

# Forecasts of monthly flows in the post-266 Itaipu using neural networks with wavelet shrinkage

## Previsões de vazões mensais no posto 266-Itaipu mediante redes neurais com encolhimento wavelet

Levi Lopes Teixeira<sup>1</sup>  
Luiz Albino Teixeira Jr<sup>2</sup>  
Paulo Henrique Siqueira<sup>3</sup>

<sup>1</sup>UTFPR / Departamento de Matemática e Estatística  
prof.levilopes@gmail.com

<sup>2</sup>UNILA / Departamento de Estatística  
luiz.a.t.junior@gmail.com

<sup>3</sup>UFPR / Departamento de Expressões Gráficas  
paulohscwb@gmail.com

**Abstract:** In this paper forecasts of the series of monthly flows were conducted in the post 266 (Itaipu) located in the Paraná River, Brazil. To this end, three methods were used, namely, Box-Jenkins, Neural Networks (ANN) and Wavelet\_ANN. The forecasts obtained with these methods were compared to those provided by the method proposed in this paper, formed by a combination of the wavelet shrinkage and the method Wavelet\_ANN. In the test sample, the mean absolute percentage error (MAPE) obtained with the proposed method was 1.30%, against 18.01%, 20.12% and 2.8% obtained by the methods Box-Jenkins, ANN and Wavelet\_ANN respectively.

**Keywords:** Wavelet, artificial neural networks, time series.

**Resumo:** Neste artigo foram realizadas previsões da série de vazões mensais no posto 266 (Itaipu) localizado no rio Paraná, Brasil. Para tanto, foram usados os métodos de Box-Jenkins, Redes Neurais Artificiais (RNA) e Wavelet\_RNA. As previsões obtidas com esses métodos foram comparadas às fornecidas pelo método proposto neste artigo, formado por uma combinação do método Wavelet\_RNA e encolhimento

wavelet. Na amostra de teste, o erro médio percentual absoluto (MAPE) obtido com o método proposto foi de 1,30%. Com os métodos individuais de Box-Jenkins e RNA ele ficou em 18,01% e 20,12, respectivamente e 2,8% com o método composto Wavelet\_RNA.

**Palavras-chave:** wavelet, redes neurais artificiais, séries temporais.

### 1 Introdução

Estimativas da Empresa de Pesquisa Energética (EPE) apontam que o consumo de energia elétrica no Brasil em 2020 será de 659.092 GWh, sendo predominante a produção hidráulica em virtude da situação privilegiada do país em relação aos recursos hídricos. A produção de energia a partir das usinas hidrelétricas tem ligação direta com a quantidade de água (vazão afluente) disponível para a movimentação das turbinas geradoras e longos períodos de seca podem levar a uma diminuição da capacidade geradora de energia da usina. O Operador Nacional do Sistema Elétrico (ONS) faz previsões de vazões naturais para os locais de aproveitamento hidrelétricos que constituem o Sistema Interligado Nacional (SIN). Um dos modelos mais utilizados nas previsões feitas pelo ONS é o PREVIVAZ, um modelo estocástico univariado e desenvolvido pelo Centro de Pesquisas de Energia Elétrica (CEPEL) [1].

Entre os métodos usados pelo ONS, ressalta-se a frequência de uso dos métodos de Box-Jenkins [2], eficientes para séries oriundas de processos lineares e com estruturas matemáticas relativamente simples. Os modelos Box-Jenkins para previsão de séries temporais são pertencentes à família dos autorregressivos médias móveis (ARMA) subdividido em dois outros modelos: o autorregressivo (AR) e médias móveis (MA). O segundo (MA) são processos de ordem  $q$  em que cada observação é gerada por uma média ponderada dos erros aleatórios  $q$  períodos no passado. O método AR modela uma autorregressão da variável observável com ela mesma, defasada para os  $p$  períodos de defasagem em que a autocorrelação parcial entre as variáveis é significativa. Quando a série é não estacionária sugere-se que seja feita a diferenciação, neste caso o modelo em questão pode ser o ARIMA  $(p,d,q)$ , onde  $d$  representa a ordem de diferenciação.

Segundo Wang *et al.*, [3], os processos de geração de vazões são sazonal e não linear. Os mecanismos não lineares são mais fortes em escalas menores de tempo, sendo as vazões diárias e mensais não lineares [4]. Neste caso, os métodos de Box-Jenkins não são recomendados, sendo indicado outro tipo de modelador, como as redes neurais artificiais (RNA), por exemplo. Revisando 20 anos (1985-2005) de pesquisa sobre previsão de séries temporais, Gooijer e Hyndman [5] concluíram que importantes progressos foram feitos nesta área. Salientam que encontraram numerosos estudos que atestam o sucesso das redes neurais artificiais (RNA) na previsão de séries temporais.

Da Teoria Wavelet advêm importantes métodos auxiliares de pré-processamento de uma série de tempo (como, por exemplo, a sua decomposição wavelet - revisada na Seção 2.1), antes de sua efetiva modelagem. Desde que sejam

utilizados de forma adequada, é possível obterem-se ganhos preditivos substanciais. Pode-se verificar tal fato, dentre muitas outras referências, em: [6–9].

A previsão de séries temporais ruidosas envolve duas abordagens básicas. Na primeira, ignoram-se os ruídos e modela-se a série na expectativa de se capturarem 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 [10]. Menezes *et al.* [11] modelaram séries temporais por meio de vários métodos preditivos, comparando os resultados de previsões oriundas das séries observadas (dados não tratados) com os resultados obtidos a partir das séries filtradas. Concluíram que a filtragem de séries temporais antes da modelagem apresenta ganho significativo nas previsões.

Neste artigo é proposto um método de previsão de séries temporais que usa encolhimento wavelet para tornar a série temporal observada menos ruidosa. A série obtida é decomposta ortogonalmente via wavelet e as componentes previstas por RNA. A previsão da série observada resulta da soma das previsões das componentes wavelet. A fim de mensurar a qualidade do método proposto foram realizadas previsões um passo à frente da série de vazões mensais no posto 266 (Itaipu) localizado no rio Paraná, Brasil. Os resultados encontrados com o método proposto foram comparados aos obtidos com os métodos de Box-Jenkins, RNA e Wave\_RNA. Constatou-se uma superioridade preditiva do método proposto em relação aos outros três métodos aplicados neste estudo.

