

Metodologia híbrida *Wavelet* – SVR para projeção de deslocamentos relativos no bloco I11 da barragem da usina hidrelétrica de Itaipu

Wavelet – SVR hybrid methodology for the projection of relative displacements in block I11 of the Itaipu hydroelectric plant dam

Tásia Hickmann¹ – Univ. Tecnológica Fed. do Paraná – Câmpus Medianeira - Dep. de Matemática e Estatística
Liliana Madalena Gramani² – Univ. Fed. do Paraná – Dep. de Matemática
Eloy Kaviski³ – Universidade Fed. do Paraná – Dep. de Hidráulica e Saneamento
Luiz Albino Teixeira Junior⁴ – Univ. Fed. da Integração Latino-Americana – Dep. de Estatística
Samuel Bellido Rodrigues⁵ – Univ. Tecnológica Fed. do Paraná, Câmpus Medianeira – Dep. de Matemática e Estatística

RESUMO Este artigo propõe uma metodologia híbrida para a previsão das séries temporais de deslocamentos relativos em um bloco da barragem da usina hidrelétrica de Itaipu, que integra os seguintes métodos numéricos: decomposição *wavelet*, *support vector Machine* e combinação linear de previsões. Todos os resultados estatísticos alcançados pela metodologia proposta foram mais acurados do que outras técnicas tradicionais (usadas aqui como *benchmark*), encorajando a sua adoção para tal finalidade.

Palavras-chave Deslocamentos relativos. Decomposição *wavelet*. *Support vector regression*. Combinação linear de previsões.

ABSTRACT *This article proposes a hybrid methodology for forecasting relative displacements in a dam block at the Itaipu hydroelectric plant, which integrates the following numerical methods: wavelet decomposition, support vector machine and linear forecast combination. All the statistical results achieved by the proposed method are more accurate than other traditional techniques (used here as a benchmark), encouraging its adoption for this purpose.*

Keywords *Relative displacements. Wavelet decomposition. Support vector regression. Forecasts.*

1. Avenida Brasil, 4232, CEP: 85884-000, Caixa Postal 271, Medianeira, Paraná, hickmann@utfpr.edu.br
2. l.gramani@gmail.com
3. eloy.dhs@ufpr.br
4. luiz.a.t.junior@gmail.com
5. samuelb@utfpr.edu.br

HICKMANN, T.; GRAMANI, L. M.; KAVISKI, E.; TEIXEIRA JUNIOR, L. A.; RODRIGUES, S. B. Metodologia híbrida *Wavelet* – SVR para projeção de deslocamentos relativos no bloco I11 da barragem da usina hidrelétrica de Itaipu. **GEPROS. Gestão da Produção, Operações e Sistemas**, Bauru, Ano 11, nº 2, abr-jun/2016, p. 103-119.

DOI: 10.15675/gepros.v11i2.1419

1. INTRODUÇÃO

As barragens de concreto de usinas hidrelétricas são, em geral, monitoradas, com o auxílio de instrumentos que realizam aferições de diferentes tipos de movimentos dos blocos que compõe a sua estrutura. Os dados aferidos, ao longo do tempo, podem ser, do ponto de vista estatístico, interpretados como séries temporais estocásticas que exibem estruturas de autodependência (visto que o valor de leitura corrente de um instrumento é afetado pelos valores passados). Com efeito, estas são tais que obedecem a um determinado regime estocástico de comportamento, o qual pode ser mapeado, de forma aproximada, por meio de um método preditivo, com a finalidade de se produzir previsões. Uma vez que tais obras demandam muita atenção devido às diversas condições de segurança estrutural e operacional, o objetivo principal é minimizar as chances consequências catastróficas em suas estruturas, ao meio ambiente ou às vidas humanas; além de custos financeiros exacerbados. Para tanto, a identificação prévia de problemas em sua estrutura de concreto demanda que a equipe técnica responsável disponha de previsões acuradas dos valores futuros dos instrumentos nela instalados.

Baseado em Morettin; Toloi,(2006) e Ehlers, (2009), uma série temporal (estocástica) consiste de um conjunto de observações aleatórias ordenadas no tempo existente nas mais variadas áreas de aplicações como, por exemplo, na engenharia, na economia, na biologia, nas ciências sociais, entre outras ciências. O foco, na maioria das vezes, porém, é a necessidade de se predizer, de forma aproximada, os seus estados futuros. Não obstante, no processo de predição de séries de tempo, tem sido frequentemente assumir que existe um método preditivo individual verdadeiro. No entanto, ao se assumir tal suposição, está-se subestimando o risco estrutural inerente aos dados estocásticos, como também a incerteza associada ao modelo adotado (FIRMINO et al., 2015). Zhang, (2003), destaca ainda que os métodos preditivo híbridos (ou combinados) tende a ser mais seguros e informativas, principalmente em predição de séries de tempo que apresentam mudança de regime estocástico, como é o caso dos dados de instrumentos de barragens.

Um exemplo da obra prima da engenharia, e um dos maiores projetos hidrelétricos do mundo, a usina hidrelétrica de Itaipu está localizada no rio Paraná, à 14 km à montante da ponte internacional que liga a cidade de Foz do Iguaçu, no Brasil, à Ciudad del Este, no Paraguai. A barragem principal de Itaipu é de concreto, do tipo gravidade aliviada, sendo composta por 20 blocos, cada um dotado de uma unidade geradora. No total são 2384 instrumentos instalados no concreto e na fundação, fornecendo dados temporais dos movimentos dos blocos. Alguns blocos são designados blocos-chave (como é o caso do Bloco I11), os quais são dotados de maior quantidade de instrumentos, devido à maior força que recebem da água no sentido montante-jusante, ou seja, no sentido do leito do rio. Esse volume de água também exerce, nas partes inferiores dos blocos, uma pressão que cria um efeito contrário ao que a sua própria massa exerce sobre a fundação. A identificação de problemas e execução de reparos diminui o fator risco, porém esses estudos requerem previsões acuradas para auxiliar à tomada de decisões de soluções adequadas.

Dado o exposto, propõe-se Equation Chapter 1 Section luma metodologia híbrida que integra as seguintes técnicas para a projeção dos deslocamentos relativos (que é um dos tipos de movimentos dos blocos de concreto) em Itaipu: *support vector regression* (SVR), análise *wavelet* e combinação linear de previsões. Diante a sua relevância na barragem de concreto em Itaipu, são previstos os movimentos relativos do Bloco I11, um ano à frente (12 meses).

