

Redes neurais temporais aplicadas ao monitoramento de barragens

Janaina Veiga Carvalho¹, Celso Romanel²

¹Universidade Severino Sombra, CECETEN, Curso de Sistemas de Informação e Matemática,
Vassouras - RJ, Brasil

janvcarvalho@globocom.com

²Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Civil,
Rio de Janeiro – RJ, Brasil

romanel@civ.puc-rio.br

Resumo

O monitoramento do desempenho de barragem é realizado pela revisão e análise de dados coletados pela instrumentação que medem indicadores críticos do comportamento da barragem. Neste trabalho emprega-se a técnica de redes neurais temporais (RNT) para análise, modelagem e previsão dos valores de vazão na barragem Funil, do sistema Furnas Centrais Elétricas, a partir dos dados de instrumentações disponíveis no período compreendido entre 02/09/1985 e 25/02/2002. As redes neurais temporais empregadas foram: RNT com arquitetura feedforward associada a técnica de janelamento, RNT recorrente Elman e RNT FIR. Neste trabalho estuda-se ainda a geração de intervalos de confiança para RNT. As previsões de vazão analisadas neste trabalho, envolvendo o comportamento da barragem Funil, apresentaram resultados satisfatórios.

Abstract

Performance monitoring of dams is accomplished by reviewing and analyzing data collected from instruments which measure critical indicators of engineering behavior. In this work the technique of temporal neural networks (TNN) is used for analysis, modeling and forecast of the water discharges values in the Funil dam, from Furnas Centrais Elétricas system, starting from the data of available instrumentation in the period between 02/09/1985 and 25/02/2002. The temporal neural networks used in this research were the following: TNN with feedforward architecture and the windowing technique, recursive TNN Elman and TNN FIR. In this research the generation of confidence intervals for TNN was also investigated. The discharge values forecasts analyzed in this work for the Funil dam presented satisfactory results.

Key-words: Funil Dam, temporal neural network, confidence intervals, dam monitoring, time series.

1 Introdução

Barragens são monitoradas para verificar se o desempenho está consistente com as expectativas do projeto original. O monitoramento da performance das barragens é geralmente realizado através da revisão e análise de dados coletados por instrumentos que medem indicadores críticos do comportamento da obra. Um programa de medição envolve o projeto do sistema de medição, instalação, operação, manutenção e avaliação dos instrumentos da barragem, incluindo estruturas e fundações. A instrumentação completa engloba medições de deslocamentos, deformações, tensões, pressões, cargas nos membros estruturais, percolação e drenagem junto com fatores ambientais que afetam o comportamento da barragem tais como temperatura, nível do reservatório e precipitação.

Neste artigo, redes neurais temporais são utilizadas para modelar séries temporais que consistem de vazões em drenos, medidas ao longo de vários anos de

observação contínua, através da fundação da barragem brasileira de Funil. Devido à quantidade considerável de dados de instrumentação disponíveis, a utilização de redes neurais artificiais é uma escolha bastante natural por apresentar a habilidade de previsão do desempenho do sistema de monitoramento da barragem. Então, qualquer medida de vazão que diferir significativamente do valor previsto pode ser uma indicação de uma condição anormal que requer controle imediato e cuidadoso por parte da equipe de engenharia que acompanha a operação da barragem.

1.1 Características da barragem

A barragem de Funil (Figura 1), situada próxima da cidade de Resende, no Estado do Rio de Janeiro, é uma barragem do tipo abóbada de dupla curvatura com 85 m de altura máxima e 385 m de comprimento. Completada em 1963, é monitorada por um sistema de instrumentação que mede temperatura, deslocamentos

em pontos específicos ao longo de várias galerias internas, pressão de água e vazão pela fundação da barragem.



Figura 1. Vista geral da barragem de Funil

2 Modelos para séries temporais

Medidas sucessivas de vazão de água através da fundação da barragem de Funil são modeladas utilizando redes neurais temporais.

As redes neurais temporais são redes dinâmicas que possuem propriedades de memória que lhes permitem realizar tarefas de caráter temporal. A maneira como a memória é incorporada determina o tipo de rede neural. Existem basicamente duas maneiras de se incluir habilidades de memória na rede neural: considerando entradas atrasadas no tempo ou laços de realimentação.

Neste trabalho duas redes neurais com atraso no tempo são utilizadas (rede neural com método de “janelamento” e rede neural FIR - Finite Response Impulse) bem como uma rede neural com laços de realimentação (rede neural Elman).

2.1 Pré-processamento das séries

As séries temporais passaram por uma etapa de pré-processamento de dados para substituir leituras consideradas anômalas por valores obtidos por interpolação com valores vizinhos. Os dados foram normalizados entre 0 e 1, com o objetivo de tornar a etapa de treinamento da rede neural mais eficiente.

2.2 Rede neural com método de “janelamento”

No método de “janelamento” a memória é introduzida nos neurônios da primeira camada oculta, fornecendo-se valores das entradas atuais e também em tempos anteriores. A Figura 2 mostra o método de “janelamento” em uma rede com memória de ordem 2 na primeira camada oculta. Além da entrada atual $x(k)$, os neurônios da primeira camada oculta recebem dois valores anteriores $x(k-1)$ e $x(k-2)$ criando, assim, sinapses novas.

