

MODELAGEM BOX & JENKINS, DECOMPOSIÇÃO WAVELET E PROGRAMAÇÃO MATEMÁTICA NA GERAÇÃO DE PREVISÕES

Luiz Albino Teixeira Júnior

Departamento de Engenharia Elétrica
Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio
juniorg3@hotmail.com

Moisés Lima de Menezes

Departamento de Estatística – Universidade Federal Fluminense - UFF
moises_lima@msn.com

Keila Mara Cassiano

Departamento de Estatística – Universidade Federal Fluminense - UFF
keilamath@hotmail.com

José Francisco Moreira Pessanha

Instituto de Matemática e Estatística – Universidade do Estado do Rio de Janeiro -
UERJ
professorjfm@hotmail.com

Reinaldo Castro Souza

Departamento de Engenharia Elétrica
Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio
reinaldo@ele.puc-rio.br

Resumo

Este artigo propõe um método que integra a modelagem de Box & Jenkins, a decomposição *wavelet* e a Programação Matemática na geração de previsões pontuais. Inicialmente, fez-se a decomposição *wavelet* da série temporal de interesse, obtendo-se as componentes *wavelets*. Em seguida, cada componente *wavelet* foi modelada através da abordagem Box & Jenkins. A previsão final é obtida por meio da combinação linear das previsões das componentes *wavelets*, na qual os pesos atribuídos às componentes são determinados por Programação Matemática. Para ilustrar a metodologia proposta, utilizou-se a série temporal de afluência da Região Sul do Brasil.

Palavras-chave: Modelagem Box & Jenkins, Decomposição *Wavelet*, Programação Matemática.

Abstract

This paper proposes a method that integrates the Box & Jenkins modeling, wavelet Theory and Mathematical Programming in the generation of point forecasts. Initially, we performed the wavelet decomposition of the time series of interest generating the wavelet components. Then, each one was modeled through Box & Jenkins modeling. Finally, the point forecasts of each wavelet component were linearly combined with the use of Mathematical Programming in order to obtain the forecasts. To illustrate the proposed methodology, we used a series of streamflow

from the Brazilian South Region.

Keywords: Box & Jenkins Modeling, Wavelet Decomposition, Mathematical Programming.

1. INTRODUÇÃO

Os modelos ARIMA foram inicialmente formulados por Box & Jenkins na década de 1970 [2]. Sob a hipótese de ergodicidade [5] procura-se estimar um processo estocástico plausível e seus respectivos momentos, a partir da realização de uma série temporal. De acordo com [2], a metodologia Box & Jenkins tem como base a Teoria Geral de Sistemas Lineares, a qual supõe que a passagem de um ruído branco por um filtro de memória infinita gera um processo estacionário de segunda ordem. Trata-se, portanto, de um modelo pertencente à classe dos modelos estatísticos lineares.

Neste artigo é proposta uma metodologia alternativa, na qual são integradas a Modelagem de Box & Jenkins com a Decomposição *Wavelet* [3, 7, 9, 11] e a Programação Matemática [15]. O desempenho do método proposto é demonstrado por meio de um experimento computacional envolvendo a modelagem da série temporal de afluência da Região Sul do Brasil [14].

O artigo está organizado em sete seções. A seguir, na seção 2, é descrita a Teoria *wavelet*. Na seção 3 tem-se uma breve apresentação da Modelagem Box & Jenkins. Na seção 4 é ilustrada a combinação linear de previsões. O método proposto é descrito na seção 5. O experimento computacional e os principais resultados obtidos são comentados na seção 6. Por fim, na seção 7, são expostas as conclusões do trabalho.

2 TEORIA WAVELET

2.1. Espaço de Hilbert, Base Ortonormal e Série de fourier

De acordo com [8], um espaço de Hilbert H (arbitrário) é qualquer espaço linear produto interno completo. A coleção de todas as sequências infinitas de números complexos quadraticamente somáveis, denotada por l^2 , munida de um produto interno, denotado por (\cdot, \cdot) , isto é, l^2 é um caso particular de espaços de Hilbert.

