VIII ERMAC-R3

5º Encontro Regional de Matemática Aplicada e Computacional
20 a 22-Novembro-2008
Universidade Federal do Rio Grande do Norte - Natal/RN

Métodos de Reamostragem de Séries Temporais Baseados em Transformada Wavelet Discreta e Transformada Wavelet Discreta por Pacotes

Ronaldo Mendes Evaristo **Luiz Antonio Baccalá**,
Depto. de Engenharia de Telecomunicações e Controle, Escola Politécnica, USP,
Av. Prof. Luciano Gualberto, Tr.3, 158
17033-360, São Paulo, SP
E-mail: ronaldo@lcs.poli.usp.br, baccala@lcs.poli.usp.br,

Resumo

Neste artigo revisamos a reamostragem de séries temporais baseada na Transformada Wavelet Discreta(DWT) e apresentamos um método que usa Transformada Wavelet Discreta por Pacotes(DWPT) para a geração de dados reamostrados em substituição aos métodos clássicos [5, 7, 8, 13] e os aplicamos na estimação da densidade espectral de potência de séries temporais geradas a partir de modelos autoregressivos.

Palavras-chave

Wavestrap, wavestrapping, estimação espectral, wavelet packets, reamostragem, intervalos de confiança.

Introdução

Idealmente, quando precisamos estimar certos parâmetros de um processo estocástico, necessitamos de várias realizações do mesmo para podermos determinar, por exemplo, através de intervalos de confiança, o grau de confiabilidade das estimativas produzidas. Ocorre que nem sempre é possível realizar um experimento mais de uma vez e, nestes casos, necessitamos dos métodos de reamostragem para abordar a problemática. Tais métodos surgiram das idéias básicas de *bootstrap*, introduzidas por Efron [6], e cabem a eles gerar realizações hipotéticas(dados reamostrados) de um determinado experimento com base em uma realização inicial fornecida. Contudo, o *bootstrap* básico pode ser aplicado somente à variáveis aleatórias i.i.d. o que limita sua aplicação direta sobre séries temporais, onde as amostras geralmente são correlacionadas. Isso têm motivado muitos pesquisadores a buscarem métodos que permitam descorrelacionar as amostras das séries para uma posterior aplicação do procedimento de *bootstrap*.

Com o surgimento da Transformada Wavelet e principalmente com a introdução do algoritmo piramidal, feita por Mallat [14], permitindo o cálculo eficiente da Transformada Wavelet Discreta(DWT), certas propriedades vieram à tona. Uma dessas propriedades, discutida por vários autores [9, 10, 17, 20], diz que seus coeficientes tendem a se tornarem aproximadamente descorrelacionados conforme o número de níveis e a ordem dos filtros utilizados na decomposição aumentam, o que levou Bullmore et.al. [4] a propôr o procedimento de *wavestrap*.

Além dos métodos de reamostragem baseados em *wavelets*, existem outros baseados nos domínios do tempo e da freqüência [13], vários deles recentemente revistos em [2].

Alguns autores [9, 10] mostraram que o *wavestrap* funciona bem para modelos onde as densidades espectrais de potência(DEP) são suaves, ou seja, provenientes de modelos que contenham mais zeros que pólos. Contudo, quando o espectro possui picos muito acentuados, que é o caso dos modelos autoregressivos, esses métodos baseados na DWT apresentam certo viés dependendo da forma como são aplicados. Nesse artigo fazemos uma revisão dos métodos baseados na DWT para estes casos e apresentamos um método que usa a DWPT como outra alternativa.