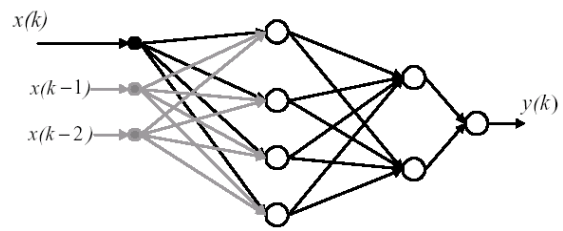


Figura 2. Método de “janelamento” [1]

A arquitetura utilizada é a feedforward multicamada, composta por uma camada de entrada, com número de neurônios dependente do número de parâmetros de entrada, por uma camada escondida, contendo uma quantidade de neurônios definida pelo algoritmo de treinamento, e por uma camada de saída com um neurônio - previsão de vazão - com função de ativação linear. Na camada escondida a função de ativação adotada foi a sigmoide logística e o algoritmo de treinamento utilizado foi a regularização bayseana associada ao método leave-one-out.

Cada rede neural foi inicializada cem vezes, com o propósito de reduzir o efeito de mínimos locais não-ótimos na superfície de erro da rede.

2.3 Rede neural FIR

O modelo da rede neural FIR é baseado na técnica de processamento de sinais. As sinapses são modeladas por um filtro linear invariante no tempo, que representa a natureza temporal do problema. Nesta modelagem a memória é considerada em todos os neurônios.

Soto [7] apresenta a formulação de uma rede neural FIR detalhadamente, considerando no modelo filtros sinápticos de tempo contínuo. A Figura 3 ilustra o modelo FIR do neurônio j com N entradas, onde $x_1(k)$, $x_2(k)$ e $x_N(k)$ são as entradas do neurônio, w_{ij} é o vetor de pesos sinápticos, $s_{ij}(k)$ é o potencial sináptico do neurônio j , $s_j(k)$ é a soma dos potenciais sinápticos, w_b representa bias e $y_j(k)$ é a saída do neurônio após passar pela função de ativação.

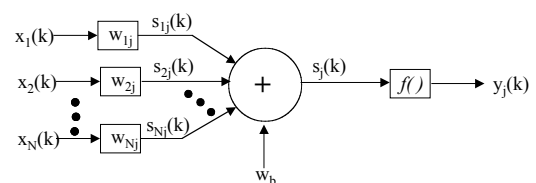


Figura 3. Modelo FIR do neurônio j [1]

Neste modelo o potencial sináptico $s_{ij}(k)$ é a resposta de um filtro FIR de ordem M , onde M indica o número total de atrasos unitários considerados no modelo do filtro FIR. Este potencial sináptico $s_{ij}(k)$ é definido por:

$$s_{ij}(k) = \sum_{n=0}^M w_{ij}(n)x_i(k-n) + w_b \quad (1)$$

onde k representa o tempo discreto.

Então a soma dos potenciais sinápticos é definida por:

$$s_j(k) = \sum_{i=1}^N \sum_{n=0}^M w_{ij}(n)x_i(k-n) + w_b \quad (2)$$

Na expressão acima, o comportamento temporal é representado pelo somatório interno e o comportamento espacial pelo somatório externo.

A Figura 4 apresenta uma rede FIR com uma camada de entrada, duas camadas escondidas e a camada de saída, enquanto que a Figura 5 mostra a ampliação das sinapses entre os neurônios da primeira camada escondida e o neurônio do meio da segunda camada escondida.

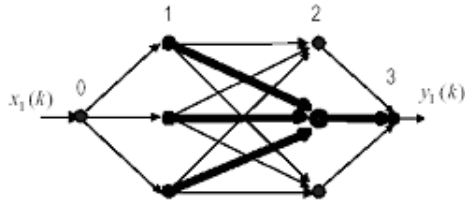


Figura 4. Rede FIR multicamada

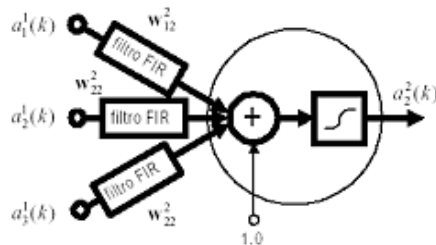


Figura 5. Ampliação do neurônio FIR

As séries são modeladas com rede neural FIR multicamada composta por uma camada de entrada, uma camada escondida (2, 3, 5, 8, 10 ou 15 neurônios) e uma camada de saída com 1 neurônio (previsão da vazão) com função de ativação linear. A função de ativação da camada escondida é a sigmoide logística e o algoritmo de treinamento utilizado é o da retropropagação temporal [10].

Cada rede neural foi inicializada somente cinco vezes, devido ao tempo de processamento bastante longo.