De acordo com [9], um subconjunto B de um espaço de Hilbert H é uma base ortonormal de H se e somente se satisfaz aos seguintes axiomas: (i) ortogonalidade, $(b_i, b_j) = \delta_{ij}$, sempre que $b_i, b_j \in B$; (ii) normalização: $\|b_i\| = 1$, para todo $b_i \in B$; e (iii) completamento, se e somente se $\overline{\text{span}} B = H$.

Pelo Teorema da série de Fourier [9], um subconjunto B é uma base ortonormal de um espaço de Hilbert H se e somente se existir, para qualquer $x \in H$, uma expansão única definida por:

A expansão do vetor x sobre B , em (1), é chamada de série de Fourier.

2.2. Função Wavelet

Considere o espaço de Hilbert H . Um elemento $f \in H$ com um produto interno (\cdot, \cdot) é chamado de função *wavelet* se e somente se as funções $\psi_{j,k}$, $j, k \in \mathbb{Z}$, formam uma base ortonormal para o espaço de Hilbert H . Com base em [9], uma função ψ em H admite a expansão através da série de Fourier, em termos de base ortonormal wavelet de l^2 , definida em (2).

O parâmetro α é chamado de parâmetro de escala e β , de parâmetro de translação. De acordo com [9], a projeção de f sobre $V_{\alpha, \beta}$ pode ser interpretada como uma variação de detalhes de f , na escala α e translação β . Segundo [10], o subespaço fechado de $V_{\alpha, \beta}$ é chamado de subespaço de detalhes, na escala α . Por sua vez, a projeção de f , sobre o subespaço de detalhes $V_{\alpha, \beta}$, é definida pela soma parcial descrita em (3).

De acordo com [9], a projeção $P_{\alpha, \beta} f$ pode ser referida como uma componente de detalhes de f , na escala α , sobre $V_{\alpha, \beta}$. Como consequência, com base na identidade em (1), segue que f pode ser interpretada como a soma de todas suas componentes de detalhes $P_{\alpha, \beta} f$, em todas as escalas inteiras α sobre $V_{\alpha, \beta}$.

Por sua vez, um elemento f - com um produto interno $\langle \cdot, \cdot \rangle$ - é chamado de função escala se e somente se as funções $\psi_{\alpha, \beta}$, são tais que $\langle f, \psi_{\alpha, \beta} \rangle = 0$, sempre que $\alpha > 0$, e $\beta \in \mathbb{R}$, caso contrário. De acordo com [10], o subespaço fechado de $V_{\alpha, \beta}$ é chamado de subespaço de aproximação, na escala α . A projeção de f sobre $V_{\alpha, \beta}$ é definida pela soma parcial descrita em (4).

De acordo com [10], $P_{\alpha, \beta} f$ pode ser referida como uma componente de aproximação de f , na escala α , sobre $V_{\alpha, \beta}$.

2.3. Transformada Wavelet

Por transformada *wavelet*, sobre $V_{\alpha, \beta}$, entende-se como sendo um produto interno, denotado por $\langle \cdot, \cdot \rangle_{\alpha, \beta}$, entre uma função f e uma função *wavelet* ou uma função escala $\psi_{\alpha, \beta}$.

De acordo com [10], as transformadas *wavelet* podem ser agrupadas em dois conjuntos disjuntos: coeficientes de detalhes, denotados por $d_{\alpha, \beta}$, e coeficientes de aproximação, denotados por $a_{\alpha, \beta}$. O coeficiente de detalhes $d_{\alpha, \beta}$ e o coeficiente de aproximação $a_{\alpha, \beta}$ são definidos, respectivamente, como $d_{\alpha, \beta} = \langle f, \psi_{\alpha, \beta} \rangle$ e $a_{\alpha, \beta} = \langle f, \phi_{\alpha, \beta} \rangle$.

