Report 99-012

Wavecov: An S-Plus software package for the analysis of univariate and bivariate time series using wavelets

Brandon Whitcher

ISSN: 1389–2355

WaveCov: An S-Plus software package for the analysis of univariate and bivariate time series using wavelets

Brandon Whitcher
EURANDOM, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

March 5, 1999

Abstract

We introduce a suite of S-Plus functions and C code in order to perform a wavelet analysis of covariance and correlation between two time series. The necessary functions are described and used to analyze the recorded height of ocean waves measured by two different instruments in Cape Henry near Virginia Beach, Virginia. The bivariate spectral analysis of Percival (1994) is also provided in order to compare and contrast the wavelet methodology with classical Fourier techniques.

Some key words: Cross-correlation; Cross-covariance; Discrete wavelet transform; Spectral analysis; Time series.

1 Introduction

The wavelet transform has been used extensively in the various fields of Statistics, such as, non-parametric regression, classification, and time series analysis. With respect to the latter field, the emphasis has been mostly on univariate time series. The discrete wavelet transform (DWT) was introduced by Lindsay, Percival, and Rothrock (1996) in order to analyze bivariate time series and expanded upon in Whitcher (1998). Several software routines were developed while working on my thesis and are provided here so that other researchers may easily implement a wavelet analysis of covariance for bivariate time series.

Two example data sets are used to illustrate the implementation of these routines and were taken from Percival (1994). They are available over the World Wide Web at

```
http://lib.stat.cmu.edu/datasets/saubts
```

or through the WaveCov software package at

```
http://lib.stat.cmu.edu/S/wavecov
```

After downloading the software, follow the instructions on unpacking it as given in the `shar` archive. First, the C code must be compiled and included as a shared library into S-Plus. This is done using the following commands:

```
## Code to make a shared library in S-PLUS Version 3.4 Release 1 for
## Silicon Graphics Iris, IRIX 5.3 : 1996
!SHLIB -o Sdwt.so Sdwt.c
progs _ c("dwt", "idwt", "modwt", "imodwt")
dyn.load.shared("./Sdwt.so", symbols=symbol.C(progs))
```

The compilation of the shared library need only be performed once, but the inclusion of the shared library using `dyn.load.shared()` must be performed each time an S-Plus session is instigated (although one may include this line in the `.First()` function).

The two series, of length $N = 4096$, are displayed in Figure 1 and are a recorded measurement of the height of ocean waves as a function of time by two different instruments. One instrument was a wire wave gauge, while the other was an infrared wave gauge. The sampling frequency for both instruments was 30 samples per second, so the sampling period is $\Delta t = 1/30$ second.

2 Analysis of Variance

2.1 Spectral Techniques

Before performing a bivariate analysis, either spectral or wavelet, let us first look at each time series individually. The usual quantity to measure the variance of a time series is the spectral density function, as estimated by the periodogram (Percival and Walden 1993, Ch. 6). Suppose the time series X_t, $t = 1, \ldots, N$, is a realization of a portion of a zero mean stationary process with sdf $S(\cdot)$ and autocovariance sequence $\{s_\tau\}$. Let $\{\hat{s}_\tau^{(p)}\} \longleftrightarrow \widehat{S}^{(p)}(\cdot)$, where $\hat{s}_\tau^{(p)}$ is the usual biased estimator of the autocovariance sequence; i.e.,

$$\hat{s}_\tau^{(p)} \equiv \frac{1}{N} \sum_{t=1}^{N-|\tau|} X_t X_{t+|\tau|} \quad \text{for} \quad 0 \le \tau \le N-1,$$

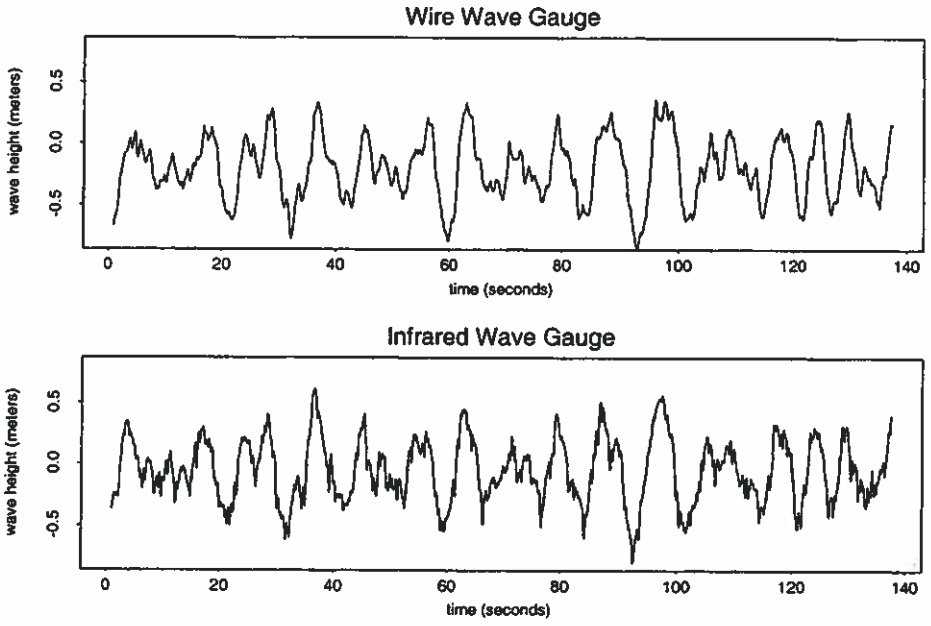

Figure 1: Plot of height of ocean waves versus time as measured by a wire wave gauge and an infrared wave gauge. Both series were collected at a rate of 30 samples per second. There are $N = 4096$ values in each series.

