

SÉRIES TEMPORAIS: COMBINAÇÃO LINEAR DE PREVISÕES WAVELET- NEURAL

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

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

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

Resumo: *NESTE ARTIGO FEZ-SE A PREVISÃO DE SEIS SÉRIES CONHECIDAS NA LITERATURA E COMPAROU COM RESULTADOS OBTIDOS EM TRABALHOS QUE USARAM AS MESMAS SÉRIES. PARA A PREVISÃO FOI USADO UMA COMPOSIÇÃO DE REDES NEURAS ARTIFICIAIS (RNA), DECOMPOSIÇÃO WAVELETS (DW) E COMBINAÇÃO LINEAR DE PREVISÕES (CL). OBTENDO COM ESTE MÉTODO PREVISÕES MELHORES AOS OBTIDOS POR MODELOS AUTORREGRESSIVOS INTEGRADO DE MÉDIAS MÓVEIS (ARIMA) E RNA, QUANDO APLICADOS ISOLADAMENTE. O USO CONJUGADO DW-RNA-CLP SE APRESENTA COMO UM MÉTODO BASTANTE PROMISSOR NO CAMPO DAS PREVISÕES DE SÉRIES TEMPORAIS.*

Palavras-chaves: *DECOMPOSIÇÃO WAVELET; COMBINAÇÃO LINEAR DE PREVISÕES; REDES NEURAS ARTIFICIAIS; PREVISÃO DE SÉRIES TEMPORAIS.*

Área: 6 - PESQUISA OPERACIONAL

Sub-Área: 6.3 - PROCESSOS ESTOCÁSTICOS

TIME SERIES: LINEAR COMBINATION OF FORECASTS WAVELET-NEURAL

Abstract: *THIS PAPER IS THE PREDICTION MADE SIX SERIES KNOWN IN THE LITERATURE AND COMPARED WITH RESULTS OBTAINED IN STUDIES THAT USED THE SAME SERIES. FOR THE FORECAST WE USED A COMPOSITION OF ARTIFICIAL NEURAL NETWORKS (ANN), WAVELET DECOMPOSITION (WD) AND LINEAR COMBINATION OF FORECASTS (LC). GETTING BETTER PREDICTIONS THAN THE AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODELS (ARIMA) AND ANN WHEN USED SEPARATELY. THE COMBINED USE OF WD-ANN-LC IS PRESENTED AS A PROMISING METHOD IN THE FIELD OF TIME SERIES FORECASTS.*

Keyword: *WAVELET DECOMPOSITION; LINEAR COMBINATION OF FORECASTS; ARTIFICIAL NEURAL NETWORKS; FORECASTING TIME SERIES.*

1 Introdução

Desde 1927 quando George Udny Yule lançou a ideia que série temporal pode ser vista como sendo gerada por um conjunto de variáveis aleatórias conjuntamente distribuídas e aplicou modelos autorregressivos (AR), muitos foram os avanços na tentativa de melhorar a acurácia na modelagem de séries temporais.

Box e Jenkins (1970) apresentam metodologia que faz 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.

Além da metodologia de Box e Jenkins é frequente a utilização de Redes Neurais Artificiais (RNA). Wang *et al.*,(2006), Valenzuela *et al.*, (2008); Abreu, *et al.*, (2012) e Teixeira, Jr *et al.*; (2012), entre outros, utilizaram Redes Neurais Artificiais (RNA) na previsão de séries temporais.

Revisando 25 anos (1985-2005) de pesquisa sobre previsão de séries temporais, Gooijer e Hyndman (2006) concluíram que importantes progressos foram feitos nesta área. Salientam que encontraram numerosos estudos que atestam o sucesso das redes neurais artificiais (RNA) na previsão de séries temporais e apontam temas que necessitam de um maior desenvolvimento. Entre eles, a combinação de previsões. Ressaltam que ainda não há uma resposta para a melhor estratégia de combinação de métodos lineares e não lineares.

