

# An Introduction to Wavelet Analysis with Applications to Vegetation Time Series

Donald B. Percival<sup>1</sup>, Muyin Wang<sup>2</sup> and James E. Overland<sup>3</sup>

<sup>1</sup>corresponding author  
Applied Physics Laboratory  
Box 355640  
University of Washington  
Seattle, WA 98195-5640  
dbp@apl.washington.edu  
206.543.1368

<sup>2</sup>Joint Institute for the Study of the Atmosphere and Ocean  
Box 357941  
University of Washington  
Seattle, WA 98195-4235

<sup>3</sup>Pacific Marine Environmental Laboratory/NOAA  
7600 Sand Point Way NE  
Seattle, WA 98115-6349

For submission to *Community Ecology*

June 24, 2004

Contribution No. 2647 from Pacific Marine Environmental Laboratory/NOAA

Contribution No. 1032 from Joint Institute for the Study of the Atmosphere and Ocean

**Keywords:** Discrete Wavelet Transform, Köppen Classification, Multiresolution Analysis, Time Series Analysis.

**Abstract:** Wavelets are relatively new mathematical tools that have proven to be quite useful for analyzing time series and spatial data. We provide a basic introduction to wavelet analysis that concentrates on their interpretation in the context of analyzing time series. We illustrate the use of wavelet analysis on time series related to vegetation coverage in the Arctic region.

**Abbreviations:** CWT–Continuous Wavelet Transform, DWT–Discrete Wavelet Transform, FD–Fractionally Differenced MODWT–Maximal Overlap Discrete Wavelet Transform, MRA–Multiresolution Analysis.

## 1 Introduction

Since their introduction in the geophysical literature in the 1980s (Goupillaud et al. 1984), wavelets have become an increasingly popular tool for analyzing time series and spatial data. Examples of wavelet analysis are in abundance in almost all areas of science and engineering. A recent search on the INSPEC bibliographic database using the word ‘wavelet’ yielded over 20,000 articles written from 1990 and on. A similar search for articles with ‘wavelet’ and ‘vegetation’ uncovered 50 articles, many of which are directly related to vegetation monitoring. Some key papers that have applied wavelets in vegetation ecology and related fields include Bradshaw and Spies (1992), Brunet and Collineau (1994), Csillag and Kabos (2002), Dale and Mah (1998), Lark and Webster (1999) and Lark and Webster (2001). It is our contention that wavelet analysis is useful in many more areas related to vegetation monitoring than what have already been explored in the literature. The purpose of this article is to provide an introduction to wavelet analysis that concentrates on the ideas behind it in the context of the analysis of time series related to vegetation coverage in the Arctic region. After a brief description of these series and a discussion of what wavelet analysis can tell us about them (§2), we discuss the notion of a wavelet (§3) and the basics of wavelet analysis (§4). We then describe the discrete wavelet transform (§5) and use this transform to analyze the vegetation coverage series (§6). We conclude with a summary of our key points and a small discussion (§7).

## 2 Time Series of Vegetation Coverage

Figure 1 shows time series of four vegetation groups over land between 50°–90° N based on the modified Köppen classification system (Köppen 1931, Trewartha and Horn 1980). Each line in the figure represents the coverage of one classification group (defined below) in the

Arctic region for the period from 1901 to 2000 in units of  $10^6$  square kilometers (each series consists of 100 values in all). Wang and Overland (2004) generated these four time series based on monthly surface air temperature data assembled by the Climate Research Unit, University of East Anglia (CRU TS2.0).

The Köppen classification system uses natural vegetation as an expression of climate, and is based upon annual and monthly means of temperature and precipitation. The classes can be used as surrogates for different types of vegetation. Köppen classification recognizes five major groups of climate and lesser subdivisions and characteristics in a quantitative system. In the modified Köppen classification system (Trewartha and Horn 1980), there are six major groups, five of which are based on the thermic zones. The sixth is the dry group, which cuts across the first four thermic zones. These thermic groups have a strong zonal orientation, and they correspond closely to zonal patterns in vegetation coverage. Over the Arctic region ( $50^{\circ}$ – $90^{\circ}$  N), the polar group (fifth in the classification) is further divided into two subgroups, namely, tundra and perpetual frost. The tundra group is defined as areas over which the monthly mean temperature of the warmest month is less than  $10^{\circ}$  C, but above freezing. Perpetual frost covers the area where the maximum temperature in the summer months is below freezing. Two other groups are also identified in the Arctic region, namely, the boreal and temperate groups. The boreal group has one to three months of mean temperatures above  $10^{\circ}$  C, whereas the temperate group has four or more months above  $10^{\circ}$  C. The polar group is dominated by treeless tundra vegetation. The boreal group is comprised mainly of boreal coniferous forests. The temperate group consists of broadleaf deciduous forests inland and temperate evergreen forests over maritime regions.

Wang and Overland (2004) compared the spatial distribution of the tundra group based on the modified Köppen classification system using temperatures generated by CRU and NCEP (National Centers for Environmental Prediction) with the circumpolar Arctic Vege-

tation Map (CVAM Team 2003) and with the spatial map from National Aeronautics and Space Administration (NASA) satellite Normalized Difference Vegetation Index (NDVI). They found good agreement among the analyses. They also noted that, since the late 1970s, there has been a downward linear trend in three time series representing the tundra group. The first series is based upon Köppen classification using the CRU temperatures; the second, on a similar classification using the NCEP analysis; and the third, on NDVI values of 0.1 to 0.4. The three slopes are quite similar ( $-7.6 \times 10^4$ ,  $-6.8 \times 10^4$  and  $-5.8 \times 10^4$  km<sup>2</sup>/yr, respectively). These trends represent about a 20% reduction in the coverage of the tundra group based on the CRU data from 1980 to 2000. Although Köppen classification is not an actual measurement of area of a given vegetation community, the agreement between the classified tundra group and satellite NDVI indicates that Köppen classification is a good approximation to the actual area of coverage in the Arctic region. In addition, the decrease in coverage by tundra in the last few decades is also supported by photographic comparisons of several places in the Arctic region (Sturm et al. 2001 and Shiyatov 2003).

The following are examples of questions that a wavelet analysis can potentially address and that we return to in §6.

1. Are the statistical variations in a given time series homogeneous across time? More generally, because of their localized nature, wavelets can address questions about other time-dependent variations such as the presence or absence of a trend.
2. In a given series, are the variations from one year to the next more prominent than the variations from one decade to the next? This question is typical of others that wavelets can answer about the dominant factors contributing to the overall variance of a time series.
3. It is obvious from Figure 1 that the boreal time series is significantly more variable

than the frost series, but do they have other statistical properties that are significantly different? Here we exploit the ability of wavelets to compare overall properties of different time series.

4. How are, say, the boreal and temperate series related on a scale by scale basis (e.g., year to year, decade to decade)? This question illustrates that wavelets can provide a useful covariance analysis; i.e., they can tell us how two time series co-vary on various scales across time.

### 3 What is a Wavelet?

In English ‘let’ is sometimes added to the end of a noun to indicate a diminutive version of the concept; e.g., piglet is a small pig (Milne 1926). To understand the term ‘wavelet,’ we need to define what is meant by a wave. Let us take a wave to be a real-valued function that is defined over the entire real axis and that oscillates back and forth about zero, with the amplitude of the oscillations being relatively the same everywhere. The top plot of Figure 2 shows part of a typical wave, namely, a function that is defined as  $\sin(u)$  for  $-\infty < u < \infty$ . By contrast, the bottom three plots show three different wavelets  $\psi(\cdot)$ . The top one is the Haar wavelet, defined by

$$\psi^{(\text{H})}(u) = \begin{cases} -1/\sqrt{2}, & -1 < u \leq 0; \\ 1/\sqrt{2}, & 0 < u \leq 1; \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

The middle wavelet  $\psi^{(\text{rdG})}(\cdot)$  is proportional to the first derivative of the probability density function  $\frac{1}{\sqrt{(2\pi\sigma^2)}}e^{-u^2/2\sigma^2}$  for a normal (Gaussian) random variable with mean zero and variance  $\sigma^2$ , while the bottom wavelet  $\psi^{(\text{Mh})}(\cdot)$  is proportional to the second derivative and, for an obvious reason, is referred to as the ‘Mexican hat’ wavelet. We can deduce from Figure 2

that the sense in which a wavelet is smaller than a wave is that, whereas a wave has substantial fluctuations over the entire real axis, a wavelet is zero (or very close to zero) outside of some finite interval (e.g.,  $(-1, 1]$  for the Haar wavelet and approximately  $(-3, 3]$  for the other two wavelets). Thus, in contrast to waves, the amplitude of a wavelet damps down as  $|u|$  gets large, and a wavelet is said to have finite width.