## 2 Análise wavelet

A análise wavelet propicia a decomposição de uma série temporal em componentes ortogonais, de forma a facilitar seu estudo. Embora a análise de Fourier seja uma das formas mais tradicionais para o tratamento de sinais, as transformadas wavelets são mais completas, fornecendo uma melhor descrição do sinal. Karthikeyan e Kumar [12], Tiwari e Chatterjee [13], Kisi [14], Kisi e Cimen [15] e Nalley *et al.* [16] utilizaram decomposição wavelet em seus trabalhos de previsão de séries temporais e concluíram o bom desempenho desta técnica.

### 2.1 Decomposição wavelet

Enquanto na análise de Fourier o sinal é decomposto em ondas senoidais de várias frequências, na análise wavelet o sinal é decomposto a partir de informações de escala (compressão e dilatação) e translação de uma função wavelet  $\omega_{m,n}(\cdot) = 2^{\frac{m}{2}} \omega(2^m(\cdot) - n)$ , onde  $m, n \in \mathbb{Z}$ . Segundo Levan e Kubrusly [17], o parâmetro  $m$  é chamado de parâmetro de escala e  $n$ , de parâmetro de translação. Ao contrário das funções trigonométricas seno e cosseno que possuem energia infinita e são usadas na decomposição de Fourier, as funções wavelet são de suporte compacto, isto é, possuem domínio diferente de zero em uma extensão finita e igual a zero em todo o resto. A função  $\omega_{m,n}$  faz o mapeamento das componentes de detalhes (alta

frequência) do sinal e com a função escala  $\phi_{m,n}(\cdot) = \frac{1}{2^{\frac{m}{2}}} \phi(2^m(\cdot) - n)$  obtém-se a componente de aproximação que mapeia a baixa frequência do sinal.

Entre as famílias de funções wavelet a mais antiga é a de Haar, definida em (1).

$$\omega(x) = \begin{cases} 1, & \text{se } x \in [0, 1/2) \\ -1, & \text{se } x \in [1/2, 1) \\ 0, & \text{caso contrário} \end{cases} \quad (1)$$

Em (2) encontra-se a definição da função escala de Haar.

$$\phi(x) = \begin{cases} 1, & \text{se } x \in [0, 1) \\ 0, & \text{caso contrário} \end{cases} \quad (2)$$

Nas figuras 1 e 2 estão representados os gráficos das funções escala e wavelet de Haar, respectivamente.

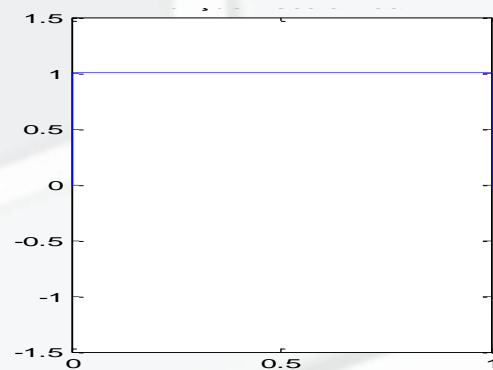


Figura 1: Função escala de Haar

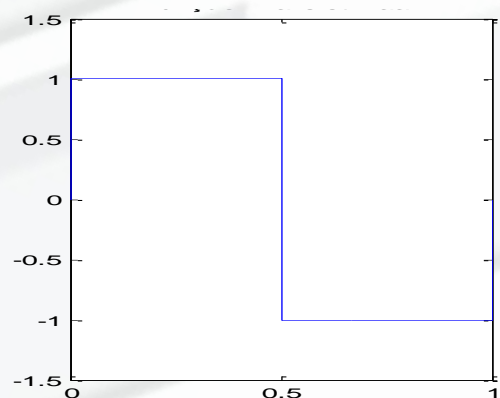


Figura 2: Função wavelet de Haar.

Baseado em Kubrusly e Levan [18] e Mallat [19], o conjunto formado pela união de  $\{\phi_{m_0,n}\}_{n \in \mathbb{Z}}$  e  $\{\omega_{m,n}\}_{m \geq m_0}$  é uma base ortonormal wavelet do espaço de Hilbert  $l^2$ . Segue que a decomposição wavelet da função  $f(\cdot) \in l^2$  pode ser escrita segundo a expansão (3).

$$f(\cdot) = \sum_{n \in \mathbb{Z}} a_{m_0, n} \phi_{m_0, n} + \sum_{m=m_0}^{\infty} \left( \sum_{n \in \mathbb{Z}} d_{m, n} \omega_{m, n} \right) \quad (3)$$

onde  $a_{m_0, n}$  é chamado de coeficiente de aproximação e é calculado a partir do produto interno entre  $f(\cdot)$  e  $\phi_{m_0, n}$ . Os coeficientes de detalhes são representados por  $d_{m, n}$  e obtidos por meio do produto interno entre  $f(\cdot)$  e  $\omega_{m, n}$ .

## 2.2 Encolhimento wavelet

No caso da série temporal ser fortemente ruidosa, recomenda-se sua filtragem, para então modelar a parte determinística obtida com a filtragem. Seja  $(y_t)_{t=1}^T$  uma série temporal, onde para todo  $t$ :  $y_t = f_t + e_t$ . Sendo  $e_t$  um estado da variável aleatória  $\varepsilon_t \sim (0, \sigma^2)$ , onde  $f_t$  é a componente determinística de  $y_t$ . Como não é possível modelar  $e_t$  e sendo a série excessivamente ruidosa, indica-se o tratamento dos dados formadores da série temporal a fim de se obter a componente determinística. O encolhimento (*shrinkage*) wavelet é uma das técnicas destinadas para este fim.

