



IV-127 - PREVISÃO DA VAZÃO AFLUENTE ANUAL NO AÇUDE BANABUIÚ COM REDES NEURAIAS ARTIFICIAIS RECORRENTES

Adbeel Goes Filho⁽¹⁾

Engenheiro Civil pela Universidade Federal do Ceará (UFC). Mestre em Ciência da Computação pela UFC. Coordenador de Gestão Estratégica do Departamento Nacional de Obras Contra as Secas (DNOCS). Professor da Universidade de Fortaleza (UNIFOR). Doutorando em Engenharia Civil – Recursos Hídricos na UFC.

Endereço⁽¹⁾: Rua Antonina do Norte, 363 – Monte Castelo - Fortaleza - CE - CEP: 60325-610 – Brasil - Tel: +55 (85) 3288-5170 - Fax: +55 (85) 3288-5200 - e-mail: adbeel.goes@dnocs.gov.br.

RESUMO

Este trabalho apresenta resultados de previsão da vazão afluente anual no Açude Banabuiú com redes neurais de Elman. A rede neural artificial (RNA) foi preparada para receber como entrada amostras da série temporal da vazão afluente do rio Banabuiú, pertencente ao sistema do rio Jaguaribe, mês a mês. Os dados usados para treinamento da rede foram os da série temporal das vazões afluentes entre 1912 e 1989. Os anos de 1990 à 1996 não fizeram parte do treinamento da RNA, foram usados para testes de generalização da rede e validação dos resultados.

PALAVRAS-CHAVE: Previsão de vazão; Redes neurais recorrentes; Redes neurais de Elman.

INTRODUÇÃO

A região semi-árida do Brasil sofre constantemente com os efeitos da má distribuição das precipitações, assim, muitos estudos proliferam rapidamente com o objetivo de resolver problemas de várias tipos e intensidades. A capacidade de armazenamento de água superficial na região está se esgotando, e medidas mitigadoras devem ser tomadas para que problemas maiores não ocorram.

Nosso sistema hídrico apresenta aspectos peculiares que o torna diferente de outros sistemas no país. Inicialmente pelas características das afluições nos rios locais, que apresentam acentuada sazonalidade e elevado nível de incerteza, ao contrário do sistema do hemisfério norte no qual o regime hidrológico é proveniente, basicamente, do degelo.

Assim, o planejamento na região depende do conhecimento prévio do volume da água nesses reservatórios, ou seja, é preciso conhecer a quantidade de água que estará disponível, com certa antecedência, para que se possa, finalmente, estimar a quantidade máxima a ser disponibilizada. Em função dos bons resultados da previsão das vazões é possível realizar o planejamento de abastecimento de modo a atingir níveis aceitáveis.

Para viabilizar esse processo existem distribuídos em pontos específicos dos rios que integram as bacias hidrográficas, pontos de medição que fornecem medidas discretas da vazão, permitindo compor as séries históricas de vazão afluente. Prever a vazão significa descobrir, por antecipação, os valores da quantidade de água que irá fluir nesses pontos de medição com base nas séries históricas disponíveis (CHATFIELD, 1991).

A previsão de vazão é um verdadeiro desafio utilizado para o gerenciamento dos recursos hídricos numa bacia hidrográfica (MORAES, 1996). Também a previsão de inundações, da umidade do solo para agricultura, os níveis de navegação de uma via, a disponibilidade hídrica para abastecimento de água, irrigação e produção de energia são usos conhecidos para a previsão antecipada da vazão numa bacia hidrográfica (TUCCI, 2002).

Segundo Tucci (2002), a previsão de vazão pode ser de curto, médio e longo prazo. A de curto prazo é usada para estimar a vazão num trecho da bacia com antecedência de horas ou poucos dias. A previsão de médio prazo envolve a previsão de vazão com antecedência de um a alguns meses e depende, fortemente, das condições conhecidas de clima e dos oceanos que podem influenciar as vazões futuras na bacia hidrográfica.



A previsão de longo prazo é a estimativa, normalmente estatística, dos riscos de ocorrência de vazões num determinado local, como por exemplo: o risco de inundação numa seção de um rio, a probabilidade de períodos úmidos e secos, dentre outros (TUCCI, 2002).

