

Scale-specific inference using wavelets

Timothy H. Keitt¹

Section of Integrative Biology

University of Texas

Austin, TX 78712 USA

Dean L. Urban

Nicholas School of the Environment and Earth Sciences

Duke University

Durham, NC 27708 USA

¹Correspondence author: tkeitt@mail.utexas.edu; address: 1 University Station C0930, Austin, Texas 78712; phone: 512-471-5004; fax: 512-471-0961

Abstract

Understanding of spatial pattern and scale has been identified as a key issue in ecology, yet ecologists lack formal tools for making inference about relationships between scale-specific patterns. We introduce wavelet-coefficient regression, in which the dependent and independent variables are wavelet transformed prior to analysis, as a means to formalize scale-specific relationships in ecological data. We apply this method to data on vegetation and environmental factors related to water availability from Sequoia-Kings Canyon National Park. We find that the wavelet transform and wavelet-coefficient regression efficiently characterize scale-specific pattern in these data. We also find that different environmental factors show up as good predictors of vegetation growth at different scales and that these differences in scale greatly facilitate interpretation of the mechanisms relating water availability to vegetation growth.

Keywords: wavelets, wavelet-coefficient regression, pattern, scale, statistical analysis, Sierra Nevada, remote sensing, vegetation analysis, environmental control

Running head: wavelets

Introduction

The problem of pattern and scale in landscapes has been identified as central to further progress in ecological theory (Levin, 1992), yet ecology lacks general methods for detecting scale-specific association between observed patterns and hypothesized predictors. Considerable attention has been given to the problem of detecting dominant scales of pattern in environmental data, going back to the pioneering work of Greig-Smith (1964). In vegetation science, multi-scale ordination (Noy-Meir and Anderson, 1971; ver Hoef and Glenn-Lewin, 1989; Wagner, 2004) was designed to find the scale(s) at which trends in multivariate species data are mostly strongly evident.

Multi-scale ordination amounts to an analysis of rescaled datasets generated via blocking the base-level data (typically, quadrats along a transect); this is essentially a multivariate extension of

two-term local covariances that evolved from blocksize ANOVA (Greig-Smith, 1964; Dale, 1999). Similarities between modern blocking techniques (e.g., two-term local quadrat covariance, TTLQC) and the analysis we describe will be apparent (Dale, 1999). More recently, general methods for the characterization and treatment of spatial pattern have arisen across a variety of disciplines, with many of these subsumed under the title of geostatistics (Cressie, 1993). In this, the contiguity relationship explicit in blocking techniques is represented by the distance between two samples (lag distance). Dale (1999) explains how blocking techniques and semivariance are related, and Wagner (2003) illustrates how, under certain circumstances and assumptions about the data, blocking techniques such as multi-scale ordination and semivariance analysis can be integrated into a more cohesive body of spatial analyses. All of these methods typically ask whether a pattern exists, at what scale or scales pattern is expressed most strongly. In many instances, spatial pattern is seen as a nuisance, and the goal is to discover how statistical procedures can avoid undesirable outcomes such as spurious correlation between variables or biased significance tests, in the presence of spatial pattern (Dale, 1999).

Less often asked, and perhaps a Rosetta Stone for translating complex spatial patterns into mechanistic understanding, is the question of whether a scale-specific pattern of interest can be predicted from patterns expressed at similar scales in measured covariates. For example, a plant ecologist might ask whether typical patch size in a plant community is determined by the scale of pattern in water, nutrient or sunlight availability, or is the result of typical herbivore group size, typical scales of pathogen outbreaks, or the typical scale of abiotic disturbances. These questions have, in the past, been addressed indirectly, for example through hierarchical quadrat blocking, with standard univariate and multivariate statistical methods. In spatial statistics, partial Mantel tests provide some leverage on this problem, but not in a scale-specific manner (Urban et al., 2002). However, direct assessment of relationships between variables at specific scales has remained a challenging problem. Here, we show that traditional statistical methods, such as multiple regression, can be adapted to address scale-specific questions in ecological data. Our approach utilizes wavelet transforms (Daubechies, 1992) to extract scale-specific information

from independent and dependent variables and then employs familiar statistical methods to detect scale-specific associations in the transformed data. We illustrate the method using geophysical and remotely-sensed data from the Sierra Nevada landscape in western North America to address the question of whether scale-specific patterns in physical drivers of ecosystem productivity generate similar scale-specific patterns in vegetation structure. In the Sierra Nevada, previous studies have shown that the physical environment is spatially structured on multiple scales and we suspect that vegetation responds accordingly (Urban et al., 2000, 2002). However, the approach is quite general and we close with a discussion of what we see as fruitful avenues in which to further develop these methods.

Wavelets

Increasingly, wavelet transforms have become the preferred representation in which to analyze pattern and scale in environmental data (Bradshaw and Spies, 1992; Dale and Mah, 1998). Wavelet transforms are similar to Fourier transforms, in that they decompose a pattern into a hierarchy of different scales (Keitt, 2000), but offer distinct advantages over Fourier transforms when analyzing complex, non-stationary patterns. This is because wavelets transforms are local transforms and thus provide information about the intensity of pattern at different scales at a particular location. Formal treatment of wavelets has been covered exhaustively elsewhere (Daubechies, 1992) and is beyond the scope of this paper. However, basic concepts are easily mastered and one can get an intuition of how the wavelet transform works through simple arguments. A particularly accessible treatment can be found in Walker (1999).

