

ANÁLISE DE MODELOS DE REDES NEURAIS APLICADOS AO PROCESSO CHUVA-DEFLÚVIO NO SEMI-ÁRIDO

Carlos Henrique Ribeiro Lima¹ & Walter Martins Ferreira Filho²

Resumo - No presente trabalho é analisada a performance de diversos modelos de Redes Neurais Artificiais (RNA) aplicados ao processo chuva-deflúvio. As diversas configurações de RNA's são obtidas com a variação do número de nós na camada de entrada, na camada intermediária e pela variação do número de anos usado para o treinamento e para a validação da rede. Os testes dos modelos foram realizados com dados de precipitação e deflúvio no posto da Fazenda Cajazeiras, no Estado do Ceará. O deflúvio mensal é previsto a partir de dados de precipitação dos meses anteriores e corrente e em alguns modelos utilizando também o deflúvio de meses anteriores. Os resultados mostram a grande influência que tem a configuração da rede neural e o número de anos de treinamento e de validação nos resultados obtidos para previsão do deflúvio mensal na bacia hidrográfica.

Abstract - In this work is analyzed the performance of several models of artificial neural networks (ANN's) applied to the process rainfall-runoff. The diverser configurations of ANN's are obtained with the diversous numbers of nodes in the input layer, in the hidden layer and by the variation of the number of years used for the training and the validation of the algorithm. The model's tests were realized with the measurements from precipitation and flow at the Fazenda Cajazeiras, in the State of Ceará. The mensal runoff is predicted with the data from precipitation of the past and present months and in some models utilizing also the runoff from the past months. The results shown the large influence in the results of the forecast of the mean flow from the watershed by the configuration and the number of years of training and validation of the neural network.

Palavras-chave - Redes Neurais Artificiais, Modelos Chuva-Deflúvio, Previsão de Deflúvio.

¹ Universidade Federal do Ceará - Departamento de Engenharia Hidráulica e Ambiental - Campus do PICI - Bloco 713, altos, Fortaleza-CE. Tel.: (0xx85) 288-9623 e (0xx85) 257-2194 - Residência . E-mail : beberibe@mail.com

² Universidade Federal do Ceará - Departamento de Engenharia Hidráulica e Ambiental - Campus do PICI - Bloco 713, altos, Fortaleza-CE. Tel.: (0xx85) 288-9623 . E-mail : walter@ufc.br

INTRODUÇÃO

Um modelo que relaciona a precipitação numa dada bacia hidrográfica e a consequente vazão gerada no exutório da bacia, doravante denominado modelo chuva-deflúvio (ou chuva-vazão), é de extrema importância no processo de avaliação de projetos de recursos hídricos, pois, na maioria dos casos, o tamanho das séries históricas de vazão é relativamente curto, enquanto as séries históricas de medição de precipitação são de tamanho razoavelmente bons. Verifica-se também a ausência de medições de vazão (e em alguns casos de precipitação) em certos meses das séries históricas, também chamadas de falhas. Esses problemas são bastante verificados na região Nordeste do Brasil. Além de tudo isso, mesmo quando se dispões de dados razoáveis de precipitação, o conjunto de pares de medição precipitação-vazão é ainda relativamente curto para parametrização de modelos clássicos chuva-deflúvio. Assim, modelos chuva-deflúvio que produzem resultados acurados para séries históricas curtas são de grande valia e utilidade em hidrologia.

Acredita-se que o processo chuva-deflúvio é extremamente não-linear, variável no tempo, espacialmente distribuído, e não facilmente descrito por modelos simples. Assim, duas aproximações tem sido bastante exploradas na literatura de forma a modelar o processo: A modelagem "Conceitual" (Física) e a modelagem do tipo "Caixa-Preta" (*Black Box*).

Modelos conceituais são designados para aproximar dentro das estruturas chuva-deflúvio os sub-processos gerais internos e os mecanismos físicos que governam o ciclo hidrológico. Normalmente tais modelos incorporam formas simplificadas de leis físicas, e são geralmente, não-lineares, invariantes no tempo e determinísticos, com parâmetros que são características representativas da bacia hidrográfica. São referidos como sendo confiáveis na previsão dos pontos mais importantes do hidrograma, como o início da curva de ascensão, o tempo e a altura da vazão de pico, e o volume do escoamento. Contudo, a implementação e a calibração de tais modelos apresentam normalmente várias dificuldades, requerendo sofisticadas ferramentas matemáticas, quantidades significativas de dados de calibração, além de um certo grau de esperteza e experiência com o modelo.

Em situações onde as dificuldades dos modelos conceituais se agravam, e o interesse não está em entender o processo hidrológico, mas numa aplicação prática, como previsão de vazões, em que o objetivo principal é com predições acuradas em locais específicos de bacias hidrográficas, os hidrologistas e demais profissionais da área normalmente preferem utilizar modelos do tipo caixa-preta. Em tais modelos, diferentes equações ou modelos de equações diferenciais são usados para identificar um mapeamento direto entre entrada e saída, sem considerações detalhadas da estrutura interna do processo físico.

Nos modelos caixa-preta tradicionais, devido a uma enorme fragilidade na identificação da estrutura não-linear do processo e na estimação de parâmetros, pouco se tem desenvolvido nos

últimos anos. Por outro lado, a aproximação sistemática, em geral, e a análise funcional não-linear, em particular, são tidos como percusores de um novo tipo de renascença, em virtude das novas ferramentas de trabalho desenvolvidas. Como exemplo, tem-se as redes neurais artificiais (RNA) e os algoritmos genéticos. Entretanto, esses modelos não substituem de forma alguma os modelos físicos conceituais do processo chuva-deflúvio, sendo alternativas viáveis em aplicações hidrológicas onde a previsão acurada de vazão é necessária e se dispõe apenas de dados de entrada e saída para as séries temporais e pouco conhecimento conceitual relativo da dinâmica hidrológica da bacia hidrográfica sobre investigação.

Uma RNA é uma estrutura matemática não-linear, capaz de representar processos complexos arbitrários e não-lineares, que relacionam a entrada e a saída de qualquer sistema. O sucesso com que RNA's tem sido usadas para sistemas dinâmicos em outros campos da Ciência e Engenharia, sugere que a aproximação por RNA é um efetivo e eficiente caminho para modelar o processo chuva-deflúvio em situações onde o conhecimento explícito dos sub-processos hidrológicos internos não é necessário. Aplicações de RNA são vistas também em previsões climáticas, problemas de água subterrânea e previsão de temperaturas média do ar.

Este artigo consiste numa aplicação prática de redes neurais artificiais, onde são previstos os deflúvios mensais no posto da Fazenda Cajazeiras, no Estado do Ceará. O presente rio é caracterizado como intermitente e apresenta um alto coeficiente de variação, sendo ideal para a aplicação e análise de modelos. A série histórica analisada contém dados da precipitação mensal na bacia e da vazão mensal no exutório, compreendidos entre o período de Janeiro de 1963 e Dezembro de 2000. O parâmetro de entrada dos modelos consiste em dados de precipitação e deflúvio. Os dados de precipitação correspondem à precipitação no mês analisado e em meses anteriores. O deflúvio corresponde aos meses anteriores ao analisado. A influência na performance dos modelos é investigada pela variação de três parâmetros: o número de anos usado na calibração do modelo (ou fase de treinamento); o número e o tipo dos parâmetros usados na entrada do modelo e a variação do número de nós na camada intermediária, que representa um parâmetro interno do modelo.