Tradicionalmente usamos métodos estatísticos para prever a vazão afluente, o método de Box-Jenkins, conforme Chatfield (1991), por exemplo, que pressupõe uma relação linear entre os valores presente e passado da vazão. Podem ser aliados aos modelos lineares auto-regressivos (AR), de média móvel (MA) e o auto-regressivo de média móvel (ARMA), que igualmente não são os mais adequados para tratar um conjunto de dados com características não lineares e não estacionárias inerentes às séries de vazão. Por outro lado, as redes neurais artificiais (RNAs) segundo Fog (1995) e Lachtermacher (1995) são modelos compostos de um certo número de elementos não lineares (neurônio artificial) com operações paralelizadas e dispostos em camadas similarmente às redes neurais biológicas. Através de algoritmos específicos as RNAs podem aprender. As RNAs podem ser agrupadas em duas categorias: redes sem realimentação (*feedforward networks*) e com realimentação (*recurrent networks*). As redes sem realimentação são estáticas, isto é, uma dada entrada pode produzir apenas um conjunto de saída e, portanto, não produzem nenhum tipo de memória. Por outro lado as redes recorrentes permitem que as características temporais da informação sejam memorizadas. Um caso típico é a rede neural artificial de Elman, parcialmente recorrente, e que será usada para fazer as previsões de vazões.

As principais características no uso de redes neurais artificiais na análise de séries temporais são os seguintes:

- Aprendem a realizar previsões precisas;
- Permitem previsões iterativas;
- Capazes de lidar com a não estacionariedade e a não linearidade das séries temporais investigadas;
- Oferecem predição paramétrica e não paramétrica.

Milioni (2001) apresentou uma pesquisa cujo objetivo é tentar abstrair a natureza física do processo, simplificando a abordagem, através de um sistema que utiliza, numa primeira etapa, modelos econométricos de regressão múltipla que expliquem a vazão de uma seção de controle a partir da observação do nível do rio a montante. Já Tucci (2002) apresenta os resultados da previsão em tempo real do volume afluente ao reservatório de Ernestina, tendo como objetivo investigar a previsão da vazão afluente, num período de 5 anos, mês a mês, de 1992 a 1996 (60 meses), visando auxiliar a decisão dos especialistas do setor elétrico, envolvidos em planejamento energético.

Não é objetivo do presente trabalho, investigar os processos de pré-processamento, usados, normalmente, no tratamento de séries temporais. Faremos uma breve fundamentação teórica dos procedimentos utilizados e, em seguida, apresentaremos os resultados obtidos, especificamente para o Açude Banabuiú que, com seus 1.800.000.000m³, é uma das principais fontes hídricas do Estado do Ceará.

FUNDAMENTAÇÃO TEÓRICA

As redes neurais estáticas como é o caso dos Perceptrons de múltiplas camadas (MLP), treinadas normalmente pelo algoritmo de retropropagação de erro, apresentam certa dificuldade em realizar mapeamentos dinâmicos. Consequentemente aprender as características temporais de um sinal, que represente a série histórica da vazão afluente de um rio, pode se tornar uma tarefa difícil. Para resolver o problema pode-se usar uma rede MLP tradicional, porém com as entradas atrasadas no tempo (HAYKIN, 1994).

A Figura 1 mostra esse esquema denominado de rede temporal, no qual uma rede MLP é alimentada por uma entrada $u(t)$ que é sucessivamente atrasada de uma unidade pela estrutura z^{-1} até, genericamente, $u(t-k)$ e que tem como saída a $y(t)=u(t)$.

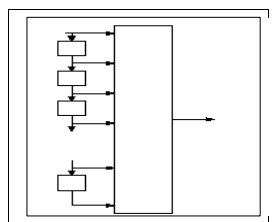


Figura 1: Rede temporal.

Nesse caso, o atraso está sendo aplicado apenas na entrada, porém poderia ser aplicado também nas camadas escondidas e na saída. Uma outra forma de resolver o problema é usar redes com elos de realimentação, denominadas de redes recorrentes. Normalmente as redes recorrentes podem incorporar uma MLP ou parte dela, fazendo uso de sua capacidade de mapeamento não linear. De maneira geral as redes recorrentes são adequadas para tratar problemas temporais, pois as conexões recorrentes ajudam a rede a memorizar informações atuais no momento seguinte da sequência investigada.

Redes neurais recorrentes podem apresentar uma ou mais malhas de realimentação. Naquelas totalmente recorrentes tem-se cada neurônio conectado a qualquer dos outros e constituem o caso mais geral de redes neurais artificiais.