and $\hat{s}_\tau^{(p)} = 0$ for $|\tau| \geq N$. The method of moments spectral estimator is the *periodogram*

$$\widehat{S}_X^{(p)}(f) \equiv \frac{\Delta t}{N} \left| \sum_{t=1}^{N} X_t e^{-i2\pi f t \Delta t} \right|^2 .$$

The periodogram has its share of problems, so it may be more appropriate to use an alternative to the periodogram, such as a multitaper spectral estimator

$$\widehat{S}_X^{(mt)}(f) = \frac{1}{K} \sum_{k=0}^{K-1} \widehat{S}_{k,X}^{(mt)}(f) \quad \text{with} \quad \widehat{S}_{k,X}^{(mt)}(f) = \Delta t \left| \sum_{t=1}^{N} h_{t,k} X_t e^{-i2\pi f t \Delta t} \right|^2 ,$$

where $\{h_{t,k} \mid t = 1, \ldots, N\}$, k ranges from 0 to $K-1$, is a set of K orthonormal data tapers; i.e., $\sum_{t=1}^{N} h_{t,j} h_{t,k} = 1$ if $j = k$ and 0 if $j \neq k$. Examples of common data tapers are the sine tapers (Riedel and Sidorenko 1995) and discrete prolate spheroidal sequences (dpss) data tapers (Slepian 1978; Thomson 1982; Percival and Walden 1993, Ch. 8). Sine tapers were designed to minimize the spectral window bias and can be approximated well using the following closed form expression

$$h_{t,k}^{(sine)} = \left\{ \frac{2}{N+1} \right\}^{\frac{1}{2}} \sin \left\{ \frac{(k+1)\pi t}{N+1} \right\} .$$

In contrast, the dpss data tapers minimize the spectral window sidelobes, using a resolution bandwidth parameter W, and must be calculated using techniques such as inverse iteration, numerical integration or a tridiagonal formulation (Percival and Walden 1993, Ch. 8). The role of any data taper is to protect against leakage, and all the sine tapers provide moderate leakage protection where the dpss data tapers offer adjustable leakage protection through the parameter W. In practice there is little difference in the multitaper spectral estimators when using either data taper. Figure 2 shows the multitaper spectral estimates for the wire and infrared wave gauges using 6 dpss data tapers ($NW = 4$).

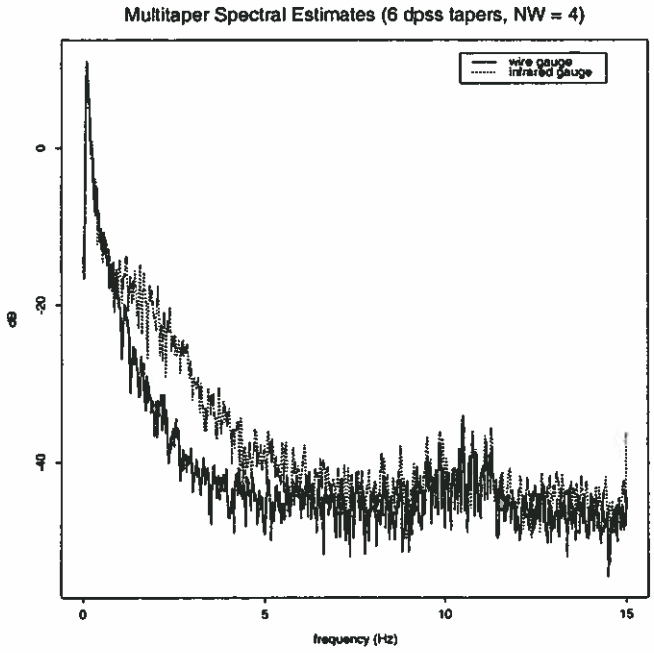

Figure 2: Multitaper spectral estimate $\widehat{S}_X^{(mt)}(\cdot)$ for the wire (solid line) and infrared (dotted line) wave gauges based upon 6 dpss data tapers.

2.2 Wavelet Techniques

It has been shown that the wavelet variance $\nu_X^2(\lambda_j)$ can decompose the variance of a time series on a scale by scale basis, instead of the frequency by frequency basis used the spectrum (Percival 1995). The wavelet variance is defined to be the variance of the wavelet coefficients associated with scale $\lambda_j \equiv 2^{j-1}$. This is equivalent to the expected value of the squared wavelet coefficients, since they are assumed to have zero mean (this is assured if the number of vanishing moments for the wavelet filter is chosen to be large enough). The DWT coefficients of $X_1, \ldots, X_N$ are denoted by $W_{j,t}^{(X)}$ for $j = 1, \ldots, J$ and $t = 1, \ldots, N/2^j$. The wavelet variance is estimated using the DWT coefficients for scale $\lambda_j \equiv 2^{j-1}$ via

