

[ARTIGO RETRATADO] REDES NEURAIAS ARTIFICIAIS APLICADAS À PREVISÃO DE ENCHENTES PARA REGIÃO DO PANTANAL NO MATO GROSSO DO SUL

Marcia Ferreira CRISTALDO¹, Leandro de JESUS¹, Paulo Tarso de OLIVEIRA², Celso Correia de SOUZA³, Hevelyne H. G. VIGANÓ¹, Carlos Roberto PADOVANI⁴

¹Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul. Rua Taquarí, 831 - Santo Antonio, Campo Grande – MS. E-mails: marcia.cristaldo@ifms.edu.br; leandro.jesus@ifms.edu.br; hevelyne.vigano@ifms.edu.br

²Universidade Federal de Mato Grosso do Sul, UFMS, Cidade Universitária, Avenida Costa e Silva - Pioneiros, MS Brasil. E-mail: paulotarsoms@gmail.com

³Universidade Anhanguera Uniderp. Avenida Ceará, 333 - Vila Miguel Couto, Campo Grande - MS E-mail: csouza939@gmail.com

⁴Empresa Brasileira de Pesquisa Agropecuária, Centro de Pesquisa Agropecuária do Pantanal (Embrapa

Os dados da Tabela 2 na página <199> e da Figura 10 na página <200> do artigo “REDES NEURAIAS ARTIFICIAIS APLICADAS À PREVISÃO DE ENCHENTES PARA REGIÃO DO PANTANAL NO MATO GROSSO DO SUL” foi publicada em outra revista em 2018 [Rev. Ambient. Água vol.13 no.4 Taubaté 2018 Epub Aug 13, 2018, <http://dx.doi.org/10.4136/ambi-agua.1988>] COM REPRODUÇÃO INTEGRAL DA Tabela 2 e Figura 10, e não deve estar publicada no periódico Revista Geociências, vol 39, p. 191-201, 2020. Em função do ocorrido, solicito a retirada do artigo de publicação.

No artigo <REDES NEURAIAS ARTIFICIAIS APLICADAS À PREVISÃO DE ENCHENTES PARA REGIÃO DO PANTANAL NO MATO GROSSO DO SUL> publicado no periódico <Revista Geociências>,

<vol>(39):<191-201>, na página <199>

onde se leia:

“<Tabela 2 - Análise comparativa dos desempenhos.>”

leia-se:

“<Retirar Tabela 2>”

No artigo <REDES NEURAIAS ARTIFICIAIS APLICADAS À PREVISÃO DE ENCHENTES PARA REGIÃO DO PANTANAL NO MATO GROSSO DO SUL> publicado no periódico <Revista Geociências>,

<vol>(39):<191-201>, na página <200>

onde se leia:

“<Figura 10 - Evolução temporal do nível previsto para 2014 (laranja), observado (azul) e respectivos diagramas (r₂). Os dados correspondentes ao período de 1995 a 2014, obtidos das previsões 1 dia (a), 2 dias (b), 3 dias (c), 4 dias (d), 5 dias (e).>”

leia-se:

“<Retirar Figura 10>”



Profa. Dra. Marcia Ferreira Cristaldo

(Solicitante)

REDES NEURAIIS ARTIFICIAIS APLICADAS À PREVISÃO DE ENCHENTES PARA REGIÃO DO PANTANAL NO MATO GROSSO DO SUL

ARTIFICIAL NEURAL NETWORKS APPLIED TO FLOOD FORECASTING FOR THE PANTANAL REGION IN MATO GROSSO DO SUL

Marcia Ferreira CRISTALDO¹, Leandro de JESUS¹, Paulo Tarso de OLIVEIRA², Celso Correia de SOUZA³, Hevelyne H. G. VIGANÓ¹, Carlos Roberto PADOVANI⁴

¹Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul. Rua Taquarí, 831 - Santo Antonio, Campo Grande – MS. E-mails: marcia.cristaldo@ifms.edu.br; leandro.jesus@ifms.edu.br; hevelyne.vigano@ifms.edu.br

²Universidade Federal de Mato Grosso do Sul, UFMS, Cidade Universitária, Avenida Costa e Silva - Pioneiros, MS Brasil. E-mail: paulotarsoms@gmail.com

³Universidade Anhanguera Uniderp. Avenida Ceará, 333 - Vila Miguel Couto, Campo Grande - MS E-mail: csouza939@gmail.com

⁴Empresa Brasileira de Pesquisa Agropecuária, Centro de Pesquisa Agropecuária do Pantanal (Embrapa Pantanal). Rua 21 de Setembro número 1880. Corumbá – MS. E-mail: carlos.padovani@gmail.com

Introdução
Rede Neural Artificial
Treinamento da Rede Neural Artificial
Rede Multilayer Perceptron (MLP)
Análise de Componentes Principais (ACP)
Método de Thiessen
Material e Métodos
Área de Estudos
Dados Disponíveis
Resultados e Discussões
Resultados
RNAs Propostas
Conclusões
Agradecimentos
Referências

RESUMO - Desde 1941 existe relatos de enchentes na cidade de Aquidauana e Anastácio. Nesse contexto, este estudo teve como objetivo prever enchentes na área urbana da cidade de Aquidauana, utilizando a técnica de Redes Neurais Artificiais (RNAs). O modelo de RNA proposto foi o Perceptron Multi Camadas (MLP), sendo usados os dados de cotas do rio Aquidauana e da precipitação acumulada do período de 1995 a 2014 em pontos da bacia do rio a montante da cidade de Aquidauana. Para reduzir o número de dados meteorológicos e melhor explicar a variância da matriz de observações, foi aplicada aos dados a técnica de Análise de Componentes Principais (ACP), com o cálculo do coeficiente de correlação r . Os melhores resultados dos coeficientes de correlação foram para um e três dias de antecedência de previsão, com os valores de r iguais a 97% e 86%, respectivamente. Os resultados indicaram que as RNAs forneceram meios confiáveis de detectar o risco de enchentes na cidade de Aquidauana, com previsão de até cinco dias de antecedência.

Palavras chaves: Rede Neural Artificial, Análise de Componentes Principais, Previsão de cheias.

ABSTRACT - Since 1941, there have been reports of flooding in the city of Aquidauana and Anastácio. In this context, this study aimed at forecast flood the urban area of the city of Aquidauana, using the technique of Artificial Neural Networks (ANNs). The ANN model adopted was the Perceptron Multi Layers (MLP), using data from the quotas of the Aquidauana river and the accumulated precipitation from 1995 to 2014 in points of the river basin in the city of Aquidauana. To reduce the number of meteorological data, was applied to the technical data of the Principal Component Analysis (ACP), with the calculation of the correlation coefficient r . The best results of the correlation coefficients were for one and three days before the forecast, with values equal to 97% and 86%, respectively. The results indicate that the RNAs provide means of detection or risk of flooding in the city of Aquidauana, with a forecast of up to five days before.

Keywords: Artificial Neural Networks, Principal Component Analysis, Flood forecast.

