

Emprego de modelo regressivo linear e rede neural artificial no controle da segurança de barragens de concreto

Romulo A. C. Lopes,⁽¹⁾ Eurípedes P. dos Santos,⁽²⁾ Guilherme J. F. Salazar,⁽³⁾
Maurício F. Gomes,⁽⁴⁾ Luis A. C. M. Veloso⁽⁵⁾

Trabalho recebido em 30/1/2009 e aprovado para publicação em 27/3/2009.

Resumo

Este trabalho apresenta um estudo sobre a modelagem empírica da série temporal de deslocamentos na direção do fluxo de blocos de concreto de barragem da UHE de Tucuruí, maior hidroelétrica genuinamente brasileira. São empregados modelos regressivos lineares nos parâmetros e redes neurais artificiais. Ambos os modelos mostram-se promissores, sendo as previsões obtidas por meio da rede neural mais condizentes com os valores observados. No artigo é descrita a metodologia empregada na obtenção de ambos os modelos, seguida da análise de resultados.

Palavras-chave: Redes neurais artificiais. Modelos regressivos lineares. Séries temporais. Barragens de concreto. Segurança estrutural.

⁽¹⁾ Graduando - UFPa - Faculdade de Engenharia Civil - lopes8007@yahoo.com.br

⁽²⁾ Professor - UFPa - Faculdade de Engenharia da Computação - epsantos@ufpa.br

⁽³⁾ Engenheiro - UFPa - Faculdade de Engenharia Civil - guisal@gmail.com

⁽⁴⁾ Mestrando - UFPa - Faculdade de Engenharia Civil - mauriciofg@ufpa.br

⁽⁵⁾ Professor Doutor - Faculdade de Engenharia Civil - lveloso@ufpa.br

1 Introdução

Grandes barragens necessitam de um monitoramento constante, pois avarias em sua estrutura podem representar riscos para o bom funcionamento da construção. A ruptura de uma grande barragem pode ocasionar a perda de inúmeras vidas, além de causar sérios impactos ambientais e sociais. Em virtude disso, técnicas rigorosas de auscultação devem ser empregadas.

A auscultação permite a identificação antecipada de anomalias que possam comprometer a integridade da estrutura. Se algum comportamento anormal for detectado, medidas corretivas devem ser tomadas, se necessárias, para sanar o problema.

Na engenharia de barragens o sistema é constituído pelo conjunto barragem, fundação e reservatório. Em termos de comportamento estrutural, a avaliação das condições de segurança é efetuada pela comparação das leituras coletadas com os valores estimados com base em modelos de previsão de comportamento (Da Mata 2007).

O presente trabalho utiliza modelos regressivos lineares nos parâmetros e redes neurais artificiais para a previsão do deslocamento de blocos de concreto da UHE Tucuruí, maior hidroelétrica genuinamente brasileira. De posse de um modelo de previsão de comportamento, os técnicos de Tucuruí poderão identificar leituras que fogem da normalidade da série histórica de dados. Assim, poderão averiguar se houve mudanças no comportamento estrutural da barragem, erro de leitura ou falha de algum instrumento de medição.

Procedimentos semelhantes aos abordados neste trabalho também podem ser empregados para previsão de outras grandezas de interesse na barragem, tais como tensões, deformações e subpressões.

2 Auscultação

A auscultação é uma atividade fundamental para avaliar o desempenho e a segurança das barragens ao longo de sua vida útil. Compreende inspeções visuais periódicas, monitoramento por instrumentos instalados nas suas estruturas e fundações e a modelagem dos dados com vistas a identificar possíveis irregularidades.

A instrumentação eficiente de uma barragem requer a elaboração de um plano de instrumentação. A calibração dos instrumentos deve ser supervisionada, tomando-se especial cuidado com a sua instalação e manutenção. Também são importantes as tarefas de aquisição, processamento e apresentação dos dados, seguidas pela comparação com valores de projeto. Vale ressaltar que o atendimento desses requisitos demanda uma equipe de profissionais bem treinados. A partir da execução criteriosa dessas ações, poderão ser definidos valores limites para as leituras, a serem empregados na detecção de uma eventual anomalia (Matos 2002).

Um importante instrumento para a auscultação e para os propósitos deste trabalho é o pêndulo direto. Os pêndulos medem deslocamentos horizontais e diferenciais

dos blocos da barragem. Sua leitura é feita em relação a uma das paredes da galeria da barragem, utilizando-se para isso um cordinômetro óptico ou eletrônico (Silveira 2003).

As tubulações que acomodam os pêndulos de uma barragem são instaladas durante o período de construção, mas a leitura dos instrumentos acontece geralmente após esse período, pois é após a construção que ocorre o enchimento do reservatório e, conseqüentemente, o deslocamento da crista da barragem.