$$\hat{\nu}_X^2(\lambda_j) \equiv \frac{1}{\widehat{N}_j} \sum_{t=L_j'}^{N} \left[W_{j,t}^{(X)}\right]^2,$$

where $\widehat{N}_j = N/2^j - L_j'$ and $L_j' = \lceil (L-2)(1-2^{-j}) \rceil$. Constructing an estimator of the wavelet variance using a variation of the DWT, the maximal overlap DWT (MODWT), has been shown to be superior to that of the DWT-based estimator (Percival 1995). The MODWT coefficients of $X_1, \ldots, X_N$ are denoted by $\widetilde{W}_{j,t}^{(X)}$ for $j = 1, \ldots, J$ and $t = 1, \ldots, N$. The wavelet variance estimated by the MODWT coefficients for scale λ_j is given by

$$\tilde{\nu}_X^2(\lambda_j) \equiv \frac{1}{\widetilde{N}_j} \sum_{t=L_j}^{N} \left[\widetilde{W}_{j,t}^{(X)}\right]^2,$$

where $\widetilde{N}_j = N - L_j$ and $L_j = (2^j - 1)(L-1)$. Figure 3 gives the MODWT estimator of wavelet variance for the wire and infrared wave gauges using the Daubechies extremal phase wavelet filter

of length $L = 4$ (Daubechies 1992, Ch. 6). We will denote this wavelet filter by D(4) from now on. It was computed in S-Plus by first obtaining the MODWT coefficients of each time series, applying "brick wall" boundary conditions, and then applying the wavelet variance estimating procedure:

```
## Compute the MODWT for each series using the D(4) wavelet filter
wire.modwt.d4 _ modwt(wire, wavelet = "d4", n.levels = 9)
ir.modwt.d4 _ modwt(ir, wavelet = "d4", n.levels = 9)

## Impose "brick wall" boundary conditions
wire.modwt.d4.bw _ modwt.brick.wall(wire.modwt.d4, wavelet = "d4", N = 4096)
ir.modwt.d4.bw _ modwt.brick.wall(ir.modwt.d4, wavelet = "d4", N = 4096)

## Compute the wavelet variance for each series
wire.modwt.d4.var _ wave.var(wire.modwt.d4.bw)
ir.modwt.d4.var _ wave.var(ir.modwt.d4.bw)
```

A special plotting routine `plot.var()` was used to produce this particular figure (its usage can be found in the S-Plus code provided). Notice that the infrared wave gauge exhibits significantly greater wavelet variance for the first four scales, associated with changes of 1/30, 2/30, 4/30, and 8/30 second.

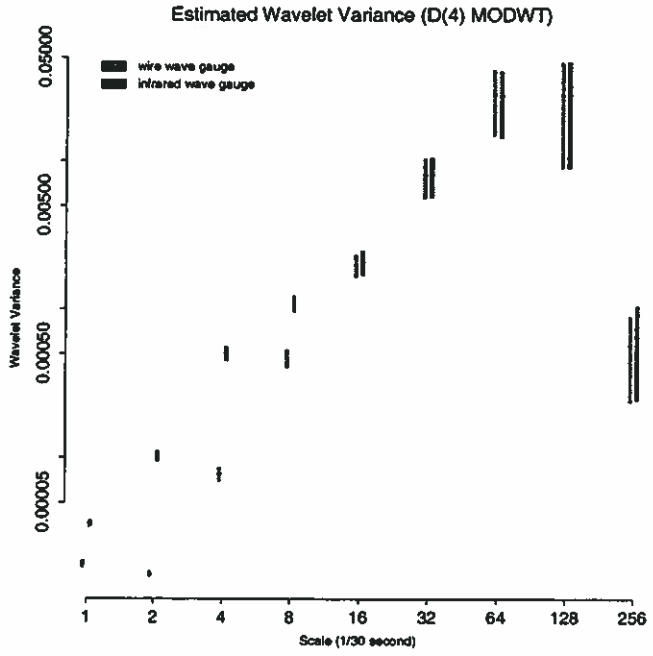

Figure 3: MODWT estimator of the wavelet variance $\tilde{\nu}_X^2(\lambda_j)$ using a D(4) wavelet filter. The shaded bars give approximate 95% confidence intervals on the estimates. The x-axis is plotted in decreasing frequencies so that scales associated with high-frequency are at the left and scales associated with low frequency ($f \to 0$) are at the right.

3 Analysis of Covariance

3.1 Spectral Techniques

The following material closely follows an introduction to bivariate spectral analysis in Percival (1994), and is a natural extension of univariate topics found in Percival and Walden (1993) using similar notation. A more thorough introduction to multivariate spectral analysis can be found in, for example, Koopmans (1974), Priestley (1981) and Brillinger (1981).

