

CADERNOS DO IME – Série Estatística

Universidade do Estado do Rio de Janeiro - UERJ
ISSN impresso 1413-9022 / ISSN on-line 2317-4536 - v.38, p.01 - 19, 2015
DOI: 10.12957/cadest.2015.18510

MELHORIA NA PREVISÃO DA SÉRIE TEMPORAL DE INSTRUMENTOS DE MONITORAMENTO DE BARRAGEM VIA COMBINAÇÃO DE MÉTODOS

Emerson Lazzarotto

Universidade Estadual do Oeste do Paraná – UNIOESTE

emerson.lazzarotto@gmail.com

Liliana Madalena Gramani

Universidade Federal do Paraná – UFPR

l.gramani@gmail.com

Anselmo Chaves Neto

Universidade Federal do Paraná – UFPR

anselmo@ufpr.br

Luiz Albino Teixeira Junior

Universidade Federal de Integração Latino Americana – UNILA

luiz.a.t.junior@gmail.com

Edgar Manuel Carreno Franco

Universidade Estadual do Oeste do Paraná - UNIOESTE

emf1@gmail.com

Resumo

Barragens de usinas hidrelétricas são avaliadas por inspeções visuais e por instrumentos de monitoramento. As leituras periódicas de um instrumento podem ser interpretadas como uma série temporal estocástica cujos valores passados fornecem informações relevantes para construção de previsões acerca de seus valores futuros, além de indicar tendências do comportamento futuro da barragem como um todo. Desta forma, é fundamental produzir previsões mais acuradas possíveis, de modo que sirvam de alerta confiável para predição de eventuais anormalidades no comportamento da barragem, permitindo a realização precoce de obras e ações de intervenção. Assim, este artigo propõe uma metodologia na qual é feita uma avaliação do desempenho por meio da previsão de séries temporais das leituras de um instrumento da barragem da usina hidrelétrica de Itaipu usando uma combinação híbrida dos previsores ARIMA-GARCH e redes neurais artificiais, com a utilização da decomposição wavelet. Os resultados mostram que a combinação híbrida proposta alcançou desempenho, em termos de acurácia, bastante superior quando comparado com o uso individual dos métodos preditivos tradicionais e de suas combinações.

Palavras-chave: Segurança de barragem; Séries Temporais Estocásticas; Modelos ARIMA-GARCH; Redes Neurais Artificiais; Decomposição Wavelet.

1. Introdução

As leituras de instrumentos de monitoramento são séries temporais que podem indicar tendências no comportamento da barragem e, com isso, sugerir ações preventivas ou corretivas da equipe de segurança da barragem, minimizando perdas de ordem financeira e operacional. É necessário, portanto, produzir previsões destas séries temporais.

Dentre as principais anomalias que a instrumentação pode detectar estão: *piping* (infiltração) através dos maciços de terra e suas fundações, eventuais trincas com infiltrações no concreto, movimentos bruscos entre os blocos da barragem, subpressões excessivas na cabeça dos blocos a montante, ou a jusante da Casa de Força, movimentos excessivos na região da calha do vertedouro (ITAIPU BINACIONAL, 2002).

Em grandes barragens, há instrumentos instalados nas estruturas de concreto e nas fundações da barragem. Em especial, o pêndulo direto é um instrumento instalado na estrutura de concreto e é responsável por medir o deslocamento da crista da barragem na direção do fluxo e normal a ele em relação a sua base.

As subpressões e infiltrações nas barragens de concreto não podem ser previstas com acurácia por modelos numéricos, como métodos de elementos finitos (MEF) que se baseiam nas leis físicas que governam o fenômeno envolvido. Por isso, a disponibilidade de dados de monitoramento tem difundido a aplicação de modelos estatísticos para a previsão do comportamento da barragem e estes têm servido como um complemento às inspeções visuais e aos modelos numéricos no suporte a tomada de decisões (SALAZAR, MORÁN, *et al.*, 2015). Portanto, como, em geral, as leituras de instrumentos de monitoramento de barragens são avaliadas com valores definidos em projeto, como MEF, o seu domínio de variação pode não ser mais realístico durante seu período operacional e são necessárias outras formas de monitoramento.

Os métodos preditivos com abordagem estocástica podem utilizar os próprios dados observados de uma série temporal a fim de se estabelecer estimativas para os seus parâmetros e gerar previsões. Os modelos autorregressivos integrados de médias móveis (ARIMA) são previsores lineares consagrados na literatura de análise e previsão de séries temporais, enquanto que os modelos autorregressivos de heterocedasticidade condicional generalizada (GARCH's) são utilizados para o mapeamento de estruturas de volatilidade (GUJARATI e PORTER, 2010; BOX, JENKINS e REINSEL, 2008).

Por sua vez, as redes neurais artificiais (RNA's) têm atraído nas últimas décadas crescente atenção para previsão de séries temporais e têm conseguido aplicações de sucesso em problemas não lineares de previsão. As RNA's possuem a capacidade de capturar estruturas de autodependência não lineares (ADHIKARI e AGRAWAL, 2013).

O uso de modelos híbridos ou a combinação de diversos modelos tem se tornado prática comum para melhorar a exatidão de previsões (KHASHEY e BIJARI, 2011). Metodologias híbridas que combinam *ARIMA* e redes neurais artificiais para a previsão de séries temporais clássicas da literatura foram propostas por Zhang (2003) e Khashey e Bijari (2011). Seus resultados mostraram uma melhoria no desempenho da previsão quando comparado aos dois modelos usados separadamente. Diversos outros autores têm empregado com sucesso métodos híbridos de previsão de séries temporais. Aplicações na previsão de preços no mercado de eletricidade espanhol foram feitas por Tan, Zhang, *et al.*(2010), Liu e Shi (2013) e Amjady e Hemmati (2009); previsão da quantidade de ozônio na superfície em regiões urbanas por Kumar e De Ridder (2010); qualidade do ar (poluentes), por Díaz-Robles, Ortega, *et al.*(2008); preços de metais, por Kriechbaumer, Angus, *et al.*(2014); deformação de uma barragem, por Cao, Qiao e Ren(2009) e velocidade dos ventos, por Liu, Tian, *et al.*(2013).

