The Pennsylvania State University
The Graduate School
College of Earth and Mineral Sciences

NEW METHODS IN WAVELET ANALYSIS AND THEIR APPLICATION TO THE UNDERSTANDING OF HYDROCLIMATE VARIABILITY IN THE MID-ATLANTIC REGION OF THE UNITED STATES

A Dissertation in
Meteorology
by
Justin Alexander Schulte

© 2015 Justin Alexander Schulte
Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2015

The dissertation of Justin Alexander Schulte was reviewed and approved∗ by the following:

Raymond G. Najjar
Professor of Oceanography
Dissertation Advisor
Chair of Committee

Sukyoung Lee
Professor of Meteorology

Chris E. Forrest
Associate Professor of Climate Dynamics

Christopher J. Duffy
Professor of Hydrology

Ming Li
Professor of Oceanography
Special Member

Johannes Verlinde
Professor of Meteorology
Associate Head of Department of Meteorology

∗Signatures are on file in the Graduate School

Abstract

New statistical and topological methods in wavelet analysis are developed to help fully understand historical hydroclimate variability. The methods improve upon existing methods and enhance confidence in results. In particular, a geometric significance test was found to produce results similar to an existing areawise significance test while being more computationally flexible and efficient. The geometric test determined that features in wavelet power of the North Atlantic Oscillation (NAO) are indistinguishable from a red-noise background, suggesting that the NAO is a stochastic, unpredictable process, which could render difficult the future projections of the NAO under a changing global system. The geometric test did, however, identify features in the wavelet power spectrum of an El Niño index (Niño 3.4) as distinguishable from a red-noise background. A topological analysis of pointwise significance patches determined that holes, deficits in pointwise significance embedded in significance patches, are capable of identifying important structures, some of which are undetected by the geometric and areawise tests. Niño 3.4 time series contains numerous phase-coherent oscillations that could be interacting nonlinearly. Another statistical test, the cumulative areawise test, was also developed and was found to have greater statistical power than the geometric test in most cases, especially when the signal-to-noise ratio is high. The new testing procedure was applied to the time series of the Atlantic Multi-decadal Oscillation (AMO), the NAO, Pacific Decadal Oscillation (PDO), and the Niño 3.4 index and determined that the NAO, PDO, and AMO are consistent with red-noise processes, whereas significant power was found in the 2-7 year period band for the Niño 3.4 index. High-order wavelet methods were developed to understand nonlinear climate phenomenon. A local autobicoherence spectrum of the QBO time series showed that the QBO time series contained a mode with a period of 28 months that was phase-coupled to a harmonic with a period of 14 months. An additional nonlinearly interacting triad was found among modes with periods of 10, 16, and 26 months. The oscillations were found to be skewed so that negative QBO regimes were preferred, and also asymmetric in the sense that phase transitions between the easterly and westerly phases occurred more rapidly than those from westerly to easterly regimes. The developed methods were applied to Mid-Atlantic region hydroclimatic time series. The results from the wavelet analysis showed that 18- and 26-year periodicities were embedded in the streamflow time series of the Delaware, Hudson, and Susquehanna Rivers. Decadal variability of streamflow was coherent with the El-Niño/Southern Oscillation and the PDO. The Southern Oscillation explained 37-54% of the 1960s drought, 33-49% of the 1970s pluvial, and 19-50% of the 2000s pluvial in the three river basins. A composite analysis with meteorological data determined that the anomalously high daily streamflow events during the recent pluvial were associated with Rossby waves emanating from the tropics and sea-level pressure (SLP) dipole pattern over eastern North America (ENA). A new ENA SLP dipole index was constructed, which could explain more daily streamflow variance than the existing climate indices. The ENA dipole index was also found to be

phase-locked to the Gulf Stream Index at a period of 74 months so that salinity variability at that timescale may have resulted from ENA dipole-related changes in precipitation.

TABLE OF CONTENTS

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to thank Dr. Raymond G. Najjar for all his help during my graduate school career. His encouraging words and constructive critiques helped me improve as a scientist and as a scientific writer. I am also thankful for his flexibility in allowing me to pursue research problems in other fields. I would also like to acknowledge my committee members for their advice and guidance throughout the dissertation process. I also give thanks to Dr. Nigel Higson from the Department of Mathematics for his advice with algebraic topology.

I would also like to acknowledge Astrid Suarez for all the great discussions of wavelet analysis and suggestions. I also thank Andrew Ross for computing programing suggestions and helping me develop the Matlab software packages. Finally, I give thanks to all my friends and family whom have inspired and helped me throughout my academic career.

This work was funded by the National Science Foundation (award number GF/02/14), the Hudson River Foundation (award number 0961423), and Pennsylvania Department of Environmental Protection.

Dedication

This work is dedicated to my mother for her never-ending support and encouragement throughout my academic career.

Chapter 1

Introduction

1.1 Background and Motivation

Hydroclimate variability across a region has impacts on many societal systems. Hydroclimate variability includes droughts and wet periods (pluvials). During droughts, water resources are stressed and agriculture is strained, which may have economic implications. Dry periods may also increase the risk of forest fires, which can devastate forests and nearby communities. Floods related to wet periods also pose threats to human safety and property. Floods, for example, can inundate towns and cities and relief efforts can be costly. In some regions, excessive rainfall can create mudslides, which are particularly destructive and can occur with little warning. In other regions, excessive precipitation results from hurricanes, powerful storms that are capable of producing enormous amounts of precipitation.

Given the potential societal impacts of hydroclimate variability, it is important to study the causes of past hydroclimate variability. Two primary ways of studying hydroclimate variability are the analysis of long data sets and modeling. In modeling studies, one has the flexibility, for example, to prescribe sea surface temperature in a oceanic region to understand their impacts on regional climate variability. Models also have the advantage of allowing the construction of data sets otherwise unavailable. For example, models may generate data at higher temporal resolutions than observed data sets. The analysis of observed data sets involves the implementation of various statistical and time series analysis methods and the assessment of statistical uncertainty of the results. The use of the time series analysis methods, however, are limited by the length of the data sets and prevalence of missing data. Many procedures in spectral analysis, for example, assume that data are evenly spaced. Some example of data analysis methods include: singular spectrum analysis, Fourier analysis, wavelet analysis, regression analysis, and correlation analysis. In the time domain, statistical tools often seek relationships between two time series. One assumption of many time domain statistical tools is stationarity, meaning that

statistical moments, such as first- and second-order moments, of a time series are time-independent. In practical applications, the inappropriate application of the methods may result in misleading results. Another problem with time domain statistics is that many climatic time series have characteristic time scales and therefore establishing relationships between time series at all timescales simultaneously could be misleading. One way to circumvent the problem is to smooth time series, reducing noise and allowing statistical relationships to emerge. The approach, however, is subjective; one must choose the degree to which the time series are smoothed. A more objective approach is Fourier analysis, which decomposes time series into frequency components. The frequency decomposition allows characteristic time scales to be extracted and also the strength of the relationship between two time series to be assessed at different frequencies. The main assumption of Fourier analysis is stationarity, which limits its application. Other problems include finding an appropriate data taper to minimize spectral leakage, bias, and variance of the spectral estimate. Problems associated with Fourier analysis can be remedied by using wavelet analysis. In wavelet analysis, the assumption of stationarity is relaxed and thus the method provides more flexibility in its application. An additional benefit is that the window width, unlike in short-time Fourier analysis, is not fixed, maximizing resolution at low-frequencies. One can also detect nonstationary relationships between two time series at different frequencies, making wavelet analysis for climate studies particularly useful.

A singular implementation of wavelet analysis, however, is inadequate for feature extraction of geophysical time series. For example, red-noise, a typical null hypothesis against which climatic time series are often tested, can produce large spectral power. To facilitate the interpretation of results from wavelet analysis, statistical methods have been developed, the first of which is pointwise significance testing (Torrence and Compo, 1998). In a pointwise significance test, results at each frequency and time are tested against a suitable null hypothesis with the assumption that all wavelet coefficients are independent. The application of the procedure yields subsets of the wavelet domain containing elements whose likelihood of being generated by noise is less than those not contained in the subsets. This procedure was found by Maraun et al. (2007) to be problematic for two reasons: (1) wavelet coefficients are not independent and (2) the number of wavelet coefficient to which the test is applied can be very large, resulting in numerous spurious results. Addressing the

limitations of the pointwise significance test, Maraun et al. (2007) developed an areawise test to dramatically reduce the spurious results. While the test is a great improvement from the pointwise test, its practical application is limited by computational inflexibility. The application of the test requires the calculation of a new critical level for different analyzing wavelets and pointwise significance levels.

The application of standard wavelet analysis, in some cases, may be inadequate for feature extraction of time series. This inadequacy is found analogously in the time domain, where the mean and variance can only partially explain the distribution of time series for non-normally distributed data. In such cases, it is useful to look at high-order moments such as skewness and kurtosis. In the frequency domain, by analogy, one computes the bispectra of time series to determine what frequency components are contributing to the third-order moment of the time series. The sunspot cycle and the quasi-biennial oscillation (QBO) are both examples of nonlinear geophysical phenomena and thus high-order spectral analysis is needed to more fully understand these time series (Moussas et al., 2005; Rusu, 2007; Lu et al., 2009).

The highly populated Northeast US is home to Hudson River, Delaware River, Chesapeake Bay estuaries, three large estuaries of economic, industrial, recreational importance, which help bolster the very populated metropolitan region. The Delaware River estuary, for example, contains the largest freshwater seaport in the world, which is situated in Philadelphia. The Delaware River estuary also contains important shipping channels through which commodities are transported. Additional industrial uses of the estuaries are the cooling of power plants provided that the water is not saline. The three estuaries also support local economies through fishing.

The Northeast US is a region that exemplifies societal impacts of droughts. In the 1960s, a severe drought afflicted the highly populated region and created conflicts between New York City and Philadelphia. As a result of the 1960s drought, the water diverted to New York City from the upper Delaware River watershed is now regulated and limited by the downstream flow conditions. The reason for the regulation is that low flow conditions near Philadelphia will allow saline water to intrude into the Philadelphia water supply. The impact of the drought has led many researchers to investigate its causes. For example, Ning

and Bradley (2015) found the 1960s drought to be related to El-Niño/Southern Oscillation (ENSO), which implies that such events may be predictable. In contrast, Seager et al. (2012) found, using climate models, that the drought was the result of intrinsic atmospheric variability, rendering future droughts impossible to predict. The contrasting results merit an additional investigation and thus the goal of the dissertation is to determine mechanisms governing historical streamflow variability in the mid-Atlantic region.

To achieve the goal, it will be necessary to implement wavelet analysis and develop new statistical tests to assess the confidence in the results. The new statistical tests will improve on the pointwise significance test. Wavelet analysis was adopted by Labat (2008) and Whitney (2010) for their investigations of mid-Atlantic streamflow variability. While Whitney (2010) detected significant variability in Susquehanna streamflow at a period of 26 years, the variability was not linked to physical mechanisms. On the other hand, the cross-wavelet analysis by Labat (2004) determined that North American continental freshwater discharge was related to ENSO at a period of 22 years. One deficiency of the aforementioned studies is a lack of rigorous statistical significance testing; only pointwise significance testing was applied in the studies. Therefore, it will be necessary to development tools in wavelet analysis to gain more confidence in the results.

The three main unanswered questions that will be addressed in the dissertation are the following:

1) Can new statistical procedures in wavelet analysis be developed that improve upon existing methods?

2) Did mid-Atlantic streamflow time series contain characteristic time scales?

3) Did climate modes have impacts on historical streamflow variability and what time scales were their influences the strongest?

Question (1) will be addressed by using ideas from geometry, topology, and existing methodologies. Question (2) will be addressed using wavelet analysis together with new significance testing procedures. Question (3) will be addressed using correlation, wavelet, and composite analyses.

1.2 Dissertation Overview

In this work, new and existing procedures in wavelet analysis will be used to understand historical hydroclimate variability in the mid-Atlantic region of the US. In particular, in Chapter 2 a new geometric significance test and topological methods are developed to remedy the problem with existing statistical procedures. In Chapter 3, a more flexible and powerful cumulative areawise test is constructed and applied to well-known geophysical time series. High-order wavelet methods are developed in Chapter 4 and are applied to the QBO time series. Chapter 5 will consist of a wavelet coherence analysis that will relate climate modes to streamflow at an array of timescales. Salinity and streamflow variability are investigated in Chapter 6 and linked to prevailing atmospheric conditions. Chapter 7 will include additional applications of wavelet analysis to streamflow time series. Concluding remarks and a discussion are provided in Chapter 8.

Chapter 2 contains work published in *Nonlinear Processes in Geophysics* (Schulte et al., 2015) that was co-authored by C. Duffy and R. Najjar. Material in Chapter 3 was published as a discussion paper in *Nonlinear Processes in Geophysics Discussions* and is currently under review. The results presented in Chapter 4 have been prepared as a manuscript and will be submitted in the near future. Chapter 5 contains work that was submitted to the *Journal of Hydrology: Regional Studies* and the co-authors of the paper are R. Najjar and M. Li. A manuscript co-authored with R. Najjar and S. Lee containing the results in Chapter 6 is currently under preparation and will be submitted in the near future. The results in Chapter 7 will likely be part of another manuscript.

Chapter 2

Geometric and Topological Approaches to Significance Testing in Wavelet Analysis

2.1 Introduction

Time series are often complex, composed of oscillations and trends. The goal of researchers is to decide whether the embedded structures in the time series are stochastic or deterministic. Such decisions can be made using Fourier analysis, with the assumption that the underlying time series is stationary (Jenkins and Watts, 1968). In many cases, however, the stationary assumption is not satisfied, making Fourier analysis an inappropriate tool for feature extraction. For non-stationary time series, wavelet analysis (Meyers, 1993; Torrence and Compo, 1998) can be used for decomposing a time series into both frequency and time components, allowing the extraction of transient features and dominant modes of variability. Once embedded structures in time series have been identified, a natural question arises: what physical mechanisms are responsible for the detected modes of variability? Linkages between the modes of variability and possible physical mechanisms can be obtained using wavelet coherence (Grinsted et al., 2004), a bivariate tool for detecting common oscillations between two time series. Together, wavelet power and coherence analyses have proven useful in climate science (Velasco and Mendoza, 2007; Muller et al., 2008), hydrology (Zhang et al., 2006; Ozger et al., 2009; Labat, 2008; Labat, 2010), atmospheric science (Terradellas et al., 2005; Schimanke et al., 2011), and oceanography (Lee and Lwiza, 2008).

The application of wavelet analysis alone is not sufficient for feature extraction of time series; indeed, random fluctuations can produce large values of spectral power or coherence related to the underlying process (e.g., red-noise) and not necessarily the time series. In Fourier analysis, one chooses a suitable noise model and assesses the significance of features relative to some analytically or empirically derived threshold. In climate science, for example, one often compares the sample power spectrum of a time series to that of a theoretical red-noise spectrum (Hasselman, 1976; Torrence and Compo, 1998). Statistical

significance testing is also necessary in the wavelet domain. Torrence and Compo (1998) were the first to assess the significance of features in wavelet power spectra using discrete red-noise background spectra. Grinsted et al. (2004), using Monte Carlo methods, extended significance testing to wavelet coherence using surrogate red-noise time series. The (pointwise) significance tests developed by Torrence and Compo (2010) and Grinsted et al. (2004), however, have multiple-testing problems, given the large number of wavelet coefficients being tested simultaneously (Maraun and Kurths, 2004). Suppose, for example, that a pointwise significance test was applied to M wavelet power coefficients at the 5% significance level. Then, on average, there will be $0.05M$ false positive results, which would make the pointwise test permissive for large M. Maraun et al. (2007) addressed these problems by developing an areawise test that sorts through contiguous regions of pointwise significance called significance patches based on their area and geometry, minimizing spurious results, and thus giving researchers more insight into the time series in question. According to the areawise test, the larger the pointwise significance patch, the less likely it was generated from a stochastic fluctuation.

In this study, significance testing in the wavelet domain is improved through the following: (1) the development of a flexible and computationally efficient geometric test capable of minimizing spurious results from the pointwise test by associating p-values to individual patches in wavelet-power and wavelet-coherence spectra; and (2) the application of topological methods that can further distinguish spurious patches from true structures that can reveal information about time series undetected by current methods. Given the deficiencies of pointwise significance testing, there is a need to improve current methods of evaluating significance of features in the wavelet domain. The areawise test, though a substantial improvement from the pointwise test has one drawback: finding the significance level of the areawise test requires a complicated root-finding algorithm, making p-values for the areawise test difficult to obtain, as it would require the repeated application of a root-finding algorithm (see Sect. 4.1 for details).

The remainder of the paper is organized as follows. A brief overview of wavelet analysis is presented in Sect. 2.2. In Sect. 2.3, the pointwise and areawise tests are discussed briefly. The development of the geometric test is presented in Sect. 2.4. In Sect. 2.5, ideas

inspired by persistence homology (Edelsbrunner, 2010) are used to show that holes, voids of pointwise significance surrounded by regions of pointwise significance, can distinguish important structures from trivial structures, linking the geometric and topological tests. Using ideas from Sect. 2.4 and Sect. 2.5, the application of a local geometric test is presented in Sect. 2 6. The new methods are applied to time series of two idealized cases, which provide important benchmarks for the methods, and to indices of two prominent climate modes, El Niño/Southern Oscillation and the North Atlantic Oscillation (NAO), to illustrate, in a geophysical setting, the insights afforded by the methods.

2.2 Definitions

In wavelet analysis, a time series is decomposed into frequency and time components by convolving the time series with a wavelet function satisfying certain conditions. There are many different kinds of wavelet functions but the most widely used is the Morlet wavelet, a sine wave damped by a Gaussian envelope expressed as

$$\psi_0(\eta) = \pi^{-1/4} e^{i\omega_0\eta} e^{-\frac{1}{2}\eta^2}, \tag{2.1}$$

where ψ_0 is the Morlet wavelet, ω_0 is the dimensionless frequency, and $\eta = s \cdot t$, where s is the wavelet scale, and t is time (Torrence and Compo, 1998; Grinsted et al., 2004). The wavelet transform of a discrete time series x_n (n = 1, ..., N) is given by

$$W_n^X(s) = \sqrt{\frac{\delta t}{s}} \sum_{n'=1}^{N} x_{n'} \psi_0[(n'-n)\frac{\delta t}{s}], \tag{2.2}$$

where δt is a uniform time step determined from the time series and $|W_n^X(s)|^2$ is the wavelet power of a time series at scale s and time index n (Torrence and Compo, 1998; Grinsted et al., 2004). Note that for the Morlet wavelet with ω_0 = 6 the wavelet scale and the Fourier period λ are approximately equal ($\lambda \approx 1.03s$).

2.3. Existing Significance Testing Methods

2.3.1 Pointwise Significance Testing

For climatic time series, the significance of wavelet power can be tested against a theoretical red-noise background (Torrence and Compo, 1998). For a first-order autoregressive (Markov) process

$$X_n = \alpha X_{n-1} + w_n \tag{2.3}$$

with lag-1 autocorrelation coefficient α, Gaussian white noise w_n, and $X_0 = 0$, the normalized theoretical red-noise power spectrum is given by

$$P_f = \frac{1-\alpha^2}{1+\alpha^2-2\alpha\cos(2\pi f/N)}, \tag{2.4}$$

where $f = 0,\ldots, N/2$ is the frequency index (Gilman et al., 1963). To obtain, for example, the 5% pointwise significance level one must multiply Eq. (2.4) by the 95% percentile of a chi-square distribution with two degrees of freedom and divide the result by 2 to remove the degree of freedom factor (Torrence and Compo, 1998). The discrete Fourier red-noise spectrum has been shown by Torrence and Compo (1998) to be adequate in estimating the significance of local wavelet power and is thus used in this paper to estimate pointwise significance. The parameter α can be estimated using standards methods such as the Burg's and the Yule-Walker methods (Kay, 1988; Hayes, 1996).

Monthly time series and normalized wavelet power spectra for the NAO index (Hurrell et al., 1995, https://climatedataguide.ucar.edu/climate-data/hurrell-north-atlantic-oscillation-nao-index-station-based) and the Niño 3.4 index (Trenberth 1997, http://www.cgd.ucar.edu/cas/catalog/climind/Nino_3_3.4_indices.html) are shown in Figs. 2.1 and 2.2. The Niño 3.4 index data were converted to anomalies by subtracting the mean monthly values for each month from the monthly values. Note that the normalized wavelet power is the wavelet power at every time and period divided by the variance of the time series, which allows different wavelet power spectra to be readily compared. Another important feature of the wavelet power spectrum is the cone of influence, the region in which edge effects become important, or more precisely, the *e*-folding time of the autocorrelation for wavelet power at each scale, where the *e*-folding time is defined by Torrence and Compo (1998) as the point at which the wavelet power for a discontinuity at the edge drops by a factor of e^{-2}. The wavelet power spectrum of the NAO index reveals numerous time periods of enhanced variance at an array of time scales, though no preferred

timescale is evident. For the Niño 3.4 index, the wavelet power spectrum detects statistically significant variance in the 16-64 month period band for the period 1960-2010. Another interesting feature emerges (labeled *H* in Fig. 2.2b): regions of no pointwise

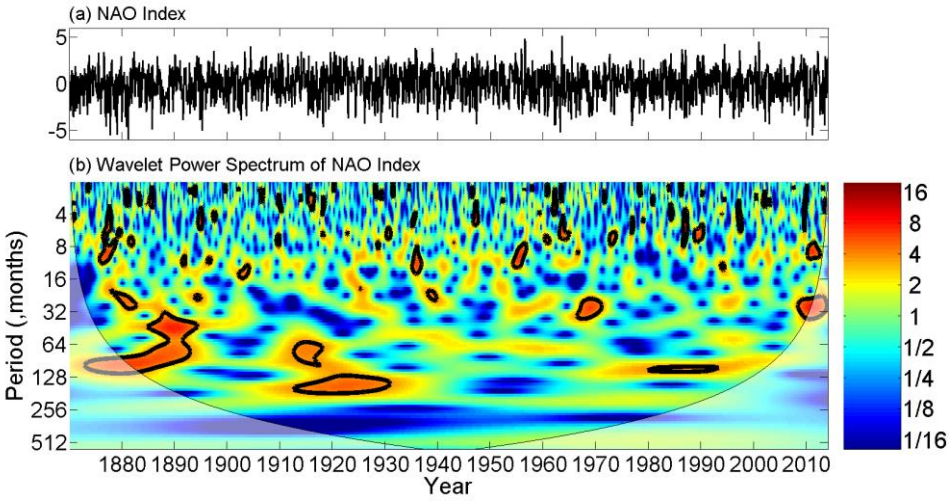

Figure 2.1. (a) The NAO index from 1870 to 2013. (b) The normalized wavelet power spectrum of the NAO index. Thick contours enclose regions of 5% pointwise significance. Light shading corresponds to the cone of influence, the region in which edge effects become important.