INTRODUÇÃO

De acordo com Joia & Anunciação (2013), os problemas com as cheias no rio Aquidauana são catalogados desde de novembro de 1941, com precipitações mensais, em sua bacia, de 192 mm e 260 mm, respectivamente. Segundo os moradores mais antigos, essas foram as maiores inundações na bacia do rio Aquidauana. Os

mesmos autores relatam que de novembro de 1946 a fevereiro de 1947 as precipitações também foram muito expressivas. Houve registros de inundações em 1997, 2001, 2006, 2010 e 2011 com níveis impactantes, sendo o 2011 o ano que desalojou o maior número de ribeirinhos.

Joia & Anunciação (2013) enfatizam que, durante as últimas duas décadas, o rio Aquidauana apresentou aumento excessivo das águas, com o transbordamento ocasionando diversos transtornos para a população local. Em maio de 1990 o nível do rio chegou a 9,5 m, medido na régua localizada na ponte que une as cidades de Aquidauana e Anastácio. Em março de 2000 esse nível alcançou a marca de quase 9,0 m; em janeiro de 2010, chegou a 9,2 m; e, no ano de 2011, o ápice se deu no dia 4 de março, quando o nível das águas ultrapassou a altura de 10 m.

Como a região situa-se na borda do Pantanal, as cheias afetam as atividades econômicas desenvolvidas na região, assim o autor Padovani (2010) completa que,

“apesar da importância da dinâmica da inundação para os diversos compartimentos ecológicos do sistema do Pantanal, esse processo ainda não é bem compreendido. Isso se deve, principalmente, à falta do emprego de métodos de análise e tecnologias que permitam caracterizar e monitorar de forma contínua o ambiente, além de permitir que se façam cenários e previsões do comportamento do sistema como um todo”.

Para realizar previsões, os modelos necessitam de calibração de grande número de parâmetros, o que pode se tornar uma tarefa tediosa e, quando automática, é dependente da eficiência de um algoritmo de calibração automático (Tucci & Collischonn, 2012).

Neste contexto, alguns modelos de previsão utilizando RNAs já estão sendo desenvolvidos conforme Thirumalaiah & Deo (2000), Fazel et al. (2014), Ruslan et al. (2014) e Roy et al (2012). Os modelos são desenvolvidos para bacias pequenas com modelos específicos. Nesse contexto, as previsões utilizando RNA têm sido apresentadas como uma alternativa aos métodos tradicionais na solução de problemas de previsão de séries temporais, como é o caso das vazões ou níveis de um rio. O uso da técnica de RNA conforme Ruslan et al. (2014) interessante na medida em que não é necessário se ter um conhecimento físico-matemático prévio das relações complexas não lineares que descrevem os modelos matemáticos de precipitação-vazão

em bacias hidrográficas que não dispõem de dados hidrológicos mais detalhados, tais como fisiografia, estado e constituição do solo e umidade antecedente.

A RNA para previsão é utilizada não só para alertas de inundações, mas também em outras aplicações, tais como: previsão de níveis para navegação, disponibilidade hídrica para irrigação, abastecimento de água e operação energética. Segundo Ruslan et al. (2014), as previsões baseadas em observações do nível d'água próximo à nascente de um rio são mais precisas, entretanto proporcionam um horizonte de previsão menor e, por isso, sua utilização é mais observada em rios de médio ou grande porte cujo tempo de deslocamento da onda de inundação é maior.

Tendo como característica desse estudo uma bacia pequena, Tucci (2005) considera que uma bacia hidrográfica pequena tem como opção somente a previsão com base na precipitação. Assim, em situações específicas, é preferível implementar um simples modelo “caixa-preta” para identificar um mapeamento direto entre entradas e saídas, sem uma consideração detalhada da estrutura interna dos processos físicos envolvidos. Nesse sentido, as RNAs são, provavelmente, a técnica de aprendizagem de máquina, modelo caixa-preta, de maior sucesso e com uma estrutura matemática flexível capaz de identificar as relações não lineares complexas entre entrada e saída, sem a necessidade de entendimento dos fenômenos naturais (Fazel et al., 2014).

Segundo Fazel et al. (2014), desde que haja relações não lineares e incertezas entre muitos parâmetros numa modelagem de previsão de vazão ou nível d'água, a calibração de modelos conceituais ou físicos é frequentemente um procedimento complicado e que consome muito tempo. O objetivo deste trabalho foi desenvolver um modelo de previsão de enchentes baseado em Redes Neurais Artificiais (RNA) do tipo *Multilayer Perceptron* (MLP) para as cidades de Aquidauana e Anastácio, localizadas na sub-bacia do rio Aquidauana, no planalto adjacente ao Pantanal Sul Mato-Grossense.

REDE NEURAL ARTIFICIAL

Estudos recentes relatam que uma RNA pode oferecer uma alternativa promissora para a previsão hidrológica de inundação (Ozgur, 2011). A RNA é um modelo computacional que foi

projetado inspirando-se no funcionamento do cérebro humano e sua capacidade de aprender tarefas (Haykin, 2013). Assim, essa técnica não é baseada em regras comuns; ela se difere de outras

formas de inteligência computacional pelo fato de ser treinada para reconhecer e generalizar a relação entre um conjunto de entradas e saídas.

As RNAs foram inspiradas pelo funcionamento do cérebro humano com as redes neurais biológicas, composta de neurônios que se comunicam por meio de sinapses neurotransmissoras através das quais os impulsos nervosos são transmitidos, propiciando a capacidade de aprender. O neurotransmissor pode excitar ou inibir a geração de pulsos nesse outro neurônio (Haykin, 2013).

A arquitetura de uma RNA é composta por camadas, com unidades (neurônios artificiais)

conectadas a unidades de outras camadas. Cada conexão está associada a um peso, que é ajustado durante o treinamento, dando ao algoritmo o comportamento inteligente. A quantidade de camadas da RNA e o número de neurônios artificiais de cada camada podem variar, mas, necessariamente, têm-se uma camada para entrada de dados, uma camada intermediária e uma camada de saída pela qual são apresentados os resultados obtidos pela rede.

Na figura 1 pode-se visualizar a RNA progressiva, composta de três camadas: camada de entrada, camada intermediária (oculta) e camada de saída.

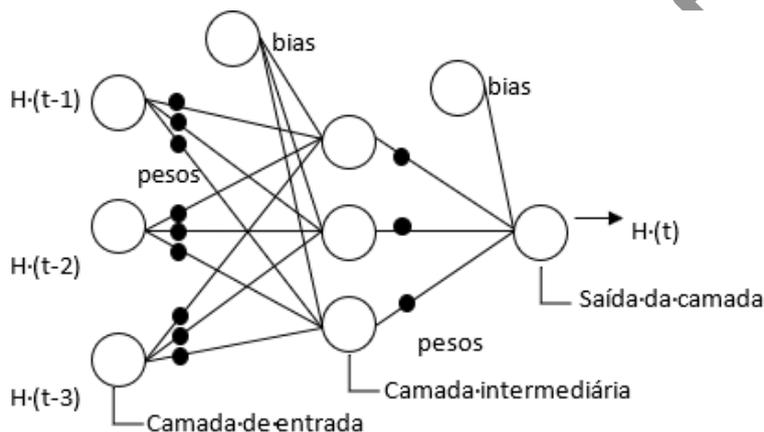


Figura 1 - Exemplo de uma RNA multicamadas de alimentação simples (Adaptado de Satarri et al., 2012).