As funções *wavelets* são usadas para decompor uma série temporal em termos de uma componente wavelet (CW) de aproximação A e de detalhes D_j . A CW de aproximação é a principal e pode ser considerada uma versão 'alisada' da série original. Por outro lado, as CW's de detalhes representam flutuações estocásticas de baixa variabilidade e alta frequência ao redor da CW de aproximação. A série original é reconstruída através da soma de todas as suas CW's geradas a partir dela mesma (KRIECHBAUMER, ANGUS, *et al.*, 2014). Teixeira Junior (2013), Teixeira Junior, Souza, *et al.*(2015), mostram aplicações nas quais ganhos preditivos relevantes foram alcançados com o uso de decomposições *wavelet*.

Neste trabalho é realizado um estudo comparativo entre diversas abordagens para previsão da série temporal de dados da leitura de um pêndulo direto, que mede o deslocamento da crista da barragem na direção do fluxo do rio, da barragem da usina hidrelétrica de Itaipu, localizada no Rio Paraná, entre o Brasil e o Paraguai. O objetivo é comparar o desempenho das previsões realizadas em cada abordagem e mostrar que o uso da decomposição *wavelet* e da combinação de métodos preditivos como *ARIMA*-

GARCH e redes neurais artificiais pode acarretar melhorias relevantes no desempenho da previsão do comportamento de barragens e auxiliar, conseqüentemente, no fornecimento de alertas prévios de eventuais anormalidades.

Este artigo está estruturado, como segue: na Seção 2, são descritos os principais conceitos teóricos utilizados; na Seção 3, os dados e os métodos empregados; na Seção 4 são detalhados os resultados obtidos; na Seção 5, é feita uma análise dos resultados; e na Seção 6, são expostas algumas conclusões do trabalho.

2. Referencial Teórico

2.1. Modelos ARIMA e ARIMA-GARCH

Uma série temporal (ST) é a realização de um *processo estocástico* que é descrito como um fenômeno estatístico que se desenvolve no tempo de acordo com leis probabilísticas, caracterizado por um modelo que busca capturar a estrutura de dependência existente no passado da série para realizar previsões.

A abordagem de Box e Jenkins é uma metodologia bastante utilizada na análise de modelos paramétricos e consiste do ajuste de modelos autorregressivos integrados de médias móveis ($ARIMA(p,d,q)$) aos dados (BOX, JENKINS e REINSEL, 2008). Sua construção se baseia num ciclo iterativo, cujas etapas são: especificar uma classe geral de modelos a serem considerados, identificar um modelo adequado baseado na análise da autocorrelação, estimar os parâmetros e diagnosticar a adequação dos resíduos para os fins desejados. Caso o diagnóstico dos resíduos seja inadequado, deve-se retomar a etapa da identificação (CHATFIELD, 1996), (MORETTIN e TOLOI, 2006).

Os modelos $ARMA(p,q)$ possuem genericamente a seguinte formulação:

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

onde μ é a média, ϕ_i os coeficientes da parte autoregressiva, θ_i os coeficientes da parte de médias móveis, y_{t-i} são as observações passadas da série e a_t é termo estocástico da série. Processos $AR(p)$, $MA(q)$ e $ARMA(p,q)$ apresentam características especiais em sua função de autocovariância e de autocovariância parcial que podem ser úteis no estágio da identificação de modelos adequados (MORETTIN e TOLOI, 2006).

Dentre os testes de diagnóstico dos resíduos mais utilizados na prática estão o teste conjunto de autocorrelações (Ljung-Box), que tem o objetivo de testar se as

primeiras m autocorrelações da série temporal (ST) são nulas, ou seja, se a ST é não correlacionada. A hipótese do teste é $H_0: \rho_1 = \rho_2 = \dots = \rho_m = 0$ e H_a : pelo menos um dos $\rho_i \neq 0$ e a estatística do teste é

$$LB = T(T+2) \sum_{k=1}^m \frac{\rho_k^2}{T-k} \sim \chi^2_{(v)}, \quad (2)$$

em que os graus de liberdade $v = m - (p + q + 1)$, e p, q são os termos do modelo $ARMA(p, q)$ e o teste de Durbin-Watson que testa a hipótese nula de inexistência de autocorrelação residual de primeira ordem, definido por

$$DW = \frac{\sum (e_t - e_{t-1})^2}{\sum e_t^2}. \quad (3)$$

Se não existe autocorrelação nos resíduos e_t de ordem 1, $\rho_1 = 0$ e $DW \approx 2$.

Quando se tem séries com a variância (condicional) evoluindo no tempo, modelos lineares como *ARIMA* não são adequados para descrever tal tipo de comportamento (MORETTIN e TOLOI, 2006). Os modelos autorregressivos com heteroscedasticidade condicional (*ARCH*), introduzidos por Engle (1982), são não lineares no que se refere à variância. Já no modelo *GARCH* (r, s) (*ARCH* generalizado), proposto por Bollerslev (1986), o choque aleatório a_t , na Equação (1), é representado por $a_t = \sigma_t \epsilon_t$, em que a_t é autodependente dos resíduos e variâncias anteriores, de modo que a variância é dada por

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^r \alpha_i a_{t-i}^2 + \sum_{i=1}^s \beta_i \sigma_{t-i}^2. \quad (4)$$

2.2. Decomposição Wavelet de uma série temporal

Uma *wavelet* é uma função capaz de decompor e representar uma série temporal de forma a permitir a análise desta série em diferentes escalas de frequência e de tempo. Por isso, as funções *wavelet* são ferramentas poderosas de processamento de sinais, muito aplicadas na eliminação de ruído, separação de componentes do sinal, identificação de singularidades e previsão de séries temporais.

Uma *wavelet* é uma função com um número determinado de oscilações que duram certo período de tempo e tendem a zero no resto do período. Uma ‘família’ *wavelet* tem uma *wavelet* ‘mãe’ ψ que produz *wavelets* ‘filhas’ por meio de expansões (dilatações) e translações (KRIECHBAUMER, ANGUS, *et al.*, 2014).

