

## ESTIMATIVA DAS INCERTEZAS DE MODELO NEURAL UTILIZADO NA PROJEÇÃO DE SÉRIES TEMPORAIS DE VAZÕES MENSAS

LEVI LOPES TEIXEIRA - prof.levilopes@gmail.com  
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ - UTFPR

PAULO HENRIQUE SIQUEIRA - paulohs@ufpr.br  
UNIVERSIDADE FEDERAL DO PARANÁ - UFPR

LUIZ ALBINO TEIXEIRA JÚNIOR - luiz.junior@unila.edu.br  
UNIVERSIDADE FEDERAL DA INTEGRAÇÃO LATINO AMERICANA

**Resumo:** DADA A RELEVÂNCIA DA PROJEÇÃO DE VAZÕES NO ÂMBITO DAS USINAS HIDRELÉTRICAS, ESTE ARTIGO PROPÕE UMA NOVA METODOLOGIA HÍBRIDA WAVELET-NEURAL INTEGRADA COM UM AMOSTRADOR BOOTSTRAP. BASICAMENTE, COMBINAM-SE OS MÉTODOS DE ENCOLHIMENTO E DE DECOMPOSIÇÃO WAVELET NO PRÉ-PROCESSAMENTO DOS DADOS E, EM SEGUIDA, UMA REDE NEURAL ARTIFICIAL (RNA) É USADA PARA PRODUZIR AS DE PREVISÕES PONTUAIS. A MEDIDA DE INCERTEZA DO MODELO RNA É OBTIDA COM A UTILIZAÇÃO DE UMA AMOSTRAGEM BOOTSTRAP DOS RESÍDUOS DO MODELO RNA AJUSTADO À SÉRIE TEMPORAL SUBJACENTE. A FIM DE SE OBTER O INTERVALO DE CONFIANÇA, CALCULOU-SE A MÉDIA PONDERADA DAS PREVISÕES DE B SÉRIES TEMPORAIS ORIUNDAS DO PROCESSO BOOTSTRAP, SENDO OS PESOS DETERMINADOS VIA OTIMIZAÇÃO DE UM PROBLEMA DE PROGRAMAÇÃO NÃO LINEAR CUJA FUNÇÃO OBJETIVO É A MINIMIZAÇÃO DA RAIZ QUADRADA DO ERRO QUADRÁTICO MÉDIO ENTRE A COMBINAÇÃO LINEAR DAS B PREVISÕES E A SÉRIE TEMPORAL SUBJACENTE. EM PARTICULAR, AS PREVISÕES PONTUAIS E INTERVALO DE CONFIANÇA FORAM ESTIMADOS PARA UMA SÉRIE TEMPORAL DE VAZÃO MÉDIA MENSAL AFLUENTE DA HIDRELÉTRICA DE ITAIPU, EM FOZ DO IGUAÇU, BRASIL. AS PREVISÕES PONTUAIS OBTIDAS COM O MÉTODO PROPOSTO APRESENTARAM BAIXO ERRO PREDITIVO E AS ESTIMATIVAS DAS INCERTEZAS DO MODELO REFLETIRAM A BAIXA VARIABILIDADE DO PROCESSO DE MODELAGEM

**Palavras-chaves:** PREVISÃO DE VAZÕES; WAVELET; REDES NEURAS ARTIFICIAIS;  
BOOTSTRAP; INTERVALO DE CONFIANÇA

**Área:** 6 - PESQUISA OPERACIONAL  
**Sub-Área:** 6.3 - PROCESSOS ESTOCÁSTICOS

# ESTIMATED UNCERTAINTIES OF NEURAL MODEL USED IN PROJECTION TIME SERIES OF MONTHLY FLOWS

**Abstract:** *GIVEN THE IMPORTANCE OF PROJECTION OF FLOWS WITHIN THE HYDROELECTRIC PLANTS, THIS ARTICLE PROPOSES A NEW HYBRID METHODOLOGY INTEGRATED WAVELET-NEURAL WITH A BOOTSTRAP SAMPLER. BASICALLY, COMBINE THE SHRINKAGE AND DECOMPOSITION WAVELET MMETHODS PROCESSING OF DATA AND THEN AN ARTIFICIAL NEURAL NETWORK (ANN) IS USED TO PRODUCE THE POINT PREDICTIONS. THE ANN UNCERTAINTY MEASUREMENT OF THE TEMPLATE IS ACHIEVED WITH THE USE OF A BOOTSTRAP SAMPLE OF ANN TEMPLATE TO THE UNDERLYING WASTE ADJUSTED TO THE TIME SERIES. IN ORDER TO OBTAIN THE CONFIDENCE INTERVAL, IT WAS CALCULATED THE WEIGHTED AVERAGE OF B TIME SERIES FORECASTS DERIVED FROM THE BOOTSTRAP PROCESS, THE WEIGHTS WERE DETERMINED THROUGH OPTIMIZATION OF A NONLINEAR PROGRAMMING PROBLEM WHOSE OBJECTIVE FUNCTION IS THE MINIMIZATION OF THE SQUARE ROOT OF MEAN SQUARE ERROR BETWEEN THE LINEAR COMBINATION OF B FORECASTS AND THE UNDERLYING TIME SERIES. IN PARTICULAR, THE POINT ESTIMATES AND CONFIDENCE INTERVALS WERE ESTIMATED FOR A TIME SERIES OF MONTHLY AVERAGE TRIBUTARY FLOW OF ITAIPU DAM, IN FOZ DO IGUAÇU, BRAZIL. THE POINT ESTIMATES OBTAINED WITH THE PROPOSED METHOD HAD LOW PREDICTIVE ERROR AND ESTIMATES OF MODEL UNCERTAINTIES REFLECTED THE LOW VARIABILITY OF THE MODELING PROCESS*