Technically three conditions are required for a function  $\psi(\cdot)$  to be a wavelet.

1.  $\psi(\cdot)$  must integrate to 0; i.e.,  $\int_{-\infty}^{\infty} \psi(u) du = 0$ . Intuitively this condition means that the oscillations undertaken by a wavelet must be balanced above and below zero.
2.  $\psi^2(\cdot)$  integrates to 1; i.e.,  $\int_{-\infty}^{\infty} \psi^2(u) du = 1$ . This condition implies that, for every small  $\epsilon > 0$ , there is an interval  $(-T, T]$  of finite width such that  $\int_{-T}^T \psi^2(u) du > 1 - \epsilon$ , i.e., that  $\int_{-\infty}^{-T} \psi^2(u) du + \int_T^{\infty} \psi^2(u) du < \epsilon$ . Intuitively this condition says that most of the fluctuations in  $\psi(\cdot)$  are contained in some interval of finite width.
3.  $\psi(\cdot)$  must be ‘admissible,’ which is a mild technical condition that rules out the use of certain pathological functions of no practical importance (Mallat 1999).

(For simplicity, we have restricted our definitions to real-valued functions; however, it is possible to define complex-valued wavelets, and in fact the wavelets that were introduced first in the geophysical literature are complex-valued (Goupillaud et al. 1984).)

#### 4 What is Wavelet Analysis?

Now that we have defined what a wavelet is, how can we use it to tell us something about a time series? Suppose we let  $x_t$  represent the  $t$ th value of a time series, where  $t$  is an integer. Suppose further that we regard  $x_t$  as a sample from a real-valued function  $x(\cdot)$  that is defined over the real axis; i.e., we have  $x_t = x(t \Delta)$ , where  $\Delta$  is the spacing between adjacent observations (for example,  $\Delta$  is one year for the four series shown in Figure 1).

Merely to facilitate our discussion, let us assume  $\Delta$  to be small compare to one (we can force this to be true for the time series of vegetation groups by taking the time variable to be measured in centuries, in which case  $\Delta = 0.01$  centuries). Let us now consider what happens if we multiply  $x(\cdot)$  and the Haar wavelet  $\psi^{(H)}(\cdot)$  together (as illustrated in Figure 3) and integrate the resulting function; i.e., we form

$$W^{(H)} = \int_{-\infty}^{\infty} \psi^{(H)}(u)x(u) du.$$

The variable  $W^{(H)}$  is called a Haar wavelet coefficient.

What does the wavelet coefficient  $W^{(H)}$  tells us about  $x(\cdot)$ ? To address this question, we can use the definition for  $\psi^{(H)}(\cdot)$  in Equation (1) to argue that

$$W^{(H)} = -\frac{1}{\sqrt{2}} \int_{-1}^0 x(u) du + \frac{1}{\sqrt{2}} \int_0^1 x(u) du.$$

Texts on elementary calculus tell us that an integral of the form

$$\frac{1}{b-a} \int_a^b x(u) du$$

defines the average value of the function  $x(\cdot)$  over the interval  $(a, b]$ . The integral  $\int_{-1}^0 x(u) du$  can be regarded as the average value of  $x(\cdot)$  over the interval  $(-1, 0]$ , and similarly  $\int_0^1 x(u) du$  is its average value over  $(0, 1]$ . The width of these intervals is commonly referred to as the scale associated with  $W^{(H)}$ . We now see that the Haar wavelet coefficient is proportional to a difference between two averages, each of unit scale. Thus, if  $|W^{(H)}|$  is large, these unit scale averages of  $x(\cdot)$  differ substantially; on the other hand, if  $W^{(H)}$  is close to zero, there is not much difference between these two averages.

Given that we only observe  $x(\cdot)$  at the points  $x_t = x(t \Delta)$ , we can obtain an approximation to  $W^{(H)}$  as follows. For any positive integer  $K$ , let

$$\bar{x}_t^{(K)} = \frac{1}{K} \sum_{l=0}^{K-1} x_{t-l};$$

i.e.,  $\bar{x}_t^{(K)}$  is the average of the  $K$  variables  $x_{t-K+1}, x_{t-K+2}, \dots, x_{t-1}, x_t$ . Now suppose that  $K$  represents the number of samples  $x_t$  in the interval  $(-1, 0]$ . Then we have

$$\int_{-1}^0 x(u) du \approx \frac{1}{K \Delta} \sum_{l=0}^{K-1} x_{-l} \propto \bar{x}_0^{(K)};$$

i.e., to some degree of approximation, the average value of the function  $x(\cdot)$  over the interval  $(-1, 0]$  is proportional to the average  $\bar{x}_0^{(K)}$  of all the observed values falling in that interval.

Similarly,

$$\int_0^1 x(u) du \approx \frac{1}{K \Delta} \sum_{l=0}^{K-1} x_{K-l} \propto \bar{x}_K^{(K)},$$

from which it follows that  $W^{(H)} \propto \bar{x}_K^{(K)} - \bar{x}_0^{(K)}$  approximately; i.e., the wavelet coefficient is approximately proportional to a difference between two adjacent averages, each spanning an interval of unit scale.

While the Haar wavelet of Equation (1) tells us about the difference in averages over intervals of unit scale centered about the origin, it would be nice to obtain similar information about averages over other scales centered about arbitrary time points. We can do so by adjusting the scale of the wavelet and then relocating it. Accordingly, let us define a Haar wavelet that is associated with an arbitrary scale  $\lambda > 0$  and location  $t$  as follows:

$$\psi_{\lambda,t}^{(H)}(u) \equiv \frac{1}{\sqrt{\lambda}} \psi^{(H)}\left(\frac{u-t}{\lambda}\right) = \begin{cases} -\frac{1}{\sqrt{2\lambda}}, & t - \lambda < u \leq t; \\ \frac{1}{\sqrt{2\lambda}}, & t < u \leq t + \lambda; \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

Note that  $\psi_{1,0}^{(H)}(\cdot)$  and  $\psi^{(H)}(\cdot)$  are identical. It is easy to check that  $\psi_{\lambda,t}^{(H)}(\cdot)$  integrates to zero and that its square integrates to unity, so it satisfies two of the basic conditions for a wavelet. As examples, the left-hand column of Figure 4 shows  $\psi_{1,-1}^{(H)}(\cdot)$ , a unit scale wavelet centered at  $t = -1$ , and  $\psi_{2,0}^{(H)}(\cdot)$ , a wavelet of scale two centered at the origin. If we now multiply

$\psi_{\lambda,t}^{(H)}(\cdot)$  and  $x(\cdot)$  together (as illustrated in the figure), we obtain a wavelet coefficient that now depends on  $\lambda$  and  $t$ :

$$W^{(H)}(\lambda, t) \equiv \int_{-\infty}^{\infty} \psi_{\lambda,t}^{(H)}(u)x(u) du. \quad (3)$$

Note that  $W^{(H)}(1, 0)$  and  $W^{(H)}$  are identical. This function of scale and location is known as the Haar wavelet transform of  $x(\cdot)$ . The particular value  $W^{(H)}(\lambda, t)$  is called the Haar wavelet coefficient at  $\lambda$  and  $t$ . In analogy to our interpretation of  $W^{(H)}$ , we can say that  $W^{(H)}(\lambda, t)$  is proportional to a difference in scale  $\lambda$  averages located just before and after time  $t$ . If the difference between these averages is large, this is reflected in a large magnitude for  $W^{(H)}(\lambda, t)$ ; on the other hand, if the difference is small, then  $|W^{(H)}(\lambda, t)|$  will be small. Given samples  $x_t$  of  $x(\cdot)$ , we can approximate  $W^{(H)}(\lambda, t)$  based upon averages of appropriate portions of  $x_t$ .