Application of the wavelet transform repeatedly splits an initial sequence into detail coefficients that quantify local fluctuations at a particular scale, and smooth coefficients that quantify remaining low-frequency variation in the signal after the high-frequency detail is removed. The simplest wavelet transform is the discrete Haar transform. In the Haar transform, detail coefficients are calculated simply by subtracting successive values in the sequence (Walker,

1999). Given a sequence $f = (f_1, f_2, \dots, f_N)$ sampled at N equally spaced intervals, the first level detail coefficients are given by

$$d_m^{(1)} = \frac{f_{2m-1} - f_{2m}}{\sqrt{2}} \quad (1)$$

where $m = 1, 2, 3, \dots, N/2$. Notice that variation in the level 1 detail coefficients only reflects local, nearest-neighbor fluctuations in the sequence. The level 1 smooth coefficients are then

$$s_m^{(1)} = \frac{f_{2m-1} + f_{2m}}{\sqrt{2}} \quad (2)$$

where $m = 1, 2, 3, \dots, N/2$. Level 2 detail coefficients are computed by differencing this trend information

$$d_m^{(2)} = \frac{s_{2m-1}^{(1)} - s_{2m}^{(1)}}{\sqrt{2}} \quad (3)$$

where $m = 1, 2, 3, \dots, N/4$. Similarly, the level 2 smooth coefficients are obtained by summing pairs of level 1 smooth coefficients. At each level, the number of smooth and detail coefficients obtained drops by 1/2. The process can be iterated until only one smooth and one detail coefficient are produced. Decimation of the sequence by a factor of 2 at each level has the effect of amplifying the output variance by a factor of $\sqrt{2}$ at each level. The tradition is to adjust the transform such that the variance is constant across levels, hence the division by $\sqrt{2}$ at each level (Walker, 1999). An example illustrating the transform is shown in Figure 1.

The relationship between the level of the transform and scale is quite simple in this case. The level 1 detail coefficients correspond to a window spanning two sample intervals; level 2 coefficients correspond to a window spanning four sample intervals, and so on. In signal theory, this is equivalent to decomposing the variation in the original sequence into components corresponding to frequencies of π radians per sample interval (level 1), $\pi/2$ radians per sample interval (level 2), $\pi/4$ radians per sample interval (level 3), and so on down to π/N radians per sample interval (Diggle, 1990). The end result of the transform is a partitioning of variance into a

hierarchy of well-defined scales.

The wavelet transform is reminiscent of blocking techniques familiar to vegetation scientists. In particular, two-term local quadrat covariances (TTLQC) compute a local variance (squared differences) from contiguous samples that are successively aggregated into larger blocks (Dale, 1999). By comparison, the wavelet filter evaluates a function (the wavelet) using the data within each window (block). Importantly, there are wavelet functions (including the Haar transform) that transform data without loss of information; the back-transformation can reconstitute the original data (see below) unlike traditional methods of pattern analysis in vegetation studies.

Multiple regression of wavelet-transformed data

Like the Fourier transform, the wavelet transform is a linear operator. This is convenient within the framework of linear regression because, for any linear function f , we require that $f(ax) = af(x)$ where a is a constant. Thus, the standard multiple linear model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \sigma^2 \epsilon \quad (4)$$

can be transformed by a linear operator Ψ such that

$$\Psi(y) = \beta'_0 + \beta_1 \Psi(x_1) + \beta_2 \Psi(x_2) + \dots + \beta_n \Psi(x_n) + \sigma^2 \Psi(\epsilon) \quad (5)$$

without altering the interpretation of the coefficients. What is different is the intercept term and, in general, the covariance structure of the error terms. By choosing the wavelet transform for our linear operator, we can fit scale-specific models in the sense that the transform is extracting scale-specific components of pattern at each level of the decomposition. We refer to this general approach as wavelet-coefficient regression.

While not commonly employed in the ecological literature, we note that spectral transformation of

data prior to data analysis is quite common in some areas of statistics (see for example Shumway and Stoffer, 2000). In these cases, wavelets are not generally used for scale-specific inference, but rather to reduce the dimensionality of complex space-time problems. Similar applications also appear in the economics and finance literature (Ramsey, 1999).

Inverse transforms and multiresolutional decomposition

If the chosen wavelet meets certain requirements, then the original data sequence can be reconstructed by applying an inverse wavelet transform to the coefficients produced by the forward transform (Daubechies, 1992). This property leads to several interesting applications of wavelets. One application is as an adaptive smoother for noisy data. Because uncorrelated errors result in relatively small wavelet coefficients, random corruptions of a signal can often be removed by replacing small coefficients obtained from the forward transform with zeros, and then applying the inverse transform to yield a smoothed version of the original sequence. Selective removal of wavelet coefficients can be used for scatter-plot smoothing, or as input to generalized additive regression models (Hastie and Tibshirani, 1990). The latter case is sometimes referred to as 'wavelet regression,' but is quite different from the approach presented here. Additive models seek smooth, non-parametric functions of independent variables that are good predictors of dependent variables, but the procedure is neither scale-specific, nor does it directly assess relationships between spatial patterns in independent and dependent variables.

