

Previsão de Vazões utilizando Redes Neurais Dinâmicas do Tipo FIR

I. Luna, S. Soares, M. H. Magalhães e R. Ballini

Resumo—Este trabalho propõe um modelo de previsão baseado em redes neurais dinâmicas do tipo *Finite Impulse Response* (FIR), ou seja, nesta estrutura as sinapses são representadas por filtros lineares do tipo FIR, ao invés de pesos escalares. Os parâmetros da rede são ajustados usando o algoritmo de retropropagação temporal. Este modelo é aplicado ao problema de previsão de séries temporais, mais especificamente séries de vazões, devido a sua habilidade para o processamento temporal e eficiência na resolução de problemas não-lineares. Os modelos de previsão de vazões são de grande importância no planejamento de sistema hidroelétrico, sendo necessário desenvolver modelos capazes de capturar a natureza não-linear das séries, que forneçam previsões satisfatórias a médio e longo prazo. Assim, o modelo FIR é aplicado para previsão um passo à frente e vários passos à frente. O desempenho deste modelo é comparado ao modelo de previsão utilizando redes neurais multicamadas (MLP) e ao modelo periódico autoregressivo (PAR). Os resultados mostram em geral, um melhor desempenho da rede FIR para o problema em estudo.

Index Terms—rede dinâmica, FIR, *temporal backpropagation*, previsão de vazões.

I. INTRODUÇÃO

Atualmente, a área de planejamento e controle de sistemas de recursos hídricos é de grande importância, uma vez que influencia diretamente na produção de energia hidroelétrica do país. Esta, por sua vez, depende da série histórica de vazões das diversas usinas do sistema, no sentido de sua utilização como dado de entrada para um modelo de simulação. Estes modelos são importantes para realizar as previsões de vazões necessárias para um planejamento energético efetivo. Entretanto, trata-se de um problema estocástico e não-linear.

O sistema hidroelétrico brasileiro encontra-se distribuído em diversas regiões geográficas, com dados coletados através de diversas redes de aquisição de dados que resulta em uma considerável incerteza na informação hidrológica coletada. Além disso, a inerente relação não-linear entre as variáveis envolvidas dificulta e afeta o desempenho dos modelos de previsão. Uma outra dificuldade na previsão de vazões é a natureza não-estacionária devido aos períodos úmidos e secos ao longo do ano [1].

A maior parte dos modelos de previsão atualmente utilizados, baseiam-se em relações lineares entre as variáveis. Estas técnicas são freqüentemente baseadas na metodologia clássica de Box & Jenkins [2]. Este modelo tem sido amplamente aplicado a problemas associados a recursos hídricos, principalmente na previsão de vazões mensais. Nos últimos anos, modelos baseados em redes neurais artificiais tem surgido como al-

ternativas de modelos de previsão [3], uma vez que estes são capazes de capturar as relações não-lineares entre as entradas e a saída do modelo [1].

Desta forma, redes neurais vêm sendo utilizadas como uma poderosa técnica para tratar problemas de previsão de séries temporais. Contudo, a aplicação destes modelos vem sendo dominada por estruturas estáticas [4]. Um dos primeiros trabalhos que utilizou redes neurais recorrentes como modelos de previsão de séries temporais foi proposto em [5].

No decorrer destes anos, diferentes arquiteturas de redes neurais recorrentes têm sido aplicadas a problemas de previsão com diversos graus de desempenho [6]. Em todas essas estruturas, as redes neurais recorrentes têm demonstrado importantes características não encontradas em redes neurais estáticas, tais como, a habilidade de armazenar informação para depois utilizar esta informação na resolução de problemas, através do processamento temporal associado à sua topologia. Além disso, embora a análise e a síntese de redes neurais dinâmicas sejam mais complexas, estes modelos têm demonstrado desempenho superior em diversas aplicações [7]. A presença de realimentação permite a criação de representações internas e mecanismos de memória capazes de processar e armazenar tanto informações temporais, como seqüenciais [8].

Neste trabalho, propõe-se a aplicação de uma rede neural dinâmica na previsão de séries temporais. A estrutura da rede é baseada em um modelo do tipo *Finite Impulse Response* (FIR), proposto em [9]. Este modelo é uma modificação do modelo clássico de redes neurais multicamadas, no qual cada peso é substituído por um filtro linear do tipo FIR. Os pesos da rede são ajustados utilizando o algoritmo de retropropagação temporal.