2.4. Expansão Wavelet

De acordo com [9], uma cadeia de subespaços fechados $\{V_{\alpha, \beta}\}_{\alpha \in \mathbb{R}^+}$ em H é uma Análise Multirresolução Wavelet (AMR *wavelet*), com função escala $\psi_{\alpha, \beta}$, se e somente se os axiomas (a), (b), (c), (d) e (e) são satisfeitos.

- $\psi_{\alpha, \beta} \in V_{\alpha, \beta}$;
- $\psi_{\alpha, \beta} \in V_{\alpha, \beta}$;
- $\psi_{\alpha, \beta} \in V_{\alpha, \beta}$;
- $\psi_{\alpha, \beta} \in V_{\alpha, \beta}$; e
- $\{\psi_{\alpha, \beta}\}_{\alpha \in \mathbb{R}^+}$ é uma base ortonormal de $V_{\alpha, \beta}$.

Em [9], mostra-se que o espaço de Hilbert H pode ser ortogonalmente expandido tal como $V_{\alpha, \beta}$, em [11], por meio dos axiomas de uma AMR *wavelet*, mostra-

se que $\{e_n\}$, onde $\{e_n\}$. Com base nas identidades de $\{e_n\}$ e $\{e_n\}$, tem-se a identidade descrita em (5).

Por consequência, segue que $\{e_n\}$ possui decomposição ortogonal (única), sobre o espaço de *Hilbert*, dada em (6).

Dadas as definições das componentes $\{e_n\}$ e $\{e_n\}$ e as identidades (1) e (6), segue que a série de Fourier da função $\{e_n\}$ sobre o espaço de Hilbert $\{e_n\}$, em termos da base ortonormal *wavelet* de $\{e_n\}$, é dada em (7).

3. MODELAGEM BOX & JENKINS

De acordo com [12], um processo estocástico estacionário de segunda ordem é definido como uma família de variáveis aleatórias tal que a média, a variância e as covariâncias são invariantes no tempo. Isto é, $\{e_n\}$, para todo t . Seja $\{e_n\}$ uma realização do processo estocástico $\{e_n\}$, isto é, a série temporal estacionária de segunda ordem. O modelo de *Box & Jenkins* para $\{e_n\}$ é dado em (8).

O modelo, em (8), pode ser representado em termos de dois polinômios, os quais obtidos com o *shift defasagem*, onde B é definido por $B^d y_t = y_{t-d}$, conforme em (9).

Onde $\{e_n\}$ e $\{e_n\}$ denotam os parâmetros do modelo de Box & Jenkins, em (8), e ε_t , o ruído em t . Em relação à verificação da estacionariedade de segunda ordem em uma série temporal, mostra-se, em [5, 12, 16], que tal verificação pode ser realizada através da análise de perfil do gráfico da função de autocorrelação simples k :

onde \bar{y} é a média da série temporal $\{e_n\}$; e k , a defasagem da autocorrelação.

A determinação das ordens dos polinômios p e q pode ser realizada por meio de uma análise do perfil dos gráficos das funções de autocorrelação (k) - ACF (*autocorrelation function*) - e de autocorrelação parcial (kk) - PACF (*partial autocorrelation function*) [12, 16].

Em particular, supondo que a série temporal $\{e_n\}$ apresente tendência (isto é, não estacionariedade na média), deve-se diferenciá-la d vezes [5, 12, 16], a fim de encontrar uma série temporal estacionária de segunda ordem. Neste caso, o modelo de Box & Jenkins para $\{e_n\}$ é dado em (11).