Quanto às técnicas escolhidas para integrar a metodologia proposta na previsão dos deslocamentos relativos, é importante destacar os seguintes aspectos, que justificam a escolha de cada uma:

- I. A *Análise Wavelet* é uma técnica de tratamento de sinais que permite a decomposição ortogonal de séries de temporais, de maneira a se obter subséries de tempo com padrões de comportamento mais regulares, o que possibilita a obtenção de ganhos preditivos relevantes (TEIXEIRA JR et al., 2015). É possível ainda listar alguns artigos recentes que utilizam *wavelet* no processo de previsão de séries de tempo, obtendo com excelentes resultados: (MAHESWARAN; KHOSA, 2015), (SEO et al., 2015), (ZHU; WANG; FAN, 2014), (ORTEGA; KHASHANAH, 2014), (LIU et al., 2013), (SANG, 2013), (TEIXEIRA JR et al., 2012), (KISI; CIMEN, 2011), entre outros.
- II. Os modelos SVR, propostos inicialmente por Vapnik (1995), consistem em métodos preditivos individuais que exigem não estacionariedade e normalidade dos dados, como é o caso dos modelos lineares de Box e Jenkins (HAMILTON, 1994), somente a existência de estrutura de autodependência (linear ou não linear). Além disso, os previsores SVR vêm alcançando resultados superiores, em problemas de previsão, quando comparados com outras abordagens tradicionais como, por exemplo, as redes neurais artificiais (HAYKIN, 2001). Em (YAO; CROOK; ANDREEVA, 2015), (ALTAMEEM et al., 2015), (PIRI et al., 2015), (CHEN et al., 2015), (LU, 2014), (BAYDAROĞLU; KOÇAK, 2014), (KAO et al., 2013), (HONG, 2011), (HE; WANG; JIANG, 2008) podem ser verificados diferentes aplicações que encorajaram a sua utilização.
- III. A combinação linear de previsões não tendenciosas, em geral, alcança ganhos de acurácia relevantes, no processo preditivo. A previsão linearmente combinada por ser vista com uma agregadora de informações oriundas de diferentes fontes (no caso, de modelos preditivos individuais). Em Granger (1989), por exemplo, pode ser verificado, na modelagem de diversas séries temporais, oriundas de diferentes áreas, que a combinação linear de previsões obteve em ganhos de acurácia preditiva relevantes. Em Wallis (2011), e em Mancuso e Werner (2013), em que são fornecidas revisões históricas sobre previsões linearmente combinadas, são referidos e comentados muitos dos principais artigos que encorajam a sua adoção em processos preditivos. Recentemente, várias outras contribuições atestam e incentivam à utilização de previsões linearmente combinadas, como, por exemplo em: (BABU; REDDY, 2014), (KOSANAN; KANTANANTHA, 2014), (RODRIGUES et al., 2014), (RODRIGUES et al., 2015), (TEIXEIRA; TEIXEIRA, J.; SIQUEIRA, 2015).

De posse destas informações, se abre um caminho de possibilidades de aplicação de técnicas matemáticas e estatísticas através de modelos híbridos, na previsão de séries temporais, sendo o objetivo deste trabalho a aplicação de um modelo híbrido para a previsão dos deslocamentos relativos de um bloco da usina hidrelétrica de Itaipu.

O presente artigo é dividido em seis seções: na Seção 1, é feita uma contextualização do problema em questão e justifica a escolha dos métodos que integram a metodologia proposta; na Seção 2, realiza-se uma sucinta revisão das abordagens SVR e *Análise Wavelet* que integram o método híbrido proposto, o qual é detalhado na Seção 3; na Seção 4, o problema é apresentado e a metodologia é aplicada; na Seção 5, são apresentados os principais resultados numéricos, e a comparação é realizada com os resultados obtidos por (RODRIGUES et al., 2015); e, finalmente, na Seção 6, o artigo é concluído.

2. REFERENCIAL TEÓRICO

2.1. Decomposição *wavelet*

De acordo com Morettin e Tolo (2006), na Análise de Fourier Canônica, uma série temporal com T observações, denotada por $y(\cdot) = (y(1), y(2), \dots, y(T))$, pode ser ortogonalmente decomposta em termos de senos e cossenos definidos em diversas bandas de frequências; enquanto que, de acordo com Teixeira, J. et al., (2015), na Análise *Wavelet*, $y(\cdot)$ pode, também de forma ortogonal, ser decomposta, simultaneamente, a partir de informações de frequência e de tempo, ambas capturadas a partir de um conjunto de funções *wavelet* $\{\omega_{m,n}(\cdot)\}_{(m,n) \in \mathbb{Z} \times \mathbb{Z}}$, onde \mathbb{Z} é o conjunto dos números inteiros, sendo que cada mapa $\omega_{m,n}(\cdot)$ é tal que $\omega_{m,n}(\cdot) = 2^{\frac{m}{2}} \omega(2^m(\cdot) - n)$, o qual é gerado por meio de convoluções de dilatação (associada ao parâmetro m) e de translação (ao parâmetro n) de uma função *wavelet* $\omega(\cdot)$ original (MALLAT, 2009). Ainda de acordo com Mallat (2009), o parâmetro m é chamado de parâmetro de escala diádica e n , de parâmetro de translação unitária.

Em Teixeira, J. et al., (2012), verifica-se, empiricamente, que cada função *wavelet* $\omega_{m,n}(\cdot)$ está associada ao mapeamento das componentes *wavelet* (CWs) de detalhe (que são as CW de alta frequência) da série temporal $y(\cdot)$; e, teórica e empiricamente, que a CW de aproximação (que é a CW de baixa frequência) da série temporal $y(\cdot)$ está associada ao conjunto de funções escala $\{\phi_{m_0,n}(\cdot)\}_{n \in \mathbb{Z}}$ (TEIXEIRA, J. et al., 2015). Cada mapa $\phi_{m_0,n}(\cdot) = 2^{\frac{m_0}{2}} \phi(2^{m_0}(\cdot) - n)$, onde m_0 é um número inteiro fixo, é gerado por meio da convolução de translação unitária n de uma função escala original $\phi(\cdot)$.

De acordo com Mallat (2009), as famílias *wavelet*, em geral, são apresentadas em pares constituídos de uma função *wavelet* $\omega(\cdot)$ e uma função escala $\phi(\cdot)$ (originais). As famílias de *Haar*, de *Daubechies*, de *Coiflet* e de *Symlet* são exemplos que, comumente, são utilizadas na decomposição de sinais temporais, antes de sua efetiva modelagem.

Baseado em Kubrusly e Levan (2006), e Teixeira, J. et al., (2015), tem-se que a decomposição *wavelet* de uma série temporal $y(\cdot)$ é dada, genericamente, pela Equação 1.

$$y(\cdot) = y_{A_{m_0}}(\cdot) + \sum_{m=m_0}^{\infty} y_{D_m}(\cdot) \quad (1)$$

Onde: (i) $y_{A_{m_0}}(\cdot) = \sum_{n \in \mathbb{Z}} a_{m_0,n} \phi_{m_0,n}(\cdot)$ é a CW de aproximação de nível m_0 do estado $y(\cdot)$; $y_{D_m}(\cdot) = \sum_{n \in \mathbb{Z}} d_{m,n} \omega_{m,n}(\cdot)$ é a CW de detalhe de nível m do estado $y(\cdot)$; (ii) $d_{m,n}$ 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 $y(\cdot)$ e $\omega_{m,n}(\cdot)$ (isto é, $d_{m,n} = \sum_{t=1}^T y(t) \omega_{m,n}(t)$) e entre $y(\cdot)$ e $a_{m_0,n} = \sum_{t=1}^T y(t) \phi_{m_0,n}(t)$, para todo $m_0, m, n \in \mathbb{Z}$.

2.2. Support Vector Regression

O método preditivo individual *support vector regression* (SVR) é fundamentado nas *support vector machines* (SVMs), propostas por Vapnik (1995). O modelo produzido pelo SVR depende apenas da própria série temporal e de um objetivo chamado de função custo (o qual é utilizado para a construção do SVR).