Another application of the inverse transform is multiresolutional decomposition (Mallat, 1989). In multiresolutional decomposition (MRD), the original sequence is decomposed into a series of new sequences, each containing only patterns found at a particular scale. This is accomplished by retaining only the coefficients at a single level of the transform (setting all others to zero) and then back-transforming. The resulting sequence only contains patterns from the original sequence specific to the scale at which the coefficients were retained. A nice property of MRD is that the decomposition is additive such that summing the resulting sequences together yields the original

data (see Figure 4, below). This also suggests an alternative scale-specific regression approach. One could fit a regression model at each level of an MRD of the independent and dependent variables rather than directly on the wavelet coefficients as in this paper: the principle disadvantage being the tendency for the MRD to produce strong autocorrelation, a potentially confounding factor in any subsequent statistical analysis.

Data and methods

To illustrate our approach, we constructed a data set composed of an index related to vegetation growth and several hypothesized physical predictors of plant growth sampled from a study site in Sequoia-Kings Canyon National Park. For our characterization of vegetation, we calculated normalized-difference vegetation index (ndvi) values from Landsat Thematic Mapper (TM) imagery. Data used for statistical analysis were subsampled from the TM scene at 1024 locations spaced at 30 m intervals along a 31 km transect (Figure 2). The transect runs across the hill slope, thus intersecting transverse draws cut by the drainage network, and not simply running uphill. We also chose several physical predictors based on current understanding of biotic response to the physical template (Stephenson, 1990, 1998; Urban et al., 2000). These included elevation (elev), topographic convergence index (tci) and analytic hillshading (sun). All of these variables influence soil water availability, a critical determinant of plant growth during the summer dry-season, albeit through different mechanisms (Stephenson, 1998; Urban et al., 2000). Increasing elevation results in lower temperatures and greater precipitation depending on local lapse rates (which are typically high in the study area). Topographic convergence index is positively related to water availability owing to drainage from upslope areas. Increased radiation tends to reduce soil moisture because of increased evaporative demand.

For our statistical analysis we used software routines available for the R statistical package (Ihaka and Gentleman, 1996). We first applied a discrete Haar-wavelet transform to the transect data,

being careful to discard coefficients obtained from wavelets overlapping the boundary of the data sequence. (In the case of the Haar wavelet, only one coefficient need be discarded at each level; a less compact wavelet would require more boundary coefficients to be discarded unless a suitable boundary correction can be applied.) We stopped the transform when the number of wavelet coefficients dropped to 15 (level 6), the minimum sample size we felt was sufficient for the results to be interpretable. We then fit a separate regression model to the coefficients at each level of the transform. We used the ndvi coefficients for the dependent variable and coefficients for elev, tci, sun and all two-way interactions for the independent variables. We then applied a step-wise model selection algorithm to arrive at a final model at each level that minimized Akaike's Information Criterion (Venables and Ripley, 1999).

We emphasize that a scale-specific approach is not restricted to our choice of linear regression with step-wise model selection. Other statistical approaches could have been used (e.g., Bayesian methods, ordination, non-linear regression) in analyzing the wavelet transformed data. In choosing widely familiar techniques, we hope to demonstrate the power of wavelets for detecting scale-specific patterns while avoiding undue confusion over the statistical methods applied at each scale of analysis.

Results

Each of the chosen variables show different patterns of variability, even as simple line transects (Figure 3). Elevation shows a broad-scale pattern of peaks and draws corresponding probably generated by water runoff. Topographic convergence exhibits a fairly noisy, fine-scale pattern. Pattern in hillshading appears to be somewhat intermediate between elevation and tci, consistent with geostatistical analyses by Urban et al. (2000). The pattern in ndvi seems to be a complex mixture of both fine- and intermediate- scale patterns.

The complexity of the ndvi pattern is further supported by the multiresolutional decomposition of

the data (Figure 4). Clearly, there are patterns exhibited at all scales of the decomposition. Furthermore, different locations in the sequence exhibit pattern at different scales. For example, the pattern is relatively intense about 1/3 of the distance down the transect at levels 2 and 3, but shifts more toward the center of the transect at levels 4 and 5 (left column, Figure 4). Wavelet transforms are ideal for characterization of these sorts of non-stationary, scale-specific patterns.

The overall intensity of pattern at different scales can be quantified by computing the variance of the wavelet coefficients at each level of the transform. Wavelet variance is computed as the sum of the squared wavelet coefficients within a level. These terms, summed over all levels, yield simple variance for the dataset. The wavelet variance within a level divided by the total variance yields the energy at that level: the proportion of the total variance at that level or scale (Percival, 1995). A plot of the wavelet variance versus scale is known as a 'scalogram.' Similarly, one can quantify scale-specific relationships between patterns observed in different variables by computing covariances of wavelet coefficients and plotting these as a function of scale. Figure 5 plots the wavelet variance-covariance matrix for the four variables in our analysis as a function of scale. Scalograms lie along the diagonal, whereas 'coscalograms' fill the off-diagonals. Typical of noisy, multiscale patterns, wavelet variances for all variables decay rapidly when moving from course-resolution (level 6 \approx 1.6km) to fine-resolution (level 1 \approx 30m). This is fairly intuitive as it is reasonable to expect that values at nearby locations will be more similar, and thus have a smaller autocovariance, than values located far apart. The large increase in wavelet variances at course scales indicates the presence of significant pattern in the data (wavelet variances for uncorrelated random noise are independent of scale).

Different patterns appear in the wavelet covariances depending on the variables compared. Both tci-elev and sun-ndvi show increasingly negative associations with increasing scale. At increasingly broad scales, it appears that higher sun intensity reduces ndvi, perhaps due to increased water stress. As expected, tci increases with decreasing elevation as ravine bottoms (where convergence is high) must be at lower elevations than ridge crests. Elevation strongly

covaries with ndvi, but only at the coarsest scale sampled (level 6). Topographic convergence also covaries negatively with ndvi at level 5, but only weakly covaries at other levels. Note, however, that larger covariances are expected at coarser scales because the variances are also larger. Small relative covariances at smaller scales do not necessarily imply a lack of correlation.