Let X_t, Y_t, $t = 1, \ldots, N$, be a realization of a portion of a zero mean stationary process $\{X_t, Y_t\}$ with cross spectrum $S_{XY}(\cdot)$ and autospectra $S_X(\cdot)$ and $S_Y(\cdot)$, respectively. Just as the periodogram was used in the univariate case (Section 2), the *cross periodogram*

$$\widehat{S}_{XY}^{(p)}(f) = \sum_{\tau=-(N-1)}^{N-1} \widehat{C}_{\tau,XY} e^{-i2\pi f\tau}$$

is utilized here to estimate the cross spectrum. The *sample cross covariance sequence* is defined to be

$$\widehat{C}_{\tau,XY} \equiv \sum_t X_t Y_{t+\tau},$$

where the summation goes from $t = 1$ to $N - \tau$ for $\tau \geq 0$ and from $t = 1 - \tau$ to N for $\tau < 0$. The cross periodogram can also be written in a more computationally friendly form as

$$\widehat{S}_{XY}^{(p)}(f) = \frac{1}{N} \left(\sum_{t=1}^{N} X_t e^{-i2\pi ft} \right)^* \left(\sum_{t=1}^{N} Y_t e^{-i2\pi ft} \right),$$

where the asterisk denotes complex conjugation.

The multitaper estimator of the cross spectrum is given by

$$\widehat{S}_{XY}^{(mt)}(f) = \frac{1}{K} \left(\sum_{t=1}^{N} h_{k,t} X_t e^{-i2\pi ft} \right)^* \left(\sum_{t=1}^{N} h_{k,t} Y_t e^{-i2\pi ft} \right),$$

where $\{h_{k,t}\}$ is the kth-order data taper for a sequence of length N normalized such that $\sum_t h_{k,t}^2 = 1$, $k = 1, \ldots, K$ (c.f. Section 2). Thus, the multitaper estimators for the phase spectrum and magnitude squared coherence are given by

$$\hat{\phi}_{XY}^{(mt)}(f) = \arg\left\{ \widehat{S}_{XY}^{(mt)}(f) \right\} \quad \text{and} \quad \left| \hat{w}_{XY}^{(mt)}(f) \right|^2 = \frac{\left| \widehat{S}_{XY}^{(mt)}(f) \right|^2}{\widehat{S}_X^{(mt)}(f) \widehat{S}_Y^{(mt)}(f)},$$

respectively. The phase spectrum $\hat{\phi}_{XY}^{(mt)}(\cdot)$ takes on values between $-\pi$ and π and, hence, is modulo 2π. This can lead to discontinuities around $\pm\pi$. Priestley (1981, p. 709) describes a method to avoid these discontinuities - by simultaneously plotting the original estimate and translated versions of it.

3.2 Wavelet Covariance and Cross-covariance

Let X_t and Y_t, $t = 1, \ldots, N$ be defined as before. For $N \geq L_j$, we can define an unbiased estimator $\tilde{\gamma}_{XY}(\lambda_j)$ of the wavelet covariance based upon the MODWT via

$$\tilde{\gamma}_{XY}(\lambda_j) \equiv \frac{1}{\tilde{N}_j} \sum_{t=L_j}^{N} \widetilde{W}_{j,t}^{(X)} \widetilde{W}_{j,t}^{(Y)}.$$

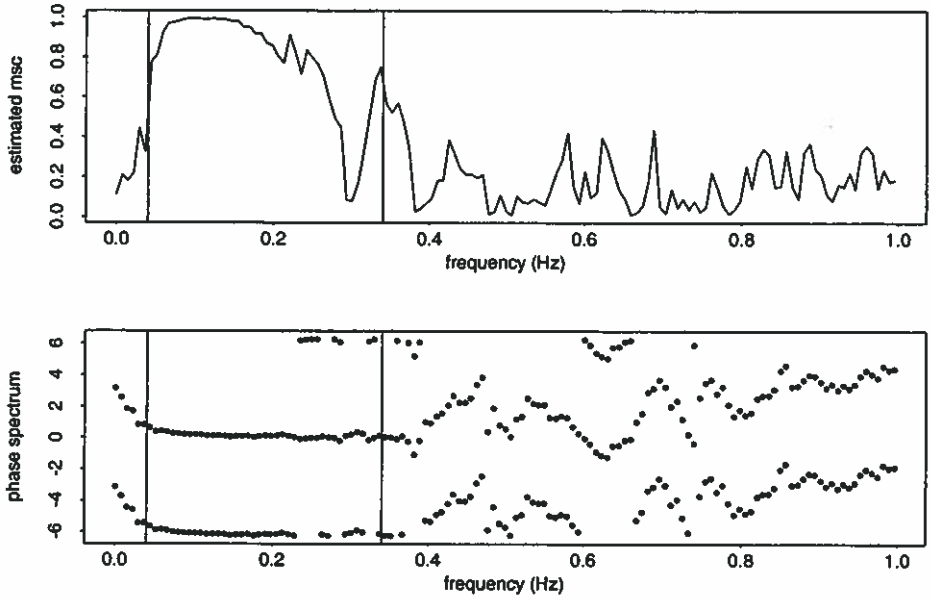

Figure 4: The estimated msc $|\hat{w}_{XY}^{(mt)}(\cdot)|$ and estimated phase spectrum $\hat{\phi}_{XY}^{(mt)}(\cdot)$ versus frequency ($0 \leq f \leq 1$ Hz) between the wire and infrared wave gauges using 6 dpss data tapers. The two vertical lines are at $f = 0.04$ Hz and $f = 0.34$ Hz, a frequency band of interest in Percival (1994). The phase spectrum is plotted using Priestley's recommendation.

Note, the estimator does not include any coefficients that make explicit use of the periodic boundary conditions. We can construct a biased estimator of the wavelet covariance by simply including the MODWT wavelet coefficients affected by the boundary and renormalizing.