Zou e Yang (2004); Deutsch *et al.*, (1994); Martins (2011); Fiordaliso (1998); Fang (2003); Cavalieri e Ribeiro (2011); Bunn (1988) são autores que desenvolveram trabalhos onde comprovam empiricamente a eficácia da combinação de previsões. Em sua pesquisa clássica de combinações de previsão, Clemen (1989) concluiu que a combinação de previsões tem se mostrado um método prático e útil.

Outra técnica que vem sendo usada com frequência nas previsões de séries temporais é a wavelet. Embora os primeiros estudos de no campo das wavelets tenham ocorridos em 1909, (HAAR, 1910). Só recentemente em, 1985, Stephane Mallat deu às wavelets um grande impulso através do seu trabalho em processamento digital de imagens, construindo a primeira wavelet não-trivial (suave) (LIMA, 2002). Daubechies (1988) construiu um conjunto de bases ortonormais de wavelet suaves, passando a ser uma das principais referências para as atuais aplicações de wavelet.

Lima (2011) utilizou wavelet no seu trabalho de previsão de séries econômicas, separando a parte determinística dos ruídos. Teixeira Jr *et al.*, (2012) optou por um método de combinação em que as componentes wavelets de uma série temporal constituem os padrões de entrada de uma RNA *feedforward* MLP, cuja saída fornece a previsão da série temporal. Karthikeyan e Kumar (2013); Tiwari e Chatterjee (2010); Kisi (2010); Kisi e Cimen (2011); Nalley *et al.*, (2012) também utilizaram wavelet em seus trabalhos de previsão de séries temporais.

MacDonald e Xu (1993) investigaram o desempenho de previsão relativa de alguns estimadores parcialmente adaptativos de modelos ARIMA baseados em distribuições de erro paramétricos flexíveis que permitem a curtose e assimetria. Estes métodos foram utilizados para obter previsões de alguns conjuntos de dados bem conhecidos, as séries de Box e Jenkins A,B,...,F.

Hansen *et al.*, (1999) utilizaram algoritmo genético como auxiliador na determinação da arquitetura da RNA usada na modelagem das séries (A,B,...,F) de Box e Jenkins (1970), e fizeram um comparativo com os resultados obtidos por MacDonald e Xu (1993).

Gnanlet e Rajendran (2009) desenvolveram um sistema para ajuste de modelos ARMA (BOX e JENKINS, 1970) baseado nas meta-heurísticas Algoritmo Genético e *Simulated Annealing*. Para validar o modelo proposto, Gnanlet e Rajendran (2009) fizeram um comparativo com os resultados obtidos por MacDonald e Xu (1993) e Hansen *et al.*, (1999).

A proposta deste artigo é fazer a previsão para as séries A,B,...,F de Box e Jenkins (1970) e comparar com os resultados alcançados por MacDonald e Xu (1993), Hansen *et al.*, (1999) e Gnanlet e Rajendran (2009), doravante designados por [1], [2] e [3], respectivamente. As previsões serão obtidas a partir de decomposição wavelet, redes neurais e combinação linear das previsões.

Em um primeiro momento as séries serão decompostas ortogonalmente via wavelet. Cada uma das componentes Wavelet será modelada a partir da RNA. As previsões obtidas com as modelagens serão combinadas a fim de se obter a previsão da série observada.

2 Decomposição Ortogonal Wavelet

A wavelet é uma forma de onda que ao contrário das senóides podem apresentar (dependendo da família) forma irregular e assimétrica. Enquanto na análise de Fourier o sinal é decomposto em ondas senoidais de várias frequências, na análise wavelet o sinal é

decomposto a partir de informações de escala (compressão e expansão) e translação de uma função wavelet.