In a similar manner we can define wavelet transforms other than the Haar by starting with a basic wavelet  $\psi(\cdot)$  that has unit scale and is centered at the origin. We then define  $\psi_{\lambda,t}(\cdot)$  as in the left-hand portion of Equation (2). The wavelet transform itself is given by the obvious analog of Equation (3). The interpretation of the resulting wavelet coefficients  $W(\lambda, t)$  can be deduced from a study of the shape of the basic wavelet  $\psi(\cdot)$ . For example, while the Haar wavelet  $\psi^{(H)}(\cdot)$  depicted in Figure 2 leads to coefficients that are proportional to the difference between adjacent averages, it is evident from the plot of  $\psi^{(rdG)}(\cdot)$  in this figure that this wavelet yields coefficients that are proportional to a difference between adjacent weighted averages. Similarly, the shape of the Mexican hat wavelet  $\psi^{(Mh)}(\cdot)$  implies that its coefficients are proportional to a difference between a weighted average centered about the origin and two weighted averages occurring before and after it.

The interpretation of wavelet coefficients as the difference between averages is fundamental for transforms based upon wavelets similar to those shown in Figure 2. The fact that we can make this physical interpretation helps explain why the wavelet transform has proven to be useful for many different types of time series. While average values of time

series are important, what is more often of scientific interest is how stable such averages are over time. To cite a well-known example, the study of a possible change in the climate leads us to determine how much change there has been over time in averages of various climate indicators. Since an effective way to determine this would be to see how much change there has been in averages associated with different scales (decades, centuries, etc.), the wavelet transform emerges as an important tool because of its physical interpretation. (The wavelet transform can be extended to handle images, in which case the coefficients can be interpreted as differences between co-located spatial averages.)

In addition to its appealing physical interpretation, let us note two other key properties of the wavelet transform. The first says that the continuous time series  $x(\cdot)$  and its wavelet transform  $W(\cdot, \cdot)$  are equivalent in the sense that, given either  $x(\cdot)$  or  $W(\cdot, \cdot)$ , we can recover the other. The ‘forward’ wavelet transform, i.e., going from  $x(\cdot)$  to  $W(\cdot, \cdot)$ , is simply a matter of definition (e.g., Equation (3)). The inverse wavelet transform, i.e., going from  $W(\cdot, \cdot)$  to  $x(\cdot)$ , is not immediately obvious, but in fact can be done as follows:

$$x(t) = \frac{1}{C_\psi} \int_0^\infty \left[ \int_{-\infty}^\infty W(\lambda, u) \frac{1}{\sqrt{\lambda}} \psi\left(\frac{t-u}{\lambda}\right) du \right] \frac{d\lambda}{\lambda^2}, \quad (4)$$

where  $C_\psi$  is a constant depending on just  $\psi(\cdot)$ . The fact that we can recover  $x(\cdot)$  from its wavelet transform is important because this means that all the information about the time series must also be contained in  $W(\cdot, \cdot)$  in some manner. We can consider  $x(\cdot)$  and  $W(\cdot, \cdot)$  to be two representations for a single mathematical entity, with  $W(\cdot, \cdot)$  offering a reexpression of the time domain formulation  $x(\cdot)$  in a manner that can bring out certain key information more succinctly. Another way of stating this equivalence is that knowledge of how  $x(\cdot)$  changes at all scales  $\lambda$  and all time locations  $t$  is equivalent to knowing  $x(\cdot)$  itself.

The second key property is that the wavelet transform can be used as the basis for an ‘energy’ decomposition of  $x(\cdot)$ . By definition, the term ‘energy’ is just short-hand for the quantity  $\int_{-\infty}^\infty x^2(t) dt$ , which we assume to be finite (for certain  $x(\cdot)$ , this definition would be

consistent with the notion of energy defined in physics). We can then write

$$\int_{-\infty}^{\infty} x^2(t) dt = \int_0^{\infty} \int_{-\infty}^{\infty} \frac{W^2(\lambda, t)}{C_{\psi}\lambda^2} dt d\lambda. \quad (5)$$

Thus, whereas a plot of  $x^2(t)$  versus  $t$  tells us how the energy in  $x(\cdot)$  is distributed with respect to time, the right-hand integrand above yields an energy distribution across both scale and time. (The above is the analog of a well-known result in Fourier theory, namely, that the energy in  $x(\cdot)$  can be decomposed across different frequencies in terms of the square magnitude of the Fourier transform for  $x(\cdot)$ .)

## 5 Discrete Wavelet Transforms

The wavelet transform we described in the previous section is usually called a continuous wavelet transform (CWT) since it is applied to a function  $x(\cdot)$  defined over the entire real axis. In practical applications, we only have a finite number  $N$  of sampled values. If we only have these samples, it is not possible to compute the CWT exactly, but we can resort to approximations.

Rather than approximating the CWT, an alternative approach is to define a wavelet transform specifically tailored for sampled values. This approach leads to the discrete wavelet transform (DWT), but it does require that the samples be of the form  $x_t = x(t\Delta)$ , where for convenience we take the integer  $t$  to range from 0 up to  $N - 1$ ; i.e., we need the data to be collected at equal intervals (if this is not the case, we can resort to an interpolation method to create a regularly sampled time series). To some degree of approximation, we can regard the DWT as being formed by taking ‘slices’ through a corresponding CWT, but it is important to note that the DWT can stand on its own, independent of any connection to a CWT. The slices are restricted to scales  $\lambda$  that assume the dyadic values  $2^{j-1}\Delta$ , where  $j = 1, 2, \dots, J_0$ . Here  $J_0$  is the total number of scales used and is usually dictated by the

sample size  $N$  and/or by the application at hand. Typically  $J_0$  is set to a value no greater than  $\log_2(N)$ , so the number of scales that we can meaningfully analyze is restricted by the amount of data available. It is convenient to define  $\tau_j = 2^{j-1}$  as a standardized dyadic scale, with the actual physical scale then being given by  $\tau_j \Delta$ .

The DWT also requires a discretization of the continuous time variable. One choice gives us  $N$  wavelet coefficients on each scale, yielding a version of the DWT called the maximal overlap DWT (MODWT); however, essentially the same transform goes by several other names, including the stationary DWT (Nason and Silverman 1995) and the non-decimated DWT (Bruce and Gao 1996). Let  $\mathbf{X}$  be an  $N$  dimensional vector whose  $t$ th element is  $x_t$ . For a given  $J_0$ , the MODWT transforms  $\mathbf{X}$  into  $J_0+1$  new vectors, each of dimension  $N$ . The first  $J_0$  of these are denoted by  $\widetilde{\mathbf{W}}_1, \dots, \widetilde{\mathbf{W}}_{J_0}$  and constitute the MODWT wavelet coefficients associated with (standardized) scales  $\tau_j$ ,  $j = 1, \dots, J_0$ . To some degree of approximation, each element in  $\widetilde{\mathbf{W}}_j$  can be associated with a CWT wavelet coefficient on scale  $\tau_j \Delta$ , so the MODWT wavelet coefficients have the same physical interpretation as the CWT coefficients. The final vector is denoted by  $\widetilde{\mathbf{V}}_{J_0}$  and contains the so-called MODWT scaling coefficients. Whereas the wavelet coefficients  $\widetilde{\mathbf{W}}_j$  are proportional to changes in averages over a scale of  $\tau_j$ , these scaling coefficients are proportional to averages over a scale of  $\tau_{J_0+1}$ . Roughly speaking, the scaling coefficients are summaries of the information in the CWT at all physical scales  $\lambda$  greater than or equal to  $\tau_{J_0+1}\Delta$ .

The MODWT is equivalent to the original time series in the sense that, given the MODWT coefficients, we can reconstruct  $\mathbf{X}$  (cf. Equation (4) for the CWT). This leads to the following additive decomposition, which is known as a multiresolution analysis (MRA):

$$\mathbf{X} = \sum_{j=1}^{J_0} \widetilde{\mathcal{D}}_j + \widetilde{\mathcal{S}}_{J_0}. \quad (6)$$

In the above,  $\widetilde{\mathcal{D}}_j$  is an  $N$  dimensional vector that depends upon just  $\widetilde{\mathbf{W}}_j$  and hence is

constructed using just those MODWT wavelet coefficients that are associated with changes of averages on a scale of  $\tau_j$ . This vector is called a ‘detail series’ and is the part of the MRA of  $\mathbf{X}$  that can be attributed to variations on a scale of  $\tau_j$ . The final term in the MRA is  $\tilde{\mathcal{S}}_{J_0}$ , which again is an  $N$  dimensional vector, but this depends just on the scaling coefficients  $\tilde{\mathbf{V}}_{J_0}$ . The vector  $\tilde{\mathcal{S}}_{J_0}$  is called the ‘smooth series’ because it is associated with averages over scales  $2\tau_{J_0}$  and longer and hence captures the slowly varying portion of  $\mathbf{X}$ . Thus an MRA is an additive decomposition that expresses a time series as the sum of several new series, each of which can be associated with variations on a particular scale.