O Princípio do encolhimento wavelet tem como objetivo a redução (ou mesmo remoção) do ruído presente num sinal, diminuindo (ou zerando) a magnitude dos coeficientes wavelet [20]. Outras referências nesta área são os trabalhos de Donoho e Johnstone [21] e Danoho *et al.* [7].

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, as mais tradicionais são o limiar duro (*hard threshold*) e limiar suave (*soft threshold*), representadas em (4) e (5), respectivamente.

$$d_{m, n}^H = \begin{cases} d_{m, n}, & \text{se } |d_{m, n}| > \lambda \\ 0, & \text{se } |d_{m, n}| \leq \lambda \end{cases} \quad (4)$$

$$d_{m, n}^S = \begin{cases} \text{sign}(d_{m, n})(|d_{m, n}| - \lambda), & \text{se } |d_{m, n}| > \lambda \\ 0, & \text{se } |d_{m, n}| \leq \lambda \end{cases} \quad (5)$$

onde  $d_{m, n}$  é o coeficiente de detalhe com ruído,  $d_{m, n}^H$  e  $d_{m, n}^S$  são coeficientes de detalhes sem ruído,  $\text{sign}(d_{m, n})$  é o sinal de  $d_{m, n}$  e  $\lambda$  o limiar. O *hard threshold* gera uma descontinuidade, um “salto” em  $\lambda$ , eliminando ou mantendo o coeficiente de detalhe. Enquanto o limiar suave atenua por  $(|d_{m, n}| - \lambda)$  os valores de  $d_{m, n}$  (se  $|d_{m, n}| > \lambda$ ), eliminando-os em caso contrário. Nos gráficos das figuras 3 e 4 podem ser observadas essas características.

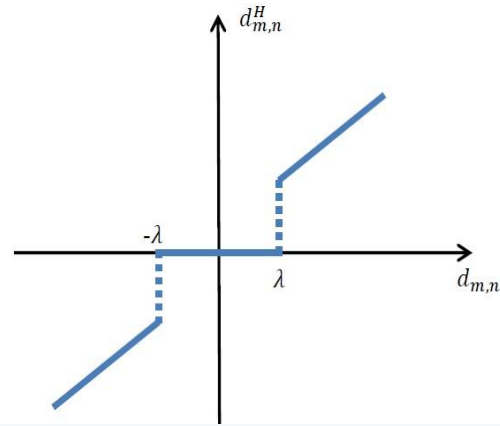


Figura 3: Gráfico da função limiar duro.

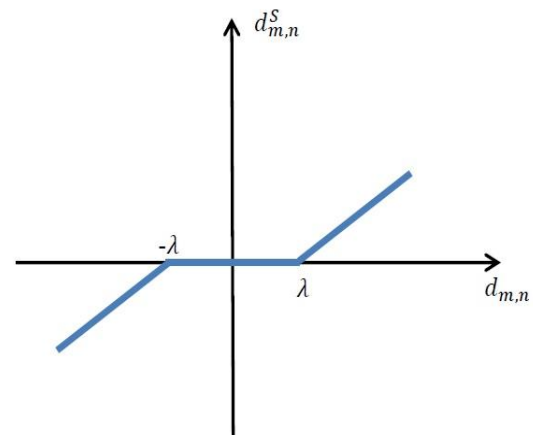


Figura 4: Gráfico da função limiar suave.

Neste artigo 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 [21], Johnstone e Silverman [22] e Morettin [20].

De acordo com Donoho e Johnstone [23] o cálculo do limiar universal é dado por (6).

$$\lambda = \sigma \sqrt{2 \log(T)} \quad (6)$$

sendo  $\sigma$  o nível do ruído e  $T$  a cardinalidade da série. O limiar universal é independente da escala  $m$ , assumindo um valor global.

Donoho e Johnstone [21] propuseram selecionar um limiar, pela minimização do estimador não-viesado do risco Stein (SURE), em cada escala  $m$ . Seja  $N_m$  o número de coeficientes na escala  $m$ , então o limiar de SURE é calculado a partir de (7).

$$\lambda_m = \underset{0 \leq \lambda \leq \sqrt{2 \log(N_m)}}{\text{argmin}} \text{SURE}(y_m, \lambda) \quad (7)$$

### 3 Redes neurais artificiais

O mais eficiente sistema de aprendizado existente na natureza é o cérebro humano. Um dos tipos de células que o compõem é o neurônio, capaz de processar e transmitir informações. Os neurônios estão ligados uns aos outros formando uma rede neural. O fluxo das informações (sinais) por esta rede depende de uma fenda existente entre os neurônios denominada de fenda sináptica, ou sinapse. Segundo Castro [24], depois que todos os sinais de entrada são recebidos e ponderados pelas sinapses eles são somados no corpo celular, gerando um sinal de entrada líquida do neurônio, e que será posteriormente propagado para a saída do neurônio.

As Redes Neurais Artificiais (RNA) foram idealizadas com base nos processos de aprendizagem do cérebro humano, de forma que a RNA é capaz de simular as conexões sinápticas. De acordo com Haykin [25] a 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. Entre as RNAs, um dos tipos mais usados é a multilayer feedforward que é composta por várias camadas e os sinais são propagados apenas da entrada para a saída, nunca ao contrário.

O principal algoritmo de treinamento de redes neurais artificiais é o backpropagation, cujo ajuste dos pesos sinápticos ocorre por meio de um processo de otimização realizado em duas fases: forward e backward. Na fase forward, é calculada a resposta fornecida pela RNA para determinado padrão de entrada. Na fase backward, o desvio (ou erro) entre a resposta da RNA e a resposta desejada é utilizado no processo de ajuste dos pesos sinápticos. Ao longo do treinamento da RNA, os vários padrões de entrada e as respectivas respostas desejadas são apresentados à RNA, de forma que os pesos sinápticos sejam tais que minimizem a soma dos erros quadráticos (MSE) [9].