To further explore the correlation structure between patterns in ndvi and patterns in the independent variables, we fit linear regression models to the wavelet coefficients at levels 1-6, as well as to the untransformed data and applied a step-wise procedure for model selection (Table 1). For the untransformed data (level "0"), we found significant relationships for each of the independent variables, and a weak interaction effect between elevation and hillshading. Interestingly, when the data were wavelet transformed, the model selection procedure highlighted quite different associations among the variables at different scales. Models selected at levels 1, 4, 5 and 6 gave relatively poor fits, whereas models fit at levels 2 and 3 (60m and 120m scales respectively) were significant, and more interestingly, suggest different variables control pattern in ndvi at different scales. At level 2, there was a strong relationship between elevation crossed with topographic convergence. As noted previously, topographic convergence depends crucially on elevation, so the interaction comes as no surprise. [When taken alone, neither elev nor tci were significantly related to ndvi at level 2, even though this model minimized AIC; Venables and Ripley (1999) note that AIC tends to over-fit and suggest further variable elimination based on marginality; however because our analysis is primarily exploratory, we did not pursue further adjustments to the models.] At level 3, our results show that hillshading (sun) is a significant determinant of pattern in ndvi. Although difficult to ascertain visually, it does appear that hillshading exhibits somewhat coarser-scale patterning than does topographic convergence (Figure 3), and this likely explains the sudden switch from topographic convergence to hillshading as a driver of pattern at level 3.

Because the linear regression method applied to the data is potentially sensitive to spatial autocorrelation (Keitt et al., 2002), we additionally tested the residuals of each fit for autocorrelation using

Mantel's test (Mantel, 1967; Manly, 1986). We detected significant spatial autocorrelation in the residuals when the data were not transformed ($p < 0.001$). Visual inspection of the autocorrelation function indicated significant correlation up to 450m lag distance. The results of the regression on the untransformed data must therefore be considered with some skepticism. (Since we are more interested here in scale-specific analysis, we did not bother to account for the autocorrelation in the residuals of the untransformed regression, although methods to do so are readily available.) Little or no autocorrelation was observed in any of the regressions on wavelet-transformed data. Only at level 3 did we observe a marginally significant autocorrelation ($p = 0.02$; two-tailed comparison). Plotting the autocorrelation function for the level 3 residuals revealed minor autocorrelation that decayed rapidly with lag distance. No pattern was observed beyond the 30m (nearest neighbor) lag scale. Thus, wavelet-transforming the data eliminated any meaningful spatial autocorrelation as a confounding factor in our analysis.

Discussion

Our initial application of wavelets to the problem of scale-specific inferences appears quite promising: the method highlights strong differences in the scales at which physical environmental factors are associated with vegetation patterns in the Sierra Nevada. Topographic effects on water availability appear to dominate vegetation patterns at a scale roughly half that of solar inputs.

This is likely because shading is strongly influenced by the larger gullies and ravines which cast shadows over large areas blocking any effect of finer scale topography. On the other hand, effects of topography and topographic convergence on water availability do not have a shadow effect, and thus can influence vegetation at the finest scales of topographic relief. It is rather appealing feature of scale-specific inference that the spatial signatures of these different mechanisms of water availability and their effects on vegetation can be easily detected and separated by scale.

There are a number of areas in ecology where we believe wavelet-coefficient regression could lead to interesting new insights. For example, environmental control over species distribution and

abundance, a core topic in ecology, would seem to be a natural area for scale-specific analysis, as patterns in the environment and species distributions are typically manifest across a wide range of scales (Kendal, 1992; Maurer, 1994). A species distribution may be controlled at fine scales by the presence of riparian habitat, but at a broader scale, by topography and rainfall patterns that determine the distribution of riparian habitats across the landscape. We would expect that testing for scale-specific associations using the methods presented here would nicely highlight important environmental controls of species distributions at different scales. An interesting question arises in the appropriate handling of presence-absence data. Logistic regression is the typical approach. However, after wavelet-transformation, the data would no longer be strictly presence-absence, but instead would contain both positive and negative real numbers ruling out the use of the logistic framework. We have yet to explore this problem in detail, but it does raise the interesting question of whether binary wavelet-like transforms can be defined for presence-absence data.

Results from our initial application of wavelet-coefficient regression appear quite promising. Nonetheless, some caution is warranted, as significant statistical and ecological issues have yet to be explored. We have illustrated applications of the Haar wavelet, but myriad other wavelets are available and these yield slightly different results as a consequence of the degree to which the shape of the wavelet matches the dominant pattern in the data. For example, the Haar square-wave is well suited to detecting abrupt edges, while a smoother wavelet might be more appropriate to topographic undulations. For our application, the nuance of wavelet shape is unlikely to have a significant effect as we are principally interested in extracting scale-specific covariance information. As long as the wavelet does a reasonable job of isolating variance in narrow frequency bands, it should perform adequately. Nonetheless, statistical packages that support wavelet analysis often do include automated functions for discovering good wavelet kernels for specific tasks, such as image compression. It is yet to be determined, and a topic of interest for future studies, whether these algorithms are of use in enhancing the detection of scale-specific associations in ecological data.