The MODWT also offers a scale-based decomposition of the energy in  $\mathbf{X}$ . Recalling that we took  $\int_{-\infty}^{\infty} x^2(t) dt$  to be the energy in  $x(\cdot)$ , let us define  $\|\mathbf{X}\|^2 = \sum_{t=0}^{N-1} x_t^2$  to be the energy in  $\mathbf{X}$ . The MODWT-based energy decomposition says that

$$\|\mathbf{X}\|^2 = \sum_{j=1}^{J_0} \|\tilde{\mathbf{W}}_j\|^2 + \|\tilde{\mathbf{V}}_{J_0}\|^2 \quad (7)$$

(cf. Equation (5) for the CWT). Thus the energy in  $\mathbf{X}$  is preserved in its MODWT coefficients, and  $\|\tilde{\mathbf{W}}_j\|^2$  represents the contribution to the energy that is attributable to variations on the scale of  $\tau_j$ . This energy decomposition in turn allows us to decompose a well-known measure of the variability in a time series, namely, the sample variance, defined as

$$\hat{\sigma}_x^2 = \frac{1}{N} \sum_{t=0}^{N-1} (x_t - \bar{x})^2 = \frac{1}{N} \sum_{t=0}^{N-1} x_t^2 - \bar{x}^2 = \frac{1}{N} \|\mathbf{X}\|^2 - \bar{x}^2$$

where  $\bar{x} = \frac{1}{N} \sum_t x_t$  is the sample mean of the time series. If we use Equation (7) to substitute for  $\|\mathbf{X}\|^2$  in the above, we obtain the following scale-based decomposition of the sample variance:

$$\hat{\sigma}_x^2 = \frac{1}{N} \left( \sum_{j=1}^{J_0} \|\tilde{\mathbf{W}}_j\|^2 + \|\tilde{\mathbf{V}}_{J_0}\|^2 \right) - \bar{x}^2. \quad (8)$$

We can use this expression to argue that the contribution to the sample variance due to variations on a scale of  $\tau_j$  is given by  $\frac{1}{N} \|\tilde{\mathbf{W}}_j\|^2$ . We can take this quantity to be an estimator

for a corresponding theoretical quantity called the wavelet variance, which gives us a scale-based decomposition of a theoretical variance associated with a population that  $x_t$  is assumed to be drawn from. These ideas extend naturally to a scale-based decomposition of the sample covariance  $\hat{\sigma}_{xy}$  between two time series  $\mathbf{X}$  and  $\mathbf{Y}$ :

$$\hat{\sigma}_{xy} = \frac{1}{N} \sum_{t=0}^{N-1} (x_t - \bar{x})(y_t - \bar{y}) = \frac{1}{N} \left( \sum_{j=1}^{J_0} \langle \widetilde{\mathbf{W}}_{x,j}, \widetilde{\mathbf{W}}_{y,j} \rangle + \langle \widetilde{\mathbf{V}}_{x,J_0}, \widetilde{\mathbf{V}}_{y,J_0} \rangle \right) - \bar{x}\bar{y}, \quad (9)$$

where  $\langle \mathbf{X}, \mathbf{Y} \rangle = \sum_{t=0}^{N-1} x_t y_t$  for any two vectors  $\mathbf{X}$  and  $\mathbf{Y}$ , and  $\widetilde{\mathbf{W}}_{x,j}$  denotes the scale  $\tau_j$  wavelet coefficients for  $\mathbf{X}$  and so forth. The contribution to the sample covariance due to variations on a scale of  $\tau_j$  is thus given by  $\frac{1}{N} \langle \widetilde{\mathbf{W}}_{x,j}, \widetilde{\mathbf{W}}_{y,j} \rangle$ .

It should be noted that the MODWT of  $\mathbf{X}$  yields  $J_0 + 1$  new vectors for a total of  $(J_0 + 1) \times N$  values, which are  $J_0 \times N$  more than we started with (this expansion is reminiscent of the CWT, which starts with a function defined over the real axis and returns a ‘larger’ function in the sense of being defined both over the real axis and over all  $\lambda > 0$ ). It is in fact possible to define a DWT consisting of exactly the same number  $N$  of coefficients as there are values in  $\mathbf{X}$ . This version of the DWT can be obtained from the MODWT by rescaling and shortening all of its  $J_0 + 1$  vectors. It could be called a decimated version of the MODWT, but it is commonly called just ‘the DWT’ in the literature. Like the MODWT, this DWT can also be used to define an MRA and to decompose the energy in a time series, but the MODWT has some technical advantages that recommend its use for time series analyses such as we describe in the next section. There are, however, important applications in which the DWT would be the discrete transform of choice, including the compression, fast simulation and approximate decorrelation of time series (see Percival and Walden 2000 for details).

## 6 Wavelet Analysis of Vegetation Coverage Series

Let us now illustrate how the MODWT can be used to address the questions that we posed about the vegetation coverage series in §2. Figures 5 and 6 shows MRAs for, respectively, the time series of the boreal and tundra groups based upon a MODWT using the Haar wavelet. Here we have set  $J_0 = 4$  so that we get four detail series  $\tilde{\mathcal{D}}_1, \dots, \tilde{\mathcal{D}}_4$  (associated with changes in averages on scales of 1, 2, 4 and 8 years) and the smooth series  $\tilde{\mathcal{S}}_4$  (associated with averages over all scales greater than 8 years). If we let  $\tilde{\mathcal{D}}_{j,t}$  and  $\tilde{\mathcal{S}}_{4,t}$  represent the  $t$ th elements of, respectively,  $\tilde{\mathcal{D}}_j$  and  $\tilde{\mathcal{S}}_4$ , and if we let  $x_t$  represent the time series for either the boreal or tundra group, we have

$$x_t = \sum_{j=1}^4 \tilde{\mathcal{D}}_{j,t} + \tilde{\mathcal{S}}_{4,t}$$

for all  $t$ , which is just an element-wise reexpression of Equation (6).

A study of these MRAs indicates that the variations at any given scale appear to be homogeneous across time; i.e., there is little evidence that the statistical characteristics of the series have changed much over the course of the twentieth century. The MRAs do point out some localized features of potential interest. For example, the detail series  $\tilde{\mathcal{D}}_1$  for the boreal data (Figure 5) shows short stretches of decreased variability in the 1920s, 1960s and 1980s, indicating time spans over which there was relatively little year-to-year variability and inviting us to seek a physical explanation. In retrospect, these features are evident in the boreal time series itself (bottom plot of the figure) as portions of the series with relatively smooth variations, but the MRA has nicely brought these to our attention. If we now turn to the  $\tilde{\mathcal{D}}_2$  series, we see enhanced fluctuations in the early 1940s, pointing out a time span in which the boreal series tended to have similar values for groups of two years, but with relatively large shifts between adjacent groups. Again this feature can be seen in  $\mathbf{X}$  itself once it has been pointed out to us by the MRA. The fact that the boreal smooth series  $\tilde{\mathcal{S}}_4$  is

relatively flat is consistent with the plot of  $\mathbf{X}$ , which shows no obvious long term trends. On the other hand, if we consider the corresponding tundra smooth series in Figure 6, we see an overall long term decrease, with an acceleration over the last two decades. This pattern is evident in the tundra series itself (plotted at the bottom of the figure), but the smooth series in the MRA offers a nice visual summary of the overall decrease.

The MRAs for the boreal and tundra groups show a decrease in variability in the detail series as the associated scale increases. We can quantify how much each scale contributions to the overall variability in a given series via the wavelet variance, which is a population version of the variance decomposition described by Equation (8). Figure 7 shows a plot of unbiased MODWT-based wavelet variance estimates versus scales ranging from  $\tau_1 \Delta = 1$  year up to  $\tau_6 \Delta = 32$  years for the four vegetation group time series (details about this estimator are given in Percival and Walden 2000, Chapter 8). Each wavelet variance estimate is denoted by an ‘o’, out of which vertical lines emanate above and below. These lines represent 95% confidence intervals for the theoretical wavelet variance. Note that the confidence intervals associated with the largest scale of  $\tau_6 \Delta = 32$  years encompass nearly four orders of magnitude, a reflection of the fact that we really cannot expect to say much about changes on a scale of 32 years with only 100 years of data.