2.4 Rede neural Elman

Um tipo de rede neural temporal com realimentação global é a rede Elman [4]. Nesta rede cada um dos neurônios da camada oculta tem realimentação para as unidades de contexto, como ilustrado na Figura 6, onde $x(k)$ é a entrada externa no tempo discreto k , $a(k)$ é o vetor contendo as saídas dos N neurônios da camada oculta no tempo discreto k , $z(k)$ é o conjunto de entradas dos neurônios da camada oculta, denominado vetor de estados e $y(k)$ é a saída da rede no tempo discreto k .

A seguir, apresentam-se as equações para uma rede neural Elman com uma entrada e uma saída (Figura 6). As entradas da camada oculta são definidas por:

$$z_i^h(k) = \begin{cases} x_i(k) & , \quad i \in I \\ a_i(k-1) & , \quad i \in H \end{cases} \quad (3)$$

onde I representa uma camada externa e H o conjunto de saídas dos neurônios da camada oculta.

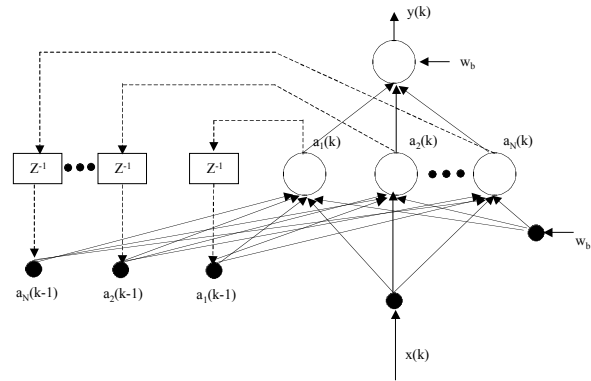


Figura 6. Esquema de rede neural Elman [1]

A saída do j -ésimo neurônio da camada oculta é dada por:

$$a_j(k) = f\left(\sum_{i \in I} w_{ij} x_i(k) + \sum_{i \in H} w_{ij} a_i(k-1) + w_{bj}\right) \quad , \quad j \in H \quad (4)$$

onde $f\left(\sum_{i \in I} w_{ij} x_i(k) + \sum_{i \in H} w_{ij} a_i(k-1) + w_{bj}\right)$ indica a função de ativação não-linear e w_{bj} o peso do viés.

Para o neurônio da camada de saída, $y(k)$ é calculado pela expressão:

$$y(k) = f\left(\sum_{i \in H} w_{io} a_i(k) + w_{bo}\right) \quad (5)$$

onde o índice “o” representa camada de saída.

As séries temporais são modeladas com uma rede Elman composta por uma camada de entrada, uma camada escondida com número de neurônios definido pelo algoritmo de treinamento, uma camada de contexto com número de neurônios igual ao número de neurônios escondidos e uma camada de saída com um neurônio (previsão de vazão) com função de ativação linear. A função de ativação da camada escondida é a sigmoide logística. O algoritmo de treinamento utilizado foi o de regularização bayseana associado ao método leave-one-out.

De maneira semelhante ao caso da rede neural com “janelamento”, cada rede neural Elman foi inicializada cem vezes.

2.5 Medidas estatísticas de erro

Há vários métodos de medidas de erro que permitem comparação do desempenho da previsão com os valores observados. Os dois mais comumente usados são o RMSE (*Root Mean-Square Error*) e o MAPE (*Mean Absolute Percentage Error*), definidos a seguir, onde \hat{f}_i é o resultado previsto, t_i o valor observado da amostra e N o número de previsões.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (f_i - t_i)^2}{N}} \quad (6)$$

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{f_i - t_i}{t_i} \right| \times 100\% \quad (7)$$

A desvantagem do RMSE é a sua sensibilidade a valores anômalos e a desvantagem da medida MAPE é que impõe uma penalidade maior nas previsões que excedem os valores atuais do que naquelas que subestimam. Visto que o RMSE e o MAPE apresentam vantagens e desvantagens, eles foram igualmente utilizados neste trabalho, admitindo-se que o menor desses valores indicará a previsão mais precisa.

Para auxiliar no julgamento do desempenho de um modelo de regressão, Theil [8] propôs um índice, conhecido na literatura como índice U-Theil,

$$\text{U - Theil} = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (f_i - t_i)^2}}{\sqrt{\frac{1}{N} \sum_{i=1}^N f_i^2 + \frac{1}{N} \sum_{i=1}^N t_i^2}} \quad (8)$$

Valores altos indicam um comportamento bastante simples da série temporal. Para casos em que o índice U-Theil for igual ou superior a 1, então o melhor resultado é simplesmente o da previsão ingênua, repetindo-se o último valor observado na série temporal.

3 Técnicas para estimar intervalos de confiança para redes neurais

Redes neurais artificiais são sistemas paralelos distribuídos formados por unidades de processamento simples (neurônios) que calculam funções matemáticas, geralmente não-linear. São utilizadas principalmente em problemas de previsão e classificação de padrões em diferentes áreas, como por exemplo, industrial, médica ou financeira.