Assim, assumamos que $\{(x_1, y_1), \dots, (x_l, y_l)\} \subset X \times \mathbb{R}$ consiste em uma amostra de dados temporais de treinamento, onde X indica o espaço dos padrões de entrada. Aqui, o objetivo é encontrar uma função $f(x)$ que apresente, no máximo, um desvio ε dos alvos y_i ($i=1, \dots, l$) de treinamento. A função linear de aproximação $f(x)$ é definida por $f(x) = \langle w, x \rangle + b$, com $w \in X$, $b \in \mathbb{R}$, onde $\langle \cdot, \cdot \rangle$ denota o produto interno usual sobre o espaço X . Não obstante, nem sempre é possível garantir a viabilidade do problema, uma vez que existem pontos que violam as restrições. Por conta disso, a formulação em (2), proposta por Vapnik (1995), utiliza variáveis de folga ξ_i e ξ_i^* .

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2} \|w\|^2 + C \left(\sum_{i=1}^l \xi_i + \xi_i^* \right) \\ \text{sujeito a:} \quad & \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad \forall i = 1, \dots, l \end{cases} \end{aligned} \quad (2)$$

Para a construção da formulação dual do problema descrito em (2), são introduzidos alguns multiplicadores de Lagrange (α_i, α_i^*) não negativos, conforme destaca Vapnik (1995), de modo que o mesmo é convertido em um problema de maximização, na sua forma dual, conforme é descrito em (3).

$$\begin{aligned} \text{Maximizar} \quad & \left(-\frac{1}{2} \right) \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \langle x_i, x_j \rangle - \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + y_i \sum_{i=1}^l (\alpha_i - \alpha_i^*) \\ \text{Sujeito a} \quad & \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\ & \alpha_i, \alpha_i^* \in [0, C], \quad \forall i = 1, \dots, l \end{aligned} \quad (3)$$

Assim o sendo, a função linear $f(x)$ de aproximação pode ser reescrita como segue:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b$$

A formulação dual do problema do SVR fornece como alternativa fazer modelagens em um espaço de alta dimensionalidade. Adicionalmente, pode-se realizar um mapeamento não linear de dados temporais de entrada para espaços de dimensão maior, nos quais a regressão linear torna-se possível, desde que sejam utilizadas abordagens do SVR baseadas em funções Kernel (não lineares), as quais são denotadas, genericamente, por $K(x, x') = \langle \phi(x), \phi(x') \rangle$ (SMOLA; SCHÖLKOPF, 1998). Consequentemente, o problema descrito em (3) é reeditado conforme a formulação em (4).

$$\begin{aligned} \text{Maximizar} \quad & -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i, x_j) - \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + y_i \sum_{i=1}^l (\alpha_i - \alpha_i^*) \\ \text{Sujeito a} \quad & \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\ & \alpha_i, \alpha_i^* \in [0, C], \quad \forall i = 1, \dots, l \end{aligned} \quad (4)$$

Com efeito, a função de aproximação da SVR consiste em um modelo não-linear e é reeditada tal como segue: $f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x, x_i) + b$. Dentre as funções *Kernel* mais utilizadas na modelagem SVR, destacam-se: *Kernel* linear, *Kernel* polinomial, *Kernel* sigmoidal e o *Kernel* gaussiano, cujas definições podem ser encontradas em Smola e Schölkopf (1998).

2.3. Combinação linear de previsões

A combinação linear de previsões, de acordo com Clemen (1989), consiste em uma metodologia atraente na obtenção de previsões mais acuradas que as oriundas de previsores individuais. Assim o sendo, ao invés de se escolher o melhor método preditivo individual, baseado em algum critério de seleção, o problema agora passa a ser definir quais os previsores individuais a serem utilizados na obtenção das previsões combinadas.

De acordo com Teixeira, J. (2013), as previsões linearmente combinadas necessitam de algum mecanismo ponderador tal que vise à minimização da variância dos resíduos combinados ou outra função residual. Ainda de acordo com o autor, os pesos adaptativos lineares (responsáveis por ponderar as previsões geradas por previsores individuais) podem ser fixos ou variáveis e, além disso, podem não necessariamente ser positivos ou somar uma unidade. Para mais detalhes, veja (FARIA; MUBWANDARIKWA, 2008) e (TEIXEIRA; TEIXEIRA, J.; SIQUEIRA, 2015).

3. METODOLOGIA HÍBRIDA

Seja uma série temporal de cardinalidade T a ser prevista, o método híbrido é realizado de acordo com os três passos seguintes.

Passo 1 – Decomposição *wavelet* da série original: uma decomposição *wavelet* de nível r , (TEIXEIRA, J. et al., 2015), na série temporal $(y(t))_{t=1}^T$ é realizada, gerando-se $r + 1$ CWs – isto é, uma CW de aproximação de nível m_0 , denotada por $(y_{A_{m_0}}(t))_{t=1}^T$, e r CWs de detalhe de níveis $m_0, m_0 + 1, \dots, m_0 + (r - 1)$, denotadas por $(y_{D_{m_0}}(t))_{t=1}^T, (y_{D_{m_0+1}}(t))_{t=1}^T, \dots, (y_{D_{m_0+(r-1)}}(t))_{t=1}^T$ respectivamente.

Na prática, usualmente se adota para o parâmetro m_0 valor igual ao nível de decomposição r . As CWs de aproximação e de detalhe são obtidas a partir do aplicativo computacional *Matlab* 8.0, que oferece várias opções de bases *wavelet*, tais como: as famílias de *Haar*, *Daubechies*, *coiflets* e *symlets*.

Passo 2 – Modelagem das componentes *wavelet* via SVR: as CW obtidas no passo 1 são modeladas através de um SVR.

Para a obtenção dos modelos individuais SVR, foi utilizada uma biblioteca do *software Matlab* 8.0 denominada LIBSVM, versão 3.2, desenvolvida por Chang e Lin (2015), disponível no link: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.

Passo 3 – Combinação linear das previsões das componentes *wavelet* previstas por SVR: Obtidas as previsões da componente *wavelet* de aproximação e as r previsões das componentes de detalhe, na etapa anterior, através do método SVR, faz-se a combinação linear das previsões, com o objetivo de se obter a previsão pontual do método.

A combinação é dada pela Equação(5).

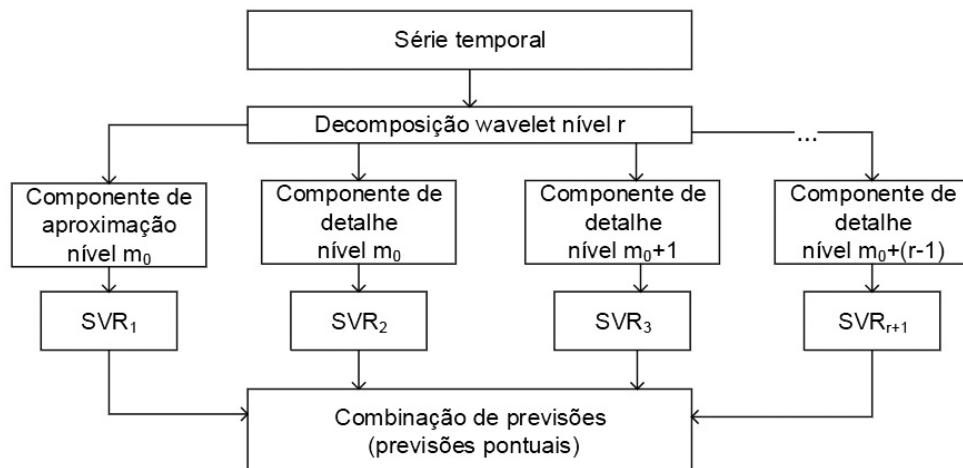
$$(y_H)_{t=T'}^T = \tau_{m_0} \times (\hat{y}_{A_{m_0-SVR_{A_{m_0}}}}(t))_{t=T'}^T + \sum_{m=m_0}^{m_0+r-1} \kappa_m \times (\hat{y}_{D_{m-SVR_{D_m}}}(t))_{t=T'}^T \quad (5)$$