Reamostragem de Séries Temporais Baseada em Wavelets

As duas transformadas em questão podem ser estudadas por meio de matrizes ortogonais [15], isto é, cada uma delas possui uma matriz ortogonal de transformação $\mathcal{W}_{N\times N}$ que transforma um sinal(série temporal)

$$\mathbf{x} = [x(0)\cdots x(N-1)]^T, \tag{1}$$

de comprimento $N = 2^J$ em um vetor $\mathbf{W}_{N\times 1}$ contendo seus respectivos coeficientes por meio da equação de análise,

$$\mathbf{W} = \mathcal{W}\mathbf{x}. \tag{2}$$

Além disso, como a matriz $\mathcal{W}$ é ortogonal, o sinal $\mathbf{x}$ pode ser recuperado de acordo com a equação de síntese

$$\mathbf{x} = \mathcal{W}^T\mathbf{W}. \tag{3}$$

Considerando que $\mathbf{x}$ seja decomposto em $J_0 \leq J$ níveis, $J_0 \in \mathbb{N}$, o vetor de coeficientes transformados $\mathbf{W}$ pode ser particionado para as duas transformadas como [15]

$$\underset{\text{DWT}}{\mathbf{W}} = \begin{bmatrix} \mathbf{W}_1 \\ \vdots \\ \mathbf{W}_{J_0} \\ \mathbf{V}_{J_0} \end{bmatrix} \quad \text{e} \quad \underset{\text{DWPT}}{\mathbf{W}} = \begin{bmatrix} \mathbf{W}_{J_0,0} \\ \mathbf{W}_{J_0,1} \\ \vdots \\ \mathbf{W}_{J_0,2^{J_0}-1} \end{bmatrix} \tag{4}$$

onde,

$$\mathbf{W}_j = \left[\begin{array}{ccc} W\left(\frac{2^{j-1}-1}{2^{j-1}}N\right) & \cdots & W\left(\frac{2^j-1}{2^j}N-1\right) \end{array} \right]^T$$
$$\mathbf{V}_{J_0} = \left[\begin{array}{ccc} W\left(\frac{2^{J_0-1}-1}{2^{J_0-1}}N\right) & \cdots & W\left(\frac{2^{J_0}-1}{2^{J_0}}N-1\right) \end{array} \right]^T$$
$$\mathbf{W}_{J_0,k} = \left[\begin{array}{ccc} W\left(\frac{k}{2^{J_0}}N\right) & \cdots & W\left(\frac{k+1}{2^{J_0}}N-1\right) \end{array} \right]^T$$

com $1 \leq j \leq J_0$ e $0 \leq k \leq 2^{J_0} - 1$.

Foi utilizando essa interpretação que Mallat [14] introduziu o algoritmo piramidal, ilustrado nas figuras 1 e 2 para $J_0 = 3$. Esse algoritmo está ligado aos conceitos de bancos de filtros [1, 18] e sua conexão com a representação matricial [15] é feita pela matriz $\mathcal{W}$ que depende explicitamente dos coeficientes do filtro $H(z)$ que por sua vez depende da *wavelet* adotada.

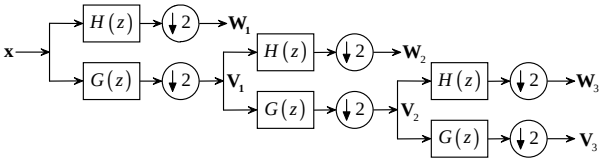

Figura 1: DWT - Decomposição por Bancos de Filtros

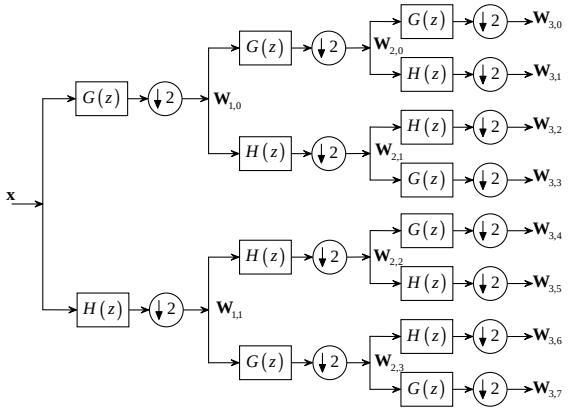

Figura 2: DWPT - Decomposição por Bancos de Filtros

Os métodos de reamostragem baseados em wavelets estão apresentados nos procedimentos 1 e 2 e a grande motivação para estarmos apresentando um método de reamostragem baseado na DWPT vem do fato mostrado na figura 3, onde vemos como o espectro está