If we are strictly interested in Gaussian processes, we can re-express the variance of the wavelet covariance (see Whitcher (1998) for details) by

$$\mathrm{Var}\{\tilde{\gamma}_{XY}(\lambda_j)\} \approx \frac{\mathcal{V}_j}{\tilde{N}_j},$$

for large $\tilde{N}_j$, where

$$\mathcal{V}_j \equiv \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{j,X}(f) S_{j,Y}(f)\, df + \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{j,XY}^2(f)\, df. \tag{1}$$

This will allow us to easily construct approximate confidence intervals for the MODWT estimator of the wavelet covariance. For non-Gaussian processes, Serroukh and Walden (1998) use a multitaper spectral estimator at frequency zero of the product of the two processes $S_{j,XY}(0)$ to obtain $\mathcal{V}_j$.

Figure 5 shows the MODWT estimator of wavelet covariance between the wire and infrared wave gauges. They were computed using the previously obtained MODWT coefficients with "brick wall" boundary conditions:

```
## Compute MODWT estimator of wavelet covariance
wire.ir.modwt.d4.cov _ wave.cov(wire.modwt.d4.bw, ir.modwt.d4.bw)
```

A special plotting routine `plot.cov()` was used to produce the figure. The two signals are virtually unrelated for small scales (high frequencies) and become highly associated for scales of 16/30 to 128/30 seconds. This corresponds with the results from the analysis of wavelet variance.

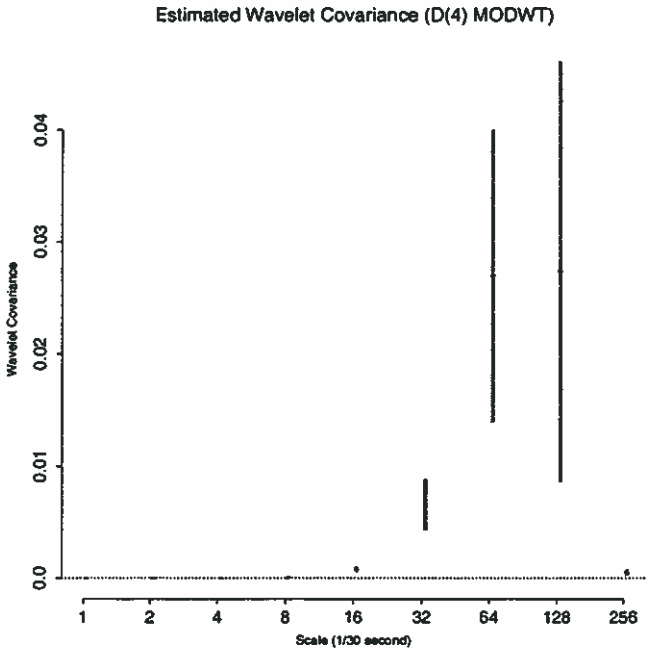

Figure 5: MODWT estimators of wavelet covariance $\tilde{\gamma}_{XY}(\lambda_j)$ for the wire and infrared wave gauges with approximate 95% confidence intervals.

Estimation of the wavelet cross-covariance follows directly from the biased estimator of the usual cross-covariance (Priestley 1981, pp. 692–693). For $N \geq L_j$, we can define a biased estimator $\tilde{\gamma}_{\tau,XY}(\lambda_j)$ of the wavelet covariance based upon the MODWT via

$$\tilde{\gamma}_{\tau,XY}(\lambda_j) \equiv \begin{cases} \frac{1}{\tilde{N}_j} \sum_{t=L_j}^{N-\tau} \widetilde{W}_{j,t}^{(X)} \widetilde{W}_{j,t+\tau}^{(Y)}, & \tau = 0, \ldots, \tilde{N}_j - 1; \\ \frac{1}{\tilde{N}_j} \sum_{t=L_j-\tau}^{N} \widetilde{W}_{j,t}^{(X)} \widetilde{W}_{j,t+\tau}^{(Y)}, & \tau = -1, \ldots, -(\tilde{N}_j - 1); \\ 0, & \text{otherwise.} \end{cases}$$

The bias is due to the denominator $1/\tilde{N}_j$ remaining constant for all lags. We are still not using wavelet coefficients which make use of the periodic boundary conditions.

The formula for an approximate $100(1-2p)\%$ confidence interval of the MODWT estimator of the wavelet covariance was established (although incorrectly, as previously noted in Serroukh and Walden (1998)) in Lindsay, Percival, and Rothrock (1996). First, the periodogram and cross-periodogram are used to produce estimators of the univariate and bivariate spectra given in Equation (1). Next, define the biased estimator of the autocovariance sequence associated with the scale λ_j MODWT wavelet coefficients of $\{X_t\}$ by

$$\hat{s}_{j,\tau,X}^{(p)} \equiv \frac{1}{\tilde{N}_j} \sum_{t=L_j-1}^{N-1-|\tau|} \widetilde{W}_{j,t}^{(X)} \widetilde{W}_{j,t+|\tau|}^{(X)},$$

with a similar definition for $\{\hat{s}_{j,\tau,Y}^{(p)}\}$. The corresponding biased estimator of the cross covariance sequence associated with the scale λ_j MODWT wavelet coefficients of $\{X_t, Y_t\}$ by

$$\hat{C}_{j,\tau,XY}^{(p)} \equiv \frac{1}{\tilde{N}_j} \sum_{t} \widetilde{W}_{j,t}^{(X)} \widetilde{W}_{j,t+\tau}^{(Y)},$$

where the summation goes from $t = L_j - 1$ to $N - 1 - \tau$ for $\tau \geq 0$ and from $t = L_j - 1 - \tau$ to $N - 1$ for $\tau < 0$. We appeal to Parseval's relation to obtain an estimator for $\mathcal{V}_j$ that uses only the autocovariance and cross-covariance sequences instead of the autospectra and cross spectrum; i.e.