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

### 4 Método Box-Jenkins

Box e Jenkins [26] apresentaram metodologia que fez a integração de conhecimentos existentes na época. A metodologia consiste em ajustar modelos autorregressivos integrados de médias móveis (ARIMA) a um conjunto de dados a partir de três estágios: identificação do modelo, estimação dos parâmetros e verificação do modelo ajustado. De acordo com Box e Jenkins [26] um modelo

ARMA plausível para a série temporal  $(y_t)_{t=1}^T$ ,  $T$ =cardinalidade, de ordens  $p$  (autorregressivo) e  $q$  (médias móveis) é descrito pela equação (8).

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} + a_t \quad (8)$$

O modelo (8) combina valores passados das entradas  $y_t$  e choques aleatórios  $(a_t)$  decorrelacionados, de média zero e variância constante. Neste modelo  $\phi_i, \theta_j \in \mathbb{R}$ , com  $i = 1, \dots, p$  e  $j = 1, \dots, q$ , denotam os parâmetros do modelo e  $\delta$  uma constante. Na hipótese da série temporal ser não estacionária, esta deve ser diferenciada e o modelo ARMA( $p, q$ ) substituído pelo ARIMA( $p, d, q$ ), sendo  $d$  a ordem de diferenciação da série. Para a identificação do modelo, as ordens  $p$  e  $q$  podem ser determinadas por meio da análise do perfil dos gráficos das funções de autocorrelação (FAC) e autocorrelação parcial (FACP). Identificado o modelo, passa-se ao estágio seguinte que é a estimação dos parâmetros. Para tanto, é necessário utilizar métodos iterativos não lineares de mínimos quadrados. Maiores detalhes podem ser encontrados em Box e Jenkins [26] e Morettin e Tolo [27]. Para a validação do modelo já com os parâmetros estimados podem-se usar testes estatísticos, tais como: teste de Box-Pierce, teste do periodograma acumulado e teste da autocorrelação cruzada. No caso da série temporal  $(y_t)_{t=1}^T$  apresentar componente sazonal, o modelo Box-Jenkins plausível é dado genericamente, em (9).

$$\phi(B)(1 - \phi_1 B - \dots - \phi_p B^p) \nabla^d (1 - B^S)^D y_t = \theta(B)(1 - \theta_1 B - \dots - \theta_q B^q) a_t \quad (9)$$

onde:

$\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$ ,  $\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ ,  $d$  é a ordem das diferenças simples,  $D$  é a ordem das diferenças sazonais,  $S$  é o período sazonal,  $\phi_k \in \mathbb{R}$  e  $\theta_j \in \mathbb{R}$  são os coeficientes dos polinômios não sazonais e  $\phi_m \in \mathbb{R}$  e  $\theta_n \in \mathbb{R}$  são os coeficientes dos polinômios sazonais.

### 6 Materiais e métodos

A seguir apresentam-se as informações referentes às séries de vazões usadas neste trabalho e o método proposto.

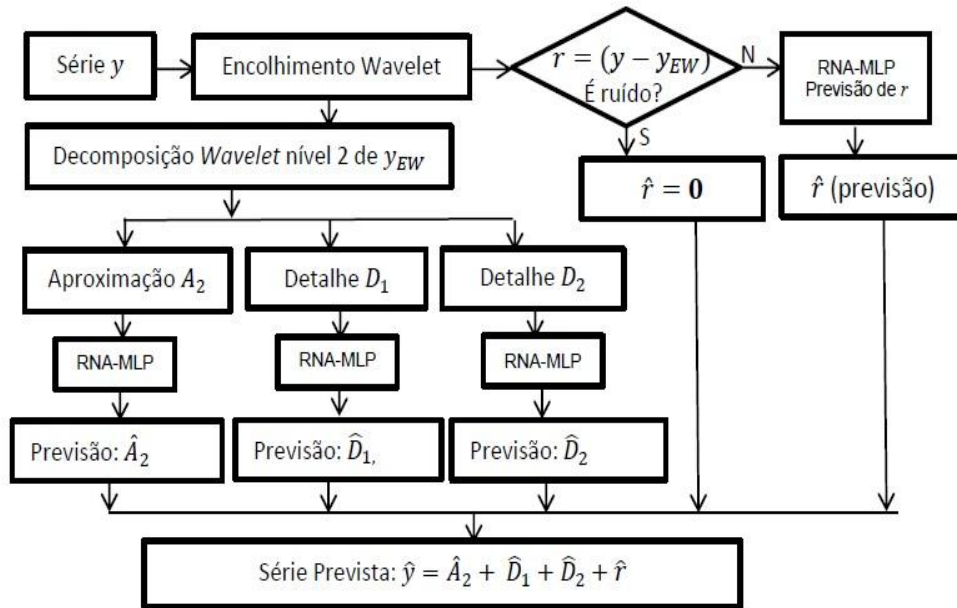


Figura 5: Fluxograma com as etapas do método proposto - descrito na seção 6.2

### 6.1 Materiais

A série de vazões mensais  $(y_t)_{t=1}^{492}$  usada neste trabalho envolve quarenta e um anos de observações no período 1970-2010 e pode ser obtida no site do ONS, com endereço eletrônico:

[http://www.ons.org.br/operacao/vazoes\\_naturais.aspx](http://www.ons.org.br/operacao/vazoes_naturais.aspx).

O primeiro registro de vazão natural média mensal, no site supracitado, é referente ao mês de janeiro de 1931. Os dados que compõem a série são provenientes das medições de vazões no posto 266, nele são obtidas as vazões afluentes à usina hidrelétrica de Itaipu, que está localizada no Rio Paraná, no trecho de fronteira entre o Brasil e o Paraguai, a 14 km ao norte da Ponte da Amizade. A área da usina se estende desde Foz do Iguaçu, no Brasil, e Ciudad del Este, no Paraguai, ao sul, até Guairá (Brasil) e Salto del Guairá (Paraguai), ao norte.