Dado o par ordenado $(l^2, \langle \cdot, \cdot \rangle)$, onde l^2 é o conjunto das sequências infinitas quadrados somáveis de números complexos, isto é $l^2 := \{f: \mathbb{Z} \rightarrow \mathbb{C} / \sum_{t \in \mathbb{Z}} |f(t)|^2 < \infty\}$ e a função $\langle \cdot, \cdot \rangle: l^2 \rightarrow \mathbb{C}$ que é um produto interno. Um elemento $\omega(\cdot)$ em l^2 é um vetor l^2 –wavelet com produto interno $\langle \cdot, \cdot \rangle: l^2 \rightarrow \mathbb{C}$, se e somente se a sequência duplamente indexada $\{\omega_{m,n}(\cdot)\}_{(m,n) \in \mathbb{Z} \times \mathbb{Z}}$ consiste de uma base ortonormal para l^2 , onde o parâmetro m é chamado de parâmetro de escala e n , é chamado de parâmetro de translação. Por sua vez, um elemento $\phi(\cdot) \in l^2$ é um vetor de escala para o produto interno $\langle \cdot, \cdot \rangle: l^2 \rightarrow \mathbb{C}$, se, e somente se, o conjunto $\{\phi_{m,n}(\cdot)\}_{(m,n) \in \mathbb{Z} \times \mathbb{Z}}$, para todo $m, n \in \mathbb{Z}$, dado por $\phi_{m,n}(\cdot) = 2^{m/2} \phi(2^m(\cdot) - n)$ é tal que $\langle \phi_{l,i}(\cdot); \phi_{j,k}(\cdot) \rangle = 0$, para $l = j$ e $i \neq k$ e $\langle \phi_{l,i}(\cdot); \phi_{j,k}(\cdot) \rangle \neq 0$, caso contrário.

Segundo Kubrusly (2011), Mallat (1998) e Kubrusly e Levan (2006), dada uma função $f(\cdot) \in l^2$, ela pode ser decomposta ortogonalmente unicamente em termos de uma base wavelet ortonormal, representada por $\{\phi_{m_0,n}(\cdot)\}_{n \in \mathbb{Z}} \cup \{\omega_{(m,n)}(\cdot)\}_{(m,n) \in \{m\}_{m=m_0}^{+\infty} \times \mathbb{Z}}$, como em

$$f(\cdot) = f_{V_{m_0}(\phi)}(\cdot) + \sum_{m=m_0}^{+\infty} f_{W_m(\omega)}(\cdot). \quad (5)$$

Onde: $f_{V_{m_0}(\phi)}(\cdot) := \sum_{n \in \mathbb{Z}} a_{m_0,n} \phi_{m_0,n}(\cdot)$ é a componente de aproximação de nível m_0 , $a_{m_0,n} := \langle f(\cdot), \phi_{(m_0,n)}(\cdot) \rangle$ (isto é, o produto interno usual entre o vetor sinal $f(\cdot)$ e o vetor escala de nível m_0 e n); e $f_{W_m(\omega)}(\cdot) := \sum_{n \in \mathbb{Z}} d_{m,n} \omega_{m,n}(\cdot)$ é a componente de detalhe de nível m , $d_{m,n} := \langle f(\cdot), \omega_{(m,n)}(\cdot) \rangle$ (isto é, o produto interno usual entre $f(\cdot)$ e o vetor wavelet de nível m e n). Os coeficientes $a_{m_0,n}$ e $d_{m,n}$ são, respectivamente, os coeficientes wavelet de aproximação e de detalhe. Em (5), existe uma decomposição wavelet, em termos de $\{\phi_{m_0,n}(\cdot)\}_{n \in \mathbb{Z}} \cup \{\omega_{(m,n)}(\cdot)\}_{(m,n) \in \{m\}_{m=m_0}^{+\infty} \times \mathbb{Z}}$, de $f(\cdot)$ sobre l^2 .

Seja $\chi_{[0,T]}$ a função tal que $\chi_{[0,T]} := 1$, se $t \in \{1, \dots, T\}$ e $\chi_{[0,T]} := 0$ se $t \in \mathbb{Z} - \{1, \dots, T\}$ e $y(\cdot) := (y_t)_{t=1}^T$ é uma série temporal de tamanho T , onde $T > 1$. Assim, a

composição $\chi_{[0,T]} \circ y(\cdot) = \tilde{y}(\cdot): \mathbb{Z} \rightarrow \mathbb{R}$ pode ser vista como uma sequencia $\tilde{y}(\cdot) := (\dots, 0, 0, y_1, y_2, \dots, y_T, 0, 0, \dots)$ em l^2 . Assumindo que $\{\phi_{m_0,n}(\cdot)\}_{n \in \mathbb{Z}} \cup \{\omega_{m,n}(\cdot)\}_{(m,n) \in \{m\}_{m=m_0}^{+\infty} \times \mathbb{Z}}$ é uma base ortonormal wavelet para l^2 , segue que a série temporal $y(\cdot)$ admite ser decomposta, aproximadamente como em (6).

$$\tilde{y}(\cdot) \simeq \tilde{\tilde{y}} = \sum_{n=1}^{n_{m_0}} a_{m_0,n} \phi_{m_0,n}(\cdot) + \sum_{m=m_0}^{m_0+(p-1)} \sum_{n=1}^{n_m} d_{m,n} \omega_{m,n}(\cdot) \quad (6)$$

A expansão em (6) é chamada de decomposição wavelet de nível p . O valor adotado para o parâmetro de nível m_0 é geralmente o mesmo de p , isto é, $m_0 = p$. Desta forma, dada uma série temporal $y_t, (t = 1, \dots, T)$, a decomposição ortogonal *wavelet* de nível p de y_t consiste de

$$y_t \cong y_{A_p,t} + \sum_{m=p}^{p+(p-1)} y_{D_m,t} \quad (7)$$

onde $p \in \mathbb{Z}$ é o parâmetro de nível de escala (resolução ou frequência), y_{A_p} é a componente *wavelet* (CW) de aproximação de nível p , y_{D_m} é uma CW de detalhe de nível $p, p+1, \dots$, e $y_{A_p,t} = \sum_{n \in \mathbb{Z}} a_{p,n} \phi_{p,n}(t)$ e $y_{D_m,t} = \sum_{n \in \mathbb{Z}} d_{m,n} \omega_{m,n}(t)$, onde $a_{p,n}$ é o coeficiente de aproximação de nível p e n , ϕ é a função escala ('pai'), $d_{m,n}$ é o coeficiente de detalhe de nível m e n , ω é a função *wavelet* ('mãe'). Para um m_0 fixo, os coeficientes de aproximação e de detalhes são dados pelo produto interno entre a série temporal y_t e as funções escala e *wavelets* filhas.