O objetivo deste trabalho é a aplicação da rede FIR para previsão de vazões incrementais médias mensais um passo à frente e vários passos à frente, da usina hidroelétrica de Furnas, localizada no Rio Grande, região Sudeste do Brasil.

O desempenho deste modelo é comparado aos resultados obtidos pelos modelos periódico autoregressivo (PAR) e redes neurais multi-camadas (MLP) com algoritmo de retropropagação dos erros. A comparação com o modelo PAR é devido este modelo ser empregado no setor elétrico para previsão de vazões. Já a rede MLP é uma das estruturas amplamente utilizadas em distintos problemas de previsão de séries temporais, como em [10], [11], entre outros.

As comparações feitas mostram que o modelo de previsão baseado na estrutura de rede dinâmica do tipo FIR possui um melhor desempenho, considerando as previsões um e vários passos a frente.

O artigo está organizado da seguinte forma. A próxima seção descreve a estrutura da rede dinâmica do tipo FIR. A seção III apresenta o algoritmo de treinamento utilizado para ajustar os

I. Luna e S. Soares, Departamento de Engenharia de Sistemas - DEN-SIS, Faculdade de Engenharia Elétrica e de Computação - FEEC, Universidade Estadual de Campinas - UNICAMP, 13083-970 Campinas, Brasil. e-mail: {iluna,dino}@cose.fee.unicamp.br

M. H. Magalhães, Departamento de Computação e Automação Industrial - DCA, Faculdade de Engenharia Elétrica e de Computação - FEEC, Universidade Estadual de Campinas - UNICAMP, 13083-970 Campinas, Brasil. e-mail: marinah@dca.fee.unicamp.br

R. Ballini, Departamento de Teoria Econômica - DTE, Instituto de Economia IE, Universidade Estadual de Campinas - UNICAMP, 13083-852 Campinas, Brasil. e-mail:ballini@eco.unicamp.br

parâmetros do modelo. Na seção IV apresenta algumas diretrizes e características do problema de previsão em estudo, assim como os resultados obtidos e as comparações com os modelos PAR e MLP. Finalmente, na seção V discute-se as conclusões do trabalho.

II. REDE NEURAL RECORRENTE FIR

A estrutura da rede dinâmica do tipo FIR foi proposta em [9]. Este tipo de rede neural tem sido utilizada em diversas aplicações, tais como a previsão de séries temporais, controle adaptativo, processamento de sinais e identificação de sistemas [12], obtendo bons resultados devido às suas características dinâmicas [13]. Este modelo é uma modificação da rede neural multicamadas, em que as sinapses são representadas por um filtro linear do tipo FIR ao invés de um peso escalar. Para um melhor entendimento, a seguir será apresentada uma breve descrição da estrutura da rede.

Sejam os coeficientes do filtro linear W , definido como:

$$W = [w(0), w(1), \dots, w(T)]^T.$$

A saída do filtro linear é definida como:

$$y(k) = \sum_{i=0}^{i=T} w(i) \cdot x(k-i) \quad (1)$$

sendo $x(k)$ a entrada no tempo discreto k . No caso da rede dinâmica do tipo FIR, \mathbf{W}_{ij}^l representa o vetor de peso entre o neurônio i da camada l e o neurônio j da camada seguinte. N_l é o número de neurônios da l -ésima camada, com $l = 1, 2, \dots, L$. A dinâmica da rede pode ser representada da seguinte maneira:

$$\mathbf{W}_{ij}^l = [w_{ij}^l(0), w_{ij}^l(1), \dots, w_{ij}^l(T_l)]^T \quad (2)$$

$$\mathbf{X}_i^l(k) = [x_i^l(k), x_i^l(k-1), \dots, x_i^l(k-T_l)]^T \quad (3)$$

$$y_{ij}^l = \mathbf{W}_{ij}^{l-1} \cdot \mathbf{X}_i^{l-1}(k) \quad (4)$$

$$y_j^l(k) = \sum_{i=1}^{N_{l-1}} y_{ij}^l(k) \quad (5)$$

$$x_j^l(k) = f(y_j^l(k)) \quad (6)$$

com $i \in [1, N_l]$, $j \in [1, N_{l+1}]$ e $l \in [1, L]$. A Figura 1 ilustra a configuração de uma camada da rede FIR.

