entre as duas simulações foi de 0,226 (RNAs: 0,898; SMAP: 0,672), e a diferença de RMSE foi de 17,5 m³/s.

AGRADECIMENTOS

Agradecimento a FINEP pelo financiamento da pesquisa - Chamada pública MCT/FINEP CT-HIDRO 01/2010 - Convênio 01.12.0396.00, Projeto - Rede de Pesquisa em Monitoramento e Modelagem de Processos Hidrossedimentológicos em Bacias Representativas Rurais e Urbanas do Bioma Mata Atlântica (RHIMA). Agradecimento ao CNPq pela bolsa de doutorado do primeiro autor deste estudo.

REFERÊNCIAS

- AKAIKE, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723.
- ALLEN, R.G., PEREIRA, L.S., RAES, K., SMITH, M. 1998. Crop evapotranspiration (guidelines for computing crop water

- requirements). Rome: FAO. 300 p. (Irrigation and Drainage Paper, 56).
- ANMALA, J., ZHANG, B., GOVINDARAJU, R.S. 2000. Comparison of ANNs and empirical approaches for predicting watershed runoff. *Journal of Water Resources Planning and Management*, Reston: ASCE, v.126, n.3, 156-166.
- BLACK, T.L. 1994. The new NMC mesoscale Eta Model: Description and forecast examples. *Weather Forecasting*, 9, 265-278.
- BLOCK, P.J., SOUZA FILHO, F.A., SUN, L., KWON, H. 2009. A streamflow forecasting framework using multiple climate and hydrological models. *Journal of the American Water Resources Association*, 45:4, 828-843.
- BOWDEN, G.J., DANDY, G.C., MAIER, H.R. 2005. Input determination for neural network models in water resources applications. Part I – background and methodology. *Journal of Hydrology*, 301, 75-92.
- CAMPOLO, M., SOLDATI, A., ANDREUSSI, P. 2003. Artificial neural network approach to flood forecasting in the River Arno. *Hydrological Sciences Journal*, 48:3, 381-398.
- CASTRO, N.M.R., AUZET, A.V., CHEVALLIER, P., LEPRUN, J.C. 1999. Land use change effects on runoff and erosion from plot to catchment scale on the basaltic plateau of Southern Brazil. *Hydrological Processes*, 13 (11), 1621-1628.
- CASTRO, N.M.R., CHEVALLIER, P., GOLDENFUM, J.A. 2000. Projeto Potiribu, atualização 1989-1998 – dados básicos de fluviometria e pluviometria. IPH: UFRGS, *Revista de Recursos Hídricos*, v. 35.
- CASTRO, N.M.R., PINHEIRO, A., FERNANDES, C.V.S., PAIVA, E.M.C.D., MERTEN, G.H., MINELLA, J.P.G., PAIVA, J.B.D., GOLDENFUM, J.A., SILVA, M.R., GASTALDINI, M.C.C., BRAGA, S.M. 2010. Relatório Final das atividades de março de 2007 a maio de 2010. Relatório de pesquisa – IPH, UFRGS, Porto Alegre, Brasil.
- DORNELLES, F.; GOLDENFUM, J.; PEDROLLO, O.; 2013. Artificial neural network methods applied to forecasting river levels. RBE. *Revista Brasileira de Engenharia*. Caderno de Recursos hídricos. ISSN: 0102-2652. Aceito para publicação no vol. 18, nº 4, outubro/dezembro/2013.
- HECHT-NIELSEN, R. 1987. Kolmogorov's mapping neural network existence theorem. *Proceedings of the First IEEE International Joint Conference on Neural Networks*, San Diego, California, pp. 11–14, IEEE, New York.
- HECHT-NIELSEN, R. 1990. *Neurocomputing*. Addison-Wesley, 1990. pp. 433.
- HORNIK, K.; STINCHCOMBE, M.; WHITE, H.. 1989. Multilayer feedforward networks are universal approximators. *Neural networks*, v. 2, n. 5, p. 359-366.
- HSU, K.L., GUPTA, H.V., SOROOSHIAN, S. 1995. Artificial neural network modeling of the rainfall-runoff process. *Water Resources Research*, 31:10, 2517-2530.
- JAIN, A., KUMAR, A.M. 2007. Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing*, 7, 585-592.
- JAIN, A.; KUMAR, S. 2009. Dissection of trained neural network hydrologic models for knowledge extraction. In: *Water Resources Research*, vol. 45; 1-17.
- JAIN, A., SUDHEER, K.P., SRINIVASULU, S. 2004. Identification of physical processes inherent in artificial neural network rainfall runoff models. *Hydrological Processes*, 18, 571-581.
- KWON, H., SOUZA FILHO, F.A., BLOCK, P.J., SUN, L., LALL, U., REIS JR, D.S. 2012. Uncertainty assessment of hydrologic and climate forecast models in Northeastern Brazil. *Hydrological Processes*, 26, 3875-3885.
- LEAHY, P., KIELY, G., CORCORAN, G. 2008. Structural optimisation and input selection of an artificial neural network for river level prediction. *Journal of Hydrology*, v. 355, 192-201.
- LEK, S., DELACOSTE, M., BARAN, P., DIMOPOULOS, I., LAUGA, J., AULAGNIER, S. 1996. Application of neural networks to modelling nonlinear relationships in ecology. *Ecological Modelling* 90, 39-52.
- LOPES, J.E., BRAGA, B.F.F., CONEJO, J.L. 1982. A Simplified Hydrologic Model in Applied Modeling in Catchment Hydrology. *Water Resources Publication*.
- MCCULLOCH, W., PITTS, W. 1943. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115-133.
- MONTEITH, J.L. 1965. Evaporation and Environment. In: *SYMPOSIA OF THE SOCIETY FOR EXPERIMENTAL BIOLOGY*, 19., 1965, Swansea. *Proceedings...Cambridge: University Press*, 235-239.
- OLIVEIRA, G.G., PEDROLLO, O.C., CASTRO, N.M.R. 2011. Metodologia de análise de sensibilidade e exclusão de variáveis de entrada em simulação hidrológica por Redes Neurais