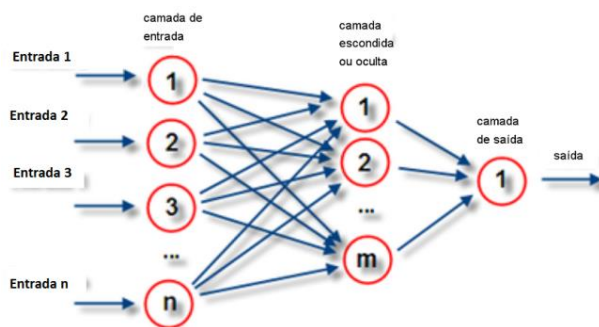
A *wavelet* de Haar $\phi = \phi_{00}$ constitui uma base ortonormal e é dada por $\phi(x) = 1$, se $x \in [0, 1/2)$, -1 , se $x \in [1/2, 1)$ e 0 caso contrário. Além disso, segundo Walnut (2002), existe um nível de decomposição em que se pode aproximar tanto quanto se queira uma função contínua por um conjunto de *wavelets* de Haar.

2.3. Redes Neurais artificiais (RNA's)

Uma RNA consiste em um sistema que conecta um conjunto de entradas até um conjunto de saídas de uma forma não linear. As conexões são feitas via uma ou mais camadas escondidas de neurônios. Uma função linear das entradas leva aos neurônios (ou nós) da camada escondida e, a estes, é aplicada uma função não linear para conduzir aos nós de saída (CHATFIELD, 1996).

Um dos modelos de RNA mais utilizados em problemas de previsão são as *feed-forward*, que se caracterizam por possuir três camadas (uma de entrada, uma escondida e uma de saída) ligadas por conexões sem ciclos (ADHIKARI e AGRAWAL, 2013), (BRAGA, CARVALHO e LUDERMIR, 2000). A arquitetura de uma rede neural *feed-forward* com três camadas é apresentada na Figura 1.

Figura 1 - Arquitetura de uma RNA *feed-forward* com uma camada oculta e uma de saída.



Fonte: Adaptado de Braga *et.al.* (2000)

A saída do modelo de RNA *feed-forward* é calculada através da equação

$$y_t = \alpha_0 + \sum_{j=1}^m \alpha_j g \left(\beta_{0j} + \sum_{i=1}^n \beta_{ij} y_{t-i} \right) + \varepsilon_t \quad (8)$$

onde y_{t-i} ($i = 1, 2, \dots, n$) são as n entradas e y_t é a saída, os inteiros n e m são número de nós de entrada e na camada escondida, respectivamente, α_j ($j = 1, 2, \dots, m$) e β_{ij} ($i = 0, 1, \dots, n; j = 0, 1, \dots, m$) são os pesos das conexões, ε_t é um ‘choque aleatório’ e α_0 e β_{0j} são os termos do *bias* (viés) e a função logística sigmoide $g(x) = 1/(1 + e^{-x})$ normalmente é utilizada com a função de ativação não linear. O modelo de RNA é equivalente a um modelo autorregressivo não linear.

Um modelo de RNA deve gerar um erro pequeno não somente dentro da amostra (conjunto de treinamento), mas, também fora da amostra (conjunto de teste). Deve-se ter muito cuidado na escolha do número de neurônios na camada de entrada e na camada escondida. Entretanto, esta é uma tarefa difícil que não possui teoria disponível para a resposta ótima desta questão (ADHIKARI e AGRAWAL, 2013; ZHANG, 2003).

2.4. Modelos Híbridos

Métodos híbridos são geralmente construídos na suposição de que uma série temporal é composta de uma componente linear e uma componente não linear, como em

$y_t = L_t + N_t$. Neste artigo, tem-se que, após se obter uma aproximação de L_t , \hat{L}_t , usando um modelo *ARIMA*, se obtêm os resíduos $e_t = y_t - \hat{L}_t$ ($t=t', \dots, T$), onde t' são os graus de liberdade perdidos na modelagem *ARIMA*, as relações não lineares nos resíduos passados podem ser modeladas por meio de uma RNA, $f(e_{t-1}, e_{t-2}, \dots, e_{t-k})$, onde $f(e_{t-1}, e_{t-2}, \dots, e_{t-k}) = \hat{N}_t$. Finalmente, as previsões híbridas, $\hat{y}_t = \hat{L}_t + \hat{N}_t$ ($t=t', \dots, T+h$), onde h denota o horizonte de previsão, são geradas. Zhang (2003) salienta que a modelagem dos resíduos via RNA's podem acarretar melhorias preditivas nas previsões oriundas de modelos *ARIMA* tradicionais.

Para Khashey e Bijari (2011), os modelos de RNA e *ARIMA* têm obtido sucesso em seus domínios não lineares e lineares, entretanto, nenhum deles é um modelo universal adequado para todas as ocasiões. Como é difícil dispor do conhecimento completo das características dos dados de um problema real, o método híbrido, que tem a capacidade de modelar linear e não linearmente, tende a ser mais adequado. A ideia básica da combinação de modelos é usar cada característica única do modelo a fim de capturar diferentes padrões nos dados.

3. Procedimentos Metodológicos

A série temporal utilizada nesta pesquisa refere-se à barragem da usina hidrelétrica de Itaipu, que iniciou sua operação nos anos 80 e foi, até 2013, a maior geradora de energia elétrica do mundo, superada apenas em 2014 pela usina de Três Gargantas, na China.

De acordo com Rosso, Fiorini *et al.* (1995) a barragem principal de Itaipu é constituída por blocos tipo gravidade aliviada com 34m de largura e 612m de extensão total e na região do leito do Rio Paraná atinge a altura máxima de 196m. Ao lado desta estrutura, a barragem principal de Itaipu é integrada por 83 blocos com 17m de largura, altura média de 60m e extensão total de 1411m e dividida em trechos. Ambas as estruturas (de concreto e fundação) possuem instrumentação detalhada e merecem acompanhamento pormenorizado do seu comportamento durante o período operacional da barragem com vistas a verificar as hipóteses de projeto e supervisionar sua segurança.