REDES NEURAIAS ARTIFICIAIS (RNA's)

A noção de redes neurais, a partir de meados da década de 40, foi motivada pela velocidade e eficiência com que o cérebro humano processa informações. Matematicamente, redes neurais são sistemas de processamento de informações para modelar o cérebro como computação paralela, segundo Freitas (1999) e Minns e Hall (1996).

Uma rede neural artificial (RNA) é composta por um grande número de unidades interconectadas de processamento (nós), arranjadas numa camada de entrada (*Input Layer*), numa camada de saída (*Output Layer* ou *Grossberg Layer*), e em uma ou mais camadas ocultas (*Hidden Layer*). Cada camada consiste de vários nós. Dois tipos de redes neurais são comumente usados em problemas de mapeamento de padrões: a Rede Perceptron Acíclica (ou sem realimentação) de Multicamadas (*Feed-Foward Multilayer Perceptron Network - MPL*) e a Rede de Contra Propagação (*Counter Propagation Network - CPN*).

A Figura 1, baseada no esquema apresentado por Smith e Eli (1995), mostra a estrutura de uma rede neural acíclica de 3 camadas, utilizada no presente trabalho. Neste tipo de rede a informação flui em uma direção, de modo que, ao ser processada por um elemento processador, não retorne a ele.

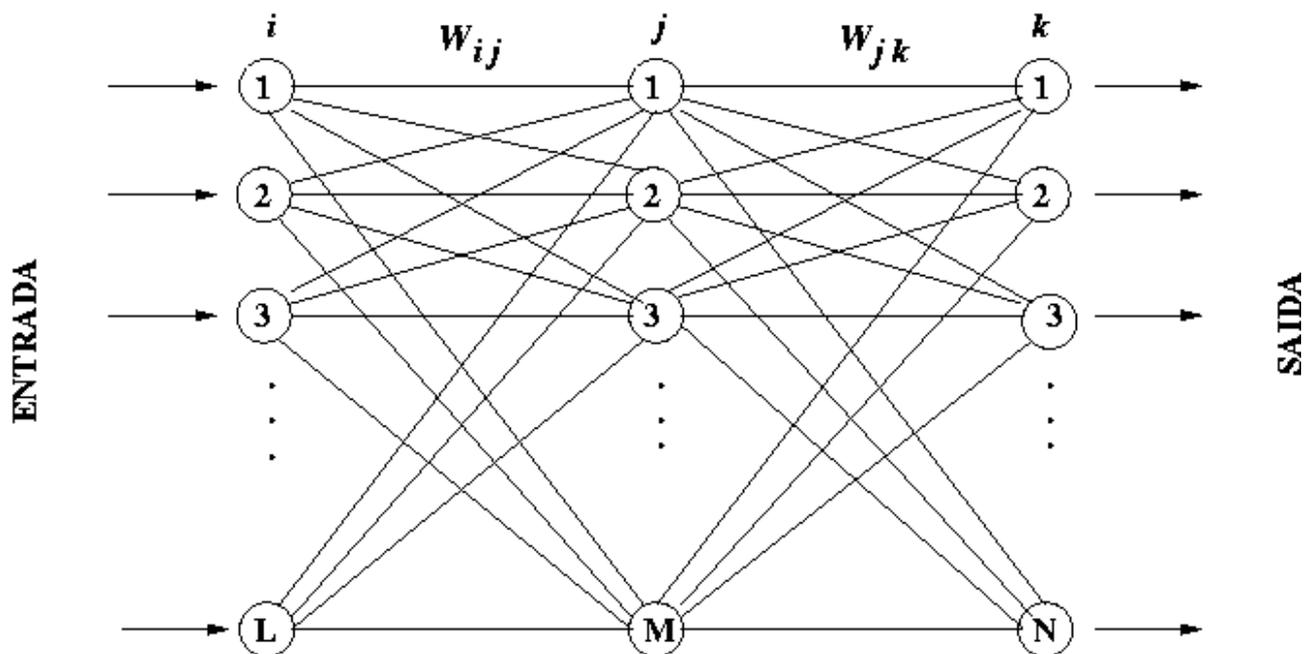


Figura 1: Configuração de uma Rede Neural de 3 camadas

Um sinal de entrada, consistindo de um vetor de números, X_i , é introduzido nos nós da camada de entrada, como mostrado na Figura 1. Os sinais são carregados ao longo das conexões para cada nó da camada seguinte, e podem ser amplificados ou inibidos por meio de pesos W_{ij} , associados a cada conexão. Os nós dessa camada adjacente (camada intermediária) agem como um somatório dos sinais de chegada ponderados pelo peso. O sinal de chegada, S_j , em um dado nó j , pode ser expresso pela equação:

$$S_j = \sum_{i=0}^L W_{ij} X_i, \quad (1)$$

onde L é o número de nós na camada de entrada, W_{0j} o peso " bias ", e $X_0 = 1$ o " bias ", que neste caso não pode ser considerado como sinal de entrada. A Figura 2 representa os sinais de entrada em um nó j .

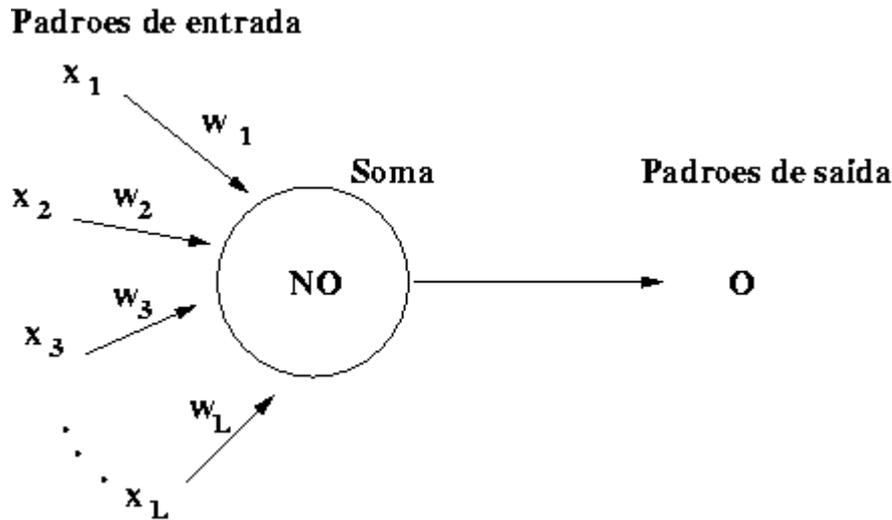


Figura 2: Entrada e saída em um dado nó

O sinal S_j é então passado através de uma função de transferência para produzir o sinal de saída O_j do nó j . A função de transferência utilizada neste trabalho é a função Sigmoid logística, dada por:

$$O_j = \frac{1}{1 + e^{(-S_j)}} \quad (2)$$

onde S_j pode apresentar qualquer valor, mas O_j é limitado no intervalo $[0,1]$.

A função de transferência Sigmoid é escolhida por conveniência matemática, pois possui limites inferiores e superiores, é monoticamente crescente e contínua e diferenciável em todo ponto.