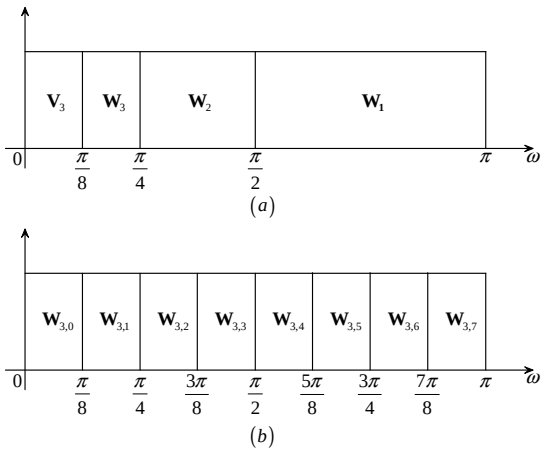

Figura 3: Divisão do Espectro: (a)DWT; (b)DWPT

divido para o caso das duas transformadas. Conforme pode ser visto, a DWT não analisa a banda de freqüências $[\pi/2, \pi]$ aos quais os coeficientes $\mathbf{W}_1$ estão associados, ao passo que a DWPT o faz. Isto se torna um grande diferencial conforme veremos nos exemplos a seguir.

Procedimento 1 Reamostragem baseada em Wavelets utilizando a DWT: Wavestrap

Entrada: $\mathbf{x} = [x(0) \cdots x(N-1)]^T$ e J_0

1. Escolha a wavelet desejada assim como sua ordem e monte a matriz $\mathcal{W}$ [15].
2. Determine a DWT de $\mathbf{x}$ através da equação de análise,
$$\mathbf{W} = \mathcal{W}\mathbf{x}.$$
3. Avalie as funções de autocorrelação dos coeficientes presentes nos subvetores $\mathbf{W}_1 \cdots \mathbf{W}_{J_0}$, aplique o *bootstrap* aos $\mathbf{W}_i$ que estiverem aproximadamente descorrelacionados, mantenha os restantes assim como $\mathbf{V}_{J_0}$ inalterados e obtenha
$$\mathbf{W}^* = \begin{bmatrix} \mathbf{W}_1^* \\ \vdots \\ \mathbf{W}_{J_0}^* \\ \mathbf{V}_{J_0} \end{bmatrix}.$$
4. Calcule o sinal reamostrado através da equação de síntese da DWT,
$$\mathbf{x}^* = \mathcal{W}^T\mathbf{W}^*.$$

Saída: $\mathbf{x}*$

Exemplos

Apresentaremos três aplicações dos procedimentos 1 e 2. Em todos os casos consideraremos sistemas

Procedimento 2 Reamostragem baseada em Wavelets utilizando a DWPT: Wavestrapping

Entrada: $\mathbf{x} = [x(0) \cdots x(N-1)]^T$ e J_0

1. Escolha a wavelet desejada assim como sua ordem e monte a matriz $\mathcal{W}$ [15].
2. Determine a DWT de $\mathbf{x}$ através da equação de análise,
$$\mathbf{W} = \mathcal{W}\mathbf{x}.$$
3. Avalie as funções de autocorrelação dos coeficientes presentes nos subvetores $\mathbf{W}_{J_0,0} \cdots \mathbf{W}_{J_0,2^{J_0}-1}$, aplique o *bootstrap* aos $\mathbf{W}_{J_0,i}$ que estiverem aproximadamente descorrelacionados, mantendo os restantes inalterados e obtenha
$$\mathbf{W}^* = \begin{bmatrix} \mathbf{W}^*_{J_0,0} \\ \vdots \\ \mathbf{W}^*_{J_0,2^{J_0}-1} \end{bmatrix}.$$
4. Calcule o sinal reamostrado através da equação de síntese da DWPT,
$$\mathbf{x}^* = \mathcal{W}^T\mathbf{W}^*.$$