In this study, we have followed Mallat's (1989) pyramid scheme (Figure 1) for computing the wavelet transform. The pyramid scheme has the limitation that the scale of analysis jumps by a power of two in each successive level of the transform. If interesting covariation occurs at scales falling between the levels analyzed, they may be assigned to the wrong scale or missed altogether. Wavelet transforms may also be calculated using a continuous (albeit discretely sampled) wavelet function (Daubechies, 1992). The continuous transform permits analysis across a continuum of scales allowing the investigator to choose how finely to explore the frequency spectrum. The cost is a large overlap in the scales analyzed as one makes finer and finer steps. It seems convenient to minimize overlap between levels so that models produced at each scale are relatively independent (and represent an additive decomposition of the original variance). However, the implications of this trade-off of independence for scale-resolution have not been explored.

We have illustrated wavelet applications in a single dimension, along transects. Even in this case, the logistical expense implied by the long transect (> 30 km) is essentially beyond field-sampling methods and instead invites applications to remotely sensed data. In this sense, it is especially compelling to consider applications to 2-dimensional data such as satellite imagery. Wavelet applications are well developed to higher-dimensional data in other disciplines, but less so in ecology.

As with any statistical analysis, some care is needed also in interpreting the results of wavelet-coefficient regressions. Notice that the residual errors are also wavelet transformed in the wavelet-coefficient model (Eq. 5), a procedure that may alter the error distribution and interdependence. Transformation of the residuals is typically less of a problem than it might seem, and can have beneficial side-effects. Generally, wavelet-transformed residuals will not show significant autocorrelation. This is apparent when one considers that the wavelet decomposition separates a sequence into low-frequency trend (the smoothed sequences) and maximally uncorrelated detail (the wavelet coefficients). Thus, wavelet coefficient-regression will tend to meet assumptions of independence, even when the input data are strongly autocorrelated.

We found exactly this case in the results presented here. However, as with any regression analysis, it is wise to test for normality, independence and homoscedacity of errors before placing great confidence in the estimated parameters.

Development and interpretation of significance tests is further complicated by the inverse relationship between the number of coefficients and the scale or level of analysis; at coarse scales, the very few coefficients make it difficult to meet parametric tests of significance because of the low degrees of freedom. Yet much of the variance, in an absolute sense, tends to occur at these coarse scales. By contrast, somewhat less variance tends to be concentrated at very fine scales, yet the large number of coefficients make it easier to meet conventional tests of significance. Some artful blending of scale-specific and global tests of significance might be in order. Related to this issue is a question about how generalizable or transportable the results of wavelet-based regression models might be. For example, can a wavelet regression be used in predictive mode using data independent of the data used to fit the model? We have not explored this issue with our data, though the results of such tests would dictate whether such regressions could be used in inferential mode for a particular dataset or in predictive mode for other datasets. If applications are intended primarily to interpret a single dataset, then this might invite applications of alternative regressions with more within-dataset flexibility (locally weighted regressions, LOESS, and so on).

Although these and other issues remain to be resolved, it is clear that scale-specific analysis using wavelets hold great promise for ecological applications concerned with multivariate, multi-scaled patterns. We offer this introduction as an indication of this promise and as a benchmark for further developments in this new approach.

Acknowledgments

We thank P. Dixon and two anonymous reviewers for constructive comments that greatly improved this manuscript.

References

- Bradshaw, G. A. and T. A. Spies. 1992. Characterizing canopy gap structure in forests using wavelet analysis. *Journal of Ecology* **80**:205–215.
- Cressie, N. A. C. 1993. *Statistics for Spatial Data*. revised edition. John Wiley & Sons, Inc., New York.
- Dale, M. R. T. 1999. *Spatial Pattern Analysis in Plant Ecology*. Cambridge University Press, Cambridge.
- 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.
- Daubechies, I. 1992. *Ten Lectures On Wavelets*. CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia.
- Diggle, P. J. 1990. *Time Series: A Biostatistical Introduction*. Oxford Science Publications, Oxford.
- Greig-Smith, P. 1964. *Quantitative plant ecology*. second edition. Butterworths, London.
- Hastie, T. J. and R. J. Tibshirani. 1990. *Generalized Additive Models*. Chapman and Hall, London.
- Ihaka, R. and R. Gentleman. 1996. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* **5**:299–314. [Software available at <http://www.r-project.org/>].
- Keitt, T. H. 2000. Spectral representation of neutral landscapes. *Landscape Ecology* **15**:479–494.

- Keitt, T. H., O. N. Bjørnstad, P. Dixon, and S. Citron-Pousty. 2002. Accounting for spatial pattern when modeling environment-abundance relationships. *Ecography* **25**:616–625.
- Kendal, W. S. 1992. Fractal scaling in the geographic distribution of populations. *Ecological Modelling* **64**:65–69.
- Levin, S. A. 1992. The problem of pattern and scale in ecology. *Ecology* **73**:1943–1967.
- Mallat, S. G. 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**:674–693.
- Manly, B. F. J. 1986. Randomization and regression methods for testing for associations with geographical, environmental and biological distances between populations. *Researches on Population Biology* **28**:201–218.
- Mantel, N. 1967. The detection of disease clustering and a general regression approach. *Cancer Research* **27**:209–220.
- Maurer, B. A. 1994. *Geographical Population Analysis: Tools for the Analysis of Biodiversity*. Blackwell Scientific Publications, Oxford.
- Noy-Meir, I. and D. J. Anderson, 1971. Multiple pattern analysis, or multiscale ordination: Towards a vegetation hologram? Pages 207–301 in G. Patil, E. C. Pilou, and W. E. Walters, editors. *Many-Species Populations, Ecosystems, and Systems Analysis*. Statistical Ecology Series 3, Penn. State Univ. Press, University Park, PA.
- Percival, D. B. 1995. On estimation of the wavelet variance. *Biometrika* **82**:771–787.
- Ramsey, J. 1999. The contribution of wavelets to the analysis of economic and financial data. *Philosophical Transactions of the Royal Society of London, Series A* **357**:2593–2606.
- Shumway, R. H. and D. S. Stoffer. 2000. *Time Series Analysis and its Applications*. Springer, New York.