Devido à grande extensão da barragem principal (sem considerar a barragem de terra e enrocamento), por sugestão da equipe de segurança de barragens de Itaipu, este

trabalho limitou-se a um pêndulo direto localizado no trecho E da barragem da usina que é o único instrumento responsável por medir a variável deslocamento da crista da barragem na direção do fluxo e normal a ele em relação a sua base. Sua escolha deu-se em função da necessidade da equipe de segurança de barragem da usina de monitorar e realizar previsões sobre seu comportamento. Entretanto, os procedimentos empregados neste instrumento podem ser estendidos a outros instrumentos da barragem. Além disso, não há o emprego regular de técnicas estocásticas para o monitoramento de instrumentos.

Os dados fornecidos por Itaipu são compostos por duzentas e oitenta observações mensais do pêndulo direto localizado no trecho E da barragem, registradas no período de julho de 1991 a outubro de 2014. A amostra foi separada em amostra de treinamento, com 240 dados do período de julho de 1991 a junho de 2011, usada para construção e estimativa dos parâmetros dos modelos e amostra de validação, com 40 dados do período de julho de 2011 a outubro de 2014, para avaliar o desempenho em termos da capacidade preditiva das diversas abordagens (modelos) propostas. Foram empregadas quatro medidas de avaliação de desempenho: MAPE (Erro absoluto percentual médio), MSE (Erro quadrático médio), MAE (Erro absoluto médio) e R^2 (coeficiente de explicação).

Com a intenção de comparar os modelos, as técnicas testadas utilizando *wavelet*, *ARIMA-GARCH* e *RNA's*, bem como as combinações destas, foram adotadas onze abordagens conforme a Tabela 1.

Os modelos *ARIMA* e *ARIMA-GARCH* foram determinados através do *software EViews*, versão 8, e a escolha do modelo final se deu em termos da ordem de diferenciação para obter estacionariedade, sazonalidade, significância dos termos autorregressivos e de médias móveis, diagnóstico dos resíduos ordinários padronizados e resíduos quadráticos (função de autocorrelação e autocorrelação parcial), e da estatística de Durbin-Watson.

Para os modelos de redes neurais foram usadas *RNA's* do tipo *feed-forward*, desenvolvidas no *software Matlab*. Os parâmetros selecionados foram: janela de previsão variando de 1 a 13 e número de neurônios na camada escondida variando de 1 a 10. Amostra de treinamento ($200/280 = 0,71$), amostra de validação ($40/280 =$

0,145) e amostra de teste ($40/280 = 0,145$). A medida do desempenho para a seleção da rede ótima foi o *MSE*.

Tabela 1 - Abordagens testadas.

Abordagem	Descrição
1	Modelagem <i>ARIMA</i> da série original.
2	Modelagem com Rede Neural da série original.
3	Modelo híbrido de combinação linear <i>ARIMA</i> + Rede Neural (em paralelo) da série original.
4	Modelo híbrido de combinação linear <i>ARIMA</i> (série original) + Rede Neural (seqüencial – rede aplicada nos resíduos <i>ARIMA</i> da série original).
5	Decomposição Wavelet (série original) + Modelagem <i>ARIMA – GARCH</i> em cada componente.
6	Decomposição Wavelet (série original)+ Modelagem com rede neural em cada componente.
7	Decomposição Wavelet (série original) + Modelo híbrido em paralelo da combinação linear ótima da soma da combinação ótima <i>ARIMA – GARCH</i> de todas as componentes com a combinação ótima de REDE NEURAL de todas as componentes.
8	Decomposição Wavelet (série original) + Modelo híbrido em paralelo da soma das combinações ótimas (<i>ARIMA – GARCH + RNA</i>) em cada componente.
9	Decomposição Wavelet (série original) + Modelo híbrido em paralelo da combinação linear ótima da soma das combinações ótimas (<i>ARIMA – GARCH + RNA</i>) em cada componente.
10	Decomposição Wavelet (série original) + Modelagem <i>ARIMA – GARCH</i> em cada componente seguida da modelagem dos resíduos de cada componente com Redes neurais.
11	Modelo híbrido da combinação da soma das previsões <i>ARIMA – GARCH</i> das componentes com as previsões de redes neurais dos resíduos em cada componente.

Fonte: Proposição dos autores.

No caso da decomposição wavelet da série original foi considerado somente um tipo de decomposição, em que, por simplicidade, se utilizou o nível de decomposição $p = 2$ e a base de *Haar*, tanto para a posterior modelagem *ARIMA-GARCH* quanto para modelagem com rede neural, da qual se obteve uma série de aproximação e duas séries de detalhes. A decomposição também foi feita no *MatLab*, v.8.

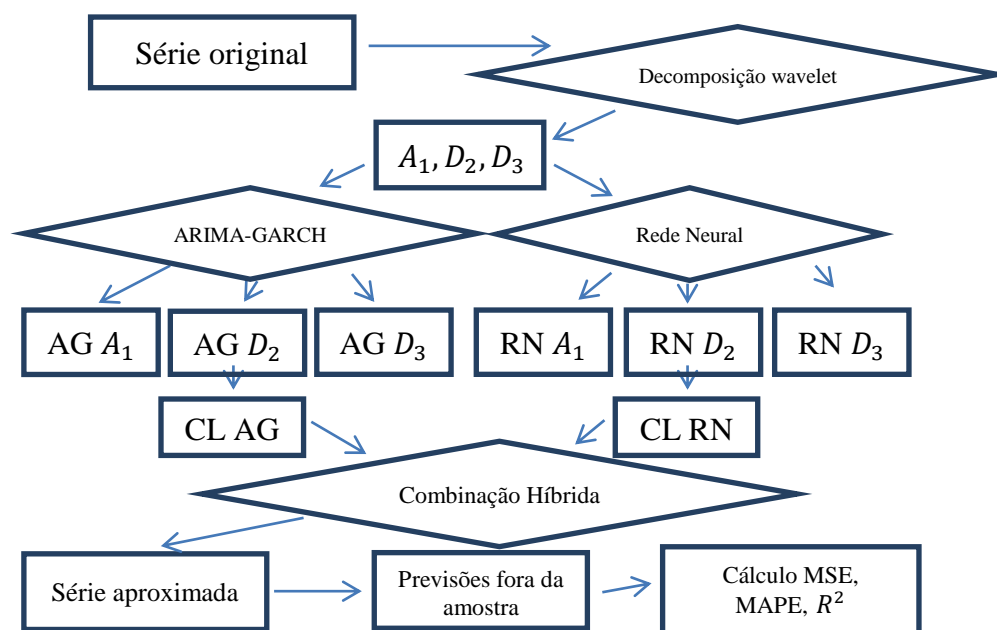
A Figura 2, a seguir, apresenta um fluxograma da abordagem 7, mas serve de ilustração para representar a maioria das abordagens, algumas das quais são parte da abordagem 7 ou bastante semelhantes a esta. Para as combinações lineares foi feita uma regressão da variável dependente sobre as variáveis independentes. Por exemplo, a