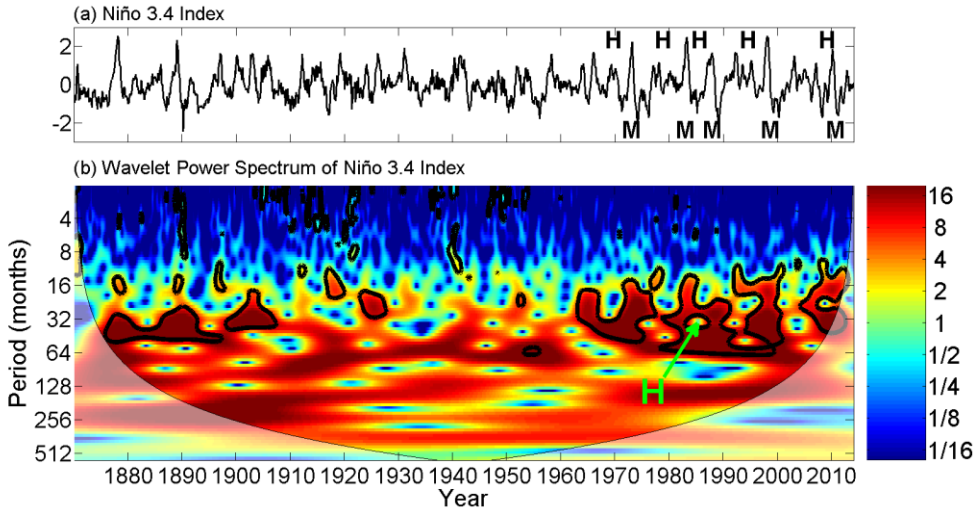

Figure 2.2. (a) The Niño 3.4 index time series from 1870 to 2013. Points labeled *M* indicate where the merging process occurred and points labeled *H* indicate where a hole was formed (see Sect. 2.5.2 for details). (b) Same as Fig. 2.1b except for the Niño 3.4 index for the period 1870-2013. *H* together with the arrow marks the location of a hole.

significance surrounded by regions of pointwise significance. These "holes" will turn out to be important structures in wavelet power spectra and are discussed thoroughly in Sect. 5.

2.3.2 Areawise Significance Testing

The idea behind the Maraun et al. (2007) areawise test (hereafter simply the "areawise test") is that correlations between adjacent wavelet coefficients arising from the reproducing kernel (see Appendix A) produce continuous regions of pointwise significance that resemble the reproducing kernel. The reproducing kernel for a given analyzing wavelet represents the time-scale uncertainty, which is related to the scale and time localization properties of the analyzing wavelet. Let (t, s) denote the location of a wavelet coefficient at scale s and time t. The correlation, $C(t, s, t', s')$, between any two wavelet coefficients located at (t, s) and (t', s') obtained from the wavelet transformation of a Gaussian white process is given by the reproducing kernel moved to t and stretched to s (Maraun et al., 2007), i.e.

$$C(t,s,t',s') = \sqrt{\frac{2s's}{(s')^2 + s^2}} \exp\left\{i\omega_0 \frac{s'+s}{(s')^2+s^2}(t'-t)\right\} \times \exp\left\{-\frac{1}{2}\frac{(t'-t)^2 + \omega_0^2(s'-s)^2}{(s')^2+s^2}\right\} \quad (2.5)$$

(Maraun and Kurths, 2004). Thus, for significance patches generated from random fluctuations, the typical patch area is the area of the reproducing kernel. The test can be described more formally as follows: Let P_{pw} be the set of all pointwise significance values and define a critical area $P_{crit}(t,s)$ as the subset of the time-scale domain for which the reproducing kernel K (corresponding to the analyzing wavelet), dilated and translated to time t and scale s, exceeds the threshold of a critical level K_{crit}. Mathematically, $P_{crit}(t,s)$ is given by

$$P_{crit}(t,s) = \{(t',s') : K(t,s;t',s') > K_{crit}\}. \quad (2.6)$$

It is noted that critical area of the areawise test is not area of significance patches but the area of the reproducing kernel at some critical level and at some scale. For a patch

of pointwise significant values, a point inside the patch is said to be areawise significant if the reproducing kernel dilated according to the scale in question entirely fits into the patch, i.e.

$$P_{aw} = \bigcup_{P_{crit}(t,s) \subset P_{pw}} P_{crit}(t,s), \tag{2.7}$$

where P_{aw} is the subset of pointwise significant values consisting of additionally areawise significant wavelet power coefficients. According to the areawise test, entire significance patches need not be areawise significant, just portions or subsets of them. That is, it is only those points that fit inside the kernel that are deemed areawise significant. The critical area is related to significance level of the areawise test by the following equation:

$$1 - \alpha_{aw} = 1 - \langle \frac{A_{aw}}{A_{pw}} \rangle, \tag{2.8}$$

where $1 - \alpha_{aw}$ is the significance level of the areawise test, A_{aw} is the area of the areawise significance patch, A_{pw} is the area of the pointwise significance patch, and $\langle \frac{A_{aw}}{A_{pw}} \rangle$ is the average ratio between the areas of area wise-significant patches and pointwise significance patches. It turns out that the calculation of α_{aw} is non-trivial, involving a root-finding algorithm that solves the equation $f(P_{crit}) - \alpha_{aw} = 0$ (see Sect. 2.4).

To illustrate the importance of the areawise significance test, the test was applied to the wavelet power spectra of the NAO and Niño 3.4 index time series (Figs. 2.3 and 2.4). Numerous 5% pointwise significance patches in the Niño 3.4 wavelet power spectrum were found to contain areawise-significant subsets, suggesting that these patches were less likely to be an artifact of multiple testing. For example, as indicated by the thick red contours, there are three areawise-significant regions located at a period of approximately 48 months, one at 1890, one at 1905, and a third one at 1985. Many more areawise-significant regions were identified at periods less than 8 months, especially before 1955. The wavelet power spectrum of the NAO index also contained pointwise significance patches with areawise-significant subsets, all at periods less than 8 months. However, it will be shown in Sect. 2.4 that they all may be artifacts of multiple testing, resulting from the large number of patches to which the areawise test was applied.

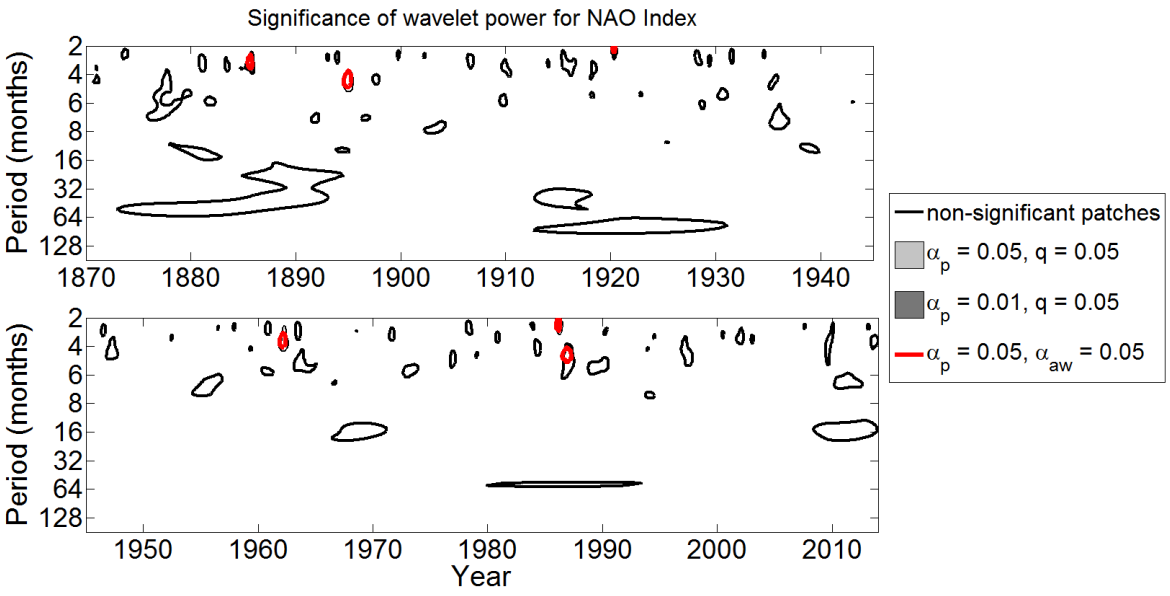

Figure 2.3. Significance of wavelet power for the NAO index mean monthly values for the period 1870-2013. Black contours enclose regions of 5% pointwise significance (see Sect. 2.3.1) and thick red contours are the 5% areawise-significant subsets (see Sect. 2.3.2). Light gray shading indicates those 5% pointwise significance patches that are geometrically significant at the q = 0.05 level and dark gray shading indicates

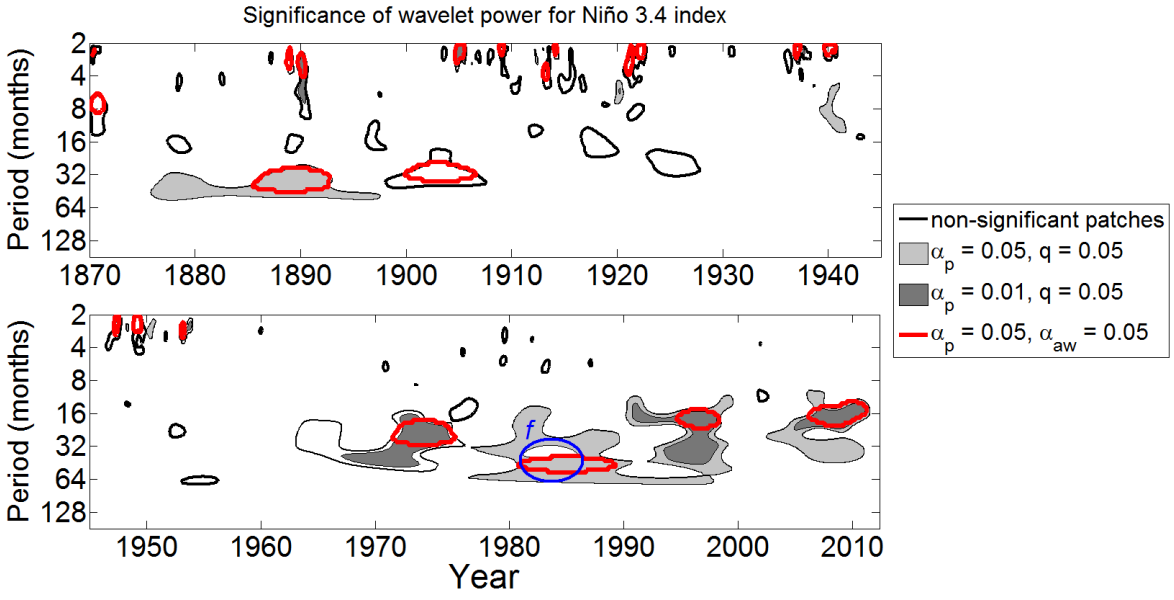

Figure 2.4. Same as Fig. 2.3 but for the Niño 3.4 for the period 1870-2013. The blue curve represents a closed path f that is not contractible to a point because it surrounds a hole (see Sect. 2.5.1 and Fig. 2.2).

2.4 Geometric Significance Testing

2.4.1 Development

A disadvantage of the areawise test is the complexity of the α_{aw} calculation, which involves a root-finding algorithm. It is therefore desirable to construct an alternative test whose significance level is easy to calculate, readily allowing the following: (1) the application of the test to patches at various pointwise significance levels; (2) the adjustments of the significance level of the test; (3) the application of the test to wavelet power spectra obtained using other analyzing wavelets; and (4) the implementation of *p*–value adjustment procedures to control the family-wise error rates and false discovery rates.

The development of a geometric significance test will require ideas from basic geometry and set theory. In wavelet analysis, the wavelet power is computed at a discrete set of time coordinates T with elements t_i for $i = 1, \dots, N$ and at a discrete set of scales S whose elements s_j (j =1,…,J) are given by

$$s_j = s_{min} 2^{j\delta j} \tag{2.9}$$

and

$$J = \delta j^{-1} log_2 \left(\frac{N\delta t}{s_{min}}\right), \tag{2.10}$$

with δt a time step and s_{min} the smallest resolvable scale (Torrence and Compo, 1998). Note that the maximum value of δj for which adequate sampling can be achieved depends on the wavelet function, being approximately equal to 0.5 for the Morlet wavelet. For the geometric test, a patch will be considered to be a polygon with vertices $v_k = (x_k, y_k)$ for k = 0,…, m-1, where x_k and y_k are, respectively, elements from T and S and m-1 is the number of vertices. It is worth noting that not all patches are closed in the sense that some are located near the edges of the wavelet domain. To remedy this problem, semi-enclosed patches are artificially closed by connecting the two vertices located on the boundary of the wavelet domain with a line segment.

Perhaps the most fundamental property of a pointwise significance patch is its area, which can be calculated using the following special case of Green's Theorem (Appendix C):

$$A = \frac{1}{2}\left|\sum_{k=0}^{m-1}(x_k\, y_{k+1} - \, x_{k+1}y_k)\right|, \tag{2.11}$$

where $y_0 = y_m$, $x_0 = x_m$ (Worboys and Duckham, 2004). For significance patches containing holes, the total area of the holes is subtracted from the area the significance patch would have if it did not contain the holes.

What will be of particular interest is the normalized area of a significance patch, not its absolute area. To compute the normalized area, the centroid of a significance patch will need to be calculated using the following formulas (Worboys and Duckham, 2004):

$$C_t = \frac{1}{6A}\sum_{k=0}^{m-1}(x_k + \, y_{k+1})\,(x_k y_{k+1} - \, x_{k+1}y_k) \tag{2.12}$$

and

$$C_s = \frac{1}{6A}\sum_{k=0}^{m-1}(y_k + \, x_{k+1})\,(x_k y_{k+1} - \, x_{k+1}y_k), \tag{2.13}$$

where C_t and C_s are the time and scale coordinates, respectively, of the centroid. Recall that the centroid is the area-weighted location of a polygon. If A_R is the area of the reproducing kernel dilated or contracted (at a certain critical level) to (C_t,C_s), then the normalized area of a significance patch is given by

$$A_n = \frac{A}{A_R}, \tag{2.14}$$

and allows one to compare sizes of significance patches across all scales simultaneously. Two idealized pointwise significance patches with equal normalized area are shown in Figs. 2.5a and 2.5b.

The idea of the geometric significance test is to generate a null distribution of A_n and use the null distribution to compute the significance of patches in the wavelet domain. In climate science, a suitable null hypothesis is red-noise so that A_n will be computed for a large ensemble of patches generated from red-noise processes. Using the null distribution of A_n, one can assign to each patch in the wavelet domain a probability p that the patch

was not generated from a random stochastic fluctuation. It is noted that the null distribution of A_n depends on the choice of null hypothesis (not shown), with, for red-noise processes, A_n increasing with increasing lag-1 autocorrelation coefficients.

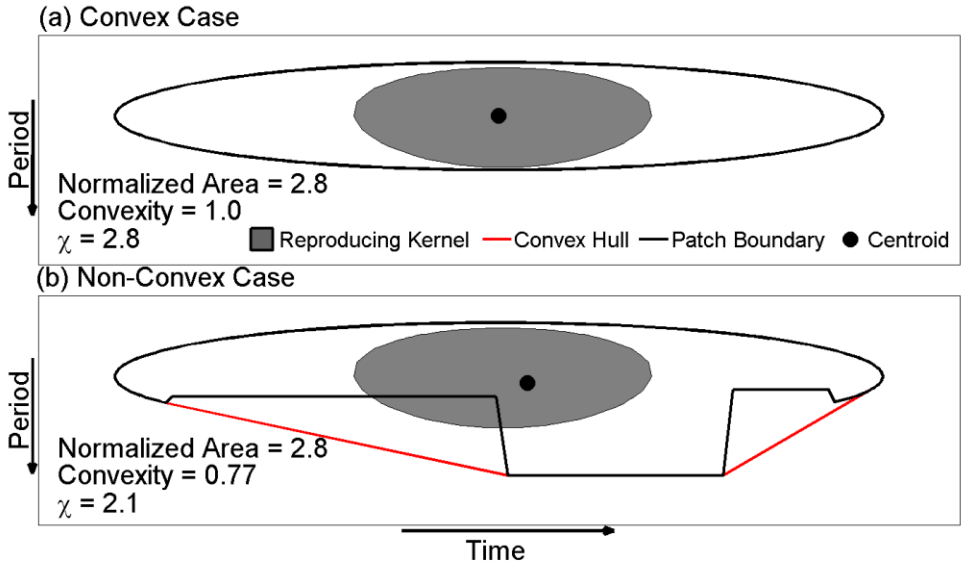

Figure 2.5. (a) An idealized convex pointwise significance patch whose boundary is indicated by the black contour and whose centroid is indicated by the black dot. For reference, the reproducing kernel associated with the areawise test is shown, which is indicated by gray shading. In this case, the reproducing kernel lies entirely inside the patch. The convexity, normalized area, and χ are displayed on the bottom left corner. (b) Same as (a) except the area of the convex hull (red curve) is not equal to the area of the patch and the reproducing kernel is unable to fit entirely inside the patch.

The calculation of the geometric significance level $1 - \alpha_g$, unlike the calculation of $1 - \alpha_{aw}$, is straightforward: for the areawise test one needs to compute α_{aw} as a function of P_{crit}, whereas for the geometric test α_g is no longer a function P_{crit}. Moreover, the estimation of P_{crit} involves a root-finding algorithm that solves the equation $f(P_{crit}) - \alpha_{aw} = 0$, where $f(P_{crit})$ is estimated using Monte Carlo simulations. Thus, the application of the areawise test to pointwise significance patches for *M* different values of α_{aw} would require *M* Monte Carlo ensembles, making *p*-values for the test difficult to obtain. For the geometric test, only a single Monte Carlo ensemble is needed, as a single choice of P_{crit} is needed to generate a null distribution, from which any desired value of

α_g can be obtained. In fact, while the choice of P_{crit} impacts the mean value of the null distribution, the geometric significance of a significance patch is left unchanged, as the significance is relative to a distribution of χ under some noise model (Appendix B).

The elimination of the P_{crit} dependence from the calculation of the geometric significance level allows the geometric test to be readily performed on patches of various pointwise significance levels. For the areawise test, a new P_{crit} must be estimated for each pointwise significance level since A_{pw}, on average, will change depending on if the pointwise significance level $1 - \alpha_p$ is increased (patches shrink) or is decreased (patches grow). For the geometric test, there is no need to find a new P_{crit} —simply compute a new null distribution based solely on the information of the pointwise significance patches at some pointwise significance level $1 - \alpha_p$.

Another advantage of eliminating the P_{crit} dependence is that the geometric test can be readily applied to wavelet coherence, partial wavelet coherence (Ng and Chan, 2012), multiple wavelet coherence, and cross-wavelet spectra. The application of the geometric test to significance patches in the aforementioned wavelet spectra only requires a single Monte Carlo ensemble to generate a null distribution, eliminating the calculation of a new P_{crit} for each wavelet spectra and for each value of α_g. For the areawise test, a new P_{crit} must be estimated for each value of α_{aw} and for each wavelet spectra, making the areawise test difficult to implement in practical applications.

It may happen that a pointwise significance patch is so large that individual oscillations embedded in the patch cannot be detected by the geometric test. However, there are two solutions to this localization problem: the first solution is to increase the significance level of the pointwise test, allowing large patches to separate, and then perform the geometric test on the smaller patches. The second solution is to examine other properties of significance patches that may indicate the presence of multiple periodicities that form large significance patches from the merging of several smaller patches. The second solution will be addressed thoroughly in Sect. 5.

Another situation that may arise in practice is the application of the geometric test to patches located both inside and outside the cone of influence (COI). In the case of the

pointwise significance test, the edge effects only influence those wavelet power coefficients that lie inside the COI; however, for the geometric test, the significance of the entire patch will be impacted even if the patch only partially lies inside the COI. The reason is that the COI will act to decrease the size of significance patches through the reduction of wavelet power in the COI and subsequently the total area of the patch. One should thus be cautious when interpreting the results of the geometric test for patches near the COI.

2.4.2 Multiple Testing

If the geometric test was performed on *K* significance patches at the α_{geo} level, then, on average, one can expect $\alpha_{geo}K$ false positive results, which would make the geometric test permissive for large *K*. It is therefore necessary to reduce the number of false positive results. There are various ways to reduce the number of false positives, including the Walker test, Bonferroni correction, and other counting procedures (Wilks, 2006). Recently, methods for controlling the false discovery rate (FDR) have been developed, where the FDR is the expected proportion of rejected local null hypotheses that are actually true (Benjamini and Hochberg, 1995). In particular, Benjamini and Hochberg (1995) developed a method for controlling the FDR based on the number of local hypotheses being tested and the degree to which the local hypotheses were rejected, contrasting with other procedures that ignore the confidence with which the local tests reject the local hypotheses (Wilks, 2006). Moreover, the method has proven to have high statistical power, especially when only a small fraction of the *K* local tests correspond to false null hypotheses (Wilks, 2006). The procedure will therefore be used to control the false discovery rate of the geometric test, which will facilitate the interpretation of results.

Suppose that *K* local hypotheses were tested, where, in the present case, the local hypotheses refer to the testing of each patch individually under the assumption that the results of the individual tests are independent. A global geometric test can be performed at the α_{global} level as follows: Let $p_{(l)}$ denote the *l*th smallest of *K* local *p*-values; then, under the assumption that the *K* local tests are independent, the FDR can be controlled at the *q*-level by rejecting those local tests for which $p_{(l)}$ is no greater than

$$p_{FDR=} \max_{r=1,\ldots,K} \left[p_{(r)} : p_{(r)} \leq q(r/K)\right] \tag{2.15}$$

$$\max_{r=1,\ldots,K}\left[p_{(r)}: p_{(r)} \leq \alpha_{global}(r/K)\right] \tag{2.16}$$

so that the FDR level is equivalent to the global test level. According to the procedure, any local test resulting in a *p*-value less than or equal to the largest *p*-value for which Eq. (2.16) is satisfied is deemed significant. If no such local *p*-values exist, then none are deemed significant and, therefore, the global test hypothesis cannot be rejected. The global geometric test will thus only deem those significant patches with *p*-values satisfying Eq. (16) as significant. Throughout the paper $q = \alpha_{global}$ will be set to 0.05.

