Communications in Statistics—Simulation and Computation®, 37: 1307–1315, 2008

ISSN: 0361-0918 print/1532-4141 online
DOI: 10.1080/03610910802050902

Taylor & Francis
Taylor & Francis Group

Inference

Undercoverage of Wavelet-Based Resampling Confidence Intervals

LIANSHENG TANG[1], WAYNE A. WOODWARD[2], AND WILLIAM R. SCHUCANY[2]

[1]Department of Statistics, George Mason University, Fairfax, Virginia, USA
[2]Department of Statistics, Southern Methodist University, Dallas, Texas, USA

The decorrelating property of the discrete wavelet transformation (DWT) appears valuable because one can avoid estimating the correlation structure in the original data space by bootstrap resampling of the DWT. Several authors have shown that the wavestrap approximately retains the correlation structure of observations. However, simply retaining the same correlation structure of original observations does not guarantee enough variation for regression parameter estimators. Our simulation studies show that these wavestraps yield undercoverage of parameters for a simple linear regression for time series data of the type that arise in functional MRI experiments. It is disappointing that the wavestrap does not even provide valid resamples for both white noise sequences and fractional Brownian noise sequences. Thus, the wavestrap method is not completely valid in obtaining resamples related to linear regression analysis and should be used with caution for hypothesis testing as well. The reasons for these undercoverages are also discussed. A parametric bootstrap resampling in the wavelet domain is introduced to offer insight into these previously undiscovered defects in wavestrapping.

Keywords Bootstrap; Discrete wavelet transform (DWT); fMRI; Inverse discrete wavelet transform (IDWT); Permutation tests; Wavestrap.

Mathematics Subject Classification Primary 62G09; Secondary 62P10.

1. Introduction and Notation

The wavelet transform has emerged as a powerful mathematical tool for decomposing a function $f(t)$ in terms of its time and frequency components. The wavelet transform is superior to the classical Fourier transform whenever one

Received June 22, 2007; Accepted March 12, 2008
Address correspondence to Liansheng Tang, Department of Statistics, George Mason University, Fairfax, VA 22030, USA; E-mail: ltang1@mail.gmu.edu

wants the localization in time as well as frequency for non stationary signals. Wavelet transforms have attracted some attention from statisticians since Donoho and his coauthors introduced signal analysis with wavelet denoising and shrinkage (See, for example, Donoho, 1993, 1995).

The discrete wavelet transform (DWT) is described in the next section. The DWT is computationally efficient using Mallat's pyramid algorithm, which can be easily generalized to higher dimensions (Mallat, 1989a,b). One desirable property of the DWT is the decorrelating property of the wavelet coefficients (Flandrin, 1992; Frazier et al., 1991; Tewfik and Kim, 1992; Zhang and Walter, 1995). This decorrelating property in the two-dimensional setting is the foundation for the "enhanced" false discovery rate, which reduces the number of local hypotheses in the wavelet domain and can result in better power than the classical false discovery rate (Shen et al., 2002).

Resampling techniques including bootstrap and permutation tests are discussed in the next section. Bootstrap methods work well, preserving good coverage properties when the data are uncorrelated (Efron and Tibshirani, 1993). However, the classical bootstrap methods fail when the data are correlated (Davison and Hinkley, 1997). The decorrelating property of wavelet coefficients suggests that resamples from correlated data can be obtained by transforming the original data to the wavelet domain and resampling decorrelated wavelet coefficients. Bullmore et al. (2001) proposed this new resampling method based on independent permutations of the wavelet coefficients of the observed time series. Their simulations show that the resampled time series have autocorrelation functions very similar to the original. Breakspear et al. (2004) extended the wavelet-based resampling method to identify dynamic interactions between brain regions from fMRI data.

Several authors have shown that the wavestrap approximately retains the correlation structure of observations. However, estimating regression parameters from the resampling observations is not only related to the correlation structure, but more importantly to the magnitudes of the series. The simplest case in regression analysis would be inference on the mean of a series of observations from wavestrap resamples. It is well known that the mean will remain the same if only the detail wavelet coefficients are resampled. Thus, no matter how many wavestrap resamples are obtained from the data, the sample means from the wavestrap resamples are the same since resampled versions of the sample mean have virtually no variation. It is therefore impossible to draw inference on the population mean. In a recent article, Feng et al. (2005) noted that the sample mean of the wavestrap resamples displays much less variability than the sample mean of the original process. Since inference based on a sample mean is a special case of inference based on regression coefficients, similar results can be expected for linear regression settings. In this article, we present simulation studies to show that wavelet-based resampling should also be used with caution in a regression modeling that is typical in functional Magnetic Resonance Imaging (fMRI). Although the autocorrelation structure of wavelet resamples may be similar to that for the original data as illustrated in Bullmore et al. (2001), Breakspear et al. (2004) and other authors, our simulation studies show that wavestrap-based confidence intervals of the slope coefficient suffer from severe undercoverage. In addition, we introduce the parametric wavestrap to offer insight into these wavestrapping defects.