III. ALGORITMO DE TREINAMENTO

Os pesos da rede FIR são ajustados utilizando o algoritmo de retropropagação temporal. Seja a saída desejada $d_i(k)$ no instante k , para $1 \leq i \leq N_l$. O erro total quadrático instantâneo é definido da seguinte forma:

$$e^2(k) = \sum_{i=1}^{N_l} e_i^2(k) \quad (7)$$

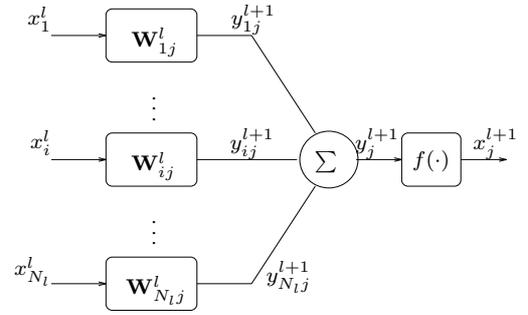


Figura 1. Estrutura da rede neural dinâmica do tipo FIR: conexões entre os neurônios da camada l e o neurônio j da camada seguinte.

com $e_i(k) = d_i(k) - x_i^L(k)$. O erro quadrático total é calculado como:

$$e^2 = \sum_{k=0}^K e^2(k) \quad (8)$$

Como o objetivo é minimizar o erro quadrático total, a variação nos valores correspondentes aos vetores de pesos pode ser re-escrita como segue:

$$\frac{\partial e^2}{\partial \mathbf{W}_{ij}^l} = \sum_{k=0}^K \frac{\partial e^2}{\partial \mathbf{W}_{ij}^l} = \sum_{k=0}^K \frac{\partial e^2}{\partial y_j^{l+1}(k)} \frac{\partial y_j^{l+1}(k)}{\partial \mathbf{W}_{ij}^l} \quad (9)$$

Considerando a versão *on-line* da equação acima, a atualização dos vetores de pesos é definida por:

$$\mathbf{W}_{ij}^l(k+1) = \mathbf{W}_{ij}^l(k) - \eta \frac{\partial e^2}{\partial y_j^{l+1}(k)} \frac{\partial y_j^{l+1}(k)}{\partial \mathbf{W}_{ij}^l} \quad (10)$$

sendo η a taxa de aprendizado com valores no intervalo $(0, 1]$. Entretanto,

$$\frac{\partial y_j^{l+1}(k)}{\partial \mathbf{W}_{ij}^l} = \mathbf{X}_i^l(k) \quad (11)$$

sendo $\mathbf{X}_i^l(k)$ o vetor de entrada aplicado a sinapse i do neurônio j . Definindo o gradiente local do neurônio j como:

$$\delta_j^l(k) = \frac{\partial e^2}{\partial y_j^l(k)} \quad (12)$$

a equação (10) pode ser re-escrita como:

$$\mathbf{W}_{ij}^l(k+1) = \mathbf{W}_{ij}^l(k) - \eta \delta_j^{l+1} \mathbf{X}_i^l(k) \quad (13)$$

sendo esta regra válida para $l = 1, \dots, L$. Existem dois casos para o cálculo de δ_j^l :

1. No caso do neurônio j pertencer a camada de saída:

$$\delta_j^L(k) = \frac{\partial e^2}{\partial y_j^L(k)} = -2e_j(k) f'(y_j^L(k)) \quad (14)$$

2. No caso do neurônio j pertencer a uma camada oculta:

$$\delta_j^l(k) = \frac{\partial e^2}{\partial y_j^l(k)} \quad (15)$$

$$= \sum_{m=1}^{N_{l+1}} \sum_{t=0}^K \frac{\partial e^2}{\partial y_m^{l+1}(t)} \frac{\partial y_m^{l+1}(t)}{\partial y_j^l(k)} \quad (16)$$

$$= \sum_{m=1}^{N_{l+1}} \sum_{t=0}^K \delta_m^{l+1}(t) \frac{\partial y_m^{l+1}(t)}{\partial y_j^l(k)} \quad (17)$$

$$= f'(y_j^l(k)) \sum_{m=1}^{N_{l+1}} \sum_{t=0}^K \delta_m^{l+1}(t) \frac{\partial y_m^{l+1}(t)}{\partial x_j^l(k)} \quad (18)$$