- Stephenson, N. L. 1990. Climatic controls on vegetation distribution: The role of water balance. *American Naturalist* **135**:649–670.
- Stephenson, N. L. 1998. Actual evapotranspiration and deficit: Biologically meaningful correlates of vegetation distribution across spatial scales. *Journal of Biogeography* **25**:855–870.
- Urban, D., S. Goslee, K. Pierce, and T. Lookingbill. 2002. Extending community ecology to landscapes. *Ecoscience* **9**:200–212.
- Urban, D. L., C. Miller, N. L. Stephenson, and P. N. Halpin. 2000. Forest pattern in Sierran landscapes: The physical template. *Landscape Ecology* **15**:603–620.
- Venables, W. N. and B. D. Ripley. 1999. *Modern Applied Statistics with S-PLUS*. Third edition. Statistics and Computing, Springer, New York.
- van der Hoef, J. M. and D. C. Glenn-Lewin. 1989. Multiscale ordination: a method for detecting pattern at several scales. *Vegetatio* **82**:59–67.
- Wagner, H. H. 2003. Spatial covariance in plant communities: Integrating ordination, geostatistics, and variance testing. *Ecology* **84**:1045–1057.
- Wagner, H. H. 2004. Direct multi-scale ordination with canonical correspondence analysis. *Ecology* **85**:342–351.
- Walker, J. S. 1999. *A Primer on Wavelets and Their Scientific Applications*. Chapman and Hall, New York.

Table 1. Results of linear models fit to wavelet coefficients.

Level	N	Model	F	AIC
0 ^a	1024	elev*** + tci** + sun** + elev:sun*	31.21***	6117
1	511	elev + sun + tci + elev:tci* + elev:sun	1.49	2168
2	255	elev + tci + elev:tci***	3.18**	1334
3	127	sun**	4.24**	793
4	63	elev + sun + elev:sun	1.45	451
5	31	intercept only	NA	234
6	15	tci + sun** + tci:sun**	2.33	125

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; ^aUntransformed data

Figure legends

Figure 1. Haar wavelet transform of an autoregressive time (or equivalently, space) series defined by $x_{t+1} = -\frac{1}{2}x_t - \epsilon$ where $\epsilon \sim N(0, 1)$. The right hand column shows detail coefficients calculated as in Eq. 1. The left column shows the original sequence and smooth coefficients calculated as in Eq. 2. In this example, the series oscillates rapidly and most of the variance is transferred to the d1 coefficients. Little variance remains by the third level of the transform. Many space and time series in nature show increasing variance at higher levels of the transform (see for example, Figure 6).

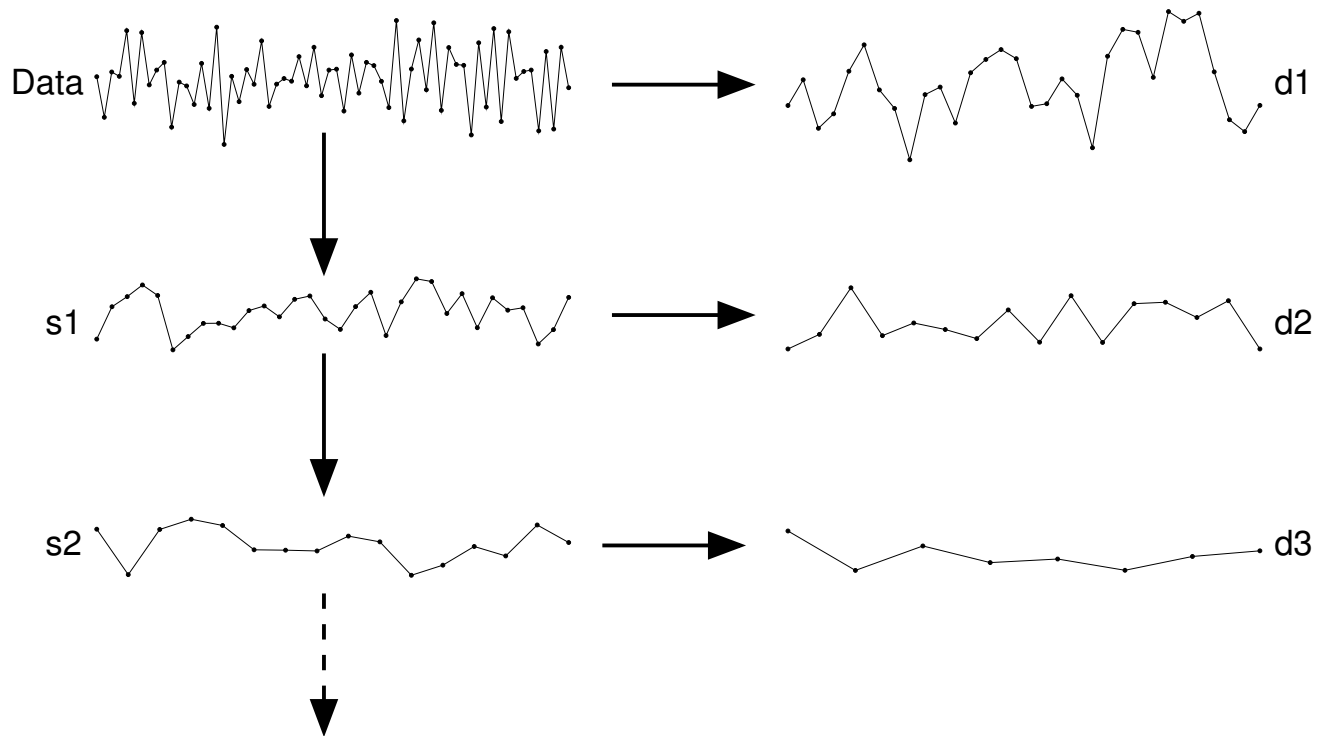
Figure 2. Image of the study site showing data transect (green line) overlaid on terrain. Blue lines show drainage network.