2. Methods

2.1. *Wavelet-Based Resampling*

Let $\mathbf{y} = \{y_0, y_1 \ldots, y_{N-1}\}$ denote a vector or signal, where $N = 2^J$ for some integer J. Using a mother wavelet ψ, the signal y_i can be expressed as:

$$y_i = S_{Jo} + \sum_{j=1}^{J_0} D_j, \tag{1}$$

where J_0 is an integer less than J,

$$D_j = \sum_{k=1}^{N/2^j} d_{jk}\psi_{jk}(i/N),$$

and

$$S_{J_0} = \sum_{k=1}^{N/2^{J_0}} s_{J_0k}\varphi_{J_0k}(i/N),$$

and φ is the associated father wavelet. The detail coefficients, d_{jk}, and the scaling coefficients, s_{J_0k}, are obtained using the highly efficient pyramid algorithm developed by Mallat (1989). S_{J_0} in (1) accounts for the "smooth" features in the data while the D_j's provide "detail" information with smaller values of j being associated with finer detail. Given a signal of length $N = 2^J$, the discrete wavelet transform (DWT) consists of the coefficients d_{jk} and s_{J_0k}. There are $N/2^j$ detail coefficients associated with D_j, $j = 1, \ldots, J_0$ and $N/2^{J_0}$ scaling coefficients associated with S_{J_0} yielding N coefficients in all. Thus, the DWT is a transformation from the N data values to these N coefficients. The original data can also be completely reconstructed from the coefficients, and this transformation is referred to as the inverse discrete wavelet transform (IDWT).

Several authors (Flandrin, 1992; Frazier et al., 1991; Tewfik and Kim, 1992; Zhang and Walter, 1995) have discussed the wavelet decorrelating property. This property is the basis for wavelet-based resampling techniques in recent articles such as Bullmore et al. (2001) and Breakspear et al. (2004). The key idea of wavelet decorrelating is that when a correlated time series is transformed into the wavelet domain via the DWT, the detail coefficients $\mathbf{d}_j$'s, where $\mathbf{d}_j = (d_{j1}, \ldots, d_{jk})$, are approximately uncorrelated within each level and between levels. The decorrelated detail coefficients are then resampled within levels and inverse transformed back to the time domain via the IDWT to obtain a simulated resample of the original process.

Let $\mathbf{y}(t)$ be a time series of N points. The wavelet-based resampling algorithm is outlined in the following steps:

1. Choose a suitable orthogonal wavelet basis, such as Daubechies D(4) or D(8).
2. Compute the DWT of $\mathbf{y}(t)$, i.e., $\mathrm{DWT}(\mathbf{y}) = \mathbf{w} = (\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_{J_0}, \mathbf{s}_{J_0})$.
3. Apply the ordinary "naive" bootstrap (resampling with replacement) or permutation (resampling without replacement) to all the wavelet detail coefficients $\mathbf{d}_j$, $j = 1, \ldots, J_0$ where $J_0 < J$, within each level to obtain the resampled coefficients $\mathbf{d}_j^*$. Leave the $N/2_0^J = 2^{J-J_0}$ scaling coefficients unchanged, $\mathbf{s}_{J_0}^* = \mathbf{s}_{J_0}$. That is, $\mathbf{w}^* = (\mathbf{d}_1^*, \mathbf{d}_2^*, \ldots, \mathbf{d}_{J_0}^*, \mathbf{s}_{J_0}^*)$.