Na auscultação o controle das medidas produzidas pela instrumentação pode ser realizado por meio de modelos estruturais ou modelos estatísticos. Este tópico recebe atenção no que segue.

3 Modelos de previsão

Em uma primeira abordagem, para efeitos de modelagem de barragens, os modelos de previsão de comportamento podem ser estruturais ou estatísticos (De Sotis and Paoliani 2007). Modelos estruturais são elaborados com base em princípios físicos; requerem conhecimento refinado e abrangente das possíveis relações de causa e efeito presentes nos fenômenos envolvidos e, em geral, tomam a forma de equações integro-diferenciais. Possuem como vantagens a exatidão e a possibilidade de interpretação direta de seus parâmetros; o lado desfavorável está no tempo de desenvolvimento e custos envolvidos, principalmente no caso de estruturas de grande porte. Já os modelos estatísticos ou empíricos não decorrem da aplicação de leis da física, embora possam utilizar conhecimento experimental *a priori* como auxílio na definição de dependências funcionais específicas para as relações de causa e efeito subtendidas. Apresentam o aspecto favorável de demandar menor tempo e custo de desenvolvimento, porém tendem a oferecer menor exatidão e seus parâmetros não são passíveis de interpretação direta.

Neste trabalho são empregados dois modelos empíricos, um paramétrico e um não paramétrico. O modelo paramétrico é do tipo regressivo linear nos parâmetros. No que toca ao modelo não paramétrico, foi empregada uma rede neural artificial do tipo perceptron multicamadas. Este modelo é não linear e não paramétrico por conter elementos não lineares em sua estrutura e pela grande quantidade de parâmetros ajustáveis, respectivamente, conferindo-lhe grande flexibilidade.

Ambos os modelos foram ajustados com base em séries históricas. O deslocamento do bloco TA-2 na direção do fluxo foi tomado como variável de resposta. As variáveis independentes são representadas pela série de nível de montante, a série do número de dias decorridos do início do ano até a data de observação no mesmo ano e a série do número de dias decorridos desde a data da primeira observação.

3.1 Modelo de previsão do comportamento com base em modelo linear nos parâmetros

Observando-se a série temporal de deslocamentos, torna-se perceptível um comportamento sazonal. Na verdade, pode-se dizer que esse comportamento é típico em praticamente todas as barragens de concreto de grande porte.

Além disso, considerando que o deslocamento da crista do bloco seja influenciado pelo nível de montante, pela temperatura ambiente e por parâmetros que levem em conta a fluência do concreto e das rochas da fundação, a função adotada para o modelo tem a forma:

$$U(h_i, T_i, F_i, \Delta_i) = U_h(h_i) + U_T(T_i) + U_F(F_i) + U_\Delta(\Delta_i) + k + \varepsilon_i \quad (1)$$

Nesta equação $U(h_i, T_i, F_i, \Delta_i)$ é o deslocamento observado no tempo i , que depende do nível de montante (h), da temperatura (T), da fluência dos materiais (F) e da variação temporal (Δ), considerados no tempo i . O termo $U_x(x_i)$ representa a contribuição da variável x para o deslocamento no tempo i . Por sua vez, k é o termo constante na regressão e ε_i é o resíduo dado pela diferença entre o valor observado e o valor calculado.

A contribuição do nível de montante para o deslocamento da barragem é costumeiramente representada por meio de uma função polinomial (Da Mata 2007) da forma:

$$U_h(h_i) = a_1 h^8 + a_2 h^7 + a_3 h^6 + a_4 h^5 + a_5 h^4 + a_6 h^3 + a_7 h^2 + a_8 h^1 \quad (2)$$

onde h_i é o nível de montante em metros observado no tempo i e $a_1, a_2, a_3, \dots, a_8$ são parâmetros ajustáveis.

A influência anual dos efeitos térmicos sobre o deslocamento do bloco pode ser representada pela somatória de funções harmônicas de periodicidade anual (Da Mata 2007). A função térmica anual pode ser dada por:

$$U_T(\theta) = b_1 \cos(\theta) + b_2 \sin(\theta) + b_3 \sin^2(\theta) + b_4 \cos(\theta) \sin(\theta) \quad (3)$$

$$\theta = \frac{2\pi t_d}{365}, \quad 1 \leq t_d \leq 365 \quad (4)$$

onde t_d é o número de dias decorridos do início do ano até a data de observação no mesmo ano e b_1, b_2, b_3, b_4 são parâmetros ajustáveis.

Segundo (Guedes & De Faria 2007), o efeito da fluência ao longo do tempo pode ser representado por meio dos elementos de uma série de Dirichlet, que tem a forma:

$$U_F(t) = c_1 [1 - \exp(-10^{-3}t)] + c_2 [1 - \exp(-10^{-4}t)] \quad (5)$$

onde c_1 e c_2 são parâmetros ajustáveis e t representa o número de dias decorridos desde a data da primeira observação.