$$\tilde{\mathcal{V}}_j \equiv \frac{\hat{s}^{(p)}_{j,0,X}\hat{s}^{(p)}_{j,0,Y}}{2} + \sum_{\tau=1}^{\tilde{N}_j-1} \hat{s}^{(p)}_{j,\tau,X}\hat{s}^{(p)}_{j,\tau,Y} + \frac{1}{2}\sum_{\tau=-(\tilde{N}_j-1)}^{\tilde{N}_j-1} \left[\hat{C}^{(p)}_{j,\tau,XY}\right]^2.$$

The estimator $\tilde{\mathcal{V}}_j$ is unbiased when the periodogram is used to estimate the spectra (Whitcher 1998). Under the assumption that the spectral estimates are close to the true values, an approximate $100(1-2p)\%$ confidence interval for $\gamma_{XY}(\lambda_j)$ is

$$\left[\tilde{\gamma}_{XY}(\lambda_j) - \Phi^{-1}(1-p)\sqrt{\frac{\tilde{\mathcal{V}}_j}{\tilde{N}_j}}\,,\;\; \tilde{\gamma}_{XY}(\lambda_j) + \Phi^{-1}(1-p)\sqrt{\frac{\tilde{\mathcal{V}}_j}{\tilde{N}_j}}\right],$$

where $\Phi^{-1}(p)$ is the $p \times 100\%$ percentage point for the standard normal distribution.

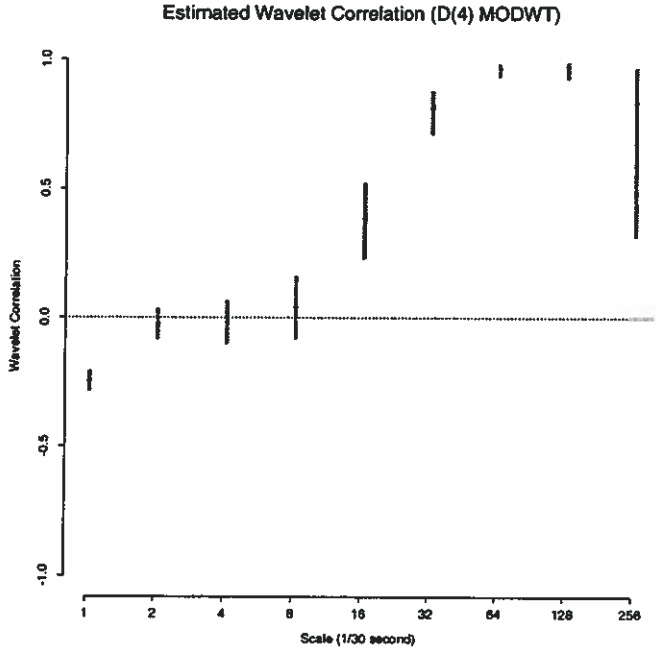

Figure 6: MODWT estimators of wavelet correlation $\tilde{\rho}_{XY}(\lambda_j)$ for the wire and infrared wave gauges with approximate 95% confidence intervals using Fisher's z-transformation.

3.3 Wavelet Correlation and Cross-correlation

Given the covariance doesn't take into account the variation of the univariate time series, a natural next-step is to introduce the concept of wavelet correlation. As with the usual estimator for correlation in time series, the wavelet correlation is simply made up of the wavelet covariance for $\{X_t, Y_t\}$ and wavelet variances for $\{X_t\}$ and $\{Y_t\}$, the MODWT estimator of the wavelet cross-correlation is simply

$$\tilde{\rho}_{\tau,XY}(\lambda_j) \equiv \frac{\tilde{\gamma}_{\tau,XY}(\lambda_j)}{\tilde{\nu}_X(\lambda_j)\tilde{\nu}_Y(\lambda_j)},$$

where $\tilde{\gamma}_{\tau,XY}(\lambda_j)$ is the wavelet covariance, and $\tilde{\nu}^2_X(\lambda_j)$ and $\tilde{\nu}^2_Y(\lambda_j)$ are the wavelet variances. When $\tau = 0$ we obtain the MODWT estimator of the wavelet correlation between $\{X_t, Y_t\}$.

Given the non-normality of the correlation coefficient for small sample sizes, a nonlinear transformation is sometimes required – Fisher's z-transformation (Fisher 1915; Kotz, Johnson, and Read 1982, Volume 3). Let

$$h(\rho) \equiv \frac{1}{2}\log\left(\frac{1+\rho}{1-\rho}\right) = \tanh^{-1}(\rho)$$

define the transformation. For the estimated correlation coefficient $\hat{\rho}$, based on n independent samples, $\sqrt{n-3}(h(\hat{\rho}) - h(\rho))$ has approximately a $N(0,1)$ distribution.

