

COMPARATIVO ENTRE ARIMA, REDES NEURAIAS ARTIFICIAIS, MÁQUINAS DE VETORES SUPORTE PARA REGRESSÃO E COMBINAÇÃO LINEAR DE PREVISÕES DE VAZÕES MENSAIS NO POSTO 266- ITAIPU

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

SAMUEL BELLIDO RODRIGUES - samuelbellido@hotmail.com
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ - UTFPR-MEDIANEIRA

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

ARINEI CARLOS LINDBECK DA SILVA - arineicls@gmail.com
UNIVERSIDADE FEDERAL DO PARANÁ - UFPR

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

Resumo: *A PREVISÃO DE VAZÕES MENSAIS E DIÁRIAS É DE SUMA IMPORTÂNCIA PARA O PLANEJAMENTO OPERACIONAL DAS USINAS HIDRELÉTRICAS, SENDO VÁRIOS OS MÉTODOS EXISTENTES PARA A RESOLUÇÃO DO PROBLEMA. NESTE ESTUDO FEZ-SE A PREVISÃO DA SÉRIE DE VAZÕES MENSAIS OBTIDAS NO POSTO 266-ITAIPU ATRAVÉS DOS MÉTODOS ARIMA, REDES NEURAIAS ARTIFICIAIS (RNA) E MÁQUINAS DE VETORES SUPORTE PARA REGRESSÃO (SVR), E COMBINAÇÃO LINEAR (CL) DAS PREVISÕES ORIUNDAS DESTES MÉTODOS. PARA A SÉRIE ESTUDADA, AS PREVISÕES PROVENIENTES DA MODELAGEM POR SVR SÃO MELHORES AOS OBTIDOS POR RNA E ARIMA. SUPERIORES AOS RESULTADOS AUFERIDOS PELO SVR FORAM OS OBTIDOS PELA COMBINAÇÃO LINEAR DAS PREVISÕES ORIUNDAS DOS TRÊS MÉTODOS (ARIMA, RNA E SVR), MOSTRANDO SER A CL UM PROCEDIMENTO ÚTIL E SE UTILIZADO LEVA A GANHOS PREDITIVOS CONSIDERÁVEIS.*

Palavras-chaves: *ARIMA; REDES NEURAIAS ARTIFICIAIS; COMBINAÇÃO LINEAR DE PREVISÕES; MÁQUINAS DE VETORES SUPORTE PARA REGRESSÃO; PREVISÃO DE SÉRIES TEMPORAIS*

Área: 6 - PESQUISA OPERACIONAL
Sub-Área: 6.1 - PROGRAMAÇÃO MATEMÁTICA

COMPARING ARIMA, ARTIFICIAL NEURAL NETWORKS, SUPPORT VECTOR REGRESSION AND LINEAR COMBINATION FORECASTS OF MONTHLY STREAMFLOWS IN STATION 266-ITAIPU

Abstract: *THE FORECAST MONTHLY AND DAILY STREAMFLOWS OF PARAMOUNT IMPORTANCE FOR THE OPERATIONAL PLANNING OF HYDROELECTRIC PLANTS, WITH SEVERAL EXISTING RESOLUTION METHODS FOR PROBLEM. THIS STUDY MADE THE PREDICTION SERIES OF MONTHLY FLOW RATES ACHIEVED IN THE STATION 266-ITAIPU METHODS THROUGH ARIMA, ARTIFICIAL NEURAL NETWORKS (ANN) AND SUPPORT VECTOR REGRESSION (SVR), AND LINEAR COMBINATION (LC) OF THE PREDICTIONS ARISING FROM THESE METHODS. FOR THE SERIES STUDIED, THE PREDICTIONS FROM THE SVR MODELING ARE BEST OBTAINED BY THE ANN AND ARIMA. HIGHER THAN THOSE EARNED BY THE SVR RESULTS WERE OBTAINED BY THE LINEAR COMBINATION OF FORECASTS DERIVED FROM THE THREE METHODS (ARIMA, ANN AND SVR), SHOWING THAT THE LC IS A USEFUL PROCEDURE AND IF USED LEADS TO CONSIDERABLE PREDICTIVE GAINS.*

Keyword: *ARIMA; ARTIFICIAL NEURAL NETWORKS; LINEAR COMBINATION OF FORECASTS; SUPPORT VECTOR REGRESSION; FORECASTING TIME SERIES*

1 Introdução

No Brasil a principal fonte de energia elétrica é oriunda da produção hidráulica, impulsionada pela 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 afluyente) disponível para a movimentação das turbinas geradoras. De forma que longos períodos de seca podem levar a uma diminuição da capacidade geradora de energia da usina. Então, fazer a previsão de vazões é de grande relevância no planejamento das usinas hidrelétricas.