The wavelet variance estimates in Figure 7 indicate that, for all four series, the dominant contribution to the sample variance is due to variations at the smallest scale (one year). Thus the year-to-year variations are more prominent than variations at other scales such as a decade (this answers one of the questions we posed in §2). The rate of decay of the wavelet variance versus scale is an indicator of the type of noise process that might be a good model for the time series. A slope of  $-1$  on a plot of the log of the wavelet variance versus the log of the scale indicates that the time series might be well modeled by a white noise process or by some other process that quickly decorrelates; i.e., any two distinct  $x_t$ ’s can be considered

to be realizations of approximately uncorrelated random variables, with the approximation rapidly improving as the distance in time between the  $x_t$ 's increases. For reference, we have drawn four thick lines with a slope of  $-1$  below the wavelet variance estimates in Figure 7. There are also thin lines drawn through the wavelet variance estimates themselves. These represent weighted linear least squares fits to the log of the wavelet variance estimates versus the log of the scale (Percival and Walden 2000, §9.5). A visual comparison of the thick and thin lines indicates that, for all four series, the wavelet variance estimates decay at a slower rate than what is consistent with a rapidly decorrelating process; however, when we take into account the sampling variability associated with the estimated slopes, we cannot reject the null hypothesis of white noise for the boreal, temperate and frost group series at a significance level of 0.05, but we can do so for the tundra series. While the coverage of the frost group is too small to merit much attention, the fact that this group and the boreal group have similar rates of decay indicates that, apart from a marked difference in their overall variance, their other statistical properties are quite similar (this addresses another question posed in §2). That the time series for the tundra group is qualitatively different from those of the other three groups is evident in Figure 1, where we can see that that this series has more prominent large scale (or low frequency) fluctuations than the other series. In fact, the rate of decay for the tundra group is consistent with what we would expect from a fractionally differenced (FD) process with a long memory parameter  $\delta = 0.3$  (Granger and Joyeux 1980, Hosking 1981). Details about FD processes are beyond the scope of this article, but suffice it to say that their physical basis has some interesting implications, namely, that they do not have a well-defined characteristic scale and that their correlation structure has certain fractal-like properties (Beran 1994, Percival et al. 2001a). A study of the rate of decay of the wavelet variance versus scale can give us considerable insight into the correlation structure of a time series.

Finally let us consider the relationship between the boreal and temperate time series. If we compute the sample cross-correlation  $\hat{\sigma}_{xy}/(\hat{\sigma}_x\hat{\sigma}_y)$  between these series, we find that it is  $-0.79$ ; i.e., the two series are anticorrelated, which is consistent with the manner in which they were constructed. Just as we used the MODWT coefficients to decompose the sample variance across different scales, so can we use them to decompose the sample cross-correlation across scales by making use of Equation (9). This decomposition is displayed in Figure 8. We see that, while the two series are in fact anticorrelated to some degree at all scales, the sample cross-correlation is mainly due to the two smallest scales (addition of these two components yields 75% of the observed anticorrelation).

## 7 Summary and Discussion

Wavelets are mathematical tools that allow us to decompose a time series with respect to two independent variables, namely, time and scale (they can also be used to decompose an image with respect to location and extent). Both the continuous wavelet transform and the closely related discrete wavelet transform can be interpreted in terms of differences between (possibly weighted) temporally adjacent averages of portions of a time series. Both transforms have two fundamental properties. The first is that, given the transform coefficients, we can reconstruct the time series perfectly; moreover, in the case of the discrete wavelet transform, this reconstruction leads to a multiresolution analysis in which the series is reexpressed as the sum of a new set of time series (details and a smooth), each of which is associated with variations at a particular scale. A multiresolution analysis can point out local features of potential interest in a time series. For example, using this type of analysis, we were able to identify certain decades in the series for the boreal group in which the year-to-year variations were noticeably smaller than usual.

The second fundamental property is that the energy in a time series is preserved in its

transform coefficients. This fact leads to a decomposition of the sample variance for a time series into components that can be associated with different scales. This scale-based analysis of variance can give us considerable insight into the correlation properties of a time series. This type of analysis allowed us to note, e.g., that, while the boreal group series is much more variable than the frost group series overall, they are both consistent with a white noise hypothesis, implying that some of their other statistical properties are identical. As a second example, this wavelet-based analysis pointed out that the tundra group series is unique amongst the four series in having a significantly slower rate of decay of its scale-based contributions to the sample variance as a function of scale. An extension of this technique shows that a significant negative cross-correlation between the boreal and temperate time series is due primarily to year-to-year variations.

This introduction to wavelets has been narrowly focused so that we could concentrate on certain key concepts. There are many other uses for wavelets beyond what we have briefly covered here in the study of time series and spatial data related to vegetation monitoring. Here are three examples.

1. For some time series, use of wavelets centers around the fact that they can approximately decorrelate highly correlated time series, which in turn can be used to assess the sampling properties of certain commonly used statistics (Percival et al. 2001b).
2. The time series that we have used here as examples are actually derived from a spatial database of monthly surface air temperatures. Whereas our present analysis considers temporal changes, we could also use wavelets to help us identify what regions are most affected by these changes (this would complement the spatial analysis considered in Wang and Overland 2004).
3. For spatial data collected in the form of images, the notion of a scale-based decom-

position of variance leads to an analysis of texture, while there are also ways of using wavelets to identify the boundaries (edges) between different types of vegetation (Mallat 1999).

In these and other cases, the usefulness of wavelets is due to the fact that they allow a localized and physically meaningful decomposition tied to the notion of differences between adjacent averages over different scales.

### **Acknowledgments**

We would like to thank the three referees of this article, all of whom offered very helpful suggestions for improvements. This work was funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement No. NA17RJ1232, Contribution #1032. This article is also Contribution No. 2647 from the Pacific Marine Environmental Laboratory/NOAA. The first author would like to thank the organizers of the International Symposium on the State of the Art in Vegetation Monitoring Approaches, March 24–26, 2003, Swiss Federal Research Institute WSL, Birmensdorf, Switzerland, for an invitation and support for giving a presentation around which this article is based.

## References

- Beran, J. 1994. *Statistics for long memory processes*. Chapman and Hall, New York.
- Bradshaw, G.A., and T.A. Spies. 1992. Characterizing canopy gap structure in forests using wavelet analysis. *Journal of Ecology* 80:205–215.
- Bruce, A.G., and H.–Y. Gao. 1996. *Applied wavelet analysis with S-PLUS*. Springer, New York.
- Brunet, Y., and S. Collineau. 1994. Wavelet analysis of diurnal and nocturnal turbulence above a maize crop. E. Foufoula–Georgiou and P. Kumar (eds), *Wavelets in Geophysics*. Academic Press, San Diego, CA, pp. 129–150.
- CAVM Team. 2003. *Circumpolar Arctic Vegetation Map*. Scale 1:7,500,000. Conservation of Arctic Flora and Fauna (CAFF) Map No. 1. U.S. Fish and Wildlife Service, Anchorage, Alaska
- Csillag, F., and S. Kabos. 2002. Wavelets, boundaries, and the spatial analysis of landscape pattern. *Ecoscience* 9:177–190.
- Dale, M.R.T., and M. Mah. 1998. The use of wavelets for spatial pattern analysis in ecology. *Journal of Vegetation Science* 9:805–814.
- Goupillaud, P., A. Grossmann and J. Morlet. 1984. Cycle-octave and related transforms in seismic signal analysis. *Geoexploration* 23:85–102.
- Granger, C.W.J., and R. Joyeux. 1980. An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis* 1:15–29.
- Hosking, J.R.M. 1981. Fractional differencing. *Biometrika* 68:165–176.
- Köppen, W.P. 1931. *Grundriss de Klimakunde*. Walter de Gruyter, Berlin.
- Lark, R.M., and R. Webster. 1999. Analysis and elucidation of soil variation using wavelets. *European Journal of Soil Science* 50:185–206.
- Lark, R.M., and R. Webster. 2001. Changes in variance and correlation of soil properties with scale and location: analysis using an adapted maximal overlap discrete wavelet transform. *European Journal of Soil Science* 52:547–562.
- Mallat, S.G. 1999. *A wavelet tour of signal processing (second edition)*. Academic Press, San

Diego.