An approximate $100(1-2p)\%$ confidence interval for $\rho_{XY}(\lambda_j)$ based on the MODWT is given by

$$\left[\tanh\left\{h[\tilde{\rho}_{XY}(\lambda_j)] - \frac{\Phi^{-1}(1-p)}{\sqrt{\widehat{N}_j - 3}}\right\},\ \tanh\left\{h[\tilde{\rho}_{XY}(\lambda_j)] + \frac{\Phi^{-1}(1-p)}{\sqrt{\widehat{N}_j - 3}}\right\}\right]$$

where $\widehat{N}_j = N_j - L'_j$ and $L'_j = \lceil (L-2)(1-2^{-j}) \rceil$ is the number of DWT wavelet coefficients associated with scale λ_j. Note that we are using the number of wavelet coefficients as if the point estimates has been computed using the DWT. This is because, under the assumptions of Fisher's z-transformation, the denominator should consist of the number of independent samples used in constructing the correlation coefficient. Since we are using the MODWT wavelet coefficients this assumption is clearly violated. If the DWT wavelet coefficients were utilized instead, then we could reasonably assume independence of the wavelet coefficients within each scale.

Figure 6 gives the MODWT estimator of wavelet correlation (lag 0) between the wire and infrared wave gauges. The approximate 95% confidence intervals were computed using Fisher's z-transformation. This is implemented using the following line of S-Plus code:

```
wire.ir.modwt.d4.cor _ wave.cor(wire.modwt.d4.bw, ir.modwt.d4.bw, N = 4096)
```

We now see a much more informative picture of the association between the two wave gauges. The first scale (1/30 second) now exhibits significant negative correlation, while scales from 16/30 second on exhibit significant positive correlations. Scales of 64/30 and 128/30 second show the highest amount of association, as with Figure 5, but the confidence intervals are significantly decreased given the amount of variability in the estimated wavelet variances. Because of the non-linear transform, these confidence intervals need not be symmetric.

Figure 7 gives the MODWT estimator of wavelet cross-correlation (± 1024 lags) between the wire and infrared wave gauges. The following S-Plus code establishes a maximum number of lags to be estimated, then a for-loop is initiated for scales $j = 1, \ldots, 9$ using `spin.cor()` (the final two lines makes the output into a nice data frame with appropriate names):

```
lmax _ 1024; J _ 9
wire.ir.cross.cor _ NULL
for(i in 1:J) {
  blah _ spin.cor(as.data.frame(wire.modwt.d4.bw)[,i],
                  as.data.frame(ir.modwt.d4.bw)[,i], lmax)
  wire.ir.cross.cor _ cbind(wire.ir.cross.cor, blah)
}
wire.ir.cross.cor _ as.data.frame(wire.ir.cross.cor)
names(wire.ir.cross.cor) _ paste("d", 1:J, sep="")
```

The figure was plotted using `stack.plot()` (a standard function provided with S+WAVELETS) with the option `same.scale=T` so that all plots are put on the same y-axis. From the plot of wavelet