2.1 Função Wavelet

Funções wavelets são elementos do espaço de Hilbert l^2 formado por funções quadraticamente somáveis. A função $\omega(\cdot)$ pertencente a l^2 é chamada de função wavelet se, e somente se, as funções

$$\omega_{m,n}(\cdot) = 2^{\frac{m}{2}} \omega(2^m(\cdot) - n) \quad (1)$$

onde $m, n \in \mathbb{Z}$, formam uma base ortonormal para o espaço de Hilbert l^2 . Segundo Levan e Kubrusly (2003), o parâmetro m é chamado de parâmetro de escala e n , de parâmetro de translação. Desta forma, a função $f(\cdot) \in l^2$ admite sua expansão ortogonal por meio de uma série de Fourier, em termos de uma base ortonormal wavelet $\{\omega_{m,n}(\cdot)\}_{(m,n) \in \mathbb{Z} \times \mathbb{Z}}$ para l^2 , definida como em (2).

$$f(\cdot) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f(\cdot), \omega_{m,n}(\cdot) \rangle \omega_{m,n}(\cdot) \quad (2)$$

De acordo com Kubrusly e Levan (2006), o subespaço fechado (3)

$$W_m(\omega) := \left(\text{span}\{\omega_{m,n}(\cdot)\}_{n \in \mathbb{Z}} \right)^- \quad (3)$$

pertencente a l^2 é chamado de subespaço de detalhes, na escala m . Segundo Levan e Kubrusly (2003), a projeção ortogonal de $f(\cdot)$, sobre o subespaço de detalhes $W_m(\omega)$, é definida pela soma parcial definida em (4).

$$f_{W_m(\omega)}(\cdot) := \sum_{n \in \mathbb{Z}} \langle f(\cdot), \omega_{m,n}(\cdot) \rangle \omega_{m,n}(\cdot) \quad (4)$$

Segundo Daubichies (1988), a projeção $f_{W_m(\omega)}(\cdot)$ pode ser referida como uma componente de detalhe wavelet de $f(\cdot)$, na escala m , sobre $W_m(\omega)$.

A função $\phi(\cdot) \in l^2$ é chamada de função *escala* se, e somente se, as funções

$$\phi_{m,n}(\cdot) = 2^{\frac{m}{2}} \phi(2^m(\cdot) - n) \quad (5)$$

onde $m, n \in \mathbb{Z}$, são tais que

$$\langle \phi_{l,i}(\cdot), \phi_{j,k}(\cdot) \rangle = 0 \quad (6)$$

sempre que $l = j$ e $i \neq k$, e

$$\langle \phi_{l,i}(\cdot), \phi_{j,k}(\cdot) \rangle \neq 0 \quad (7)$$

no caso contrário.

De acordo com Mallat (1998), o subespaço fechado

$$V_m(\Phi) := \left(\text{span}\{\phi_{m,n}(\cdot)\}_{n \in \mathbb{Z}} \right)^- \quad (8)$$

de l^2 é chamado de subespaço de aproximação, na escala m . Segundo Levan e Kubrusly (2003), a projeção ortogonal de $f(\cdot)$ sobre $V_m(\Phi)$ é definida pela soma parcial definida em (9).

$$f_{V_m(\Phi)}(\cdot) := \sum_{n \in \mathbb{Z}} \langle f(\cdot), \phi_{m,n}(\cdot) \rangle \phi_{m,n}(\cdot) \quad (9)$$

De acordo com Daubichies (1988), a projeção $f_{V_m(\Phi)}(\cdot)$ pode ser referida como uma componente de aproximação wavelet de $f(\cdot)$, na escala m , sobre $(V_m(\Phi), \langle ; \rangle)$.

2.2 Transformada Wavelet

Uma transformada wavelet, sobre l^2 , é definida por um produto interno $\langle ; \rangle$ entre uma função $f(\cdot) \in l^2$ e uma função wavelet $\omega_{m,n}(\cdot)$ ou uma função escala $\phi_{m,n}(\cdot)$, onde $(m, n) \in \mathbb{Z} \times \mathbb{Z}$. De acordo com Morettin (2012), as transformadas wavelet podem ser agrupadas em dois conjuntos disjuntos: o dos coeficientes de detalhe, denotado por $\{d_{m,n}\}_{(m,n) \in \mathbb{Z} \times \mathbb{Z}}$, e o dos coeficientes de aproximação, denotado por $\{a_{m,n}\}_{(m,n) \in \mathbb{Z} \times \mathbb{Z}}$, onde os coeficientes de detalhe $d_{m,n}$ e os de aproximação $a_{m,n}$ são definidos, respectivamente, em (10) e (11).