Problemas de previsão são complexos já que as informações disponíveis do problema podem ser limitadas e o sistema pode ser incerto. A grande questão é qual a precisão da previsão. Isto é importante devido ao resultado da previsão ser normalmente utilizado na tomada de decisão. A precisão da previsão permite aos usuários da rede neural determinar a confiança da saída da rede neural. Também permite incluir a saída estimada da rede como parte de um esquema de estimação global.

O conjunto de dados de entrada utilizado na previsão geralmente é disperso e com erros de medição. Estes dados utilizados como entradas do modelo neural geram incertezas denominadas incertezas na entrada. Existe ainda o erro na saída da rede originado por ruídos na saída e pela escolha de modelos de rede imperfeitos (definição dos pesos sinápticos inadequados). Estes erros são responsáveis pelas incertezas dos pesos. A

incerteza total da previsão é a combinação das incertezas na entrada com a incerteza nos pesos.

Então a estimativa de intervalos de confiança associada a previsão aumentam a confiabilidade na rede neural. Diversos métodos para estimar os intervalos de confiança têm sido apresentados na literatura.

Chryssoulouris, Lee & Ramsey [3] desenvolveram um método para estimar intervalos de confiança baseado em um modelo para prever intervalos de confiança que considera uma distribuição normal para os erros (usando a distribuição t-student) em lugar de covariância para as saídas. Este método difere de outros métodos existentes devido não ser necessário informações sobre as segundas derivadas da saída da rede neural.

Rivals & Personnaz [6] apresentam resultados teóricos da construção de intervalos de confiança para uma regressão não linear, baseado na estimação do mínimo quadrado e utilizando a expansão linear de Taylor da correspondente saída do modelo não linear. Eles aplicam a metodologia desenvolvida em um modelo de rede neural. Um problema real é analisado e simulado. Os trabalhos mostram ainda que a expansão linear de Taylor não fornece somente um intervalo de confiança em qualquer ponto de interesse, mas também fornece uma ferramenta para detectar overfitting.

Townsend & Taransenko [9] analisam o problema de estimativa de precisão das saídas da rede neural através de um modelo de perturbação. Neste trabalho, as fontes de ruídos modelados inicialmente estão no vetor de entrada e nos pesos. O modelo de perturbação é aplicado a redes de bases de funções radiais.

Papadopoulos, Edward & Murray [5] comparam três métodos de estimativa de intervalos de confiança. Os três métodos são probabilidade máxima, aproximação bayesiana e técnica bootstrap. Os métodos são testados com problemas artificiais e problemas reais.

Alves da Silva & Moulin [1] e Alves da Silva & Moulin [2] utilizam três técnicas para cálculo de intervalos de confiança. As técnicas são: saída de erro, re-amostragem dos erros e regressão multilinear adaptada para redes neurais. O problema analisado é previsão de cargas de curto tempo. A previsão das cargas é obtida com o auxílio de redes perceptrons multi-camadas.

Zhang & Luh [11] e Zhang, Luh & Kasiviswanathan [12] estudam intervalos de confiança para a previsão gerada por uma rede neural em cascata utilizando bayesian inference framework. Este método considera ruídos nos pesos, ruídos dos dados de entrada medidos e ruídos de entradas gerados no processo de previsão. Neste método a distribuição de saída é aproximada para uma distribuição gaussiana. A variância da saída pode ser calculada pelo método metrópole ou por um método memoryless Quasi-Newton. O método memoryless Quasi-Newton é rápido e com boas características computacionais.

Neste trabalho, os intervalos de confiança serão gerados utilizando as técnicas: saída de erro e re-

amostragem dos erros. Nas seções seguintes serão apresentadas estas técnicas.

3.1 Saída de Erro

Na técnica de saída de erro, a rede neural possui duas saídas. A primeira saída corresponde a previsão da vazão e a outra saída ao erro de previsão da vazão. Deste modo, os intervalos de confiança são gerados durante o processo de previsão. A idéia proposta por Alves da Silva & Moulin [2] é que seja possível capturar possíveis padrões existentes na previsão do erro, assim como é possível também na previsão da vazão. A figura 7 mostra um exemplo de rede neural para a técnica de erro de saída.

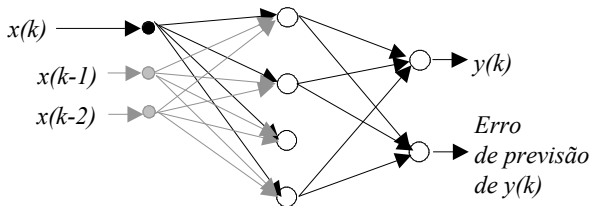


Figura 7. Rede neural para a técnica erro de saída (Alves da Silva & Moulin [2]).

O treinamento da rede neural consiste em se calcular a cada época de treinamento os padrões para o neurônio de saída do erro de previsão. Então para cada padrão utilizado para previsão da vazão, o erro da previsão da vazão é calculado. Deste modo, em cada época um padrão de treinamento diferente será usado para a previsão do erro.

O processo de treinamento deve convergir para um conjunto de pesos sinápticos com erros de previsão de vazão baixo. É esperado que os erros da saída de erro sejam baixos também. Isto acontece porque os padrões de treinamento para a saída de erro tornam-se mais estável ao longo das interações. Caso contrário, o processo de treinamento divergiria.