2.4.3 Comparisons with the Areawise Test

With a formal geometric significance test now developed, it is useful to compare the areawise and geometric significance tests, where comparisons will be made using an empirically derived quantity. Let N_{sig} be the number of pointwise significance patches in a given wavelet power spectrum, N_a the number of patches containing an areawise-significant region, N_g the number of geometrically significance patches, and N_{ag} the number patches that are both geometrically significant and that contain areawise-significant regions. The quantity

$$I_{sim} = \frac{N_{sig} - N_a - N_g + 2N_{ag}}{N_{sig}} \tag{2.17}$$

then measures the similarity between the two tests. The interpretation of I_{sim} is as follows: if I_{sim} = 1 then all patches containing areawise-significant regions are also geometrically significant and all patches which do not contain areawise-significant regions are also not geometrically significant. On the other hand, for values of I_{sim} less than 1 some patches containing areawise-significant regions may not be geometrically significant, with the converse also being true.

To better compare the similarity between the two tests, distributions of I_{sim} were constructed by generating 1000 synthetic wavelet power spectra of red-noise processes with fixed autocorrelation coefficients and length N = 1000 (arbitrary units) and computing I_{sim} for each of the synthetic wavelet power spectra. The experiment was performed for red-noise processes with different lag-1 autocorrelation coefficients to determine if I_{sim} depends on the AR1 model. The results are shown Fig. 2.6a. With a mean value of 0.90, a

strong agreement was found between the areawise and geometric tests, differences arising from the fact that the areawise test is a local test, finding significant regions within patches, whereas the geometric test assigns a significance value to entire patches (see discussion below). Since I_{sim} was often less than 1.0, some patches containing areawise-significant regions were not found to be geometrically significant, and, conversely, some patches were geometrically significant without containing areawise-significant regions.

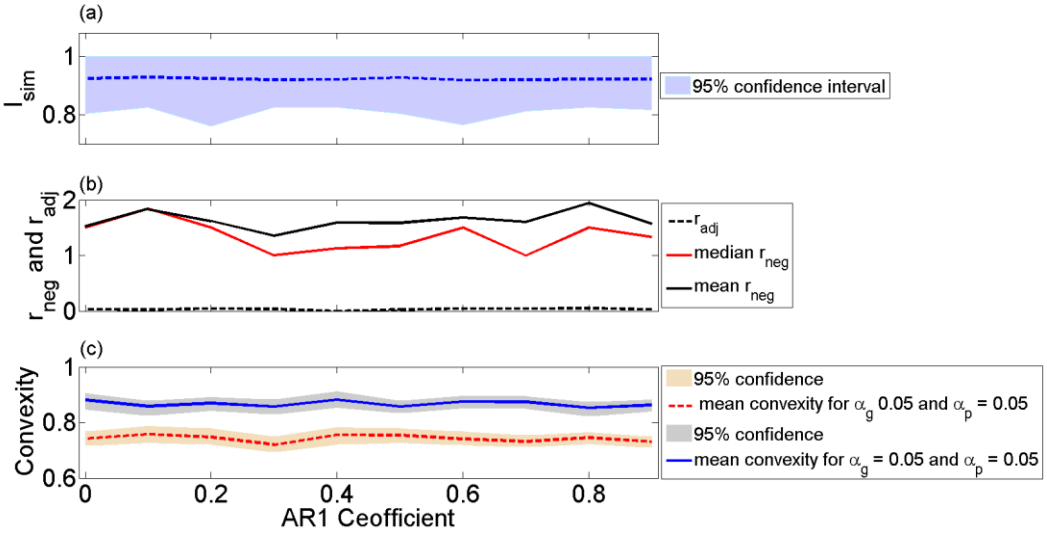

Figure 2.6. (a) Similarity index between the geometric and areawise tests for different lag-1 autocorrelation coefficients for red-noise processes (see text). (b) Same as (a) except for the ratio between the false positive results of the geometric and areawise tests. The dotted black line represents the ratio of false positive between the two tests when the false discovery rate of the geometric test is controlled at the 0.05 level. (c) Same as (a) but for the mean convexity of 5% pointwise significance patches that are geometrically significant at the 5% level and for the mean convexity of 5% pointwise significance patches that are areawise significant at the 5% level.

The quantity $r_{neg} = N_g/N_a$, which measures the ratio of false positive results between both tests, was also computed for case when both the geometric and areawise test levels were set to 0.05 (Fig. 2.6b). In this case, the mean value of r_{neg} was found to range from 1.0 to 2 and the median value was found to be generally greater than 1.0, ranging from 1 to 1.8. No dependence on the lag-1 autocorrelation coefficients was identified. The results indicate that the geometric test is generally less conservative than the areawise test for a given wavelet power spectrum. The lack of conservativeness, however, can be remedied by controlling the FDR of the geometric test at the q = 0.05 level. Fig. 6b shows

r_{adj}, the ratio of false positive results between the areawise tests and the geometric test but with FDR controlled for the geometric test. As indicated in Fig. 2.6b, by controlling the FDR the geometric test is much more conservative than the areawise test, resulting in fewer false positive results, with a typical value of r_{adj} ranging from 0.02 to 0.05.

To explain the differences between the areawise and geometric tests, it will be necessary to consider the convexity of a patch, the degree to which a polygon or point set lacks concavities. The reason for considering convexity is illustrated by considering the two significance patches shown Fig. 2.5, which have equal values of A_n but different geometries: one is convex (i.e., has no concavities, Fig. 5a) and the other is not convex (Fig. 2.5b). Suppose that the areawise test was performed on the two patches at the α_{aw} level. For the convex patch shown Fig. 2.5a, the reproducing kernel is capable of fitting entirely inside the patch but is unable to fit inside the non-convex patch as a result of the concavity. Thus, although having equal area, the two patches differ in their areawise significance, where the difference in significance is related to their geometry. Thus, $p_{aw} = g(\mathcal{C}, A; H_0)$ for some function g, where p_{aw} is the areawise test p-value associated with a patch calculated under the null hypothesis H_0 and $\mathcal{C}$ is the convexity of the patch, which is now formally defined.

Rigorously, convexity is defined as follows: Let x and y be any two points in a set Z; then the set Z is convex if for all t the line segment

$$[x, y] = \{tx + (1 - t)y: 0 \leq t \leq 1\} \tag{2.18}$$

is in Z (Ziegler, 1995). Equivalently, a set is convex if it contains any line segment joining any pair of points in Z. Under this definition, for example, patches with thin bridges as described by Maraun et al. (2007) are not convex.

To quantify convexity, another idea from set theory, the convex hull, will be needed, which for a point set Z is defined as the intersection of all convex sets containing Z (Ziegler, 1995). In other words, it is the smallest convex set containing Z constructed from the intersection of all convex sets containing Z. Mathematically, the convex hull of a point set Z is expressed as

$$\text{conv}(Z) = \bigcap\{Z' \subseteq \mathbb{R}^2: Z \subseteq Z', Z' \text{ convex}\}. \tag{2.19}$$

In practical applications, the convex hull of a set can be easily computed using existing algorithms (Barber et al., 1996). It is noted that all holes are ignored in the computation of the convex hull because the computation of the convex hull assumes that there are no holes in the polygon. A patch containing a hole can never have a smallest convex set containing the set because holes allow line segments to leave the patch regardless of the size of the convex hull.

A metric for convexity will now be defined using the area of a significance patch together with the area of its convex hull as follows: If A_k is the area of the convex hull of a significance patch whose area is *A*, then the convexity is

$$\mathcal{C} = \frac{\mathrm{A}}{A_k}, \tag{2.20}$$

where $0 \leq \mathcal{C} \leq 1$. High values of $\mathcal{C}$ correspond to significance patches with relatively small concavities, whereas small values of $\mathcal{C}$ correspond to patches with relatively large concavities, as in the case of significance patches with thin bridges.

According to the areawise test, patches with smaller values of $\mathcal{C}$ are less likely to be areawise significant so that it is expected that patches deemed significant by the areawise test will be primarily convex. To test this hypothesis, 10,000 patches arising from red-noise processes with different lag-1 autocorrelation coefficients were generated and the convexity of those patches deemed areawise significant at the $\alpha_{aw} = 0.05$ level was calculated. The results in Fig. 2.6c show the mean convexity as a function of the lag-1 autocorrelation coefficients, together with the 95% confidence bound. The mean convexity of the patches was found to be approximately 0.8, regardless of the lag-1 autocorrelation coefficient. An identical experiment was also performed for geometrically significant patches but with the convexity of patches that are geometrically significant at the $\alpha_{geo} = 0.05$ being computed. In contrast to areawise-significant patches, patches that were found to be geometrically significant, on average, had lower convexity, the reason for which is that the calculation of α_{geo} makes no assumption about convexity. The *p*-value for the geometric test is thus $p_{geo} = f(A; H_0)$ for some function *f*, contrasting with p_{aw} that depends on convexity. The results of the experiments are consistent with Figs. 2.5a and

2.5b, where both the ideal patches have the same geometric significance but the ideal patch in Fig. 2.5b has a larger p_{aw} so that $p_{aw} > p_{geo}$.

Convexity cannot fully explain the differences between p_{aw} and p_{geo} for a given patch. More generally, $p_{aw} = g(\mathcal{C}, A, S_1, \ldots, S_R; H_0)$, where S_1 to S_R are shape parameters of the patch, such as aspect ratio and symmetry. Consider, for example, a convex patch whose length in the time direction is long with respect to the reproducing kernel (at some critical level) but thin in the scale direction with respect to the reproducing kernel. Such a patch would be deemed insignificant by the areawise test, though it may have an area much larger than the critical area of the areawise test. Asymmetry with respect to the scale axis, as another example, may also result in a patch being deemed insignificant by the areawise test if, for example, the width of the patch in the scale direction decreases with time. If the normalized areas of such patches are larger than the critical level of the geometric test, the patches will be geometrically significant, though may not be areawise significant if the reproducing kernel is unable to fit inside the narrow portion of the patch. The above arguments suggest that $f(A; H_0) \neq g(\mathcal{C}, A, S_1, \ldots, S_R; H_0)$ and thus the significance of patches as determined by the geometric and areawise tests need not be equal.

2.4.4 Geometric Significance Testing of Climatic Data

For climatic time series, significance is often tested against a red-noise background and therefore it is reasonable to expect that the areawise and geometric tests behave similarly when applied to climatic time series. As such, the areawise and geometric tests were applied to the NAO and Niño 3.4 time series. For the wavelet power spectrum of the NAO index time series (see Fig. 2.3), not a single patch was found to be geometrically significant after controlling the FDR at the 0.05 level, suggesting the NAO index time series is composed of stochastic fluctuations. In fact, the NAO has already been shown to be consistent with a first-order Markov process (Feldstein, 2002). Recent work by Hanna et al. (2014) claimed that the NAO variability has increased over the past 30 years; however, the results from this analysis suggest that such changes cannot be distinguished from stochastic fluctuations, which could render difficult projections of future changes of the NAO.

The wavelet power spectrum of the Niño 3.4 index (see Fig. 2.4) was found to contain numerous geometrically significant patches in the period band 16-64 months, especially after 1960. The 5% pointwise significance patch extending from 1980 to 2000, as an example, was found to be significant, as well as the patch centered at 2008. The significance patch centered at 1985 and at a period of 32 months, however, is so large that individual oscillations could not be identified. To remedy the problem, the geometric significance was applied to 1% (α_p = 0.01) pointwise significance patches with q = 0.05, resulting in 1% pointwise significance patches at 1970, 1995, and 2007 being deemed significant, all of which also contained areawise-significant regions. Patches located at a period less than 8 months were also found to be geometrically significant, though only before 1955.

2.5. Topological Significance Testing

2.5.1 Topological Significance Testing of Ideal Time Series

Topology is a branch of mathematics concerned with properties of spaces that remain unchanged after continuous deformations. So far only geometric aspects of significance patches have been discussed. Area of a significance patch, as an example, is a geometric property in the sense that stretching the patch in both the scale and time direction would increase its area. There are properties, however, that would be unaffected by stretching the significance patch. As a motivating example, consider the significance patches shown in Fig. 2. 4 corresponding to the wavelet power spectrum of the Niño 3.4 index (see Fig. 2.2), where there is a hole or void of pointwise significance located within a significance patch at 1985. This feature is topological, as the hole would remain under a continuous deformation such as stretching. A more formal definition of a hole will require some notions from topology. Let I = [0,1] be the closed unit interval. Then a path from a point a to a point b in a significance patch P is a continuous function $f: I \to P$ with $f(0) = a$ and $f(1) = b$, where in the case that $f(0) = f(1) = c$ the path is said to be closed (Hatcher, 2002). Note that a point is a special kind of closed path called the constant path. A patch will be said to contain a hole if there exists a path in the significance patch such that it cannot be continuously deformed into a point, where the feature obstructing the path from such a deformation is a hole. The definition is consistent with notions of simply-connectedness in

topology (Hatcher, 2002). Figure 2.4 shows an example of a closed path (blue curve) in a patch that cannot be contracted to a point because it surrounds a hole located in the patch.

For a patch with a hole, there will be two boundaries, an external boundary and an internal boundary representing the boundary between the hole and the patch. Thus, if a patch contains an internal boundary or contour it will contain a hole, whereas a patch without a hole will contain no internal contours. In practical applications, the existence of a hole can be determined by orienting external contours in the clockwise direction and internal contours in the counter-clockwise direction, a procedure automatically implemented by the Matlab contour routine. The number of counter-clockwise oriented contours is thus the number of holes in the wavelet power spectrum at a given pointwise significance level.

To begin the topological analysis, the topology of time series with known structures will be analyzed. Given the importance of red-noise processes in the spectral analysis of climatic time series, the topology of patches generated from red-noise processes is first considered to determine if pointwise significance patches can be distinguished from those generated from red-noise processes solely based on their topology. To answer this question, 10,000 wavelet power spectra of red-noise processes were generated and the number of holes (denoted by N_h hereafter) at a finite set of pointwise significance levels was computed for each wavelet power spectra (Fig. 2.7). It was found that N_h is not a random function of the pointwise significance level, as indicated by the 95% confidence bounds. Most importantly, for pointwise significance levels less than 10%, few patches contained holes, suggesting that holes are an uncommon feature of significance patches generated from red-noise processes (Table 2.1) and therefore can be used to distinguish spurious patches from important structures. It also noted that neither the shape nor the amplitude of the curve in Fig. 2.7 depends on the lag-1 autocorrelation coefficient of the red-noise process. Table 2.1 also suggests that patches containing more than a single hole are unlikely to be the result of red-noise, even for a modest pointwise significance level of 20%. For pointwise significance levels of 1% and 5%, no more than a single hole was identified in a

given patch.

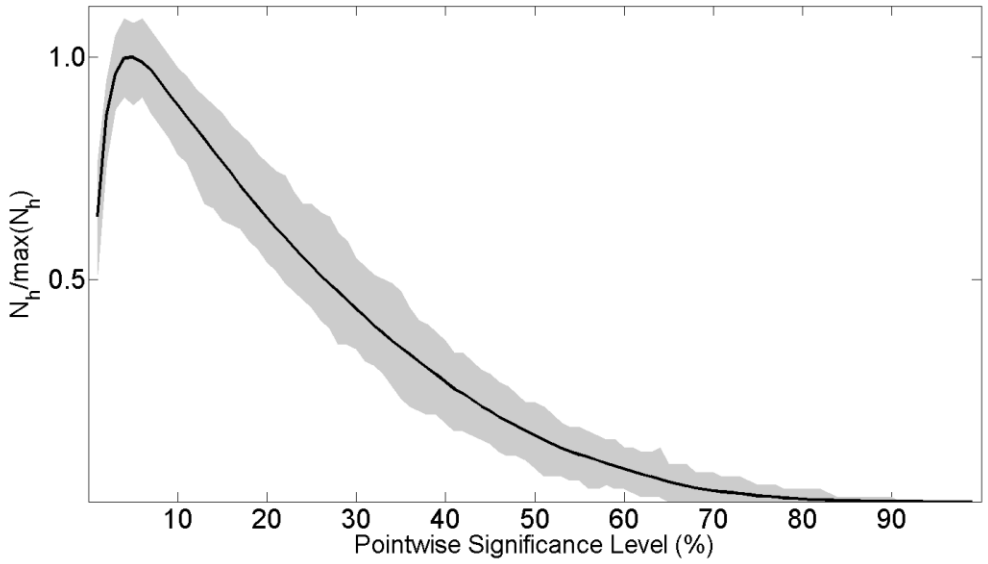

Figure 2.7. Normalized mean number of holes as a function of pointwise significance level. The number of holes was calculated by generating 10,000 synthetic wavelet power spectra of red-noise processes with fixed autocorrelation coefficients of 0.5 and computing the number of holes Gray shading represents the 95% confidence interval.

A simple algorithm for assessing the significance of holes is therefore developed. To find the significance of holes, plot the centroids of holes at a finite set of pointwise significance levels and project the centroids onto the wavelet domain, resulting in a topological wavelet diagram. The number of holes contained in a patch should also be computed, as patches with more holes are less likely to result from red-noise. In accordance with Fig. 2.7 and Table 2.1, regions in the wavelet domain where holes exist below the 20% pointwise significance level will be considered regions with significant topological features.

Table 2.1. Fraction of pointwise significance patches containing at least N_h holes as a function of the pointwise significance level calculated from an ensemble of 200,000 significance patches generated from red-noise processes with fixed autocorrelation coefficients equal to 0.5.

Significance level (%)	$N_h \geq 1$	$N_h \geq 2$	$N_h \geq 3$	$N_h \geq 4$
20	2.3×10^{-2}	2.6×10^{-3}	4.0×10^{-3}	0
15	1.0×10^{-2}	5.0×10^{-3}	1.0×10^{-3}	0

10	2.0×10^{-3}	1.0×10^{-3}	0	0
5	3.4×10^{-4}	0	0	0
1	0	0	0	0

With a method for assessing the significance of holes, it is reasonable to analyze different ideal time series, both linear and nonlinear, to determine what types of time series produce holes in significance patches. Perhaps the simplest case is a single sinusoid with additive white noise (not shown), where the time series power spectrum in tested against a white-noise background spectrum. In this case, no evidence was found that a single sine wave, regardless of amplitude and signal-to-noise ratio, is capable of generating holes in 5% pointwise significance patches. A similar experiment was repeated but the power spectra of the sine waves were tested against red-noise spectra. The results also indicated that a single sine wave is incapable of producing holes in 5% pointwise significance patches, implying holes arise from a richer structure embedded in time series. Thus, two more complex cases are considered.

To derive the Case 1 time series, first consider the nonlinear system

$$X_{out}(t) = \ bX_{in}(t) + \ \gamma X_{in}^2(t), \tag{2.21}$$

where $X_{in}(t)$ is the input into the system, $X_{out}(t)$ is the output of the system, b is a linear coefficient, and γ is a nonlinear coefficient. The output from this system will be quadratically phased coupled (King, 1996), where quadratic phase coupling indicates that for frequencies f_1, f_2, and f_3 and corresponding phases ϕ_1, ϕ_2, and ϕ_3 the sum rules $f_1 + f_2 = f_3$ and $\phi_1 + \phi_2 = \phi_3$ are satisfied. In Case 1, $X_{in} = cos\, 2\pi ft$ so that

$$X_{out}(t) = \frac{\gamma}{2} + b \cos 2\pi ft - \frac{\gamma}{2} \cos 4\pi ft, \tag{2.22}$$

indicating that the output contains an additional frequency component at the harmonic $2f$ (harmonic generation) and the mean value of the output has shifted (rectification) with respect to the input. Figures 2.8a and 2.8b show the time series of X_{out} and the significance of the wavelet power for the case when $f = 1/64 = 1/\lambda_1$, b = 1, $\phi_1 = \pi/2$, $\phi_2 = \pi/3$, and γ = 0.25 (arbitrary units) and with Gaussian white noise added to the output. In this

case, the significance of the wavelet power was tested against a red-noise background spectrum. Figure 2.8 shows numerous pointwise significance patches, all of which are spurious except for the one at $\lambda_1 = 64$. The areawise and geometric test correctly identified the pointwise significance patch at $\lambda_1 = 64$ to be significant but deemed a spurious patch as significant at time 140 and at $\lambda = 3$. It is noted that the geometric test only deemed the 1% pointwise significance patch at $\lambda_1 = 64$ as significant. Also note that the pointwise significance test was unable to detect the harmonic with period $\lambda_2 = 32$ using a red-noise background spectrum.

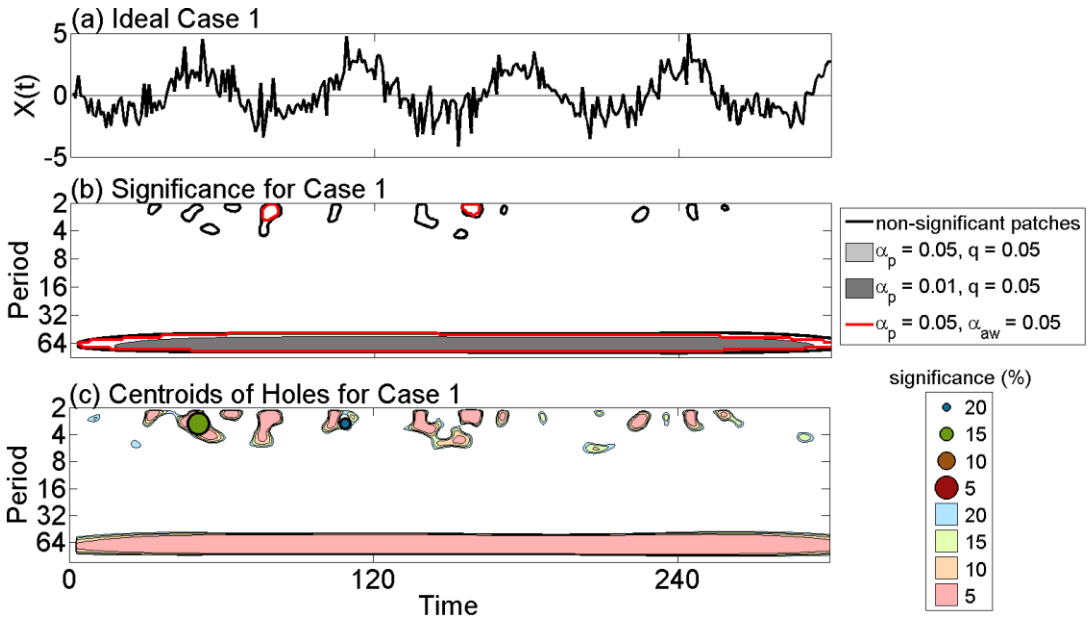

Figure 2.8. (a) Time series of Case 1, which results from passing a single sinusoidal input with period $\lambda = 64$ through Eq. (2.16). Gaussian additive white noise with a signal-to-noise of 2 was added to the output response. (b) The significance of wavelet power for Case 1 (see Fig. 2.3 for details). (c) Topological wavelet diagram corresponding to (b). Points are the centroids of the holes at a given pointwise significance level, where both the color and size of the dots indicate the pointwise significance level at which the hole existed. The shading of the patches corresponds to the pointwise significance level at which the wavelet power coefficient existed, with the color of the shading lighter than the dots for clarity.