Diferentemente da maioria dos modelos matemáticos, uma RNA não é “programada”; pelo contrário, é “ensinada” para dar uma resposta aceitável a um problema particular. Valores de entrada e de saída são enviados para a RNA, onde os pesos inicialmente atribuídos às ligações entre os neurônios são ajustados iterativamente até que a RNA calcule os valores de saída que corres-

pondam aos valores esperados (Haykin, 2013). O ajuste dos pesos é calculado a partir da função de ativação adotada para a RNA.

Um recurso que ajuda a compreender o mecanismo que ocorre dentro da caixa preta, representada pela RNA, é considerar a rede neural como uma super-forma de regressão múltipla (Haykin, 2013).

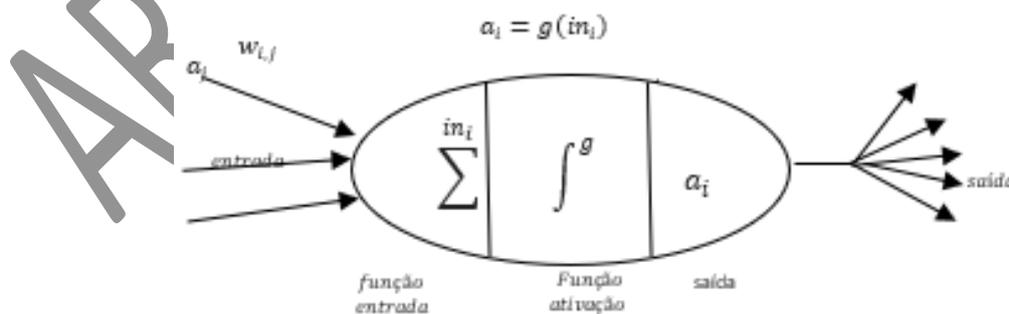


Figura 2 - Modelo de neurônio onde o valor a_j é igual ao valor de ativação da unidade j ; $w_{i,j}$ é igual ao peso na ligação da unidade j à unidade i ; in_i é igual à soma ponderada das entradas da unidade i ; a_i é igual ao valor de ativação da unidade i (conhecido como valor de saída); e g é igual à função de ativação.

A atividade das unidades de entrada representa as informações que são alimentadas para dentro da rede. A atividade de cada unidade escondida é determinada a partir das atividades das unidades de

entrada e dos pesos nas ligações entre as unidades de entrada e as unidades escondidas. O comportamento da saída depende da atividade das camadas escondidas e dos pesos entre as camadas

escondidas e a camada de saída (Haykin, 2013).

Treinamento da Rede Neural Artificial

Uma vez que uma RNA foi estruturada para uma determinada aplicação, ela está pronta para ser treinada. Para iniciar esse processo, os pesos iniciais são escolhidos aleatoriamente; em seguida inicia-se o treinamento ou aprendizagem.

Existem duas abordagens de treinamento principais: supervisionado e não supervisionado. O treinamento supervisionado exige o fornecimento da saída desejada para a rede, seja por modo manual de “classificação” do desempenho da rede, seja fornecendo os resultados desejados com as entradas. O treinamento sem supervisão é onde as entradas e as saídas da rede têm que fazer sentido sem ajuda externa (Haykin, 2013).

A complexidade computacional de uma rede e a sua capacidade de propagação (generalização) dependem diretamente da sua topologia. Assim, uma rede maior do que o necessário leva a um desempenho ruim da propagação. Em contrapartida, uma rede menor do que o necessário enfrenta dificuldade em aprender os dados de treinamento porque reduz a flexibilidade da rede.

Rede Multilayer Perceptron (MLP)

De acordo com Haykin (2013), a característica das RNAs do tipo MLP é possuir pelo menos uma camada intermediária ou escondida, apresentando um poder computacional muito maior do que as redes sem camadas escondidas, uma vez que possuem a capacidade de tratar problemas que não são linearmente separáveis, como é o caso da previsão de enchentes em rios.

A rede MLP convencional utiliza a arquitetura *feedforward* com múltiplas camadas, sendo o treinamento realizado de forma supervisionada ou não supervisionada. Neste projeto foi utilizada a forma supervisionada, na qual para cada amostra dos sinais de entrada também é disponibilizada a saída desejada. Dessa forma, cada amostra de treinamento é composta pelos sinais de entrada e suas respectivas saídas (Haykin, 2013).

A forma supervisionada é a mais utilizada em problemas de previsão hidrológica, inclusive para previsão de inundações. Assim, durante o treinamento da RNA, os sinais de entrada, geralmente nível do rio e/ou vazão e/ou precipitação, são disponibilizados juntamente com a saída.

O treinamento da MLP mais tradicionalmente utilizado é o algoritmo de retro propagação do erro ou *backpropagation*, conhecido também como regra delta generalizada. A utilização do algoritmo

backpropagation convencional, na prática, tende a convergir de forma lenta, exigindo, assim, um elevado esforço computacional para a etapa de treinamento. Algumas das alternativas para tornar o processo de convergência da rede mais eficiente são: a inserção do termo de *momentum* e a utilização de variantes do *backpropagation*, como o algoritmo de *Levenberg-Marquardt* e a Regularização Bayesiana (Haykin, 2013).

Análise de Componentes Principais (ACP)

As técnicas de análises multivariadas são ferramentas poderosas para analisar um número grande de variáveis. A ACP, permite reduzir a dimensão da matriz de observações sem perder as informações importantes dos dados originais, investigando o comportamento espaço-temporal das variáveis envolvidas no problema e detectando grupos de variáveis que apresentam comportamento homogêneo (Lima et al., 2010).

A ACP consiste essencialmente em reescrever as coordenadas das amostras em outro sistema de eixo mais conveniente para a análise. As n variáveis originais geram, através de suas combinações lineares, n componentes principais, cuja principal característica, além da ortogonalidade, é serem obtidos em ordem decrescente de máxima variância (Guedes et al., 2010).

O método da ACP é descrito por Johnson & Wichern (2007); através da Equação 1, realizou-se a padronização das variáveis meteorológicas recomendada por Amanajás & Braga (2012), ocorrendo a formação de uma matriz com observações versus variáveis padronizadas Z_{ij} .

$$Z_{ij} = \frac{X_{ij}}{S(X_j)} \quad (1)$$

onde X_{ij} é o valor da correlação entre o período i e a característica j ; $S(X_j)$ é o desvio padrão da característica j ; e i é o período observado. Para a criação da matriz de dados (tempo de observação x variáveis padronizadas), primeiramente obteve-se uma matriz de correlação, através da Equação 1, a partir da matriz de dados $X_{n \times p}$, onde n é o número de indivíduos de p características. A matriz das variáveis padronizadas Z_j é igual à matriz de correlação da matriz de dados X .