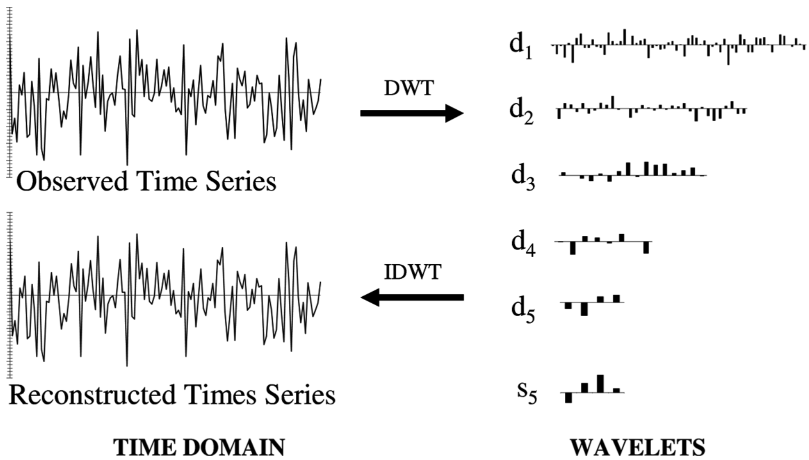

Figure 1. The wavestrap algorithm. The left panel shows observed and resampled time series. The right panel gives the plot of wavelet coefficients of different levels using the wavelet basis D(8). Here $J = 7$, and the DWT is shown for $J_0 = 5$.

4. Apply the inverse discrete wavelet transform to obtain the resampled time series $\mathbf{y}^* = \text{IDWT}(\mathbf{w}^*)$. We also renormalize $\mathbf{y}^*$ using a scale of $\sqrt{1.1}$ as recommended in Breakspear et al. (2004) to maintain the autocorrelation of resamples.
5. Calculate the statistic of interest $\hat{\theta}^*$ using $\mathbf{y}^*$.
6. Repeat Steps 3–5 B times to estimate the sampling distribution of $\hat{\theta}^*$.

The wavestrap algorithm is illustrated in Fig. 1. The time series consists of 128 time points, and the wavelet basis used in the plot is D(4). The figure is similar to one in Bullmore et al. (2001). When a time series is not very long, typically 128 or 256 time points, Bullmore et al. (2001) recommend using D(4) to reduce the boundary effect of wavelets. Both D(4) and D(8) are used as wavelet bases in our simulation study. Approximate confidence intervals use the B values of $\hat{\theta}^*$.

The $100(\alpha/2)$th percentile, $\hat{\theta}^*_{((B+1)\alpha/2)}$, and $100(1-\alpha/2)$th percentile, $\hat{\theta}^*_{((B+1)(1-\alpha/2))}$, are calculated from the resampling distribution. The "basic" confidence interval is obtained using the limits (L_B, U_B) (Davison and Hinkley, 1997, Sec 5.2.1), where

$$L_B = CI_{\alpha/2} = 2\hat{\theta} - \hat{\theta}^*_{((B+1)(1-\alpha/2))} \quad \text{and} \quad U_B = CI_{1-\alpha/2} = 2\hat{\theta} - \hat{\theta}^*_{((B+1)\alpha/2)}.$$

2.2. *Parametric Wavestrap*

A potential problem with the wavestrap approach is that whenever J_0 is close to J, there are very few wavelet coefficients in $\mathbf{d}_{J_0}$ to be resampled. For example, if $J = 7$, and $J_0 = 5$ as in Fig. 2, then there are only 4 wavelet coefficients in $\mathbf{d}_5$. This can result in resampled series with too much of the structure repeated. To demonstrate the effect of this problem we introduce the parametric wavestrap (PW). Since the Daubechies wavelet bases are orthonormal, the detail coefficients at all levels and the scaling coefficients theoretically follow independent normal distributions if the original data are standard normal white noise. In this special case, therefore, instead of resampling the observed detail coefficients, the detail coefficients can be generated from $N(0, 1)$. While this parametric wavestrap approach is not one that could be

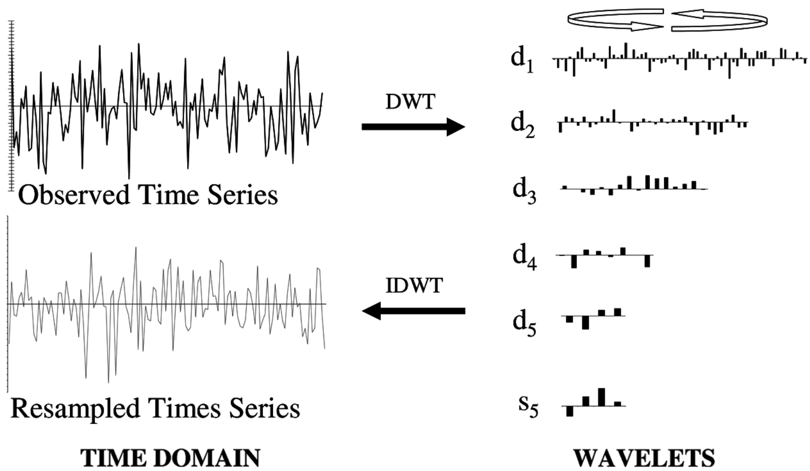

Figure 2. The wave length algorithm. The left panel shows observed and resampled time series. The right panel gives the plot of wavelet coefficients of different levels using the wavelet basis $D(8)$. Here $J = 7$, and the DWT is shown for $J_0 = 5$.