O Operador Nacional do Sistema Elétrico (ONS) utiliza as informações sobre as vazões na otimização da geração de energia e elaboração de estratégias que minimizem efeitos negativos de eventuais enchentes ou secas. Um dos principais modelos de previsão de vazões utilizado pelo ONS é o PREVIVAZ, fundamentado na metodologia apresentada por Box e Jenkins (1970), que consiste em ajustar modelos lineares auto-regressivos integrados de médias móveis. Entre os outros métodos utilizados pelo ONS pode-se citar o SMAP-MEL na bacia do rio Paraná e redes neurais artificiais (RNA) na bacia do rio Iguaçu.

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 $AR(p)$ 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.*,(2006a), 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 (Wang, *et al.*, 2006b). São vários os trabalhos existentes na literatura que discutem a previsão de séries oriundas de processos não lineares, sendo predominante para este caso o uso de Redes Neurais Artificiais (RNA). Ballini (2000), Wang *et al.*,(2006a), Valenzuela *et al.*, (2008), Ferreira (2012); Abreu, *et al.*, (2012) e Teixeira, Jr *et al.*; (2012), entre outros, utilizaram Redes Neurais Artificiais (RNA) na previsão de séries temporais.

Support vector regression (SVR) é outra técnica utilizada com frequência na previsão de séries temporais, é baseado na teoria do aprendizado estatístico ou teoria VC (SMOLA e SCHÖLKOPF, 1998). Fazendo uma revisão da literatura pode-se encontrar com facilidade estudos comparativos entre o desempenho do SVR e outros métodos preditivos. He *et al*, (2014) e Lafdani, *et al*, (2013) mostraram, para os casos estudados, que tanto RNA como SVR apresentaram boas previsões e ligeira superioridade do SVR em relação a RNA.

A proposta deste artigo é fazer previsões séries de vazão mensal no posto 266 (Itaipu) localizado no rio Paraná. Para tanto, foram usados os modelos ARIMA, RNA e SVR, e combinação linear das previsões oriundas destes três métodos. Os resultados foram comparados a partir das estatísticas raiz do erro quadrático médio (RMSE) e média dos erros absolutos (MAE).

2 Support Vector Regression

A regressão de vetores de suporte (SVR, *Support Vector Regression*) é fundamentada em uma máquina de vetor de suporte (SVM, *Support Vector Machine*), (VAPNIK, 1995). O modelo produzido por SVR depende apenas de um subconjunto dos dados de formação e por sua função de custo que é utilizada para a construção do modelo, ignorando os dados de formação perto do modelo de previsão.

As descrições a seguir sobre o SVR baseiam-se no tutorial de (SMOLA e SCHÖLKOPF, 1998), mais detalhes, podem ser encontrados nesse guia.

Seja uma amostra de dados de treinamento $\{(x_1, y_1), \dots, (x_l, y_l)\} \subset X \times \mathbb{R}$, onde X indica o espaço de padrões de entrada, o objetivo é encontrar uma função $f(x)$, que apresente no máximo um desvio ε dos alvos y_i , obtido para todos os dados de treinamento. Logo, descreve-se a função linear de aproximação como sendo $f(x) = \langle w, x \rangle + b$ com $w \in X$, $b \in \mathbb{R}$, onde $\langle \cdot, \cdot \rangle$ denota o produto interno em X . Contudo, nem sempre é possível garantir a viabilidade do problema, já que existem pontos que violam as restrições, assim chegamos à formulação mencionada em (VAPNIK, 1995), com variáveis de folga:

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

Para a formulação dual, serão introduzidos alguns multiplicadores de Lagrange (α_i, α_i^*) não negativos conforme relata (VAPNIK, 1995) e (SMOLA e SCHÖLKOPF, 1998).

Logo o problema, agora de maximização, na sua forma dual é dado por:

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

de aproximação passa a ser $f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b$.