It should be noted, however, that if the parameter γ were increased to a value greater than 1, the oscillation with period $\lambda_2 = 32$ would become more prominent. In fact, it was found that for $\gamma \geq 1$ the areawise and geometric tests perform better (not shown), correctly identifying the oscillation with period $\lambda_2 = 32$, with the result also depending on the noise

level of the white noise. Case 1 thus only serves as an illustrative example of a situation that may arise when a wavelet analysis is applied to a geophysical (often noisy) time series.

To extract more information from the wavelet power spectrum, the centroids of holes were plotted as a function of the pointwise significance level (Fig. 2.8c). Figure 2.8c shows that holes only existed at pointwise significance levels of at most 15% and 20% and therefore not all nonlinear time series can generate holes at the 5% pointwise significance level, suggesting that the relative difference between the primary frequency components or the resulting frequency combinations is important, as discussed below. The amplitudes of the coefficients b and γ, as well as the signal-to-noise ratio of the Gaussian white noise, turn out to be also important, which is discussed below.

Case 2 is the quadratically phase-coupled time series

$$X(t) = a\cos(2\pi f_1 t + \phi_1) + b\cos(2\pi f_2 t + \phi_2) + \gamma\cos[2\pi(f_1 + f_2)t + \phi_1 + \phi_2], \quad (2.23)$$

which consists of three frequency components: $f_1 = 1/20 = 1/\lambda_1$, $f_2 = 1/30 = 1/\lambda_2$, and $f_1 + f_2 = 1/12 = 1/\lambda_3$, and γ is assumed to be 0.5. It is noted that Case 1 is a special case of Case 2. Like Case 1, wavelet power was also tested against a red-noise background. Unlike the significance patches in Fig. 2.8c corresponding to Case 1, holes have appeared in 5% pointwise significance patches between periods $\lambda_1 = 20$ and $\lambda_2 = 30$ (Fig. 2.9b). Moreover, the 5% pointwise significance patch containing the hole (labeled P_1) was found to be geometrically significant but was not found to contain an areawise-significant subset. It is also worth noting that the areawise and geometric tests failed to detect a significant periodicity at $\lambda_1 = 20$ despite the fact that it is known to exist by construction. Figure 2.9c shows that a few holes existed at low pointwise significant levels ($\leq$ 20%), though only one was found at the 5% pointwise significance level (light red shading). However, if one applies the pointwise significance test to the wavelet power at the 20% significance level a feature emerges that can hardly be produced from red-noise (see Table 2.1), namely a large 20% significance patch (light blue shading) containing four holes located in the period band 20-30. One can thus have confidence that the feature is significant. Furthermore, by constructing a patch topologically unlike those generated from red-noise,

significant wavelet power extending from time 20 to 300, undetected by the pointwise, areawise, and geometric tests, has been recovered, whereas only applying the 5% pointwise test would result in two patches that are seemingly indistinguishable from red-noise (labeled P_2 and P_3), with only one at $\lambda_2 = 30$ being geometrically significant.

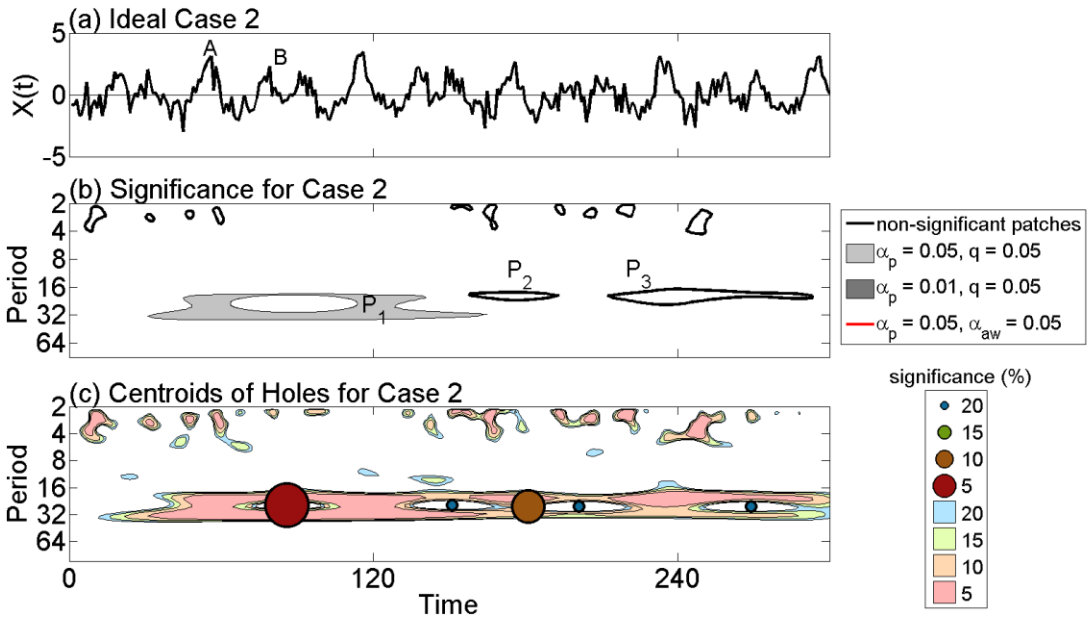

Figure 2.9. (a) Time series of Case 2. Gaussian additive white noise with a signal-to-noise ratio of 8 was added to the time series. At the point labeled A, two oscillations resonate, merging two pointwise significance patches in the wavelet domain. At the point labeled B no such resonance occurs and the two significance patches separate. (b) The significance of wavelet power (see Fig. 2.3 for details). The pointwise significance patch labeled P_1 contains a hole and the pointwise significance patches labeled P_2 and P_3 were falsely deemed insignificant by the geometric and areawise tests. (c) Same as Fig. 2.8c except for Case 2.

The ability of the pointwise, areawise, and geometric tests to detect significant structures inevitably depends on the parameters a, b, γ, f_1, and f_2. In fact, Maruan et al. (2007) has already determined that the pointwise test and areawise test are sensitive to the signal-to-noise level. It was hypothesized that the results of the topological method also depend on the parameters a, b, γ, f_1, and f_2. To test the hypothesis, several experiments were performed, the first of which investigated the relationship between f_1, f_2, and the number of holes. The experiment is described below.

Though both ideal time series contain a quadratic nonlinearity, the nonlinear interaction in Case 2 contained oscillations with nearby frequency components, allowing

the formation of holes, whereas for Case 1 no significant holes appeared in significance patches. It appears that the presence of holes depends on the relative location of two oscillations in the frequency domain, and thus it is reasonable to suspect that there exists a critical frequency difference Δf_{crit}, measuring the maximum frequency difference for which holes will appear in a wavelet power spectrum. An empirically derived Δf_{crit} was determined by generating a large ensemble of time series of the form

$$x(t) = \cos 2\pi f_1 t + \cos 2\pi f_2 t + w(t), \tag{2.24}$$

where $f_2 > f_1 > 0$ were generated at random, $w(t)$ is additive white noise, and all the time series were of a fixed length. The signal-to-noise ratio was fixed to 20 and each wavelet power spectrum was tested against a red-noise background spectrum. Figure 2.10 shows the mean value of N_h as a function of $\Delta r = (f_2 - f_1)/f_2$, the relative fractional change. For $\Delta r = 0.5$, holes never appeared, whereas for $\Delta r = 0.3$ holes appeared frequently. There is therefore a preferred frequency combination for which holes are more likely to appear. It was estimated that the upper critical value of Δr is $\Delta r_{crit} = 0.45$. Using the definition of Δr, one can write $\Delta f_{crit} = 0.45 f_2$ and therefore the critical frequency difference is a function of f_2.

It turns out that even if the above experiment (not shown) was repeated using white-noise rather than red-noise background spectra Δr_{crit} would still be equal to 0.45, though more holes were found to appear at signal-to-noise ratios less than 2. It was expected, however, that Δr_{crit} also depends on the amplitudes of the cosines in Eq. 2.24. Thus, a third experiment was conducted in which the amplitudes of the cosines were allowed to vary from 1 to 50 and f_1 and f_2 were allowed to vary from 0 to 0.5. The experiment was repeated for signal-to-noise ratios from 1 to 20. The results from the experiments (not shown) indicate that for red-noise background spectra and for a signal-to-noise ratio of 20 that $\Delta r_{crit} = 0.53$, contrasting with the case for white-noise background spectra where Δr_{crit} was found to be 0.51.

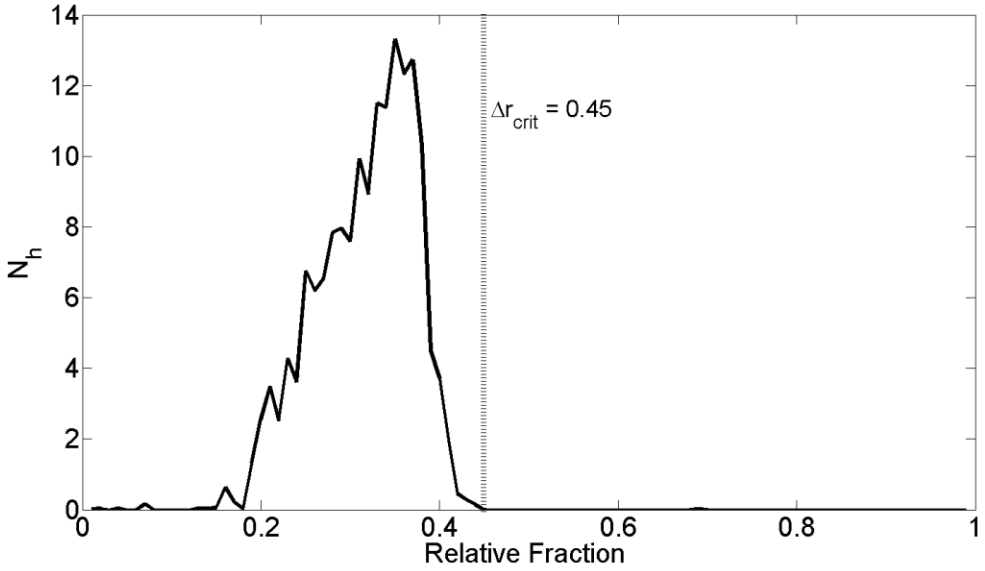

Figure 2.10. Mean number of holes found in 5% pointwise significance patches as a function of $\Delta r = (f_2 - f_1)/f_2$ for a sum of two sinusoids with amplitudes equal to unity and frequency components f_1 and f_2 such that $f_2 > f_1 > 0$. Additive white noise with a signal-to-noise ratio of 30 was added to the sum of sinusoids. Pointwise significance was tested against a red-noise background. Dashed line represents the critical value of Δr, the value beyond which holes will rarely occur between oscillations of equal amplitude (set to unity) with frequencies f_1 and f_2.

The empirical results shown in Fig. 2.10 have theoretical implications. Suppose that a time series contained two oscillations of equal amplitude such that frequency components of the two oscillations were such that $f_2 = 2f_1$. Furthermore, suppose that the wavelet power of the oscillations were computed and the significance was tested against a red-noise or white-noise background spectrum. In this case, $\Delta r = 0.45$ and therefore holes will almost never appear in 5% pointwise significance patches, making the detection of quadratic phase coupling using topological methods more difficult in the case of self-interactions. More generally, suppose that a single sinusoid $X_{in}(t) = cos2\pi ft$ is passed through the nonlinear system

$$X_{out}(t) = \ bX_{in}(t) + \gamma X_{in}^{2n}(t), \tag{25}$$

where, after using the power-reduction for a cosine (Beyer, 1987), the output is given by

$$X_{out}(t) = b\cos 2\pi t \ + \frac{\gamma}{2^{2n}}\binom{2n}{n} + \frac{\gamma}{2^{2n-1}}\sum_{k=0}^{n-1}\binom{2n}{k}\cos 4\pi f(n-k)t\ , \tag{26}$$

where n is a positive integer and $\binom{n}{q}$ is a binomial coefficient. For the cosines in the summation, the frequency difference between any two cosines is

$$\Delta f = 4\pi f(n-p) - 4\pi f(n-m) = 4\pi f(m-p), \tag{27}$$

where $0 \leq p < m \leq n-1$. Thus,

$$\Delta r = (f_2 - f_1)/f_2 = \frac{4\pi f(m-p)}{4\pi f(n-p)} = \frac{m-p}{n-p}. \tag{28}$$

Using the fact that holes can only appear between oscillation pairs with $\Delta r \leq 0.53$ for a red-noise background spectrum, one can show that for large n more holes are able to appear in wavelet power spectra, with the likelihood of holes appearing depending on b and γ, with larger values of b and γ producing more holes. In this case, holes can form in the wavelet spectrum since, for example, if $m = 6$ and $p = 5$ with $n = 10$ the condition $\Delta r \leq 0.53$ will be satisfied. The result also holds if the order of the nonlinear interaction was odd and if the cosine function $X_{in}(t)$ was replaced by a sine function. For an odd order nonlinear interaction, however, $\Delta r = (2m-2p)/(2n+1-2p)$, where $0 \leq p < m \leq n$.

2.5.2 Topological Significance Testing of Climatic Time Series

With a better understanding of the origins of holes contained in significance patches, the wavelet power spectra shown in Figs. 2.1 and 2.2 are now analyzed more closely. Shown in Fig. 2.11a is the topological wavelet diagram corresponding to the wavelet power spectrum of the Niño 3.4 index, which shows the existence of numerous holes at low ($\leq 20\%$) pointwise significance levels, indicating that these patches are significant features (see Table 2.1). For example, the rather large patch extending from 1960 to 2013 in the period band 16 to 64 months contains a hole located at 1985 and at a period of 32 months that existed at the 5% pointwise significance level. In the same patch, three more holes existed at the 10% pointwise significance level, one located at 1975 and at a period of 48 months, a second one located at 1995 and at a period of 64 months, and a third one located at 2008 and at a period of 24 months. According to Table 2.1, three holes in a single 10% pointwise significance patch under the null hypothesis of red-noise is extremely unlikely, if not impossible. On can thus conclude with high confidence that the patch was not generated from a random stochastic fluctuation. Moreover, the discussion in Sect. 2.5.1

suggests that at the very least phase-coherent oscillations were likely present in the Niño 3.4 time series, where phase coherency implies that two oscillations have a stable relative phase relationship but are not necessarily interacting nonlinearly.

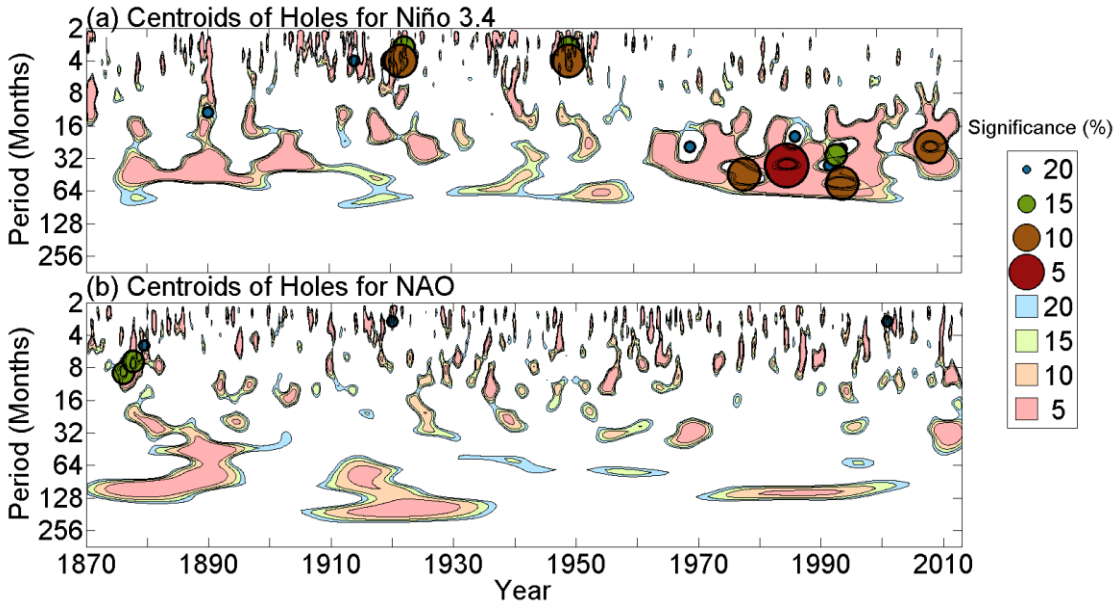

Figure 2.11. Same as Fig. 2.8c but for the mean monthly (a) Niño 3.4 and (b) NAO index anomalies for 1870-2013.

The wavelet topological diagram (Fig. 11b) corresponding to the wavelet power spectrum of the NAO is less interesting, containing few holes at high pointwise significance levels. At 1875, however, a patch contained holes at the 10% pointwise significance level, suggesting that the patch is a significant feature.

2.6. Summary and Discussion

A geometric significance test was developed for more rigorously assessing the significance of features in the wavelet domain. The geometric test, although related to the existing areawise test, was found to be more flexible in the sense that *p*-values could be readily calculated, involving a single Monte Carlo ensemble. Another strength of the geometric test is that the false discovery rate can be controlled at a desire level, minimizing the number of false rejections of the null hypothesis. On the other hand, the geometric test had the disadvantage of being less local than the areawise test.

It is noted that the geometric test was only applied to patches arising from the convolution of the Morlet wavelet with a time series. The results presented in this paper are not valid for wavelet power spectra obtained using other analyzing wavelets, the reason for which is that each wavelet function has different time- and scale-localization properties that inevitably impact the geometry of patches. For example, patches found in the wavelet power spectrum obtained using a Paul wavelet are elongated in the scale direction relative to those obtained using a Morlet wavelet with $\omega_0 = 6$, resulting in nearby patches at different scales merging together. The merging of patches at different scales will alter their geometry with respect to the relatively thin (in scale) patches obtained using the Morlet wavelet.

One disadvantage of the geometric and areawise tests is that they require a binary decision in which pointwise and geometric significance levels must be chosen. The binary decision can be circumvented by applying a *p*-value adjustment procedure to the wavelet power coefficients directly. For example, one could apply the Benjamini and Hochberg (1995) procedure to the wavelet power coefficients or a modified version of the procedure developed by Benjamini and Yekutieli (2002), which is valid for any dependency structure among the local test statistics. The latter procedure would seem most appropriate given the autocorrelation structure of wavelet power coefficients; however, it is noted that the procedure has less statistical power than the original procedure valid for independent local test statistics, though Wilks (2006) found the Benjamini and Hochberg (1995) procedure to remain powerful even when the assumption of independence is violated.

The topology of significant patches was also analyzed. Holes in significant patches, a topological notion, were capable of distinguishing spurious patches from true structures. The holes were identified as arising from phase-coherent oscillations with nearby frequency components and may indicate the existence of a nonlinear interaction. Patches arising from different analyzing wavelets can differ topologically. For the Paul wavelet, the shrinking of patches in time, for example, was found, after a preliminary investigation, to reduce the number of holes in wavelet power spectra. The reduction in the number of holes can be attributed to the tearing of a patch in the time direction. The results, however, require further investigation and are a subject of future work.

The new methods introduced in this paper were applied to the NAO and Niño 3.4 indices, two well-known but contrasting time series. For the Nino 3.4 index, the methods detected geometrically significant structures as well as topological structures unlike that of red-noise, which provide evidence of some predictability of El Niño/Southern Oscillation, which has become of increasing importance in climate science given that its future state is uncertain under a changing global climate system (Latif and Keenlyside, 2008). For the NAO index, the new methods were unable to detect features that are distinguishable from background noise, suggesting that the NAO is a stochastic process with little predictability. The methods developed in this paper will give researchers the tools needed for a better understanding of features found in wavelet power spectra.

Chapter 3

Cumulative Areawise Testing in Wavelet Analysis and its Application to Geophysical Time Series

3.1. Introduction

In many research fields, it is of interest to understand the behavior of time series in order to achieve a deeper understanding of physical mechanisms or relationships. Such a task can be formidable given that time series are composed of oscillations, non-stationarities, and noise. Fortunately, many tools have been developed to extract information from time series, including singular spectrum analysis (Vautard et al., 1992), Fourier analysis (Jenkins and Watts, 1968), and wavelet analysis (Meyers et al., 1993; Torrence and Compo, 1998). The goal of each of these tools is to assess whether deterministic features are embedded in a time series. Fourier analysis, as an example, is a method by which a time series is decomposed into frequency components so that embedded oscillations can be detected. However, the underlying assumption in Fourier analysis is that time series are stationary. The limitation can be circumvented by using a windowed Fourier analysis but with the caveat that the window width is fixed, which can lead to poor resolution at low-frequencies (Lau and Weng, 1995). Wavelet analysis, relaxing the assumption of stationarity, offers an alternative method to Fourier analysis in which the window width is no longer fixed, minimizing aliasing (Meyers et al., 1993; Torrence and Compo, 1998). Wavelet analysis has been demonstrated to be useful in the understanding of the North Atlantic Oscillation (NAO; Higuchi et al., 2003; Olsen et al., 2012), applications to oceanographic problems (Meyers et al., 1993; Lee and Lwiza, 2008; Whitney, 2010; Wilson et al., 2014), assessments of historical hydroclimate variability (Labat, 2004; Labat, 2008), and many other geophysical applications (Grinsted et al., 2004; Velasco and Mendoza, 2008).

When using any time series extraction procedure it is important to assess the statistical significance of the computed test statistic against some null hypothesis. In geophysical

applications, for example, red noise is typically chosen as the null hypothesis. Torrence and Compo (1998) were the first to apply wavelet analysis in a statistical framework using pointwise significance testing, allowing deterministic features to be distinguished from stochastic features. In a pointwise significance test, one tests each estimated wavelet power coefficient against a stationary theoretical red-noise background spectrum. Despite the insights gained from the statistical procedure, it has many deficiencies, as noted by Maruan and Kurths (2004), who showed that it can lead to many spurious results simply due to multiple testing. Addressing the multiple-testing problem, Maraun et al. (2007) developed an areawise test that assesses the significance of so-called pointwise significance patches, contiguous regions of pointwise significance in a wavelet power spectrum. The areawise test, though dramatically reducing the number of spurious results, is computationally inefficient, involving a root-finding algorithm to estimate a critical area of the reproducing kernel corresponding to the desired significance level of the test. Furthermore, the critical area needs to be computed for different analyzing wavelets and for their associated parameters, such as the central frequency for the Morlet wavelet and the order in the case of the Paul wavelet.

A simpler procedure for addressing multiple testing problems is the geometric test developed by Schulte et al. (2015). Like the areawise test, the test statistic for the procedure is based on patch area, or more specifically, the normalized area of the patch, which allows patches at different periods to be compared simultaneously. The calculation of the critical level for the geometric test is much simpler than that for the areawise test, involving the computation of the normalized area for a large ensemble of patches under a null hypothesis that results in a null distribution from which the desired critical level can be calculated.