Das equações (2)-(4), tem-se:

$$y_{jm}^{l+1} = \sum_{u=0}^{T_l} w_{jm}^l(u) x_j^l(t-u) \quad (19)$$

Destá última equação, pode-se conferir que para $u \leq t \leq T_l + u$:

$$\frac{\partial y_{jm}^{l+1}(k)}{\partial x_j^l(u)} = w_{jm}^l(t-u) \quad (20)$$

e para qualquer outro valor de t , $\partial y_{jm}^{l+1}(k)/\partial x_j^l(u) = 0$. Assim, substituindo a equação (20) na equação (18), tem-se:

$$\delta_j^l(k) = f'(y_j^l(k)) \sum_{m=1}^{N_{l+1}} \sum_{t=k}^{T_l+k} \delta_m^{l+1}(t) w_{jm}^l(t-k) \quad (21)$$

$$= f'(y_j^l(k)) \sum_{m=1}^{N_{l+1}} \sum_{n=0}^{T_l} \delta_m^{l+1}(k+n) w_{jm}^l(n) \quad (22)$$

$$= f'(y_j^l(k)) \sum_{m=1}^{N_{l+1}} \Delta_m^{l+1} \mathbf{W}_{jm}^{l+1} \quad (23)$$

sendo $\Delta_m^l(k)$ definido como:

$$\Delta_m^l(k)s = [\delta_m^l(k), \delta_m^l(k+1), \dots, \delta_m^l(k+T_{l-1})] \quad (24)$$

O algoritmo de retropropagação temporal é definido pelas equações (13), (14) e (23). Como a atualização dos pesos no instante k depende de p valores futuros, é necessário armazenar os valores de $\delta_j^l(k)$, o número de iterações p necessárias para construir o vetor $\Delta_m^l(k)$, para $k = k + p$. Maiores detalhes podem ser encontrados em [3] e [9].

IV. PREVISÃO DE VAZÕES

O planejamento e operação dos sistemas de energia e recursos hídricos envolve características de produção não-lineares e requer a utilização de inúmeras variáveis de entrada, sendo

uma delas a previsão de vazões, necessárias para o propósito de simulação, otimização e tomada de decisões. Por este motivo, as previsões ao longo do período de planejamento devem ser as mais acuradas possíveis.

Dado um conjunto de amostras da série de vazões médias mensais $x_{t-h} \in \mathfrak{R}$, $h = 1, \dots$, o objetivo deste trabalho é estimar valores de x_t utilizando um conjunto de valores passados de x_t obtidos através do histórico de vazões, ou seja, estimar a previsão um passo à frente. Além disso, também é considerado a previsão vários passos à frente, isto é, estimar x_t utilizando a informação de um conjunto de valores já previstos x_{t-h} , $h = 1, \dots, k$. Neste trabalho, adotou-se $k = 12$, que corresponde a previsão de um ano à frente. Estes dois tipos de previsão foram escolhidos devido a sua relevância no planejamento hidroelétrico a médio e longo prazo, respectivamente.

Nesta seção, o modelo de rede neural dinâmica do tipo (FIR) descrito nas seções anteriores é aplicado a previsão de vazões mensais da usina hidroelétrica de Furnas, situada na região Sudeste do Brasil. Para o ajuste do modelo, foram utilizados dados de vazões incrementais médias mensais, correspondentes ao período de 1931 a 1995. Como dados de teste, foram utilizadas as vazões correspondentes ao período de 1996 a 2001. O desempenho do modelo foi comparado a outros dois modelos de previsão já conhecidos na literatura, uma rede neural multicamada (MLP) e um modelo periódico autoregressivo (PAR).

Os três modelos são comparados utilizando a raiz do erro quadrático médio **REQM** (m^3/s), o erro absoluto médio **EAM** (m^3/s) e o erro relativo médio percentual **ERM** (%), que são definidos como segue:

$$\mathbf{REQM} = \left(\frac{1}{N} \sum_{k=1}^N (d_k - \hat{y}_k)^2 \right)^{1/2} \quad (25)$$

$$\mathbf{EAM} = \frac{1}{N} \sum_{k=1}^N |d_k - \hat{y}_k| \quad (26)$$

$$\mathbf{ERM} = \frac{100}{N} \sum_{k=1}^N \frac{|d_k - \hat{y}_k|}{d_k} \quad (27)$$

sendo N o número de dados de teste, \hat{y}_k a saída do modelo e d_k a saída desejada.

Outro critério importante de avaliação é o coeficiente de correlação definido pela equação:

$$\rho = \frac{\sum_{k=1}^N (d_k - \bar{d})(\hat{y}_k - \bar{y})}{\left(\sum_{k=1}^N (d_k - \bar{d})^2 (\hat{y}_k - \bar{y})^2 \right)^{1/2}} \quad (28)$$

O coeficiente de correlação é um critério utilizado para medir o grau de associação entre a série de vazões previstas e as vazões do histórico. Ou seja, no caso da previsão de vazões, quanto mais próximo esteja o coeficiente de correlação da unidade, significa que em média, a previsão foi melhor. Na equação (28), \bar{d}

e \bar{y} são os valores médios das vazões previstas e das vazões do histórico, respectivamente.