A formulação dual do problema SVR fornece como alternativa trabalhar em um espaço de alta dimensionalidade. Assim, pode-se realizar um mapeamento não linear dos dados de entrada para um espaço de dimensão maior, onde a regressão linear torna-se possível. Para isso, utiliza-se a abordagem baseada em funções kernel $K(x, x') := \langle \phi(x), \phi'(x') \rangle$, cuja sua introdução no problema de otimização faz com que este passe a ser descrito da seguinte forma:

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

aproximação da SVR não linear passa a ser $f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x) + b$.

Dentre as funções kernel mais utilizadas no algoritmo SVR destacam-se o kernel: linear, polinomial homogêneo, polinomial não homogêneo, sigmoïdal e gaussiano.

3 Redes Neurais Artificiais

O mais fantástico sistema de aprendizado existente na natureza é o cérebro humano, este é constituído por neurônios, que são células capazes de armazenar e processar informações. Os neurônios estão ligados uns aos outros formando uma rede neural e o fluxo das informações (sinais) por esta rede dependerá de uma fenda existente entre os neurônios denominada de fenda sináptica, ou sinapse. Segundo Castro (2010), 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 (2001) 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) (TEIXEIRA Jr *et al.*, 2012).

O problema de previsão de séries temporais com T observações $\{y_t\}_{t=1}^T$ através de RNA consiste na utilização de dados observados para se prever $y(t+k)$, onde k é o passo de tempo de previsão para o futuro. Define-se o tamanho da janela (n) e o valor de k , com isto, montam-se as sequências $\{y(t), y(t-1), \dots, y(t-n+1), y(t+1), \dots, y(t+k)\}$ que formaram 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 Metodologia Box e Jenkins

Box e Jenkins (1970) 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. Segundo Box e Jenkins (1970) 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 (1):

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

O modelo (1) combina valores passados das entradas Z_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 δ 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) (HAMILTON 1994). 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 (1970) e Morettin e Tolo (2006). Para a validação do modelo já com os parâmetros estimados pode-se usar testes estatísticos, tais como: teste de Box-Pierce, teste do periodograma acumulado, teste da autocorrelação cruzada, entre outros. No caso da série temporal $(y_t)_{t=1}^T$ apresentar componente sazonal, o modelo de Box & Jenkins plausível é dado genericamente, em (2):

$$\phi(B)(1-\Phi_1 B - \dots - \Phi_p B^{ps}) \nabla^d (1-B^S)^D y_t = \theta(B)(1-\Theta_1 B - \dots - \Theta_q B^{qs}) \varepsilon_t \quad (2)$$

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.

4 Combição De Previsões

Os precursores na área das combinações de previsões são Bates e Granger (1969), que sugeriram uma combinação, expressa em (3), para a previsão da série temporal $\{y_t\}_{t=1}^T$.

$$\hat{y}_t = \rho_1 \times \hat{y}_{1,t} + \rho_2 \times \hat{y}_{2,t} \quad (3)$$

Em (3), a previsão combinada linearmente é representada por \hat{y}_t , enquanto as previsões obtidas através dos métodos 1 e 2 são $\hat{y}_{1,t}$ e $\hat{y}_{2,t}$, respectivamente. Sendo os coeficientes ρ_1 e ρ_2 calculados a partir da matriz de covariância de erros de previsões individuais.

Gupta e Wilton (1977) usaram em seus trabalhos a média aritmética das previsões individuais. Granger e Ramanathan (1984) propuseram, entre outros métodos, uma combinação linear sem restrição para os coeficientes e adição de uma constante. Terui e Dijk (2002) sugerem o uso de filtros de Kalman para o cálculo dos coeficientes e que estes sejam variáveis ao longo do tempo.

Teixeira Jr (2013) utilizou programação matemática para calcular os coeficientes da combinação linear. O autor otimizou um problema de programação não linear, onde a função objetivo era constituída pela soma dos erros de previsão ao quadrado.

Segundo Granger (1989), a combinação de previsões possibilita a geração de melhores previsões. De acordo com Hollauer *et al* (2008), a diversificação de previsões leva à diminuição do erro. Faria e Mubwandarikwa (2008) salientam que a previsão combinada é uma agregadora de informações oriundas de diferentes fontes (no caso, os métodos preditivos base) sobre a flutuação estocástica da série temporal a ser modelada. Em Wallis (2011), é feita uma revisão histórica, na qual são mencionados muitos dos principais artigos sobre combinação de métodos preditivos individuais, publicados em quarenta e dois anos, os quais atestam a sua utilização e a sua eficiência, além de mostrarem uma enorme diversidade de aplicações e abordagens.