Uma maneira elegante de representar uma rede neural recorrente é através de um modelo em espaço de estados. A noção de estado representa papel principal na formulação matemática de um sistema dinâmico. O estado de um sistema dinâmico é formalmente definido como um conjunto de quantidades que sintetiza toda a informação sobre o comportamento passado do sistema que é necessário para descrever unicamente seu comportamento futuro, exceto pelos efeitos externos advindos da entrada ou da excitação aplicada. Seja o vetor de dimensão $[q \times 1]$ $x(k)$ o estado de um sistema não linear discreto. Seja o vetor de dimensão $[m \times 1]$ $u(k)$ a entrada aplicada ao referido sistema, e o vetor $y(k)$ a sua saída correspondente. Em termos matemáticos, o comportamento dinâmico do sistema, sem ruído, pode ser descrito pelo seguinte par de equações não lineares.

$$x(k+1) = \phi(Wa x(k) + Wb u(k)) \quad (1)$$

$$y(k) = C(x) \quad (2)$$

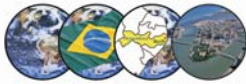
onde Wa é uma matriz de dimensão $[q \times q]$, Wb é uma matriz de dimensão $[q \times m]$, C é uma matriz de dimensão $[p \times q]$, e $\phi: \square q \rightarrow \square q$ é um mapeamento descrito por:

$$\phi: \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{bmatrix} \rightarrow \begin{bmatrix} \phi(x_1) \\ \phi(x_2) \\ \vdots \\ \phi(x_q) \end{bmatrix}$$

para alguma não linearidade sem memória
 $\phi: \square \rightarrow \square$
atuando componente a componente.

Os espaços $\square m$, q , e p são chamados de espaços de entrada, de estado e de saída respectivamente. A dimensão do espaço de estado, q , constitui a ordem do sistema.

Assim, a rede neural recorrente representada pelas equações (1) e (2) é um sistema dinâmico com m entradas e p saídas de ordem q . A equação (1) forma a equação de processo do modelo e (2) é a equação de medida. Com relação às matrizes Wa , Wb e C , e a função não linear $\phi(x)$ podemos fazer as seguintes considerações: a matriz Wa contém os pesos sinápticos dos q neurônios que são conectados aos nós de realimentação na entrada. A matriz Wb contém os pesos sinápticos desses neurônios que são conectados aos nós fontes também na entrada. A matriz C indica a forma da combinação linear dos estados que caracterizarão a saída da rede neural. A função não linear $\phi(x)$ caracteriza a função de ativação de qualquer neurônio na rede. Esta função de ativação é muitas vezes definida pela função tangente hiperbólica descrita na equação (3).



$$\phi(x) = \tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad (3)$$

Uma propriedade importante de uma rede neural recorrente descrita pelas equações de espaço de estado (1) e (2) é a sua capacidade de aproximar uma classe ampla de sistemas não lineares dinâmicos. A Figura 2 representa uma rede neural recorrente com três entradas, três estados e uma saída, ou seja, $m=3$, $q=3$ e $p=1$.

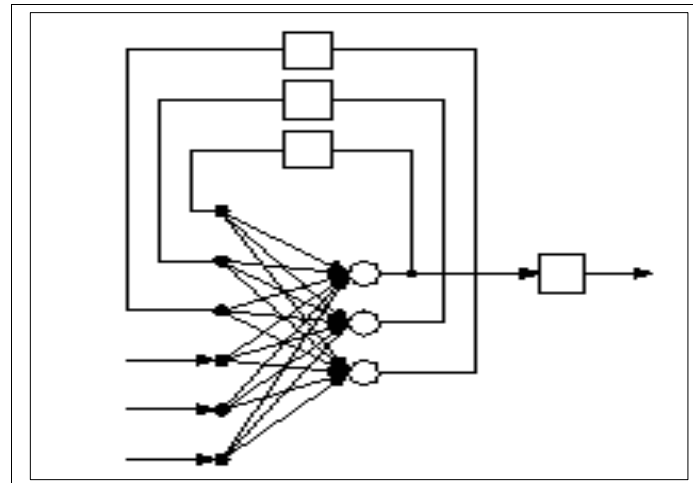


Figura 2: Rede totalmente recorrente.

As matrizes W_a e W_b são definidas como:

$$W_a = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix}$$

$$W_b = \begin{bmatrix} w_{14} & w_{15} & w_{16} \\ w_{24} & w_{25} & w_{26} \\ w_{34} & w_{35} & w_{36} \end{bmatrix}$$

A matriz C é um vetor linha definido por:

$$C = [1 \ 0 \ 0]$$

A Figura 3 representa um modelo auto-regressivo não linear com entradas exógenas. Na Figura 4 mostramos uma memória de linha de atraso com n unidades que é aplicada à entrada. A saída também é realimentada à entrada por outra memória de atraso (HAYKIN, 1994).

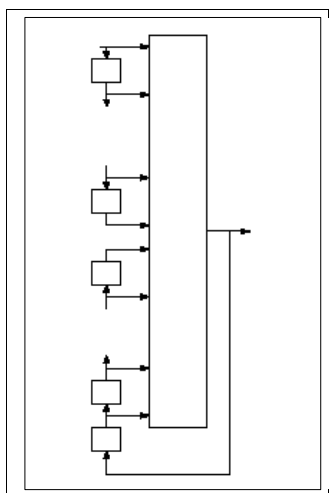


Figura 3: Modelo auto-regressivo.