$$d_{m,n} := \langle f(\cdot), \omega_{m,n}(\cdot) \rangle = \sum_{t \in \mathbb{Z}} f(t) \omega_{m,n}(t) \quad (10)$$

$$a_{m,n} := \langle f(\cdot), \phi_{m,n}(\cdot) \rangle = \sum_{t \in \mathbb{Z}} f(t) \phi_{m,n}(t) \quad (11)$$

Finalmente, a expansão de $f(\cdot)$ sobre o espaço de Hilbert l^2 , em termos de uma base ortonormal wavelet, é definida por:

$$f(\cdot) = \sum_{n \in \mathbb{Z}} a_{m_0,n} \phi_{m_0,n}(\cdot) + \sum_{m=m_0}^{+\infty} \sum_{n \in \mathbb{Z}} d_{m,n} \omega_{m,n}(\cdot). \quad (12)$$

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 (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 Combinação de Previsões

Os precursores na área das combinações de previsões são Bates e Granger (1969), que sugeriram uma combinação, expressa em (13), para a previsão da série temporal $\{y_t\}_{t=1}^T$.

$$\hat{y}_t = \rho_1 \times \hat{y}_{1,t} + \rho_2 \times \hat{y}_{2,t} \quad (13)$$

Em (13), 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 previsão.

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

Como mencionado anteriormente, as series manipuladas neste trabalho são as conhecidas séries A,B,...,F de Box e Jenkins (1970). Na tabela 1, estão registrados o número de observações de cada série, o tipo da série e as ordens do modelo ARIMA(p,d,q) sugerido por Box e Jenkins (1970). Sendo p e q as ordens dos modelos autorregressivos e médias móveis, respectivamente. Na hipótese da série temporal ser não estacionária, esta deve ser diferenciada, sendo d a ordem de diferenciação da série.

Tabela 1. Conjunto de dados: séries de Box e Jenkins (1970)

Série	Cardinalidade	Tipo	ARIMA(p,d,q)
A	197	Concentração em processos químicos	ARIMA(0,1,1)
B	369	Preços das ações da IBM (1961)	ARIMA(0,1,1)(0,0,1) ₆
C	226	Temperatura de processos químicos	ARIMA(0,2,0)(0,0,1) ₁₇
D	310	Viscosidade em processos químicos	ARIMA(0,1,1)
E	100	Manchas solares	ARIMA(2,0,0)(0,0,1) ₁₁
F	70	Rendimentos em processos químicos	ARIMA(2,0,0)

Fonte: Gnanlet e Rajendran (2009) – adaptado.

As séries mencionadas na tabela 1 podem ser encontradas no site <http://www.stat.wisc.edu/~reinsel/bjr-data/>.

Para a decomposição wavelet foi utilizado bases da família Daubechies, mas especificamente a db8. Realizada em nível $p=2$, que resulta em duas componentes de detalhes (D_1 e D_2) e uma de aproximação (A_2). Isto foi feito para cada uma das séries $\{y_t^S\}_{t=1}^T$, onde T é a cardinalidade da série e S = série A,B,....,F.

Uma vez obtidas as componentes wavelets (A_2^S, D_1^S e D_2^S), estas foram modeladas via rede neural artificial *feedforward multilayer perceptron* (RNA-MLP). Testes indicaram que, de maneira geral, apresentaria resultados satisfatórios a RNA-MLP com a seguinte estrutura topológica: tamanho de janela de entrada igual a 2, uma (1) camada escondida composta por 5 neurônios artificiais com função de ativação *tansig* e um (1) neurônio na camada de saída com função de ativação *purelin* (HAYKIN, 2001). O algoritmo de treinamento utilizado foi o *Levenberg & Marquardt*. A decomposição wavelet e treinamento da rede foram realizados no aplicativo computacional Matlab (8.0).