used in most practical settings, it is considered here to ascertain the degree to which the wavestrap undercoverage can be explained by the resampling from a small number of detail coefficients.

3. Inference on a Linear Regression Coefficient

We consider a regression example under a simplified setting of fMRI experiments. We focus our interest on two null models. The first null model is based on a Gaussian white noise sequence, $\mathbf{y}$, of length $N = 128$. For simplicity, the design matrix X is considered to be a 128×1 matrix composed of 8 segments each consisting of eight -1's and eight 1's. While a subject's response to a series of visual stimulation blocks actually involves a hemodynamic response function, we use the following simplified model to illustrate our main point concerning wavestrap undercoverage. For our purposes, we use the simple linear regression model without an intercept term

$$y_j = x_j\beta + \epsilon_j, \quad j = 1, \ldots, N, \tag{2}$$

where β is the slope coefficient and the ϵ_j's are uncorrelated errors with zero means and equal variances, σ^2. In the simplest null model, β is zero and the ϵ_j's are independent normals with unknown constant variance.

Testing the null $\beta = 0$ is of primary interest in this setting. Following the classic bootstrap algorithm for linear regression in Davison and Hinkley (1997), the estimate of the slope coefficient $\hat{\beta}$ is obtained from fitting the linear model (2) by minimizing the sum of squares $\|\mathbf{y} - X\beta\|^2$. The raw residuals, r_j, are then obtained from $r_j = y_j - x_j\hat{\beta}$. The residuals are centered by subtracting the average of the raw residuals to obtain the zero-mean residuals $\hat{\epsilon}$'s and are then rescaled as recommended in Thombs and Schucany (1990). The $\hat{\epsilon}$'s are then bootstrapped to produce the random resample ϵ^*. After setting $\mathbf{y}^* = X\hat{\beta} + \hat{\epsilon}^*$, the model (2) is fit to $(X, \mathbf{y}^*)$ and bootstrap replicates of the slope coefficients $\hat{\beta}_b^*$ are obtained for $b = 1, \ldots, B$.

The wavestrap approach presumed to be suitable even in the case of correlated errors involves first transforming these zero-mean residuals, $\hat{\epsilon}$'s, to the wavelet domain. Then the wavestrapped residuals, $\hat{\epsilon}_w^*$, are used to produce the resample $\mathbf{y}_w^* = X\hat{\beta} + \hat{\epsilon}_w^*$. The slope coefficients, β_w^*, are obtained by fitting the regression model (2) to $(X, \mathbf{y}_w^*)$.

The second model of interest involves fractional Brownian motion sequences. Such sequences have an inherent long memory property that several authors have detected in fMRI data (Aguirre et al., 1997; Bullmore et al., 2001). That is, we assume a regression model as in (2), but now the ϵ_j's are fractional Brownian noise. For this type of correlated data, the classical bootstrap method is no longer valid, and the wavestrap method is recommended by Bullmore et al. (2001) to obtain resamples. In the simulations that follow, we investigate the performance of the wavestrap method and use the classical bootstrap and PW methods for comparison.

4. Simulation Studies

4.1. *White Noise Sequences*

We first investigate the wavestrap approach in the special case in which the ϵ's are independent. For each sample we also used PW to generate the ideal detail coefficients and eventually obtain $\hat{\beta}_{PW}^*$.