Onde: τ_{m_0} , κ_m com $m = m_0, m_0 + 1, \dots, m_0 + r - 1$ são parâmetros a serem determinados com a otimização do problema de programação não linear apresentada na Equação(6), cuja função objetivo consiste na minimização do erro quadrático médio (MSE).

$$\text{Min RMSE} = \left[\frac{1}{T-T'} \sum_{t=T'}^T (y_t - \hat{y}_t)^2 \right]^{\frac{1}{2}} \quad (6)$$

Na Figura 1 apresenta-se o fluxograma com os passos do método.

Figura 1 – Fluxograma com os passos do método.



Fonte: Proposição dos autores (2015).

4. MATERIAIS E MÉTODOS

A Itaipu é uma empresa que se preocupa com a segurança de suas operações, mantendo um monitoramento do desempenho de sua estrutura, dentre outras formas, através da análise de dados coletados pela instrumentação civil da barragem. Dentre os diversos fatores que se pode considerar em uma barragem de concreto os movimentos relativos entre blocos, e entre estes e suas fundações, surgem como um assunto de extremo interesse e análise.

Os blocos estão sujeitos, principalmente à ação do nível do lago e das variações de temperatura. Por exemplo: no verão ocorrem dilatações no concreto, que provocam uma tendência de deformação do bloco para a direção montante, que por sua vez pode fazer com que aumentem as tensões de compressão no pé de montante dos blocos. No inverno o concreto se contrai, e provoca uma tendência de deformação do bloco à jusante. Pode-se então identificar um comportamento cíclico da estrutura, intimamente condicionado a condições ambientais da região (OSAKO, 2002).

Então fazer a previsão de deslocamentos mais precisamente dos deslocamentos relativos entre blocos é de grande relevância para ajudar a compreender e prever o comportamento futuro da barragem de concreto de Itaipu.

O comportamento dos deslocamentos diferenciais entre blocos se dá segundo três direções ortogonais entre si, recalques diferenciais entre blocos, deslocamentos cisalhantes horizontais e movimentos de abertura-fechamento das juntas. O instrumento denominado base de alongâmetro, é utilizado para medir os deslocamentos horizontais e verticais entre juntas, aberturas de fissuras de barragens, galerias, túneis, maciços rochosos e edificações.

Para verificarem-se as deformações oriundas dos deslocamentos relativos na barragem da usina de Itaipu, são feitas duas leituras em cada base de alongâmetro: uma das medidas corresponde à abertura ou fechamento das juntas, ao passo que a outra medida indica o deslizamento entre os blocos. O deslizamento, por sua vez, pode ser interpretado de duas formas, conforme as bases de alongâmetro estejam instaladas na parede ou no piso das galerias da estrutura. Quando no piso, o deslizamento indica um movimento relativo no sentido montante-jusante. Quando na parede, isto corresponderia a um recalque relativo entre os blocos, mais detalhes podem ser vistos nos trabalhos de Matos (2002), Osako (2002) e Buzzi (2007).

Do instrumento base de alongâmetro, se obtêm as seguintes medidas: duas no piso, abertura entre juntas e deslizamento no sentido montante-jusante, e duas na parede, recalque entre os blocos e abertura entre os blocos.

Para esta aplicação, os dados que formam a série temporal a ser modelada são provenientes do instrumento denominado base de alongâmetro que mede a abertura e o fechamento, o recalque e o deslizamento entre os blocos.

A série escolhida para aplicação do método é bivariada e contém os dados referentes ao período de 1982 – 2013, é são provenientes do instrumento base de alongâmetro JS-I35, localizado na parede entre os blocos I10 e I11, e que mede a abertura (JS-I35-X) e os recalques (JS-I35-Y), tais blocos estão localizados na barragem de ligação esquerda, entre a estrutura de desvio e a barragem de enrocamento. A seta de cor preta indica na Figura 2, a localização dos blocos I10-I11.

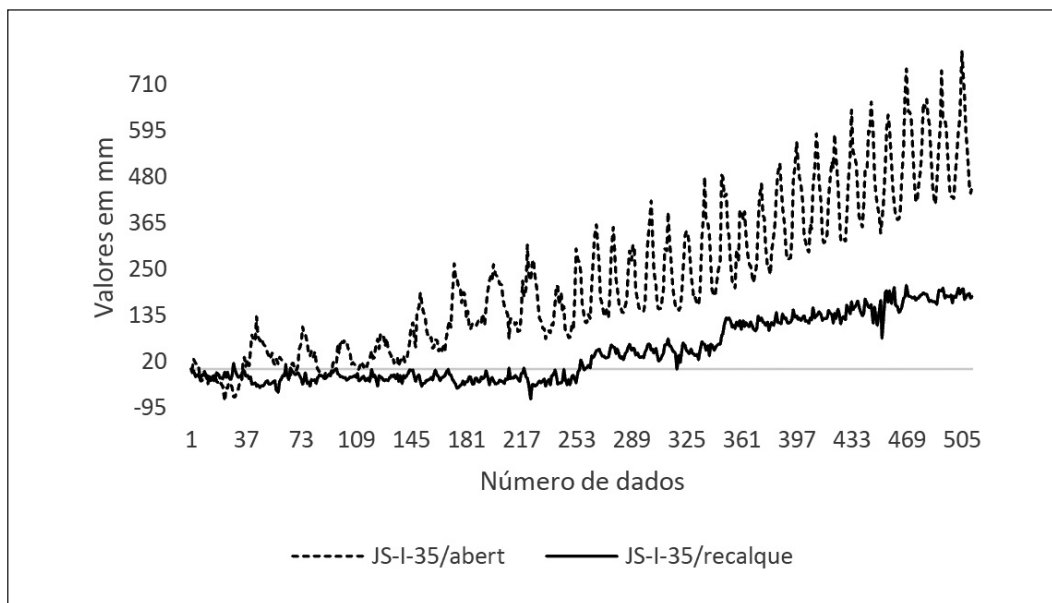
Figura 2 – Localização dos blocos I10/I11.



Fonte: <https://www.itaipu.gov.br/> (2015).

A série dos deslocamentos relativos do bloco I11, nas direções de abertura entre juntas e deslizamento (que são os deslocamentos horizontais), do instrumento base de alongômetro JS-I35 pode ser visto na Figura 3, sendo os valores dados em milésimos de milímetros.

Figura 3 – Deslocamentos relativos do bloco I11.



Fonte: Proposição dos autores (2015).

5. RESULTADOS

Primeiramente foi feito o tratamento das informações, para serem estruturadas de acordo com a metodologia de previsão a ser aplicada, optou-se por trabalhar com o período de 1992-2013, com valores mensais. Para ambas as séries foram deixadas para treino o período de 1992-2012 totalizando 252 valores e os 12 últimos valores que representa o ano de 2013, para teste.