Da modelagem por RNA resultaram as previsões \hat{A}_2^S, \hat{D}_1^S e \hat{D}_2^S das componentes wavelets obtidas da decomposição de S . No passo seguinte combinam-se as previsões \hat{A}_2^S, \hat{D}_1^S e \hat{D}_2^S a fim de se obter a previsão da série S ($\{\hat{y}_t^S\}_{t=1}^T$), conforme (14):

$$\{\hat{y}_t^S\}_{t=1}^T = \alpha_{A_2,S} \times \hat{A}_2^S + \alpha_{D_1,S} \times \hat{D}_1^S + \alpha_{D_2,S} \times \hat{D}_2^S \quad (14)$$

onde $\alpha_{A_2,S}, \alpha_{D_1,S}$ e $\alpha_{D_2,S}$ são parâmetros a serem determinados com a otimização do problema de programação não linear (15), cuja função objetivo é a raiz do erro quadrático médio (RMSE).

$$\begin{aligned} \text{Min } RMSE &= \sqrt{\frac{1}{T} \sum_{t=1}^T (\{y_t^S\}_{t=1}^T - \{\hat{y}_t^S\}_{t=1}^T)^2} \\ \text{s. a. } & \alpha_{A_2,S}, \alpha_{D_1,S}, \alpha_{D_2,S} \text{ são irrestritas} \end{aligned} \quad (15)$$

Para a otimização do problema de programação não linear (15) foi usado o aplicativo computacional LINGO 11.

6 Resultados Obtidos

Para que fosse possível fazer um comparativo com os resultados obtidos por [1], [2] e [3], reservou-se para amostras de testes as dez últimas observações dos totais registrados na tabela 1 das séries A, B, C e D e as cinco últimas para as séries E e F. A soma dos erros ao quadrado (SSE) e a soma dos desvios absolutos (SAD) foram as medidas de acurácia usadas por [1], [2] e [3]. Neste artigo, optou-se por fazer o comparativo segundo a medida SAD, sendo as previsões feitas a um passo à frente.

A seguir, para efeitos de ilustração, são apresentados os resultados completos para a série F. Na figura 1 pode-se observar o gráfico da série F, as componentes wavelets de

aproximação (A2) e detalhes (D1 e D2). A componente de aproximação mapeia as baixas frequências da série, enquanto as componentes de detalhes faz o mapeamento das Altas frequências.