**Keyword:** *STREAMFLOW FORECASTING; WAVELET; ARTIFICIAL NEURAL NETWORKS; BOOTSTRAP; CONFIDENCE INTERVAL*

## 1 Introdução

No Brasil, o Operador Nacional do Sistema Elétrico (ONS) utiliza as informações sobre as vazões na otimização da geração de energia e elaboração de estratégias que minimizem efeitos negativos de eventuais enchentes ou secas. A produção de energia a partir das usinas hidrelétricas tem ligação direta com a quantidade de água (vazão afluyente) disponível para a movimentação das turbinas geradoras. De forma que longos períodos de seca podem levar a uma diminuição da capacidade geradora de energia da usina. A relação entre vazões afluentes e geração de energia implica em um cuidado especial com os dados que cercam o problema.

Segundo WANG *et al.*(2006a), os processos de geração de vazões são sazonais e não lineares. Os mecanismos não lineares são mais fortes em escalas menores de tempo, sendo as vazões diárias e mensais não lineares (Wang, *et al.*, 2006b). São vários os trabalhos existentes na literatura que discutem a previsão de séries oriundas de processos não lineares, sendo predominante para este caso o uso de Redes Neurais Artificiais (RNA). Ballini (2000); Wang *et al.*,(2006a); Valenzuela *et al.* (2008); Teixeira, Jr *et al.* (2012), entre outros, utilizaram Redes Neurais Artificiais (RNA) na previsão de séries temporais.

No caso da série temporal ser ruidosa, há duas abordagens básicas. Na primeira, ignoram-se os ruídos e modela-se a série na expectativa de se capturar características determinísticas da série temporal. Na segunda, faz-se a filtragem da série procurando identificar as componentes de ruídos e determinística. De maneira geral, a segunda abordagem é mais eficiente que a primeira (HASSANI; ZHIGLJAVSKY, 2009).

Da Análise *Wavelet* advém importantes métodos auxiliares de pré-processamento (decomposição, filtragem ou alisamento) de uma série de tempo, antes de sua efetiva modelagem com um método preditivo base plausível, de forma a gerar um novo método para se prever séries temporais (AQUINO *et al.* 2009). Desde que sejam utilizados de forma adequada com os métodos preditivos, é possível se obter ganhos preditivos relevantes. Pode-se verificar este fato, dentre muitas outras referências, em: Donoho e Jonhstone (1994); Donoho *et al.* (1995).

A aplicação de modelos de regressão implica na avaliação das incertezas que cercam estimadores de parâmetros. O intervalo de confiança (IC) é uma medida estatística valorosa

que desempenha esta importante função. O IC mede as variações concernentes aos parâmetros do modelo de regressão. Tibshirani (1996) utilizou *bootstrap pairs* e *residual* na avaliação das incertezas decorrentes de previsões obtidas por meio de redes *multi-layer perceptron*. Concluiu que o *bootstrap* apresentou desempenho satisfatório, capturando de forma eficiente a variabilidade do processo em consequência da aleatoriedade dos pesos iniciais das redes *bootstrap*.

Diante do exposto, propõe-se, neste artigo, um método para previsões pontuais de séries temporais e estimativa das incertezas do modelo regressor. O método proposto possui as seguintes etapas: (i) alisamento da série temporal por encolhimento *wavelet* (EW); (ii) decomposição ortogonal *wavelet*; (iii) modelagem das componentes *wavelet* por RNAs; e (iv) estimativa do intervalo de confiança com *bootstrap residual*.

## 2 Análise Wavelet

De acordo com Kubrusly e Levan (2006) e Mallat (2009), tem-se que a decomposição *wavelet* de uma série temporal  $f(\cdot)$  é dada, genericamente, em (1).

$$f(\cdot) = f_{A_{m_0}}(\cdot) + \sum_{m=m_0}^{\infty} f_{D_m}(\cdot), \quad (1)$$

onde:  $f_{A_{m_0}}(\cdot) = \sum_{n \in \mathbb{Z}} a_{m_0,n} \phi_{m_0,n}(\cdot)$  é a componente de aproximação de nível  $m_0$  de  $f(\cdot)$ ;