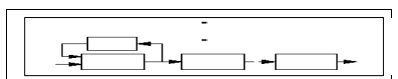


Figura 4: Rede neural de Elman.

Para uma entrada $u(t)$ a saída $y(t+1)$ estará adiantada de uma unidade em relação à entrada. Os valores atuais e passados da entrada são: $u(t)$, $u(t-1)$, $u(t-2)$... $u(t-k+1)$ e os valores atrasados da saída são: $y(t)$, $y(t-1)$, $y(t-2)$... $y(t-k+1)$ sobre os quais é feita a regressão da saída do modelo, onde \square definida na equação (4) é uma função não linear.

$$y(t+1) = \square(y(t), \dots, y(t-k+1), u(t), \dots, u(t-k+1)) \quad (4)$$

A rede de Elman é uma arquitetura de RNA considerada parcialmente recorrente, pois os elos de realimentação ocorrem entre a saída e a entrada da primeira camada oculta. O elo recorrente é feito através de uma unidade de contexto, normalmente uma estrutura de atraso do tipo z^{-1} , armazenando a saída da primeira camada oculta por um passo de tempo. Assim, essa estrutura garante a geração de padrões variáveis no tempo, habilitando essa configuração neural para aplicações envolvendo séries temporais.

Além da camada recorrente a rede pode apresentar várias outras camadas formando uma tradicional MLP, com uma ou várias saídas. O algoritmo usado para o treinamento é o de retropropagação de erro (*backpropagation*), intensivamente discutido na literatura. Na verdade existe apenas um cuidado que deve ser tomado na aplicação do algoritmo: num dado instante a rede recebe não só as entradas externas, normalmente aplicadas à ela, mas também realimentações da saída da primeira camada oculta, obtidas no instante anterior.

A Figura 4 apresenta uma rede de Elman, na qual se procura mostrar, com nitidez, o mecanismo de funcionamento da estrutura de realimentação. No caso da utilização em previsão de séries temporais é possível adotar soluções híbridas, fazendo-se com que as entradas sejam devidamente atrasadas.

MODELAGEM

Adotamos um modelo semelhante ao mostrado na Figura 4, no qual as entradas são apresentadas a rede de 12 em 12 meses e encontram-se defasadas uma da outra de 1 mês, conforme malha defasadora representada na Figura 5. A disponibilidade de dados é de 85 anos de vazão mensal relativa a vazão afluente no rio Banabuiú, pertencente ao sistema do rio Jaguaribe.

Desse total usaremos para treinamento da rede os dados dos anos de 1912 à 1989 de vazão, deixando para teste de generalização e validação os anos de 1990 à 1996. Os dados de treinamento estão arranjados segundo uma matriz de entrada contendo 78 linhas cada uma defasada da outra de 1 mês e 12 colunas relativas ao período de 1 ano.

A Figura 5 mostra um modelo da rede utilizada, sem a malha defasadora, representada na Figura 4. A matriz de teste de generalização e validação dos resultados é construída de forma semelhante a matriz de entrada. O número ótimo de neurônios nas diversas camadas escondidas (N...R), vistos na mesma figura, é escolhido durante o treinamento, visando-se obter o melhor resultado da rede, especialmente no tocante a generalização. A camada de saída tem um único neurônio que apresentará o vetor representativo da vazão que se deseja prever. O algoritmo de treinamento utilizado foi o de retropropagação de erro, adequado à essa arquitetura de rede.