5 Materiais e Métodos

A série de vazões mensais $(y_t)_{t=1}^{492}$ usada neste trabalho envolve dez 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. Os dados que compõem a série são provenientes das medições de vazões no posto 266. Neste posto 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 Guaíra (Brasil) e Salto del Guairá (Paraguai), ao norte.

As séries temporais de vazões possuem como característica a não-estacionariedade, com a presença de componente sazonal. Ballini (2000) sugere extrair tais componentes do processo de previsão através de um tratamento estatístico, onde a série original é transformada em outra a partir de uma padronização. Quando as séries estudadas apresentam indícios de não-estacionariedade, Box; Jenkins e Reinsel (1994) sugerem a diferenciação da série a fim de se obter séries estacionárias, neste caso o modelo escolhido é o ARIMA(p,d,q), onde d é a ordem de diferenciação da série. Como a série $(y_t)_{t=1}^{492}$ possui componente sazonal, foi necessário o uso do modelo ARIMA multiplicativo, representado ARIMA(p,d,q)x(P,D,Q). A estimativa dos parâmetros foi realizada por máxima verossimilhança a partir do aplicativo computacional *EViews* 8. A definição das ordens do modelo ocorreu com a análise dos gráficos da FAC e FACP, resíduos e testes com diversas opções de ordens p,d e q.

A fim de validar o modelo ajustado foram analisados os resíduos com a construção de gráficos de autocorrelações, buscando mostrar que os mesmos são não autocorrelacionados. A validação é reforçada com o periodograma integrado, segundo Morettin e Toloí (2006) se o modelo não for adequado, haverá desvios sistemáticos da reta no gráfico do periodograma acumulado $C(f_j)$ pela frequência (f_j) .

Para testar a hipótese que todos os coeficientes de autocorrelação ρ_k são iguais a zero, foi utilizado a estatística $Q = n \sum_{k=1}^m \hat{\rho}_k^2$, desenvolvida por Box e Pierce, em que n é o tamanho da amostra e m a duração da defasagem. A estatística Q tem distribuição qui-quadrado com m graus de liberdade. Quando Q excede o valor crítico, rejeita-se a hipótese nula de que todos os ρ_k são iguais a zero (GUJARATI, 2000).

A RNA usada na modelagem da série $(y_t)_{t=1}^{492}$ foi *feedforward multilayer perceptron* (RNA-MLP). Na busca pela RNA que melhor ajustasse os dados, identificou-se a RNA-MLP com a seguinte estrutura topológica: tamanho de janela de entrada igual a cinco, uma camada escondida composta por cinco neurônios artificiais com função de ativação *tansig* e um neurônio na camada de saída com função de ativação *purelin* (HAYKIN, 2001). O algoritmo de treinamento utilizado foi o *Levenberg & Marquardt* e realizado no aplicativo computacional Matlab.

Para o uso do SVR foi utilizado uma interface Matlab para SVM denominada de LIBSVM (CHANG e LIN). Neste estudo, foi utilizado ε -SVR para investigar a previsão da vazão. Foram feitos vários testes e estes demonstraram um melhor ajustamento para o modelo de previsão. Os dados foram divididos em duas partes. Um deles é usado para treinar o modelo, enquanto o outro, chamado de conjunto de validação, é usado para validar o modelo. Nesse caso o tratamento é não linear e os seguintes valores foram usados para os parâmetros:

gamma: $g = \frac{1}{443}$, Custo: $C = 100$, Kernel = RBF, epsilon: $p = 0,001$.