Os modelos de *Box & Jenkins* também podem ser utilizados na modelagem de séries temporais que apresentam sazonalidade [5, 12, 16]. Neste caso, segue que o

modelo *Box & Jenkins* para \hat{y}_t é dado em (12).

onde d é a ordem das diferenças simples; D é a ordem das diferenças sazonais; S é o período sazonal; ϕ , θ e Φ são, respectivamente, os coeficientes dos polinômios não sazonais; e ϕ_s e θ_s são, respectivamente, os coeficientes dos polinômios sazonais.

4. COMBINAÇÃO DE MODELOS PREDITIVOS

O termo “combinação de modelos” concerne tanto à combinação de previsões quanto à combinação de densidades preditivas. Desse modo, seja uma série temporal. Suponha que exista uma coleção \mathcal{M} , de métodos preditivos plausíveis para a geração de previsões para t , dentro e fora da amostra. No entanto, existe o interesse do tomador de decisão em determinar um método único para geração de previsões. De acordo com [1, 4, 13, 17], existem duas possibilidades de determiná-lo: escolher um método no conjunto baseado em critérios de seleção ou tomar k métodos em \mathcal{M} , e combiná-los.

Considere K métodos em \mathcal{M} . Assuma que \mathcal{M} seja um conjunto com todas as previsões pontuais para t , na amostra e fora da amostra (com horizonte de previsão igual a h passos à frente), dos k métodos em \mathcal{M} . Considere ainda que \mathcal{C} seja o conjunto com todas as previsões pontuais para t , na amostra e fora da amostra (com horizonte de previsão igual a h passos à frente), de todas as combinações de previsões pontuais dos K métodos em \mathcal{M} . Por combinação de previsões pontuais entende-se como sendo uma função f que faz o mapeamento de um vetor de previsões pontuais em uma previsão pontual combinada, onde t é um inteiro no intervalo $[1, n+h]$.

Em particular, uma combinação linear de previsões pontuais é uma função que faz o mapeamento de um vetor de previsões pontuais em uma previsão pontual linearmente combinada definida em (13).

onde w_t é o peso adaptativo associado à previsão pontual em t do método i ; e c_i e d_i são as constantes adaptativas multiplicativa e aditiva.

5. METODOLOGIA PROPOSTA

Tome a função característica definida por [10] e a série temporal y_t , a qual pode ser interpretada como o mapa T . Com efeito, segue que a composição $T \circ T$ é um vetor no espaço de Hilbert H . Tome uma base ortonormal *wavelet* de H .

De acordo com a Seção 2.4, a série de Fourier do mapa T sobre H , em termos da base ortonormal *wavelet* de H , é definida por:

Em [7], tem-se que, para a decomposição *wavelet* de nível p , onde $\{x_t\}$, a série de Fourier aproximada é dada conforme em (15).

O valor ótimo do parâmetro de escala λ visa à minimização de uma métrica dos desvios entre \hat{x}_t e x_t .

A partir da expansão em (15), se obtém para λ : uma componente de aproximação \hat{x}_t^A , para $\lambda = 2^p$; e uma componente de detalhes \hat{x}_t^D , para cada $k < p$, onde $\lambda = 2^k$. Com efeito, dada a decomposição *wavelet* de nível p , a decomposição ortogonal (única) \hat{x}_t é dada em (16).

Uma vez obtida a expansão em (16), realiza-se a modelagem de cada componente *wavelet* em (16) via modelos de Box & Jenkins, obtendo-se os conjuntos de previsões: \hat{x}_t^A e \hat{x}_t^D , para cada valor inteiro t no intervalo $[1, T-h]$, sendo h o horizonte de previsão.

Em seguida, estes conjuntos são combinados de forma linear, para cada t , conforme em (17).

onde \hat{x}_t^A é a previsão para $t+h$; \hat{x}_t^D é a previsão para $t+h-k$; α é a previsão para $t+h-k$; β é a constante multiplicativa; γ é a constante aditiva; ω_k é o peso associado às previsões \hat{x}_t^D ; e ω_A é o peso associado às previsões \hat{x}_t^A . Com a decomposição *wavelet*, o peso ω foi substituído pelos pesos atribuídos às componentes de detalhe e de aproximação.