Durante o processo de treinamento o erro de porcentagem absoluto da previsão de vazão é utilizado como padrão de treinamento para o neurônio do erro de saída. Este erro é usado no lugar do erro relativo devido ser mais fácil de ser aprendido. Depois do treinamento da rede neural, o erro de saída, é somado e subtraído da previsão da vazão, para gerar um intervalo de confiança simétrico.

Nesta técnica, o grau de confiança do intervalo de confiança não é pré-definido. Deve ser calculado verificando o sucesso da estimação do intervalo de confiança para o conjunto de teste.

3.2 Re-amostragem dos erros

A técnica de re-amostragem dos erros dos erros de previsão para cada previsão um passo a frente pode ser feita do modo descrito por Alves da Silva & Moulin [2]. O conjunto utilizado para re-amostragem dos erros deve ser representativo das vazões futuras. Considera-se ainda, que erros das amostras são independentes um dos

outros, embora a distribuição de probabilidade seja desconhecida.

Figura 8 representa o conjunto de dados de teste disponíveis. O processo recursivo de previsão considerando dois lags de entrada para a previsão três passos a frente, é considerado. O valor da vazão conhecida para os tempos 1 e 2 são utilizados para previsão da vazão para o tempo 3. Como o valor da vazão verdadeiro para o tempo 3 é conhecido, o erro de previsão para este um passo a frente pode ser calculado. Em seguida, usando o valor conhecido para 2, e a previsão prévia para o tempo 3, o valor dois passos à frente é encontrado, permitindo o cálculo do erro de previsão correspondente. Os valores previstos para o tempo 3 e 4 são usados para encontrar a previsão da vazão para o tempo 5, e assim por diante. Medidas de previsão do erro para cada tempo foram obtidas, uma vez que a distância de previsão máxima desejada, instante 5, é encontrada.

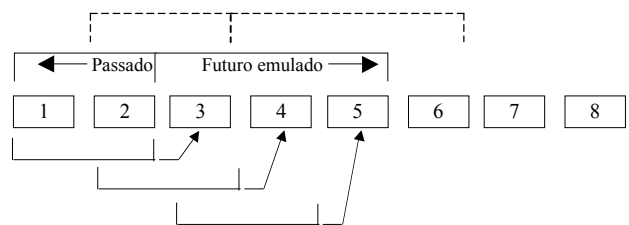


Figura 8. Exemplo da técnica de re-amostragem dos erros (Alves da Silva & Moulin [2]).

O procedimento descrito anteriormente é repetido para colecionar mais amostras para cada tempo, com os valores de vazão conhecidos dos tempos 2 a 6 (linha pontilhada superior). Este processo é repetido até, para uma certa janela, a distância máxima desejada de predição alcance o fim da série conhecida.

Em seguida, os n erros são organizados em ordem crescente (considerando os sinais) e são representados por $z_{(1)}, z_{(2)}, \dots, z_{(n)}$, a função de distribuição cumulativa dos erros de previsão das amostras pode ser calculada como mostrado a seguir:

$$S_n(z) = \begin{cases} 0 & , \quad z < z_{(1)} \\ \frac{r}{n} & , \quad z_{(r)} \leq z < z_{(r+1)} \\ 1 & , \quad z_{(n)} \leq z \end{cases} \quad (9)$$

$S_n(z)$ é a fração do conjunto de erros menor ou igual a z . Quando n é grande o suficiente, $S_n(z)$ é uma boa aproximação da distribuição de probabilidade cumulativa $F(z)$. Então, o intervalo de confiança pode ser calculado mantendo os $z(r)$'s valores intermediários e eliminando as extremidades, a quantidade de valores das extremidades descartados depende do grau de confiança desejado. O intervalo de confiança é calculado para ser simétrico em probabilidade (geralmente não simétrico em z). O número de casos para eliminar em cada extremidade da distribuição de erro da previsão é np onde p é a probabilidade em cada extremidade. Considerando que np geralmente é um

número fracionário, é truncado de modo conservador, e $(np-1)$ é levado como o número de casos para eliminar em cada extremidade. Z_p denotando tal que $F(Z_p)$ é igual a p , isto é, há uma probabilidade p que um erro seja menor ou igual a Z_p , indica que Z_p é o intervalo de confiança inferior para os erros de previsão futuras. Então, Z_{1-p} é o limite superior do intervalo de confiança e há um $(1-2p)$ intervalo de confiança para erros futuro.

4 Resultados para a barragem de Funil

As series temporais consideradas referem-se às leituras de vazão dos drenos, do nível do reservatório de água e das temperaturas. As medidas de instrumentação foram obtidas de setembro de 1985 a fevereiro de 2002, com frequência semanal, totalizando 736 amostras. Dois conjuntos de dados foram considerados: o primeiro composto de 588 amostras, usadas durante a fase de treinamento e as 148 amostras restantes formando o conjunto utilizado no processo de teste das previsões da vazão pela fundação da barragem.