Saída: $\mathbf{x}*$

autoregressivos, ou seja, sistemas do tipo

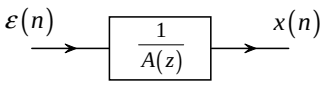

Figura 4: Sistema Autoregressivo

com
$$x(n) = \frac{1}{A(z)}\varepsilon(n), \quad 0 \le n \le N-1 \tag{5}$$
onde $\varepsilon(n)$ é um ruído branco gaussiano de média nula e variância unitária.

Em todos os casos consideraremos $N = 1024$, $J_0 = 4$ e como *wavelet* utilizaremos a $D(40)$ [15]. Além disso, utilizaremos o método de Burg [11] para estimar as densidades espectrais de potência(DEP) das séries geradas.

Exemplo 01: Seja

$$A(z) = 1+0,4z^{-1}+0,2z^{-2}+0,1z^{-3}+0,1z^{-4}+0,1z^{-5}$$

e $x_1(n)$ calculado de acordo com a equação 5. Uma estimativa da densidade espectral de potência de $x_1(n)$ pode ser vista na figura 5(a).

As figuras 6 e 7 mostram a autocorrelação dos coeficientes da DWT e DWPT de $x_1(n)$, respectivamente. Podemos ver que todos os coeficientes em todos os níveis são praticamente descorrelacionados. Com isso, se aplicarmos os procedimentos 1 e 2 à série $x_1(n)$ 10 vezes, obtendo as versões reamostradas $x_1^*(n)$ teremos como resultado o mostrado nas figuras 5(b) e 5(c) para a DWT e DWPT, respectivamente.

Exemplo 02: Seja

$$A(z) = 1 + 1,6018z^{-1} + 0,9800z^{-2}$$

e $x_2(n)$ calculado de acordo com a equação 5. Uma estimativa da densidade espectral de potência de $x_2(n)$ pode ser vista na figura 8(a).

As figuras 9 e 10 mostram a autocorrelação dos coeficientes da DWT e DWPT de $x_2(n)$, respectivamente. Se fizermos 10 vezes a reamostragem em todos os níveis da DWT e em todos exceto, por exemplo, $\mathbf{W}_{4,10}$, $\mathbf{W}_{4,12}$, $\mathbf{W}_{4,13}$ e $\mathbf{W}_{4,14}$ da DWPT teremos o resultado mostrado nas figuras 8(b) e 8(c), respectivamente. Podemos ver que o resultado apresentado pela DWT é bem viesado comparado ao da DWPT. Isso se deve ao fato de termos reamostrado os coeficientes de $\mathbf{W}_1$ que apresentam uma autocorrelação alta, ao passo que os coeficientes da DWPT são praticamente descorrelacionados.

Exemplo 03: Seja

$$A(z) = 1 - 0,0001z^{-1} + 0,9802z^{-2} - 0,0002z^{-3} + \\ +0,9605z^{-4} - 0,0001z^{-5} + 0,9414z^{-6} - 0,0001z^{-7} + \\ +0,9227z^{-8}$$

e $x_3(n)$ calculado de acordo com a equação 5. Uma estimativa da densidade espectral de potência de $x_3(n)$ pode ser vista na figura 11(a).