O sinal de saída é subsequentemente carregado ao longo das conexões ponderadas até a camada seguinte de nós, e o processo é repetido até que o sinal atinga a camada de saída. O sinal de saída pode ser interpretado como a resposta da RNA à um dado estímulo de entrada. No caso de se utilizar 3 camadas, o sinal de chegada S_k para um dado nó k na camada de saída, é dado por:

$$S_k = \sum_{i=0}^M W_{ik} O_i, \quad (3)$$

onde M é o número de nós na camada intermediária, W_{0k} o peso " bias ", e $O_0 = 1$ o " bias ".

O sinal de saída final será então:

$$O_k = \frac{1}{1 + e^{(-S_k)}}. \quad (4)$$

Note que no presente estudo tem-se apenas um nó na camada de saída, que representa o deflúvio previsto para um dado mês t , portanto $k = I$.

Matemáticos tem mostrado que uma RNA de multicamadas acíclica tem a incrível capacidade de serem "aproximadores universais de funções". Hecht-Nielsen (1987), provou que uma RNA de três camadas contém todos os requisitos para ser um mapeador de funções universais, e qualquer função multivariada pode ser aproximada por uma RNA que tenha somente um número finito de nós na camada intermediária. Esse resultado é chamado de teorema de Kolmogorov da existência de redes neurais mapeadoras.

Estudos recentes também mostraram que uma RNA de três camadas acíclica, que faz uso da função de transferência Sigmoid, pode implementar qualquer mapeamento de funções multivariadas contínuas e delimitadas. Esses resultados demonstram que uma RNA de 3 camadas pode ser usada para modelar relações funcionais reais desconhecidas ou pouco definidas, e bastante complexas, como é o caso do processo chuva-deflúvio.

Fase de Treinamento

Assim como ocorre na modelagem conceitual de processos reais, a modelagem por meio de redes neurais também requer uma fase de calibração (ou parametrização). Este procedimento é conhecido como "treinamento da rede neural" (*Network Training*), onde os parâmetros do modelo são representados pelos pesos W .

A RNA pode ser treinada para produzir respostas de saída conhecidas ou esperadas para um dado estímulo de entrada. Primeiramente são atribuídos valores randômicos para os pesos W_{ij} e W_{jk} interconectados. Em seguida é introduzido um sinal padrão p na entrada, e o valor de saída O_{pk} é calculado. Esse sinal de saída é comparado com o valor de saída T_{pk} esperado. O erro total E_p , baseado na diferença quadrática entre a saída prevista e a esperada para o padrão p , é calculado como:

$$E_p = \sum_{k=1}^N (T_{pk} - O_{pk})^2, \quad (5)$$

onde $N=I$ para o modelo aqui utilizado.

O objetivo do processo de treinamento é apresentar um número P suficiente de pares padrões únicos de entrada e saída, que quando acoplados com uma metodologia que corrige a cada iteração os pesos interconectados, produza um conjunto final de pesos que minimize o erro global, E , definido como:

$$E = \sum_{p=1}^P E_p. \quad (6)$$

No presente artigo, P representa o número total de meses usados no treinamento.

Na prática, qualquer um de vários critérios pode ser adotado para finalizar o processo de treinamento. Uma tolerância pode ser selecionada com base na magnitude do erro global, E , ou nos erros padrões individuais, E_p , ou ainda o processo pode ser finalizado após um certo número de iterações. Independente do critério de parada, precauções devem ser tomadas para não "supertreinar" a RNA, podendo resultar numa redução de sua habilidade de generalizar. O "supertreinamento" pode ser comparado com memorização rotativa, que resulta numa ótima performance com pares padrões usados no treinamento, mas numa performance significativamente baixa com padrões que não são apresentados no treinamento.

Neste trabalho o ajuste dos pesos interconectados durante a fase de treinamento emprega um método bastante conhecido e utilizado, chamado de Erro Retro-propagado (*Error Back-propagation*), primeiramente formulado por Rumelhart *et al.* (1986). Após a atribuição de valores aleatórios para os pesos, o padrão apresentado na entrada é propagado até a camada de saída (*Forward-propagation*), onde é calculado o erro entre a saída real e a saída desejada de cada elemento processador, dado pela Equação (5). Este erro é então propagado de volta (*Backward Propagation*) por meio dos respectivos pesos das conexões, até atingir a camada de entrada. Matematicamente, cada peso, W_{yz} , é ajustado pela fórmula:

$$W_{yz} + \Delta p W_{yz}, \quad (7)$$

onde $\Delta p W_{yz}$ é a mudança nos pesos interconectados entre as camadas arbitrárias y e z para o padrão p , e pode ser expressado por:

$$\Delta p W_{yz} = \eta \delta_{pz} O_{py}, \quad (8)$$

onde η é a razão de aprendizado, δ_{pz} o termo do erro associado ao neurônio z para o padrão p , e O_{py} a saída do neurônio conectado a camada anterior y para o padrão p . O valor absoluto dos pesos

individuais não pode ser interpretado como tendo algum significado físico profundo (Minns (1995)).

O termo δ_{pk} é calculado diferentemente, dependendo se o mesmo está associado com um neurônio na camada de saída k , ou com um neurônio na camada intermediária j . O resultado na forma final para cada termo do erro para a rede neural de três camadas mostrada na Figura 1, é dado por:

$$\delta_{pk} = (1 - O_{pk}) O_{pk} (T_{pk} - O_{pk}) \quad (9)$$

para neurônios na camada de saída k , e

$$\delta_{pj} = (1 - O_{pj}) O_{pj} \sum_k \delta_{pk} W_{jk} \quad (10)$$

para os neurônios na camada intermediária j .

O próximo par padrão entrada-saída é apresentado, e os pesos das conexões são reajustados para minimizar o novo erro. Nesse sentido, o algoritmo de retro-propagação pode ser visto como sendo da forma gradiente descendente (*gradient descent*), pois encontra o mínimo valor da função erro multi-dimensional. Este procedimento é repetido até que todos os dados de treinamento tenham sido aplicados. O processo global é então reiniciado, começando mais uma vez pelo primeiro par entrada-saída, e continuado até que o erro para todos conjuntos de dados seja suficientemente pequeno e ajustes subsequentes nos pesos não influenciem mais neste erro.

A razão η de aprendizado deve ser ajustada, inicialmente, para pequenos valores, com o objetivo de evitar oscilações na correção dos pesos durante o treinamento. No modelo aqui apresentado, é usado uma razão η variável, conforme descrito em Math Works (2001). Inicia-se com uma razão η baixa, e, a partir da segunda iteração, desde que o novo erro calculado seja menor que o erro da iteração anterior, a razão é aumentada por um fator 1,05. Caso o erro seja superior ao anterior (normalmente por um fator de 1,04), os novos pesos e "bias" são descartados, e a razão η é diminuída de um fator 0,7.

Para melhorar a velocidade e o desempenho do processo de treinamento, foi usado aqui o algoritmo de treinamento de Levenberg-Marquardt, apresentado em Math Works (2001). O algoritmo consiste numa técnica numérica de otimização em que a matriz Hessiana não é computada no cálculo da velocidade de treinamento de segunda ordem. Quando a função objetivo tem a forma de soma de raízes quadradas (como no treinamento de redes neurais acíclicas), a matriz Hessiana pode ser aproximada por:

$$H = J^T J, \quad (11)$$

e o gradiente pode ser calculado como:

$$g = J^T e, \quad (12)$$

onde J é a matriz Jacobiana que contém as primeiras derivadas do erro da rede neural com relação aos pesos e "bias", e e é o vetor erro da rede. A matriz Jacobiana pode ser calculada por um técnica padrão de retro-propagação que é bem menos complexa que o cálculo da matriz Hessiana.