$f_{D_m}(\cdot) = \sum_{n \in \mathbb{Z}} d_{m,n} \omega_{m,n}(\cdot)$  é a componente de detalhe de nível  $m$  de  $f(\cdot)$ ;  $d_{m,n}$  e  $a_{m_0,n}$  são,

respectivamente, os coeficientes de detalhe e de aproximação e são calculados, ainda nesta ordem, a partir do produto interno usual entre  $f(\cdot)$  e a função *wavelet*  $\omega_{m,n}(\cdot)$  (isto é,

$$d_{m,n} = \sum_{t=1}^T f(t) \omega_{m,n}(t) \text{ e entre } f(\cdot) \text{ e a função escala } \phi_{m_0,n}(\cdot) \text{ (isto é, } a_{m_0,n} = \sum_{t=1}^T f(t) \phi_{m_0,n}(t)),$$

para todo  $m_0, m, n \in \mathbb{Z}$ , onde  $\mathbb{Z}$  é o conjunto dos números inteiros. São várias as famílias de funções *wavelet* existentes na literatura, tais como: as famílias *daubechies* (db), *coiflets* (coif) e *symlets* (sym), desenvolvidas por Ingrid Daubechies (DAUBECHIES, 1992).

Se a série temporal  $f(\cdot) = (f(1), f(2), \dots, f(T))$  for fortemente ruidosa, é recomendada a sua filtragem para então modelá-la. Note que  $f(\cdot)$  é tal que, para cada instante  $t$ , pode ser teoricamente decomposta como:  $f(\cdot) = \tilde{f}(t) + e_t$ , onde: (i)  $e_t$  é um estado (ou realização) de uma variável aleatória independente e identicamente distribuída (iid), com

média zero e variância constante; e (ii)  $\tilde{f}(t)$  é a componente determinística (que, de fato, é passível de ser prevista pontualmente ao contrário de  $e_t$ ) de  $f(t)$ .

Uma das técnicas destinadas à filtragem de séries temporais é o encolhimento (*shrinkage*) *wavelet*, que tem como objetivo a redução (ou mesmo remoção) do ruído presente num sinal, diminuindo (ou zerando) a magnitude dos coeficientes *wavelet* (MORETTIN, 1999). Outras referências nesta área são os trabalhos de Donoho e Johnstone (1994) e Donoho *et al.*, (1995).

As três etapas do encolhimento *wavelet* são: 1) decomposição *wavelet* do sinal ruidoso; 2) selecionar limiares (*thresholds*) que eliminam (ou reduzem) os ruídos; 3) aplicar a transformada *wavelet* inversa a fim de se obter o sinal menos ruidoso. Das várias regras de limiarização encontradas na literatura, talvez as mais tradicionais sejam: o limiar duro (*hard threshold*) e limiar suave (*soft threshold*). Em ambas as regras, eliminam-se os coeficientes de detalhes ( $d_{m,n}$ ) que em módulo são menores ou iguais a um limiar  $\lambda$ . Na regra limiar duro  $d_{m,n}$  é mantido sempre que  $|d_{m,n}| > \lambda$  e atenuado por  $|d_{m,n}| - \lambda$  na limiar suave. Neste trabalho o cálculo do limiar  $\lambda$  foi realizado a partir de duas propostas: universal e SURE (*Estimation Stein's Unbiased Risk Estimate*). Maiores detalhes dessas técnicas podem ser encontradas em Donoho e Johnstone (1993), Johnstone e Silverman (1996) e Morettin (1999).

### 3 Redes Neurais Artificiais

De acordo com Haykin (2001), uma RNA é um sistema de processamento paralelo formado por unidades capazes de armazenar conhecimento e disponibilizá-lo para o futuro, sendo formada basicamente por neurônios artificiais distribuídos em camadas. Neste artigo foi utilizada RNA do tipo *multilayer feed-forward*, que são formadas por várias camadas e os sinais são propagados apenas da entrada para a saída, nunca ao contrário. A ligação entre os neurônios de uma camada e neurônios de camadas adjacentes ocorre por meio de pesos, que são parâmetros ajustáveis numericamente e que representam a “força” das conexões. O somatório das entradas linearmente ponderadas para um neurônio é transformado por uma função de ativação, que pode ser linear ou não linear (por exemplo, a tangente hiperbólica, a sigmoide). Neste tipo de RNA, não há conexões entre neurônios da mesma camada. As informações de dados são transmitidas de forma que as saídas dos neurônios de uma camada correspondam às entradas da camada seguinte. Para o treinamento da RNA os algoritmos

*backpropagation* e *Levenberg-Marquardt* são amplamente utilizados e recomendados (HAYKIN, 2001).

O problema de previsão de séries temporais, com  $T$  observações  $(y_t)_{t=1}^T$ , através de RNA consiste na utilização de dados observados para se prever  $y(t+l)$ , onde  $l$  é o passo de tempo de previsão para o futuro. Uma vez definido o tamanho da janela  $n$  e o valor de  $l$ , montam-se a sequência  $(y(t), y(t-1), \dots, (y(t-n+1), (y(t+1), \dots, (y(t+l))$  de dados, que forma um conjunto de padrões de treinamento, onde:  $(y(t), y(t-1), \dots, y(t-n+1))$  e  $(y(t+1), \dots, y(t+l))$  representam as entradas e saídas desejadas, respectivamente.