Artificiais (RNAs): resultados preliminares. In: Anais XIX Simpósio Brasileiro de Recursos Hídricos, 2011, Maceió.

OLIVEIRA, G.G., PEDROLLO, O.C., CASTRO, N.M.R. 2013a. Simplifying Artificial Neural Network models of river basin behaviour by an automated procedure for input variable selection. Artigo submetido à Water Resources Research.

OLIVEIRA, G.G., PEDROLLO, O.C., CASTRO, N.M.R. 2013b. Simulações hidrológicas com diferentes proporções de área controlada na bacia hidrográfica. Artigo aceito para publicação na Revista Brasileira de Recursos Hídricos.

PENMAM, M.L. 1948. Evaporation: an introductory survey. Neth. Jour. Of Agric. Science, 4, 9-29.

RAJURKAR, M.P., KOTHYARI, U.C., CHAUBE, U.C. 2004. Modeling of daily rainfall-runoff relationship with artificial neural network. Journal of Hydrology, London: Elsevier, v.285, n.1-4, 96-113.

RUMELHART, D.E., HINTON, G.E., WILLIAMS, R.J. 1986. Learning representations by back-propagating errors. Nature, 323, 533-536.

SAJIKUMAR, N., THANDAVESWARA, B.S. 1999. A non-linear rainfall-runoff model using an artificial neural network. Journal of Hydrology, 216, 32-55.

SCS. 1975. Urban hydrology for small watersheds. Washington. US Dept. Agr. Technical Release, n. 55.

SEDIYAMA, G.C. 1996. Evapotranspiração: necessidade de água para os cultivos. Brasília: ABAES. 167p.

SHAMSELDIN, A.Y. 1997. Application of a neural network technique to rainfall-runoff modelling. Journal of Hydrology, 199, 272-294.

SIBSON, R. 1981. A brief description of natural neighbor interpolation. In: Barnett, V. (Ed.). Interpreting Multivariate Data, Wiley, Chichester, 21-36.

SILVA, V.S.V. 2011. Estimativa de precipitação pontual em diferentes escalas para uso em modelo concentrado chuva-vazão. Dissertação de Mestrado – Instituto de Pesquisas Hidráulicas da Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil.

VOGL, T.P.; MANGIS, J.K.; ZIGLER, A.K.; ZINK, W.T.; ALKON, D.L. 1988. Accelerating the convergence of the backpropagation method. Bio. Cybern., vol. 59, pp. 256-264, Sept. 1988. WIDROW, B.; HOFF, M. E. 1960. Adaptive switch-

ing circuits. IRE WESCON Convention Record, New York: IRE Part 4, pp. 96-104.

WIDROW, B., HOFF, M.E. 1960. Adaptive switching circuits. In: 1960 IRE WESCON Convention Record, New York: IRE Part 4, 96-104.

The Performance Of Artificial Neural Networks (ANNs) For Monthly Hydrological Simulation

ABSTRACT

The aim of this study was to evaluate the performance of hydrological models based on ANNs for hydrologic simulation of mean monthly flows, using only meteorological and rainfall data. This study was performed in Rio Ijuí Basin, in the area of contribution to the Santo Ângelo river gauging station (5,414 km²), in Rio Grande do Sul, Brazil, with average flow of 138 m³/s in the period between 1941 and 2005. The following research procedures were adopted: i) organization and consistency analysis of data, ii) spatial interpolation of meteorological and rainfall data, iii) estimation of reference evapotranspiration, iv) monthly rainfall, evapotranspiration and flow extraction, v) hydrological modeling with ANNs. The evaluation of the tested models was performed through visual comparative analysis of hydrographs and permanence curves (observed and simulated), the quantitative performance indicators such as the Nash-Sutcliffe coefficient (NS), and a sensitivity analysis to understand and interpret the functioning of ANN. Furthermore, a comparison was performed with the hydrological model SMAP, a conceptual model adapted for monthly simulations, proposed by Lopes et al. (1982). The ANN model that was most suitable for flow simulation presented only three input variables, with 16 synaptic weights. Using this model an NS equal to 0.904 was obtained, with RMSE equal to 37.1 m³/s. Sensitivity analysis showed that the ANN related correctly chosen input variables to the network output, respecting the physical principles involved in the hydrological system. Furthermore, it was found that the ANN model presented a reasonable extrapolation capability. A large difference was observed in NS (ANNs: 0.898, SMAP: 0.672), comparing the performance of the simulations with the ANN model and SMAP model, between 2002 and 2009.

Key-words: Rio Ijuí Basin; hydrological modelling; monthly streamflow; sensitivity analysis.