Das 492 observações da série analisada, quarenta e nove foram reservadas para testes com valores previstos um passo à frente. Tendo feitas as previsões (\hat{y}_t) com o modelo ajustado a partir das primeiras 443 observações, obteve-se o RMSE (equação (4)) e MAE (equação (5)) relativos à amostra de teste.

$$RMSE = \sqrt{\frac{1}{49} \sum_{t=444}^{492} (y_t - \hat{y}_t)^2} \quad (4)$$

$$MAE = \frac{1}{49} \sum_{t=444}^{492} |y_t - \hat{y}_t| \quad (5)$$

Da modelagem pelos métodos SVR, RNA e ARIMA resultaram as previsões \hat{y}_t^{SVR} , \hat{y}_t^{RNA} e \hat{y}_t^{ARIMA} . No passo seguinte obteve-se a previsão de y_t através da combinação linear das previsões \hat{y}_t^{SVR} , \hat{y}_t^{RNA} e \hat{y}_t^{ARIMA} , conforme (6):

$$\{\hat{y}_t^{CL}\}_{t=1}^T = \alpha_{SVR} \times \hat{y}_t^{SVR} + \alpha_{RNA} \times \hat{y}_t^{RNA} + \alpha_{ARIMA} \times \hat{y}_t^{ARIMA} \quad (6)$$

Onde: α_{SVR} , α_{RNA} e α_{ARIMA} são parâmetros a serem determinados com a otimização do problema de programação não linear (7), cuja função objetivo é a raiz do erro quadrático médio (RMSE).

$$\begin{aligned} \text{Min } RMSE &= \sqrt{\frac{1}{T} \sum_{t=1}^T \left(\{y_t\}_{t=1}^T - \{\hat{y}_t^{CL}\}_{t=1}^T \right)^2} \\ \text{s.a. } &\alpha_{SVR}, \alpha_{RNA}, \alpha_{ARIMA} \text{ são irrestritas} \end{aligned} \quad (7)$$

Para a otimização do problema de programação não linear (7) foi usado o aplicativo computacional LINGO 11.

6 Resultados Obtidos

A análise do correlograma (figura 1(a)) da série observada mostra a existência de componente de sazonalidade com período igual a doze e autocorrelações não decrescentes nos lags múltiplos de doze, indicando não-estacionariedade. A indicação de não-estacionariedade foi reforçada pelo teste de Dickey Fuller aumentado para um nível de significância de 5% (p-

valor obtido = 0,0026). A utilização do teste de Dickey Fuller é válida quando o processo gerador da série temporal é um processo auto-regressivo (GUJARATI, 2000).

A fim de eliminar o efeito de não-estacionariedade e sazonalidade foi aplicada uma diferenciação de ordem $D = 12$. Com a série diferenciada, fez-se a análise dos gráficos da FAC e FACP, podendo assim montar possíveis modelos para o ajuste dos dados. Testes com os resíduos indicaram que entre os modelos candidatos o modelo $ARIMA(1,0,2) \times (0,1,2)_{12}$ é o que melhor se ajusta aos dados da série em estudo.

A (figura 1 (b)) mostra o correlograma dos resíduos provenientes da modelagem. Não considerando a autocorrelação no lag 15, pode-se afirmar pelo teste da autocorrelação residual que o modelo ajustado é adequado. Visto que o correlograma para a série de resíduos mostra valores situados num intervalo no qual a autocorrelação é considerada nula.

Os testes de Box-Pierce e periodograma (figura 2) corroboram com a conclusão obtida com o teste da autocorrelação residual. Como o p-valor (0,203) obtido com o teste de Box-Pierce é maior que 0,05, não se pode rejeitar, a hipótese que a série de resíduos é randômica ao nível de 5%. O periodograma (figura 2) reforça os testes anteriores, confirmando que o modelo $ARIMA(1,0,2) \times (0,1,2)_{12}$ é adequado aos dados, não tendo nos resíduos parte sistemática a ser modelada, estando os resíduos próximos ao ruído branco.

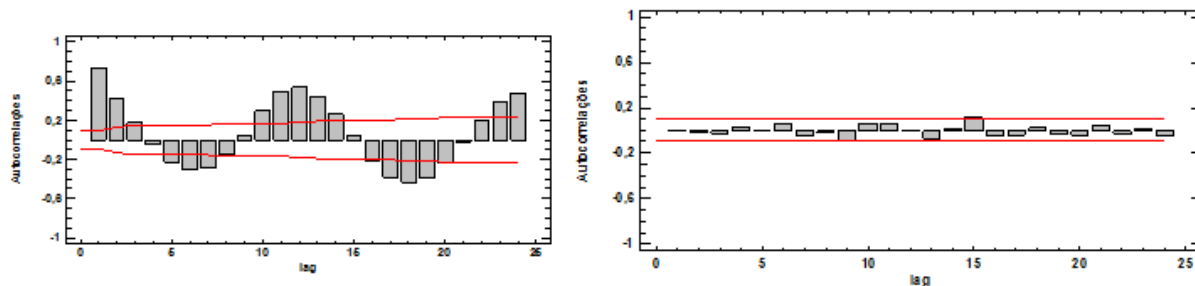


FIGURA 1 – (a) Correlograma das vazões mensais. (b) Autocorrelações residuais modelo $ARIMA(1,0,2) \times (0,1,2)_{120}$.
 Fonte: autor.