Foi utilizada a metodologia de previsão de 1 passo à frente, num horizonte de 12 valores, e os resultados comparados com os obtidos por Rodrigues et al., (2015).

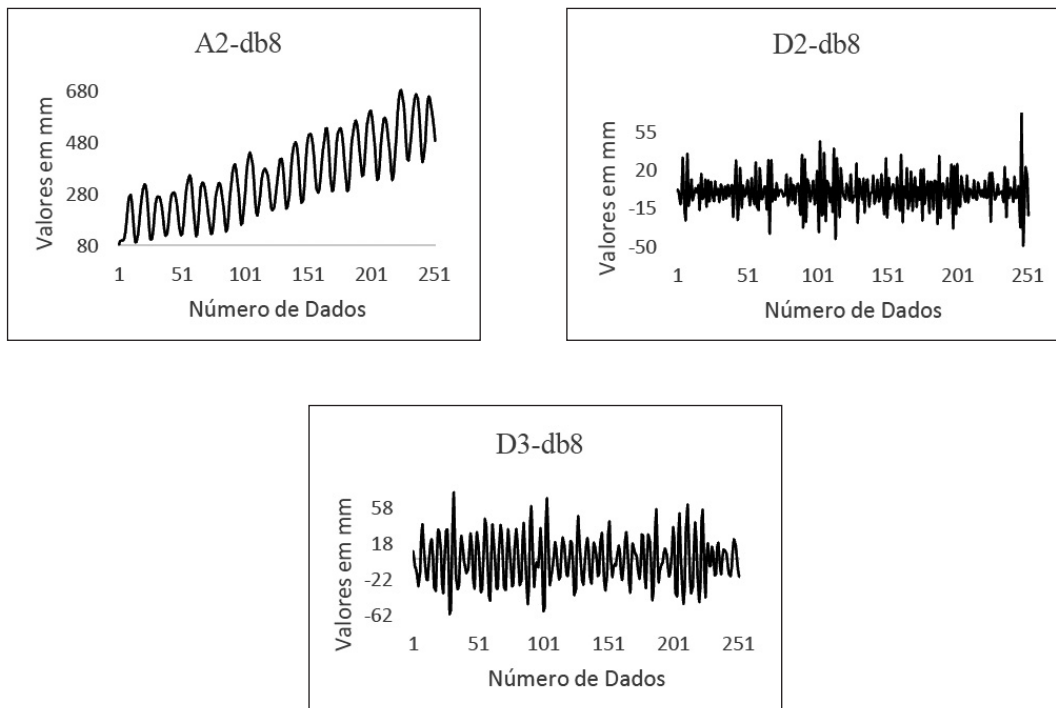
A série JS-I35-X, e a série JS-I35-Y possuem cardinalidade $T = 1, 2, \dots, 252$, e são representados respectivamente por $(y_{I35X}(t))_{t=1}^{252}$ e $(y_{I35Y}(t))_{t=1}^{252}$.

Seguindo as etapas da metodologia híbrida, tem-se:

Etapa 1: Foi realizada a decomposição *wavelet* de nível 2, gerando uma CW de aproximação de nível 2 (cuja notação é A2) e duas CWs de detalhe de níveis 2 e 3 (cujas notações são respectivamente D2 e D3). A base *wavelet* escolhida para a referida decomposição foi a da família de *daubechies* com momento nulo igual a 8 (cuja notação é dada por “db 8”).

Para a série $(y_{I35X}(t))_{t=1}^{252}$ gerou-se as subséries: $(\tilde{y}_{I35X-A_2}(t))_{t=1}^{252}$, $(\tilde{y}_{I35X-D_2}(t))_{t=1}^{252}$ e $(\tilde{y}_{I35X-D_3}(t))_{t=1}^{252}$, onde as 3 componentes *wavelet* geradas podem ser visualizadas na Figura 4.

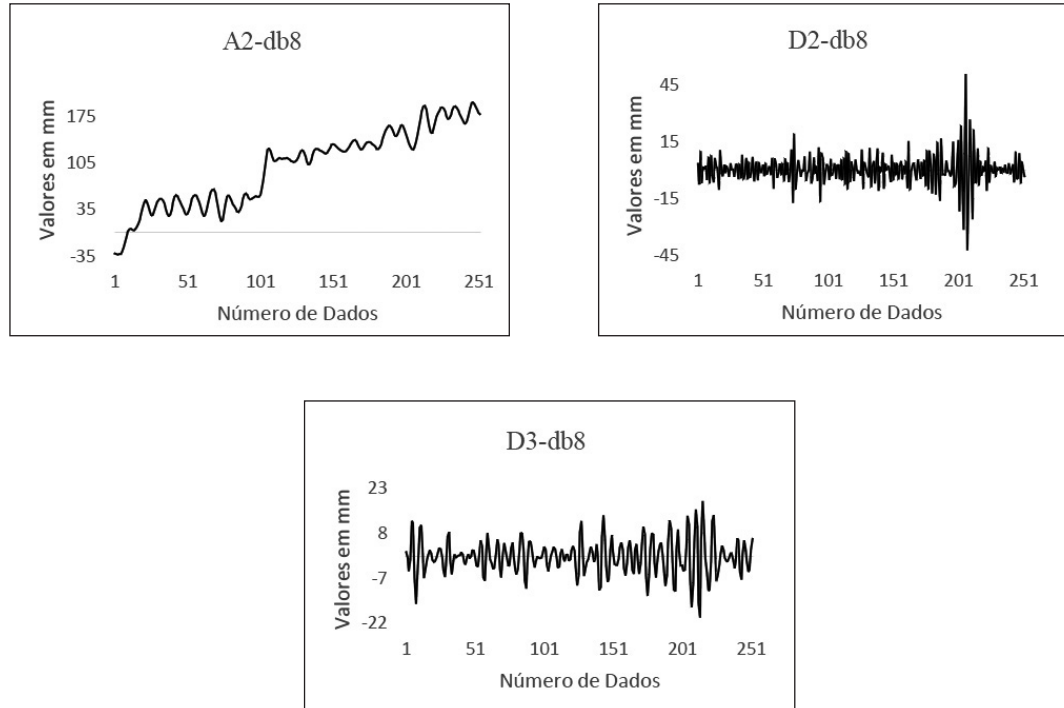
Figura 4 – Decomposição *wavelet* de nível 2, com base ortonormal *daubechies* 8 para série I35X.



Fonte: Proposição dos autores (2015).

Para a série $(y_{I35Y}(t))_{t=1}^{252}$ gerou-se as subséries: $(\tilde{y}_{I35Y-A_2}(t))_{t=1}^{252}$, $(\tilde{y}_{I35Y-D_2}(t))_{t=1}^{252}$, $(\tilde{y}_{I35Y-D_3}(t))_{t=1}^{252}$ e sendo o gráfico de cada subsérie representado na Figura 5.

Figura 5 – Decomposição *wavelet* de nível 2, com base ortonormal *daubechies* 8 para série I35Y.



Fonte: Proposição dos autores (2015).