### 6.2 Método Proposto

Seja a série temporal  $\{y_t\}_{t=1}^T = \{f_t\}_{t=1}^T + \{\varepsilon_t\}_{t=1}^T$ , onde  $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$ . Inicialmente, submete-se a série  $\{y_t\}_{t=1}^T$  ao encolhimento wavelet, obtendo-se a série  $\{y_{t,EW}\}_{t=1}^T \cong \{f_t\}_{t=1}^T$ . Os testes de independência estatística BDS e de estacionariedade *Dickey-Fuller* são aplicados à componente total de resíduo  $(r = \{y_t\}_{t=1}^T - \{y_{t,EW}\}_{t=1}^T)$  a fim de se obter indicação que ela pode ser classificada estatisticamente como ruído branco. No caso de  $r$  não ser classificado como ruído, a sua previsão (obtida por RNA) irá compor a previsão da série observada.

Considere a sequência infinita  $\{\tilde{y}_t\}_{t \in \mathbb{Z}}$  obtida da série temporal  $\{y_{t,EW}\}_{t=1}^T$  a partir da igualdade  $\tilde{y}_t = \{y_{t,EW}, t \in \{1, \dots, T\}\}$ . Segundo Teixeira Jr *et al.* [9] a decomposição wavelet de nível  $p$  de uma série temporal  $\tilde{y}$  em relação a uma determinada base ortonormal wavelet é representada por (10).

$$\tilde{y} \cong \tilde{\tilde{y}} = \sum_{n \in \Gamma_{V_{m_0}}} a_{m_0,n} \phi_{m_0,n} + \sum_{m=m_0}^{m_0+(p-1)} \left( \sum_{n \in \Gamma_{W_m}} d_{m,n} \omega_{m,n} \right) \quad (10)$$

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

O encolhimento wavelet e as componentes de aproximação  $\sum_{n \in \Gamma_{V_{m_0}}} a_{m_0,n} \phi_{m_0,n}$  e detalhe  $\sum_{n \in \Gamma_{W_m}} a_{m_0,n} \phi_{m_0,n}$  foram obtidas a partir do aplicativo computacional Matlab 8.0, que oferece várias opções de bases wavelets, tais como: as famílias daubechies (dbN), coiflets (coifN) e symlets (symN), desenvolvidas por Ingrid Daubechies [28].

Neste trabalho a série de vazões mensais foi decomposta ortogonalmente via wavelet de nível  $p = 2$ , que resulta em duas componentes de detalhe ( $D_1$  e  $D_2$ ) e uma de aproximação ( $A_2$ ), utilizando base da família Daubechies com momento igual a 10 (db 10).

Uma vez obtidas as componentes wavelet de aproximação e detalhe, estas foram modeladas via três redes neurais artificiais *feedforward multilayer perceptron* (RNA-MLP),

resultando as previsões das componentes wavelet de detalhe ( $\hat{D}_{1,RNA}$  e  $\hat{D}_{2,RNA}$ ) e aproximação ( $\hat{A}_{2,RNA}$ ). A previsão da série observada é dada pela soma das previsões das componentes wavelet e previsão do resíduo ( $\hat{r}$ ), como mostra (11).

$$\{\hat{y}_t\}_{t=1}^T = \{\hat{A}_{2,RNA}(t)\}_{t=1}^T + \{\hat{D}_{1,RNA}(t)\}_{t=1}^T + \{\hat{D}_{2,RNA}(t)\}_{t=1}^T + \{\hat{r}_t\}_{t=1}^T \quad (11)$$

Para a avaliação dos métodos foram utilizados os Erros MAPE (Erro Médio Percentual Absoluto) e MAE (Erro Médio Absoluto), representados em (12).

$$MAPE = \frac{1}{T} \sum_{t=t_0}^T \left| \frac{y_t - \hat{y}_t}{y_t} \right| e$$

$$MAE = \frac{1}{T} \sum_{t=t_0}^T |y_t - \hat{y}_t| \quad (12)$$

Na figura 5 podem ser observadas as etapas do método proposto.

## 7 Resultados obtidos

Apresentam-se a seguir os resultados obtidos com as abordagens investigadas e com o método proposto neste trabalho.

### 7.1 Métodos Individuais Box & Jenkins e RNA

Aplicando-se a metodologia Box-Jenkins na análise da série de vazões mensais, chegou-se à conclusão que o modelo ARIMA(6,0,6)x(1,1,2)<sub>12</sub> é o que melhor se ajusta aos dados da série em estudo transformada pela aplicação de logaritmo natural.

A RNA usada na modelagem da série de vazões é do tipo *feedforward multilayer perceptron* com tamanho da janela de entrada igual a 5, uma camada escondida composta por cinco neurônios artificiais com função de ativação tangente hiperbólica sigmoidal (*tansig*) e um neurônio na camada de saída com função de ativação linear (*purelin*).

A amostra de teste, formada pelas últimas 49 observações, apresentou MAPE igual a 18,01% com o modelo ARIMA e 20,12% com a modelagem por RNA. Nas figuras 6 e 7 estão os gráficos de dispersão da amostra de teste versus previsão por ARIMA e RNA, respectivamente. Os gráficos em questão são um indicativo da baixa qualidade das previsões obtidas com os métodos individuais ARIMA e RNA.

### 7.2 Método Composto Wavelet\_RNA

Inicialmente, a série de vazões foi decomposta ortogonalmente via wavelet de nível  $p = 2$ , utilizando a base daubechies com momento 10 (db 10). As três componentes (duas de detalhes e uma de aproximação) obtidas com a decomposição foram ajustadas por três RNA e as previsões somadas a fim de se obter a série observada.

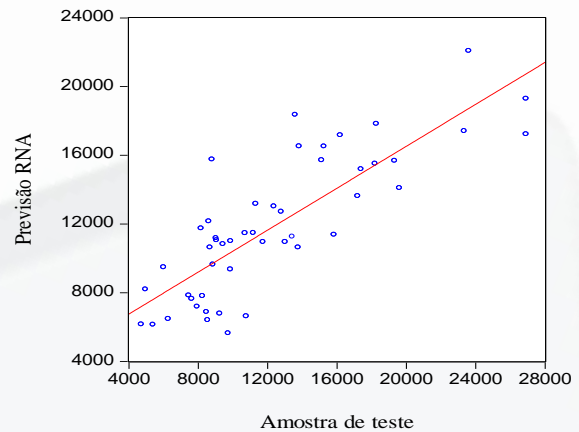


Figura 6: Diagrama de dispersão entre a amostra de teste e a previsão por RNA.