O número de pares padrão entrada-saída requeridos na fase de treinamento é altamente dependente do sistema a ser modelado. Widrow (1989), recomenda um conjunto de dados de tamanho igual a $10 \times L$ (Dez vezes a dimensão do padrão de entrada) como ponto inicial. Aqui foram testados seis diferentes números para o conjunto padrão entrada-saída, correspondendo a 5, 10, 15, 20, 25 e 30 anos de treinamento.

Após toda essa fase, a RNA é dita como tendo "aprendido" a relação entre o conjunto dos dados de entrada-saída, estando pronta para ser validada, onde novos valores de entrada são utilizados, e os resultados da saída são comparados com os desejados, sem que haja modificações nos pesos.

Considerações sobre a normalização da série histórica

Pelas próprias características da função de transferência Sigmoid, os dados de entrada e saída devem ser normalizados entre o intervalo $[0,1]$. Entretanto, na maioria dos métodos de normalização, como a divisão dos valores pelo máximo observado no grupo, duas dificuldades principais são bastante encontradas:

i) A convergência é lenta, e o processo de aprendizado é ineficiente se os dados de saída conterem valores próximos de zero ou um, conforme reportaram Ooyen e Nichhuis (1992).

ii) Quando não se conhece o processo sobre investigação, o melhor estimador do comportamento do mesmo é a magnitude observada, o que restringe o máximo valor que possa ser previsto a ser o máximo valor presente no conjunto de treinamento, ou seja, extrapolações não são permitidas. Entretanto, apesar de não ser formidável, extrapolações na vizinhança dos limites muitas vezes são essenciais.

Dessa forma, adota-se aqui a normalização sugerida por Sajikumar e Thandaveswara (1999), onde, além dos valores na vizinhança $[0,1]$ serem evitados, extrapolações num certo limite são permitidas.

Os valores de entrada e esperados são normalizados num intervalo $FMIN$ e $FMAX$, ao invés de zero e um ($FMIN > 0$ e $FMAX < 1$), tal que:

$$X_n = FMIN + \frac{(X_u - fact_min)}{(fact_max - fact_min)} (FMAX - FMIN), \quad (13)$$

onde X_u e X_n representam a variável a ser normalizada e o seu valor normalizado, respectivamente, $fact_max$ e $fact_min$ são os valores máximo e mínimo presentes no vetor X .

Os valores adotados como $FMIN$ e $FMAX$ devem ser tomados de tal forma que não reduzam o intervalo a um valor pequeno que possa influenciar o treinamento, ou ainda permitam um grande intervalo de interpolação. No presente artigo foram selecionados os valores $FMIN = 0,1$ e $FMAX = 0,9$, como indicados por Hsu *et al.* (1995).

SÉRIE HISTÓRICA UTILIZADA

Os dados de vazão mensal utilizados foram coletados do posto Fazenda Cajazeiras, localizado na Bacia do Rio Acaraú, no Semi-Árido Cearense. Os dados de precipitação mensal foram coletados de algumas estações pluviométricas da referida Bacia, e o método de Thiessen foi usado para o cálculo da precipitação média mensal na Bacia. Não foram observadas falhas na séries históricas de precipitação e vazão.

O período utilizado foi de janeiro de 1963 a dezembro de 2000, usando a nomenclatura mês 1 para janeiro de 1963 e assim sucessivamente, finalizando no mês 456, que equivale a dezembro de 2000.

O cálculo do deflúvio mensal na Bacia foi efetuado a partir da área da Bacia ($A = 1489 \text{ km}^2$) e do número de dias do mês calculado.

CRITÉRIOS DE PERFORMANCE

Um bom modelo deve conter três características principais: acurácia, consistência e versatilidade (Kachroo (1992)). O termo acurácia refere-se a capacidade do modelo em ter erros pequenos no período de calibração. A consistência de um modelo é verificada quando o nível de acurácia e de estimação dos valores dos parâmetros persistem para diferentes amostras dos dados históricos. A versatilidade do modelo é definida quando o mesmo é acurado e consistente em aplicações que não sejam diretamente baseadas na função objetivo usada na calibração do modelo.

Neste trabalho foram analisadas as características acurácia e consistência na avaliação dos diversos modelos de redes neurais. Três critérios foram utilizados:

i) Volume acumulado no período.

ii) Erro médio quadrático *RMSE* (*Root Mean Square Error*), definido como:

$$RMSE = \sqrt{\frac{\sum_{k=1}^K (Q_k - \hat{Q}_k)^2}{K}}, \quad (14)$$

onde K é o tamanho da amostra, Q_k e \hat{Q}_k são, respectivamente, os deflúvios observados e calculados para o k -ésimo intervalo.

iii) Eficiência R^2 , expressa por:

$$R^2 = \frac{F_0 - F}{F_0}, \quad (15)$$

onde

$$F_0 = \sum_{k=1}^K (Q_k - \bar{Q})^2, \quad (16)$$

$$F = \sum_{k=1}^K (Q_k - \hat{Q}_k)^2, \quad (17)$$

sendo \bar{Q} o valor médio do deflúvio para o período analisado.

É importante ressaltar que o *RMSE* indica a performance relativa de diferentes modelos para um mesmo período de calibração e verificação, não podendo de fato ser utilizado como critério de comparação para modelos com diferentes períodos de calibração e verificação. Nesta situação, o critério R^2 de eficiência é melhor indicado. Entretanto, são apresentados os valores *RMSE* pois os mesmos indicam erros dimensionais. O valor ideal para o erro *RMSE* é zero, e para R^2 deve ser um.

RESULTADOS E DISCUSSÃO

Foram realizadas mais de 600 simulações, de modo a combinar e melhor avaliar os modelos de redes neurais para três diferentes configurações:

i) número e parâmetro dos nós na camada de entrada;

ii) número de nós na camada intermediária;

iii) período de calibração e verificação.

Pode-se assumir que o deflúvio D numa bacia hidrográfica para um dado mês t , é função da precipitação P no mês t e de precipitações de meses anteriores ($t-1$, $t-2$, ...), e também do deflúvio de meses anteriores. Assim, a camada de entrada pode ser composta por diversas combinações de modo a verificar a influência de cada um desses termos na previsão do deflúvio.

Neste trabalho, para um dado período de treinamento e verificação, e um dado parâmetro na camada de entrada, o número de nós na camada intermediária recebeu os valores 3, 5, 8 e 10. Observou-se nitidamente que a influência dos mesmos nos resultados era muito baixa, não chegando a ser significativa. Assim, os resultados mostrados na coluna 2 da Tabela 3 indicam o número de nós (3, 5, 8 ou 10) na camada intermediária em que foi observado o melhor resultado. Não são mostrados os resultados para os outros três valores de nós na camada intermediária, pois iria demandar bastante espaço, sem acréscimo de informações significativas. Claramente o número de nós na camada intermediária para o período de treinamento é o mesmo para o período de validação.

Foram testados seis períodos diferentes de treinamento, e, conseqüentemente, seis períodos diferentes de validação. Optou-se nesta primeira fase da pesquisa por escolher períodos sequenciais. A Tabela 1 mostra os seis períodos de treinamento e validação, assim como as estatísticas dos períodos (média \bar{Q} e desvio padrão S) e o deflúvio acumulado (DA) em cada período. Observa-se nesta tabela o alto coeficiente de variação. Para todos os períodos, seu valor ficou acima de 2, que é comum em rios intermitentes do Semi-Árido Nordeste, onde são observados valores de vazão zero para um longo período sazonal, e no mesmo ano apresentam picos relativamente altos de vazão. Na validação equivalente ao período 01/68 a 12/00, por exemplo, o coeficiente de variação ficou em torno de 3,58, valor relativamente alto.