A determinação dos componentes principais foi calculada através da resolução da equação característica da matriz S (matriz de variância e covariância) ou R (matriz de correlação). A matriz de variância e covariância S é dada pela Equação 2.

$$S = \frac{1}{x} X_c^t \cdot X_c \quad (2)$$

onde X_c é a matriz com valores das observações e a sua transposta é dada por X_c^t .

Normalizando a matriz Z , a matriz variância e covariância será igual à matriz de correlação R (Guedes et al., 2010) dada pela Equação 3.

$$R = \frac{1}{n} Z \cdot Z^t \quad (3)$$

onde a matriz Z é a matriz normal de X , sendo a matriz de X composta por 'p' variáveis e 'n' indivíduos, e sua característica observada é x_1 a x_p e sua ordem é de $n \times p$:

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2p} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \dots & x_{np} \end{bmatrix}$$

Uma nova matriz Z e sua matriz transposta Z^t são obtidas, consequentemente determinando a matriz R sendo simétrica, positiva e diagonalizável.

Com isso, a matriz diagonal λ possui elementos diagonais que são os autovalores de R (Guedes et al., 2010) obtidos da Equação 4.

$$\lambda = U^{-1} R \cdot U \quad (4)$$

Após a obtenção de λ e U da matriz S ou R , o produto de cada autovetor pela matriz gera a componente principal. As componentes principais Y são obtidas por combinações lineares entre a transposta dos autovetores (U^t) e a matriz de observações X , conforme exposto na Equação (5).

$$Y = U^t \cdot X \quad (5)$$

Neste estudo, a Análise de Componentes Principais (ACP) e seus respectivos fatores de correlação foram calculados com o *Software* WEKA (Mark et al., 2013), utilizando o princípio

MATERIAL E MÉTODOS

Área de estudos

Este estudo foi realizado na bacia hidrográfica do rio Aquidauana, que se limita ao Sul e a Leste com a bacia do rio Paraná, ao Norte com a bacia do rio Negro e Taquari e a Oeste com a bacia do rio Miranda, do qual é tributário (Fernandes et al., 2012). O rio Aquidauana tem 286 km de extensão, sendo o principal afluente do rio Miranda e um dos principais cursos d'água da bacia do rio Paraguai.

O rio Aquidauana tem seu nível médio na cota

de redução de informações de um amplo conjunto de dados, o que tornou possível a sua análise de maneira simplificada, tendo como principal parâmetro a dispersão dos dados de uma média central. Essa análise, quando realizada em um grupo de variáveis sob situações distintas, funciona como uma ferramenta que efetua uma comparação estatística em ambos os grupos, sendo, aqui neste estudo, grupos distintos de dados pluviométricos e fluviométricos. Dessa maneira, a ACP foi utilizada para tornar mais convenientes os dados para a rede neural MLP.

Método de Thiessen

Para utilizar os dados dos postos pluviométricos, houve a necessidade de se determinar qual posto pluviométrico tinha maior ou menor influência sobre o nível do rio Aquidauana. Assim, utilizou-se o método Thiessen (empregado pelo meteorologista americano Alfred H. Thiessen) também relacionado com o diagrama de Voronoi.

Ainda segundo Breda (2008), o método de Thiessen se caracteriza pela atribuição de pesos relativos aos postos considerados mais representativos para uma área específica.

O método de Thiessen assume que, em qualquer ponto da bacia, o valor da precipitação é igual àquele medido no posto pluviométrico mais próximo. Assim sendo, o registro da altura em um dado posto é aplicado em outros pontos, desde que estes estejam até a metade da distância de um outro posto. A expressão matemática para o cálculo da precipitação média P sobre uma bacia, pelo método dos polígonos de Thiessen para uma dada área A , é dada pela Equação 6.

$$P = \frac{\sum_{i=1}^n A_i P_i}{A} \quad (6)$$

onde a variável P_i representa o i -ésimo posto com dados de precipitação em milímetros; A_i é a área de influência de cada posto, dada em km^2 ; e A é a área total.

de 3,0 m, e a cota de extravasamento do rio para a cidade está no nível 8,0 m. Para este trabalho houve a necessidade de realizar um levantamento no banco de dados da Agência Nacional das Águas (ANA), identificando apenas as cotas acima de 8,0 m, e relacionando-as à quantidade de dias de precipitação do mês correspondente. Assim, o ponto da cota foi da estação de medição número 66945000 (Aquidauana), localizada na ponte velha que liga Aquidauana e Anastácio. Para os postos pluviométricos utilizaram-se os

postos denominados pela ANA Palmeiras, Fazenda Lajeado, Rochedo, Santa Elisa e Bandeirantes. Na Figura 3 encontra-se o levantamento das cotas acima de 8,0 m e a quantidade de dias de chuva acumulada referente ao mês.

Nesse gráfico verifica-se que a maior cheia

registrada foi em março de 2011, com 10 metros, sendo que nesse mês foi registrada registrou uma média entre os postos selecionados de 2119,2 mm de chuva acumulada. Esse estudo serviu para verificar o quantitativo de chuva acumulada que extravasa a cota de 8,0 m, sendo essa informação utilizada no teste e treinamento da RNA.

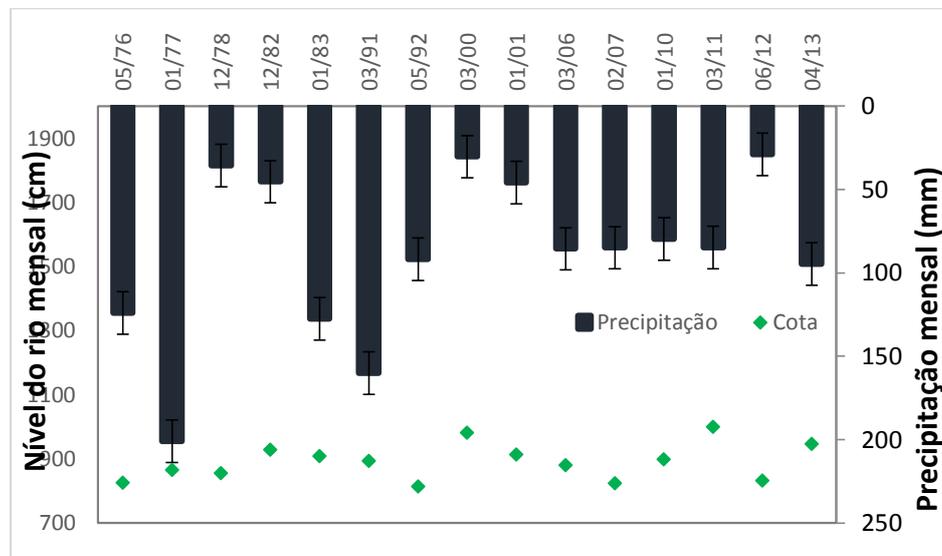


Figura 3 - Comparativo entre as maiores cheias e os dias de chuva acumulada referente ao pico de cheia. Em verde é disponibilizada a cota do rio Aquidauana no posto 66945000, representando picos acima de 8,0 m. Em azul é apresentada a chuva acumulada em mm no mês que apresentou cota acima de 8,0 m.