O efeito das variações temporais sobre o deslocamento do bloco é representado por:

$$U_{\Delta}(\Delta_t) = d_1\Delta_1(t) + d_2\Delta_2(t) \quad (6)$$

sendo Δ_1 e Δ_2 as funções degrau unitário definidas por:

$$\Delta_1 = \begin{cases} 0, & t < 2545 \\ 1, & t \geq 2545 \end{cases} \quad \Delta_2 = \begin{cases} 0, & t < 5202 \\ 1, & t \geq 5202 \end{cases} \quad (7)$$

Os tempos de transição nas funções Δ_1 e Δ_2 decorrem da constatação visual de que houve duas variações temporais nítidas na série de deslocamentos, uma a partir de 13/5/1991 e outra a partir 21/8/98. Na Eq. (6) os parâmetros d_1 e d_2 são ajustáveis.

No presente estudo o ajuste dos parâmetros definidos na Eq. (1) é feito por meio do método dos mínimos quadrados (Vuolo 1996).

4 Modelo de previsão do comportamento com base em redes neurais artificiais

Redes neurais artificiais são modelos computacionais inspirados na constituição e organização do sistema nervoso biológico (Haykin 1998). A Figura 1(a) ilustra de maneira simplificada um neurônio biológico. A unidade funcional de uma rede neural artificial é o neurônio artificial. A Figura 1(b) ilustra o modelo de neurônio artificial mais difundido na literatura conexionista, idealizado por McCulloch e Pitts (Haykin 1998) e definido pelas seguintes equações:

$$y = f(s) \quad (8)$$

$$s = \sum_{i=1}^m w_i x_i \quad (9)$$

Neste modelo as entradas $x_1, x_2, \dots, x_m$ exercem o papel dos dendritos. A ligação com o corpo celular artificial é feita por meio dos pesos $w_1, w_2, \dots, w_m$, que representam as sinapses. A resposta do neurônio artificial é determinada pela aplicação da função de transferência ou de ativação $f(\cdot)$ à soma ponderada das entradas, s .

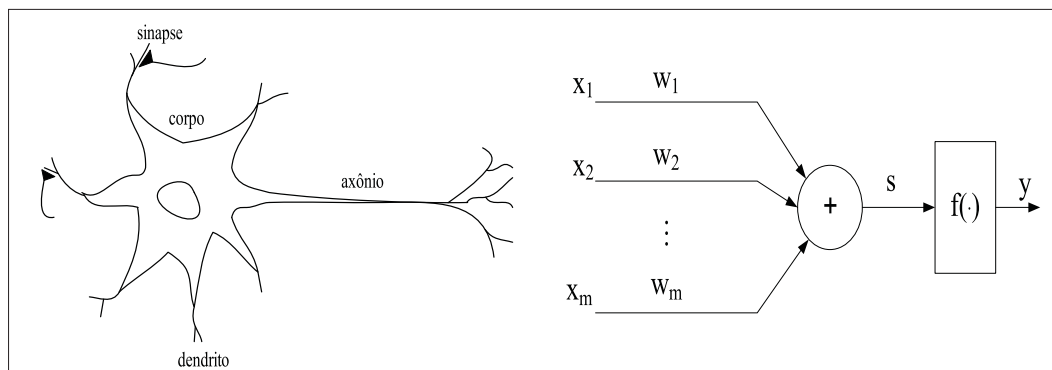


Figura 1: (a) neurônio biológico (b) neurônio artificial

As funções de ativação mais comumente encontradas na literatura são a sigmoide, a função de base radial, a tangente hiperbólica e a linear. As duas últimas são empregadas no presente trabalho, sendo definidas, respectivamente, pelas equações abaixo:

$$f(s) = \tanh(s) = \frac{\exp(s) - \exp(-s)}{\exp(s) + \exp(-s)} \quad (10)$$

$$f(s) = s \quad (11)$$

Uma rede neural artificial é formada pela interconexão de um conjunto de vários neurônios artificiais. O conhecimento adquirido na fase de treinamento da rede é armazenado nos pesos das ligações entre os neurônios. As redes neurais são classificadas de acordo com a arquitetura e o algoritmo de aprendizado que usam. Para este trabalho foram usadas redes neurais diretas do tipo perceptron multicamadas (multilayer perceptron – MLP) treinadas pelo algoritmo de retropropagação do erro (*error backpropagation algorithm*).

4.1 Redes neurais artificiais do tipo perceptron multicamadas

A Figura 2 ilustra uma rede MLP formada por duas entradas externas, uma camada oculta contendo dois neurônios e uma camada de saída contendo dois neurônios. Nesta figura cada círculo representa um neurônio segundo o modelo de McCulloch e Pitt.