#### 4 Bootstrap

O *bootstrap* é um procedimento estatístico não paramétrico introduzido por Efron (1979). A técnica constitui na reamostragem de mesmo tamanho e com reposição dos dados da amostra original, com o objetivo de se estimar uma medida de interesse. O princípio básico do método *bootstrap* estabelece que na ausência de informações sobre a verdadeira distribuição de uma variável aleatória, pode-se obter uma distribuição empírica a partir de reamostragens de uma amostra observada. Considerando-se que esta amostra possui informações da distribuição base,  $B$  amostras daquela permitirão estimar parâmetros da população da qual foi retirada.

Considere  $X = (X_1, X_2, \dots, X_n)$  uma amostra aleatória de uma distribuição  $F$  e  $\hat{\theta} = s(X)$  o estimador de um parâmetro  $\theta$  desta distribuição. Uma amostra *bootstrap*  $X^* = (X_1^*, X_2^*, \dots, X_n^*)$  é obtida da amostra original  $X$  rearranjando os seus  $n$  elementos. Assim, uma das possibilidades para  $X^*$  é  $X^* = (X_1^* = X_2, X_2^* = X_n, \dots, X_n^* = X_1)$ , fixando a probabilidade de escolha em  $X$  de  $1/n$  para cada uma das suas observações. A partir deste procedimento determinam-se  $B$  amostras *bootstrap* independentes da distribuição empírica  $\hat{F}$ , obtendo-se então o estimador *bootstrap*  $\hat{\theta}^{*(b)} = s(X^{*(b)})$ , com  $b = 1, 2, \dots, B$ , do parâmetro  $\theta$ .

De acordo com Montgomery e Runger (2003), a definição do número  $B$  de amostras *bootstrap* pode ser realizada com a observação da variação do desvio padrão do estimador *bootstrap*. A baixa variabilidade ou estabilidade desse valor indica o valor de  $B$  mais adequado.

## 5 Materiais e Métodos

Neste artigo foi utilizada uma série temporal de vazão de afluentes mensal média da usina hidrelétrica de Itaipu, cujo período é de janeiro de 1970 até dezembro de 2010 (492 observações). A série temporal utilizada pode ser visualizada em [http://www.ons.org.br/operacao/vazoes\\_naturais.aspx](http://www.ons.org.br/operacao/vazoes_naturais.aspx).

Inicialmente, submete-se a série  $(y_t)_{t=1}^T$  ao encolhimento wavelet, obtendo-se a série de tendência  $(y_{t,EW})_{t=1}^T$  e a componente total de resíduo  $(y_{t,R})_{t=1}^T = (y_t)_{t=1}^T - (y_{t,EW})_{t=1}^T$ . Os testes de autocorrelação residual, independência estatística BDS e de estacionariedade *Dickey-Fuller* são aplicados à componente  $y_{t,R}$  a fim de se obter indicação que a mesma não pode ser classificada estatisticamente como ruído branco, de forma que os resíduos devem possuir estruturas de autodependência linear e/ou não linear que podem ser modeladas. Neste caso, a previsão da série observada resultará da soma das previsões dos resíduos e tendência ( $\hat{y}_t = \hat{y}_{t,EW} + \hat{y}_{t,R}$ ).

Considere a sequência infinita  $(\tilde{y}_{t,S})_{t \in Z}$ , onde  $S = EW$  ou  $R$ , obtida da série temporal  $(y_{t,S})_{t=1}^T$  a partir da igualdade  $\tilde{y}_{t,S} = \begin{cases} y_{t,S}, t \in \{1, \dots, T\} \\ 0, t \in Z \setminus \{1, \dots, T\} \end{cases}$ . Segundo Teixeira Jr *et al.*, (2012) a decomposição wavelet de nível  $p$  de uma série temporal  $\tilde{y}_{t,S}$  em relação a uma determinada base ortonormal wavelet é representada por (2).

$$\tilde{y}_{t,S} \cong \tilde{\tilde{y}}_{t,S} = \sum_{n \in \Gamma_{V_{m_0}}} a_{m_0,n} \phi_{m_0,n} + \sum_{m=m_0}^{m_0+(p-1)} \left( \sum_{n \in \Gamma_{W_m}} d_{m,n} \omega_{m,n} \right), \quad (2)$$

onde  $\Gamma_{V_{m_0}}$  e  $\Gamma_{W_m}$  são subconjuntos de índices em  $Z$ . A determinação de tais conjuntos ocorre através da minimização do erro quadrático médio (MSE) entre  $\tilde{y}_{t,S}$  e  $\tilde{\tilde{y}}_{t,S}$ .

As componentes de aproximação  $\sum_{n \in \Gamma_{V_{m_0}}} a_{m_0,n} \phi_{m_0,n}$  e detalhe  $\sum_{n \in \Gamma_{W_m}} d_{m,n} \omega_{m,n}$ , provenientes da decomposição das séries de tendência e resíduo, e o encolhimento *wavelet* de  $y_t$  foram obtidos a partir do aplicativo computacional Matlab 8.0.