Os dados de vazões oscilam entre valores mínimos e máximos ao longo da variação sazonal no período de 12 meses. Assim, a sazonalidade das vazões mensais sugere o uso de 12 modelos diferentes, um para cada mês do ano tanto na previsão um passo à frente, como na previsão 12 passos à frente. Além disso, deve-se lembrar que as séries de vazões apresentam uma alta variabilidade nos meses úmidos (de Outubro a Março), que dificulta a modelagem e as previsões associadas, isto devido à natureza não-estacionária [1].

O número de entradas para os três modelos ajustados foi definido de acordo como o Critério de Informação de Bayes (BIC) [14]. A Tabela I mostra o número de entradas para cada mês, o qual variou de 1 a 5. Nota-se que os meses de Maio e Outubro são ajustados utilizando 5 entradas. Isso pode ser devido a esses meses corresponderem a períodos de transição de um período úmido para um período de seca, no caso do mês de Maio ou, no caso do mês de Outubro, de um período de seca para um período úmido.

TABELA I
NÚMERO DE ENTRADAS PARA OS MODELOS DE PREVISÃO.

Mês	Núm. Entradas	Mês	Núm. Entradas
Janeiro	1	Julho	2
Fevereiro	1	Agosto	4
Março	1	Setembro	4
Abril	3	Outubro	5
Maio	5	Novembro	1
Junho	1	Dezembro	2

As funções de ativação para os neurônios da camada oculta tanto para o modelo FIR como para o modelo MLP foram definidas como funções logísticas, enquanto os neurônios da camada de saída possuem funções de ativação linear. Ambos os modelos neurais, possuem uma única camada oculta, ou seja, $0 \leq l \leq 2$.

Os valores de atraso T_l para cada modelo mensal FIR foram selecionados de forma empírica, variando de 0 a 4, e o número de neurônios na camada intermediária variaram de 10 a 20, dependendo do mês a ser previsto. Utilizando estas características, os modelos FIR foram ajustados utilizando o algoritmo descrito na seção anterior.

A rede neural MLP foi ajustada usando o algoritmo de retropropagação do erro, com uma taxa de aprendizado de 0,1. O número de neurônios intermediários na rede neural MLP variou de 8 a 20 e após 1000 épocas o processo de aprendizado foi encerrado pois verificou-se que a variação do erro quadrático médio de uma época para outra assumiu um valor menor que 10^{-3} .

Os parâmetros dos modelos periódicos autoregressivos (PAR) foram estimados através do método de máxima verossimilhança proposto por [15].

As vazões previstas um passo à frente e doze passos à frente, para os três modelos ajustados, são ilustradas nas Figuras 2 e 3, respectivamente.

As Tabelas II e III mostram os erros globais das previsões um e doze passos à frente dos três modelos.

TABELA II
ERROS GLOBAIS DA PREVISÃO UM PASSO À FRENTE.

Modelo	REQM (m^3/s)	EAM (m^3/s)	ERM (%)	ρ
FIR	173,12	104,88	18,47	0,90
MLP	198,84	126,97	23,50	0,87
PAR	212,98	131,99	25,01	0,85

Para a previsão um passo à frente (Tabela II), observa-se que a rede FIR apresenta um melhor desempenho em relação a todas as estatísticas. O erro relativo médio percentual foi, aproximadamente, 5% menor em relação à rede MLP e 7% menor em relação ao modelo PAR. O desempenho da rede neural MLP foi melhor que o desempenho do modelo de séries temporais.

TABELA III
ERROS GLOBAIS DA PREVISÃO DOZE PASSO À FRENTE.

Modelo	REQM (m^3/s)	EAM (m^3/s)	ERM (%)	ρ
FIR	220,43	153,25	29,90	0,83
MLP	232,20	172,86	37,30	0,82
PAR	242,77	165,22	38,20	0,81

A Tabela III mostra que, os erros do modelo FIR são menores que os erros dos modelos MLP e PAR. Assim, o modelo recorrente FIR apresenta um desempenho superior aos demais modelos para previsão doze passos à frente, com erros relativos médios percentuais, aproximadamente, 7% menores que os demais modelos.