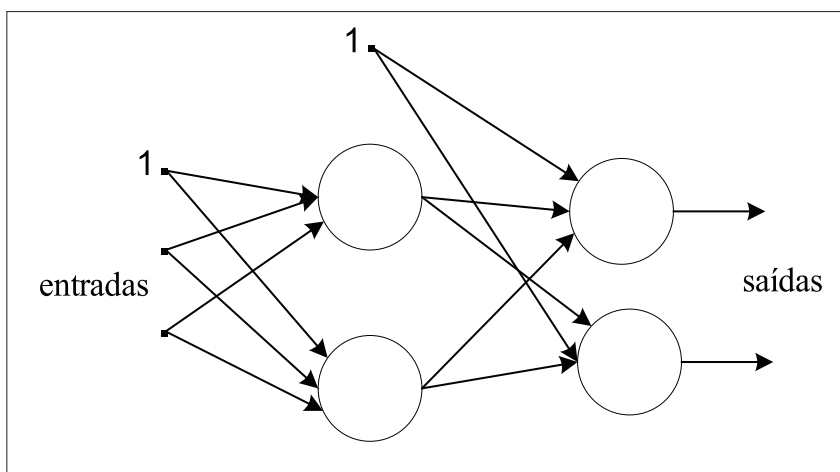


Figura 2: Rede MLP com uma camada oculta

Pode-se dizer que a grande atenção que Arquitetura MLP recebe em estudos teóricos e em aplicações práticas deve-se a dois fatores. O primeiro é o teorema de aproximação universal, segundo o qual uma rede MLP contendo pelo menos uma camada oculta e com neurônios ocultos em número suficiente é capaz de aproximar, com precisão arbitrária qualquer mapa contínuo, se forem fornecidos padrões em quantidade e qualidade apropriadas (Hush and Horne 1993). Em segundo, a existência do algoritmo de retropropagação do erro para o ajuste de seus pesos.

Neste trabalho é empregada uma arquitetura MLP de uma única camada oculta. Os neurônios nesta camada possuem função de ativação do tipo tangente hiperbólica e o neurônio de saída possui função de ativação linear.

4.2 A aprendizagem em redes MLP e o algoritmo de retropropagação do erro

Dito de maneira simples, a aprendizagem em redes neurais artificiais pode ocorrer na forma supervisionada ou não supervisionada. No primeiro caso está disponível um conjunto de exemplos ou padrões, denominado de “conjunto de treinamento”, na forma de pares entrada-saída, $(\mathbf{x}^n, \mathbf{d}^n)$, $n = 1 \dots N$, onde N é número total de padrões disponíveis. A aprendizagem supervisionada da rede neural consiste em ajustar seus pesos com base no conjunto de treinamento, de modo a minimizar o erro entre a saída da rede, $\mathbf{y}^n$, e a saída desejada, $\mathbf{d}^n$. Como resultado desse processo, deseja-se que a rede neural generalize bem, ou seja, que também produza saídas apropriadas para novas entradas, que respeitem a distribuição de probabilidades que caracteriza o conjunto de treinamento.

No caso da aprendizagem não supervisionada, os padrões de treinamento não estão acompanhados de saídas desejadas e o objetivo está em formar partições ou agrupamentos no espaço de entradas, possibilitando a categorização de padrões (Haykin

1998). Neste trabalho o treinamento da rede neural é realizado por meio da aprendizagem supervisionada.

Durante a fase de treinamento sob supervisão, o desempenho da rede neural é avaliado com base em uma função de custo, comumente definida como sendo o erro quadrático médio, dado por:

$$E = \frac{1}{N} \sum_{n=1}^n E^n \quad (12)$$

$$E^n = \frac{1}{2} (e^n)^2 = (d^n - y^n)^2 \quad (13)$$

Nestas equações considera-se o caso de uma única saída, em concordância com o problema tratado neste trabalho. O caso mais geral, saída multivariável, pode ser tratado de forma similar (Haykin 1998). O ajuste iterativo dos pesos da rede de modo a minimizar E pode ser feito com base na retropropagação do erro, conforme descrição a seguir.

4.3 O algoritmo de retro-propagação do erro

Possivelmente, este é o algoritmo de aprendizagem supervisionada mais empregado no treinamento de rede neurais artificiais, que oferece um meio para o cálculo do gradiente da função de custo definida na Equação (12) em relação aos pesos da rede. Considere-se uma rede neural contendo H neurônios na camada oculta e um único neurônio na camada de saída, vide Figura 2. Seja w_{ji}^1 o peso ligando a i -ésima entrada ao j -ésimo neurônio oculto e w_p^2 o peso ligando o p -ésimo neurônio oculto ao único neurônio de saída. O ajuste dos pesos da rede neural é comumente realizado de acordo com o método da descida em gradiente (steepest descent), segundo as equações:

$$w_p^2(k+1) = w_p^2(k) + \alpha \frac{\partial E^n}{\partial w_p^2(k)} \quad (14)$$

$$w_{ji}^1(k+1) = w_{ji}^1(k) + \alpha \frac{\partial E^n}{\partial w_{ji}^1(k)} \quad (15)$$

onde k define a iteração atual, o parâmetro α é o passo ou taxa de aprendizagem e o índice n refere-se a um dado padrão do conjunto de treinamento. As expressões que permitem o cálculo das derivadas parciais indicadas nas equações acima são calculadas por meio do algoritmo de retropropagação do erro (Haykin 1998).