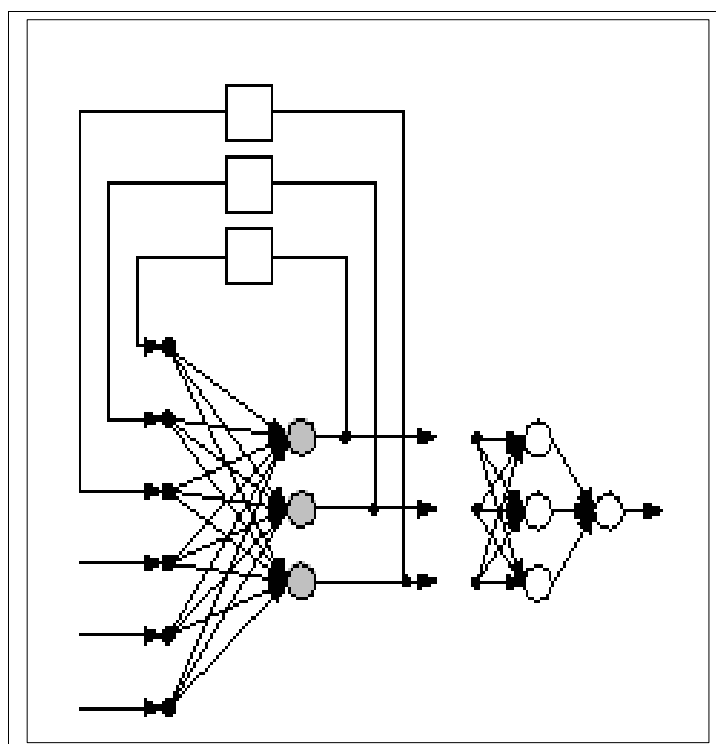


Figura 5: Modelo sintético da RNA.

RESULTADOS

Visando obter as melhores características de generalização foram testadas diversas arquiteturas, usando-se os dados conhecidos referentes aos anos de 1990 à 1996 e que não tomaram parte no treinamento da rede. A arquitetura que resultou a melhor qualidade engloba três camadas ocultas (10-5-3) e uma camada de saída, com apenas um neurônio, e usando a função tangente hiperbólica como função de ativação. Foram testadas várias funções de ativação sendo a tangente hiperbólica a que trouxe os melhores resultados. O método numérico usado para a minimização do erro foi o do gradiente decrescente, com taxa de aprendizado adaptativa e coeficiente de momento, que minimiza as oscilações. Depois de devidamente treinada a arquitetura, os parâmetros de treinamento e as matrizes de pesos foram armazenados em arquivos digitais para futuramente serem utilizados na previsão mês a mês.

Foram utilizados nos cálculos os *softwares* QNet e Java NNS para ajustes dos conjuntos de treinamento, e o *software* SATIS VMath J7x módulo RNA, conforme Figura 6, para comparações, tendo sido desenvolvido especificamente para este trabalho. A Figura 7 representa o modelo da RNA e a Figura 8 o resultado do processamento da RNA com o software QNet.

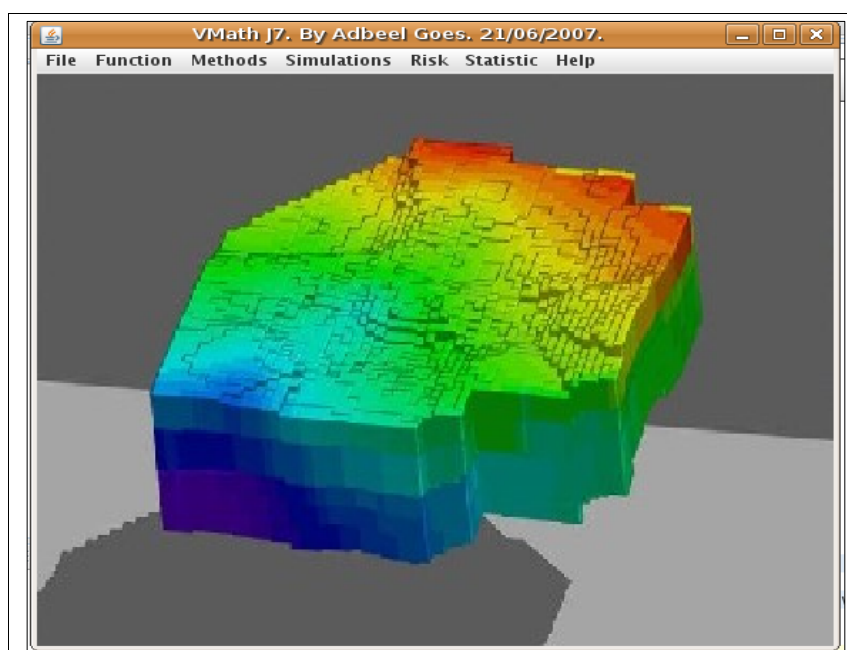


Figura 6: Software SATIS VMATH J7x.

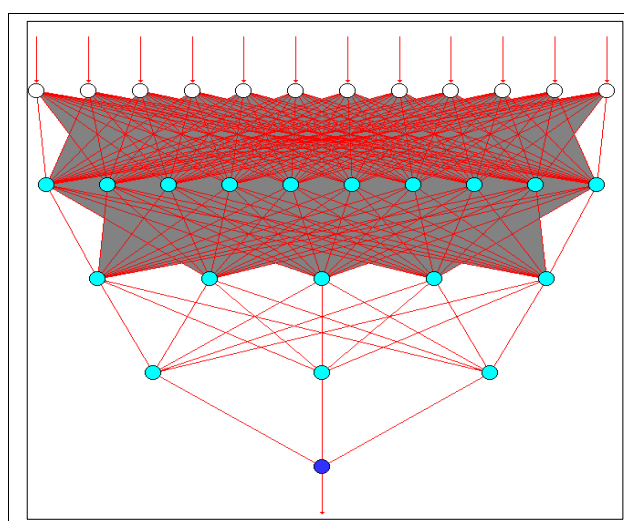


Figura 7: Modelo da RNA.