Figure 3. Values of measured variables sampled at 30m intervals along the transect show in Figure 2.

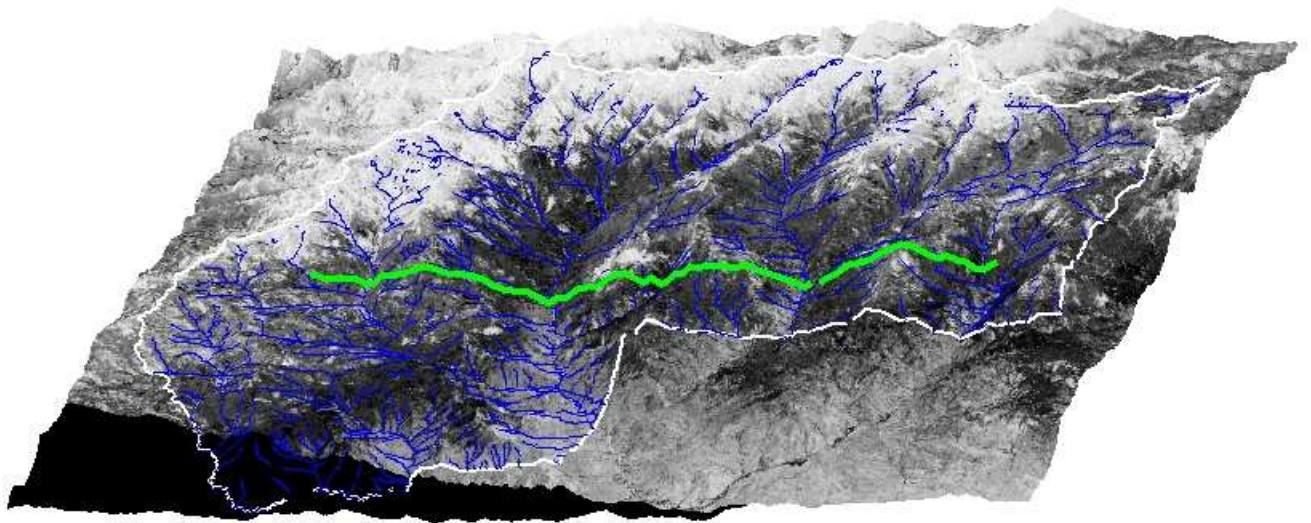
Figure 4. Multiresolutional decomposition of the ndvi data. The left-hand column shows the decomposition at each scale (1-6). The right-hand column shows the sum of the preceding scale components such that $s_k = s_{k+1} + d_{k+1}$.

Figure 5. Wavelet variance-covariance matrix plotted as a function of scale. Variances or 'scalograms' run along the diagonal; covariances or 'coscalograms' fill the off-diagonals. Dotted lines indicate zero (co)variance.