As figuras 12 e 13 mostram a autocorrelação dos coeficientes da DWT e DWPT de $x_3(n)$, respectivamente. Se fizermos 10 vezes a reamostragem em todos os níveis, exceto $\mathbf{W}_1$, da DWT e em todos, exceto $\mathbf{W}_{4,9}$, $\mathbf{W}_{4,10}$ e $\mathbf{W}_{4,13}$, da DWPT teremos o resultado mostrado nas figuras 11(b) e 11(c), respectivamente. Podemos ver que o resultado apresentado pela DWT é bem viesado comparado ao da DWPT. Isso se deve ao fato de termos reamostrado os coeficientes de $\mathbf{W}_2$ que apresentam uma autocorrelação alta, ao passo que os coeficientes da DWPT são praticamente descorrelacionados, exceto pelos desprezados na reamostragem.

Conclusões

Através dos exemplos apresentados pudemos ver que os métodos baseados em Wavelets são eficientes para a geração de dados reamostrados, bastando escolher adequadamente os níveis aonde se deseja realizar a reamostragem antes da reconstrução. O fato do *wavestrap* utilizando a DWT ter apresentado resultados viesados para os exemplos 2 e 3 está vinculado aos picos presentes em $\frac{\pi}{2} \le \omega \le \pi$, visto que ela não decompõe os sinais nessa banda. Já os coeficientes da DWPT se apresentaram aproximadamente descorrelacionados nos três casos e conforme o nível

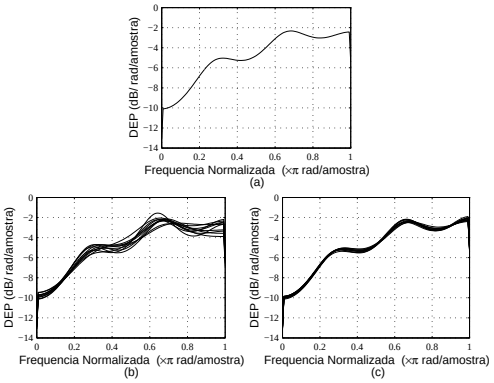

Figura 5: DEP: (a) $x_1(n)$; (b) $x_1^*(n)$(DWT); (c) $x_1^*(n)$(DWTP)

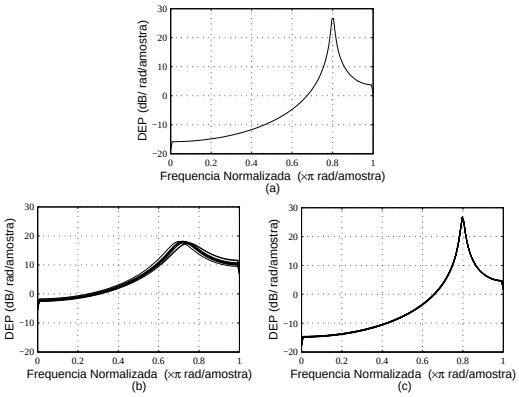

Figura 8: DEP: (a) $x_2(n)$; (b) $x_2^*(n)$(DWT); (c) $x_2^*(n)$(DWTP)