Network Definition		Training Controls	
Vazões Afluentes no Açude		Max Iterations:	10000
Network Layers:	5	Learn Control Start:	1
Input Nodes:	12	Learn Rate:	0.001001
Output Nodes:	1	Learn Rate Max:	0.001001
Hidden Nodes:	18	Learn Rate Min:	0.001000
Transfer Functions:	Tanh	Momentum:	0.800
Connections:	188	Patterns per Update:	78
Training Patterns:	86	FAST-Prop:	0.000
Test Patterns:	0	Screen Update:	200
Network Size (Bytes):	18198	AutoSave Rate:	0
Training Mode:	standard	Tolerance:	3.00000
Net Training/Total:	1/1	Quit at RMS Error:	0.00200
Training Results			
Iteration:	10000	Training Speed (CPS):	34225K
Percent Complete:	100.0%	Time Remaining:	0:0:0
	RMS Error	Correlation	Tol. Correct
Training Set:	0.005804	0.998655	87.2%
Test Set:			
Max Iterations Reached!			

Figura 8: Resultado do processamento.

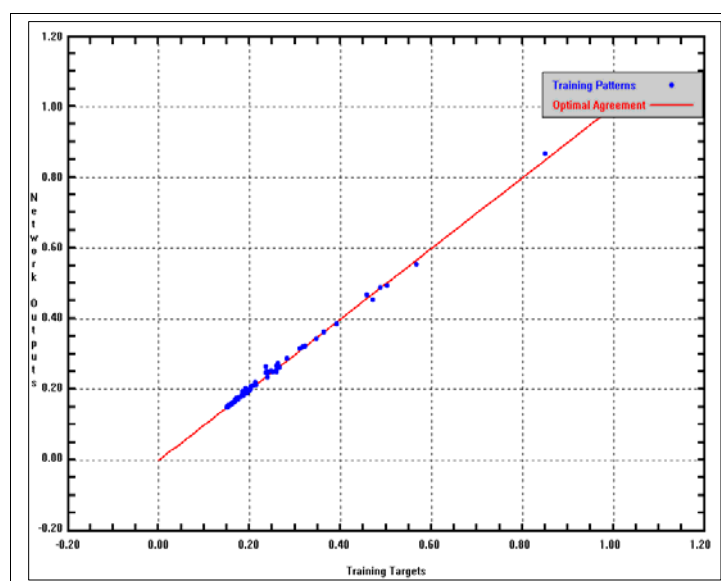


Figura 9: Normalização x saídas.