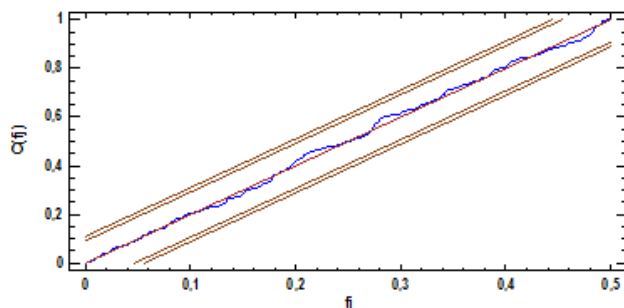


FIGURA 2 - Periodograma para os resíduos.

A seguir, para efeitos de ilustração, são apresentados na tabela 1 os resultados completos para as cinco primeiras observações da amostra de teste. Nesta tabela estão

exemplificadas algumas previsões obtidas pelos métodos SVR, ARIMA e RNA, sendo a última coluna (CL) calculada a partir da combinação linear destas previsões.

TABELA 1 - Previsões um passo à frente para as cinco primeiras observações da amostra de teste

Mês	Amostra de teste	Previsão por			
		SVR	ARIMA	RNA	CL
Dez/2006	15827,00	14578	11293,49	11375,53	16430,53
Jan/2007	26904,64	21053	18753,05	17227,14	23725,08
Fev/2007	26906,40	21054	24032,33	19288,41	23660,76
Mar/2007	17401,18	15658	20837,19	15190,94	17614,23
Abr/2007	11746,38	11291	14935,68	10956,25	12696,26
⋮	⋮	⋮	⋮	⋮	⋮

Fonte: autoria própria.

Utilizando (7), chegou-se aos coeficientes α_{SVR} , α_{ARIMA} e α_{RNA} , cujos valores arredondados são 1,220; 0,056 e -0,175, respectivamente. De forma que a última coluna da tabela 1 foi determinada pela equação (6).

Nas figuras 3(a), 3(b) e 3(c) estão os gráficos de dispersão da série observada (amostra de teste) versus previsão por SVR, RNA e ARIMA, respectivamente. Onde se observa uma maior correlação entre a série observada e previsão por SVR, indicativo de uma melhor qualidade nas previsões.

A tabela 2 informa a RMSE e MAE correspondentes às previsões da amostra de teste. Onde se pode observar uma superioridade do método SVR em relação ao ARIMA e RNA na previsão da série de vazões mensais no posto 266-Itaipu.

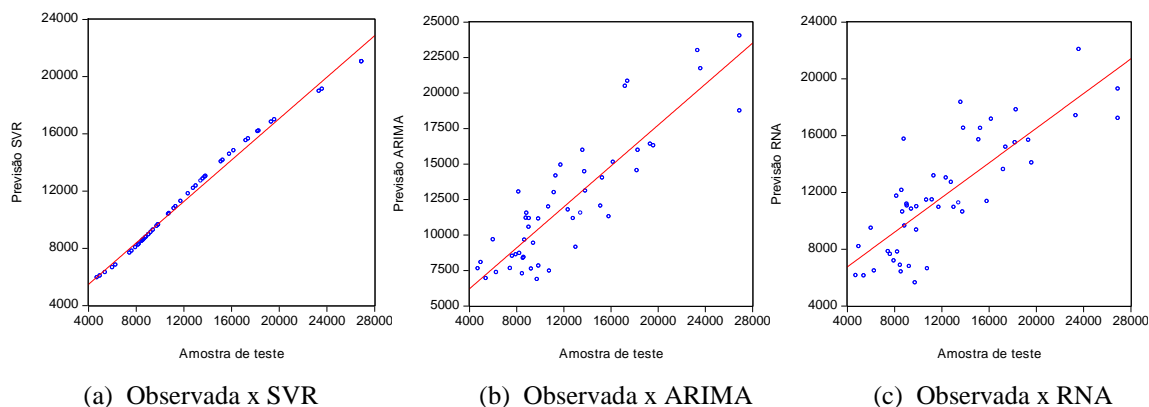


FIGURA 3 - Diagramas de dispersão entre a amostra de teste e previsões por SVR, ARIMA e RNA.