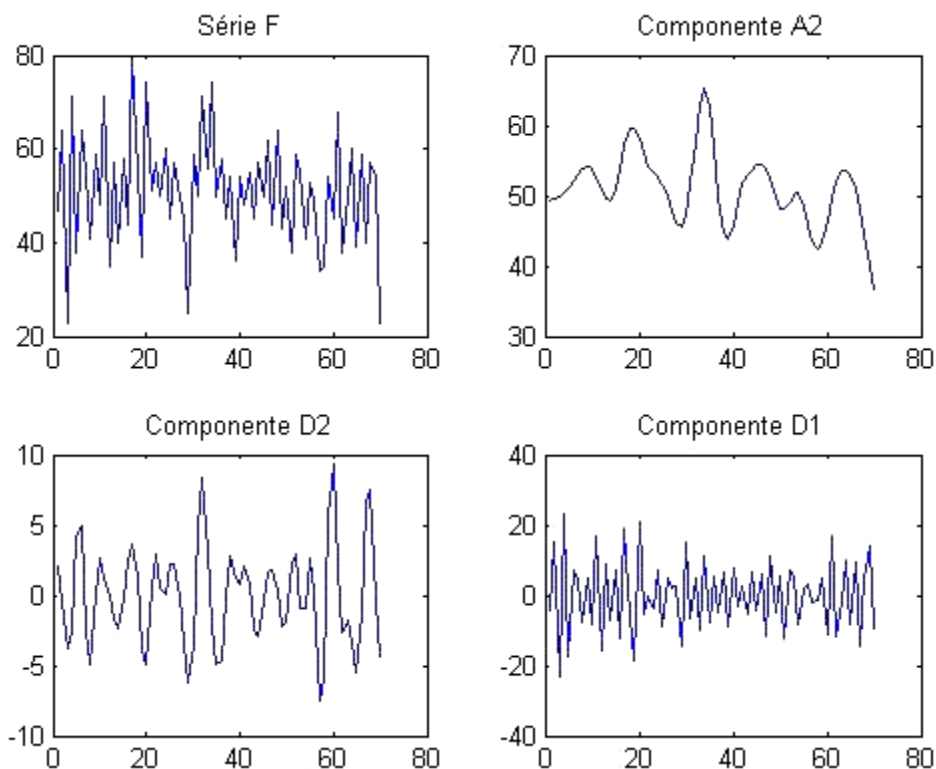


Figura 1. Gráficos da série F e componentes wavelets.

A tabela 2 informa os valores da amostra de teste para a série F, correspondentes aos últimos cinco valores da série. Além, das componentes wavelets A_2^F, D_1^F e D_2^F e as respectivas previsões \hat{A}_2^F, \hat{D}_1^F e \hat{D}_2^F . Na última coluna (\hat{y}_t^F) da tabela 2 estão registradas as combinações lineares de \hat{A}_2^F, \hat{D}_1^F e \hat{D}_2^F , em outras palavras a previsão de y_t^F .

Tabela 2. Previsões um passo à frente para a amostra de teste da série F.

t	Série: y_t^F	A_2^F	D_1^F	D_2^F	\hat{A}_2^F	\hat{D}_1^F	\hat{D}_2^F	\hat{y}_t^F
6	59	50,9872	9,41678	-	51,7869	7,49149	-	60,0618
6		8	1	1,40406	3		2,08238	1
6	40	48,1946	-	6,58659	49,2011	-	3,50231	38,1084
7		8	14,7813		4	7,07352	6	4
6	57	44,4814	5,02339	7,49519	46,3865	15,2392	9,00528	57,9873
8		2	2	2	9	7	4	6
6	54	40,2131	14,1358	-	44,0596	6,67656	-3,5931	52,3711
9		4		0,34894	2	2		3
7	23	36,7735	-	-	43,1285	-	-	24,4849
0		0	9,37195	4,40155	7	17,1763	3,97839	7

Fonte: autoria própria.

Utilizando (15), chegou-se aos coeficientes $\alpha_{A2,S} = 0,974533$, $\alpha_{D1,S} = 1,139267$ e $\alpha_{D2,S} = -0,508533$. De forma que a última coluna da tabela 2 foi calculada a partir da equação (14).

A soma dos desvios absolutos (SAD) foi obtida através de (16)

$$SAD = \sum_{t=66}^{70} |y_t^F - \hat{y}_t^F| \quad (16)$$

resultando no valor 7,054. Reduzido em aproximadamente 80% o valor do SAD em relação a melhor previsão obtida por [1], [2] e [3].

MacDonald e Xu (1993) [1] apresentaram resultados obtidos a partir dos modelos “ARIMA ML”, “OLS”, “LAD”, “GT” e “EGB2”, todos fundamentados nos modelos ARIMA. Hansen *et al.*, (1999) [2] fizeram previsões para as séries A,B,...F utilizando os modelos “Heuristic NN” e “GA NN”, constituídos por RNA e algoritmo genético (AG). Gnanlet e Rajendran (2009) [3] usaram o sistema “GA-ARMA/BNLS”, formado pelas meta-heurísticas AG e *Simulated Annealing*, para determinar as ordens e parâmetros do modelo ARMA que melhor se ajustasse aos dados das séries estudadas. Os modelos citados neste parágrafo estão na tabela 3, juntamente com a medida SAD e as respectivas séries A,B,...F. Nesta tabela, a primeira linha é formada pelos resultados obtidos neste artigo, com a denominação “Wavelet-Neural”, as linhas restantes foram obtidas em Gnanlet e Rajendran (2009).