Algumas amostras não apresentaram intervalo de tempo constante (semanal), e por esta razão a data de amostragem também foi considerada como dado de entrada. O aspecto temporal das séries foi incorporado através de uma função senoidal para representação das datas, conforme indicado pela equação (10).

$$data = \text{sen}\left(\frac{2\pi(data - data_origem)}{365}\right) \quad (10)$$

Seis diferentes tipos de configuração de redes neurais temporais, com diferentes números de parâmetros de entrada e neurônios na camada escondida, foram treinados e testados. A Tabela 1 mostra estas configurações, onde o símbolo T representa a data, NA o nível de água do reservatório, TMP a temperatura e VZ a vazão de água medida.

Na rede FIR, os valores de entrada foram considerados associados à ordem M dos filtros. Por exemplo, a rede III possui como parâmetros de entrada data e vazão, com filtros de memória de ordem 1 ($M=1$), enquanto a rede IV apresenta filtros com memória 2 ($M=2$) e as mesmas entradas da rede III.

| Rede | Entradas | Saída |
|------|---|--------|
| I | $T_{t-1}, NA_{t-1}, TMP_{t-1}, VZ_{t-1}, T_t$ | VZ_t |
| II | $T_{t-2}, NA_{t-2}, TMP_{t-2}, VZ_{t-2}, T_{t-1}, NA_{t-1}, TMP_{t-1}, VZ_{t-1}, T_t$ | VZ_t |
| III | T_{t-1}, VZ_{t-1}, T_t | VZ_t |
| IV | $T_{t-2}, VZ_{t-2}, T_{t-1}, VZ_{t-1}, T_t$ | VZ_t |
| V | $T_{t-1}, NA_{t-1}, VZ_{t-1}, T_t$ | VZ_t |
| VI | $T_{t-2}, NA_{t-2}, VZ_{t-2}, T_{t-1}, NA_{t-1}, VZ_{t-1}, T_t$ | VZ_t |

Tabela 1. Entradas e saídas adotadas

Para cada rede neural, os erros de treinamento e de teste foram listados nas Tabelas 2 a 4, sendo o melhor desempenho ressaltado em negrito.

A tabela 5 apresenta a porcentagem de acerto dos intervalos de confiança.

As Figuras 9, 10 e 11 comparam os valores observados com os previstos durante as fases de treinamento e teste. Como no conjunto de dados de teste há várias amostras com uma frequência de tempo diferente do intervalo de tempo regular (uma semana), o teste de generalização apresenta erros relativamente maiores que da fase de treinamento.

| Rede | Treinamento | | | Teste | | |
|-----------|-------------|-------------|---------------|-------------|-------------|---------------|
| | MAPE | RMSE | UTHEIL | MAPE | RMSE | UTHEIL |
| I | 3.88 | 0.39 | 0.9666 | 5.58 | 0.56 | 0.9742 |
| II | 3.67 | 0.37 | 0.9136 | 5.18 | 0.49 | 0.8543 |
| III | 3.89 | 0.39 | 0.9770 | 5.57 | 0.57 | 0.9844 |
| IV | 3.64 | 0.37 | 0.9160 | 5.25 | 0.51 | 0.8879 |
| V | 3.87 | 0.39 | 0.9672 | 5.55 | 0.56 | 0.9717 |
| VI | 3.65 | 0.37 | 0.9127 | 5.16 | 0.50 | 0.8637 |

Tabela 2. Desempenho da rede com “janelamento”

| Rede | Treinamento | | | Teste | | |
|------------|-------------|-------------|---------------|-------------|-------------|---------------|
| | MAPE | RMSE | UTHEIL | MAPE | RMSE | UTHEIL |
| I | 4.47 | 0.49 | 1.2160 | 5.98 | 0.56 | 0.9670 |
| II | 4.94 | 0.56 | 1.4080 | 6.71 | 0.63 | 1.1009 |
| III | 4.05 | 0.46 | 1.1426 | 5.84 | 0.55 | 0.9596 |
| IV | 4.59 | 0.54 | 1.3413 | 7.30 | 0.66 | 1.1402 |
| V | 4.38 | 0.48 | 1.2003 | 6.18 | 0.57 | 0.9886 |
| VI | 4.79 | 0.55 | 1.3794 | 6.99 | 0.65 | 1.1351 |

Tabela 3. Desempenho da rede FIR

| Rede | Treinamento | | | Teste | | |
|-----------|-------------|-------------|---------------|-------------|-------------|---------------|
| | MAPE | RMSE | UTHEIL | MAPE | RMSE | UTHEIL |
| I | 3.91 | 0.39 | 0.9680 | 5.67 | 0.56 | 0.9684 |
| II | 3.71 | 0.37 | 0.9235 | 5.22 | 0.49 | 0.8432 |
| III | 3.90 | 0.39 | 0.9760 | 5.69 | 0.57 | 0.9827 |
| IV | 3.68 | 0.37 | 0.9224 | 5.33 | 0.50 | 0.8740 |
| V | 3.91 | 0.39 | 0.9693 | 5.64 | 0.56 | 0.9659 |
| VI | 3.70 | 0.37 | 0.9230 | 5.21 | 0.49 | 0.8501 |