combinação ótima *ARIMA-GARCH* da abordagem 7, chamada de Combinação Linear *Arima-Garch* (*CL AG*) na Figura 2, usa a série original como variável dependente e a modelagem com *Arima-Garch* de cada componente (A_1, D_2, D_3) da DW feita, denotadas por $AG A_1, AG D_2, AG D_3$, como variáveis independentes.

4. Resultados

Inicialmente a série original foi decomposta, segundo a base de *Haar*, nas CW A_1, D_2 e D_3 , do que se obtiveram os gráficos representados na Figura 3, onde a variação no eixo y é dada em *mm*. Em relação à modelagem *ARIMA* da série original (abordagem 1) o melhor modelo encontrado consiste de uma diferenciação estacional, uma sazonal de ordem 12, com os termos autorregressivos de ordem 1, 2 e 4 e sazonal de ordem 12 e de médias móveis de ordem 1 e sazonal de ordem 12.

Figura 2 - Fluxograma para a abordagem 7.

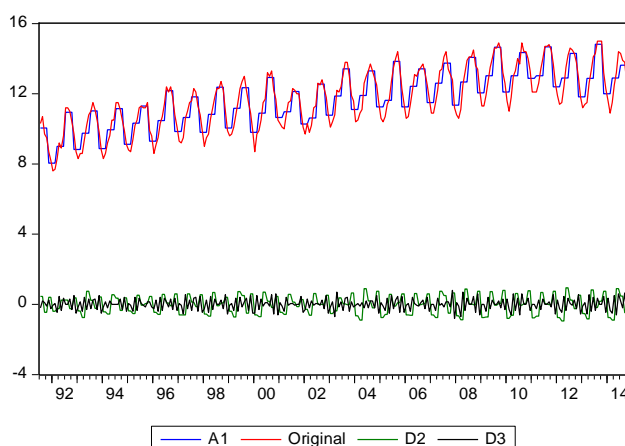


Fonte: Proposição dos autores.

Foram ajustados modelos de RNA's *feed-forward* para a série original, resíduos do modelo *ARIMA* da série original, CWA_1, D_2 e D_3 , resíduos da modelagem *ARIMA-GARCH* das CW, denominadas séries res_{A_1}, res_{D_2} e res_{D_3} , com janela de previsão (J) variando de 1 a 13 e número de neurônios (N) na camada oculta variando de 1 a 10. Na Tabela 2, estão as configurações ótimas encontradas, em qual abordagem foram utilizadas as redes e o valor do *MSE*.

Para a modelagem *ARIMA-GARCH* das *CW*'s A_1 , D_2 e D_3 os melhores modelos encontrados estão descritos na Tabela 3, onde *DE* indica ordem de diferenciação estacional, *DS* indica que houve diferenciação sazonal de ordem 12, *AR* coeficientes autorregressivos, *MA* coeficientes de médias móveis, *SAR* autorregressivos sazonais, *SMA* médias móveis sazonais, *r* e *s* as respectivas ordens *GARCH*, conforme equação(4) e a distribuição de referência dos choques aleatórios do modelo.

Figura 3 - Gráficos da série original e suas componentes wavelet.



Fonte: dados da pesquisa – saída do *software EViews*.

Tabela 2–RNA' s ótimas encontradas para as abordagens estabelecidas.

Série	Abordagem	J	N	MSE	Série	Abordagem	J	N	MSE
<i>Original</i>	2,3	12	6	0,0916	D_3	6,7,8,9	13	7	0,0134
<i>Res_ARIMA</i>	4	6	9	0,0619	<i>res_A1</i>	10,11	12	10	0,0141
A_1	6,7,8,9	10	6	0,0201	<i>res_D2</i>	10,11	8	10	0,0029
D_2	6,7,8,9	13	4	0,0039	<i>res_D3</i>	10,11	12	5	0,0147

Fonte: dados da pesquisa – saída do *software MatLab*.

Tabela 3 - Modelos *ARIMA-GARCH* ótimos encontrados para as componentes.

Série	DE	DS	AR	MA	SAR	SMA	<i>r</i>	<i>s</i>	Dist.
A_1	1	Sim	4,20	4,8	12	—	0	3	t_8
D_2	1	Sim	4,8,16	2,4	12	12	0	2	t_{49}
D_3	0	Sim	—	1	—	12	1	1	t_{19}

Fonte: dados da pesquisa – saída do *software EViews*.

Na Tabela 4 são apresentados os valores das medidas de desempenho, testadas no período de validação da amostra, nas abordagens testadas. Observa-se, em geral, melhora no desempenho da previsão com o uso da DW e da combinação de métodos.