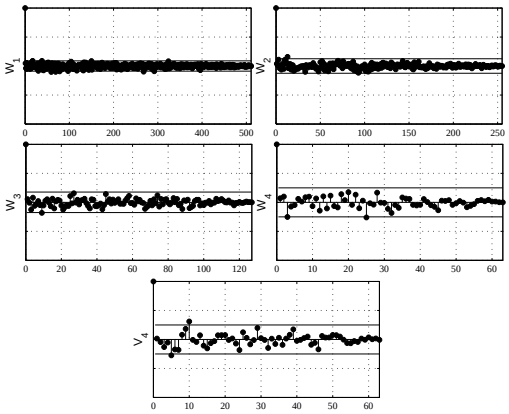

Figura 6: Autocorrelação dos Coeficientes da DWT de $x_1(n)$

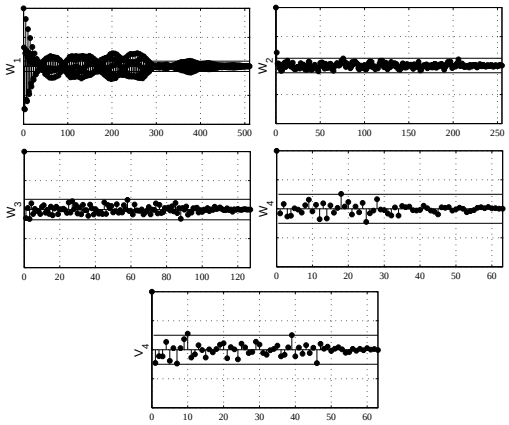

Figura 9: Autocorrelação dos Coeficientes da DWT de $x_2(n)$

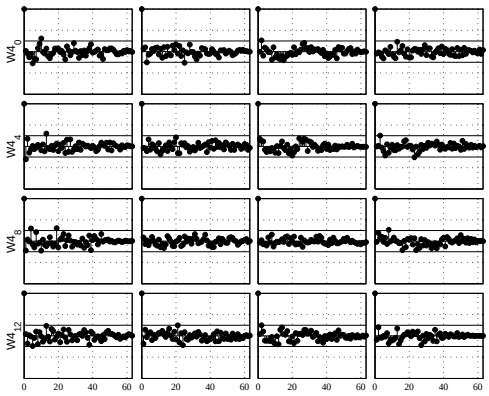

Figura 7: Autocorrelação dos Coeficientes da DWPT de $x_1(n)$

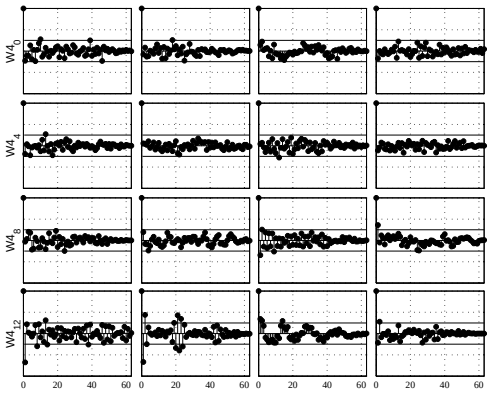

Figura 10: Autocorrelação dos Coeficientes da DWPT de $x_2(n)$

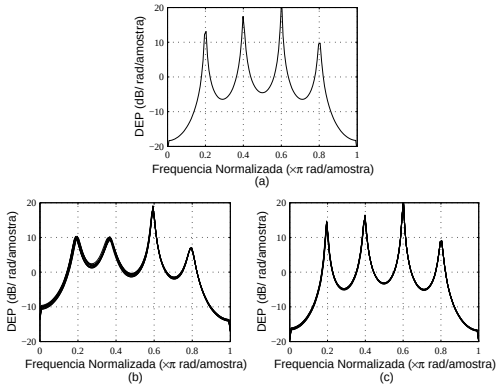

Figura 11: DEP: (a) $x_3(n)$; (b) $x_3^*(n)$(DWT); (c) $x_3^*(n)$(DWTP)