Tabela 4. Desempenho da rede Elman

| Técnica | Treinamento | Teste |
|-------------------------|-------------|-------|
| Saída de erro | 100% | 99% |
| Re-amostragem dos erros | 93% | 93% |

Tabela 5. Desempenho dos intervalos de confiança

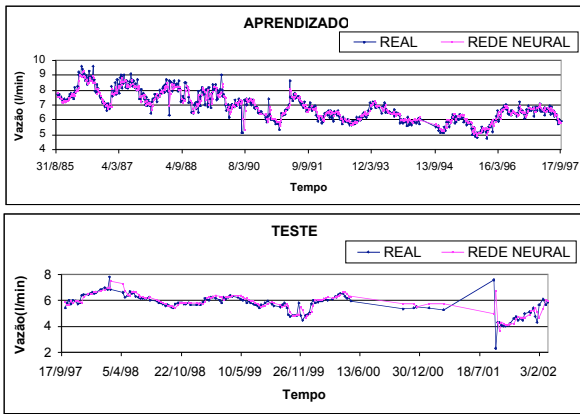


Figura 9. Comparação entre os valores observados e previstos na rede IV com método de “janelamento”.

Na Figura 12 é apresentado o intervalo de confiança gerado pela técnica de saída de erro. Enquanto que, na figura 13 o intervalo de confiança gerado pela técnica de re-amostragem dos erros. A técnica de re-amostragem dos erros foi utilizada para determinação de intervalo que compreenda 95% de probabilidade.

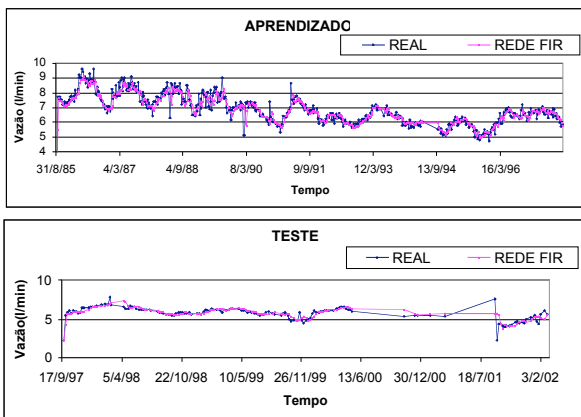


Figura 10. Comparação entre os valores observados e previstos na rede III do tipo FIR.

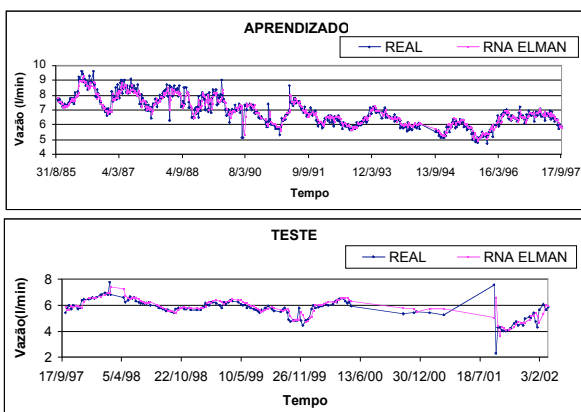


Figura 11. Comparação entre os valores observados e previstos na rede IV do tipo Elman elementos.

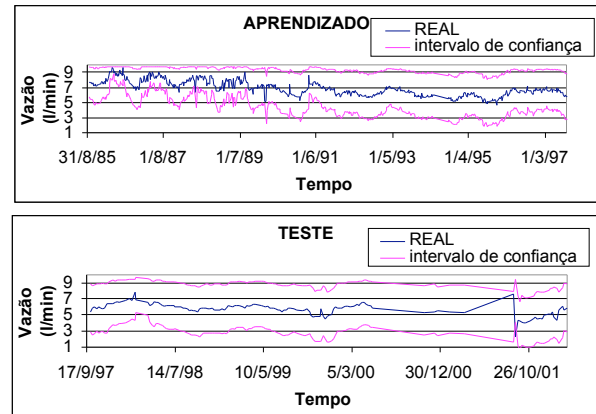


Figura 12. Intervalo de Confiança gerado pela técnica de saída do erro com a previsão obtida pela RNT com método de “janelamento”.

5 Discussão e Conclusões

Neste trabalho os resultados de instrumentação da barragem Funil, de Furnas Centrais Elétricas, em relação aos valores de vazão na fundação foram analisados com o auxílio de redes neurais temporais. As redes neurais temporais empregadas foram: RNT com arquitetura feedforward associada a técnica de janelamento, RNT recorrente Elman e RNT FIR.

Nesta pesquisa estuda-se também a geração de intervalos de confiança para RNT. As previsões obtidas pela RNT com arquitetura feedforward associada à técnica de janelamento são utilizadas para a geração de intervalos de confiança pelos métodos de saída do erro e re-amostragem dos erros.