Tabela 3. Soma dos desvios absolutos (SAD)

Método	Séries					
	A	B	C	D	E	F
Wavelet-Neural	1,273	32,845	0,498	0,667	28,732	7,054
GA-ARMA/BNLS	3,125	65,585	1,340	2,212	64,927	35,850
Heuristic NN	4,519	88,312	9,138	2,942	98,873	43,966
GA NN	3,705	72,398	6,684	2,952	69,536	36,400
ARIMA ML	4,005	78,858	11,247	3,114	116,485	49,161
OLS	3,937	83,170	10,740	3,080	114,800	45,500
LAD	3,960	79,470	10,300	3,066	117,600	44,460
GT	3,937	80,680	10,250	3,064	106,500	44,590
EGB2	4,017	81,010	10,300	3,066	111,800	44,500

Fonte: Gnanlet e Rajendran (2009) – Adaptado.

Da tabela 3, tira-se que o método usado neste artigo (wavelet-neural) possibilitou uma redução na medida SAD de 60% em média em relação ao método “GA-ARMA/BNLS”. Sendo a menor redução de 49,9% na série B e a maior de 80,3% na série F.

7 Considerações Finais

No presente trabalho foram feitas previsões para algumas séries de Box e Jenkins (1970), através de um método que englobava decomposição wavelet, RNA-MLT e combinação de previsões, denominado neste artigo por “wavelet-neural”. Comparações com resultados de previsões obtidos através de modelos ARIMA, mostraram a eficácia e superioridade da metodologia “wavelet-neural” para as séries em estudo. Kisi, (2010), combinou decomposição wavelet e regressão linear (WR) e comparou os seus resultados aos

obtidos com modelagem ARIMA. Constatou que o primeiro método (WR) apresentou melhores resultados de previsão que o segundo.

Na comparação com métodos envolvendo RNA e meta-heurísticas, novamente, a metodologia “wavelet-neural” apresentou ganhos preditivos consideráveis. Teixeira Jr *et al.*, (2012), apresentaram um método em que as componentes *wavelets* de uma série temporal constituem os padrões de entrada de uma RNA *feedforward* MLP, cuja saída fornece a previsão da série temporal. Com esta metodologia, chegaram a resultados melhores aos obtidos por RNA convencional (sem considerar componentes wavelet como padrões de entrada).

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

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.
- 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*. San Francisco: Holden-Day, 1970.
- BUNN, D. W. Combining forecasts. *European Journal of Operational search*, p. 223-229, 1988.
- CASTRO, L. N. *Computação Natural*. São Paulo: Editora Livraria da Física, 2010.
- CAVALIERI, R; RIBEIRO, E. P. Combinação de previsões de volatilidade: um estudo. *Revista Economia*, 2011.
- CLEMEN, R. T. Combining forecasts: a review and annotated bibliography. *International Journal of Forecasting*, p. 559-583, 1989.
- DAUBECHIES, I. Orthonormal basis of compactly supported wavelets. *Comm. Appl. Math.*, p. 909-996, 1988.
- DEUTSCH, M. *et al.* The combination of forecasts using changing weights. *International Journal of Forecasting*, p. 47-57, 1994.