Dados disponíveis

Para realizar a aplicação da RNA, houve a coleta de dados fornecidos pelos postos fluviométricos e pluviométricos da ANA e da Companhia de Recursos Minerais (CPRM), situados na bacia do rio Aquidauana (ANA, 2015).

Na escolha dos postos utilizaram-se, como critério, os postos com série de dados superiores

a 15 anos. Nessas condições, a bacia do rio Aquidauana possui três postos fluviométricos em operação conforme a figura 4. Na figura 5 é apresentado o conjunto das informações levantadas pelas estações pluviométricas mais representativas que influenciam o nível do rio Aquidauana. Os dados levantados levaram em consideração somente o registrador de chuvas que estão em operação.

Código	Estação	Início da observação	Município de Mato Grosso do Sul	Altitude (m)	Coordenadas	
					Latitude	Longitude
66945000	Aquidauana	Jan/1968	Aquidauana	133	20°28'41''	55°48'05''
66941000	Palmeiras	Abr/1965	Anastácio	162	20°26'53''	55°25'41''
66926000	Ponte do Grego	Jan/1982	Terenos	203	20°09'21''	55°05'24''

Figura 4 - Postos fluviométricos da bacia do rio Aquidauana com mais de 15 anos de medição.

Código	Estação	Início da observação	Município de Mato Grosso do Sul	Altitude (m)	Coordenadas	
					Latitude	Longitude
01954002	Rochedo	Nov/1967	Rochedo	447	19°57'09	54°53'34
01954005	Bandeirantes	Jun/1969	Bandeirantes	456	19°55'04	54°21'31
02055002	Palmeiras	Nov/1971	Aquidauana	175	20°26'56	55°25'51
02055003	Fazenda Lajeado	Jan/1953	Dois Irmãos do Buriti	104	20°17'32	55°26'43
02054009	Santa Elisa	Jan/1978	Terenos	411	20°29'26	54°52'18

Figura 5 - Postos pluviométricos da bacia do rio Aquidauana com mais de 15 anos de medição.

Para a análise dos postos fluviométricos e pluviométricos, verificaram-se, primeiramente, as falhas de leitura de cotas e de precipitação em

todo o período existente no banco de dados da ANA entre jan/1953 e fev/2014, para serem usadas no treinamento e testes das RNAs. Depois

de verificadas as falhas, foi extraído do banco de dados a relação das cotas e as chuvas acumuladas efetivamente medidas em ordem cronológica, utilizando somente os dados consistidos. Para a avaliação inicial dos dados e de sua relação com o alagamento na subbacia do rio Aquidauana,

optou-se por trabalhar apenas a bacia que está dentro dos limites do município. Devido à grande quantidade de falhas de leitura dos dados pluviométricos anterior a 1995, optou-se por realizar treinamento com a RNA de 1995 a 2013 e teste para 2014.

RESULTADOS E DISCUSSÕES

Resultados

Neste trabalho foi criado um *software*, programado na linguagem *Java*, para tratar os dados da ANA; e para teste e treinamento dos dados utilizou-se o *software* WEKA (Mark et al., 2013). A Rede Neural Artificial é testada e treinada no WEKA. Dessa forma, a nomenclatura adotada para a análise de componentes principais, funções de ativação e algoritmos de treinamento foi a mesma usada pelo *software*.

A primeira tarefa, após obtenção dos dados, foi criar um *software* que buscasse os dados na planilha da ANA e colocasse no formato aceito pelo WEKA, que é um arquivo com extensão ‘arff’. Assim, na figura 6 pode-se visualizar como ficaram dispostos os dados para entradas no ambiente WEKA, tendo como entrada a quantidade de dados pluviométricos e fluviométricos antecedentes e, na saída, a cota do rio. Dependendo dessas variáveis, pode-se aumentar ou diminuir a quantidade de colunas.

Data	Nível do Aquidauana	Bandeirantes	Lajeado	Palmeiras	Rochedo	Santa Elisa	Outras Cotas	Aquidauana
01/01/1945	t-1	t-1; t-2; t-3; ...	t-1; t-2; t-3; ...	t-1; t-2; t-3; ...	t-1; t-2; t-3; ...	t-1; t-2; t-3; ...	t-1; t-2; t-3; ...	t+x
02/01/1945	t-1	t-1; t-2; t-3; ...	t-1; t-2; t-3; ...	t-1; t-2; t-3; ...	t-1; t-2; t-3; ...	t-1; t-2; t-3; ...	t-1; t-2; t-3; ...	t+x

Figura 6 - Dados de entrada e saída da Rede Artificial Neural.

Na figura 6, na primeira linha, tem-se um exemplo com data $t = 01/01/1945$; na segunda coluna o valor é igual ao do nível do rio na data $t-1$. As próximas colunas apresentam os valores dos índices medidos pelo pluviômetro de Bandeirante nas datas $t-1$; $t-2$; $t-3$; ... (número máximo de dias anteriores utilizados). O mesmo é feito para todas as outras estações pluviométricas, como feito para previsão com um dia de antecedência.

Após a disposição dos dados em uma planilha, com o acoplamento temporal entre a variável de saída e as entradas adequadas, a primeira ação de pré-processamento foi a de remoção dos espaços em branco nas amostras, pois em alguns períodos não existem medições. Essas lacunas podem ser

de dias ou até mesmo semanas. Essa ação é indispensável, pois, caso não fossem removidas essas lacunas, seriam causadas distorções durante o ajuste dos pesos dos neurônios no treinamento das RNAs, implicando previsões com maiores erros. Após essa etapa, a base de dados contava com 5.368 amostras.

Na Tabela 1 pode-se verificar o peso de influência de cada estação pluviométrica, calculado pelo método de Thiessen. Para esse cálculo foi utilizado o *software* Quantum Gis 2.6 (Sherman, 2015).

A bacia do rio Aquidauana, figura 7, mostra as estações ativas monitoradas pela Agência Nacional das Águas (ANA), com estações pluviométricas e fluviométricas, (Ana, 2016).

Tabela 1 - Peso de cada estação usando o método de Thiessen.