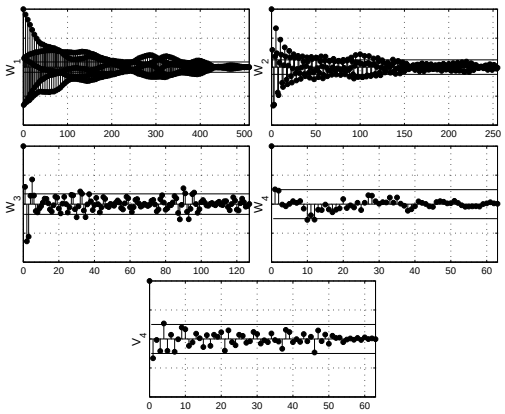

Figura 12: Autocorrelação dos Coeficientes da DWT de $x_3(n)$

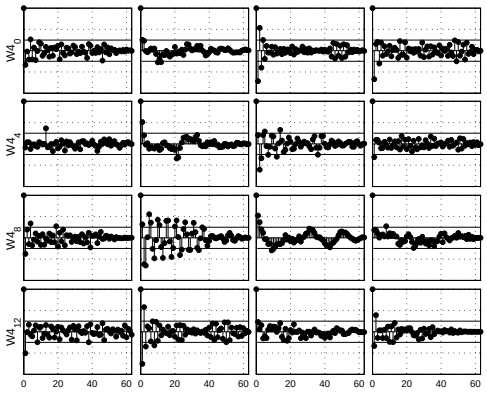

Figura 13: Autocorrelação dos Coeficientes da DWPT de $x_3(n)$

de decomposições aumente, a tendência é que fiquem ainda mais descorrelacionados.

Agradecimentos

Ao CNPq pelo apoio financeiro.

Referências

[1] A. N. Akansu and R. A. Haddad. *Multiresolution Signal Decomposition: Transforms, Subbands and Wavelets.* Academic Press, San Diego, 2001.

[2] L. A. Baccalá, D. Y. Takahashi, and K. Sameshima. Computer intensive testing for the influence between time series. In B. Schelter, M. Winterhalder, and J. Timmer, editors, *Handbook of Time Series Analysis: Introduction and Overview.* WILEY-VCH Verlag GmbH & KGaA, 2006.

[3] M. Breakspear, M. Brammer, E. Bullmore, P. Das, and L. Williams. Spatiotemporal wavelet resampling for functional neuroimaging data. *Human Brain Mapping*, 23:1–25, 2004.

[4] E. Bullmore, C. Long, J. Suckling, J. Fadili, G. Calvert, F. Zelaya, T. Carpenter, and M. Brammer. Colored noise and computational inference in neurophysiological time series analysis: resampling methods in time and wavelet domains. *Human Brain Mapping*, 12:61–78, 2001.

[5] A. C. Davison and D. V. Hinkley. *Bootstrap Methods and their Application.* Cambridge University Press, Cambridge, 1997.

[6] B. Efron. Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7:1–26, 1979.

[7] B Efron. The jackknife, the bootstrap and other resampling plans. In *CBMS-NSF Regional Conference Series in Applied Mathematics.* SIAM, 1982.

[8] B. Efron and R. Tibshirani. *An Introduction to the Bootstrap.* Chappman & Hall, London, 1993.

[9] H. Feng, T. Willemian, and N. Shang. Wavelet-based bootstrap for time series analysis. *Communications in Statistics - Simulation and Computation*, 34:393–413, 2005.

[10] P. Flandrin. Wavelet analysis and synthesis of fractional brownian motion. *IEEE Transactions on Information Theory*, 38:910–917, 1992.

[11] S. M. Kay. *Modern Spectral Estimation: Theory and Application.* Prentice Hall, New Jersey, 1988.

[12] S. N. Lahiri. Sieve bootstrap with variable-length markov chains for stationary categorical time series. *Journal of American Statistical Association*, 97:460–462, 2002.

[13] S. N. Lahiri. *Resampling Methods for Dependent Data*. Springer, New York, 2003.

[14] S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:674–693, 1989.

[15] D. B. Percival and A. T. Walden. *Wavelet Methods for Time Series Analysis*. Cambridge University Press, New York, 2000.

[16] D. N. Politis. The impact of bootstrap methods on time series analysis. *Statistical Science*, 18:219–230, 2003.

[17] A. H. Tewfik and M. Kim. Correlation structure of the discrete wavelet coefficients of fractional brownian motion. *IEEE Transactions on Information Theory*, 38:904–909, 1992.

[18] P. P. Vaidyanathan. *Multirate Systems and Filter Banks*. Prentice Hall, New Jersey, 1993.

[19] B. Vidakovic. *Statistical Modeling by Wavelets*. John Wiley & Sons, Inc, New York, 1999.

[20] G. Zhang, J. Walter. A wavelet-based kl-like expansion for wide-sense stationary random processes. *IEEE Transactions on Signal Processing*, 42:1737–1745, 1994.