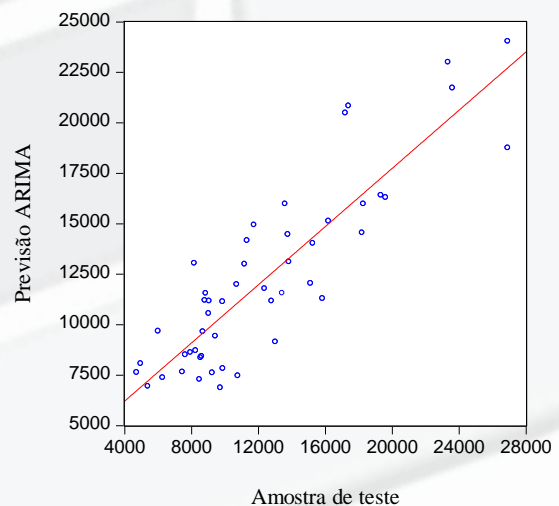


Figura 7: Diagrama de dispersão entre amostra de teste e previsão por ARIMA.

O método wavelet-RNA se diferencia do método proposto basicamente em duas etapas: encolhimento wavelet e modelagem do resíduo. Na figura 8 podem ser observadas as etapas para a execução do método Wavelet\_RNA.

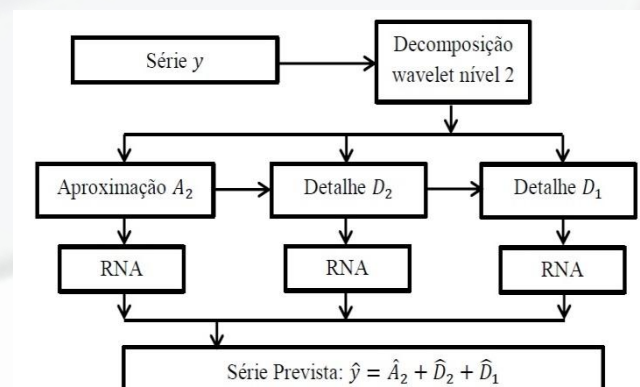


Figura 8: Fluxograma com as etapas do método Wave\_ARIMA.

Nas figuras 9, 10, 11 e 12 podem-se observar os gráficos da série de vazões mensais e as componentes wavelets de aproximação (A2) e detalhes (D1 e D2) obtidas com a base ortonormal wavelet de Daubechies com momento 10 (db 10). Sendo a série de vazões mensais aproximadamente igual à soma de A2, D1 e D2.

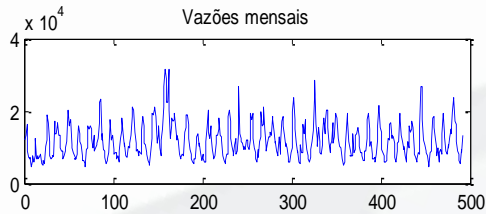


Figura 9: Vazões mensais – série observada.

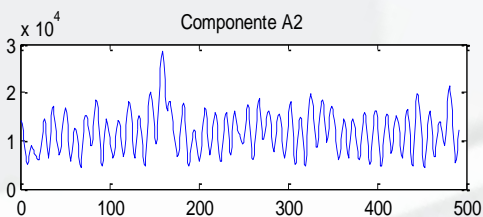


Figura 10: Componente wavelet de aproximação A2.

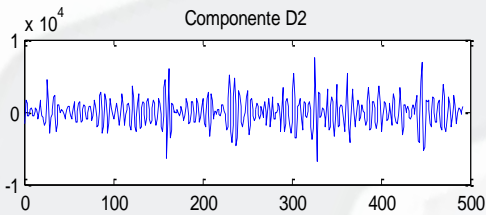


Figura 11: Componente wavelet de detalhe D2.

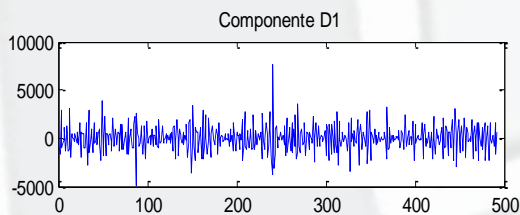


Figura 12: Componente wavelet de detalhe D1.

O quadro 1 traz as configurações das RNA utilizadas na modelagem das componentes wavelet A2, D1 e D2, onde CE é a quantidade de neurônios da camada escondida com função de ativação tangente hiperbólica (*tansig*) e J o tamanho da janela de entrada de dados. A camada de saída possui um único neurônio e a função de ativação é linear (*purelin*).

O MAPE para a amostra de teste obtido com a aplicação do método Wavelet\_RNA foi de 2,80%, mostrando a eficiência desse método em relação aos individuais RNA e ARIMA.

Quadro 1: Configurações das RNA para as componentes A2, D1 e D2.

<b>Componente de aproximação de nível 2 (A2):</b>
<b>RNA_1:</b> <i>feedforward</i> , J=20, CE=9, <i>tansig</i> (camada escondida), <i>purelin</i> (camada de saída).
<b>Componente de detalhe de nível 1 (D1):</b>
<b>RNA_2:</b> <i>feedforward</i> , J=20, CE=5, <i>tansig</i> (camada escondida), <i>purelin</i> (camada de saída)
<b>Componente de detalhe de nível 2 (D2):</b>
<b>RNA_3:</b> <i>feedforward</i> , J=21, CE=9, <i>tansig</i> (camada escondida), <i>purelin</i> (camada de saída)

### 7.3 Método Proposto e Comparativo Entre os Métodos

Para o encolhimento wavelet foram usadas as bases de Daubechies com momentos 1, 5 e 10 (db1, db5 e db10) em níveis  $p=2$  e 3, com regra de limiarização suave e limiar universal e SURE. As componentes utilizadas na previsão da série  $y$  resultaram da decomposição ortogonal wavelet com a base daubechies (db10) em nível  $p=2$ . Os quatro melhores resultados estão expostos na tabela 1.