For each simulated white noise sequence, both the wavestrap and parametric wavestrap resampling methods as well as the classical bootstrap were repeated $B = 199$ times (as recommended in Davison and Hinkley, 1997) to obtain approximate 95% confidence intervals on the slope β. The sequences were replicated 1,000 times and the coverages of 1,000 confidence intervals were obtained for each bootstrap method. Bullmore et al. (2001) recommend resampling the first five levels of detail coefficients. In our simulation studies, we evaluated the wavestrap intervals in the case of resampling from the first 5–7 levels of detail coefficients (i.e., letting $J_0 = 5, 6, 7$, respectively). Since Breakspear et al. (2004) recommended reinflating the wavestrap resamples by multiplying $\sqrt{1.1}$ to maintain the autocorrelation level, we show wavestrap results both with and without reinflation. In Table 1, we present the results of several resampling strategies in this setting. Because the independence assumption actually holds, the classical bootstrap approach is applicable and has a coverage of 94.7%, very close to the nominal level. The wavestrap with or without inflation significantly undercovers when five or more levels of detail coefficients are used. It is interesting to note that PW provides coverage close to the nominal 95% in these cases tabled here. In simulation runs not tabled here, the undercoverage of the two methods (wavestrap and parametric wavestrap) is significantly worse due to greater shrinkage for even lower levels of detail coefficients, e.g., $J_0 = 2$ and 3.

4.2. *Fractional Brownian Noise Sequences*

We simulated fractional Brownian noise sequences with Hurst exponent $H = 0.7$ using an algorithm from Percival and Walden (2000). We used the same estimation procedures as in the previous section including the "inappropriate" classical bootstrap on these correlated noises. The subsequent coverage results are shown in

Table 1
Simulation empirical coverage percentages for wavestrapping methods using D(4) and D(8) wavelet bases and for several values of J_0 in white noise sequences. The first column lists the wavelet bases and the second column lists the levels of detail coefficients used in the wavestrap methods. "w/Inflation" means reinflating the wavestrap resamples by $\sqrt{1.1}$ as recommended by Breakspear et al. (2004). The nominal confidence level is 95%. For each entry the nominal standard error from the 1,000 replications is about 0.7%

	Levels of detail coef.	Wavestrap			Classical
	$(1–J_0)$	w/Inflation	w/o Inflation	PW	bootstrap
D(4)	1–5	91.0%	89.1%	95.9%	
	1–6	91.9%	89.4%	94.8%	
	1–7	91.5%	89.8%	95.5%	94.7%
D(8)	1–5	90.8%	90.2%	94.5%	
	1–6	91.5%	89.5%	94.6%	
	1–7	91.0%	89.1%	94.9%	

Table 2. Again, the coverage percentages of the regular wavestrap are well below the nominal 95% for all the settings with and without the inflation. Also, in this case the PW method does not show marked improvement over the regular wavestrap since the assumption of $N(0, 1)$ detail coefficients in each level is not valid. Not surprisingly, the classical bootstrap did not perform well, but actually performed about as well as the wavestrap in this setting.

Table 2
Simulation empirical coverage percentages for wavestrapping methods using D(4) and D(8) wavelet bases and for several values of J_0 in fractional Brownian noise sequences. The first column lists the wavelet bases and the second column lists the levels of detail coefficients used in the wavestrap methods. "w/Inflation" means reinflating the wavestrap resamples by $\sqrt{1.1}$ as recommended by Breakspear et al. (2004). The nominal confidence level is 95%. For each entry the nominal standard error from the 1,000 replications is about 0.7%

	Levels of detail coef.	Wavestrap			Classical
	$(1–J_0)$	w/Inflation	w/o Inflation	PW	bootstrap
D(4)	1–5	90.0%	88.3%	90.6%	
	1–6	90.5%	88.4%	90.9%	
	1–7	90.9%	88.3%	90.6%	89.9%
D(8)	1–5	88.2%	87.4%	92.0%	
	1–6	89.3%	87.5%	92.1%	
	1–7	89.8%	88.0%	92.4%	

5. Concluding Remarks

In the DWT, there are fewer detail coefficients for higher levels. For example, for a series of length $N = 128 = 2^7$, the number N_j of the detail coefficients in the jth level of the DWT is given by $N_j = 128/2^j$, for $j = 1, \ldots, 6$. At the 5th and 6th levels, there are only 4 and 2 detail coefficients, respectively. Those few conditionally fixed coefficients cannot adequately represent random samples from the corresponding distribution. Therefore, the wavestrap fails to yield correct coverage when the nonparametric resample is obtained from the coefficients within each level. We have demonstrated this effect using a parametric wavestrap which uses generated coefficients from the appropriate normal distribution. In the special case of $N(0, 1)$ white noise errors, the PW gave appropriate coverage. However, it is not applicable in general and the coverage for PW based intervals did not provide improvement in the case of fractional Brownian noise. We also applied permutation within each level of wavelet coefficients and found similar results as bootstrapping wavelet coefficients.