Neste trabalho, as séries de tendência e resíduos foram decompostas ortogonalmente via *wavelet* de nível  $p = 2$ , que resulta em duas componentes de detalhe ( $D_{1,S}$  e  $D_{2,S}$ ) e uma de aproximação  $A_{2,S}$ .

Uma vez obtidas as componentes *wavelet* de aproximação e detalhe para as séries de tendência e resíduos, estas foram modeladas via redes neurais artificiais *feedforward multilayer perceptron* (RNA-MLP), resultando as previsões das componentes *wavelet* de detalhe ( $\hat{D}_{1,S}$  e  $\hat{D}_{2,S}$ ) e aproximação ( $\hat{A}_{2,S}$ ). A previsão da série observada é dada pela soma das previsões das componentes *wavelet*, representado em (3).

$$\hat{y}_t = \hat{A}_{2,EW}(t) + \hat{D}_{1,EW}(t) + \hat{D}_{2,R}(t) + \hat{A}_{2,R} + \hat{D}_{1,R}(t) + \hat{D}_{2,R}, t = 1, \dots, T. \quad (3)$$

Tendo previsões pontuais  $\hat{y}_t$  da série temporal observada  $y_t$ , geram-se  $B$  amostras *bootstrap* a partir da amostragem residual sugerida por Tibshirani (1996). Uma vez obtidos os resíduos de treinamento  $r_t = y_t - \hat{y}_t$ , estes são reamostrados  $B$  vezes segundo o procedimento *bootstrap*, sendo que o resultado para a  $b$ -ésima amostra representado por  $r_1^{*b}, r_2^{*b}, \dots, r_t^{*b}$ , com  $b = 1, \dots, B$ . Na sequência, obtêm-se as séries temporais *bootstrapadas*  $y_t^{*b} = \hat{y}_t + r_t^{*b}$ , com  $t = 1, \dots, T$ . A fim de se determinar as previsões pontuais das séries temporais *bootstrapadas*  $y_t^{*b}$ , utiliza-se a equação (3).

Para cálculo do erro padrão do  $t$ -ésimo valor previsto, Tibshirani (1996) sugere o uso da equação (4).

$$\hat{s}_{boot}(\hat{y}_t) = \left[ \frac{1}{B-1} \sum_{b=1}^B (\hat{y}_t^{*b} - \bar{y}_t^*)^2 \right]^{1/2}, \quad (4)$$

onde  $\bar{y}_t^*$  denota uma média aritmética simples e é dada por  $\bar{y}_t^* = \frac{1}{B} \sum_{b=1}^B \hat{y}_t^{*b}$ , sendo  $\hat{y}_t^{*b}$  a previsão de  $y_t^{*b}$ . Neste artigo, propõe-se calcular  $\bar{y}_t^*$  a partir de uma média ponderada, conforme (5), cujos pesos  $\delta_b$  são obtidos através da otimização de um problema de programação não linear, descrito em (6), onde a função objetivo é a raiz do erro quadrático médio entre a combinação linear das previsões  $\hat{y}_t^{*b}$  e a série temporal observada  $\hat{y}_t$ , na amostra de treinamento.

$$\bar{y}_t^* = \sum_{b=1}^B \delta_b \hat{y}_t^{*b} / \sum_{b=1}^B \delta_b, \quad (5)$$

$$\text{Min} \sqrt{\frac{1}{T} \sum_{t=1}^T \left( \sum_{b=1}^B \delta_b \hat{y}_t^{*b} - y_t \right)^2} \quad \text{s.a.}, \delta_b \geq 0 \quad (6)$$

No modelo de regressão neural  $y(\bar{x}) = f(\bar{x}, \bar{w}) + e(\bar{x})$ ,  $f(\bar{x}, \bar{w})$  é a regressão verdadeira e  $\hat{y}(\bar{x}) = g(\bar{x}, \hat{w})$ , uma aproximação dela. Para uma RNA, a discussão do intervalo de confiança (IC) passa pela distribuição da quantidade  $f(\bar{x}, \bar{w}) - \hat{y}(\bar{x})$ . No caso do uso da técnica *bootstrap*, tem-se  $B$  RNAs e passa a ser de interesse a distribuição da quantidade  $f(\bar{x}, \bar{w}) - \bar{y}_t^*$ , onde  $\bar{y}_t^*$  é determinado através de (5). Assumindo que  $P(f(\bar{x}, \bar{w}) | \bar{y}_t^*(\bar{x}))$  e sua inversa  $P(\bar{y}_t^*(\bar{x}) | f(\bar{x}, \bar{w}))$  possuem distribuição gaussiana, o procedimento *bootstrap* permite gerar uma distribuição empírica  $P(\hat{y}(\bar{x}) | \bar{y}_t^*(\bar{x}))$  que pode estimar  $P(\bar{y}_t^*(\bar{x}) | f(\bar{x}, \bar{w}))$  (HESKES, 1997). Com efeito, a estimativa da variância ( $\hat{\sigma}_g^2(\bar{x})$ ) da distribuição de  $P(f(\bar{x}, \bar{w}) | \bar{y}_t^*(\bar{x}))$  pode ser representada pelo quadrado do resultado obtido em (4).