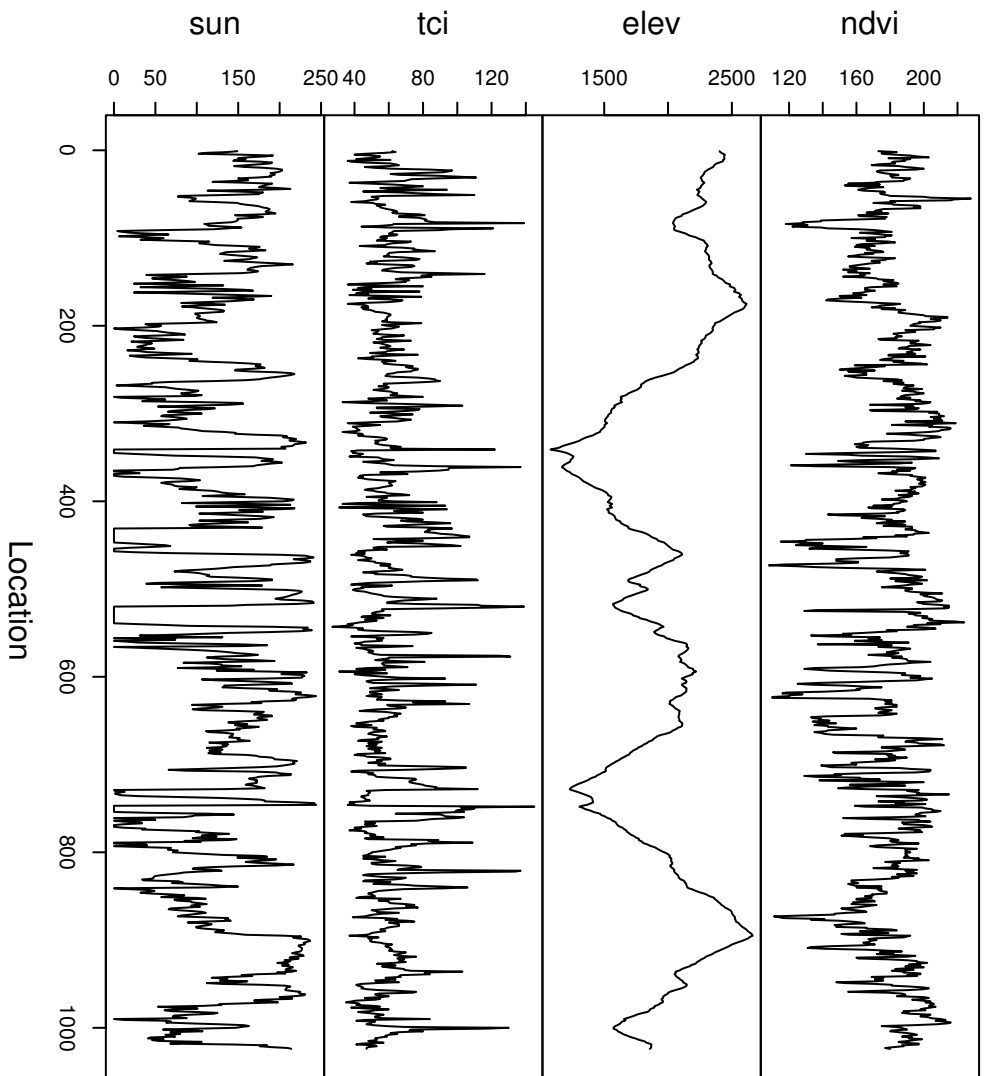
[Figure 1.]



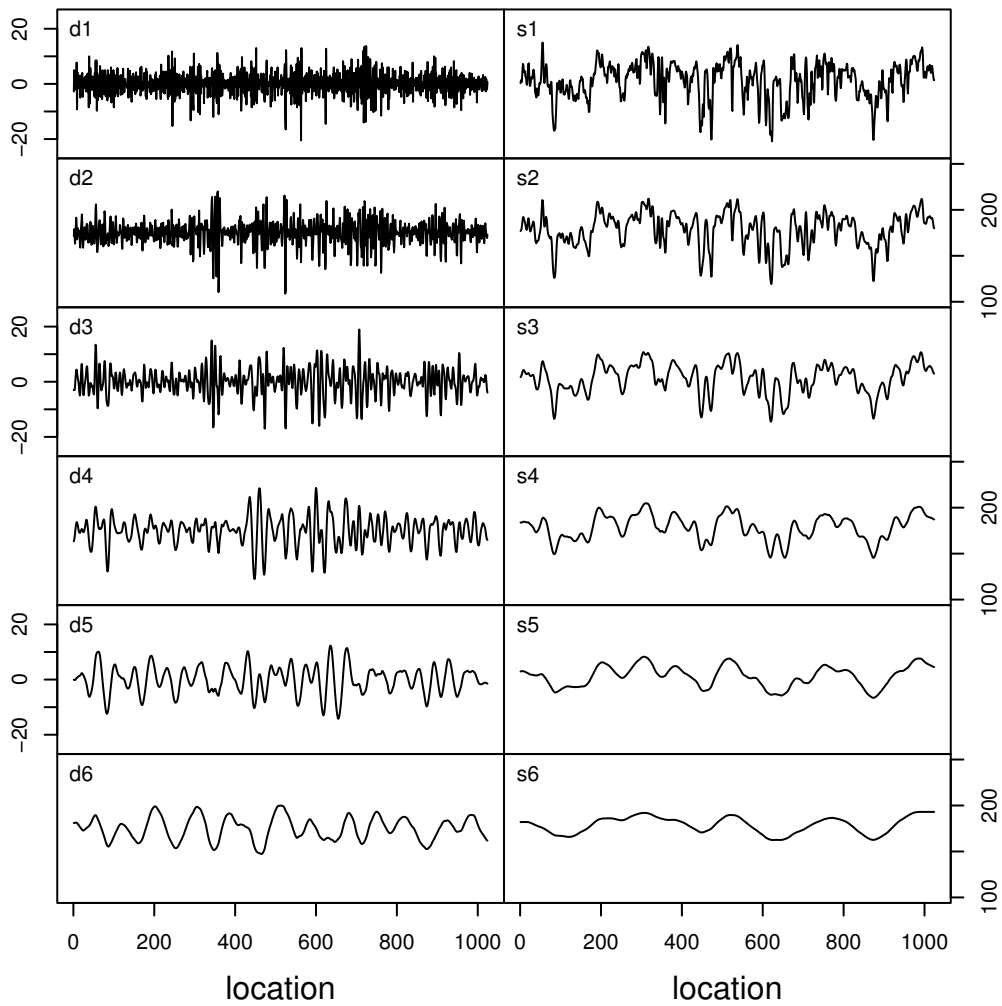
[Figure 2.]



[Figure 3.]



[Figure 4.]



[Figure 5.]