Both the geometric and areawise tests, however, suffer from a binary decision: one must choose both a pointwise and areawise or geometric significance level. The problem with such a statistical construction is that the outcomes of the testing procedure may depend on the chosen pointwise significance level. For an ideal test, there is a single significance level that is chosen and the results of the testing procedure depend only on that significance level so that a test statistic, for example, that is 1% significant is guaranteed to be 5% significant. In the present case, however, there is no such guarantee: a 1% geometrically

significant patch at one pointwise significance level may not be 5% significant at another pointwise significance level. In such cases, the statistical significance of patches is ambiguous and may preclude further scientific investigation. This sensitivity problem underscores the need to develop a computationally efficient testing procedure free of binary decisions. The approach taken here will consider the areas of patches over all pointwise significance levels, and hence the method is called the *cumulative* areawise test. This test has the important feature that the significance of the wavelet power coefficients is a monotonic increasing function of the pointwise significance level. In other words, a wavelet power coefficient, for example, that is 1% significant under the new procedure will be guaranteed to be 5% significant, a consistent statistical construction.

The paper is organized as follows: in Section 2, a brief description of wavelet analysis is provided together with a discussion of existing statistical testing procedures, including the sensitivity of the geometric test to the chosen pointwise significance, motivating the construction of the cumulative areawise test. Before proceeding to the development of the new testing procedure, the topological properties of red noise are analyzed in Section 3.3. In Section 3.4, the cumulative areawise test is developed and is followed by a comparison of the test in terms of statistical power to the existing geometric test. Applications of the test to prominent climate indices are presented in Section 3.5, followed by concluding remarks in Section 3.6.

3.2. Existing Wavelet Analysis Significance Tests

3.2.1 Wavelet Analysis

The wavelet transform of a time series is defined as the convolution of the time series with a wavelet function ψ_0. The wavelet transform of a time series x_n (n = 1, ... , N) with a wavelet function ψ_0 is given by

$$W_n(s) = \sqrt{\frac{\delta t}{s}} \sum_{n'=1}^{N} x_{n'} \psi_0[(n' - n)\frac{\delta t}{s}], \tag{3.1}$$

where s is the wavelet scale, δt is a time step determined by the data, and N is the length of the time series. There are many kinds of wavelets, but perhaps the most common are the

Morlet, Paul, and Dog wavelets. For geophysical applications, the Morlet wavelet is often used and is given by

$$\psi_0(\eta) = \pi^{-1/4} e^{i\omega_0 \eta} e^{-\frac{1}{2}\eta^2}, \tag{3.2}$$

where ω_0 is the dimensionless frequency, $\eta = s \cdot t$, t is time, and the wavelet scale is related to the Fourier period by $\lambda = 1.03s$ if ω_0 = 6. This particular wavelet balances both frequency and time-localizations. Throughout the paper, ω_0 = 6. The Paul wavelet, which is also a complex wavelet, is more localized in time, less localized in frequency space, and is given by

$$\psi_0(\eta) = \frac{(2i)^m m!}{\sqrt{\pi(2m)!}} (1 - i\eta)^{-(m-1)}, \tag{3.3}$$

where *m* is the order of the Paul wavelet, which controls the localization properties of the analyzing wavelet. In this case, the Fourier period is related to the wavelet scale by the following equation:

$$\lambda = \frac{4\pi s}{2m+1}. \tag{3.4}$$

If even more time-localization is desired, one can use the Dog wavelet, a real wavelet given by

$$\psi_0(\eta) = \frac{(-1)^{m+1}}{\sqrt{\Gamma\left(m+\frac{1}{2}\right)}} \frac{d^m}{d\eta^m} e^{-\eta^2/2}, \tag{3.5}$$

where *m* represents the order of the derivative and Γ is the gamma function. For the Dog wavelet, the Fourier period is related to the wavelet scale by the equation

$$\lambda = \frac{2\pi s}{\sqrt{m+\frac{1}{2}}}. \tag{3.6}$$

In this paper, the Paul wavelet is used with *m* = 4 and the Dog wavelet is used with *m* = 2.

The wavelet power is given by

$$|W_n(s)|^2 \tag{3.7}$$

and represents the wavelet power spectrum of the time series. Inherent in the wavelet transform are edge effects due to the finite time series. In particular, in a wavelet power spectrum there exists a region called the cone of influence, which is defined as the *e*-folding time of the autocorrelation for wavelet power at each scale. The *e*-folding time is defined as the point at which the wavelet power for a discontinuity at the edge drops by a factor of e^{-2} (Torrence and Compo, 1998).

3.2.2 Pointwise Significance Test

In spectral analysis, it is important to assess the statistical significance of spectral power against a noise background. In geophysical applications of wavelet analysis, one often tests each individual wavelet power coefficient against a stationary red-noise background to determine their statistical significance (Torrence and Compo, 1998). For a first-order autoregressive (Markov) process

$$x_n = \rho x_{n-1} + w_n, \tag{3.8}$$

where ρ is the lag-1 autocorrelation coefficient and w_n is Gaussian white noise with $x_0 = 0$, the normalized theoretical stationary red-noise power spectrum is given by

$$p_f = \frac{1-\rho^2}{1+\rho^2-2\rho\cos(2\pi f/N)}, \tag{3.9}$$

where $f = 0, \ldots, N/2$ is the frequency index (Gilman et al., 1963). To obtain, for example, the 5% pointwise significance level ($\alpha = 0.05$), one must multiply Eq. (3.9) by the 95th percentile of a chi-square distribution with two degrees of freedom and divide the result by 2 to remove the degree-of-freedom factor (Torrence and Compo, 1998). The result of the so-called pointwise testing procedure is a subset of wavelet power coefficients whose values exceed the specified background noise spectrum. Recall from Section 3.1 that significant wavelet power coefficients often occur in clusters or contiguous regions of pointwise significance called pointwise significance patches (referred to as patches, hereafter).

Consider the time series of the Atlantic Multi-decadal Oscillation (AMO), North Atlantic Oscillation (NAO), Pacific Decadal Oscillation (PDO), and Niño 3.4 indices shown in Fig. 3.1. The PDO index describes detrended sea surface temperature (SST)

variability in the North Pacific poleward of 20°N latitude (Mantua and Hare, 2002) and the AMO index captures the detrended SST variability in the Atlantic Ocean basin (Kerr, 2002). As shown in Figs. 3.1a and 3.1b, the PDO and AMO indices exhibited multi-decadal variability, with periods of 20-60 years, from, respectively, 1856 to 2014 and 1900 to 2014. The reason for the low-frequency variability of the PDO is subject to debate. Some studies suggest it is the reddened response to white noise atmospheric forcing, whereas other studies hypothesize that it is also the integrated response of the El-Niño/Southern Oscillation (ENSO) signal (Newman et al., 2004).

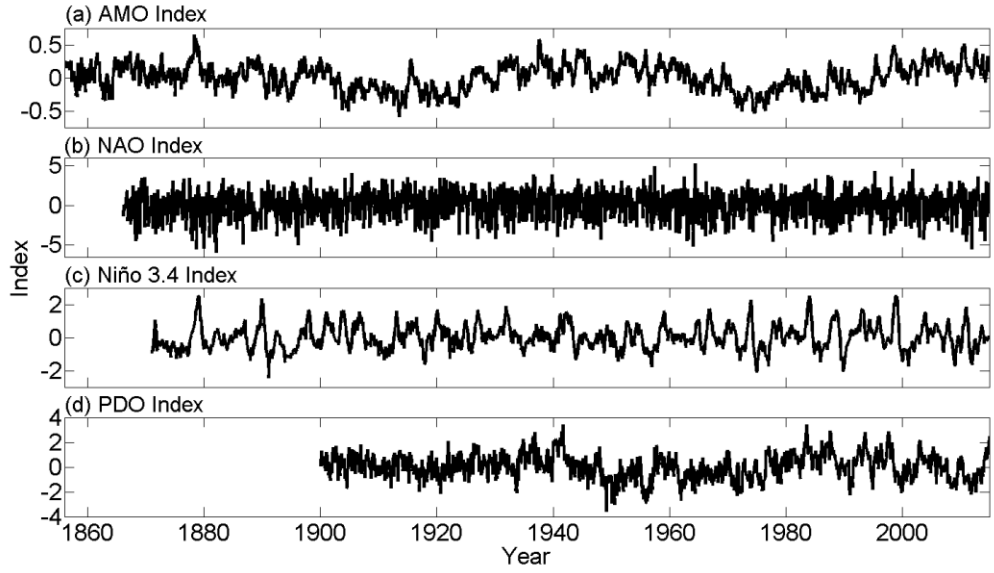

Figure 3.1. The monthly (a) AMO, (b) NAO, (c) Niño 3.4, and (d) PDO indices. Data sources are the Climate Prediction Center for the AMO index (http://www.esrl.noaa.gov/psd/data/climateindices/list/), National Center for Atmospheric Research for the NAO (https://climatedataguide.ucar.edu/data-type/climate-indices) and Niño 3.4 (http://www.cgd.ucar.edu/cas/catalog/climind/TNI_N34/) indices, and University of Washington for the PDO index (http://research.jisao.washington.edu/pdo/PDO.latest). The Niño 3.4 index was converted to monthly anomalies by subtracting off the mean annual cycle.

The NAO index, an atmospheric index, quantifies the difference in sea-level pressure of the Icelandic Low and the Azores High and is related to the strength and position of the jet stream across the North Atlantic (Hurrell et al., 2003). As shown in Fig. 3.1b, the NAO mainly operated on time scales of months and season, and the raw time series is rather noisy. The Niño 3.4 index is an oceanic metric for quantifying the strength of ENSO and is defined as SST anomalies in the Equatorial Pacific in the region bounded

by 120°W-170°W and 5°S-5°N (Trenberth, 1998). The Niño 3.4 index time series exhibited variability on an array of time scales, especially in the 2-7 year period band. Various physical interpretations for the 2-7 year oscillation have been proposed, including the Unified Oscillator, Delayed Oscillator, and the Recharge Oscillator (Wang et al., 2004).

Shown in Fig. 3.2 are the wavelet power spectra of the AMO, NAO, Niño 3.4, and PDO indices. The wavelet power spectrum of the AMO detected enhanced variance at a period of 512 months, as indicated by the thin contour that encloses a region of 5% pointwise significance. All of the other patches are located at periods less than 32 months. The wavelet power spectrum of the NAO indicated that the NAO exhibited enhanced variability on an array of time scales. For example, there is a patch located at a period of 64 months and 1910.

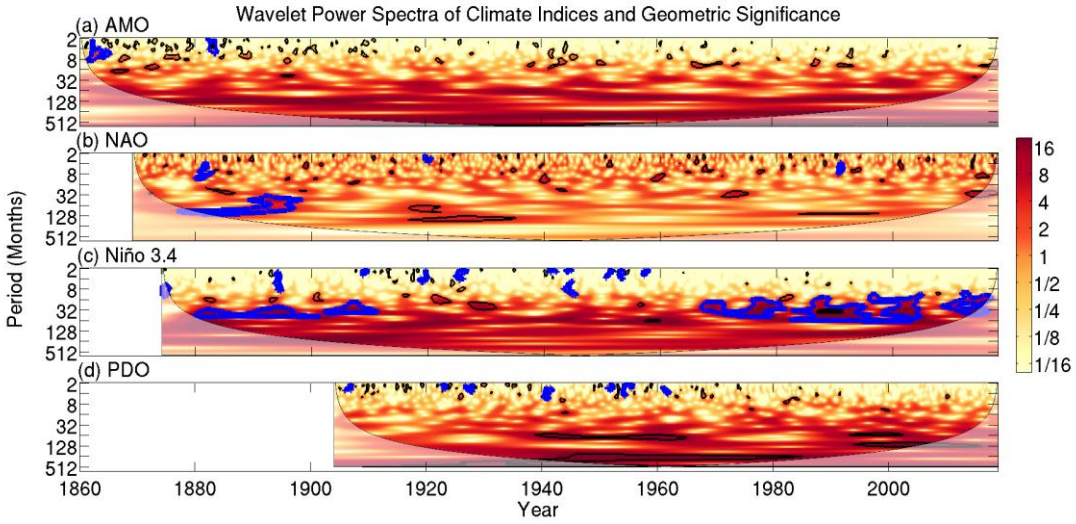

Figure 3.2. Wavelet power spectra of the (a) AMO, (b) NAO, (c) Niño 3.4, and (d) PDO indices. Thin black contours enclose regions of 5% pointwise significance and thick blue contours indicate those patches that are geometrically significant at the 5% level. Light shading represents the cone of influence (COI), the region in which edge effects cannot be ignored.

Like the wavelet power spectrum of the AMO, numerous patches were also found at periods less than 32 months. Large regions of enhanced variance were found in the wavelet power spectrum of the Niño 3.4 index. The largest of these regions was located in the time period 1950-2014 and the period band 16-32 months, consistent with how the ENSO varies with periods of 2-7 years. In the same patch, there are holes as described by Schulte et al. (2015) that may indicate the presence of nonlinearities. Holes are defined

formally as follows. For the closed unit interval I = [0 1], let $f: I \rightarrow P$ be a continuous closed path in a significance patch P. A patch is said to contain a hole if there exists a path that cannot be continuously deformed into a point, where the feature obstructing such a deformation is the hole. Two patches in the same period band were also identified from 1870 to 1890. For the wavelet power of the PDO index, a large patch centered at a period of 512 months extending from 1910 to 2013 was detected. Most of the patches, however, were located at periods less than 8 months, time scales not typically associated with the PDO.

3.2.3 Geometric Significance Test

To determine if the results from the pointwise test are artifacts of multiple testing, a geometric test was applied to the patches located in the wavelet power spectra (Schulte et al., 2015). The test statistic for the geometric test is given by a normalized area

$$A_n = \frac{A}{A_R}, \tag{3.10}$$

where A is the area of the patch and A_R is the area of the reproducing kernel dilated and translated according to the centroid of the patch. Regarding patches as polygons with vertices (x_k, y_k) with k = 1, … , m-1, the area of the patch is determined by a simple formula given by

$$A = \frac{1}{2}\left|\sum_{k=0}^{m-1}(x_k\, y_{k+1} - \; x_{k+1}y_k)\right|, \tag{3.11}$$

where $y_0 = y_m$ and $x_0 = x_m$. Simiarly, the centroids of the polygons are given by

$$C_t = \frac{1}{6A}\sum_{k=0}^{m-1}(x_k + \; y_{k+1})\,(x_k y_{k+1} - \; x_{k+1}y_k) \tag{3.12}$$

and

$$C_s = \frac{1}{6A}\sum_{k=0}^{m-1}(y_k + \; x_{k+1})\,(x_k y_{k+1} - \; x_{k+1}y_k), \tag{3.13}$$

where C_t and C_s are the time and scale coordinates, respectively, of the centroid (Schulte et al., 2015).

To determine the critical level of test, a large ensemble of patches under a noise model is generated and A_n is computed for each patch, resulting in a null distribution from which the desired critical level of the test can be obtained.

The application of the geometric test to the wavelet power spectra of the AMO, NAO, PDO, and Niño 3.4 time series is also shown in Fig. 3.2. For the AMO index 5% geometrically significant patches were identified in the period band 2-16 months prior to 1880, after which no patches were identified as geometrically significant. Note that the large patch centered around the period of 512 months was not found to be geometrically significant at the 5% level, suggesting that the multi-decadal variability is stochastic. A few geometrically significant patches were identified in the wavelet power spectrum of the NAO: one centered at a period of 64 months and 1910, a second one centered at a period of 8 months and 1910, and several others centered at a period of 4 months throughout the record length. Many geometrically significant patches were identified in the wavelet power spectrum of the Niño 3.4 index. For example, a rather large geometrically significant patch is located in the 16-64 month period band from 1950 to 2014. There are also many geometrically significant patches located in the 2-8 month period band from 1900 and 1950. Numerous geometrically significant patches were identified in the wavelet power spectrum for the PDO index, all of which were located in the 2-8 month period band. There was also a large patch centered on a period of 512 months, but the patch was not found to be geometrically significant, suggesting that, like the AMO, the multi-decadal variability arose from stochastic processes.

3.2.4 Sensitivity of the Geometric Test to the Chosen Pointwise Significance Level

To show that the geometric test is sensitive to the pointwise significance level chosen, it will be useful to compute the quantity

$$r = \frac{N_{\alpha_1,\alpha_2}}{N_{\alpha_1}}, \tag{14}$$

where N_{α_1,α_2} is the number of geometrically significant patches at pointwise significance level α_1 that are also geometrically significant at pointwise significance level α_2 and N_{α_1} is the number of patches at α_1 that are geometrically significant, where the geometric

significance level for all cases is fixed α_{geo}. In the ideal situation, r = 1, indicating that geometrically significant patches never lose their geometric significance as the pointwise significance level is increased. This case, however, is optimistic, as the calculation of geometric significance is rather stochastic. To demonstrate the stochastic nature of the geometric test, r was computed for 1000 wavelet power spectra of red-noise processes with lengths 1000 and lag-1 autocorrelation coefficients equal to 0.5 under four different scenarios, the first of which (Scenario 1) is the case in which $\alpha_1 = 0.1$, $\alpha_2 = 0.05$, and $\alpha_{geo} = 0.05$ (Fig. 3.3a). With the mean of r (denoted by $\bar{r}$ hereafter) being 0.3, it can hardly be expected that a geometrically significant patch at $\alpha_1 = 0.1$ to remain significant when the pointwise significance level is changed to $\alpha_2 = 0.05$, at least in the case of red-noise processes. Scenario 2, shown in Fig. 3.3b, is the same as Scenario 1 except that $\alpha_{geo} = 0.01$. In this case, $\bar{r} = 0.15$, suggesting that the geometric test is even more sensitive to the chosen pointwise significance level for smaller α_{geo}. Also note that, unlike the distribution shown in Fig. 3.3a, the distribution is skewed, favoring lower values and supporting the idea that the geometric test is more sensitive to the chosen pointwise significance level for $\alpha_{geo} = 0.01$.

In Scenario 3, $\alpha_1 = 0.05$ and $\alpha_2 = 0.01$, with $\alpha_{geo} = 0.01$. The distribution shown in Fig. 3c is even more skewed than that corresponding to Scenario 2, with $\bar{r} = 0.05$. Also note that in many cases $r = 0$, indicating that there are patches that are not geometrically significant for both $\alpha_1 = 0.05$ and $\alpha_2 = 0.01$. The reason is that some patches existed at $\alpha_1 = 0.05$ but did not exist at $\alpha_2 = 0.01$ so that their normalized areas are zero.

Scenario 4 was similar to Scenario 3 except that $\alpha_{geo} = 0.05$. Although Scenarios 3 and 4 used the same pointwise significance levels, the results differ, with $\bar{r} = 0.22$, suggesting that the geometric test is less sensitive for larger α_{geo}. The results are similar to that of Scenarios 1 and 2, where increasing the pointwise significance level increased the sensitivity of the geometric test to the chosen pointwise significance level.

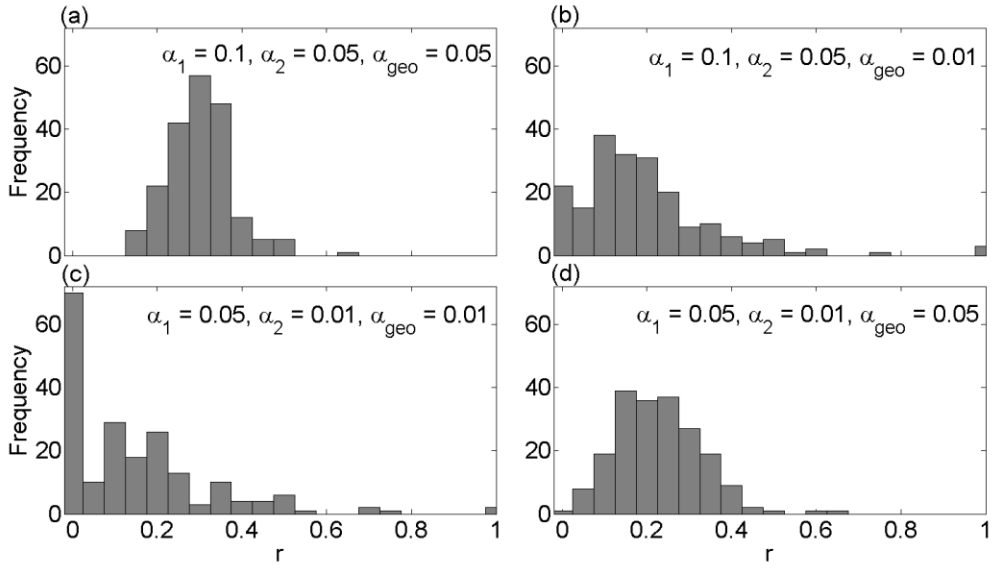

Figure 3.3. (a) A histogram of $r = \frac{N_{\alpha_1,\alpha_2}}{N_{\alpha_1}}$ for $\alpha_1 = 0.1$, $\alpha_2 = 0.05$, and $\alpha_{geo} = 0.05$ obtained from the generation of 300 wavelet power spectra of red-noise processes of length 1000 with lag-1 autocorrelation coefficients equal to 0.5. (b) Same as (a) but with $\alpha_{geo} = 0.01$. (c) Same as (a) but with $\alpha_1 = 0.05$ and $\alpha_2 = 0.01$, and $\alpha_{geo} = 0.01$. (d) Same as (c) but with $\alpha_{geo} = 0.05$.

3.3. Persistent Topology

3.3.1 Persistent Homology

Before developing the cumulative areawise test, it will be necessary to understand the topology of features found in a typical wavelet power spectrum. It will be especially important to understand how the features evolve as the pointwise significance level is increased or decreased. Such information can be obtained using persistent homology, a tool in applied algebraic topology (Edelsbrunner, 2004). Persistent homology will provide a formal setting for calculating the lifetimes of patches and holes, where lifetimes describe when features first appear and when they disappear. For example, a patch appearing at the 5% pointwise significance level and vanishing at the 1% pointwise significance level would have a lifetime of 4. In this paper, the intuitive foundation of persistent homology will be described rather than giving a formal mathematical exposition.

Suppose that A is a patch at the pointwise significance level $\alpha = \alpha_1$ as shown in Fig. 4.4a. One can increase the size of the patch A by increasing α to α_2, which lowers the threshold for significance, resulting in the new geometric realization A' of A shown in Fig.