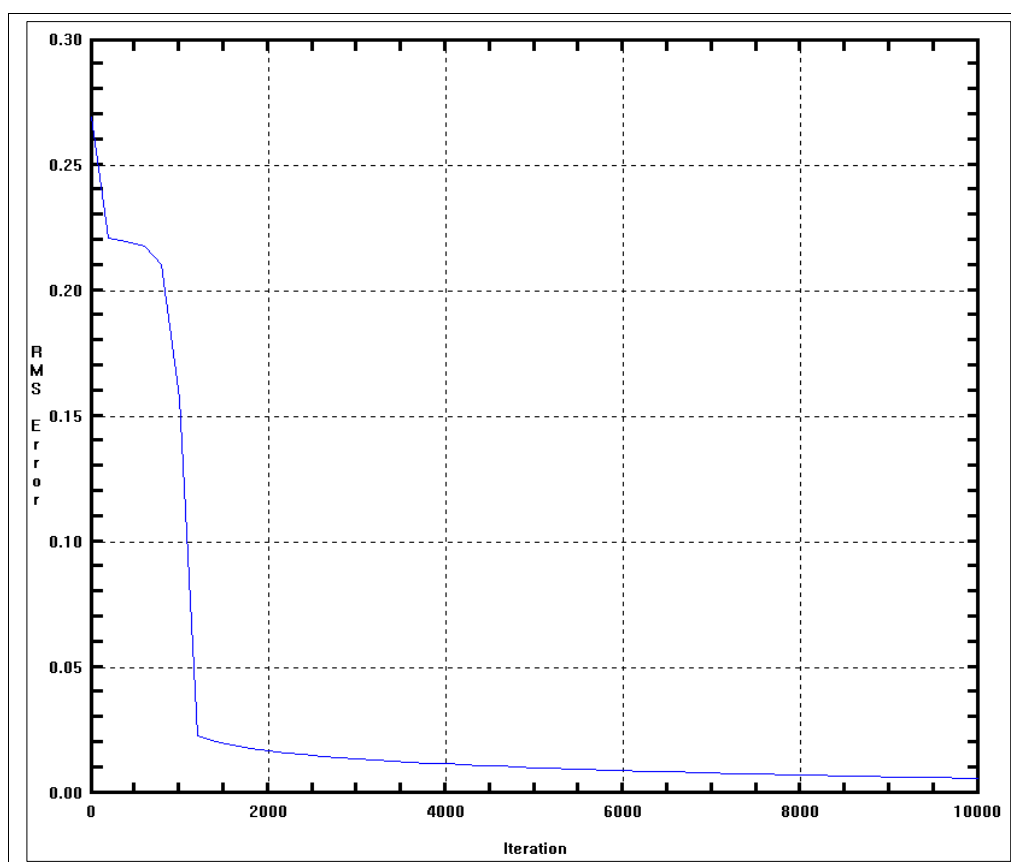


Figura 10: Erro médio quadrático (RMS) do conjunto de treinamento.

A Figura 9 representa a curva de erro percentual de previsão feita pela RNA, cujo valor médio não ultrapassa 0,59 %.

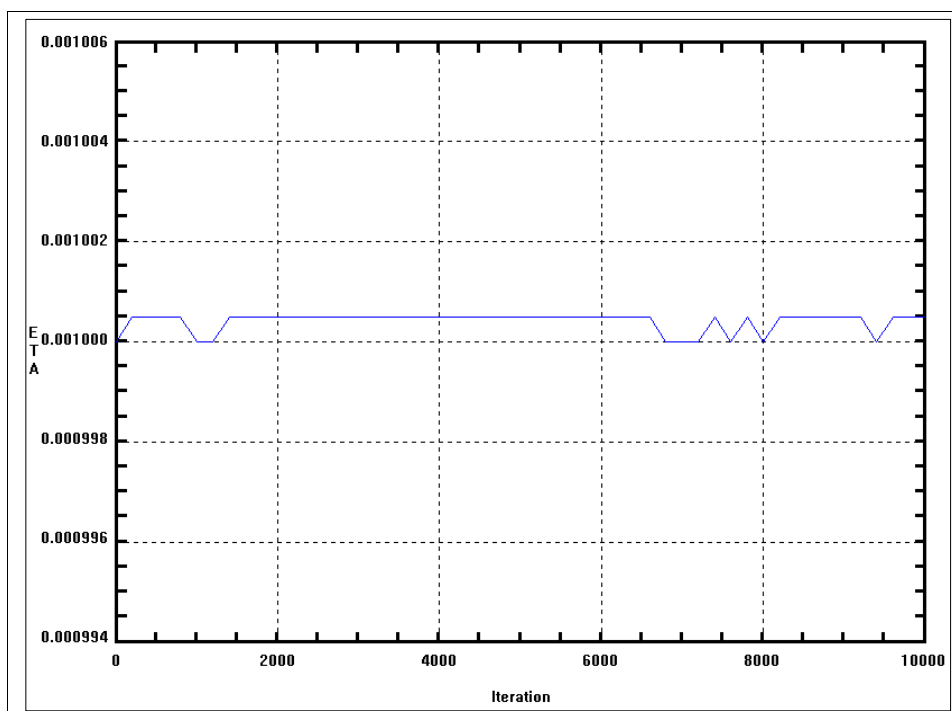


Figura 11: Histórico da taxa de aprendizagem.

A Figura 10 representa o erro médio quadrático do conjunto de treinamento e a Figura 11 representa o histórico da taxa de aprendizagem, referentes aos anos entre 1990 e 1996. Superposta a esta curva encontra-se também a curva de previsão (pontilhada) obtida mês a mês pela RNA, apresentando um erro médio menor que 0,59%.

CONCLUSÕES

O resultado obtido através da metodologia proposta com o uso de redes parcialmente recorrentes de Elman, com malha defasadora na entrada, indica serem adequadas para previsão de séries temporais. No caso investigado, o erro médio de previsão de cerca de 0,59% é muito menor que os obtidos por RNAs tradicionais, que atingem a casa de 5% (LACHTERMACHER, 1995).

Quando comparado com erros obtidos por processos estatísticos e suas variações, onde o erro médio é maior que 10%, a superioridade da rede de Elman ainda é mais significativa. A Figura 12 mostra a equivalência gráfica entre os valores reais e previstos.