- Milne, A.A. 1926. *Winnie-the-Pooh*. E.P. Dutton & Company, New York.
- Nason, G.P., and B. Silverman. 1995. The stationary wavelet transform and some statistical applications. In: A. Antoniadis and G. Oppenheim (eds), *Wavelets and Statistics (Lecture Notes in Statistics, Volume 103)*. Springer-Verlag, New York, pp. 281–299.
- Percival, D.B., J.E. Overland and H.O. Mofjeld. 2001a. Interpretation of North Pacific variability as a short- and long-memory process. *Journal of Climate* 14:4545–4559.
- Percival, D.B., S. Sardy and A.C. Davison. 2001b. Wavestrapping time series: adaptive wavelet-based bootstrapping. W.J. Fitzgerald, R.L. Smith, A.T. Walden and P.C. Young (eds), *Non-linear and Nonstationary Signal Processing*. Cambridge University Press, Cambridge, UK, pp. 442–470.
- Percival, D.B., and A.T. Walden. 2000. *Wavelet methods for time series analysis*. Cambridge University Press, Cambridge, UK.
- Shiyatov, S. G. 2003. Rates of change in the upper treeline ecotone in the polar Ural mountains. *Past Global Changes (PAGES) News*, 11(1):8–10.
- Sturm, M., C. Racine, and K. Tape. 2001. Increasing shrub abundance in the Arctic. *Nature*, 411:546–547.
- Trewartha, G.T., and L.H. Horn. 1980. *An introduction to climate (fifth edition)*. McGraw-Hill, New York.
- Wang, M., and J.E. Overland. 2004. Detecting Arctic climate change using Köppen climate classification. To appear in *Climatic Change*.

## Figure Captions

**Figure 1.** Four vegetation time series based upon Köppen classification. Each series represents the amount of area (measured in  $10^6 \text{ km}^2$ ) in the Arctic region covered during a given year by a certain vegetation group (boreal, temperate, tundra and frost). There is one value for each year from 1901 to 2000.

**Figure 2.** A wave (top plot) and three wavelets (bottom three plots). The wave is just a sine function with a period of  $2\pi$ . The wavelets are, from top to bottom, the Haar wavelet  $\psi^{(\text{H})}(\cdot)$ , a wavelet  $\psi^{(\text{fdG})}(\cdot)$  based upon the first derivative of the standard Gaussian (normal) probability density function  $f(u) = \frac{1}{\sqrt{2\pi}}e^{-u^2/2}$  and the Mexican hat wavelet  $\psi^{(\text{Mh})}(\cdot)$ , which is based upon the second derivative of  $f(\cdot)$ .

**Figure 3.** Product of the Haar wavelet  $\psi^{(\text{H})}(\cdot)$  and a time series  $x(\cdot)$ .

**Figure 4.** Products of rescaled and relocated Haar wavelets and a time series  $x(\cdot)$ .

**Figure 5.** Multiresolution analysis (MRA) for time series of boreal group based upon the Haar MODWT wavelet transform. The bottom-most curve shows the time series  $\mathbf{X}$  itself, while the other five curves show the components of the MRA. Four of these components are detail series, denoted by  $\tilde{\mathcal{D}}_j$ ,  $j = 1, 2, 3$  and 4. The values in the vector  $\tilde{\mathcal{D}}_j$  represent the portion of  $\mathbf{X}$  that is attributable to changes of averages on a scale of  $2^{j-1}$  years. The fifth component is the smooth series  $\tilde{\mathcal{S}}_4$ , which is the portion of  $\mathbf{X}$  attributable to averages on a scale of 16 years. An MRA is an additive decomposition, so we have  $\mathbf{X} = \sum_{j=1}^4 \tilde{\mathcal{D}}_j + \tilde{\mathcal{S}}_4$  exactly; i.e., the value in  $\mathbf{X}$  for, e.g., the year 1946 is equal to the sum of the five values in the four  $\tilde{\mathcal{D}}_j$ 's and  $\tilde{\mathcal{S}}_4$  for this same year. This particular MRA employs reflection boundary conditions (Percival and Walden 2000, p. 140).

**Figure 6.** As in Figure 5, but now for time series of tundra group.

**Figure 7.** Haar MODWT wavelet variance estimates. In each plot the o's indicate MODWT-based unbiased estimates of the Haar wavelet variance for scales  $\tau_j \Delta$  equal to 1, 2, 4, 8, 16 and 32

years. The vertical lines emanating from the  $\circ$ 's give 95% confidence intervals for a corresponding theoretical wavelet variance. The thin line through the  $\circ$ 's depicts a weighted regression fit to the log of the wavelet variance estimates versus the log of the scale. Below the  $\circ$ 's on each plot is a thick reference line with a slope of  $-1$ . This slope indicates the rate of decay we would expect to have for a white noise process. When sampling variability is taken into account, the slopes for the boreal, temperate and frost groups are not significantly different from  $-1$ , whereas that for the tundra group is significantly greater. A slope exceeding  $-1$  is indicative of a process possessing 'long memory' or certain fractal-like properties.

**Figure 8.** Scale-based decomposition of sample cross-correlation between boreal and temperate group time series. The sample cross-correlation between these two time series is  $-0.79$ ; i.e., the boreal and temperate groups are anticorrelated. The  $\circ$ 's indicate the contribution to this cross-correlation attributable to scales of 1, 2, 4 and 8 years, along with a composite of scales greater than 8 years. This scale-based decomposition says that the anticorrelation is mainly due to small scale variations.

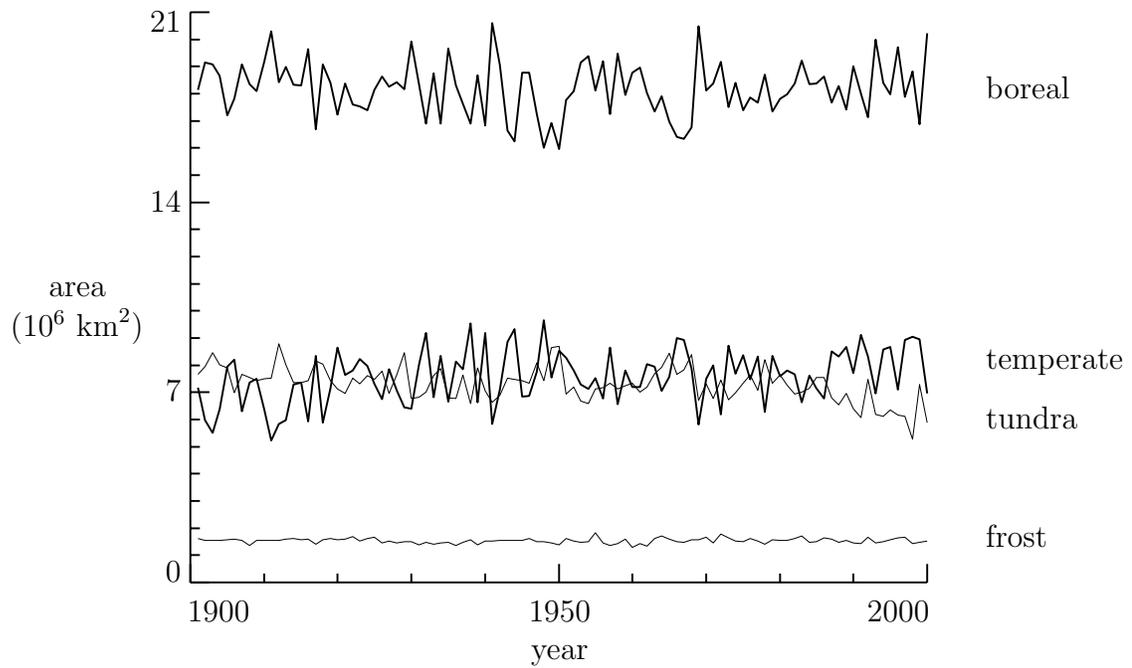


Figure 1: Four vegetation time series based upon Köppen classification. Each series represents the amount of area (measured in  $10^6 \text{ km}^2$ ) in the Arctic region covered during a given year by a certain vegetation group (boreal, temperate, tundra and frost). There is one value for each year from 1901 to 2000.