- FANG, Y. Forecasting combination and encompassing tests. *International Journal of Forecasting*, p. 87–94, 2003.
- FARIA, A. E.; MUBWANDARIKWA, E. (2008). Multimodality on the Geometric Combination of Bayesian Forecasting Models. *International Journal of Statistics and Management System*, p. 1-25, 2008.
- FIORDALISO, A. A nonlinear forecasts combination method based on Takagi–Sugeno fuzzy systems. *International Journal of Forecasting*, p. 367–379, 1998.
- GNANLET, A. e RAJENDRAN, C. Meta-Heuristics in ARMA Forecasting. *California Journal of Operations Management*. v.7, n.1, p.38-48, feb.2009.
- GOOIJER, J. G.; HYNDMAN, R. J. 25 years of time series forecasting. *International Journal of Forecasting*, p.443– 473, 2006.
- GRANGER, C. W. J. Invited Review: Combining Forecasts - Twenty Years Later. *Journal of Forecasting*, p. 167-173, 1989.
- GRANGER, C. W. J.; RAMANATHAN, R. Improved methods of combining forecasts. *Journal of Forecasting*, p. 197–204, 1984.
- GUPTA, S.; WILTON, P. C. Combination of forecasts: An extension. *Management Science*, p. 356–372, 1977.
- HAAR, A. Zur theorie der orthogonalen funktionen-systeme. *Math. Ann*, p.331-371, 1910.
- HANSEN, J. V.; McDONALD, J.B. E NELSON, R.D. Time Series Prediction With Genetic Algorithm Designed Neural Networks: An Empirical Comparison With Modern Statistical Models. *Computational Intelligence*. V. 15, n.3, p.171-184, aug.1999.
- HAYKIN, S., *Redes Neurais Princípios e Aplicações*, 2. Ed. Porto Alegre: Bookman, 2001.
- 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.
- KISI, O. Wavelet regression model for short-term streamflow forecasting. *Journal of hidrology*, p. 344-353, 2010.
- KISI, O.; CIMEN, M. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *Journal of hidrology*, p. 132-140, 2011.
- LEVAN, N. ; KUBRUSLY, C. S. A Wavelet Time-Shift-Detail Decomposition. *Mathematics and Computers in Simulation*. p. 73-78, 2003
- LIMA, P. C. Wavelets: uma introdução. *Matemática universitária*, n. 33, p. 13-44, 2002.
- KARTHIKEYAN, L; KUMAR, D. N. Predictability of nonstationary time series using wavelet and EMD based ARMA models. *Journal of hydrology*, p. 103-119, 2013.

- KUBRUSLY, C. S. AND LEVAN, N. Abstract *Wavelets* Generated by Hilbert Space Shift Operators. *Adv. Math. Sci. Appl.* 14, p. 643-660, 2006.
- MALLAT S. *A Wavelet Tour of Signal Processing*. San Diego: Academic Press, 1998.
- MARTINS, V. L. M. *Comparação de combinação de previsões correlacionadas e não correlacionadas com as suas previsões individuais: um estudo com séries industriais*. Porto Alegre, 100 p. Dissertação de Mestrado em Engenharia de Produção. Universidade Federal do Rio Grande do Sul, Porto Alegre, 2011.
- McDONALD, J. B.; XU, Y. Some forecasting applications of partially adaptive estimators of ARIMA models. *Economics letters*, p.155-160, 1994.
- MORETTIN, P.A. *Ondas e Ondaletas*. 2. Ed. São Paulo: EDUSP, 2012.
- NALLEY, D. *et al.* Using discrete wavelet transforms to analyze trends in streamflow and precipitation in Quebec and Ontario (1954–2008). *Journal of hydrology*, p. 204-228, 2012.
- TEIXEIRA Jr., L. A. *et al.* Redes neurais artificiais wavelet causal na previsão da radiação solar direta. *Anais...Rio de Janeiro: Simpósio Brasileiro de Pesquisa Operacional 2012*.
- TEIXEIRA Jr., L. A. *Combinação SSA-Wavelet de métodos preditivos com ajuste numérico MINIMAX, na geração de previsões e de cenários*. Rio de Janeiro, 114 p. Tese (Doutorado em Engenharia Elétrica). Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2013.
- TERUI, N. & DIJK, H. K. Combined forecasts from linear and nonlinear time series models. *International Journal of Forecasting*, p. 421–438, 2002.
- TIWARI, M. K.;CHATTERJEE, C. Development of an accurate and reliable hourly flood forecasting model using wavelet–bootstrap–ANN (WBANN) hybrid approach. *Journal of Hydrology*, p. 458-470, 2010.
- 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, 2006.
- ZOU, H.; YANG, Y. Combining time series models for forecasting. *International Journal of Forecasting*, p. 69– 84, 2004.