Uma vez obtida a estimativa para variância da distribuição  $P(f(\bar{x}, \bar{w}) | \bar{y}_t^*(\bar{x}))$ , constrói-se o IC para  $f(\bar{x}, \bar{w})$  conforme (7).

$$\bar{y}_t^* - t_{1-\frac{\alpha}{2}, gl} \hat{\sigma}_g \leq f(\bar{x}, \bar{w}) \leq \bar{y}_t^* + t_{1-\frac{\alpha}{2}, gl} \hat{\sigma}_g, \quad (7)$$

onde  $t_{1-\frac{\alpha}{2}, gl}$  provêm da distribuição  $t$  de Student com nível de significância  $\alpha$ , graus de liberdade  $gl$  igual à diferença entre o número de amostras de treinamento e o número de parâmetros do modelo RNA (KHOSRAVI *et al.*, 2014).

Considerando o método proposto para previsões pontuais (Equação (3)),  $\hat{y}_t$  é obtida da modelagem neural das componentes *wavelet* de aproximação e detalhes provenientes das séries temporais de tendência  $y_{t,EW}$  e resíduo  $y_{t,R}$ . Então, o procedimento descrito para a estimativa do IC deve ser aplicado a cada uma das componentes *wavelet* e suas respectivas previsões, somando-se os intervalos de confiança dos regressores dessas componentes.

Seja o caso da série de tendência  $y_{t,EW}$ , que foi decomposta ortogonalmente nas componentes *wavelet*  $A_{2,EW}, D_{1,EW}$  e  $D_{2,EW}$ . De posse das suas previsões  $(\hat{A}_{2,EW}, \hat{D}_{1,EW}, \hat{D}_{2,EW})$ , obtém-se o conjunto das séries *bootstrapadas*  $A_{2,EW}^{*b}, D_{1,EW}^{*b}$  e  $D_{2,EW}^{*b}$  e suas previsões  $\hat{A}_{2,EW}^{*b}, \hat{D}_{1,EW}^{*b}$  e  $\hat{D}_{2,EW}^{*b}$ . No passo seguinte determinam-se as médias ponderadas  $\bar{A}_{2,EW}^*, \bar{D}_{1,EW}^*$  e  $\bar{D}_{2,EW}^*$  (Equações (5) e (6)) e com elas as estimativas dos erros padrões  $\hat{\sigma}_{A_2}, \hat{\sigma}_{D_1}$  e  $\hat{\sigma}_{D_2}$  (Equação (4)). Por fim, o intervalo de confiança oriundo da modelagem da

série de tendência ( $IC(y_{t,EW})$ ), que resulta da soma dos ICs, obtidos segundo (7), para as modelagens de  $A_{2,EW}$ ,  $D_{1,EW}$  e  $D_{2,EW}$ .

Para o cálculo do IC referente à modelagem da série de resíduos *wavelet* ( $y_{t,R}$ ) utiliza-se o mesmo procedimento descrito no último parágrafo e, a soma deste com o IC de  $y_{t,EW}$  fornece o IC da série observada, ou seja:  $IC(y_t) = IC(y_{t,EW}) + IC(y_{t,R})$ . Observando que a previsão de  $y_t$  resulta da soma das médias  $\bar{A}_{2,EW}^*$ ,  $\bar{D}_{1,EW}^*$ ,  $\bar{D}_{2,EW}^*$ ,  $\bar{A}_{2,R}^*$ ,  $\bar{D}_{1,R}^*$ , e  $\bar{D}_{2,R}^*$ .

## 6 Resultados Obtidos

Os primeiros 70% dos dados da série de vazões foram utilizados no treinamento, separando 20% para a validação e 10% dos remanescentes, na amostra de teste. Realizando-se previsões um passo à frente, em um horizonte de 49 meses. Para a avaliação do método foram utilizados o Erro Médio Percentual Absoluto e Erro Médio absoluto, representados por

$$MAPE = \frac{1}{T} \sum_{t=1}^T \left| \frac{y_t - \hat{y}_t}{y_t} \right| \text{ e } MAE = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t|, \text{ respectivamente.}$$

A fim de se escolher a configuração relativa ao encolhimento *wavelet* (EW), foram realizadas quinze modelagens distintas, utilizando-se mais de um tipo de base *wavelet* (daubechies, symlets e coiflets), níveis de decomposição, limiar (universal e SURE) e regra de limiarização (suave e duro). A melhor configuração, em termos de erro preditivo, era constituída por base *wavelet* de haar, nível de decomposição ortogonal wavelet  $p = 2$ , regra de limiarização suave e limiar universal.

Com o encolhimento *wavelet* da série  $y$  obteve-se a série  $y_{EW}$  (tendência) supostamente menos ruidosa que a série observada e o resíduo *wavelet*  $y_R = y - y_{EW}$ . Testes estatísticos, ao nível de 5% de significância, indicaram que o resíduo ( $y_R$ ) proveniente do EW da série  $y$  não pode ser classificado como ruído branco, de forma que é passível de modelagem.