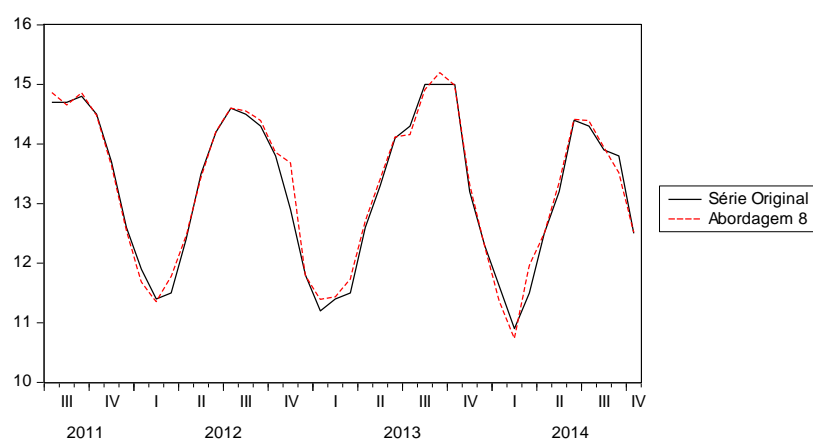
Tabela 4 - Resultados finais do desempenho das previsões por abordagem.

Abordagem	1	2	3	4	5	6	7	8	9	10	11
<i>MAPE (%)</i>	3,13	1,93	2,41	1,65	1,42	1,00	0,97	0,95	0,96	1,17	1,22
<i>MSE</i>	0,2649	0,1176	0,1569	0,0952	0,1098	0,0367	0,0372	0,0351	0,0357	0,0504	0,0608
<i>MAE</i>	0,4005	0,2525	0,3102	0,2163	0,1841	0,1281	0,1239	0,1199	0,1217	0,1530	0,1592
<i>R²(%)</i>	0,9805	0,9880	0,9851	0,9898	0,9912	0,9938	0,9940	0,9941	0,9941	0,9927	0,9924

Fonte: dados da pesquisa.

Na Figura 4 é exibido o gráfico no período de validação da série original e da melhor abordagem segundo todas as medidas de desempenho, conforme a Tabela 4.

Figura 4 - Desempenho da melhor abordagem no período de validação.

Fonte: dados da pesquisa, saída do software *EViews*.

Também foi comparado o desempenho das abordagens no período fora da amostra (julho/2011 a outubro/2014), mês a mês, para os modelos construídos no período total da amostra, de modo que as 40 observações do período fora da amostra tornaram-se 4 para os meses de julho a outubro e 3 para os demais meses. As abordagens, segundo a Tabela 1, que apresentaram o melhor desempenho em cada mês estão contidas na Tabela 5.

Tabela 5 - Melhores abordagens para cada mês.

Mês	1	2	3	4	5	6	7	8	9	10	11	12
Abordagem	2	5	4	5	4	5	8	5	7	5	2	5

Fonte: Dados da pesquisa.

A abordagem 5, decomposição wavelet seguida de modelagem *ARIMA* das componentes apresentou melhor desempenho em termos do MAE, MSE e MAPE em 6 dos 12 meses, mas nunca em períodos consecutivos, as redes neurais ‘puras’

(abordagem 2), a combinação de *ARIMA* e RNA's (abordagem 4) foram melhores em 2 meses cada e modelos híbridos (abordagem 7 e 8) tiveram melhor desempenho 1 mês cada.

5. Discussões

Os modelos 'puros', que constam nas abordagens 1 e 2 da Tabela 1 apresentam o pior desempenho nas previsões, conforme Tabela 4, desta forma o método híbrido que combina estas duas abordagens na abordagem 3, também não apresenta bons resultados. Já o modelo híbrido sequencial, abordagem 4, consegue uma melhoria nas previsões comparada aos seus antecessores. Todas estas abordagens (1 a 4) têm em comum a não utilização da decomposição *wavelet*, ou seja, a Tabela 4, mostra que todas as abordagens que contemplam a decomposição *wavelet* (5 a 11) apresentam melhor desempenho de previsão.

Das abordagens com decomposição *wavelet* (5 a 11), as abordagens 5 e 6 que só levam em conta modelagem *ARIMA-GARCH* ou RNA apresentam desempenho inferior às abordagens híbridas 7, 8 e 9. Isto confirma que, dentre as abordagens híbridas, pós-decomposição *wavelet*, os modelos híbridos melhoram a previsão. As abordagens 5 e 6 ainda superaram os modelos híbridos 10 e 11, possivelmente, porque estas modelavam com redes neurais aos resíduos do modelo *ARIMA-GARCH* e esta estratégia nem sempre produz melhorias.

As abordagens que apresentaram os melhores desempenhos foram 7, 8 e 9. Estas têm em comum, além da DW, o uso de combinações lineares híbridas dos modelos *ARIMA-GARCH* e de RNA's, combinadas em paralelo, ou seja, quando os modelos *ARIMA-GARCH* e de RNA's são desenvolvidos independentes e combinados linearmente ao final.

O fato do desempenho das redes neurais ser melhor em seis dos doze meses não permite afirmar que esta técnica seja melhor, uma vez que os modelos foram construídos com base em toda a amostra e apenas foi comparado o desempenho mês a mês. A construção de modelos mensais é uma alternativa possível, mas teria a desvantagem de reduzir drasticamente o tamanho da amostra o que poderia enfraquecer a capacidade preditiva.

6. Conclusões

Este artigo buscou promover a determinação e a melhoria no desempenho da previsão da série temporal de dados de leitura de um pêndulo direto que faz o monitoramento de deslocamentos da crista de uma barragem de usina hidrelétrica, na direção do fluxo, usando métodos estocásticos e de aprendizado de máquina.

Foram construídas onze abordagens de comparação nos quais se variava a utilização ou não de decomposição *wavelet*, modelagem *ARIMA-GARCH*, com redes neurais e combinações híbridas destas técnicas. O melhor desempenho, para todo o período de validação, foi apresentado pela abordagem 8, decomposição *wavelet* da série original seguida de um modelo híbrido, em paralelo, da soma das combinações ótimas *ARIMA-GARCH* com RNA em cada componente.

Para a série temporal estudada, de modo geral, o uso da decomposição *wavelet* tem desempenho superior de previsão a não utilização de *wavelets*, a combinação do uso das técnicas *ARIMA-GARCH* e de RNA's tem desempenho superior ao uso de somente uma delas e as abordagens com melhor desempenho são aquelas que consideram a combinação de todas as técnicas. As combinações híbridas nem sempre tem desempenho superior, isto depende dos desempenhos individuais de cada técnica.