Finalmente, se forem incorporadas ao processo às características inerentes a temperatura do oceano e evaporações como mais entradas da RNA, deveremos ter resultados ainda melhores. Observamos a importância de maiores estudos das redes de Elman e suas variações, na previsão de séries temporais, tão usadas na área de recursos hídricos, e de alto valor para pesquisas complexas na região do semiárido brasileiro.

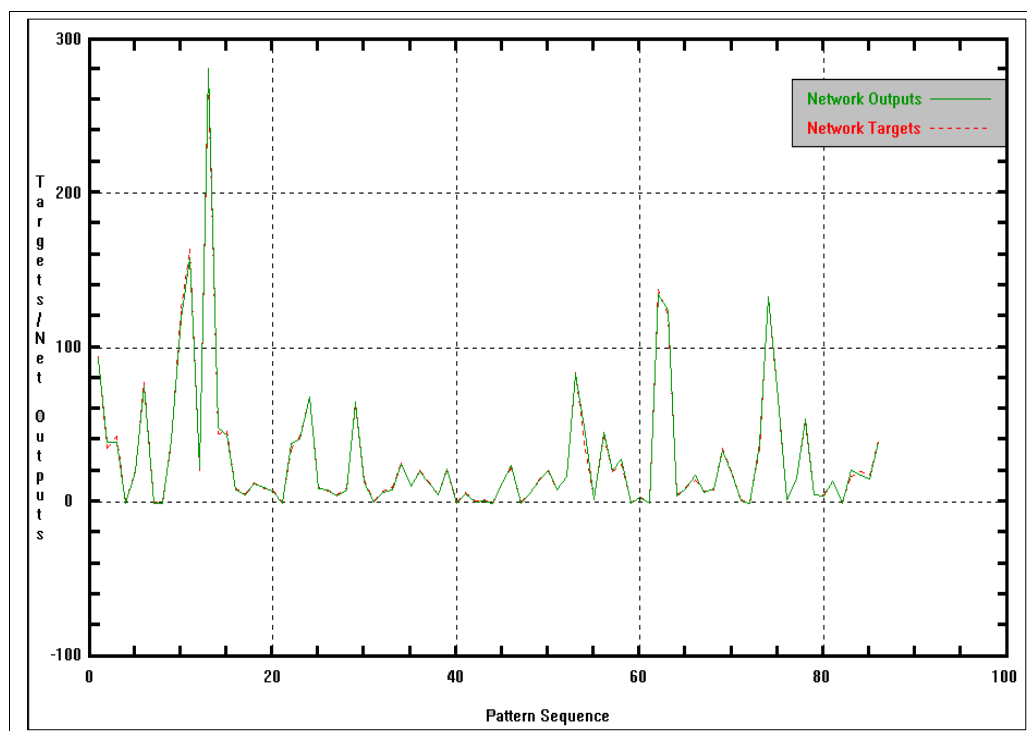


Figura 12: Gráfico de valores reais e previstos.

REFERÊNCIAS BIBLIOGRÁFICAS

1. CHATFIELD, E. *The Analysis of Time Series*. New York, USA: Chapman and Hall, 4th ed, 1991.
2. FOG, T.L. *Training and Evaluation of Neural Networks for Multi-Variate Time Series Processing*. In: Proceedings of IEEE International Conference on Neural Networks. IEEE Press, 1995.
3. HAYKIN, S. *Neural Networks, a comprehensive Foundation*. London: Macmillan College Publishing Co, 1994.
4. LACHTERMACHER, G. & FULLER, J.D. *Backpropagation in Time series Forecasting*. In: Journal of Forecasting. Vol 14, 381-393, 1995.
5. MILIONI, A. Z. & ACIOLI A. O. *Modelos Econométricos para a Previsão de Cheias Fluviais*. In: XXXIII SBPO - Simpósio Brasileiro de Pesquisa Operacional, Campos do Jordão, SP. Anais do XXXIII SBPO, v. 1. p. 534-542, 2001.
6. MORAES, J.M.; GENOVEZ, A.M.; MORTATTI, J.; PELLEGRINO, G.; BALLESTER, M.V. & MARTINELLI, L. A. *Análises e Modelagem de uma Série Temporal de Vazão sob a Influência de Ações Naturais e Antrópicas*. In: Anais do XVII Congresso Latino Americano de Hidráulica, 1996.
7. TUCCI, C.E.M.; ROBIN T.C.; DIAS P.L. DA S. & COLLISCHONN W. *Previsão de médio prazo da afluência de reservatórios com base na previsão climática*. In *Relatório Final: BRA/00/029*, Instituto de Astronomia, Geofísica e Ciências Atmosféricas e Instituto de Pesquisas Hidráulicas. UFRS, 2002.