The decorrelating property of the DWT seems attractive. Hence, bootstrapping wavelet coefficients of one-dimensional time series or higher-dimensional images have been proposed by several researchers in neuroimaging and other application areas. This article has demonstrated that for a simple linear regression model with normally distributed white noise errors, the wavestrap produces undercoverage of the confidence intervals of the slope coefficient, which is of primary scientific interest. This result comes from underestimation of the sampling variation of the slope coefficient in the fitted model. Thus, coverage on regression coefficients from wavestrap resamples may not be accurate. This defect extends to more complicated data such as fractional Brownian noise as well.

A technique analogous to the wild bootstrap (Mammen, 1993) for regression residuals may correct for the shrinkage inherent in the wavestrap. We also applied the one-dimensional wavelet packet resampling method using DWPT from the *R* package known as Waveslim, developed by Brandon R. Whitcher, to obtain the coverage of the regression parameter β with white noise and fractional Brownian noise errors. We used the wavelet basis of D(8) and got 94.5% coverage for white noise and 91.5% coverage for the fractional Brownian motion process. It seems that wavelet packet resampling methods (which sometimes are also referred to as wavestrapping methods) might be able to improve the coverage, but more development is needed for correlated observations.

In summary, the wavestrap procedures based on nonparametric resampling methods are not reliable for obtaining confidence intervals or hypothesis tests for regression coefficients with either uncorrelated or correlated errors and should be used with caution. This is the case even when the inflation factor recommended by Breakspear et al. (2004) is applied.

Acknowledgment

The authors would like to thank an associate editor and two referees for their constructive comments and suggestions. This study was supported by the U.S. Army Medical Research and Material Command numbers DAMD17-97-7-7024 and DAMD17-01-1-0741 through a consortium agreement with the University of Texas Southwestern Medical Center at Dallas, U.S. Public Health Service grant MO1-RR00633, and by a grant from the Perot Foundation, Dallas, Texas. The content of this paper does not necessarity reflect the position or the policy of the U.S. government, and no official endorsement should be inferred.

References

Aguirre, G., Zarahn, E., DEsposito, M. (1997). Empirical analyses of BOLD fMRI statistics II. Spatially smoothed data collected under null hypothesis and experimental conditions. *NeuroImage* 5:199–212.

Breakspear, M., Brammer, M., Bullmore, E., Das, P., Williams, L. (2004). Spatio-temporal wavelet resampling for functional neuroimaging data. *Human Brain Mapping* 23:1–25.

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

Davison, A. C., Hinkley, D. V. (1997). *Bootstrap Methods and Their Application*. Cambridge: Cambridge University Press.

Donoho, D. (1993). Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and noisy data. *Proceedings of Symposia in Applied Mathematics* 47:173–205.

Donoho, D. (1995). De-noising by soft-thresholding. *IEEE Transactions on Information Theory* 41:613–627.

Efron, B., Tibshirani, R. (1993). *An Introduction to the Bootstrap*. London: Chappman & Hall.

Feng, H., Willemain, T., Shang, N. (2005). Wavelet-based bootstrap for time series analysis. *Communications in Statistics—Simulation and Computation* 34:393–413.

Flandrin, P. (1992). Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Transactions on Information Theory* 38:910–917.

Frazier, M., Jawerth, B., Weiss, G. (1991). *Littlewood-Paley Theory and the Study of Function Spaces*. Providence: American Mathematical Society.

Mallat, S. (1989a). Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$. *Transactions of the American Mathematical Society* 315:69–87.

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

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics* 21:255–285.

Percival, D. B., Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge, UK: Cambridge University Press. Wavestrappingtime series: Nonstationary Signal Eds.

Ramanathan, J., Zeitouni, O. (1991). On the wavelet transform of fractional Brownian motion. *IEEE Transactions on Information Theory* 37:1156–1158.

Shen, X., Huang, H. C., Cressie, N. (2002). Nonparametric hypothesis testing for a spatial signal. *Journal of the American Statistical Association* 97:1122–1140.

Tewfik, A. H., Kim, M. (1992). Correlation structure of the discrete wavelet coefficients of fractional Brownian motion. *IEEE Transactions on Information Theory* 38:904–909.

Thombs, A. L., Schucany, R. W. (1990). Bootstrap prediction intervals for autoregression. *Journal of the American Statistical Association* 85:486–492.

Zhang, J., Walter, G. (1995). A wavelet-based KL-like expansion for wide-sense stationary random processes. *IEEE Transactions on Signal Processing* 42:1737–1745.