Etapa 2: Cada uma das componentes *wavelet* obtidas na etapa anterior, foram modeladas individualmente, através de SVRs distintos, gerando as séries: $(\hat{y}_{I35X-A_2,SVR_{A_2}}(t))_{t=9}^{252+12}$, $(\hat{y}_{I35X-D_2,SVR_{D_2}}(t))_{t=10}^{252+12}$, $(\hat{y}_{I35X-D_3,SVR_{D_3}}(t))_{t=15}^{252+12}$ e $(\hat{y}_{I35Y-A_2,SVR_{A_2}}(t))_{t=14}^{252+12}$, $(\hat{y}_{I35Y-D_2,SVR_{D_2}}(t))_{t=9}^{252+12}$, $(\hat{y}_{I35Y-D_3,SVR_{D_3}}(t))_{t=14}^{252+12}$, gerando previsões dentro e fora da amostra, onde $h = 12$ é o horizonte de previsão.

Quanto à modelagem SVR, foi utilizado o *Kernel* polinomial, definido por $(\gamma \langle u, v \rangle + \text{coef0})^{\text{degree}}$ e o *Kernel* RBF, dado por $\exp(-\gamma \|u, v\|^2)$. É conhecido que o desempenho de generalização de um modelo SVR depende de um bom ajuste do parâmetro (C) de regularização e do parâmetro ϵ da função perda, bem como dos parâmetros relacionados ao *Kernel*, escolhido conforme apresentado em Smola e Schölkopf (1998).

A biblioteca utilizada aqui permite que os parâmetros sejam definidos pelo usuário. A fim de melhor avaliar o desempenho da abordagem proposta, para cada CW a ser modelada, os parâmetros C e ϵ , e os parâmetros do *Kernel* foram configurados em uma busca por grade (*grid search*) conforme (HSU; CHANG; LIN, 2010), e selecionado aquele que apresentou um melhor desempenho de validação cruzada 10 vezes 10-fold no conjunto de treino, (VIANA et al., 2009).

Os parâmetros ótimos do SVR, após a aplicação no treinamento podem ser vistos na Tabela 1.

Tabela 1 – Parâmetros SVR.

série temporal	Kernel	gamma	coef0	degree	cost	epsilon	window	cvmse
$\hat{Y}_{I35X-A_2,SVRA_2}$	RBF	0,25	-	-	99	0,01	9	$5,79 \times 10^{-5}$
$\hat{Y}_{I35X-D_2,SVRD_2}$	Polinomial	0,20	3	3	52	0,01	10	$1,31 \times 10^{-3}$
$\hat{Y}_{I35X-D_3,SVRD_3}$	Polinomial	0,12	3	3	200	0,01	15	$3,080 \times 10^{-4}$
$\hat{Y}_{I35Y-A_2,SVRA_2}$	Polinomial	0,70	3	3	10	0,001	14	$6,74 \times 10^{-6}$
$\hat{Y}_{I35Y-D_2,SVRD_2}$	Polinomial	0,10	3	3	120	0,001	9	$8,24 \times 10^{-4}$
$(\hat{Y}_{I35Y-D_3,SVRD_3})$	Polinomial	0,18	3	3	94	0,001	14	$4,65 \times 10^{-4}$

Fonte: Proposição dos autores (2015).

Etapa 3: Nesta etapa foi realizada a combinação linear das componentes *wavelet* previstas na etapa 2, para cada uma das séries. Da combinação das componentes *wavelets* correspondentes à série JS-I35-X resultou nos parâmetros $\tau_2 = 1,000185$, $\kappa_2 = 1,014481$ e $\kappa_3 = 0,989692$, e para a série JS-I35-Y nos parâmetros $\tau'_2 = 0,99974$; $\kappa'_2 = 0,985243$ e $\kappa'_3 = 1,014124$.

Quando comparado com os resultados de outros métodos nas mesmas condições, a metodologia híbrida aplicada alcançou melhores resultados, em relação às medidas de acurácia RMSE (*root mean square error*) e MAE (*mean absolute error*), que são definidas por Hamilton (1994). Note na Tabela 2 que, em relação ao segundo melhor método competidor (a saber, Arima-SVR), o ganho relativo na estatística RMSE foi de 72%; enquanto que foi de 73% na métrica MAE, para a série JS-I35-X e para a série JS-I35-Y de 65% na estatística RMSE e 69% na métrica MAE.

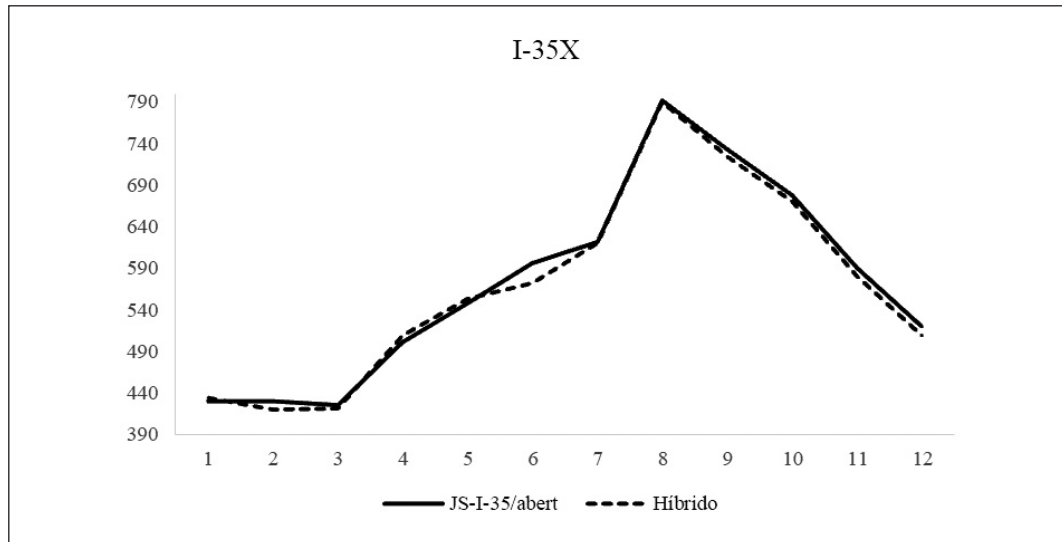
Tabela 2 – Resultados.

Referências	Métodos	JS-I35-X		JS-I35-Y	
		RMSE	MAE	RMSE	MAE
(RODRIGUES et al., 2015)	ARIMA	45,648	33,865	12,344	10,094
	SVR	55,189	37,608	11,490	8,980
	Arima-SVR	35,346	29,515	6,094	5,278
Os autores	Wavelet-SVR	09,822	07,984	02,129	01,607

Fonte: Proposição dos autores (2015).