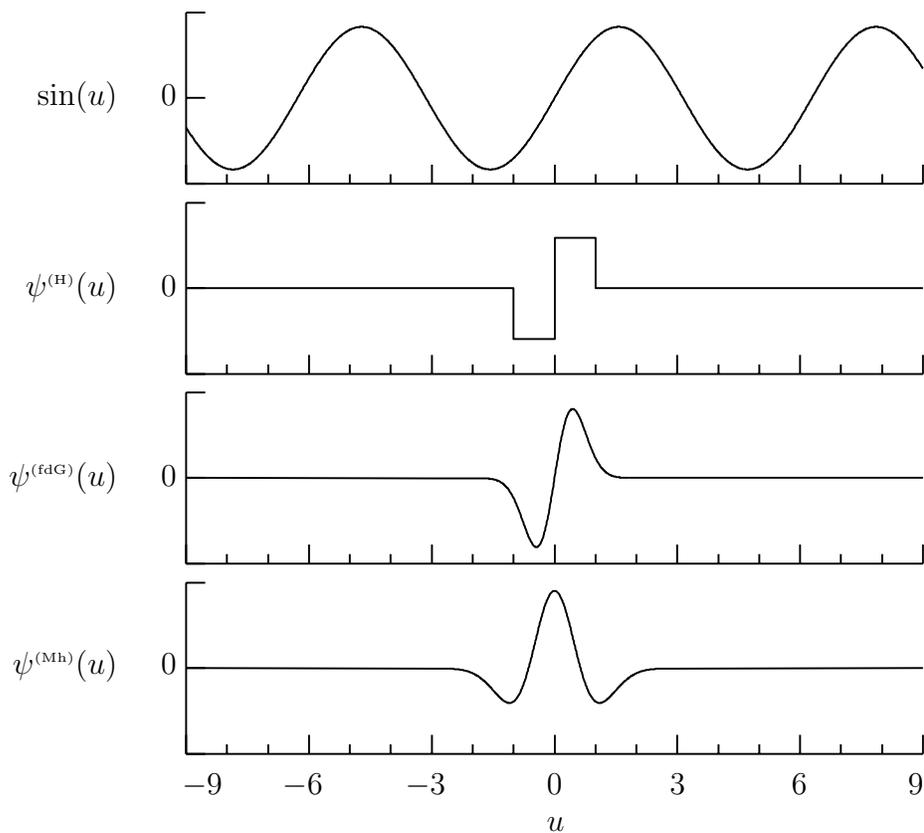


Figure 2: A wave (top plot) and three wavelets (bottom three plots). The wave is just a sine function with a period of  $2\pi$ . The wavelets are, from top to bottom, the Haar wavelet  $\psi^{(H)}(\cdot)$ , a wavelet  $\psi^{(\text{fdG})}(\cdot)$  based upon the first derivative of the standard Gaussian (normal) probability density function  $f(u) = \frac{1}{\sqrt{2\pi}}e^{-u^2/2}$  and the Mexican hat wavelet  $\psi^{(\text{Mh})}(\cdot)$ , which is based upon the second derivative of  $f(\cdot)$ .

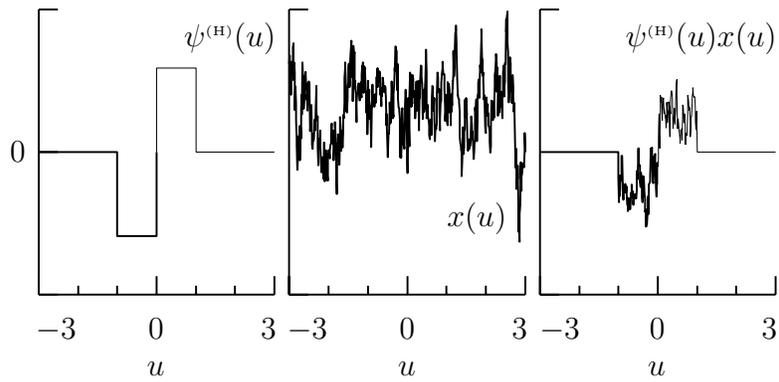


Figure 3: Product of the Haar wavelet  $\psi^{(H)}(\cdot)$  and a time series  $x(\cdot)$ .

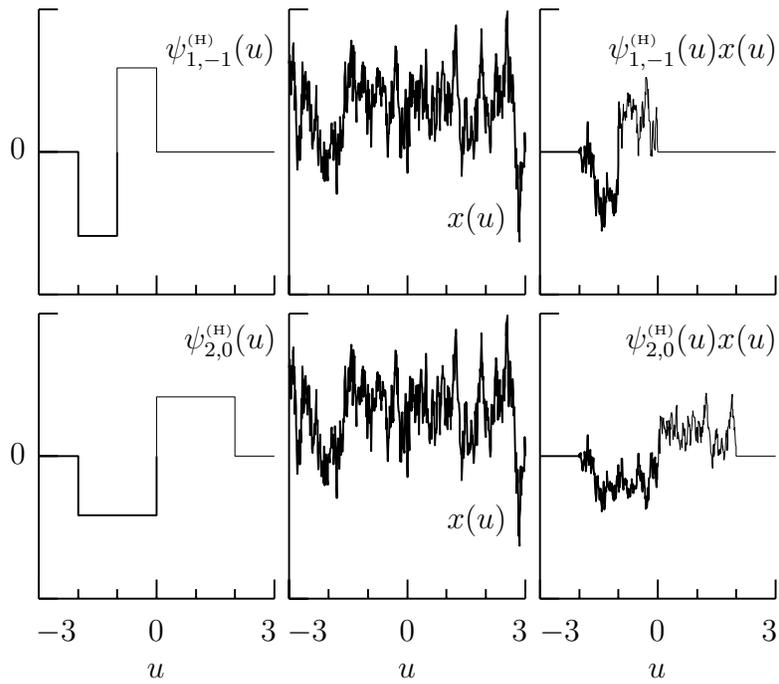


Figure 4: Products of rescaled and relocated Haar wavelets and a time series  $x(\cdot)$ .