Código	Localidade	Peso de cada localidade
1954005	Bandeirantes	0,1274
2055003	Fazenda Lajeado	0,1660
2055002	Palmeiras	0,1518
1954002	Rochedo	0,3588
2054009	Santa Elisa	0,1960
Total		1,0000

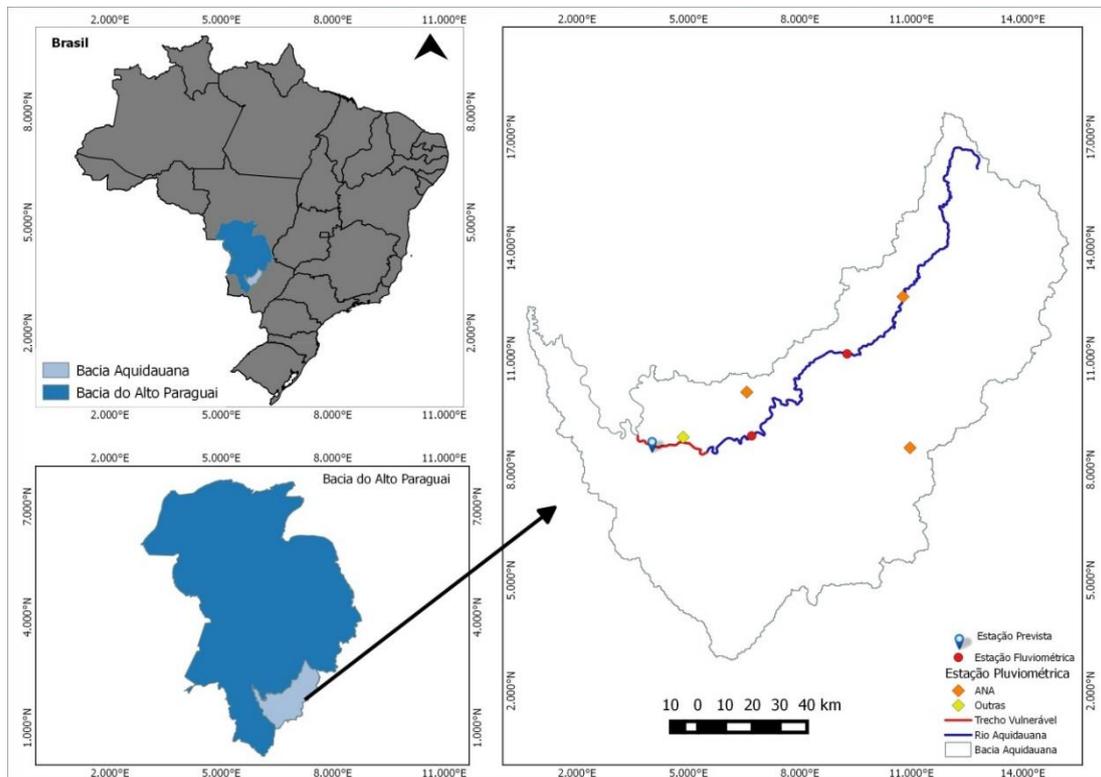


Figura 7 - Mapa da bacia hidrográfica e das estações fluviométricas (vermelho) e pluviométricas (laranja e amarelo) da bacia utilizada.

RNAs propostas

As RNAs construídas neste trabalho são do tipo MLP e possuem oito entradas, uma camada oculta com n neurônios e um neurônio na camada de saída. A figura 8 exibe como ficou a configuração das entradas e saída da RNA. Os dados disponíveis na região selecionada são das estações de medição localizadas na sub-bacia do rio Aquidauana, com cinco postos

pluviométricos e três postos fluviométricos e, como saída, o posto fluviométrico de Aquidauana. O desempenho do modelo da RNA variou com a estrutura e cenários utilizados. Portanto, foi realizada a previsão de um até cinco dias de antecedência, sendo que, para cada dia de antecedência, foram utilizados os dias de chuvas acumuladas e as cotas anteriores, variando entre 10 a 50 dias.

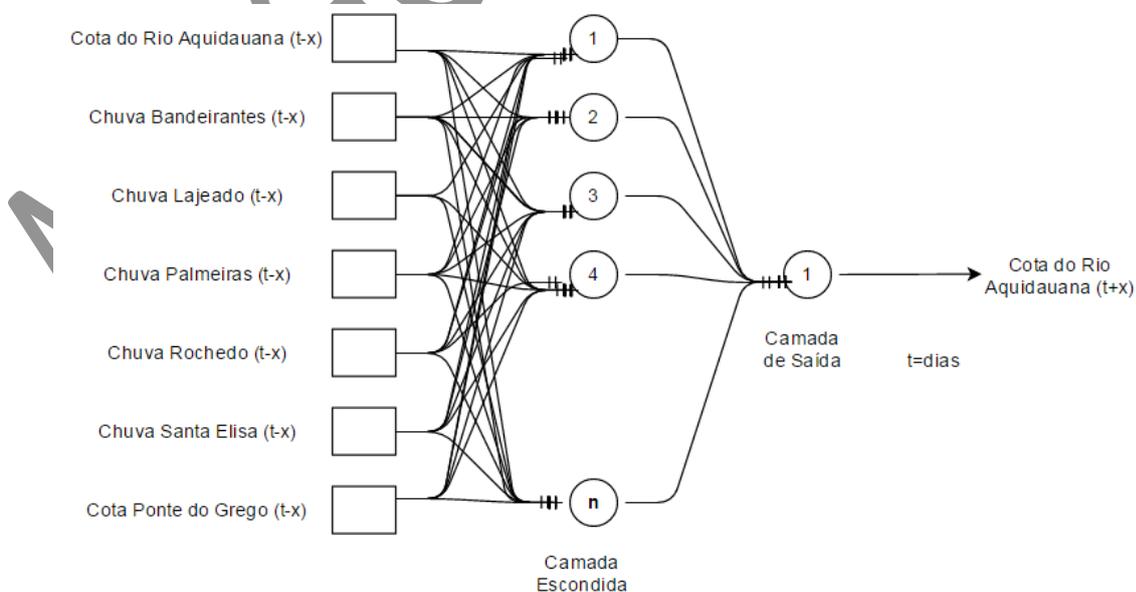


Figura 8 - Modelo de RNA para previsão de alagamento para a sub-bacia do rio Aquidauana.

As RNAs utilizadas nesta pesquisa foram construídas, treinadas e testadas utilizando-se as Redes Neurais MLP do *software* WEKA. Assim,

o parâmetro em todos os modelos, como o Learning rate (-L), Momentum (-M), Decay (-D), HiddenLayers (-H), Seed (-S), and Reset (-R)

foram calibrados para cada previsão, assim como os parâmetros das componentes principais. Sua configuração encontra-se na figura 9.

Na tabela 2 apresenta as medidas de erro para os cinco modelos, sendo que a variável t representa o tempo de previsão de 1 a 5 dias, sendo a janela de treinamento (1995-2013) e teste (2014), o local previsto é dentro da cidade de Aquidauana para a estação 66945000

(Aquidauana). Os erros são analisados por 7 critérios de avaliação: R_{tr} - coeficiente de correlação para dados de treinamento; R_{tst} - coeficiente de correlação para dados de teste; MAE - Erro Absoluto Médio; RAE - Erro Relativo Absoluto; RMSE - Erro Quadrático Médio Relativo; MSRE - Erro Quadrático Relativo; NSE - Coeficiente de Eficiência Nash-Sutcliffe.