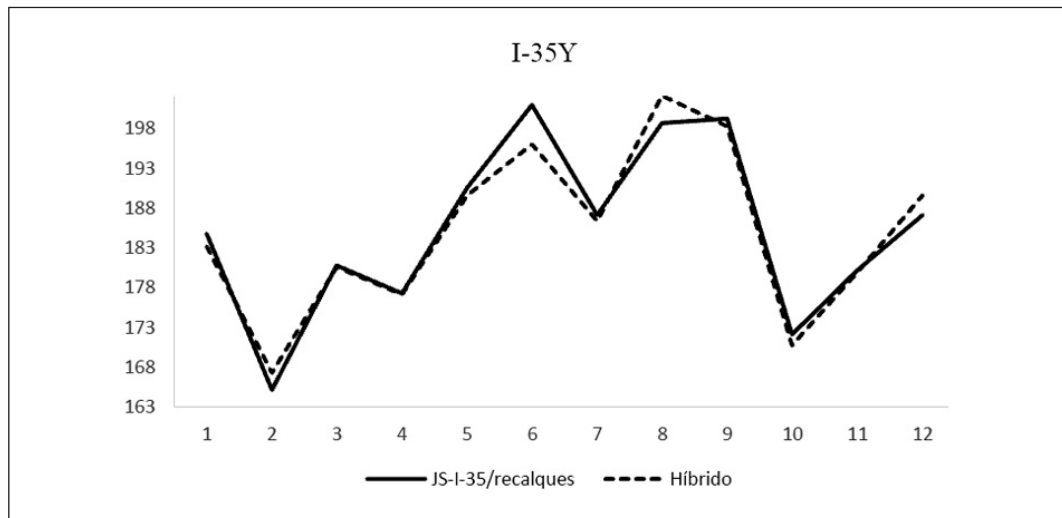
Segue na Figura 6 os valores de previsão na amostra de teste, com as últimas 12 observações, para a série JS-I35-X e na Figura 7 os valores para JS-I35-Y. A linha contínua representa os dados reais da série subjacente, enquanto que a linha pontilhada, as previsões produzidas pelo método aplicado. É possível verificar que as previsões e as observações exibem alta correlação, evidenciado pela proximidade de seus pontos no plano.

Figura 6 – Observações e previsões, na amostra de teste série I35X.



Fonte: Proposição dos autores (2015).

Figura 7 – Observações e previsões, na amostra de teste série I35Y.



Fonte: Proposição dos autores (2015).

6. CONCLUSÃO

Neste artigo, é proposta uma metodologia híbrida que integra os seguintes métodos numéricos: a Análise *Wavelet*, o predictor SVR e a combinação linear de previsões. Os dados que formam a série temporal a ser projetada são provenientes do instrumento denominado base de alongâmetro, que mede os deslocamentos relativos (ou seja, as aberturas (JS-I35-X) e os recalques (JS-I35-Y) entre os blocos 10 e 11 (referido como “Bloco I11”) da barragem da Usina Hidrelétrica de Itaipu.

Na Tabela 2 e na Figuras 6 e 7, pode-se verificar que as previsões produzidas pela metodologia proposta alcançaram alto nível de acurácia, em um horizonte de previsão de um ano. Ou seja, o comportamento futuro do Bloco I11 doze passos à frente, do ponto de vista dos desvios relativos, foi projetado com um nível de precisão elevado, o que permitiria à equipe técnica de segurança de barragens da Itaipu antever a “saúde” das condições estruturais da barragem, bem como identificar previamente os pontos em sua superfície que devem ser reparados e quando fazê-los. Além disso, as referidas previsões de deslocamento podem ser utilizadas em outras abordagens também utilizadas no monitoramento dos movimentos dos blocos de concreto e de sua fundação, como, por exemplo: análise de confiabilidade (BAIN; ENGLEHARDT, 1991), árvores de falha (MOBLEY, 1999), cartas de controle de processos (SMITH, 2013). Como consequência, fornecer uma melhoria em suas análises.

É importante salientar ainda que, dado que a método híbrido proposto é aplicável a qualquer séries temporal que exiba estrutura de autodependência – não exigindo normalidade ou estacionariedade dos dados subjacentes –, segue que pode ser utilizada na previsão dos dados futuros de outros instrumentos que também monitoram a barragem, como, por exemplo: pêndulo (invertidos e diretos), medidores de vazão, extensômetros, piezômetros (BUZZI, 2007) e (MATOS, 2002).

Em última análise, convém destacar que, embora os métodos numéricos que a integram possuem fundamentação teórica relativamente complexa, a sua operacionalização, em casos práticos, é relativamente simples, quanto utilizados os programas e pacotes mencionados no texto.

7. AGRADECIMENTOS

Às instituições: Itaipu Binacional, Programa de Pós Graduação em Métodos Numéricos em Engenharia (PPGMNE), Centro de Estudos Avançados em Segurança de Barragens (CEASB), Fundação Parque Tecnológico Itaipu (FPTI), e aos revisores pelas contribuições que possibilitaram a melhoria deste artigo.

REFERÊNCIAS

- ALTAMEEM, T. A.; NIKOLIC, V.; SHAMSHIRBAND, S.; PETKOVIC, D.; JAVIDNIA, H.; KIAH, M. L. M.; GANI, A. Potential of support vector regression for optimization of lens system. **Computer-Aided Design**, v. 62, p. 57–63, 2015.
- BABU, C. N.; REDDY, B. E. A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data. **Applied Soft Computing**, v. 23, p. 27–38, 2014.
- BAIN, L.; ENGLEHARDT, M. Statistical Analysis of Reliability and Life-Testing Models: Theory and Methods. **CRC Press**, 1991.

BAYDAROĞLU, Ö.; KOÇAK, K. SVR-based prediction of evaporation combined with chaotic approach. **Journal of Hydrology**, v. 508, p. 356–363, 2014.

BUZZI, M. F. **Avaliação das correlações de séries temporais de leituras de instrumentos de monitoração geotécnico-estrutural e variáveis ambientais em barragens estudo de caso de Itaipu**. 2007. Dissertação (Mestrado em Ciências) - Programa de pós-graduação em métodos numéricos em engenharia, Universidade Federal do Paraná, Curitiba, 2007. Disponível em: <https://www.researchgate.net/publication/26979944_Avaliacao_das_correlacoes_de_series_temporais_de_leituras_de_instrumentos_de_monitoracao_geotecnico-estrutural_e_variaveis_ambientais_em_barragens_estudo_de_caso_de_Itaipu>. Acesso em 2 jan.2015.

CHANG, C - C.; LIN, C - J. A Library for Support Vector Machines. **Transactions on Intelligent Systems and Technology**, v. 2, n. 3, p.27, 2011.

CHEN, R.; LIANG, C. - Y.; HONG, W. - C.; GU, D. - X. Forecasting holiday daily tourist flow based on seasonal support vector regression with adaptive genetic algorithm. **Applied Soft Computing**, v. 26, p. 435–443, 2015.

CLEMEN, R. T. Combining forecasts: A review and annotated bibliography. **International Journal of Forecasting**, v. 5, n. 4, p. 559-583, 1989.

EHLERS, R. S. **Análise de séries temporais**. Laboratório de Estatística e Geoinformação. Universidade Federal do Paraná, 2007.

FARIA, A. E.; MUBWANDARIKWA, E. The geometric combination of bayesian forecasting models. **Journal of Forecasting**, v. 27, n. 1, p. 519–535, 2008.

FIRMINO, P. R. A.; DE MATTOS NETO, P. S. G.; FERREIRA, T. A. E. Error modeling approach to improve time series forecasters. **Neurocomputing**, v. 153, p. 242–254, 2015.

GRANGER, C. W. J. Combining forecasts – twenty years later. **Journal of Forecasting**, v. 8, n. 3, p. 167– 173, 1989.

HAMILTON, J. D. **Time Series Analysis**. Princeton: Princeton University Press, 1994.

HAYKIN, S. S. **Redes Neurais**. Porto Alegre: Bookman, 2001.

HE, W.; WANG, Z.; JIANG, H. Model optimizing and feature selecting for support vector regression in time series forecasting. **Neurocomputing**, v. 72, n. 13, p. 600–611, 2008.