O coeficiente de correlação em ambas as Tabelas II e III, não apresentam diferenças significativas, indicando que os modelos tendem a um valor de vazão médio mensal, tornando mais acentuada esta tendência ao valor médio, quando o horizonte de previsão aumenta (Tabela III).

Nota-se que, embora os modelos consigam, em geral, seguir a tendência da curva de vazões, a rede estática multicamadas (MLP) e o modelo PAR têm mais dificuldades para atingir os valores dos picos da série, quando comparado ao modelo FIR. Ou seja, é mais difícil prever as vazões nos períodos úmidos, uma vez que os picos não apresentam freqüências similares, embora as séries de vazões sejam consideradas periódicas. Isto pode ser claramente observado nas Figuras 2 e 3, principalmente na previsão doze passos à frente, no segundo e terceiro pico da curva de vazão.

Em termos gerais, a rede FIR apresenta um bom desempenho na previsão um e doze passos à frente. Os erros globais para a previsão doze passos à frente foram maiores quando comparados aos obtidos na previsão um passo à frente, por ser este

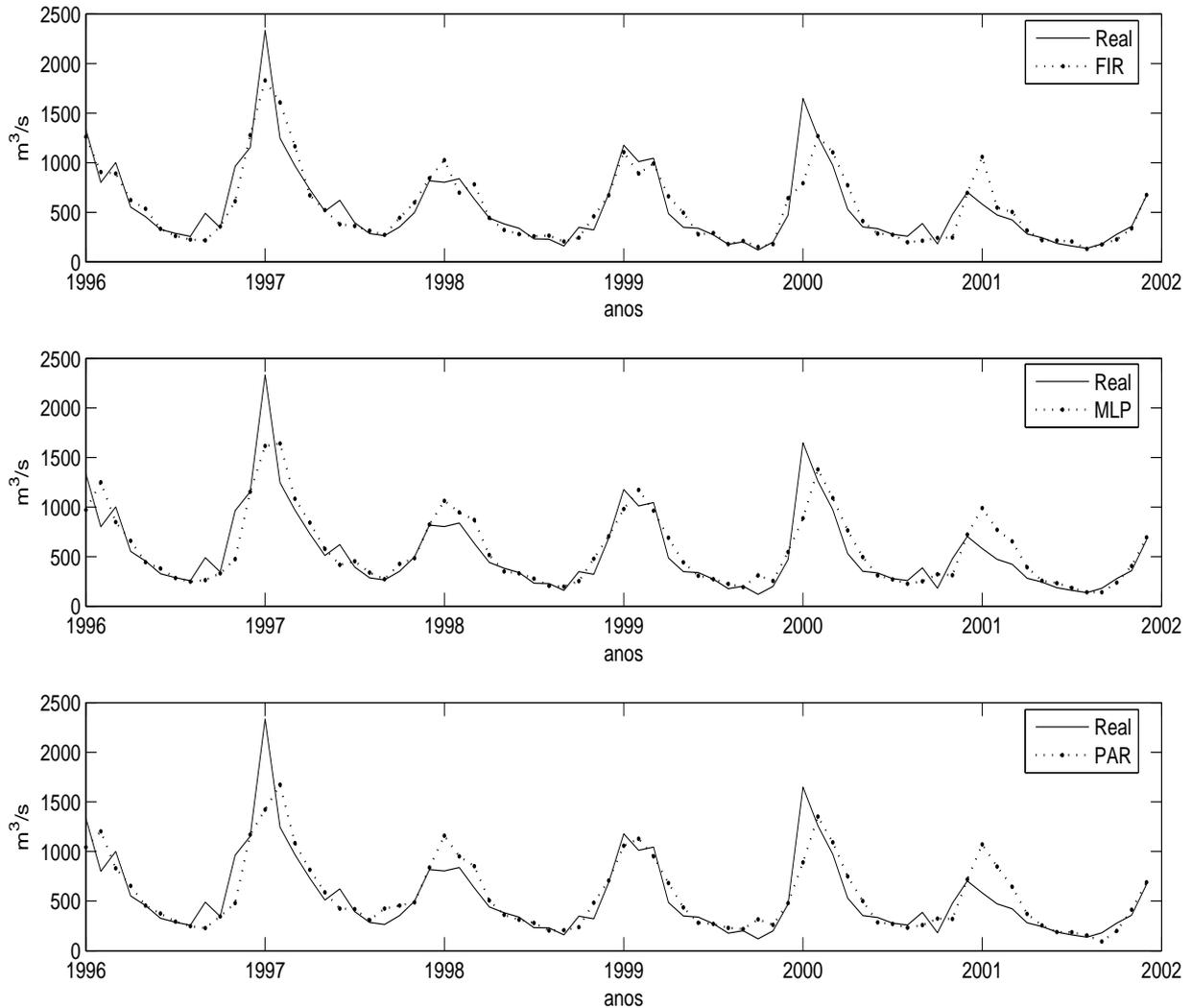


Figura 2. Previsão um passo à frente para a usina de Furnas, período de 1996 a 2001: modelos FIR, MLP e PAR.