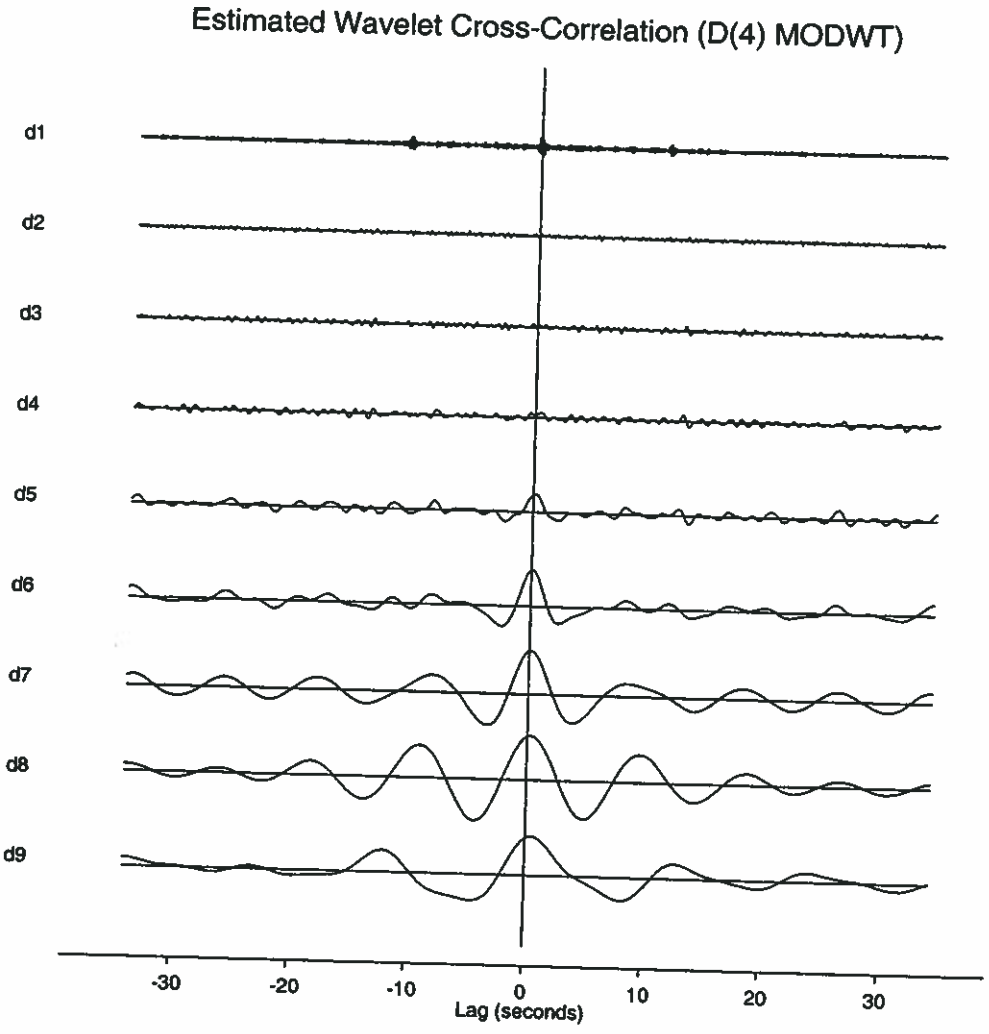

Figure 7: MODWT estimator of wavelet cross-correlation $\tilde{\rho}_{XY}(\lambda_j)$ between the wire and infrared wave gauges.

correlation, the wavelet cross-correlation for scales 2/30, 4/30 and 8/30 second ($d1, d2, d3$ in the figure) show little association. The first scale (1/30 second) contains a burst of non-zero coefficients at small positive and negative lags and then symmetric bursts around ± 11 seconds. This is not visible through bivariate spectral analysis. For higher scales, the wavelet cross-correlation is an approximately even function with the exception of scale λ_9.

Percival (1994) used the estimated phase spectrum to determine a delay between the two wave gauges of approximately $-5/30$ seconds. The likely reason for this was that the wave gauges were 6 meters apart and the prevalent direction of the waves being approximately perpendicular to a line drawn between the two gauges. The appropriate scale(s) of wavelet coefficients is λ_7 and λ_8 which correspond with pass-bands $0.12 \leq f \leq 0.23$ and $0.6 \leq f \leq 0.12$ Hz, respectively. The lags where a maximum is attained is +4 units (4/30 seconds) for λ_7 and +8 units (8/30 seconds) for λ_8. The sign change is due to the reversal of the wire and infrared series in our analysis with respect to Percival (1994).

4 Conclusions

We have introduced a suite of wavelet estimators for the variance, covariance, and correlation for bivariate time series. Approximate confidence intervals are also provided for all three types of estimator. Some S+WAVELETS functions are utilized in order to expedite computations and also display some of the results. These functions will hopefully allow researchers in any field to more easily analyze time series using wavelets.

This research was performed while the author was a graduate student at the University of Washington, Department of Statistics, under the supervision of Peter Guttorp and Donald B. Percival.

References

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory.* Holden-Day Series in Time Series Analysis. San Francisco: Holden-Day. Expanded edition.

Bruce, A. and H.-Y. Gao (1996). *Applied Wavelet Analysis with S-PLUS.* New York: Springer.

Daubechies, I. (1992). *Ten Lectures on Wavelets*, Volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics.* Philadelphia: Society for Industrial and Applied Mathematics.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika 10*, 507–521.

Koopmans, L. H. (1974). *The Spectral Analysis of Time Series.* New York. Academic Press.

Kotz, S., N. L. Johnson, and C. B. Read (Eds.) (1982). *Encyclopedia of Statistical Sciences.* New York: Wiley.

Lindsay, R. W., D. B. Percival, and D. A. Rothrock (1996). The discrete wavelet transform and the scale analysis of the surface properties of sea ice. *IEEE Transactions on Geoscience and Remote Sensing 34*(3), 771–787.

Percival, D. B. (1994). Spectral analysis of univariate and bivariate time series. In J. L. Stanford and S. B. Vardeman (Eds.), *Statistical Methods for Physical Science*, Volume 28 of *Methods of Experimental Physics*, pp. 313–348. Boston: Academic Press, Inc.

Percival, D. B. (1995). On estimation of the wavelet variance. *Biometrika 82*(3), 619–631.

Percival, D. B. and A. T. Walden (1993). *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques.* Cambridge: Cambridge University Press.

Priestley, M. B. (1981). *Spectral Analysis and Time Series.* London: Academic Press, Inc.

Riedel, K. S. and A. Sidorenko (1995). Minimum bias multiple taper spectral estimation. *IEEE Transactions on Signal Processing 43*(1), 188–195.

Serroukh, A. and A. T. Walden (1998). The scale analysis of bivariate non-Gaussian time series via wavelet cross-covariance. Technical Report 98–02, Department of Mathematics, Imperial College of Science, Technology & Medicine.

Slepian, D. (1978). Prolate spheroidal wave functions, Fourier analysis, and uncetainty – V: The discrete case. *Bell System Technical Journal 57*, 1371–1430.

Thomson, D. J. (1982). Spectrum estimation and harmonic analysis. *IEEE Proceedings 70*(9), 1055–1096.

Whitcher, B. (1998). *Assessing Nonstationary Time Series Using Wavelets.* Ph. D. thesis, University of Washington.