Na Tabela 2 é apresentado o número do modelo e a correspondente característica de sua camada de entrada. Essa numeração será utilizada na tabela 3. O intuito foi uma melhor visualização dos resultados. Na indicação do modelo mostrada na coluna 1 da Tabela 2, tem-se antes da seta \rightarrow a indicação do número de nós na camada de entrada e a sua característica, sendo P a letra representativa para precipitação e D a letra representativa para deflúvio, e o sub-escrito refere-se ao período do mês t e à períodos anteriores. Após a seta, o símbolo mostrado é referência para a saída da rede neural, ou seja, o deflúvio estimado para o período t .

Tabela 1: Dados referentes aos períodos de treinamento e verificação

Calibração					Verificação				
Período	Meses	\bar{Q} [mm]	S^a	DA^b	Período	Meses	\bar{Q} [mm]	S	DA
01/63 a 12/67	01-60	22,9	51,6	1375	01/68 a 12/00	61-456	10,1	36,2	4017
01/63 a 12/72	01-120	14,1	38,6	1698	01/73 a 12/00	121 - 456	11,0	38,8	3695
01/63 a 12/77	01-180	16,2	47,4	2910	01/78 a 12/00	181 - 456	9,0	31,6	2483
01/63 a 12/82	01-240	13,1	42,1	3144	01/83 a 12/00	241 - 456	10,4	34,7	2249
01/63 a 12/87	01-300	14,9	45,6	4482	01/88 a 12/00	301 - 456	5,8	18,5	911
01/63 a 12/92	01-360	13,8	42,8	4966	01/93 a 12/00	361 - 456	4,4	13,5	427

^a Desvio Padrão no dados observados [mm]

^b Deflúvio observado acumulado no período [mm]

Tabela 2: Número e significados dos modelos testados

Modelo	Característica da entrada → Saída
1	$P_{t-1}^a, P_t \rightarrow D_t^b$
2	$P_{t-1}, P_t, D_{t-1} \rightarrow D_t$
3	$P_{t-2}, P_{t-1}, P_t \rightarrow D_t$
4	$P_{t-2}, P_{t-1}, P_t, D_{t-1} \rightarrow D_t$
5	$P_{t-1}, P_t, D_{t-2}, D_{t-1} \rightarrow D_t$
6	$P_{t-2}, P_{t-1}, P_t, D_{t-2}, D_{t-1} \rightarrow D_t$

^a Precipitação

^b Deflúvio

Os resultados apresentados na Tabela 3 referem-se ao desempenho de cada modelo, sendo apresentado seu número na coluna 1, de acordo com a nomenclatura mostrada na Tabela 2. Na coluna 2 é mostrado o número de nós (3, 5, 8 ou 10) na camada intermediária em que foram obtidos os melhores resultados. As colunas subsequentes referem-se ao período utilizado, ao erro *RMSE* [mm] e ao erro R^2 , correspondendo as fases de treinamento e validação.

Três análises distintas podem ser feitas a partir da Tabela 3:

- i) para um dado período de treinamento e validação, avalia-se o melhor modelo;
- ii) para um dado modelo, seu desempenho é avaliado para os diferentes períodos de treinamento e validação, ou
- iii) faz-se uma avaliação geral do melhor resultado, encontrando-se o modelo e os períodos de treinamento e validação correspondentes.

Tabela 3: Resultados obtidos para os diversos modelos testados

Modelo	N _h ^a	Treinamento			Validação		
		Período	RMSE	R ²	Período	RMSE	R ²
1	5	01-60	186	0,90	61-456	337	0,39
2	5		176	0,91		179	0,82
3	10		157	0,93		334	0,40
4	8		162	0,93		198	0,79
5	3		185	0,91		173	0,84
6	5		180	0,91		259	0,64
1	3	01-120	185	0,83	121-465	321	0,52
2	8		185	0,84		191	0,83
3	5		185	0,83		247	0,71
4	8		163	0,87		192	0,83
5	5		178	0,85		208	0,80
6	3		179	0,85		232	0,75
1	5	01-180	189	0,88	181-456	195	0,73
2	8		172	0,90		131	0,88
3	8		180	0,90		157	0,83
4	3		171	0,90		136	0,87
5	5		159	0,92		128	0,88
6	5		185	0,89		143	0,86
1	8	01-240	186	0,86	241-456	200	0,77
2	3		175	0,88		133	0,89
3	8		163	0,89		175	0,82
4	10		167	0,89		164	0,84
5	3		181	0,87		144	0,88
6	3		184	0,86		139	0,88
1	8	01-300	186	0,88	301-456	193	0,24
2	10		153	0,92		116	0,72
3	10		175	0,89		162	0,46
4	8		181	0,89		141	0,59
5	5		179	0,81		142	0,59
6	5		179	0,89		145	0,57
1	8	01-360	183	0,87	361-456	214	-0,75
2	5		162	0,90		180	-0,24
3	10		168	0,89		189	-0,37
4	5		148	0,91		163	-0,02
5	5		179	0,88		167	-0,02
6	3		185	0,87		187	-0,34

^a Número de nós na camada intermediária

Pelo primeiro método de análise, para um dado período de treinamento e validação, que corresponde a cada intervalo de linhas numeradas de 1 a 6, analisamos o desempenho de cada modelo com base no valor RMSE e também pelo coeficiente R^2 . Independente do modelo, podemos observar que nos períodos de treinamento 1-60, 1-300 e 1-360, correspondentes respectivamente aos períodos de validação 61-456, 301-456 e 361-456, o desempenho geral na fase de validação foi ruim, com valores baixos para R^2 . Isso deve-se ao aprendizado da rede, que no

primeiro período, não consegue realmente "aprender" a relação entrada-saída, ou seja, assimilar os padrões estabelecidos (*underfitting*), enquanto nos dois últimos períodos a rede neural memoriza os padrões de treinamento (*overfitting*), gravando todas as suas peculiaridades e ruídos, o que leva a uma redução na capacidade de generalização, conforme afirma Furtado (1998). Nos outros três períodos, observa-se bons resultados dos modelos 1 e 3 na fase de treinamento, mas resultados relativamente ruins na fase de validação. Esses dois modelos representam a camada de entrada somente com dados de precipitação. Os outros 4 modelos apresentaram resultados sem diferenças significativas para todos os períodos de treinamento e validação. Apenas uma discreta superioridade pode ser observada no modelo 2 para a fase de validação em todos três períodos analisados. Com relação ao número de nós na camada intermediária, não foi observada nenhuma relação funcional entre esse número e as peculiaridades da camada de entrada.

No segundo tipo de análise, não são considerados os períodos problemáticos, mencionados no parágrafo anterior. Observa-se que todos os modelos, para a fase de treinamento, tem seu melhores resultados no período 01-180 seguido do período 01-240, enquanto para a fase de validação alguns modelos tiveram os melhores resultados no período 181-456 e outros no período 241-456.

O terceiro tipo de análise é feita a partir da Tabela 4, onde são mostrados os melhores resultados estatísticos obtidos por cada modelo, e o respectivo deflúvio acumulado no período DA, e seu erro percentual EDA quando comparado ao observado, mostrado na Tabela 1.