Nas Eq. (14) e Eq. (15) os pesos são ajustados com base no gradiente parcial, calculado a partir Eq. (13). Por isso, esse modo de ajuste é chamado de *on-line* ou sequencial. Há também o modo de ajuste em batelada (*batch*) no qual os pesos são atualizados com base no gradiente acumulado, calculado a partir da Eq. (12).

Neste trabalho o ajuste dos pesos da rede neural foi feito no modo batelada com base em um método da classe Quase-Newton (Battiti 1992), por oferecer taxa de convergência superior ao método do gradiente.

4.4 Método de parada antecipada

Em virtude de as redes neurais, tais como as arquiteturas perceptron multicamadas, terem quase sempre um espaço de parâmetros muito grande, tais arquiteturas são suscetíveis ao sobreajuste (*overfitting*). Em algum ponto durante o treinamento, enquanto o erro sobre o conjunto de treinamento continua decrescendo, a capacidade de generalização começa a sofrer degradação, isto é, o erro para exemplos não experimentados pela rede aumenta (Bishop 1995).

Uma das técnicas mais usadas para lidar com o sobreajuste é a parada antecipada. Esta técnica distribui os padrões disponíveis em três conjuntos de dados: treino, teste e validação. O conjunto de validação é usado para verificar a capacidade de generalização da rede ao longo do treinamento. Basicamente, a parada antecipada tenta detectar o início do sobreajuste e, ao encontrá-lo, interrompe o treinamento, o que é feito com base no indício de aumento contínuo do erro de validação (Prechelt 1998).

O treinamento com parada antecipada é empregado neste trabalho em conjunto com redes MLP. O erro de validação é avaliado a cada época ao longo de todo o processo de ajuste dos pesos. A solução adotada é aquela que corresponde ao mínimo do erro de validação. Uma época compreende a propagação dos padrões na rede, o cálculo do gradiente e o ajuste dos pesos em batelada.

5 Desempenho dos modelos de previsão e comparações

A série histórica analisada corresponde aos deslocamentos medidos pelo pêndulo direto instalado no bloco TA-2 da UHE Tucuruí, no período compreendido entre o enchimento do reservatório, realizado no mês de maio de 1984, até o mês de abril do ano de 2007, resultando num total de 384 leituras. Para teste dos modelos de previsão foram guardadas à parte medições correspondentes ao período de maio de 2007 a abril de 2008, num total de 13 observações. As amostras restantes foram empregadas no ajuste dos modelos. Inicialmente, são apresentados os resultados associados com o modelo regressivo linear nos parâmetros, conforme a Eq. (1), seguidos dos resultados obtidos com base em redes MLP. Por último é feita uma análise dos resultados.

Os parâmetros definidos na Eq. (1) foram determinados pelo método dos mínimos quadrados. Abaixo está a equação obtida:

$$\begin{aligned}
 U_D = & 0,3150 + 3,878.H^8 - 0,216.\cos\left(\frac{2\pi t}{365}\right) - 1,2359.\sin\left(\frac{2\pi t}{365}\right) \\
 & - 0,36184.\cos\left(\frac{2\pi t}{365}\right)\sin\left(\frac{2\pi t}{365}\right) - 0,14659.\sin^2\left(\frac{2\pi t}{365}\right) - 7,3627[1 - \exp(-10^{-3}t)] \\
 & - 14,5104[1 - \exp(-10^{-4}t)] - 0,000978\Delta(t - 2545) - 0,00073942\Delta(t - 5202)
 \end{aligned} \quad (16)$$

Na Figura 3 pode-se observar a série temporal dos deslocamentos medidos e os calculados com base na Eq. (16). Constata-se um bom ajuste entre os dados medidos e as previsões, principalmente se forem considerados aspectos qualitativos associados com tendências e sazonalidades. O modelo também foi capaz de acompanhar o perfil dos deslocamentos sofridos pela barragem quando referentes ao conjunto de dados reservado para teste, compreendendo o período de 16/4/2007 até 14/4/2008.