Parâmetros da MLP						
-L	0.10972987204251648	-M	0.43416608789994227	-H	i	-R -D -S 1
-L	0.6798397856131904	-M	0.10554916235673212	-H	a	-D -S 1
-L	0.105958700926245	-M	0.3574674403845104	-H	a	-R -D -S 1
-L	0.5423057915249041	-M	0.2462631974094704	-H	o	-D -S 1
-L	0.2317503826097306	-M	0.8688734180371707	-H	t	-D -S 1

Figura 9 - Parâmetros utilizado no Software WEKA.

Tabela 2 - Análise comparativa dos desempenhos.

t	R_{tr}	R_{tst}	RMSE (cm)	MAE (cm)	RAE(cm)	NSE
1	0,95	0,93	30	0,71	0,6328	0,93
2	0,89	0,87	45	0,66	0,4356	0,87
3	0,79	0,76	57	0,60	0,3796	0,76
4	0,69	0,67	65	0,57	0,339	0,67
5	0,56	0,54	69	0,63	0,301	0,54

A tabela 2 indica os parâmetros estatísticos analisados neste processo e se observam melhores resultados no processo de teste em relação ao treinamento. Nota-se, que os índices de erro associado às frequências seguiu a tendência de piora em função do aumento do horizonte de previsão para os 5 dias, além de apresentar valores aceitáveis para uma previsão de nível, onde o erro variou de cerca de 30 cm para o horizonte de 1 dia e 69 cm para o horizonte de 5 dias. Estes índices, erros associados à frequência, possibilitam um entendimento mais intuitivo e realista da incerteza do modelo de previsão, constituindo uma margem de garantia para uso do tomador de decisão em um sistema de alerta de inundações. Os dados de valores diários de NSE tendem a ser superiores aos dados de valores mensais, o que contradiz as descobertas de alguns estudos individuais (Van Liew et al., 2007). Esta anomalia é

potencialmente devido ao aumento tamanhos de amostra para dados diários. Como esperado, NSE e R os valores da previsão foram melhores para os períodos de calibração do que os períodos de validação.

Para a bacia uma previsão ótima seria de 3 a 5 dias, pois um sistema de alerta precisa de tempo para ser acionado e tomado as providências em relação a inundação. Neste estudo para esta bacia os resultados são satisfatórios a previsão de até 5 dias, observando os erros medidos pela RNA (Moriassi et al., 2007).

A Figura 10 mostra as estimativas do nível do rio fornecidas pelo modelo de predição da BP prevendo 1 a 5 dias (1D, 2D, 3D, 4D e 5D), sendo as nomenclaturas descritas no canto inferior esquerdo da figura 3. Pode-se observar hidrogramas e gráficos de dispersão, com os valores previstos próximos dos observados, confirmando as estatísticas RMSE na Tabela 1.

CONCLUSÕES

O método de previsão de alagamento em curto prazo utilizando Redes Neurais Artificiais é vantajoso, pois, com poucas variáveis, consegue-se realizar uma saída para RNA, ao passo que outros modelos requerem uma série de variáveis para produzir previsões precisas.

Para reduzir a quantidade de informações do conjunto de dados pluviométricos e fluviométricos, simplificando a análise, foi utilizada a técnica de análise de componentes principais, tornando mais aderente a utilização dos dados na RNA (MLP).

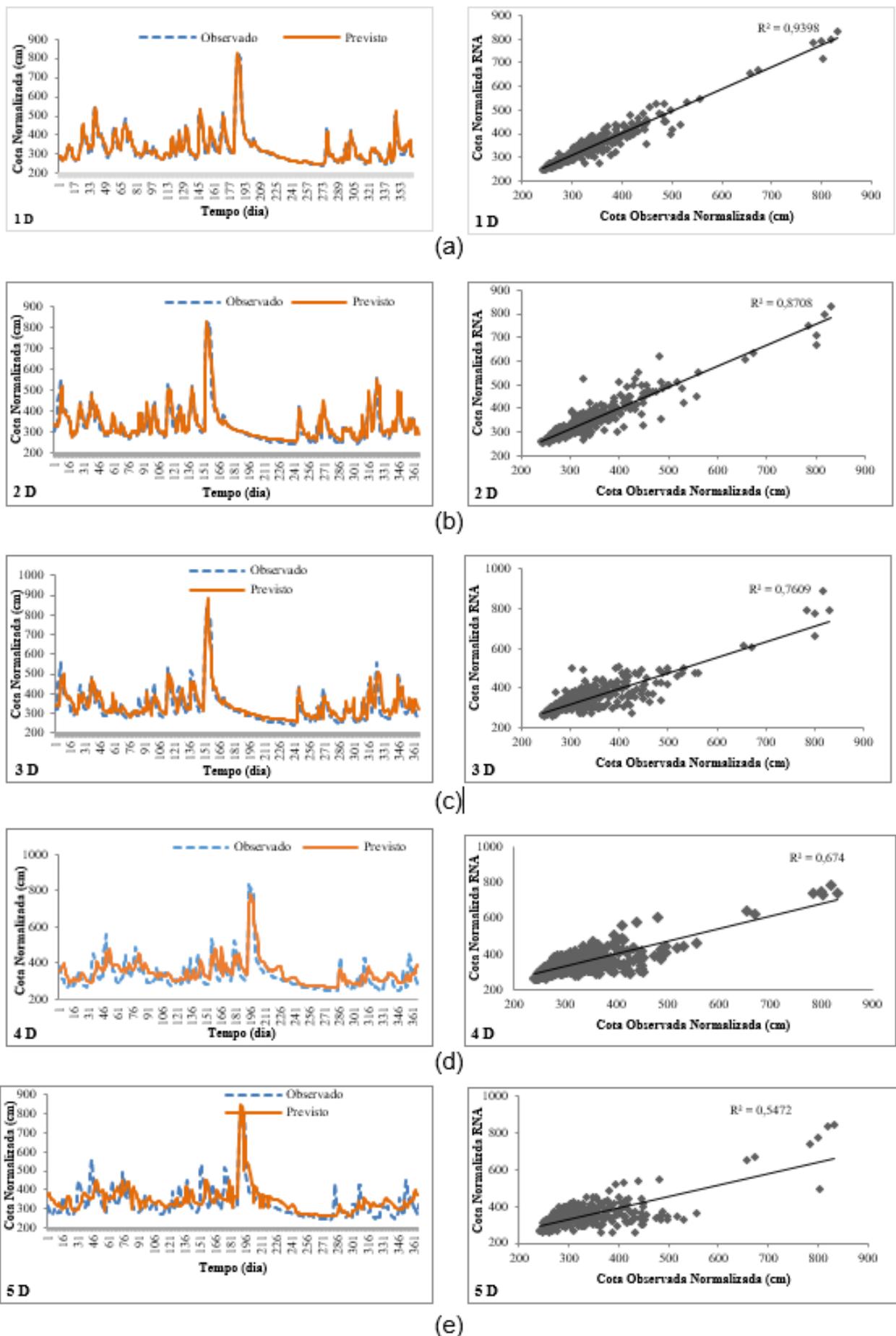


Figura 10 - Evolução temporal do nível previsto para 2014 (laranja), observado (azul) e respectivos diagramas (r^2). Os dados correspondentes ao período de 1995 a 2014, obtidos das previsões 1 dia (a), 2 dias (b), 3 dias (c), 4 dias (d), 5 dias (e).