Para a decomposição ortogonal *wavelet* das séries de tendência e resíduos EW escolheu-se as bases de Daubechies dbN em nível  $p = 2$ , com o valor de N definido a partir do erro quadrático médio entre as séries decompostas e as respectivas somas das três componentes *wavelet* (TEIXEIRA, 2013).

As modelagens das componentes *wavelet* (CW) foram realizadas por RNA *feedforward* através do *software* Matlab 8.0. As melhores configurações encontradas estão expostas no Quadro 1.

QUADRO 1 - MAPE e configurações das RNA usadas na modelagem das componentes *wavelet* provenientes das séries tendência e Resíduos EW.

Séries	Componentes <i>Wavelet</i>	Nº de neurônios na camada escondida	Janela de entrada de dados	MAPE: amostra de validação
Tendência	Aproximação A2	10	12	0,41
	Detalhes D1	14	12	28,4
	Detalhes D2	11	14	12,8
Resíduos EW	Aproximação A2	10	12	0,035
	Detalhes D1	14	14	35,2
	Detalhes D2	12	12	17,6

Fonte: o autor (2015).

As redes neurais do Quadro 1 eram constituídas por uma camada escondida com função de ativação tangente hiperbólica sigmoideal e uma camada de saída com um neurônio e função de ativação linear

A estimativa do intervalo de confiança foi alcançada por meio do processo *bootstrap residual*, conforme método proposto na Seção 5, que foi aplicado às séries (CW) do Quadro 1. A quantidade de reamostragens obtidas foi igual a 20 ( $B = 20$ ), justificada pela baixa variabilidade do desvio padrão do estimador. As séries *bootstrapadas*, 20 para cada CW do Quadro 1, foram modeladas de acordo com (3). Para cada configuração de rede utilizada foram realizadas dez simulações, escolhendo-se a rede com menor erro quadrático médio na amostra de validação, observando que as configurações das RNAs utilizadas na modelagem das séries *bootstrapadas* são semelhantes às mencionadas no Quadro 1, consequência da proximidade das séries desse quadro e suas previsões, levando a valores relativamente pequenos para os resíduos usados na montagem das séries *bootstrapadas*.

Com o intuito de se analisar a dispersão entre a amostra de teste e respectivas previsões, calculou-se o coeficiente de correlação (0,99) entre esses dois conjuntos de dados. Para a amostra de teste, obteve-se MAPE e MAE iguais a 0,89% e 107,38, respectivamente. Os valores de MAPE, MAE e coeficiente de correlação são indicativos da acuracidade das previsões obtidas com o método proposto.

Na Figura 1, estão representadas as previsões das vazões (oriundas do procedimento *bootstrap*) e respectivo intervalo de confiança.

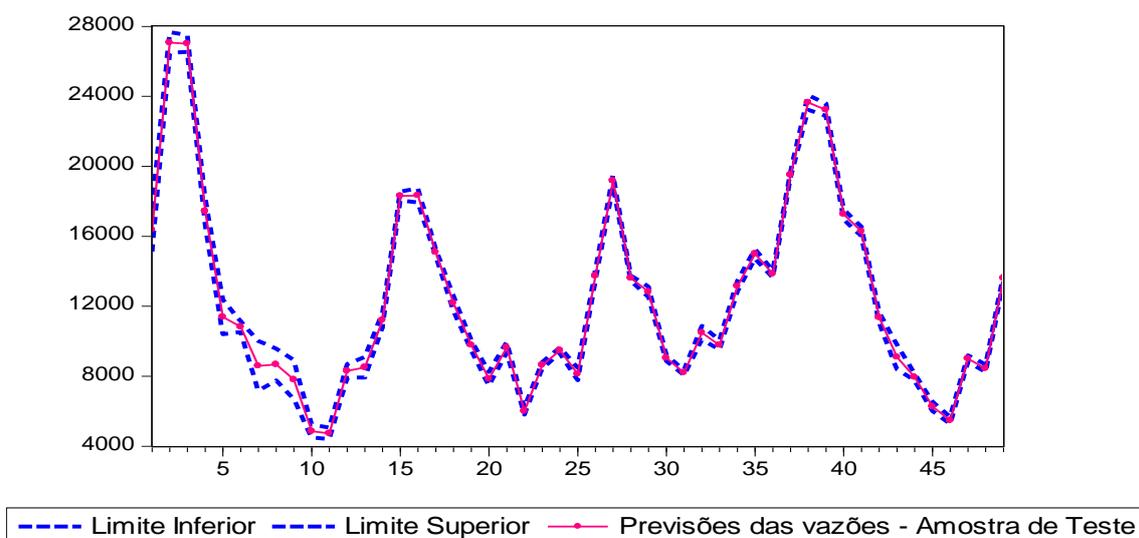


FIGURA 1 - IC previsões da série de vazões – amostra de teste. Fonte: o autor (2015).