Na tabela 2, os menores valores de RMSE e MAE são devidos à combinação linear (CL) das previsões oriundas dos métodos preditivos SVR, ARIMA e RNA, mostrando a eficiência deste procedimento. Na figura 4 são mostrados os gráficos da amostra de teste (série observada) e as respectivas previsões por SVR e CL.

TABELA 2 - MAE e RMSE: métodos CL, SVR, ARIMA e RNA.

Método	MAE	RMSE
CL	907,3	1156,5
SVR	1039,9	1745,5
ARIMA	2138,1	2607,9
RNA	2418,6	3187,9

Fonte: autoria própria.

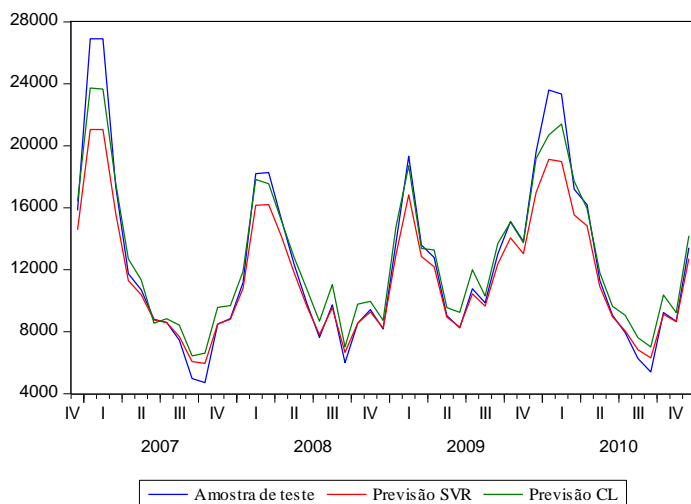


FIGURA 4 - Gráficos da amostra de teste e previsões por SVR e CL.

7 Considerações Finais

Neste artigo foram feitas previsões para a série de vazões mensais (posto 266-Itaipu) através dos métodos ARIMA, RNA e SVR, sendo os resultados combinados linearmente (CL).

Comparando os resultados obtidos através dos três primeiros métodos, observou-se uma supremacia do SVR para a série estudada. Yoon *et al*, (2011) fizeram um estudo comparativo de redes neurais artificiais e máquinas de vetores de suporte para a previsão de níveis de águas e concluíram que para os dados modelados o SVR é mais eficiente que RNA. Okasha, (2014) utilizou ARIMA, RNA e SVR na previsão de séries temporais financeiras e concluiu que SVR fornece um modelo mais preciso e uma técnica de previsão mais eficiente para os dados financeiros usados do que os modelos ARIMA e RNA.

Foi possível melhorar a previsão obtida pelo SVR a partir da combinação linear deste com os métodos RNA e ARIMA, apresentando ganhos preditivos consideráveis. Makridakis e Winkler (1983) propuseram a combinação linear de previsões oriundas de K métodos e verificaram que as previsões combinadas foram mais acuradas que as dos previsores individuais.

Resultados encontrados na literatura especializada e obtidos neste artigo apontam para a eficiência do SVR em previsões de séries temporais e ressaltam os ganhos preditivos que

podem ser alcançados com o uso da combinação linear de previsões provenientes de métodos diferentes.