HONG, W.-C. Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm. **Energy**, v. 36, n. 9, p. 5568–5578, 2011.

HSU, C. - W.; CHANG, C. - C.; LIN, C. - J. **A Practical Guide to Support Vector Classification**. National Taiwan University. Taiwan, 2010. Disponível em: <<http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>>. Acesso em 02 jan. 2015.

KAO, L. -J.; CHIU, C. C.; LU, C. J.; YANG, J. L. Integration of nonlinear independent component analysis and support vector regression for stock price forecasting. **Neurocomputing**, v. 99, p. 534–542, 2013.

- KISI, O.; CIMEN, M. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. **Journal of Hydrology**, v. 399, n. 1-2, p. 132–140, 2011.
- KOSANAN, O.; KANTANANTHA, N. Thailand's Para Rubber Production Forecasting Comparison Proceedings of the International MultiConference of Engineers and Computer Scientists 2014. **Anais... IAENG**, Hong Kong, 2014.
- KUBRUSLY, C. S.; LEVAN, N. Abstract wavelets generated by Hilbert space shift operators. **Advances in mathematical and applications**, v. 16, p. 643–660, 2006.
- LIU, H.; TIAN, H.; PAN, D.; LI, Y. Forecasting models for wind speed using wavelet, wavelet packet, time series and Artificial Neural Networks. **Applied Energy**, v. 107, p. 191–208, 2013.
- LU, C.-J. Sales forecasting of computer products based on variable selection scheme and support vector regression. **Neurocomputing**, v. 128, p. 491–499, 2014.
- MAHESWARAN, R.; KHOSA, R. Wavelet Volterra Coupled Models for forecasting of nonlinear and non-stationary time series. **Neurocomputing**, v. 149, p. 1074–1084, 2015.
- MALLAT, S. A wavelet tour of signal processing: the sparse way. **Academic Press**, 2008..
- MANCUSO, A. C. B.; WERNER, L. Review of Combining Forecasts Approaches. **Independent Journal of Management & Production**, v. 4, n. 6, p. 248–277, 2013.
- MATOS, SI. F. **Avaliação de instrumentos para auscultação de barragem de concreto**. estudo de caso: deformímetros e tensômetros para concreto na barragem de Itaipu. Dissertação (Mestrado) - Universidade Federal do Paraná, 2002.
- MOBLEY, R. K. **Root Cause Failure Analysis**. Butterworth-Heinemann, 1999.
- MORETTIN, P. A.; TOLOI, C. M. DE C. **Análise de séries temporais**. Edgard Blucher, 2006.
- ORTEGA, L.; KHASHANAH, K. A neuro-wavelet model for the short-term forecasting of high-frequency time series of stock returns. **Journal of Forecasting**, v. 33, n. 6, p. 134–146, 2014.
- OSAKO, C. I. **A manutenção dos drenos nas Fundações de Barragens o caso da Usina Hidrelétrica de Itaipu**. Dissertação (Mestrado) - Pós-Graduação em Construção Civil, Universidade Federal do Paraná, Curitiba, 2002.
- PIRI, J.; SHAHABODDIN, S.; PETKOVIC, D.; TONG WEN, C.; REHMAN UR, M. H. Prediction of the solar radiation on the Earth using support vector regression technique. **Infrared Physics & Technology**, v. 68, p. 179–185, 2015.
- RODRIGUES, S. B.; TEIXEIRA, L. L.; SIQUEIRA, P. H.; SILVA, A. C. L.; TEIXEIRA JR., L. A. Comparativo Entre Arima, Redes Neurais Artificiais, Máquinas De Vetores Suporte Para Regressão E Combinação Linear De Previsões De Vazões Mensais No Posto 266- Itaipu. *In: SIMPÓSIO DE ENGENHARIA DE PRODUÇÃO*, 2014, Bauru, São Paulo. **Anais...Bauru: XXI SIMPEP**, 2014.
- RODRIGUES, S. B.; SILVA, A. C. L.; TEIXEIRA JR., L. A.; TEIXEIRA, L. L.; SIQUEIRA, P. H.; FRANCO, E. M. C. Forecasting of the relative displacements in block I11 of the Itaipu hydroelectric plant dam through a ARIMA-SVR hybrid approach. **Espaço Energia**, v. 22, n. 22, p. 28–37, 2015.

- SANG, Y.-F. A review on the applications of wavelet transform in hydrology time series analysis. **Atmospheric Research**, v. 122, p. 8–15, 2013.
- SEO, Y.; KIM, S.; KISI, O.; SINGH, V. P. Daily water level forecasting using wavelet decomposition and artificial intelligence techniques. **Journal of Hydrology**, v. 520, p. 224–243, 2015.
- SMITH, E. S. **Control Charts: An Introduction to Statistical Quality Control**. Literary Licensing, LLC, 2013.
- SMOLA, A. J.; SCHÖLKOPF, B. A tutorial on support vector regression. **Statistics and Computing**, v. 14, n. 3, p. 199–222, 1998.
- TEIXEIRA JR., L. A.; PESSANHA, J. F. M.; MENEZES, M. L.; CASSIANO, K. M.; SOUZA, R. C. Redes neurais artificiais e decomposição wavelet na previsão da radiação solar direta. *In*: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 2012, Rio de Janeiro. **Anais... CLAIO/ SBPO**, 2012.
- TEIXEIRA JR, L. A. **Combinação SSA-Wavelet de Métodos Preditivos com Ajuste Numérico MINIMAX, na Geração de Previsões e de Cenários**. Tese (Doutorado) - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, 2013.
- TEIXEIRA, J. L. A.; SOUZA, R. M.; MENEZES, M. L.; CASSIANO, K. M.; PESSANHA, J. F. M.; SOUZA, R.C. Artificial Neural Network and Wavelet decomposition in the Forecast of Global Horizontal Solar Radiation. **Pesquisa Operacional**, v. 35, n. 1, p. 73-90, 2015.
- TEIXEIRA, L. L.; TEIXEIRA JUNIOR, L. A.; SIQUEIRA, P. H. Forecasts of monthly flows in the post-266 Itaipu using neural networks with wavelet shrinkage. **Espaço Energia**, v. 22, n. 22, p. 8–17, 2015.
- VAPNIK, V. N. The nature of statistical learning theory. **Springer Science & Business Media**, 2013.
- VIANA, F. A C.; HAFTKA, R. T.; STEFFEN, V. Multiple surrogates: How cross-validation errors can help us to obtain the best predictor. **Structural and Multidisciplinary Optimization**, v. 39, n. 4, p. 439–457, 2009.
- WALLIS, K. F. Combining forecasts : forty years later. **Applied Financial Economics**, v. 21, n. 1-2, p. 33–41, 2011.
- YAO, X.; CROOK, J.; ANDREEVA, G. Support vector regression for loss given default modelling. **European Journal of Operational Research**, v. 240, n. 2, p. 528-538, 2015.
- ZHANG, G. P. Time series forecasting using a hybrid ARIMA and neural network model. **Neuro-computing**, v. 50, p. 159-175, 2003.
- ZHU, L.; WANG, Y.; FAN, Q. MODWT-ARMA model for time series prediction. **Applied Mathematical Modelling**, v. 38, n. 5-6, p. 1859-1865, 2014.