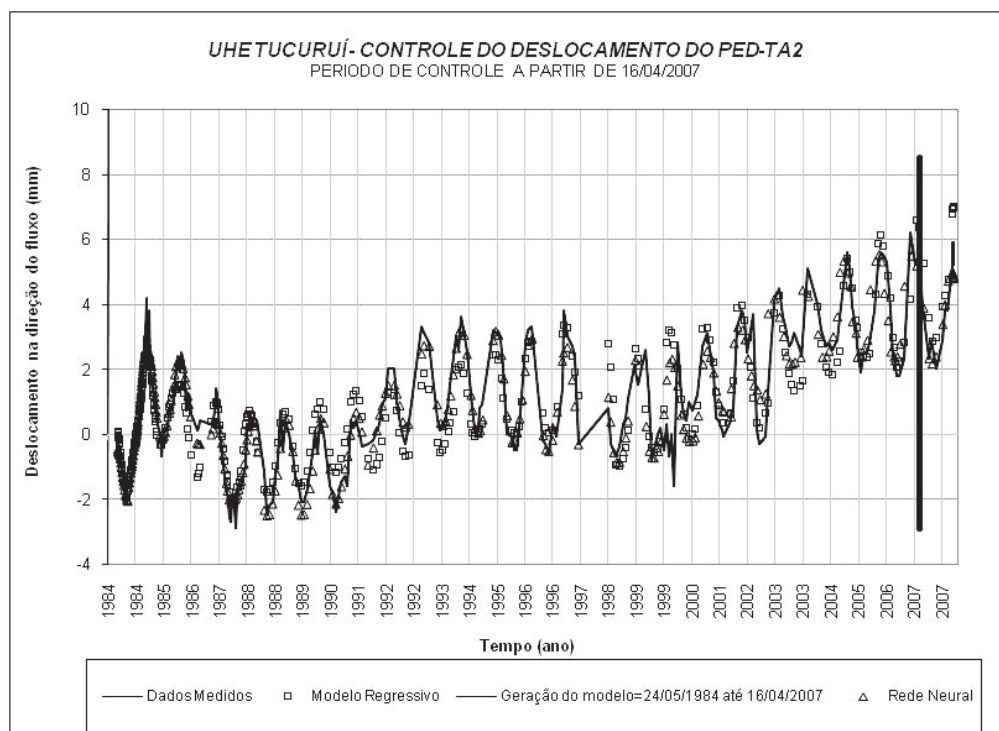


Figura 3: Série temporal de deslocamentos do bloco TA-2, valores medidos e previsão

No treinamento das redes MLP's foram investigadas arquiteturas com número de neurônios na camada oculta variando de 1 até 15. Para cada arquitetura foram treinadas redes a partir de dez condições iniciais diferentes, dado que é comum a presença de múltiplos ótimos locais nas superfícies de erro. As entradas tomadas para a rede neural foram: nível de montante, variação térmica anual representada pelas funções $\text{sen}(\theta)$ e $\text{cos}(\theta)$, com θ definido na Eq. (4), e a contribuição da fluência na forma das duas funções exponenciais que aparecem na Eq. (5). Em cada seção de treinamento foram permitidas no máximo 500 épocas. Os dados disponíveis para ajuste dos modelos foram distribuídos de modo aleatório em dois conjuntos, um de treinamento e outro de validação, nas proporções de 80% e 20%, respectivamente. As variáveis foram escalonadas linearmente para o intervalo [-1,1]. Considerando as arquiteturas com 15 neurônios ocultos, as maiores investigadas, o tempo médio de treinamento de cada uma ficou em torno de 5 s. Os programas foram executados em uma máquina Pentium 4, 3GHz e 1Gb de memória RAM, rodando Matlab 7.0. A arquitetura de rede neural escolhida, com base no erro de validação, possui cinco neurônios na camada oculta.

As previsões obtidas por meio da rede neural podem ser vistas na Figura 3. Valem as mesmas considerações feitas com relação ao desempenho do modelo regressivo linear nos parâmetros. Porém, pode-se observar que a curva de previsão produzida pela rede neural é mais suave e mais próxima dos valores observados.

Para facilitar a análise visual das curvas de previsão produzidas por ambos os modelos, estas são exibidas na Figura 4 em um horizonte de tempo reduzido. Nela se pode observar que de fato as previsões obtidas por meio da rede neural são superiores às do modelo regressivo linear nos parâmetros, tanto no período de treinamento quanto no período de teste.