Os valores ótimos das constantes α e β , bem como os pesos ω_k e ω_A , são obtidos por meio de um Problema de Programação Matemática (PPM) não linear [6, 15], cuja função objetivo consiste em minimizar a estatística MSE (*mean square error*):

$$(18.1)$$

Sujeito a duas restrições:

e com $3 + p$ variáveis de decisão irrestritas (18.2)

As previsões para são geradas por meio da equação (13) substituindo-se as constantes e pesos por seus valores ótimos determinados pela solução do PPM em (18.1 e 18.2).

6. EXPERIMENTO COMPUTACIONAL

No experimento computacional, considerou-se a série temporal mensal afluyente da Região Sul do Brasil [14], ilustrada na Figura 1.

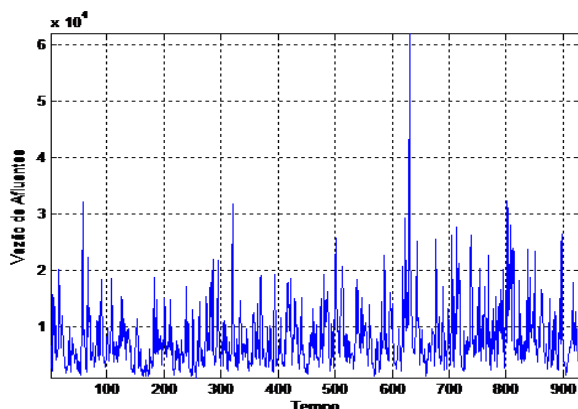


Figura 1 - Série Temporal de Vazão de Afluentes da Região Sul do Brasil.

A referida série temporal possui 936 observações mensais, sendo 924 observações na amostra usada no ajuste do modelo de previsão (período *in sample*) e 12 observações na amostra de teste (período *out sample*). Os softwares utilizados no experimento computacional são o E-Views, o MATLAB, o AIMMS e o Excel. A finalidade do experimento é avaliar o ganho preditivo fora da amostra incorrido na modelagem de *Box & Jenkins* pela utilização da decomposição *wavelet* integrada com a Programação Matemática.

6.1. Modelagem Box & Jenkins

Utilizando o software *EViews*, foi realizada a modelagem da série temporal, que denota a série de vazão de afluentes da Região Sul do Brasil. O melhor modelo foi o ARIMA (2,0,1). Na Tabela 1, tem-se o teste de *Dickey-Fuller* [5], o qual consiste em testar a hipótese de não estacionariedade de segunda ordem. Considerando um nível de significância de 5%, a hipótese de não estacionariedade de segunda ordem da série temporal de vazão de afluentes da região Sul foi rejeitada.

Tabela 1 - Teste de Estacionariedade de Segunda Ordem de *Dickey-Fuller*.

	Estatística <i>t</i>	<i>P</i> -valor	Valor Crítico (5%)
Vazão de Afluentes	-2,86	0,000000	-16,62