Com o encolhimento wavelet da série  $y$  obteve-se a série  $y_{EW}$ , supostamente menos ruidosa que a observada. Esta última, por sua vez, foi decomposta via wavelet (db10 e  $p=2$ ) e as componentes obtidas ajustadas por RNA. A soma das previsões das componentes wavelet e resíduo determinou a previsão de  $y$ , conforme descrito na seção 6.2.

A melhor RNA identificada na modelagem 1 (tabela 1) tem a seguinte estrutura topológica: tamanho da janela de entrada igual a 12, uma camada escondida composta por dez neurônios artificiais com função de ativação *tansig* e um neurônio na camada de saída com função de ativação *purelin*. Doravante, os resultados encontrados com a modelagem 1 representaram o método proposto no artigo.

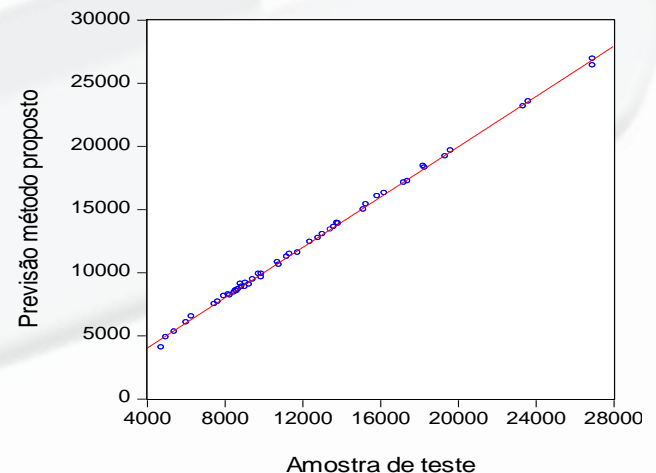


Figura 13: Diagrama de dispersão entre amostra de teste e previsão pelo método proposto.

Tabela 1: Wavelet usadas no encolhimento da série de vazões e MAPE da modelagem das respectivas séries obtidas no encolhimento.

Modelagem	Base	Nível $p$	MAPE		
			Treino	Validação	Teste
1	db1	2	0,48%	1,08%	1,30%
2	db1	3	0,40%	1,27%	2,15%
3	db5	2	1,26%	1,87%	2,44%
4	db10	2	1,21%	2,24%	2,69%

Tabela 2: MAPE e MAE para amostras retiradas da série de vazões posto 266-Itaipu.

Método	MAPE			MAE		
	Treino	Validação	Teste	Treino	Validação	Teste
ARIMA	15,27%	16,27%	18,01%	1962,43	2006,33	2417,88
RNA	17,52%	19,15%	20,12%	1667,91	1946,70	2104,40
Wavelet_RNA	2,11%	2,47%	2,80%	224	232,73	290,79
Método proposto	0,48%	1,08%	1,30%	49,88	105,60	127,48

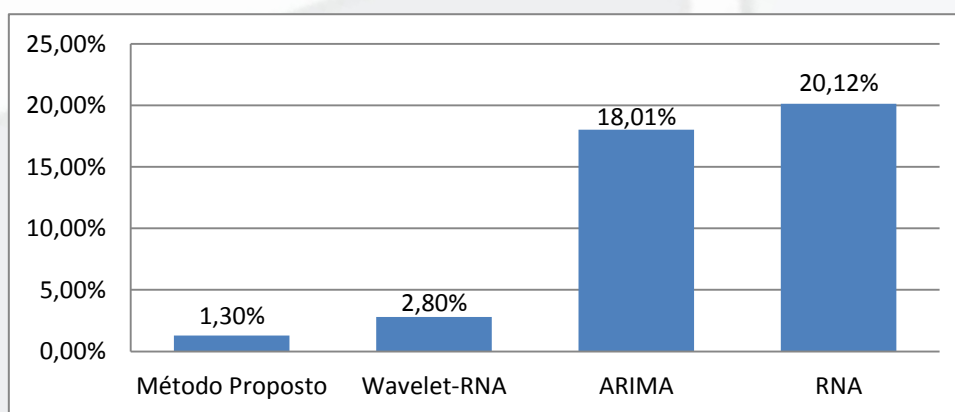


Figura 14: MAPE da amostra de teste para os métodos RNA, ARIMA, wavelet-RNA e proposto.

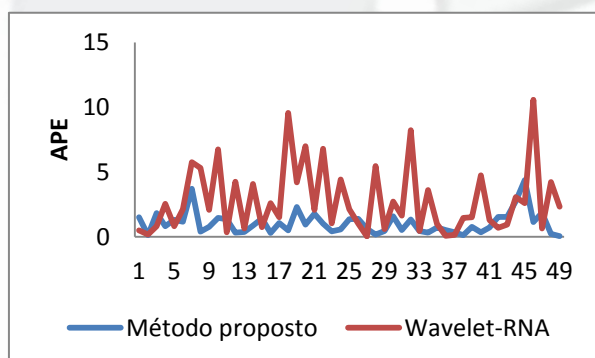


Figura 15: APE na amostra de teste para o método proposto e wavelet-RNA.

No diagrama de dispersão apresentado na figura 13 se observa alta correlação entre a amostra de teste e as previsões alcançadas com o método proposto, sendo um indicativo da qualidade das previsões obtidas.

A tabela 2 informa o MAPE e MAE correspondentes às previsões da série observada feitas um passo à frente, onde se pode observar uma superioridade do método proposto em relação aos outros métodos aplicados neste trabalho.

Na Figura 15, tem-se a comparação entre o método proposto e a abordagem wavelet-RNA. Nota-se que, em todos os 49 instantes, na amostra de teste, o método proposto apresenta evolução temporal dos valores dos erros percentuais absolutos (APE) mais estáveis que os do método wavelet-RNA. Nesse método, os valores do APE violaram em vários instantes a faixa dos 5%, enquanto que em nenhum instante isso ocorre no método proposto. Além disso, na maioria dos instantes, o método proposto apresenta valores de APE inferiores aos do wavelet-RNA.