Referências

- ABREU, T. *et al.* Metodologia híbrida utilizando os modelos ARIMA e redes neurais artificiais para previsão de cargas elétricas. *Anais. Campina Grande: XIX Congresso Brasileiro de Automática*, 2012.
- BALLINI, R. *Análise e Previsão de Vazões Utilizando Modelos de Séries Temporais, Redes Neurais e Redes Nebulosas*. 2000. 169p. Tese (Doutorado em Engenharia Elétrica e Computação) – UNICAMP, Campinas, 2000.
- BATES, J. M. and GRANGER, C. W. J. The Combining of Forecasts. *Operational Research Quarterly*, v.20, n.4, p. 451-468, 1969.
- BOX, G.; JENKINS, G. *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, 1970.
- BOX, G.; JENKINS, G.; REINSEL, G. C. *Time Series Analysis, Forecasting and Control*. 3.ed. Englewood Cliffs:Prentice-Hall, 1994. 598p.
- CASTRO, L. N. *Computação natural – uma jornada ilustrada*. São Paulo: Livraria da Física, 2010.
- CHANG, C.; LIN, C. J. LIBSVM: a library for support vector machines. Disponível em: <<http://www.csie.ntu.edu.tw/~cjlin/libsvm>>. Acesso em: 5 Janeiro 2014.
- FARIA, A. E.; MUBWANDARIKWA, E. Multimodality on the Geometric Combination of Bayesian Forecasting Models. *International Journal of Statistics and Management System*, p. 1-25, 2008.
- FERREIRA, C. C. *Previsão de vazões naturais diárias afluentes ao reservatório UHE Tucuruí utilizada a técnica de redes neurais artificiais*. Goiânia, 117 p. Dissertação (Mestrado Engenharia da Computação) – Universidade Federal de Goiás, Goiânia, 2012.
- GRANGER, C. W. J.; RAMANATHAN, R. Improved methods of combining forecasts. *Journal of Forecasting*, p. 197–204, 1984.
- GUJARATI, D. N. *Econometria básica*. São Paulo: Pearson Makron Books, 2000.
- GUPTA, S.; WILTON, P. C. Combination of forecasts: An extension. *Management Science*, p. 356–372, 1977.
- HAMILTON, J. *Time series analysis*. Princeton University Press, 1994.
- HAYKIN, S., *Redes Neurais Princípios e Aplicações*, 2. Ed. Porto Alegre: Bookman, 2001.
- HE, Z. *et al.* A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. *Journal Hydrology*, v. 509, p. 379-386, 2014.
- HOLLAUER, G., *et al.* Prevendo o crescimento da produção industrial usando um número limitado de combinações de previsões. *Economia Aplicada*, v. 12, n. 2, p. 177-198, 2008.
- LAFDANI, E. K. *et al.* Daily suspended sediment load prediction using artificial neural networks and support vector machines. *Journal Hydrology*, v.478, p. 50-62, 2013.
- MAKRIDAKIS, S. and WINKLER, R. L. Averages of Forecasts: Some Empirical Results. *Management Science*, v.29, n.9, p.987-996, 1983.
- MORETTIN, P. A.; TOLOI, C. M. *Análise de series temporais*. São Paulo: Blucher, 2006.
- OKASHA, M. K. Using Support Vector Machines in Financial Time Series Forecasting. *International Journal of Statistics and Applications*, P. 28-39, 2014.

- SMOLA, A. J.; SCHÖLKOPF, B. *A Tutorial on Support Vector Regression*. NeuroCOLT Technical Report NC-TR-98-030, 1998.
- TEIXEIRA Jr., L. A. *et al.* Redes neurais artificiais *wavelet* causal na previsão da radiação solar direta. Rio de Janeiro: Simpósio Brasileiro de Pesquisa Operacional 2012.
- TEIXEIRA Jr., L. A. *Combinação SSA-Wavelet de métodos preditivos com ajuste numérico MINIMAX, na geração de previsões e de cenários*. Rio de Janeiro, 114 p. Tese (Doutorado em Engenharia Elétrica). Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2013.
- TERUI, N. & DIJK, H. K. Combined forecasts from linear and nonlinear time series models. *International Journal of Forecasting*, p. 421–438, 2002.
- VAPNIK, V. N. *The Nature of Statistical Learning Theory*. New York: Springer - Verlag, 199
- VALENZUELA, O. Hybridization of intelligent techniques and ARIMA models for time series prediction. *Fuzzy Sets and Systems*, v.159, p. 821-845, 2008.
- WALLIS, K. F. Combining Forecasts - Forty Years Later. *Applied Financial Economics*, p. 33-41, 2011.
- WANG, A.B.W. *et al.* Forecasting daily streamflow using hybrid ANN models. Amsterdam: *Journal of Hydrology*, v. 324, p. 383-399, 2006a.
- WANG, A.B.W. *et al.* Testing for nonlinearity of streamflow processes at different time scales. Amsterdam: *Journal of Hydrology*, v. 322, p. 247-268, 2006b.
- YOON, H, *et al.* A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal Hydrology*, v.396, p. 128-138, 2011.