4.4a. The evolution of the patch can be monitored using a barcode (Ghrist, 2008), which is a collection of horizontal lines representing the birth and death of features. Following the convention of persistent homology, the *y*-axes of barcodes will be denoted by H_0 for patches. In algebraic topology, H_0 are called homology groups and measure the path-connectedness (Appendix D) of sets. The patch *A*, being created at α_1, results in the line segment beginning at α_1 in Fig. 4.4e. Furthermore, the patch neither vanishes nor merges with another patch at α_2 so that the horizontal line continues to α_2. Note that a new patch *B* is created at α_2 because the pointwise significance test is less stringent. The creation of the patch results in a new line starting from α_2. A more complicated situation occurs at α_3, where the patches A' and *B* merge and result in a new patch *C*. In this case, it is unclear if *C* is a geometric realization of A' or *B*. It will therefore be necessary to use the so-called Elder rule from persistent homology (Edelsbrunner and Harer, 2009).

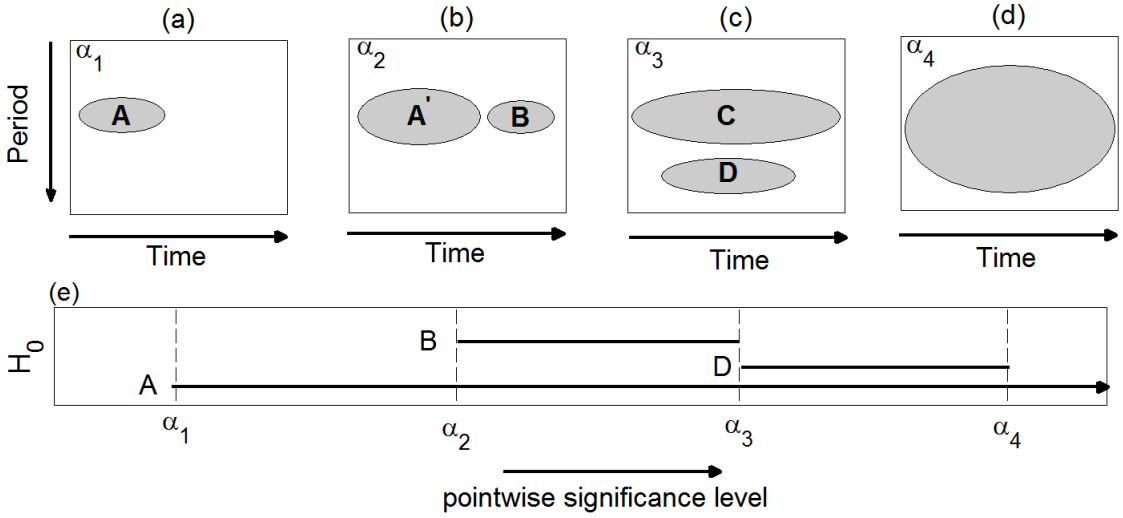

Figure 3.4. (a) - (d) The topological evolution of patches across four pointwise significance levels. (e) The barcode showing the birth and death of patches throughout the evolution process. Horizontal lines with arrows indicate those patches that never die, the so-called essential classes.

According to this rule, the oldest patch will continue to live and the younger patch will die entering the merger point. In the present case, A' is the older patch because it first appears as *A* at α_1 and *B* is the younger patch, being created after *A* at α_2. Therefore, according to the Elder rule, the horizontal line corresponding to *A* in the barcode continues to α_4, but the line corresponding to *B* terminates at α_3, as it dies entering α_3. Also note the creation of a new patch *D* at α_3 and the corresponding line segment in the barcode. Another merger

occurs at α_4 and the Elder rule determines that the line segment for D ends and the horizontal line for A continues, where the arrow indicates that A never dies.

Persistent indices of patches can also be obtained, which will be defined as the difference between the pointwise significance level for which the features die entering and the level at which the features were born. For patches that never die, their persistent indices, by convention, will be set to infinity.

3.3.2 Persistent Homology of Red-noise

To understand the topology of patches generated from red-noise processes, it is useful to use Monte Carlo methods to determine, for example, the number of patches at a particular pointwise significance level, or similarly, the number of holes. Shown in Fig. 3.5a is the ensemble mean of the number of patches (denoted by β_0, hereafter) as a function of α for the Moret, Paul, and Dog wavelets obtained from generating 100 wavelet power spectra of red-noise processes of length 300 and computing β_0 for each of the wavelet power spectra at each pointwise significance level.

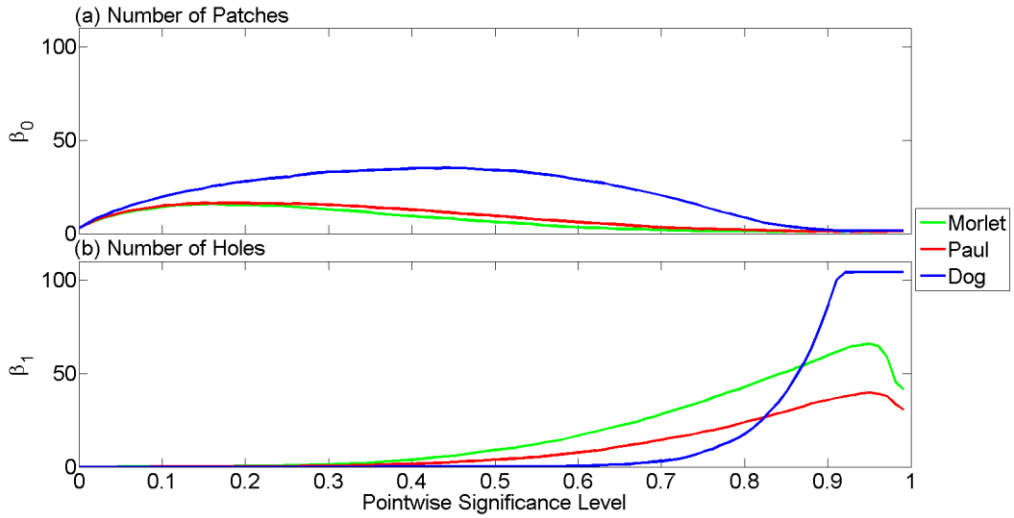

Figure 3.5. (a) Number of patches and (b) the number of holes for three analyzing wavelets as a function of α.

For the the Morlet and Paul wavelets, the number of patches reached minima at $\alpha = 0.01$ and $\alpha = 0.99$ and maxima at $\alpha = 0.18$. The minima in the number of patches for the Dog wavelet was the same as the Morlet and Paul wavelets, but the maximum occurred at $\alpha = 0.5$. Perhaps more interesting are the number of holes for all three

wavelets: very few holes existed at low pointwise significance levels and the number increased rapidly until $\alpha = 0.9$, at which point β_0 no longer increased.

To understand more fully the curves shown in Fig. 3.5a, the persistent homology of patches generated from red-noise processes of length 150 was computed as α varies from 0.01 to 0.99, and barcodes representing the evolution of patches in the wavelet power spectra were computed. In each case the lag-1 autocorrelation coefficients were set to 0.5, but the results are identical for other autocorrelation coefficients. Shown in Fig. 3.6a is a barcode corresponding to a wavelet power spectrum obtained using the Morlet wavelet. Recalling that the beginning of the line segment represents the birth of patches, the barcode indicates that a few patches were born at $\alpha = 0.02$. As α increases to $\alpha = 0.3$ more patches are born, consistent with the fact that more spurious results occur for larger pointwise significance levels. Note that, for $\alpha > 0.2$, patches begin to die, representing the merger of smaller patches into larger patches. The merging process occurs until $\alpha = 0.7$, at which point all patches have merged into a single patch, which is called the essential class. To show that the distribution of the lifetimes for patches generated from red-noise processes is not random, 100 wavelet power spectra of red-noise processes were generated and the persistence indices for all patches in each wavelet power spectra were computed (Fig. 3.6b). The resulting distribution indicates that lifetimes of patches is typically 0.1 and relatively few patches live longer than 0.6. Overall, the distribution characterizes patches generated from red-noise processes as short-lived.

A typical barcode corresponding to a wavelet power spectrum of a red-noise process whose wavelet power spectrum was obtained using the Paul wavelet is shown in Fig. 3.6c. The barcode is similar to that of the Morlet wavelet, with many patches being born before $\alpha = 0.3$ and generally merging for $\alpha > 0.3$. The distribution of persistence indices shown in Fig. 3.6d obtained from Monte Carlo methods as described above suggests that the lifetime patches are typically longer than that generated from the Morlet wavelet; the distribution has smaller negative skewness, and there are more persistence indices within the range 0.6 - 0.85.

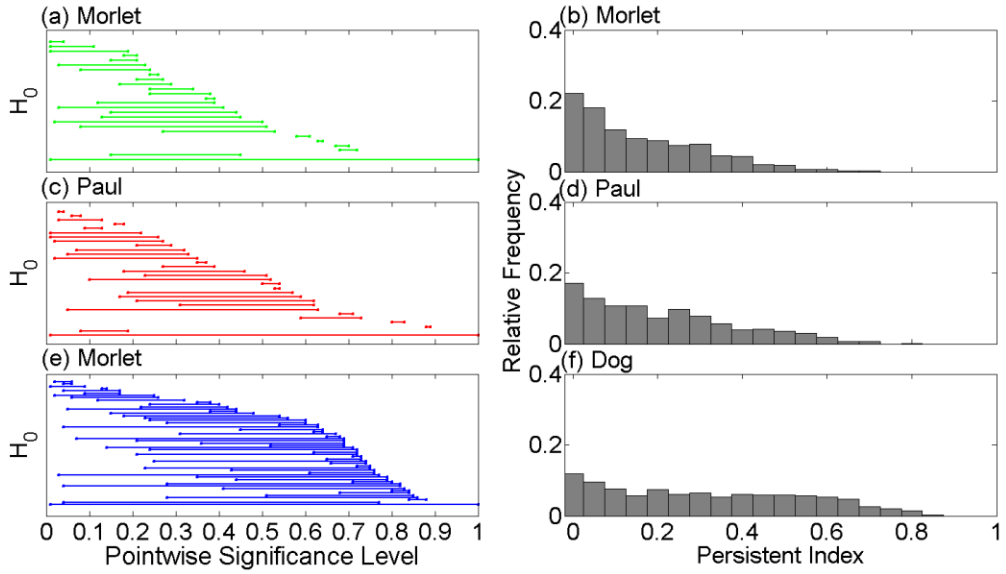

Figure 3.6. (a) Example barcodes for H_0 corresponding to a wavelet power spectrum obtained from the (a) Morlet, (b) Paul, and (c) Dog wavelets. A red-noise process with length 150 and lag-1 autocorrelation coefficient equal to 0.5 was used to create the barcodes. Distribution of persistence indices representing the lifetimes of patches for the (b) Morlet, (d) Paul, and (f) Dog wavelets. Distribution was obtained by generating 1000 wavelet power spectra of red-noise processes with lengths 500 and lag-1 autocorrelation coefficients equal to 0.5. Essential classes have been removed from the distributions.

A barcode corresponding to a wavelet power spectrum of a red-noise process obtained using a Dog wavelet is shown in Fig. 6e. Consistent with Fig. 3.6, the barcode differs from that of the other wavelets, with many patches merging for $\alpha > 0.4$. In fact, unlike the case for the Morlet wavelet, many patches were found to have merged after $\alpha > 0.8$. The merging for large α suggests that patches generated using the Dog wavelet tend to be smaller so that it takes longer to form the essential class. In other words, larger patches represent a larger fractional area of the wavelet domain so that fewer of them can reside in the wavelet domain, whereas a greater number of small patches can exist in a wavelet domain of equal size. The results for the Monte Carlo methods shown in Fig. 3.6f show that the distribution of persistence indices is indeed different. Compared to the Morlet and Paul wavelets, fewer patches were found to have persistence indices less than 0.15, and persistence indices are more uniformly distributed in the range 0.2 - 0.7.

The non-random evolution of patches and holes for red-noise processes suggests that a test can be developed that uses the information of patches at many pointwise significance levels. One would expect that patches arising from signals would behave

differently, breaking apart less frequently when the pointwise significance level is increased. In the extreme case that a patch is the result of taking the wavelet transform of a pure sinusoid, there would be a single patch at all pointwise significance levels and that patch would not break apart. Suppose, however, that white noise was added to the sinusoid such that the patch breaks apart for small α. In this case, the area of the patch would be smaller with respect to the pure sinusoid case, the difference in area arising from the splitting of the patch. On the other hand, if α were increased, then the patch would become a single patch by merging, and the area will be closer to that of the pure case. These facts suggests that a test can be constructed whose test statistic is calculated over a set of pointwise significance levels, capturing the behavior of patches as they evolve. In particular, the test should make use of how the evolution of patches and holes under the null hypothesis of red noise is not random and how the evolution of patches arising from signals may differ from that of the null hypothesis.

3.4. Development

3.4.1 Geometric Pathways

Unlike the geometric test that assesses the significance of patches at a single pointwise significance level, the cumulative areawise test (referred to as the areawise test, hereafter) will assess the significance of a patch as it evolves under a changing pointwise significance level. The goal of the method is to remove the binary decision from which the geometric test suffers. The idea behind the test statistic will be that a patch that is consistently geometrically significant for different pointwise significance levels is more significant than a patch that is only geometrically significant at a single pointwise significance level. In other words, the geometric test simply assesses the significance of patches at a single pointwise significance level but does not take into account that the patch could have been significant at that particular pointwise significance level by chance (Fig. 3.3).

The first step of the areawise test is to select a finite set of pointwise significance levels that remains fixed throughout the testing procedure. Although there are infinite number of pointwise levels to choose from, it will be shown through empirical arguments

that the selection need only be limited to a finite set of pointwise significance levels. Furthermore, it will be shown that one needs only to limit the test to a certain range of pointwise significance levels, making the test more computationally feasible. The selection of pointwise significance levels will be discussed in depth in Sections 3.4.2. and 3.4.5. For the theoretical development of the testing procedure, the pointwise significance levels will be left unspecified.

The evolution of a patch under a changing pointwise significance level will be made precise by introducing the notion of a geometric pathway. A geometric pathway will be defined as a collection $\mathcal{P}$ of r patches at the corresponding pointwise significance levels $\alpha_1 < \alpha_2 < \ldots < \alpha_r$ such that

$$P_1 \subset P_2 \subset P_3 \subset \cdots \subset P_r \tag{3.15}$$

and

$$g_1 < g_2 < g_3 \ldots < g_r, \tag{3.16}$$

where each g_j is a normalized area corresponding to the patch P_j. For this testing procedure, the normalized area will be calculated by dividing the patch area by the scale coordinate of the centroid squared. The inequalities (3.16) are guaranteed to hold for any nested sequence (3.15) (Appendix E). Viewing the α_j's as time parameters, one can think of the patch being in its initial configuration at α_1 and at its final configuration at α_r. The length of a pathway will be given by r, the number of elements in the pathway. If the computation of the geometric pathways is limited to an interval $I = [\alpha_{min}, \alpha_{max}]$, then all pathways will end at the same pointwise significance level but need not begin at the same pointwise significance level. The reason why all pathways end at α_{max} is because, once a pathway is generated, it can never die, as elements of $\mathcal{P}$ grow in size relative to its initial element.

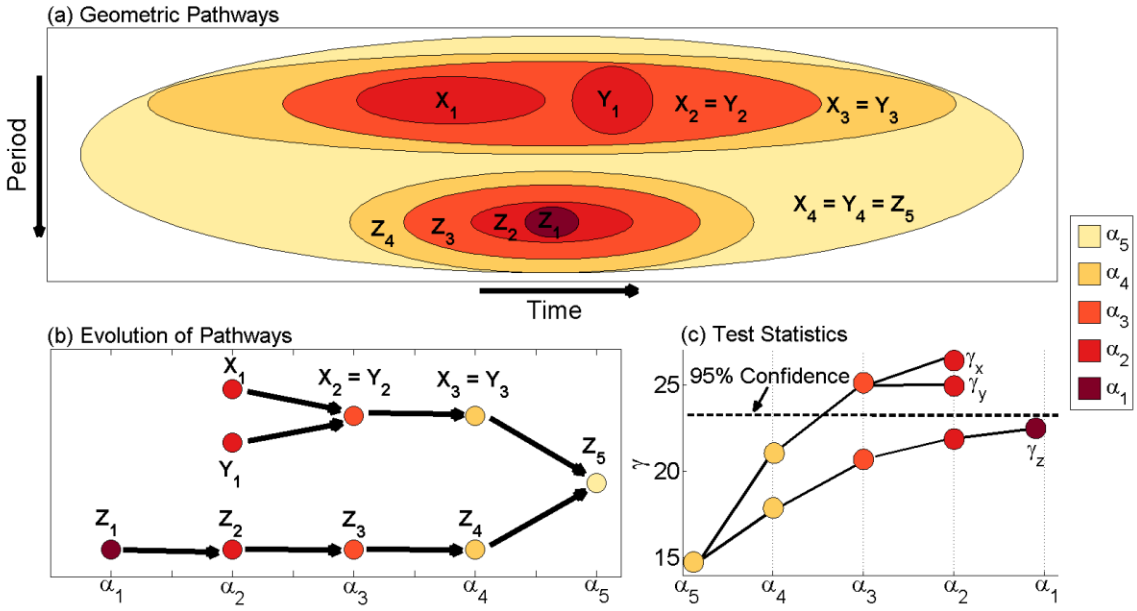

Figure 3.7. (a) Geometric pathway of three significance patches, X, Y, and Z in the interval I = [α_1, α_5]. (b) The geometric evolution of the pathways showing how Z_5 was created from the merging of X_3 and Z_4 as α changed from α_4 to α_5. (c) The cumulative areas at each step of pathway for each geometric pathway, where the summation begins at α_5, the initial point of the pathway, and the dotted line represents the critical level of cumulative areawise test.

To illustrate the idea of a pathway, it is perhaps best to consider an ideal case (Fig. 3.7). Consider, for example, three pathways *X*, *Y*, and *Z* whose lengths are, respectively, $r_x = 4$, $r_y = 4$, and $r_z = 5$. The pathway *X* can be written explicitly as

$$X_1 \subset X_2 \subset X_3 \subset X_4, \tag{17}$$

indicating that the patch exists at $\alpha_1^x = \alpha_2$, $\alpha_2^x = \alpha_3$, $\alpha_3^x = \alpha_4$, and $\alpha_4^x = \alpha_5 = \alpha_{max}$. There also exists another pathway *Y* such that $X_2 = Y_2$, $X_3 = Y_3$, and $X_4 = Y_4$ but whose initial element Y_1 is distinct from *X*. The pathways are thus distinct until α_3, at which point the pathways merge, resulting in the remaining elements being identical (Fig. 3.7b). The third pathway Z, on the other hand, shares only one element with *X* and *Y*, merging at α_5, so that *Z* represents a distinct pathway. Unlike the other pathways, the pathway *Z* begins at $\alpha_1 = \alpha_{min}$, implying that the pathway exists over the entire interval *I*.

The development of the areawise test will require the calculation of a test statistic for all pathways in the wavelet domain that end at the same pointwise significance level for a selected interval *I*. The test statistic used in this procedure will be the total sum of normalized areas

$$\gamma = \sum_{j=1}^{r} g_j, \tag{3.18}$$

one for each pathway.

The calculation of the critical level for the test can be computed using Monte Carlo methods as follows: (1) fix *I* and a partition of the interval with uniform spacing and generate red-noise processes with the same autocorrelation coefficients as the input time series; (2) for each red-noise process generate synthetic wavelet power spectra and all corresponding pathways; and (3) for every pathway, compute γ, resulting in a null distribution from which the desired critical level of test can be obtained. The critical level corresponding to the 5% significance level of the test, as an example, is the 95% percentile of the null distribution.

3.4.2 Pointwise Significance Level Selection: Maximization Method

It will often happen that pathways under consideration are of different lengths so that it is not obvious what pointwise significance patches to report after the implementation of the testing procedure. The problem can be circumvented, however, using the following procedure: Let

$$\gamma_j = \sum_{i=0}^{r-j} g_{r-i} \tag{3.19}$$

be the cumulative sum associated with the jth element P_j of a pathway with length *r*, and let γ_{crit} be the critical level of the test; then, the appropriate pointwise significance level to use for a pathway is determined by the following rule:

$$\gamma_{\max} = \max_{j=1,2\ldots,r} \gamma_j > \gamma_{crit}, \tag{3.20}$$

where the statistic satisfying the above rule is denoted by γ_{max}. The element of the pathway corresponding to $\gamma_{\max}$ is the output of the testing procedure. Note that the output elements need not be located at the same pointwise significance levels, contrasting with the pointwise and geometric tests. If, for a given pathway the statistic γ_{max} does not exist, then the pathway is deemed insignificant. It is also important to note that using the inequality (20) guarantees that the appropriate number of pathways will be deemed insignificant if the null hypothesis is known to be true, the reason for which is that if *M*

pathways have statistics γ such that $\gamma > \gamma_{crit}$, then there also must exist *M* statistics γ_{max} satisfying the inequality (3.20). In other words, the sums γ must have crossed the critical level at some point in the pathways, or, otherwise, they would have not been deemed significant.

3.4.3 Application to Ideal Pathways

The testing procedure is now demonstrated using an ideal case. For the pathway *X* shown in Fig. 3.7 the test statistic at α_5 is equal to the area of a single patch in the pathway, i.e.

$$\gamma_x = A_4^x, \tag{3.21}$$

where A_j^x denotes the area of a pathway element at α_j^x. The test statistic at α_1^x, on the other hand, is the sum of four areas

$$\gamma_x = A_1^x + A_2^x + A_3^x + A_4^x \tag{3.22}$$

so that the test statistic reaches its maximum value at the smallest pointwise significance level for which the pathway exists. In fact, all pathways will satisfy this property. The evolution of the patch can be viewed as a shape changing with time as shown in Figure 3.7b. Viewing α_5 as a time parameter, one can say that *X* merges with a large patch at time α_5 and first appears at time α_2. The evolution of the test statistic is shown in Figure 3.7c, in which the test statistic increases with decreasing α_j such that its maximum value is attained at α_2 .

Now consider the pathway *Y*, whose maximal test statistic is similar to that of *X* except that A_2^y differs from A_2^x. According to Fig. 3.7b, the pathways evolved identically from α_3 to α_5, merging at α_3. As shown in Fig. 3.7c, the evolution of γ_y only differs slightly from that of γ_x and also represents a significant pathway. For both *X* and *Y*, the condition $\gamma_{max} > \gamma_{crit}$ is satisfied at α_3 so that a single patch at α_3 will be the output of the testing procedure, not the two patches representing distinct elements of *X* and *Y* at α_2. As will be shown in Sect. 3.4.6, the overall effect of this merging process is to enhance the detection of signals.

For the pathway *Z*, the maximum value of the test statistic can be decomposed into five summands, i.e.

$$\gamma_z = A_1^z + A_2^z + A_3^z + A_4^z + A_5^z \,. \tag{3.23}$$

The pathway *Z*, being longer than *X* and *Y*, therefore has an additional pointwise significance level to allow the test statistic to exceed the critical level of the test. However, the elements of the pathway have smaller normalized areas so that the pathway is not significant. The length of a pathway is thus not the only factor influencing the significance of a pathway, as the size of elements is also important.