## 7 Considerações Parciais

O presente trabalho apresentou um método de modelagem de séries temporais que envolve combinação de filtragem e decomposição *wavelet* no pré-processamento do dados, modelando os resultados deste estágio por meio de redes neurais artificiais e com estimativa do intervalo de confiança por *bootstrap*.

Os baixos valores das medidas de acurácia e a alta correlação entre os valores observados e previsões apontam para a relativa eficiência preditiva do método proposto para as projeções pontuais. Por último, estimou-se as incertezas que cercam o modelo regressor neural. Tibshirani (1996) utilizou no seu trabalho a média aritmética simples das previsões das séries *bootstrapadas*. Nesta artigo foi proposto o uso da média ponderada, procurando com essa estratégia penalizar os resultados com maior erro preditivo. Os resultados obtidos mostraram que se fazendo modelagens adequadas das séries *bootstrapadas* podem levar a uma redução dos erros de previsão do modelo inicialmente ajustado no procedimento *bootstrap residual*. Em contrapartida o custo computacional para a execução desse procedimento é bastante alto, onerando o tempo para execução do mesmo.

## Referências

AQUINO, R. R. B.; LIRA, M. M. S.; OLIVEIRA, J. B.; CARVALHO J.; M. A.; NETO, O. N.; ALMEIDA, G. J. Application of *Wavelet* and Neural Network Models for Wind Speed and Power Generation Forecasting in a Brazilian Experimental Wind Park. *Anais...* International Joint Conference on Neural Networks. Atlanta, USA, 2009.

- BALLINI, R. *Análise e Previsão de Vazões Utilizando Modelos de Séries Temporais, Redes Neurais e Redes Nebulosas*. 2000. 169p. Tese (Doutorado em Engenharia Elétrica e Computação) – UNICAMP, Campinas, 2000.
- DAUBECHIES, I. Ten Lectures on *Wavelets* (C B M S N S F Regional Conference Series in Applied Mathematics), v. 61, *Soc for Industrial & Applied Math*, Philadelphia, 1992.
- DONOHO, D. L.; JONHSTONE, I. M. Ideal Spatial Adaptation by *Wavelet Shrinkage*. **Biometric**, 81,425-455, 1994.
- DONOHO, D. L.; JONHSTONE, I. M.; KERKYACHARIAN, G.; and PICARD, D. *Wavelet Shrinkage: Asymptopia?* (with discussion) *Journal Royal Statistical Society. B*, 57:301-369, 1995.
- EFRON, B. *Bootstrap methods: another look at the Jackknife*. *The Annals of Statistics*, v.7, n. 1, p. 1-26, 1979.
- HAYKIN, S., *Redes Neurais Princípios e Aplicações*, 2. Ed. Porto Alegre: Bookman, 2001.
- HASSANI, H.;ZHIGLJAVSKY, A. Singular Spectrum Analysis: Methodology and Application to Economics Data. *Journal of Systems Science and Complexity*, v. 22, p.372-394, 2009.
- HESKES, T. Practical confidence and prediction intervals. *Advances in neural information processing systems*, v.9, p. 466-472, 1997.
- JOHNSTONE, I.M.; SILVERMAN, B. *Wavelet Threshold Estimators for data with Correlated Noise*. **Stanford University Technical Report**, 1996.
- KHOSRAVI, A. *et al.* Constructing Optimal Prediction Intervals by Using Neural Networks and *Bootstrap Method*. *Neural networks and learning systems*, p. 1-6, 2014.
- KUBRUSLY C. S., LEVAN, N. Abstract *Wavelets* Generated by Hilbert Space Shift Operators. *Adv. Math. Sci. Appl.* 14, 643-660,2006.
- MALLAT, S. *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, 1998.
- MONTGOMERY, D. C.; RUNGER, G. C, *Estatística aplicada e probabilidade para engenheiros*, 2.ed . Rio de Janeiro: LTC, 2003.
- MORETTIN, P. A. *Ondas e Ondaletas: Da análise de Fourier à Análise*. São Paulo: EDUSP, 1999.
- TEIXEIRA Jr., L. A. *et al.* Redes neurais artificiais *wavelet* causal na previsão da radiação solar direta. **Anais...** Rio de Janeiro: Simpósio Brasileiro de Pesquisa Operacional 2012.
- TEIXEIRA Jr., L. A. *Combinação SSA-Wavelet de métodos preditivos com ajuste numérico MINIMAX, na geração de previsões e de cenários*. Rio de Janeiro, 114 p. Tese (Doutorado em Engenharia Elétrica). Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2013.
- TIBSHIRANI, R. A comparison of some error estimates for neural network models. *Neural Computation*, v.8, p. 152-163, 1996.
- VALENZUELA, O. Hybridization of intelligent techniques and ARIMA models for time series prediction. *Fuzzy Sets and Systems*, v.159, p. 821-845, 2008.

WANG, A.B.W. *et al.* Forecasting daily streamflow using hybrid ANN models. Amsterdam: *Journal of Hydrology*, v. 324, p. 383-399, 2006a.

WANG, A.B.W. *et al.* Testing for nonlinearity of streamflow processes at different time scales. Amsterdam: *Journal of Hydrology*, v. 322, p. 247-268, 2006b.