Neste estudo analisa-se o problema considerando o intervalo de tempo variável. Então para os modelos conseguirem captar a distância entre as amostras, a data da amostra foi considerada como variável explicativa. Esta data foi utilizada como entrada para uma função seno, com o intuito de considerar o caráter cíclico do fenômeno estudado, onde a saída foi utilizada como entrada das RNT. Nesta análise consideram-se ainda, outras variáveis explicativas o nível do reservatório à montante e a temperatura. Observando as métricas obtidas constatou-se que as previsões foram satisfatoriamente obtidas. Nesta análise, os desempenhos das redes neurais com “janelamento” e do tipo Elman mostraram-se semelhantes e melhores que o apresentado pela rede temporal FIR.

Quanto ao estudo dos intervalos de confiança foram necessários utilizar técnicas que permitissem a geração de intervalos de confiança para estimativas com distribuição do erro não gaussiana. As técnicas de saída do erro, re-amostragem dos erros apresentam esta característica, sendo utilizadas nas previsões geradas pelos modelos das RNT. A aplicação de metodologias para determinação de intervalos de confiança é de grande importância, já que o processo de previsão sempre estará associado a erros de estimativa. O intervalo de confiança permite que ao invés de fornecer um valor único para a estimativa, seja fornecido um intervalo e a probabilidade que o valor desconhecido

esteja dentro deste intervalo. Neste estudo, os obtidos pela técnica de saída do erro apresentam grau de confiança elevado, mas os intervalos são muito “largos”. Enquanto que os obtidos pela técnica de re-amostragem dos erros não atenderam ao grau de confiança esperado.

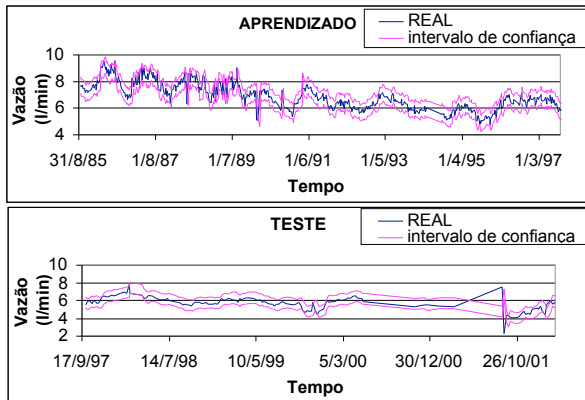


Figura 13. Intervalo de Confiança gerado pela técnica de re-amostragem dos erros com a previsão obtida pela RNT com método de “janelamento”.

Agradecimentos

Os autores expressam gratidão a Furnas Centrais Elétricas, proprietária da barragem de Funil, por permitir acesso aos dados e relatórios técnicos.

Referências

- [1] ALVES, A. P. & MOULIN, L. S. , “Neural Network based short-term electric load forecasting with confidence intervals” Proceedings of the IV Brazilian Conference on Neural Networks. ITA, São José dos Campos. pp. 7-12, 1999.
- [2] ALVES, A. P. & MOULIN, L. S., “Confidence intervals for neural network based short-term load forecasting” IEEE Trans. Power Syst., Vol. 15, pp. 1191-1196, 2000.
- [3] CHRYSOLOURIS, M. L. & RAMSEY, A., “Confidence Interval prediction for neural network models” IEEE Trans. Neural Networks, Vol. 7, pp. 229-232, 1996.
- [4] HAYKIN, S., *Neural Networks: A Comprehensive Foundation*, MacMillan Publishing Co, 1994.
- [5] PAPAPOPOULOS, G. ; EDWARDS, P. J. & MURRAY, A. F., “Confidence estimation methods for neural networks: A practical comparison” IEEE Trans. Neural Networks, Vol. 12, pp. 1278-1287, 2001.
- [6] RIVALS, I. and PERSONNAZ, L., “Construction of confidence intervals for neural networks based on least squares estimation” Neural Networks, Vol. 13, pp. 463-484, 2000.
- [7] SOTO, C. P., “Redes neurais temporais para o tratamento de sistemas variantes no tempo”, Dissertação de Mestrado, Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro, 1999.
- [8] THEIL, H., *Applied Economic Forecasting*, North-Holland, Amsterdam, 1966.
- [9] TOWNSEND, N. W. & TARASSENKO, L., “Estimations of error bounds for neural-network function approximators” IEEE Trans. Neural Networks, Vol. 10, pp. 217-230, 1999.
- [10] WAN, E. A., "Time Series Prediction by Using a Connectionist Network with Internal Delay Lines". Time Series Prediction. Forecasting the Future and Understanding the Past.. SFI Studies in the Sciences of Complexity, Proc. Vol XVII, Addison Wesley, 1994.
- [11] ZHANG, L. & LUH, P. B., “Confidence regions for cascaded neural network prediction in power markets” IEEE Power Eng. Soc. Winter Meeting, Columbus, OH, pp. 533-538, 2001.
- [12] ZHANG, L.; LUH, P. B. & KASIVISWANATHAN, K. , “Energy clearing price prediction and confidence interval estimation with cascaded neural networks” IEEE Trans. on Power Systems, Vol. 18, n. 1, pp. 99-105, 2003.