6.2. Modelagem Box & Jenkins com Decomposição Wavelet

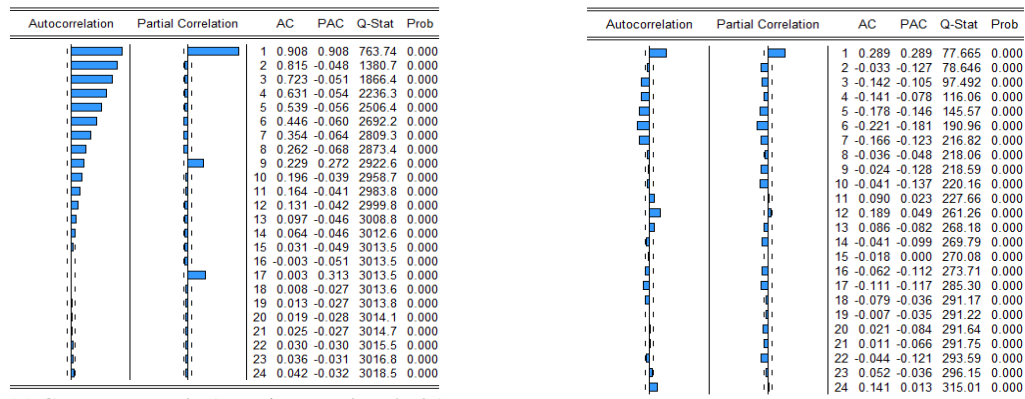
Inicialmente, com o auxílio do software MATLAB, foi realizada a

decomposição *wavelet* de nível 1 tal que . Foi utilizada a base ortonormal de funções *wavelet* de Haar [3, 9]. Na Tabela 2, tem-se o teste de *Dickey-Fuller* que testou a hipótese de não estacionariedade de segunda ordem da componente *wavelet* de aproximação, denotada por , e da componente *wavelet* de detalhes, denotada por .

Tabela 2 - Teste de Estacionariedade de Segunda Ordem de *Dickey-Fuller*.

<i>Subséries Temporais</i>	<i>Estatística t</i>	<i>P-valor</i>	<i>Valor Crítico (5%)</i>
Subsérie Temporal de Aproximação	-2,8645	0,000000	-16,17
Subsérie Temporal de Detalhes	-2,8644	0,000000	-4,85

Note, na Tabela 2, que a hipótese de não estacionariedade de segunda ordem foi rejeitada a 5% de significância.



(a) Componentes de Aproximação do Nível 1.

(b) Componente de Detalhes do Nível 1.

Figura 2 - Correlogramas da FAC e FACP das Componentes *Wavelet* da Decomposição de Nível 1.

Na Figura 2, em (a), os *lag's* 8 e 17 foram considerados como espúrios, de modo que o modelo Box & Jenkins adequado para a componente de aproximação foi o AR (1). Para a componente de aproximação, ajustou-se o modelo ARIMA (2,0,2). Na Tabela 3, têm-se os valores das estatísticas MAPE (*mean absolute percentual error*) e MSE (*mean square error*) das previsões de cada componente *wavelet* na decomposição de nível 1 de y , dentro e fora da amostra.

Tabela 3 – Estatísticas de Aderência.

<i>Subséries Temporais</i>	<i>Amostra de Treino</i>			<i>Amostra de Teste</i>		
	MAPE	MAE	MSE	MAPE	MAE	MSE
Subsérie Temporal de Aproximação	6,49%	481,05	2.049.898,13	16,99%	1.664,12	3.833.980,49
Subsérie Temporal de Detalhes	185,12%	2.731,03	15.252.307,60	74,27%	2.884,41	16.588.307,97

6.3. Método Proposto

Considere as previsões e da Seção 6.2, onde h é o horizonte de previsão. As previsões e foram combinadas, conforme em (13), e suas constantes e os pesos adaptativos ótimos foram obtidos com o uso do *software* AIMMS.

Na Tabela 5, têm-se as constantes e os pesos ótimos da combinação linear de

componentes *wavelet* de nível 1 obtidas através do *solver* CONOPT [<http://www.gams.com/solvers/solvers.htm>] do *software* AIMMS.

Tabela 4 - Constantes e Pesos Ótimos da Combinação Linear de Componentes *Wavelet* de Nível 1

<i>Peso da Componente de Aproximação</i>	<i>Peso da Componente de Detalhes</i>
0,842970654	0,787250152
<i>Constante Multiplicativa</i>	<i>Constante Aditiva β</i>
1,198827394	-74,96849044

6.4. Comparação de Resultados

Na Tabela 4, são apresentadas as estatísticas MAPE (*mean absolute percentual error*) e MSE (*mean square error*), nas amostras dos períodos *in sample* e *out sample*.