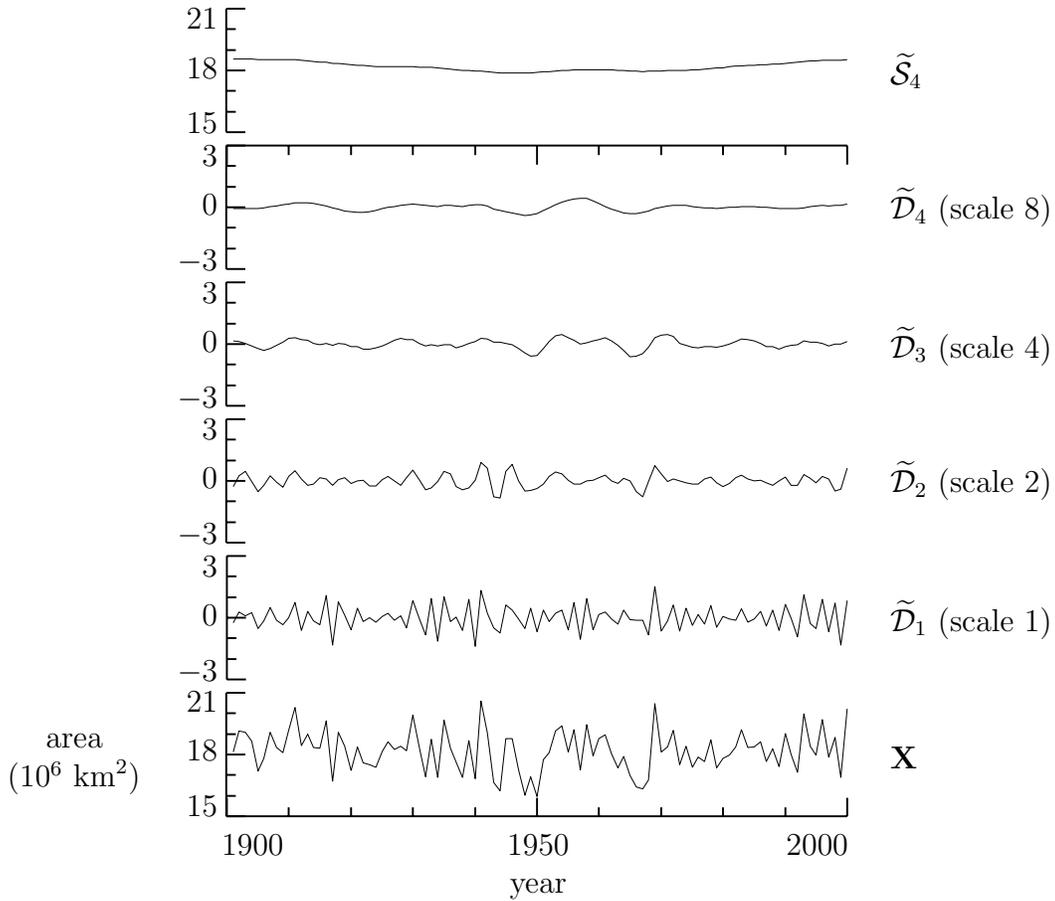


Figure 5: Multiresolution analysis (MRA) for time series of boreal group based upon the Haar MODWT wavelet transform. The bottom-most curve shows the time series  $\mathbf{X}$  itself, while the other five curves show the components of the MRA. Four of these components are detail series, denoted by  $\tilde{\mathcal{D}}_j$ ,  $j = 1, 2, 3$  and 4. The values in the vector  $\tilde{\mathcal{D}}_j$  represent the portion of  $\mathbf{X}$  that is attributable to changes of averages on a scale of  $2^{j-1}$  years. The fifth component is the smooth series  $\tilde{\mathcal{S}}_4$ , which is the portion of  $\mathbf{X}$  attributable to averages on a scale of 16 years. An MRA is an additive decomposition, so we have  $\mathbf{X} = \sum_{j=1}^4 \tilde{\mathcal{D}}_j + \tilde{\mathcal{S}}_4$  exactly; i.e., the value in  $\mathbf{X}$  for, e.g., the year 1946 is equal to the sum of the five values in the four  $\tilde{\mathcal{D}}_j$ 's and  $\tilde{\mathcal{S}}_4$  for this same year. This particular MRA employs reflection boundary conditions (Percival and Walden 2000, p. 140).

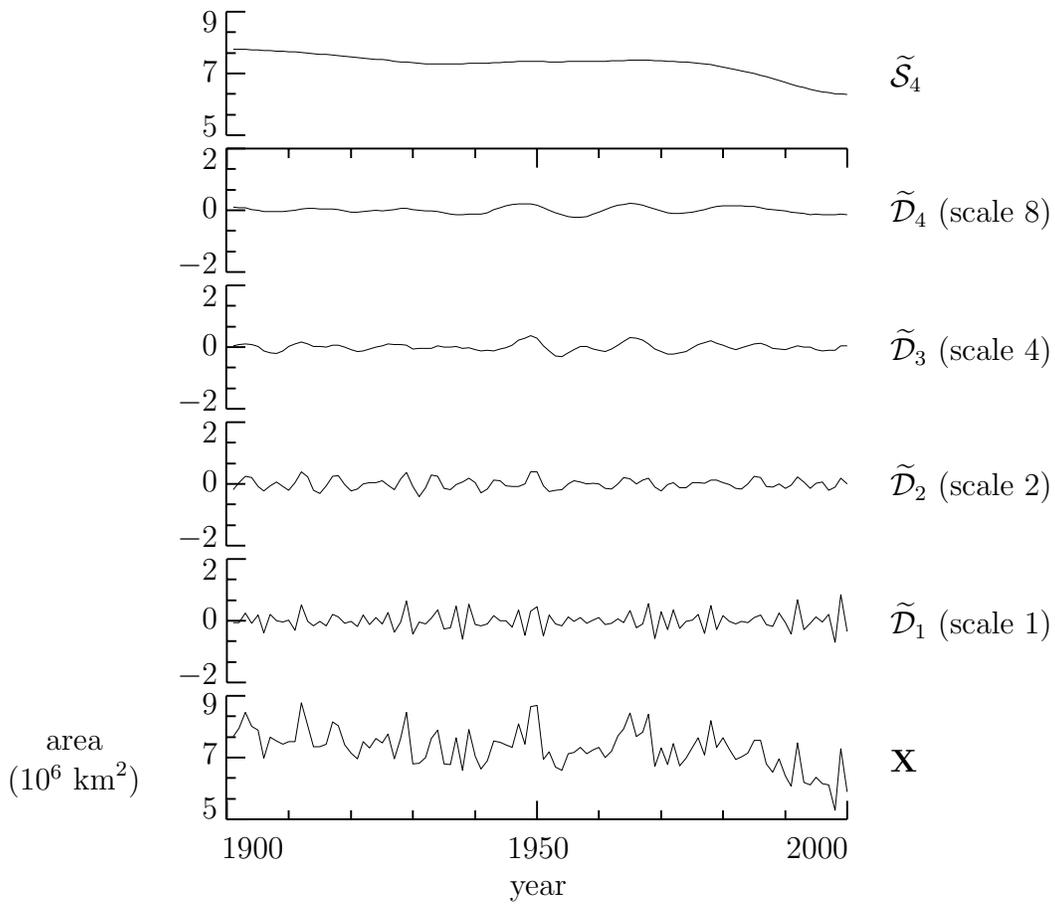


Figure 6: As in Figure 5, but now for time series of tundra group.

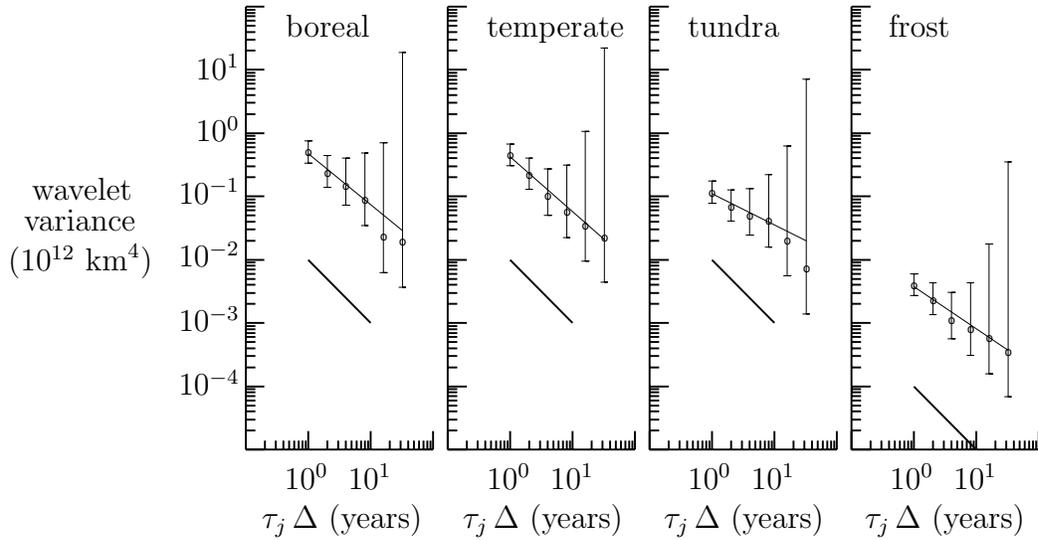


Figure 7: Haar MODWT wavelet variance estimates. In each plot the  $\circ$ 's indicate MODWT-based unbiased estimates of the Haar wavelet variance for scales  $\tau_j \Delta$  equal to 1, 2, 4, 8, 16 and 32 years. The vertical lines emanating from the  $\circ$ 's give 95% confidence intervals for a corresponding theoretical wavelet variance. The thin line through the  $\circ$ 's depicts a weighted regression fit to the log of the wavelet variance estimates versus the log of the scale. Below the  $\circ$ 's on each plot is a thick reference line with a slope of  $-1$ . This slope indicates the rate of decay we would expect to have for a white noise process. When sampling variability is taken into account, the slopes for the boreal, temperate and frost groups are not significantly different from  $-1$ , whereas that for the tundra group is significantly greater. A slope exceeding  $-1$  is indicative of a process possessing 'long memory' or certain fractal-like properties.

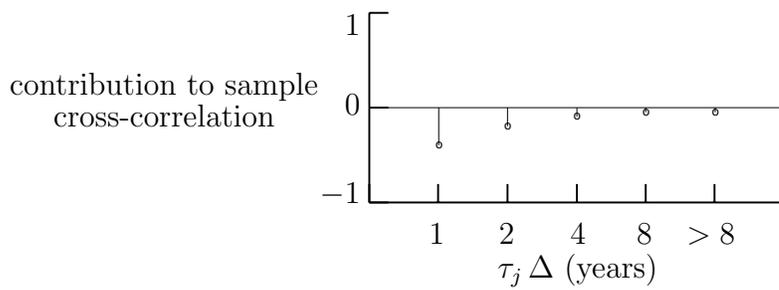


Figure 8: Scale-based decomposition of sample cross-correlation between boreal and temperate group time series. The sample cross-correlation between these two time series is  $-0.79$ ; i.e., the boreal and temperate groups are anticorrelated. The o's indicate the contribution to this cross-correlation attributable to scales of 1, 2, 4 and 8 years, along with a composite of scales greater than 8 years. This scale-based decomposition says that the anticorrelation is mainly due to small scale variations.