tipo de previsão uma tarefa mais difícil, pois ao invés de utilizar valores do histórico para calcular a previsão seguinte, utiliza-se os valores previstos. Já o desempenho da rede MLP e do modelo PAR na previsão doze passos à frente, viu-se afetado pela propagação do erro, tanto na previsão nos períodos úmidos, como nos períodos secos, principalmente no terceiro e quarto período seco, como ilustra a Figura 3.

V. CONCLUSÕES

Neste trabalho, uma rede neural dinâmica foi utilizada como modelo de previsão de vazões, um e doze passos à frente. A rede neural dinâmica possui uma topologia onde os pesos sinápticos são substituídos por um filtro linear do tipo FIR, sendo estes vetores de pesos ajustados via o algoritmo de retropropagação temporal. O desempenho do modelo proposto foi comparado a outros dois modelos clássicos da literatura, uma rede neural multicamadas com algoritmo de retropropagação do erro e um

modelo periódico autoregressivo.

O modelo neural dinâmico do tipo FIR apresentou um melhor desempenho, quando comparado aos outros dois modelos de previsão, obtendo melhores resultados principalmente em picos e na previsão doze passos à frente. O desempenho deste modelo é devido a sua habilidade para lidar com problemas não-lineares e as suas características de processamento dinâmico e temporal. A memória embutida na sua estrutura fornece ao modelo a capacidade de ser, em geral, um modelo de previsão satisfatório, tanto na previsão de vazões um passo à frente como na previsão doze passos à frente.

AGRADECIMENTOS

Os autores agradecem à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Financiadora de Estudos e Projetos (FINEP), pelo auxílio.