Tabela 4: Melhores resultados obtidos com os modelos

Modelo	N _h ^a	Treinamento					Validação				
		Período	RMSE	R ²	DA ^b	EDA ^c	Período	RMSE	R ²	DA	EDA
1	5	01-240	186	0,86	3142	0,1	241-456	200	0,77	2533	12,6
2	8	01-180	172	0,90	3687	26,7	181-456	131	0,88	3490	40,6
3	8	01-180	180	0,90	2661	8,6	181-456	157	0,83	2659	7,1
4	3	01-180	171	0,90	2767	4,9	181-456	136	0,87	2693	8,5
5	5	01-180	159	0,92	3205	10,1	181-456	128	0,88	3023	21,7
6	3	01-240	184	0,86	3246	3,2	241-456	139	0,88	2612	16,1

^a Número de nós na camada intermediária

^b Deflúvio observado acumulado no período [mm]

^c Erro percentual com relação deflúvio observado acumulado no período

Com relação à fase de treinamento, o modelo que apresentou o melhor resultado para o coeficiente R² foi o modelo 5. Em contrapartida, o volume acumulado é melhor previsto pelo modelo 1. Na fase de validação, os melhores resultados do coeficiente R² são obtidos com os modelos 2, 4, 5 e 6, sem nenhuma diferença significativa entre eles. Coincidentemente, são os modelos que apresentam pelo menos um termo de deflúvio na camada de entrada. O deflúvio acumulado é melhor previsto pelo modelo 3, que contém somente 3 termos de precipitação na camada de entrada, com um

erro da ordem de 7 %. Os modelos 2, 5 e 6, que apresentam pelo menos um termo de deflúvio na camada de entrada, apresentaram um erro relativo alto no volume acumulado no período.

Com o objetivo de melhor visualizar os resultados obtidos, são apresentados nos gráficos das figuras 3 a 6 o valor do deflúvio observado e o deflúvio previsto para os modelos 3 e 4 nas fases de treinamento e validação. Esses modelos foram considerados aqui, levando em conta o conjunto acurácia e consistência, os que melhor estimam o deflúvio mensal na bacia considerada.

Na fase de treinamento, os valores de picos são previstos pelos dois modelos, com erros em relação ao valor observado muito baixos. O modelo 4 apresenta uma melhor previsão. Para deflúvios zero ou próximos de zero, todos os dois modelos estimam com ótima acurácia.

A fase de validação também é marcada com boa previsão para picos de deflúvios, assim como para valores baixos. Novamente, o modelo 4 é o que melhor prevê o deflúvio mensal.

CONCLUSÕES

Foram realizadas simulações com o objetivo de testar diferentes modelos de redes neurais aplicados ao processo chuva-deflúvio. Três parâmetros foram tomados como definidores dos modelos testados: número e características dos nós na camada de entrada, número de nós na camada intermediária e número de meses usado na fase de treinamento.

Observou-se que o número de nós na camada intermediária tem pouca influência nos resultados, sendo o número ideal a ser utilizado determinado somente pelos resultados das simulações numéricas, não existindo, a priori, um método que determine esse número ideal.

A característica dos nós de entrada influencia significativamente nos resultados. As melhores estatísticas na previsão do deflúvio no mês t foram obtidas com dados na entrada de precipitação, correspondentes aos períodos t , $t-1$ e $t-2$, em conjunto com o deflúvio no mês $t-1$. Entretanto, o volume acumulado no período foi melhor previsto com os mesmos dados de precipitação mas a ausência do deflúvio no mês $t-1$. Isso indica que o acréscimo de deflúvio do período $t-2$ nos dados de entrada não altera os resultados estatísticos para a Bacia aqui estudada, e o volume acumulado é melhor previsto quando se utiliza somente dados de precipitação no vetor de entrada.

Finalmente, fica evidente que a maior influência no desempenho dos modelos está relacionada ao período de treinamento e validação utilizado. Períodos muito curtos, tendem a prejudicar o aprendizado da rede, enquanto períodos longos de treinamento e curtos de validação podem supertreinar a rede. Em ambos os casos, a capacidade de generalização da rede é drasticamente reduzida. O período ideal de treinamento encontrado está basicamente na metade da séries histórica, sendo a outra parte destinada à validação. Espera-se em trabalhos futuros aprofundar o estudo aqui realizado, desenvolvendo novos modelos de redes neurais com mais de 3 camadas, além de testar modelos chuva-deflúvio diários.