3.4.4 The Null Distribution

Recall from Section 3.3.2 that, for patches generated from red noise, the merging of patches is not random, with typical lifetimes of patches following a non-uniform distribution, favoring shorter lifetimes. Therefore, if the test statistic is proportional to the lifetime of patches, one can expect the test statistic to follow a similar distribution to that of persistence indices, where the smallest values of γ are preferred. To test this hypothesis, 1000 wavelet power spectra of red-noise processes with fixed autocorrelation coefficients were generated and the cumulative area of all pathways in each wavelet power spectra was computed. In the experiments, $\alpha_{min} = 0.02$, $\alpha_{max} = 0.82$, and $\Delta\alpha$, the discrete spacing between adjacent pointwise significance levels, is set to 0.02 to make calculations less computationally expensive. The results for the Morlet, Paul, and Dog wavelets are shown

in Figure 3.8.

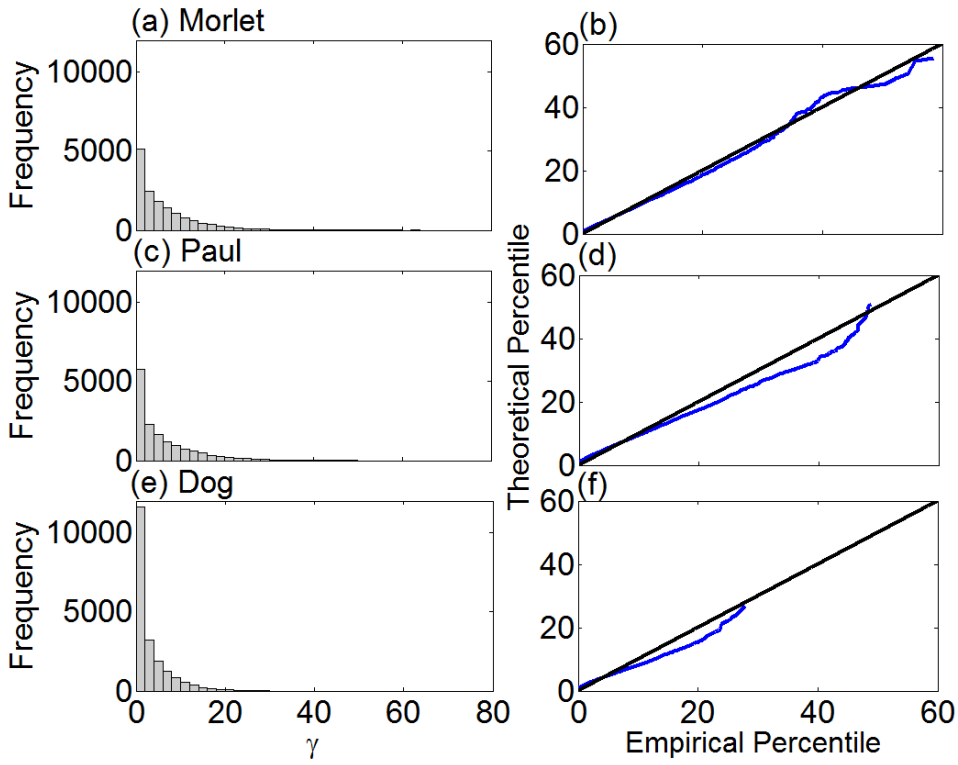

Figure 3.8. (a) Null distribution of γ for the Morlet wavelet obtained by generating 10,000 geometric pathways under the null hypothesis of red-noise, where the red-noise processes were of length 1000 and had lag-1 autocorrelation coefficients equal to 0.5. (b) Percentiles of a theoretical exponential distribution with mean 6.5 plotted as a function of the percentiles calculated from the distribution shown in (a). (c) – (d) Same as (a) – (b) but for the Paul wavelet. (e) – (f) Same as (a) – (b) except for the Dog wavelet.

The distribution of γ for the Morlet wavelet is generally similar to the shape of the distribution for the persistence indices for H_0 (Fig. 3.6b), where the smallest values of γ are preferred. It turns out that the distribution of γ for the Morlet wavelet can be well described by an exponential distribution. Using the method of maximum likelihood (Weerahandi, 2003), a theoretical exponential distribution was fitted to the empirical distribution, where the empirical distribution was found to be best described by an exponential distribution with mean 6.5. To show that the theoretical distribution models the empirical distribution, the percentiles of a theoretical exponential distribution with mean 6.5 were plotted as a function of the percentiles of the empirical distribution (Fig. 3.8b). The linear relationship between the percentiles shown in Fig. 3.8b indicates that the theoretical distribution well models the empirical distribution, with the 95% percentiles only differing by 1.0.

The results for the Paul wavelet are generally similar except that the mean value of the test statistic was found to be smaller with a value of 6.2. The distribution of the test statistic, unlike that for the Morlet wavelet, was found to be poorly modeled by an exponential distribution, with the empirical distribution generally having larger values than those predicted by a theoretical exponential distribution; this difference was more pronounced for larger γ.

The results for the Dog wavelet are shown in Fig. 8e. The null distribution differs from that for the Morlet and Paul wavelets, with the mean value of the distribution being 3.4. This result is perhaps not surprising; Fig. 3.6 shows that the lifetimes of patches generated using the Dog wavelet have larger persistence indices attributed to how patches die at larger α compared to that of the Morlet and Paul wavelets, resulting from the lack of merging with the essential class. The delayed merging implies that the area of patches must be smaller relative to that of the Morlet and Paul wavelets for large α to prevent merging with other patches. The effect is to generate a smaller cumulative sum. The null distribution, unlike that for the Morlet wavelet, was not found to be well-described by a theoretical exponential distribution.

As will be shown in Section 3.4.6, the implementation of the proposed testing procedure will allow small patches that are seemingly indistinguishable from noise to emerge, allowing the recovery of a signal that has been contaminated by noise.

3.4.5 Computational Remarks

An important parameter in the areawise test is $\Delta\alpha$, the spacing between pointwise significance levels. It is critical that $\Delta\alpha$ is chosen to be small enough to sample the merging of patches so that, for example, the null distribution can be representative of the null hypothesis. If $\Delta\alpha$ is set too large, then a pathway may be missed entirely if it is born between two adjacent pointwise significance levels. On the other hand, if $\Delta\alpha$ is too small, then the test will become computationally expensive, requiring the calculation of more normalized areas. The distribution of the persistence indices for H_0 suggests that $\Delta\alpha = 0.01$ because the modes of the distributions are 0.01. However, it will be shown in Section 3.4.6 that the results of the areawise test do not differ if one chooses $\Delta\alpha = 0.02$.

Another important parameter is α_{max}. It would be computationally inefficient if the testing procedure had to be performed on pathways whose end points are, for example, $\alpha_{max} = 0.99$. Moreover, as shown in Fig. 3.5b, the number of holes increases for increasing α, which would increase the computational costs associated with the calculation of patch areas for large α. The reason for the increase in computational costs is that the areas of the holes need to be subtracted from the area the patches would have if they did not contain holes. To circumvent the problem, properties of the areawise test can be assessed at peaks of the curves shown in Fig. 3.5a so that results for other end points can be inferred. For the Morlet wavelet, the peak occurs at α = 0.18 so that $\alpha_{max} = 0.18$. To the left of the maximum, the average number of patches are equal to that of some pointwise significance level to the right of the maximum. However, patches to the right of the maximum have larger areas so that choosing $\alpha_{max} > 0.18$ would result in the testing procedure detecting larger patches as significant. The same reasoning holds for the Paul and Dog wavelets except that the maximum of the curve for the Dog wavelet is $\alpha_{max} = 0.5$. The sensitivity of the testing procedure to α_{max} is discussed in Section 3.4.6.

Fig. 3.6 also suggests that α_{max} need not be any larger than 0.7 for the Morlet wavelet because all patches have merged with the essential class before that point, at least for red-noise processes. Thus, for $\alpha_{max} > 0.7$, patches arising from signals cannot be distinguished from those generated from noise because all patches have merged into a single, large patch. A similar argument holds for the Paul wavelet, but, for the Dog wavelet, all patches merge at a larger α_{max}.

3.4.6 Comparison with geometric test

With the areawise test now developed, it will be useful to assess the statistical power of the test relative to that of the geometric test. The first aspect of the assessment will be to quantify how well both tests detect true positive results. To do so, let

$$x(t) = A\sin(2\pi f t) + w(t) \tag{3.24}$$

be a sinusoid with amplitude *A*, frequency *f*, and additive Gaussian white noise *w*(*t*). The goal will be to evaluate the ability of both tests to detect true positives within a particular period band. A theoretical patch to which the ability of the geometric and areawise tests

were compared was constructed as follows: (1) the time series $x(t)$ for all $t \in [0, 500]$ was generated but with no additive white noise; (2) the wavelet power spectra of $x(t)$ was computed and the 5% pointwise significance test was performed on the wavelet power spectrum; and (3) the width of the significance patch in the wavelet power spectrum was calculated at t = 250 where edge effects are negligible. The theoretical patch calculated using the Morlet wavelet is indicated by dotted lines in Fig. 3.9, where the theoretical patch is a rectangle of fixed width extending from t = 0 to t = 500. In all experiments, α_{max} = 0.18 and $\Delta\alpha = 0.02$, but implications of other choices are discussed at the end of the section.

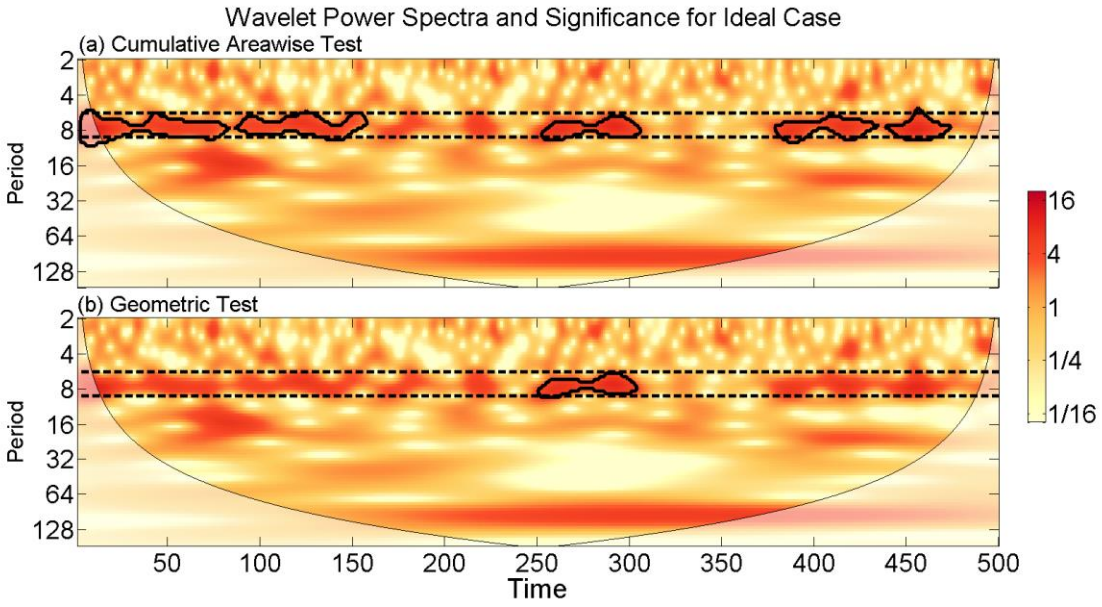

Figure 3.9. (a) Cumulative areawise test applied to a sinusoid with a frequency of 0.8 and amplitude equal to 0.8. Signal-to-noise ratio is 1.0. Contours represent patches that are elements of 5% significant pathways. Dotted lines represent the upper and lower boundaries of a theoretical patch obtained by generating the wavelet power spectrum of a pure sine wave and calculating the width of the patch at t = 250. (b) Same as (a) except for the geometric test with $\alpha = 0.05$ and $\alpha_{\text{geo}} = 0.05$. Contours represent patches that are geometrically significant.

To assess the ability of the tests to detect true positives, the area of patches deemed significant by the tests were compared to the total area of the theoretical patch. More specifically, if P_{geo} is the union of all pointwise significance patches at α that are geometrically significant at the α_{geo} level and P_{theory} is the theoretical patch, then

$$r_a = \frac{A_{P_{geo} \cap P_{theory}}}{A_{P_{thoery}}} \tag{3.25}$$

represents the areal fraction of P_{theory} detected by the geometric test, where $A_{P_{geo} \cap P_{theory}}$ denotes the area of $P_{geo} \cap P_{theory}$ and $A_{P_{thoery}}$ denotes the area of P_{theory}. If r_a = 1, then the test detected all of the true positive results that are known by construction. Small values of r_a indicate that the tests performed poorly, detecting only a fraction of the theoretical patches to be significant. A similar construction can be made for the areawise test by replacing P_{geo} with P_c. Fig. 3.9a illustrates the procedure for the areawise test when f = 0.8, A = 0.8, and the signal-to-noise ratio (defined below) equals 1.0. As indicated by the thick black contours, the areawise test was able to detect 30% of the true positives comprising the theoretical patch, whereas Fig. 3.9b shows that the geometric test was only able to detect 20% of the true positives. It will be necessary to compute N = 1000 values of r_a for different values of f and signal-to-noise ratios of the Gaussian white noise to determine if the tests truly perform differently. In the following experiments, the signal-to-noise ratio is defined as

$$\sigma = 10\log\left(\frac{p_{signal}}{p_{noise}}\right), \tag{3.26}$$

where

$$p_{signal} = \frac{A^2}{2}, \tag{3.27}$$

p_{noise} is the average power of the Gaussian white noise, and σ is measured in decibels (DB). In the discussion below the results for the Morlet are presented first; results for the Paul and Dog wavelets are discussed at the end of the section. It is also noted that because σ and A do not vary independently there is no need to perform different experiments for different values of A. For the experiments, A was set to 1.0.

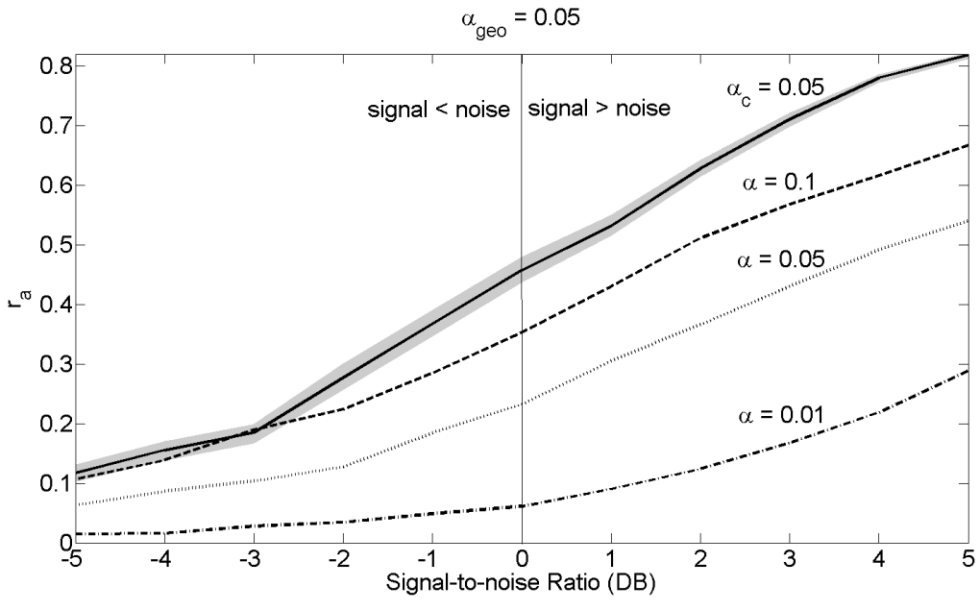

Figure 3.10. The ensemble mean r_a as a function of the signal-to-noise to ratio for the areawise test with $\alpha_c = 0.05$ and the geometric test with $\alpha_{geo} = 0.05$. Gray shading represents the 95% confidence interval and all means for the geometric test are significantly different at the 5% level from the means for the areawise test except for those corresponding to the $\alpha = 0.01$ curve for signal-to-noise ratios less than -1.5. The confidence intervals and statistical significance were obtained by the bootstrap method (Efron, 1979). The data for each signal-to-noise-ratio were sampled with replacement 1000 times to generate a distribution of bootstrap replicates, from which 95% confidence intervals were obtained. Two ensemble means were said to significantly different at the 5% level if their 95% confidence intervals did not intersect.

In the first experiment, the areawise significance (denoted by α_c, herreafter) was set to 0.05, α_{geo} = 0.05, and α = 0.01, 0.05, 0.1. The signal-to-noise ratio for the Gaussian white noise was varied from -5 DB to 5 DB. The results are shown in Fig. 3.10. For both tests, the ability to detect true positives increased with increasing signal-to-noise level. At low signal-to-noise ratios, the tests performed similarly, detecting on average 10% of true positives. On the other hand, differences between the test performances became larger as the signal-to-noise ratio was increased; in fact, the areawise test outperformed the geometric test regardless of the chosen pointwise significance levels when $\sigma \geq -2.5$ DB. The results indicate that the areawise test is particularly useful in low noise situations but one can expect the test to detect more true positives even in high noise conditions. It also worth noting that the performance of the geometric test depended strongly on the chosen pointwise significance level, especially when the signal power was high.

A second experiment was conducted where α_c = 0.01 and α_{geo} = 0.01. The same pointwise significance levels as the first experiment were chosen and the range of signal-

to-noise ratios was also the same. The results are presented in Fig. 3.11. Note that the shape of the curves for both tests are the same as the first experiment, where greater true positive detection is preferred for large signal-to-noise ratios. However, both tests detected fewer true positives, consistent with how the significance levels of the tests were increased. The relationships between the tests, like for experiment 1, depended on the signal-to-noise ratio; for low signal-to-ratios the performance of the tests are similar, whereas for high signal-to-noise ratios differences between r_a for the areawise test and r_a for the geometric increases. All the results were found to be statistically significant.

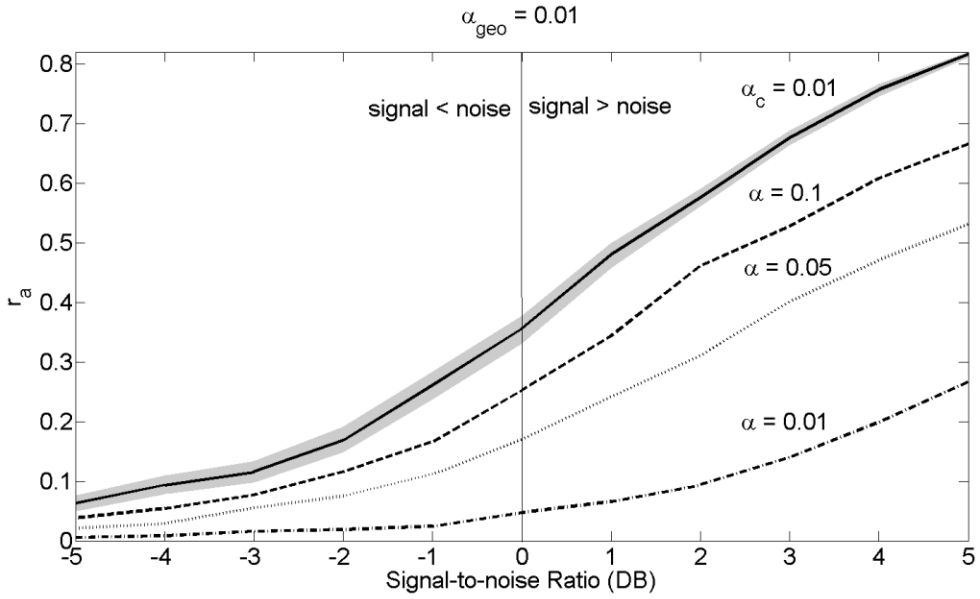

Figure 3.11. Same as Figure 3.10 except with $\alpha_c = 0.01$ and $\alpha_{geo} = 0.01$. All means for the geometric test are significantly different at the 5% level from the means for the areawise test

Additional experiments were performed using different values of *f* to determine if the frequency at which patches are located affects the performances of the areawise and geometric tests. True positive detection, for a fixed σ, was generally found to increase for larger *f*, though the areawise test was still found to detect more true positives. The behavioral assessments of both tests for $f < 0.8$ required the use of synthetic time series whose lengths were greater than 500, as patches lengthened in the time direction for lower frequencies and the COI impacts became more pronounced. The array of experiments concluded that the areawise test should be the preferred method for signal detection.

Another set of experiments were conducted to evaluate how the areawise test detects false positives relative to the geometric test. In the first experiment, 1000 wavelet power spectra arising from red-noise processes of length 1000 with equal lag-1 autocorrelation coefficients were generated. For each wavelet power spectra, the geometric and areawise tests were applied at the 0.05 level and the pointwise significance level was set to 0.05. The results were found to be independent of the chosen lag-1 autocorrelation coefficient, so only the results for the case when the lag-1 autocorrelation coefficient was set to 0.5 are presented. The ratio between the number of false positive results for the geometric and areawise tests for each wavelet power spectra was then computed. The average ratio was found to be 0.08, implying that one can expect more false positive results for the areawise test relative to the geometric test. However, in most cases, the increase in the number of false positive results with respect to the geometric test was small compared to the increase in true positive detection (Figs. 3.10 and 3.11), suggesting that the areawise test had greater overall statistical power. Note that spurious patches arising from the areawise test will have generally larger areas so that results from the test should be interpreted carefully.

The above experiment was repeated for different pointwise significance levels. The experiments indicated that the difference between false positive detection between both tests decreased as α increased, the reason for which is that the number of false positive results for the geometric test for a fixed α_{geo} will increase until $\alpha = 0.18$, the point at which the number of patches generally peaks (Fig. 3.5a). For $\alpha > 0.18$, the number of spurious patches may decrease, but the areas of the spurious patches resulting from the geometric test will be larger. A similar argument would hold if α_c and α_{geo}were increased or decreased by the same amount.

All the above experiments were performed with the Paul and Dog wavelets and the theoretical 5% pointwise significance patch was adjusted to account for the different scale- and time-localization properties of the wavelets, where the corresponding theoretical patches were found to be wider in scale compared to the Morlet wavelet. The results from the experiments were qualitatively similar to that of the Morlet wavelet, as the areawise test detected more true positives than the geometric test for both the Paul and Dog wavelets.

Like for the Morlet wavelet, the differences in performances were more pronounced for larger σ. The array of experiments provides evidence that the areawise test has greater statistical power than the geometric test regardless of the chosen analyzing wavelet.

All the above experiments were conducted for different values of $\Delta\alpha$. For $\Delta\alpha < 0.02$, the results of the experiments were virtually identical. On the other hand, if $\Delta\alpha > 0.03$, the percentage of true positives detected by the areawise test decreased. The results suggest that $\Delta\alpha$ should be chosen to be no larger than 0.02 to ensure that true positive detection is maximized.