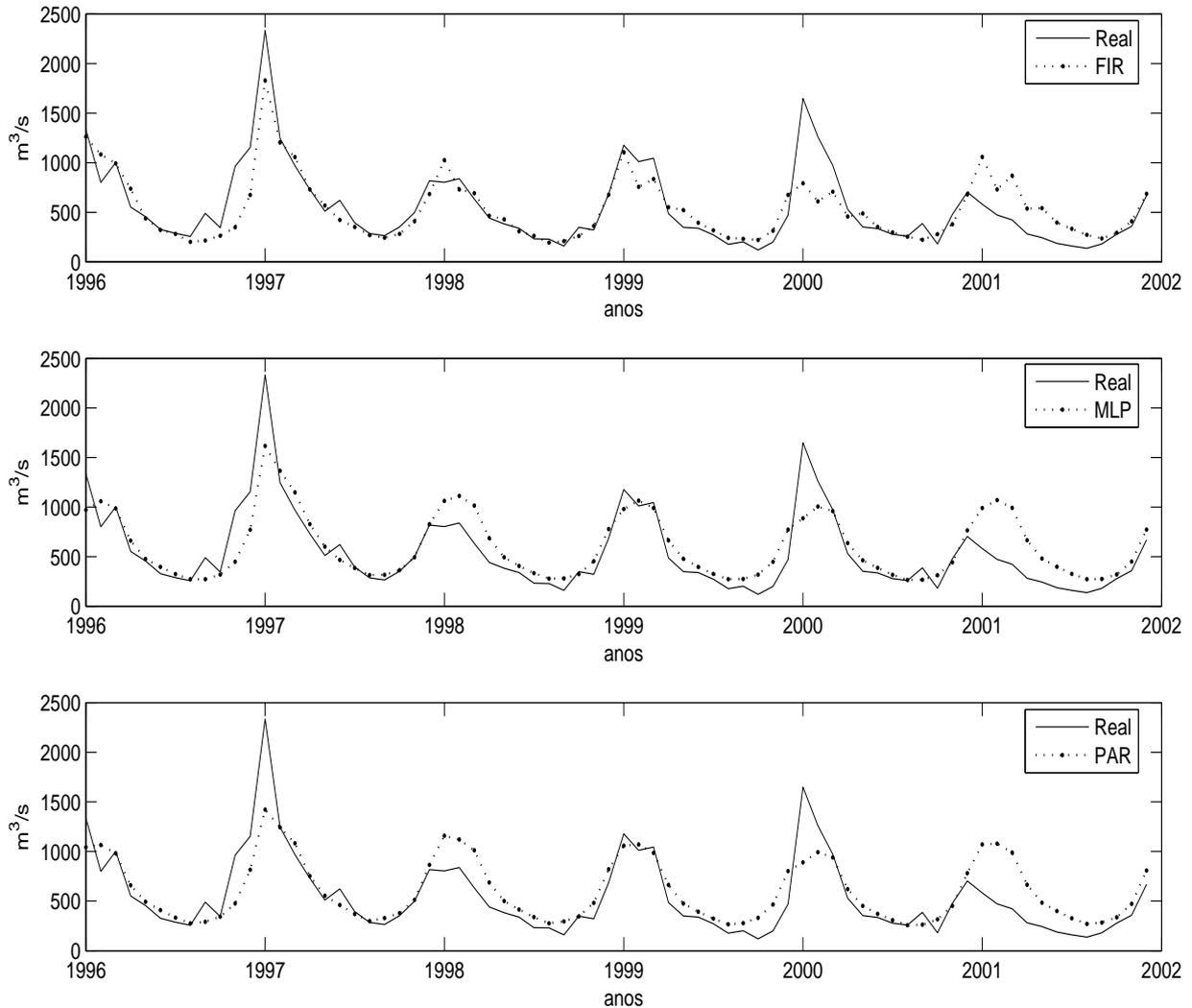


Figura 3. Previsão 12 passos à frente para a usina de Furnas, período de 1996 a 2001: modelos FIR, MLP e PAR.

REFERÊNCIAS

- [1] H. Maier and G. Dandy, "Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications," *Environmental Modelling & Software*, vol. 15, pp. 101–124, 2000.
- [2] G. Box, G. Jenkins, and G. C. Reinsel, *Time Series Analysis, Forecasting and Control*, 3rd ed. Oakland, California, EUA: Holden Day, 1994.
- [3] A. S. Weigend and N. A. Gerstenfeld, *Time Series Prediction: Forecasting the Future and Understanding the Past*, 1992.
- [4] C. Slim and A. Trabelsi, "Neural network for modeling financial time series: A new approach," in *ICCSA (3)*, 2003, pp. 236–245.
- [5] A. Lapedes and R. Farber, "Nonlinear Signal Processing Using Neural Networks: Prediction and System Modelling," Los Alamos National Laboratory, Tech. Rep. LA-Ur-&-2662, 1987.
- [6] H. J. Kim, W. D. Lee, and H. S. Yang, "A Modified FIR Network for Time Series Prediction," in *Proceedings of the 9th International Conference on Neural Information Processing*, vol. 5, 2002, pp. 2597–2600.
- [7] C. Ku and K. Lee, "Diagonal recurrent neural networks for dynamic systems control," *IEEE Transactions on Neural Networks*, vol. 6, no. 1, pp. 144–156, 1995.
- [8] F. J. Von Zuben, *Modelos Paramétricos e Não-Paramétricos de Redes Neurais Artificiais e Aplicações*. FEEC-UNICAMP, Brasil: Tese de Doutorado, Fevereiro 1996.
- [9] E. Wan, "Temporal backpropagation for FIR neural networks," *International Joint Conference on Neural Networks*, vol. 1, pp. 575–580, June 1990.
- [10] R. Ballini, *Análise e Previsão de Vazões Utilizando Modelos de Série Temporais, Redes Neurais e Redes Neurais Nebulosas*. FEEC-UNICAMP, Brasil: Tese de Doutorado, Setembro 2000.
- [11] C. M. Zealand, D. H. Burn, and S. P. Simonovic, "Short term streamflow forecasting using artificial neural networks," *Journal of Hydrology*, no. 214, pp. 32–48, 1999.
- [12] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Prentice Hall, Inc., 2001.
- [13] Li-Feng Wang and Xheng-Xi Li, "Model Identification of Time Delay Nonlinear System with FIR Neural Network," *Proceedings of the Second International Conference on Machine Learning and Cybernetics*, vol. 2, pp. 872–875, November 2003.
- [14] G. Schwarz, "Estimating the Dimension of a Model," *The Annual of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.
- [15] A. Vecchia, "Maximum likelihood estimation for periodic autoregressive-moving average models," *Technometrics*, vol. 27, no. 4, pp. 375–384, 1985.