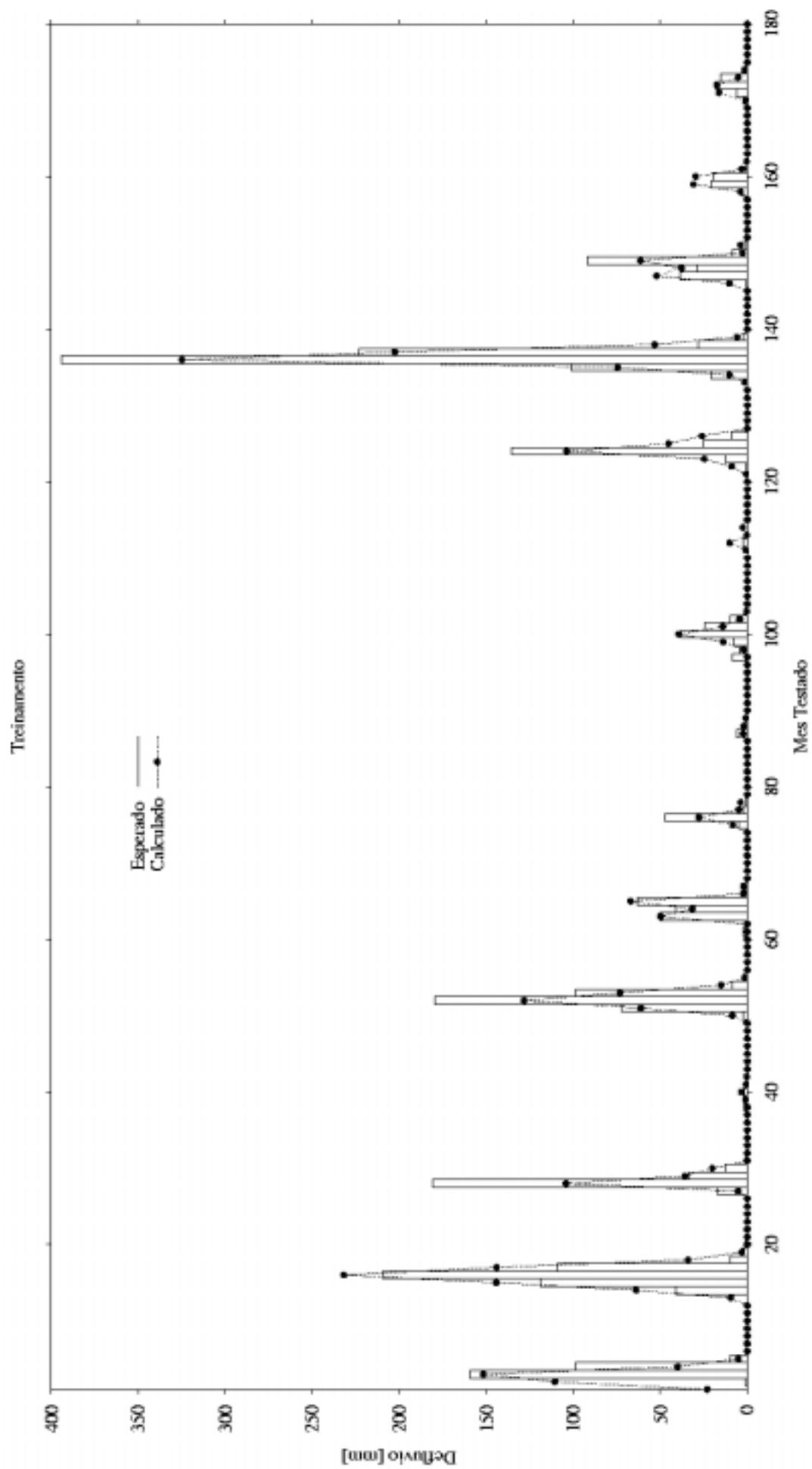


Figura 3: Resultados para a fase de treinamento do modelo 3

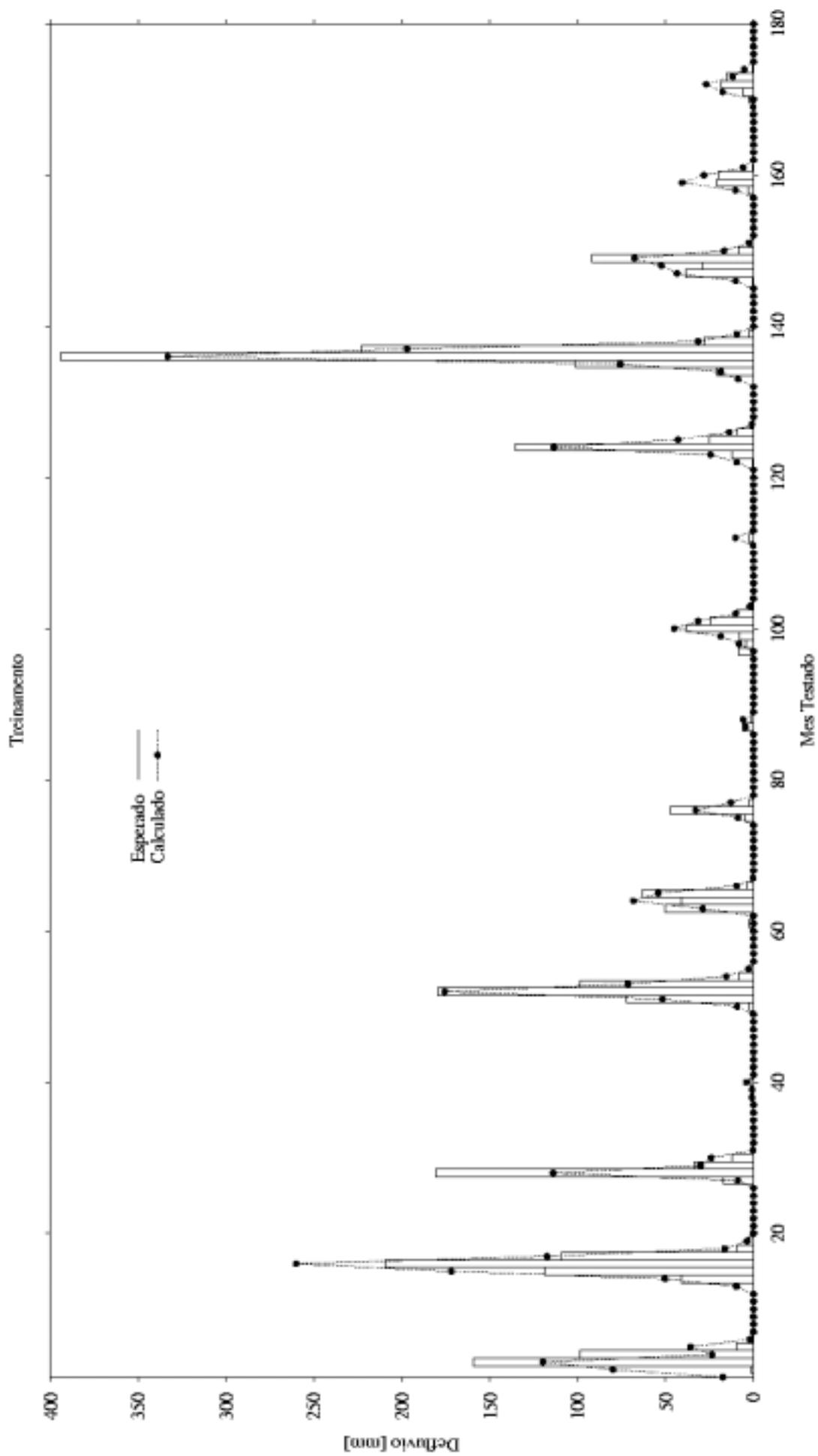


Figura 4: Resultados para a fase de treinamento do modelo 4

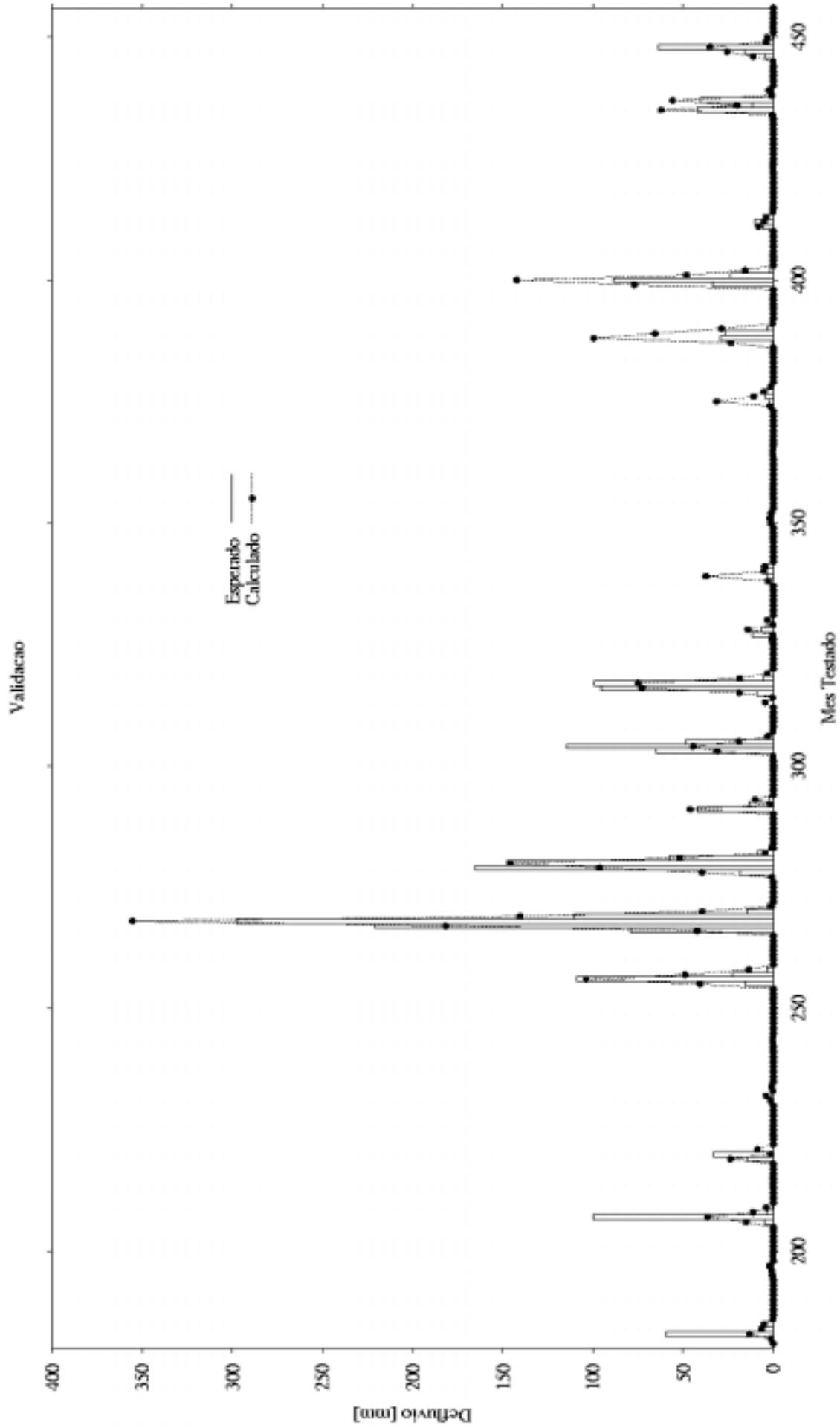


Figura 5: Resultados para a fase de validação do modelo 3

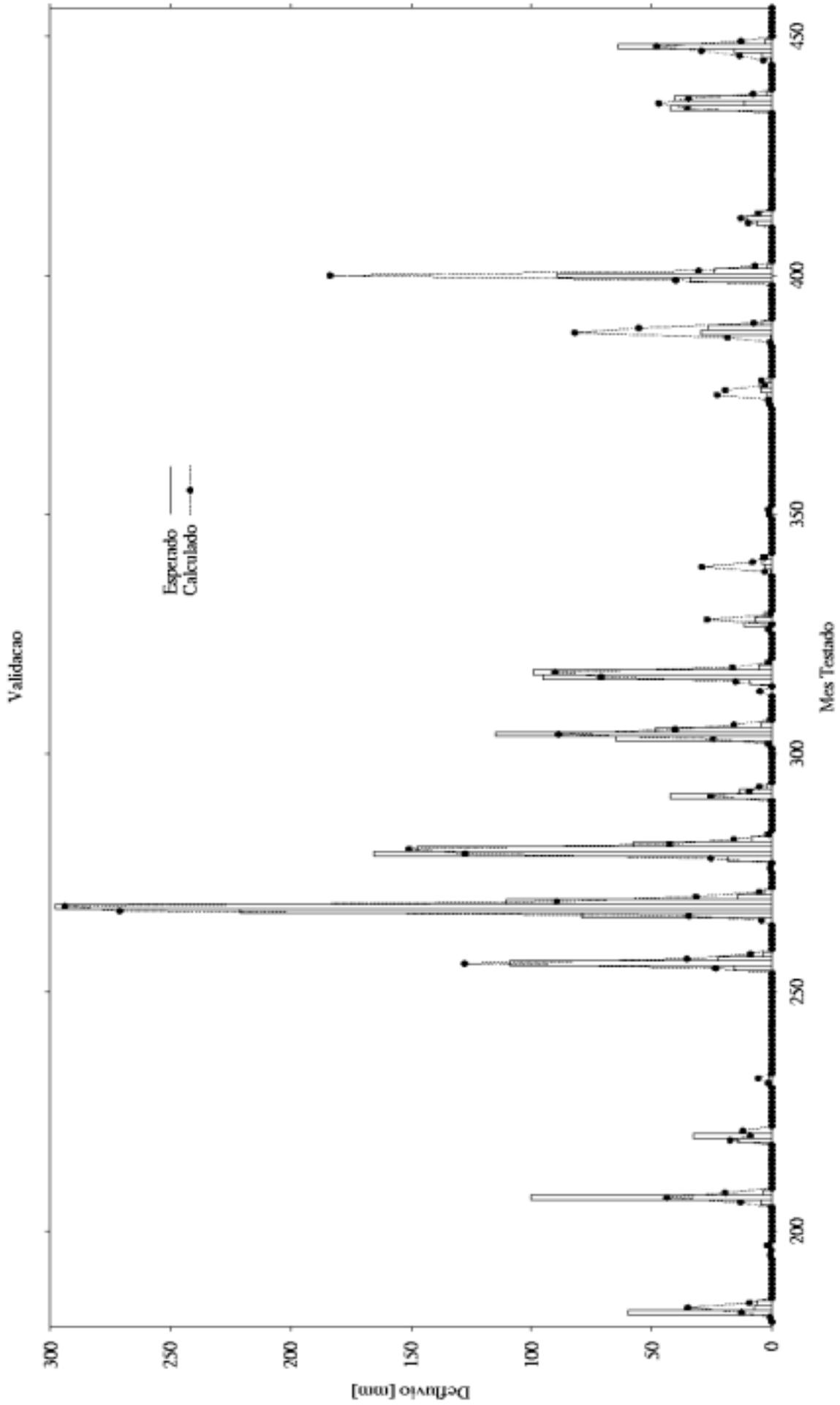


Figura 6: Resultados para a fase de validação do modelo 4

REFERÊNCIAS BIBLIOGRÁFICAS

- FREITAS, M.A.S. Neurocomputação Aplicada. **Apostila de Estudos**. Universidade de Fortaleza, 1999.
- FURTADO, A.N.D. Uma nova abordagem na Avaliação de Projetos de Transporte: O Uso das Redes Neurais Artificiais para Avaliar e Ordenar Alternativas. **Tese EESC-USP**, São Paulo, 1998.
- HECHT-NIELSEN, R. Counter-propagation networks. **Journal of Applied Optics**, v. 26 (3), p. 4979-4984, 1987.
- HSU, K-L., GUPTA, V.K., SOROOSHIANA, S. Artificial neural network modeling of the rainfall-runoff process. **Water Resour. Res.**, v. 31 (10), p. 2517-2530, 1995.
- KACHROO, R.K. River flow forecasting. Part 1. A discussion of the principles. **Journal of Hidrology**, v. 133, p. 1-15, 1992.
- MATH WORKS, INC. Neural Network Tool Box. **User's Guide**, Version 4, Natick, Mass., 2001.
- MINNS, A.W., HALL, M.J.. Artificial neural network as rainfall-runoff models. **Hydrological Science-Journal**, v. 41 (3), p. 399-419, 1996.
- OOYEN, A.V., NICHHUIS, B. Improving the convergence of back propagation problem. **Neural Networks**, v. 5, p. 465-471, 1992.
- RUMELHART, D.E., HINTON, G.E., WILLIAMS, J. Learning internal representation by error propagation. **Parallel Distributed Processing**, v. 1, p. 318-362, MIT Press, Cambridge, Mass., 1986.
- SAJIKUMAR, N., THANDAVESWARA, B.S. A non-linear rainfall-runoff model using an artificial neural network. **Journal of Hydrology**, v.216, p. 32-35, 1999.
- SMITH, J., ELI, R.N.. Neural-Network models of rainfall-runoff process. **Journal of Water Res. Plann. and Management**, p. 499-508, 1995.
- WIDROW, B. Current and past applications of neural networks to adaptative control problems. **Invited Lecture**. West Virginia Univ., College of Engineering, 1989.