Como sugestão de trabalho futuro pode se avaliar as previsões usando outras bases *wavelet* e níveis de decomposição maiores do que dois, usado neste trabalho. Acredita-se que, com isso, poderá ser melhorado ainda mais o desempenho. Esta questão é uma limitação atual do trabalho, pois a busca de modelos 'ótimos' *ARIMA-GARCH* nas componentes ainda não pode ser automatizada, pois o diagnóstico dos resíduos requer a avaliação de diversos fatores. A busca de modelos mensais ou trimestrais também pode ser uma alternativa a ser tentada.

Referências

- ADHIKARI, R.; AGRAWAL, R. K. **An introductory study on time series modelling and forecasting**. [S.l.]: Lambert Academic Publishing, 2013.
- AMJADY, N.; HEMMATI, M. Day-ahead price forecasting of electricity markets by a hybrid intelligent system. **European Transactions on Electrical Power**, v. 19, p. 89-102, 2009.
- BOLLERSLEV, T. Generalized Autorregressive Conditional Heteroscedasticity. **Journal of Econometrics**, v. 31, n. 3, p. 307-326, Apr. 1986. doi:10.1016/0304-4076(86)90063-1.
- BOX, G. E. P.; JENKINS, G. W.; REINSEL, G. C. **Time-series analysis - forecasting and control**, 4^a ed. Hoboken, New Jersey: John Wiley & Sons, Inc., 2008.

BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. **Redes Neurais Artificiais: Teoria e aplicações**. Rio de Janeiro : LTC, 2000.

CAO, M.; QIAO, P.; REN, Q. Improved hybrid wavelet neural network methodology for time-varying behavior prediction of engineering structures. **Neural Comput. & Applic.**, v. 18, p. 821-832, 2009.

CHATFIELD, C. **The analysis of time series**. 5ª. ed. London: Chapman & Hall/CRC, 1996.

DÍAZ-ROBLES, L. A. et al. A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas; The case of Temuco, Chile. **Atmospheric Environment**, v. 42, p. 8331-8340, 2008.

ENGLE, R. F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom inflation. **Econometrica**, v. 50, n. 4, p. 987-1007, July 1982. <http://www.jstor.org/stable/1912773>.

GUJARATI, D. N.; PORTER, D. C. **Econometría**. 5ª. ed. México: McGRAW-HILL/INTERAMERICANA, 2010.

ITAIPU BINACIONAL. **Valores de Controle para os instrumentos integrantes do "ADAS" e critérios para a geração de alarmes**. Foz do Iguaçu. 2002.

KHASHEY, M.; BIJARI, M. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. **Applied soft computing**, v. 11, p. 2664-2675, 2011.

KRIECHBAUMER, T. et al. An improved wavelet-ARIMA approach for forecasting metal prices. **Resources Policy**, v. 39, p. 32-41, 2014.

KUBRUSLY, C. S. **The Elements of Operator Theory**. 2nd. ed. Boston: Birkhauser, 2011.

KUBRUSLY, C. S.; LEVAN, N. Abstract Wavelets Generated by Hilbert Space Shift Operators. **Adv. Math. Sci. Appl.**, v. 16, p. 643-660, 2006.

KUMAR, U.; DE RIDDER, K. GARCH modelling in association with FFT e ARIMA to forecast ozone episodes. **Atmospheric Environment**, v. 44, p. 4256-4265, 2010.

LIU, H. et al. Forecasting models for wind speed using wavelet, wavelet packet, time series and artificial neural networks. **Applied Energy**, v. 107, p. 191-208, 2013.

LIU, H.; SHI, J. Applying ARIMA-GARCH approaches to forecasting short-term electricity prices. **Energy Economics**, v. 37, p. 152-166, 2013.

MALLAT, S. **A wavelet tour of signal processing**. 2nd. ed. San Diego: Academic Press, 1998.

MORETTIN, P. A.; TOLOI, C. M. **Análise de séries temporais, 2ª edição rev. e ampl.** São Paulo: Edgard Blucher, 2006.

ROSSO, J. A. et al. **Barragem de Itaipu - Lições aprendidas com o comportamento térmico das estruturas tipo gravidade aliviada e contraforte**. Foz do Iguaçu. 1995.

SALAZAR, F. et al. Data-Based Models for the Prediction of Dam Behaviour: A Review and Some Methodological Considerations. **Archives of Computational Methods in Engineering**, p. 1134-1160, 25 July 2015. DOI 10.1007/s11831-015-9157-9.

TAN, Z. et al. Day ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. **Applied energy**, v. 87, p. 3606-3610, 2010.

TEIXEIRA JR., L. A. **Combinação SSA-wavelet de métodos preditivos com ajuste numérico MINIMAX, na geração de previsões e cenários.** Rio de Janeiro: Tese de Doutorado - Departamento de Engenharia Elétrica - PUC, 2013.

TEIXEIRA JR., L. A. et al. Artificial neural network and wavelet decomposition in the forecast of global horizontal solar radiation. **Pesquisa Operacional**, Rio de Janeiro, v. 35, n. 1, p. 73-90, Jan./Apr. 2015.

WALNUT, D. F. **An introduction to wavelet analysis.** Boston : Birkhauser, 2002.

ZHANG, G. P. Time series forecasting using a hybrid ARIMA and neural network model. **Neurocomputing**, v. 50, p. 159-175, 2003.

IMPROVING IN TIME SERIES FORECASTING OF DAM MONITORING INSTRUMENTS VIA COMBINATION OF METHODS

Abstract

Hydroelectric dams are evaluated by visual inspections and monitoring instruments. Periodic readings of an instrument can be interpreted as a stochastic time series whose past values provide relevant information to build predictions about their future values, and indicate trends of future dam behavior as a whole. Thus, it is essential to produce more accurate forecasts possible, so as to serve as a reliable prediction alert to any abnormalities in the dam behavior, allowing for early execution of works and intervention actions. So, this article proposes a methodology in which an assessment is made of the performance by time series forecasting the readings of an instrument of the dam Itaipu hydroelectric plant using a hybrid combination of the ARIMA-GARCH forecasters and artificial neural networks with the use of the wavelet decomposition. The results show that the proposed hybrid combination achieved performance in terms of accuracy, considerably higher when compared with the single use of traditional predictive methods and your combinations.

Key-words: *Dam safety; Stochastic Time Series; ARIMA-GARCH Models; Artificial Neural Networks; Wavelet Decomposition.*