## 8. Considerações finais

Neste artigo foi proposto um método para previsões de séries que envolve encolhimento wavelet, decomposição wavelet, e modelagem das componentes wavelet via RNA. Comparações com resultados de previsões obtidos por meio de modelos Box-Jenkins e RNA, quando usados individualmente, constata consideráveis ganhos preditivos auferidos com o uso do método proposto. Reduzindo o erro preditivo em 90%, aproximadamente. Na comparação com o método composto Wavelet\_RNA a redução do MAPE foi da ordem de 50%.

Embora os resultados oriundos dos métodos individuais ARIMA e RNA apresentem grandes diferenças em relação aos compostos (wavelet-RNA e proposto), resíduos provenientes daqueles métodos passaram por testes estatísticos que não descartaram a hipótese de ruído branco.

Kisi [14] combinou decomposição wavelet e regressão linear (WR) e comparou os seus resultados aos obtidos com modelagem ARIMA. Constatou que o primeiro método (WR) apresentou melhores resultados de previsão que o segundo. Teixeira Jr *et al.* [8], apresentaram um método em que as componentes *wavelets* de uma série temporal constituem os padrões de entrada de uma RNA *feedforward* MLP, cuja saída fornece a previsão da série temporal. Com essa metodologia, chegaram a resultados melhores aos obtidos por RNA convencional (sem considerar componentes wavelet como padrões de entrada).

Resultados alcançados neste trabalho e corroborados por outros encontrados na literatura, já citados anteriormente, fortalecem a conclusão que a combinação de métodos de filtragem com incorporação da decomposição wavelet pode auferir ganhos consideráveis em relação a outros que não o fazem.

## Referências

- [1] MACEIRA, M.E.P. *et al.* Periodic ARMA models applied to weekly streamflow forecasts. Proceedings of the 1999 IEEE Power tech, Budapest, 1999.
- [2] BOX, G.; JENKINS, G.; REINSEL, G. C. **Time Series Analysis, Forecasting and Control**. 3.ed. Englewood Cliffs:Prentice-Hall, 1994. 598p
- [3] WANG, A.B.W. *et al.* Forecasting daily streamflow using hybrid ANN models. Amsterdam: **Journal of Hydrology**, v. 324, p. 383-399, 2006a.
- [4] 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.
- [5] GOOIJER, J. G.; HYNDMAN, R. J. 25 years of time series forecasting. **International Journal of Forecasting**, p.443– 473, 2006.
- [6] DONOHO, D.L.; JOHNSTONE, I.M. Ideal spatial adaptation by wavelet shrinkage. **Biometrika**, v.81, n.3, p.425-455, 1994.
- [7] DONOHO, D.L. *et al.* **Universal near Minimality of Wavelet Shrinkage**. Stanford University Technical Report, 1995.
- [8] LEI, C. and RAN, L. Short-term Wind Speed Forecasting Model for Wind Farm Based on *wavelet* Decomposition DRPT. **Anais...** Nanjing, China, pp 2525-2529, 2008.
- [9] 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.
- [10] 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.
- [11] MENEZES, L.M. *et al.* Modelagem e Previsão de Demanda de Energia com Filtragem SSA. **Revista de Estatística UFOP**, v.III(2), 2014.
- [12] KARTHIKEYAN, L; KUMAR, D. N. Predictability of nonstationary time series using wavelet and EMD based ARMA models. **Journal of hydrology**, p. 103-119, 2013.
- [13] TIWARI, M. K.;CHATTERJEE, C. Development of an accurate and reliable hourly flood forecasting model using wavelet–bootstrap–ANN (WBANN) hybrid approach. **Journal of Hydrology**, p. 458-470, 2010.
- [14] KISI, O. Wavelet regression model for short-term streamflow forecasting. **Journal of hidrology**, p. 344-353, 2010.
- [15] KISI, O.; CIMEN, M. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. **Journal of hidrology**, p. 132-140, 2011.
- [16] NALLEY, D. *et al.* Using discrete wavelet transforms to analyze trends in streamflow and precipitation in Quebec and Ontario (1954–2008). **Journal of hydrology**, p. 204-228, 2012.
- [17] LEVAN N. and KUBRUSLY, C. S. A Wavelet Time-Shift Detail Decomposition. **Mathematics and Computers in Simulation**. 63, 73-78, 2003.
- [18] KUBRUSLY C. S., LEVAN, N. Abstract Wavelets Generated by Hilbert Space Shift Operators. **Adv. Math. Sci. Appl.** 14, 643-660,2006.
- [19] MALLAT, S. **A Wavelet Tour of Signal Processing**. Academic Press, San Diego, 1998.
- [20] MORETTIN, P. A. **Ondas e Ondaletas: Da análise de Fourier à Análise**. São Paulo: EDUSP, 1999.
- [21] DONOHO, D.L.; JOHNSTONE, I. M. **Adapting Unknown Smoothness via Wavelet Shrinkage**. Division of



- Biostatistics/Stanford University Technical Report, n.161,1995.
- [22] JOHNSTONE, I.M.; SILVERMAN, B. **Wavelet Threshold Estimators for data with Correlated Noise**. Stanford University Technical Report, 1996.
- [23] DOHONO, D.L.; JOHNSTONE, I.M. Minimax Estimation via Wavelet Shrinkage. **The Annals of Statistics**, v.26, n.3, p.879-921. 1998.
- [24] CASTRO, L. N. **Computação Natural**. São Paulo: Editora Livraria da Física, 2010.
- [25] HAYKIN, S., **Redes Neurais Princípios e Aplicações**, 2. Ed. Porto Alegre: Bookman, 2001.
- [26] BOX, G.; JENKINS, G. **Time Series Analysis: Forecasting and Control**. San Francisco: Holden-Day, 1970.
- [27] MORETTIN, P. A. e TOLOI, L. M. C. **Análise Séries Temporais**. 2ª Ed. ABE Projeto Fisher. São Paulo: Edgard Blucher, 2006.
- [28] 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.