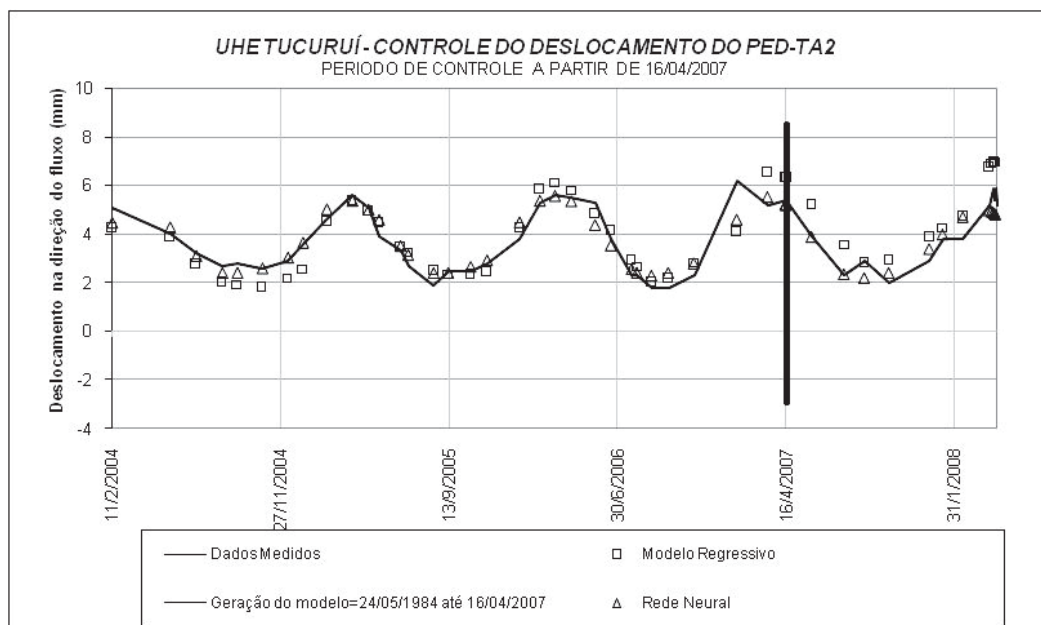


Figura 4: Série temporal de deslocamentos do bloco TA-2, valores medidos e previsão

Para fornecer uma medida quantitativa dos erros associados com cada modelo, calculou-se a raiz do erro quadrático médio (*root mean square error* – RMSE) para cada um deles, para ambos os conjuntos, de treinamento e teste. Os valores são mostrados na Tabela 1. Os números indicam que o erro da rede neural sobre o conjunto de teste foi, em média, 50% inferior ao do modelo regressivo. Tal melhoria pode ser atribuída ao controle do sobreajuste por meio da parada antecipada e ao fato de a rede neural poder capturar outras relações não lineares não contempladas pela parametrização adotada no modelo regressivo.

Tabela 1: RMSE para os modelos

	RMSE	
	Treinamento	Teste
Modelo regressivo	0,8181	1,2199
rede neural	0,7494	0,6017

6 Conclusões

Os resultados obtidos indicam a possibilidade de serem realizadas previsões do deslocamento na direção do fluxo dos blocos de concreto da barragem de Tucuruí por meio de modelo empíricos. Tais previsões são de grande importância para avaliação do comportamento da estrutura e também servem de auxílio na correção de erros de medição ou mesmo na substituição de medidas faltantes.

Modelos regressivos semelhantes ao utilizado neste trabalho vêm sendo muito empregados como ferramentas na auscultação de barragens, tendo-se como vantagens a fácil implementação e a produção de bons resultados. Por outro lado, a utilização de redes neurais com esta finalidade é bem mais recente, tendo-se como empecilhos a necessidade de conhecimento sobre aspectos da inteligência computacional para definir a arquitetura da rede. Além disso, em alguns casos a solução pode requerer um tempo de processamento elevado. No entanto, no estudo realizado, o tempo de processamento para as redes com maiores quantidades de neurônios ocultos foi da ordem de 5 segundos em microcomputadores convencionais, graças à eficiência dos algoritmos empregados.

Quanto ao estudo comparativo entre os dois tipos de modelo, verificou-se que o erro cometido nos resultados da rede neural é cerca de 50% inferior ao do modelo regressivo. Esse melhor desempenho pode ser atribuído ao controle do sobreajuste por meio da parada antecipada e ao fato de a rede neural poder capturar outras relações não lineares não contempladas pela parametrização adotada no modelo regressivo.

Agradecimentos

Os autores agradecem à ELETRONORTE, FINEP e FAPESPA o apoio financeiro dado à pesquisa.