Tabela 5 - Qualidade do Ajuste Dentro e Fora da Amostra.

<i>Modelagem</i>	<i>Período in sample</i>		<i>Período out sample</i>	
	MAPE	MSE	MAPE	MSE
Box & Jenkins	54,15 %	22.453.582,2 2	57,41 %	22.844.731,7 9
Box & Jenkins com Decomposição	44,01 %	15.904.972,7 0	30,25 %	20.626.493,8 0
Método Proposto	23,65 %	7.644.912,04	28,87 %	20.399.290,2 5

7. CONCLUSÕES

Neste artigo, foi proposta um método que integra à modelagem de *Box & Jenkins* a Teoria *Wavelet* e a Programação Matemática (PM) no processo de geração de previsões pontuais.

Na Tabela 4, tem-se que a decomposição *wavelet* incorreu em ganhos preditivos consideráveis em relação à abordagem ARIMA tradicional. Por sua vez, o método proposto (que combina linearmente as previsões das componentes *wavelet* via PM) acarretou ainda uma aderência maior à amostra do período *in sample* e aumentou o poder de generalização das previsões *out sample*.

REFERÊNCIAS

- Bates, J.M., Granger, C.W.J (1969). Combination Forecasts. *Operations Research Quarterly*.
- Box, G.E.P., Jenkins, G.M (1976). *Time Series Analysis Forecasting and Control*. Ed. San Francisco: Holden-Day.
- Daubechies, I. (1988) Orthonormal Bases of Compactly Supported *wavelets*. *Comm. Pure and Applied Math.*, 41, pp.909-996.
- Flores, B.E., White, E.M (1988). A Framework for the Combination of Forecasts, *Journal Academic Marketing Science*, v.16, (3-4), p-95-103.
- Hamilton, J (1994). *Time Series Analysis*. Princeton University Press.
- Junior, L. A. T, Pessanha, J. F. M., Souza, R. C. (2011). Combinação Linear de Previsões com Ajuste MINIMAX. XIV Simpósio de Pesquisa Operacional e

Logística da Marinha, Rio de Janeiro, Brasil.

- Levan N., Kubrusly, C. S. (2003). A *wavelet* “Time-Shift-Detail” Decomposition. *Mathematics and Computers in Simulation* 63, 73-78.
- Kubrusly, C. S. (2001). *Elements of Operator Theory*. Birkhäuser. Boston.
- Kubrusly, C. S., Levan, N. (2006). Abstract *wavelets* Generated by Hilbert Space Shift Operators. *Adv. Math. Sci. Appl.* 14, 643-660.
- Kubrusly, C. S., Levan N. (2002), Dual-Shift Decomposition of Hilbert Space. *Semigroups of Operators: Theory and Application* 2, 145-157.
- Mallat S. (1998), A *wavelet* Tour of Signal Processing. Academic Press. San Diego.
- Morettin, P.A., Toloi, L.M.C (2006). *Análise Séries Temporais*, 2ª Ed. ABE. Projeto Fisher. Editora: Edgard Blucher.
- Newbold, P., and Granger, C.W. J. (1974). Experience with Forecasting Univariate Times Series and the Combination of Forecasts. *Journal of the Royal Statistical Society*. V.137, pp.131-165.
- ONS - Operador Nacional do Sistema Elétrico. Site:<http://www.ons.com.br>.
- Ragsdale, C.T. (2009) *Modelagem e Análise de Decisão*, São Paulo, Cengage Learning.
- Souza, R.C., Camargo, M.E (1996). *Análise e Previsão de Séries Temporais: Os Modelos ARIMA*. Ijuí: SEDIGRAF.
- Winkler, R.L., Makridakis, S (1983). The combination of Forecasting, *Journal of the Royal Statistical Society*, series A, v.146, 1983, p.150-157.