Foi realizada a previsão de um a cinco dias de antecedência, utilizando-se dados de chuvas acumuladas e cotas com 10 até 50 dias anteriores. Dentre todas as configurações, o melhor desempenho foi o da rede MLP com a configuração -L 0.10972987204251648 -M 0.43416608789994227 -H i -R -D -S 1e os parâmetros de componentes principais -A 1 -C -R 0.9485751970524982, com um dia de antecedência e 40 dias de chuvas acumuladas e cotas anteriores. Essa rede produziu as previsões de melhor qualidade dentre todas as MLPs com coeficiente de correlação (r) e erro quadrático médio de 0,95 e 30 cm, respectivamente.

Para se chegar ao melhor modelo, o desempenho de cada cenário foi verificado, e os parâmetros foram configurados para cada

previsão desejada, para evitar qualquer redundância. RNAs têm a vantagem da simplicidade, quando comparadas com outros modelos mais sofisticados. Por isso, em situações onde a informação é inexistente ou difícil de se obter, a RNA é um método que fornece a opção mais viável para a previsão de enchentes. As RNAs oferecem um meio de reduzir os custos analíticos da informação topográfica e hidrológica através da redução da quantidade de tempo gasto com a análise dos dados.

Após esses resultados, será realizado um estudo de previsão para a bacia do rio Miranda utilizando a técnica de RNAs; assim, essas duas bacias farão parte do modelo de previsão de enchentes para o Pantanal, uma região afetada pelos rios Aquidauana e Miranda.

AGRADECIMENTOS

Este é um projeto de pesquisa apoiado pela FUNDECT (Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul) e pela Empresa Brasileira de Pesquisa Agropecuária – EMBRAPA Pantanal.

REFERÊNCIAS

- AMANAJÁS, J.C. & BRAGA, C.C. **Padrões espaço-temporal pluviométricos na Amazônia Oriental utilizando análise multivariada**. Revista Brasileira de Meteorologia. v. 27, n. 4. São Paulo. 2012.
- ANA. Agência Nacional da Água. Disp. em: <http://www.ana.gov.br/>. Acesso em: 20 de jun 2015.
- BREDA, A. **Avaliação de melhorias para um sistema de previsão hidrológica horária**. Curitiba. 2008. 159 p. Dissertação (Mestrado em Engenharia de Recursos Hídricos e Ambiental), Universidade Federal do Paraná.
- FAZEL, S.A.A.; BLUMENSTEIN, M.; MIRFENDERESK, H.; TOMLINSON, R. Estuarine flood modelling using artificial neural networks. In: Neural Networks (IJCNN), INTERNATIONAL JOINT CONFERENCE, p. 631-637, 2014.
- FERNANDES, E.F.L. & ANUNCIACÃO, V.S. da. Impactos de eventos climáticos extremos e seus reflexos na cidade de Aquidauana-MS. Manaus/AM, **Revista Geonorte**, v. 1, n. 4, edição especial. p. 707 – 720, 2012.
- GUEDES, R. V.; LIMA, F.L.; AMANAJÁS, J.C.; BRAGA, C.C. Análise em componentes principais da precipitação pluvial no estado do Piauí e agrupamento pelo método de Ward. **Revista de Geografia**, Recife, v. 1, n. 27, p. 218-233, 2010.
- HAYKIN, S. **Adaptive Filter Theory**. 5. ed. Pearson Education: Prentice Hall. 368 p., 2013.
- JOHNSON, R.A & WICHERN, D. W. **Applied Multivariate Statistical Analysis**. 6 ed. Prentice Hall. 328 p., 2007.
- JOIA, P.R. & ANUNCIACÃO, V.S. Inundações urbanas e vulnerabilidade socioespacial na cidade de Aquidauana. Londrina/PR: **Geografia**, v. 22, n. 2, p. 5-23, 2013.
- LIMA, F.J.L.; AMANAJÁS, J.C.; GUEDES, R.V.S.; LIMA, E.M.S. Análise de componente principal e de agrupamento para estudo de ventos para a geração de energia eólica na região do Ceará, Paraíba, Pernambuco e Rio Grande do Norte, Brasil. **Ambiente & Água – An Interdisciplinary Journal of Applied Science**, v. 5, n. 2, 2010.
- MARK, H.; RANK, E.F.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I.H. The WEKA Data Mining Software: An Update. **SIGKDD Explorations**. v. 11. Issue 1. 2013.
- OZGUR, K. A combined generalized regression neural network wavelet model for monthly stream flow prediction. **KSCE Journal of Civil Engineering**, v. 15.p. 1469–1479, 2011
- PADOVANI, R.C. **Dinâmica espaço-temporal das inundações do Pantanal**. São Paulo. 2010. 175 p. Tese (Doutorado em Ecologia Aplicada) – Escola Superior de Agricultura “Luís de Queiroz”. Centro de Energia Nuclear na Agricultura, Universidade de São Paulo.
- ROY, S.; PALLAVI, M.; SWAPTIK, C.; NIHHIL, B.; ROSHAN S. Dual Artificial Neural Network for Rainfall-Runoff Forecasting. **Journal of Water Resource and Protection**. v. 4, p. 1024-1028, 2012.
- RUSLAN, F.A.; SAMAD, A.M.; ZAIN, Z.M.; ADNAN, R. Flood water level modeling and prediction using NARX neural network: Case study at Kelang river. In: 10TH INTERNATIONAL COLLOQUIUM. **Signal Processing e its Applications (CSPA)**, 2014, p. 204-207.
- SATARRI, M. T.; YUREKLI, K.; PAL, M. Performance Evaluation of Artificial Network Approaches in forecasting reservoir inflow. **Applied Mathematical Modelling**. v. 36, p. 2649-2657, 2012.
- SHERMAN, E.G. **Software Quantum Gis**. v. 2.6. GNU General Public License. 2015.
- THIRUMALAI, K. & DEO, M. Hydrological Forecasting Using Neural Networks. **Journal of Hydrologic Engineering**. v. 5, p. 180-189, 2000.
- TUCCI, C.E.M. **Modelos Hidrológicos**. 2 ed. Porto Alegre: Ed. da Universidade/UFRGS/ABR, 2005.
- TUCCI, C.E.M. & COLLISHONN, W. Coupled Hydrologic-Hydraulic Modeling of the Upper Paraguay River Basin. **Journal of Hydrologic Engineering**. v. 17, p. 635-646, 2012.
- VAN LIEW, M.W.; VEITH, T.L.; BOSCH, D. D.; ARNOLD, J.G. Suitability of SWAT for the conservation effects assessment project: A comparison on USDA-ARS experimental watersheds. **Journal of Hydrologic Engineering**, Nova Iorque, v. 12, n. 2, p. 173-189, 2007.

Recebido em 20 de janeiro de 2020
Aceito em 17 de abril de 2020