Referências

- Battiti, R. First and second order methods for learning: between steepest descent and newton's method. *Neural Computation*, v. 4, p. 141-166, 1992.
- Bishop, C.M. Neural networks for pattern recognition. *Clarendon Press*, v. 1, 1995.
- Da Mata, J.T.C. Aplicação de redes neuronais ao controlo de segurança de barragens de Betão. Lisboa, 2007, Tese (Mestrado), Universidade técnica de Lisboa.
- De Sortis, A., Paoliani, P. Statistical analysis and structural identification in concrete dam monitoring. *Engineering Structures*, v. 29, p. 110-120, 2007.
- Guedes, Q.M., De Faria, E.F. Modelo estatístico de controle do deslocamento monitorado na barragem casca da UHE Funil. Belém, 2007, XXVII Seminário nacional de grandes barragens.
- Haykin, S. Neural Networks: A Comprehensive Foundation. Prentice-Hall, 1998.
- Hush, D.R., Horne, B.G. Progress in supervised neural networks. *IEEE Signal Processing Magazine*, p. 8-39, 1993.
- Matos, S.F. (2002). Avaliação de instrumentos para auscultação de barragem de concreto. Estudo de caso: deformímetros e tensômetros para concreto na barragem de Itaipu. Curitiba, 2002, Tese (mestrado), Universidade Federal do Paraná.
- Prechelt, L. Automatic early stopping using cross validation: Quantifying the criteria. *Neural Networks*, v. 11, p. 303-305, 1998.
- Silveira, J.F.A. Instrumentação e comportamento de fundações de barragens de concreto. São Paulo, 2003, Oficina de textos.
- Vuolo, J.H. Fundamentos da Teoria de Erros. Editora Edgard Blücher Ltda. 2. ed. São Paulo, 1996.

Applying linear regressive model and artificial neural network to safety control of concrete dams

Abstract

This work presents empirical modeling of time series collected from concrete dam pendulum displacements. The data are from Tucuruí Hydroelectric Power, the biggest Brazilian plant. Two different models were employed, a linear in the parameters regressive model and a artificial neural network. Both models exhibited good predictions, with the ones from the neural network showing better fit to the observed measurements. All details concerning the building of both models are described.

Key words: Artificial neural networks. Linear regressive models. Time series. Concrete dams. Structural safety.

Introduction

The detection of possible anomalies in large concrete dams requires the availability of prediction models. It can be achieved through the comparison of the predictions with the data collected from sensors. Examples of possible anomalies are significant deviation of time sequences from usual patterns, gross errors in the measurements and sensors failure.

This work presents the application of linear regressive models and neural networks in the prediction of displacements of concrete blocks of the UHE Tucuruí, the biggest Brazilian Hydroelectric.

Behavioral prediction models

The models employed in this work are empirical and were calibrated based on time series. The dependent variable is the crest displacement of the block TA-2 in the flow direction, and the independent variables are the upstream storage level, the environment temperature and the creep of both, concrete and rocks of the foundation. The temperature and creep effects were represented by functional relations defined on the time variable counted in days.

Prediction Model linear in the parameters

In the linear model the regressors were quantities resulting from applying specific parametric functions on the aforementioned dependent variables, vide Eq. (1). The free parameters were adjusted by the least squares method (Vuolo 1996).

Prediction model based on neural networks

We adopted a feedforward multilayer perceptron, with a single hidden layer. The hidden neurons have activation function of the type hyperbolic tangent and the output neurons linear one. The inputs to the network were the upstream storage level and a sub-set of regressors used in the linear model. Specifically, the ones associated with the environment temperature and creep effects. The training of the network was done by a Quasi-Newton method, with gradients calculated by the backpropagation algorithm. All the programs were coded in Matlab by the authors.

Multilayer perceptrons usually have a great space of adjustable parameters, making them sensible to overfitting. To overcome it, we employed the technique of early-stopping. After carried out many simulations we selected a network with 5 hidden neurons, based on validation error.

Performance of the models and comparisons

The time series studied has length 384 samples and corresponds to the displacements measured by means of the direct pendulum placed in the TA-2 block of the UHE Tucuruí. This series was used in the fitting of the models, but the last 13 samples were kept apart, to posterior assessment of the models performance.

Figures 3 to 4 show the predictions produced by both models, there is a good fitting between the predictions and the observed series, mainly regarding regular behavior, like trends and seasonality. In addition, the predictions of the neural network are more smooth and precise compared with those of the linear model. In fact, as Table 1 shows, the mean square root error evaluated for the neural network is about 50% below the one corresponding to the linear regressive model.

Conclusion

In this study good predictions of the displacements of concrete blocks of the UHE Tucuruí were obtained by means of empirical models. The ability to perform this task has great importance to the safety control of concrete dams. Also the models could be used as an auxiliary tool in the correction of measurements errors or even as virtual sensors in the replacement of damaged physical sensors.

Linear regressive models similar to the one applied in this work have been used as tools in the auscultation of dams. They have the advantages of easy implementation and useful predictions. On the other hand, the employment of neural networks in this concern is more recent. They are a powerful tool, but require from the user the knowledge of computational intelligence skills to train the neural network and to choose an appropriate architecture. To perform these requirements, it could demand a long time of computational resources. Nevertheless, in the present study the time required to train the neural networks was short, due to the efficiency of computational routines employed.

We found that the error of the neural network was about 50% below the one for the linear regressive model. This superior performance could be attributed to the control of overfitting by early stopping and to the flexibility of the neural network to capture other non-linear relations that are beyond the functions adopted as regressors for the linear model.

Acknowledgements

The authors acknowledge to the ELETRONORTE, FINEP AND FAPESPA to their support to the present research.