The parameter α_{max} was found to strongly influence the performance of the areawise test. For $\alpha_{max} > 0.18$ and $\alpha_c = 0.05$, the number of true positives detected increased and the number of false positives decreased relative to the geometric test with $\alpha_{geo} = 0.05$. However, the normalized areas of the spurious patches resulting from the areawise test with $\alpha_c = 0.05$ or $\alpha_c = 0.01$ were found to be 2-10 times larger than those resulting from the geometric test, making spurious features appear significant. The problem was remedied by decreasing α_c, but this adjustment was found to be the same as increasing α_{max}. For $\alpha_{max} < 0.18$, the true and false positive detection of the areawise test for $\alpha_c = 0.05$ approached that of the geometric test until $\alpha_{max} = 0.05$, at which point they were approximately equal.

3.5. Climate Applications

To determine if any features in the wavelet power spectra of the AMO, NAO, Niño 3.4, and PDO time series are deterministic or stochastic, the areawise test was performed on geometric pathways in the wavelet power spectra at the 0.01 level. A red-noise background spectrum was used for each, with $\alpha_{max} = 0.18$, $\alpha_{min} = 0.02$, and $\Delta\alpha = 0.02$. The wavelet power spectrum for the AMO index is shown in Fig. 3.12a. Although the AMO is usually characterized by its multi-decadal variability, no areawise significant wavelet power coefficients were detected at periods greater than 24 months, suggesting that the variability results from stochastic processes. In contrast, from 1860 to 1940, areawise-significant wavelet power coefficients were detected, primarily in the 2-16 month

period band. After 1940, no areawise-significant wavelet power coefficients were identified at any time scale.

The wavelet power spectrum of the NAO index is shown in Fig. 3.12b. Only a few regions of areawise significance were detected: one at a period of 2 months and 1900, a second one at a period of 6 months and 1880, and a third one located at a period of 100 months and 1870. The lack of areawise significance suggests that the NAO is a red-noise process with no preferred time scale, consistent with Feldstein (2002).

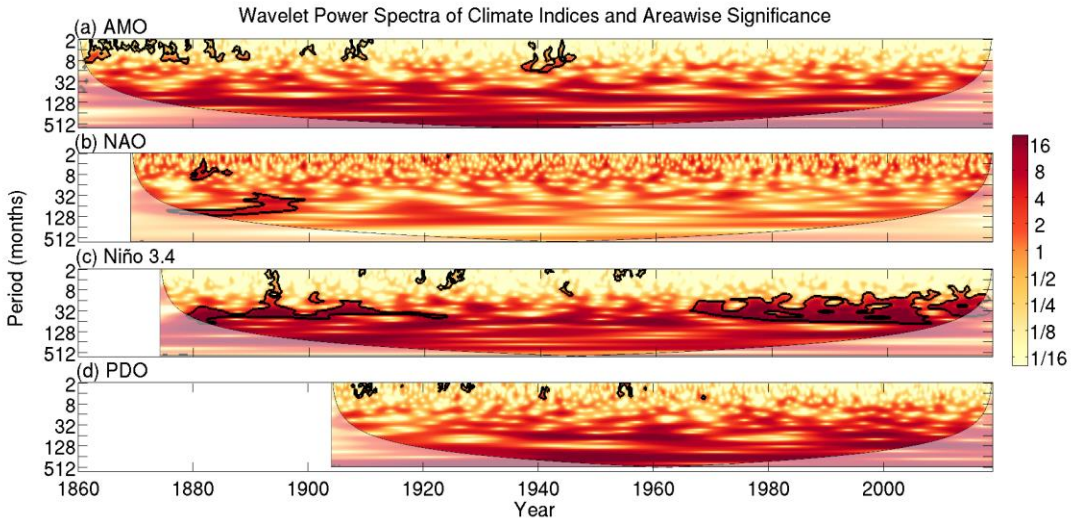

Figure 3.12. (a) The application of the cumulative areawise test to the AMO index (a) the NAO index, (c) the Niño 3.4 index, and (d) the PDO index. α_c was set to 0.01 in all cases and contours enclose regions of 1% areawise significance.

The wavelet power spectrum for the Niño 3.4 index, on the other hand, does indicate potential predictive capabilities (Fig. 3.12c). There are two notable features, one extending from 1870 to 1920 in the 16-64 month period band and another one extending from 1960 to 2014 in the 8-64 month period band. Perhaps just as interesting is the deficit in areawise significance from 1920 to 1960 in the 8-64 month period band. The deficit could be the result of the 2-7 year mode being modulated by a decadal ENSO mode, a nonlinear paradigm (Timmermann, 2002). Such a modulation would imply that the behavior of the 2-7 year mode is determined by the phase of the decadal mode, where, for example, more extreme El Niño phases would be favored if the decadal mode is in a positive regime. On the other hand, results shown in Fig. 3.12c show that neither the

decadal nor the multi-decadal variability exceed a red-noise background so modulations would be difficult to predict.

The wavelet power spectrum of the PDO index is shown in Fig. 3.12d. Like the AMO index, there is enhanced variance at multi-decadal time scales but the variance does not exceed a red-noise background. Areawise-significant regions, however, were detected in the 2-8 month period band from 1900 to 1960. The results indicate that the PDO is a red-noise process, consistent with prior work showing that the PDO results from the oceanic integration of atmospheric white-noise stochastic forcing (Newmann et al., 2003).

3.6. Conclusions

An areawise test was developed for assessing the significance of features in wavelet power spectra. The test was generally found to have greater statistical power than the geometric test except possibly under high-noise situations, in which case the tests were found to perform similarly. The main advantage of the new testing procedure is that the results are no longer dependent on the chosen pointwise significance level. In contrast, the geometric test results were found to be very sensitive to the chosen pointwise significance level, making it difficult for researchers to decide what patches are significant and what patches are not significant. In particular, the cumulative areawise test was found to detect more true positives relative to the geometric test for some common pointwise and geometric significance levels. The large increase in true-positive detection of the cumulative areawise test was also accompanied by a small increase in false-positive detection compared to the geometric test performed at the 5% level, resulting in the areawise test having greater statistical power. The difference between the tests, however, was found to decrease for low signal-to-noise ratios, indicating that there are still deterministic features that are, in principle, indistinguishable from background noise.

The results from the climate-mode analysis suggest that the predictability of the AMO, PDO, and NAO is limited and that the multi-decadal variability of the AMO and PDO is the result of a stochastic process. The Niño 3.4 index, by contrast, was found to have deterministic features, implying that future states of ENSO may be predictable. Such predictability is important given that ENSO has regional- to global-scale impacts on precipitation and temperature. The ability to predict future changes of regional climate

thus, to some extent, depends on the ability to predict ENSO. However, currently, the future state of ENSO as determined by climate models is uncertain, with some climate models suggesting large changes and others indicating no change at all (Latif and Keenlyside, 2008).

A Matlab software package written by the author to implement the cumulative areawise test is available at justinschulte.com.

Chapter 4

Wavelet Analysis for Non-stationary, Non-linear Time Series

4.1. Introduction

Spectral analysis is a tool for extracting embedded structures in a time series. In particular, Fourier analysis has been used extensively by researchers for extracting deterministic structures from time series but is incapable of detecting nonstationary features often present in geophysical time series. Wavelet analysis can extract transient features embedded in time series, with a wavelet power spectrum representing variance (power) of a time series as a function of time and period. Since the seminal work of Torrence and Compo (1998), wavelet analysis has been applied extensively to geophysical time series such as the indices for the North Atlantic Oscillation (Olsen et al., 2012), Arctic Oscillation (Jevrejeva et al., 2003), Pacific Decadal Oscillation (Macdonald and Case, 2005; Newmann et al., 2003), El-Niño/Southern Oscillation (ENSO; Torrence and Webster, 1999), Pacific-North American Pattern, and West Pacific pattern (Gan et al., 2007). The application of wavelet coherence and cross-wavelet analyses (Grinsted et al., 2004), moreover, has proven useful in relating geophysical time series to other time series (Jevrejeva et al., 2003; Gan et al., 2007; Labat, 2010; Lee and Lwiza, 2008).

Many statistical methods, including power and cross-spectral analyses, rely on the assumption that the variable in question is Gaussian distributed (King, 1996). For a linear system in which the output is proportional to the input, the first- and second-order moments, the mean and variance, can fully describe the distribution of a process. In the frequency domain, by analogy, the variable can be fully described by the power spectrum, the decomposition of variance as a function of frequency. Suppose, however, that the distribution is non-Gaussian so that higher-order moments such as skewness and kurtosis exist. In this case, the mean and variance, while useful, are unable to fully describe the distribution in question. In a time series context, non-Gaussian distributions can arise from nonlinear systems, systems for which the output is no longer simply proportional to the input. For a nonlinear system, if the input is the sum of two sinusoids with different frequency components the output will contain additional frequency components

representing the sum and difference of the input frequencies (King, 1996). In such cases, it is necessary to examine the decomposition of higher-order moments in frequency space.

The frequency decomposition of the third-order moment, for example, results in a bispectrum or skewness function that measure deviations from Gaussianity (Nikias and Raghuveer, 1987; King, 1996). In fact, Hinich (1985) developed a bispectral test to determine if a time series is non-Gaussian and nonlinear. In some situations, higher-order nonlinearities such as cubic nonlinearities may exist, in which case the trispectrum or other polyspectra would have to be used (Collis et al., 1998).

Another advantage of higher-order spectral analysis is that the cycle geometry of oscillations, such as asymmetry with respect to a horizontal axis (skewed oscillation) or with respect to a vertical axis (asymmetric oscillation) can be quantified using the biphase. A pure sine wave, for example, is neither skewed nor asymmetric, whereas a time series resembling a saw-tooth is asymmetric. Skewed and asymmetric cycle geometry can identify, for example, abrupt climatic shifts, sudden shifts in the climate system that exceed the magnitude of the background variability (King, 1996). Abrupt climate shifts have occurred numerous times in the past and have dire impacts on ecological and economic systems (Alley et al., 2005). An understanding of past abrupt climate shifts is essential to understanding future climate change and so there is a need to quantify nonlinearities present in climatic oscillations.

The Quasi-biennial Oscillation (QBO), as another example, has been shown to behave nonlinearly, transitioning from easterly phases to westerly phases more rapidly than from westerly to easterly phases (Lu et al., 2009). Another source of asymmetry in the QBO time series arises from the westerly shear zone descending more regularly than the easterly shear zone. Asymmetries in the QBO time series are not well-captured by linear methods such as linear principal component and singular spectrum analyses (Lu et al., 2009) but are better captured using, for example, nonlinear principal component analysis (Hamilton and Hsieh, 2002). Another example of a nonlinear time series is the sunspot cycle, which undergoes an 11-year oscillation characterized by asymmetric cycle geometry, with solar maxima generally rising faster than they fall, indicating the presence of nonlinearities (Moussas et al., 2005; Rusu, 2007). ENSO, a climate phenomenon with regional- to global-

scale impacts, has also been shown to exhibit nonlinearities (Timmermann, 2003). The presence of nonlinearities and possible nonstationarities in the QBO, ENSO, and sunspot time series makes traditional Fourier and wavelet analysis inadequate for feature extraction, underscoring the need to develop methods for quantifying nonlinearities in a nonstationary geophysical setting.

The application of higher-order wavelet analysis has been rather limited compared to traditional wavelet analysis (van Millagan et al., 1995; Elsayed, 2006). One geophysical application of higher-order wavelet analysis is to oceanic waves (Elsayed, 2006), which was found to be capable of identifying nonlinearities in wind-wave interactions. However, the study lacked rigorous statistical significance testing, highlighting the need to develop significance testing methods for higher-order wavelet analysis to aid physical interpretation of results. The finite length of time series, as an example, would result in nonzero values of bicoherence even for a Gaussian process. Thus, one needs to assess the confidence with which the estimated bicoherence exceeds a noise background. However, the number of bicoherence estimates to which the statistical test is applied will be large and multiple artifacts will result. The multiple-testing problem was already identified for traditional wavelet analysis (Maraun et al., 2007; Schulte et al., 2015). There is therefore a need to apply statistical methods controlling false positive detection. It is also noted that the bicoherence spectra calculated are only sample estimates of the true bicoherence spectra. As a result, it is important to calculate confidence intervals corresponding to the sample estimates, which represent a range of plausible values for the sample estimates.

Another problem with the application higher-order wavelet analysis is selection of a time interval on which to calculate the high-order wavelet quantities. Such an approach is subjective and the result of the analysis may depend on the time interval chosen. The time interval selection problem highlights the need to develop a local bicoherence spectrum, which can be computed without choosing a time interval. Such an approach has already been adopted in wavelet coherence analysis (Grinsted et al., 2004).

Additionally, properties of the biphase have only been examined for Fourier-based bispectral analysis (Elgar and Sebert, 1989; Maccarone, 2013) and its usefulness in higher-order wavelet analysis has yet to be examined. For nonstationary time series, the biphase

and cycle geometry corresponding to the time series may change with time, underscoring the need to develop a local biphase spectrum.

In this paper, higher-order wavelet analysis is put in a statistical framework and applied to the QBO time series to demonstrate the insights afforded by the methods. Before describing higher-wavelet analysis, a brief overview of wavelet analysis is first presented in Sect. 4.2. Higher-order wavelet analysis is described Sect. 4.3 and a new local autobicoherence spectrum is introduced, eliminating the selection of a time interval on which to calculate nonlinearities properties of time series. The new and existing methods are applied to an ideal time series and the QBO index. In Section 4.4, a new procedure for estimating confidence intervals of global autobicoherence quantities is developed to estimate uncertainties in the sample autobicoherence spectra. The application of the new procedure to the sample autobicoherence spectrum of the QBO time series is then used to further assess confidence in results.

4.2. Wavelet Analysis

The idea behind wavelet analysis is to convolve a time series with a function satisfying certain conditions. Such functions are called wavelets, of which the most widely used is the Morlet wavelet, a sinusoid damped by a Gaussian envelope:

$$\psi_0(\eta) = \pi^{-1/4} e^{i\omega_0\eta} e^{-\frac{1}{2}\eta^2}, \tag{4.1}$$

where ψ_0 is the Morlet wavelet, ω_0 is the dimensionless frequency, and η is the dimensionless time (Torrence and Compo, 1998; Grinsted et al., 2004). In practical applications, the convolution of the wavelet function with a time series $X = (x_n; n = 1, \ldots, N)$ is calculated discretely using

$$W_n^X(s) = \sqrt{\frac{\delta t}{s}} \sum_{n'=1}^{N} x_{n'} \psi_0[(n' - n)\frac{\delta t}{s}], \tag{4.2}$$

where δt is a uniform time step, s is scale, $\eta = s \cdot t$, and $W_n^X(s)$ is the wavelet transform. The wavelet power is given by $|W_n^X(s)|^2$ (Torrence and Compo, 1998; Grinsted et al., 2004). For the Morlet wavelet with $\omega_0 = 6$, the wavelet scale and the Fourier period λ are

approximately equal ($\lambda = 1.03s$). A more detailed discussion of wavelet analysis can be found in Torrence and Compo (1998).

Shown in Fig. 4.1a is the time series of the QBO index and shown in Fig. 4.11b is the corresponding wavelet power spectrum. The QBO data from 1950-2013 were obtained from the Climate Prediction Center. The QBO index is defined as the zonal average of the 30 hPa zonal wind at the equator. As such, a positive index indicates westerly winds and a negative index indicates easterly winds. The most salient feature of the time series is the rather regular periodicity of approximately 28 months. Also note the asymmetry between the negative and positive phase, with the negative phases generally being stronger. The periodic behavior of the QBO was corroborated by examining the wavelet power spectrum. A well-defined 28-month periodicity is evident, with the associated wavelet power changing little throughout the study period.

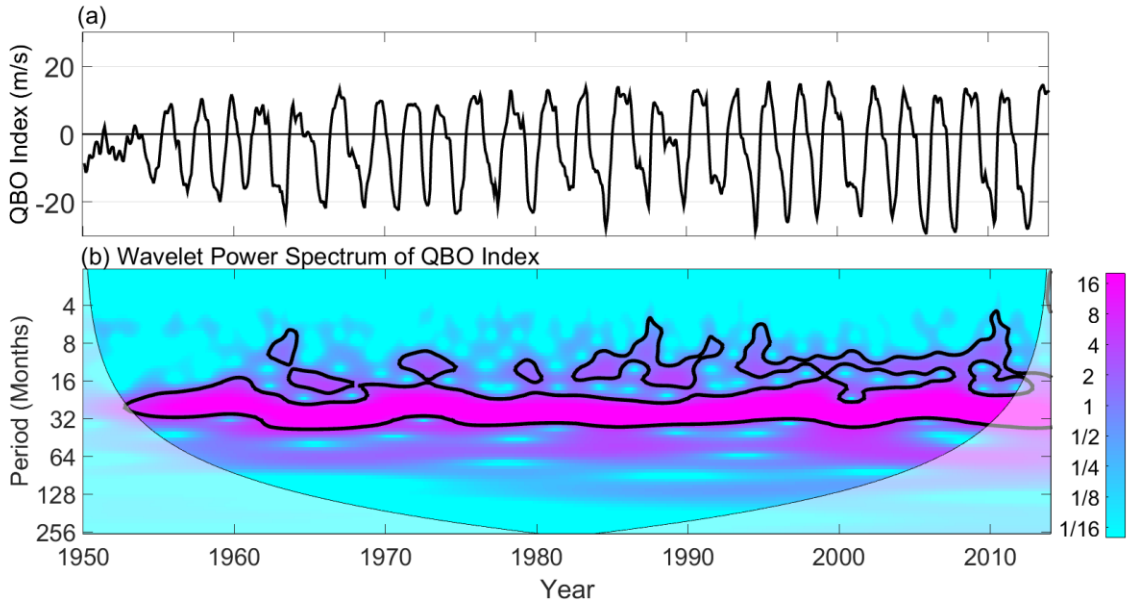

Fig. 4.1. (a) The QBO index and (b) the corresponding wavelet power spectrum. Contours enclose regions of 5% statistical pointwise significance (Torrence and Compo, 1998). Light shading represents the cone of influence, the region in which edge effects cannot be ignored.

There are also secondary features located at a period of approximately 14 months, primarily from 1985 to 2013. The appearance of significant power at a period of 14 months also coincides with most of the largest negative phases of the QBO. Such a correspondence may not have been a coincidence; the 14-month mode and the 28-month mode may have

interacted constructively to generate large negative events but interacted destructively to create smaller positive events. However, additional tools are needed to confirm if the periodicities are interacting and to understand how the interactions were related to the behavior of the QBO.

4.3. Higher-order Wavelet Analysis

4.3.1 Wavelet-based Autobicoherence

Higher-order spectral analysis provides the opportunity to quantify nonlinearities and allows the detection of interacting oscillatory modes within a time series. More specifically, nonlinearities are quantified using bicoherence, a tool for measuring quadratic nonlinearities, where quadratic nonlinearities imply that for frequencies f_1, f_2, and f_3 and corresponding phases ϕ_1, ϕ_2, and ϕ_3 the sum rules

$$f_1 + f_2 = f_3 \tag{4.3}$$

and

$$\phi_1 + \phi_2 = \phi_3 \tag{4.4}$$

are satisfied. Whereas Eq. (4.3) implies frequency coupling, Eq. (4.4) implies phase coupling. To see from where Eqs. (4.3) and (4.4) originate, let

$$X(t) = \sin(2\pi f_1 t + \phi_1) + \sin(2\pi f_2 t + \phi_2) \tag{4.5}$$

be the input into a system whose output is related to the input by

$$Y(t) = X(t) + \varepsilon X(t)^2 + w(t). \tag{4.6}$$

The multiplicative factor ε is used to represent the contribution of the nonlinear component of the signal and $w(t)$ is Gaussian white noise. Note that if $\varepsilon = 0$, then the system is linear because the output contains the same frequency components as the input. The substitution of Eq. (4.5) into Eq. (4.6) results in

$$\begin{aligned} Y(t) = \sin(2\pi f_1 t + \phi_1) + \sin(2\pi f_2 t + \phi_2) + \frac{\varepsilon}{2}[1 - \cos(2(2\pi f_1 t + \phi_1)) \\ - \cos(2(2\pi f_2 t + \phi_2)) + \cos(2\pi(f_2 - f_1)t + \phi_2 - \phi_1) \\ - \cos(2\pi(f_1 + f_2)t + \phi_1 + \phi_2)] + w(t) \end{aligned} \tag{4.7}$$

and thus the output has sinusoids with additional frequency components $2f_1$, $2f_2$, $f_2 - f_1$, and $f_2 + f_1$, which arise from the second term in right-hand side of Eq. (4.6).

Unlike the power spectrum, which is the Fourier transform of the second-order moment of a time series, the bispectrum is defined as the double Fourier transform of the third-order moment, or, more generally, the third-order cumulant, i.e.,

$$b_{xxx}(f_1, f_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C(t_1, t_2) e^{-i2\pi(f_1 t_1 + f_2 t_2)} dt_1 dt_2, \tag{4.8}$$

where *C* is the third-order cumulant, defined as

$$C(t_1, t_2) = M_3(t_1, t_2) + M_1[M_2(t_1) + M_2(t_2) + M_2(t_1 - t_2)] + 2M_1^3 \tag{4.9}$$

and the t_i are lags. If *X*(*t*) is zero-mean, then in Eq. (4.9), $M_1 = E[X(t)] = 0$ denotes the first-order moment (mean), $M_2 = E[X(t)X(t + t_1)]$ denotes the second-order moment (autocorrelation), and $M_3(t_1, t_2) = E[X(t)X(t + t_1)X(t + t_2)]$ denotes the third-order moment (Nidal and Malik, 2013). Also note that for a zero-mean process, the third-order cumulant reduces to the third-order moment (Collis et al., 1998). A more useful quantity is the normalized version of the bispectrum, the autobicoherence spectrum (Collis et al., 1998), which can be computed using the following:

$$b^2(f_1, f_2) = \frac{|b_{xxx}(f_1, f_2)|^2}{E\left[\left|X_f(f_1)X_f(f_2)\right|^2\right] E\left[X_f|(f_1 + f_2)|^2\right]}, \tag{4.10}$$

where $b^2(f_1, f_2)$ is bounded by 0 and 1 by the Schwarz inequality and X_f denotes the Fourier transform of *X*. $b^2(f_1, f_2)$ can be interpreted as the fraction of power at $f_1 + f_2$ due to quadratic phase coupling among f_1, f_2, and $f_1 + f_2$ such that the sum rule $f_1 + f_2 = f_3$ is satisfied (Elgar and Chandran, 1993). For a more in-depth discussion of higher-order spectral analysis the reader is referred to Nikias and Raghuveer (1987).

Phase information and cycle geometry can be obtained from the biphase, which is given by

$$\psi = tan^{-1}\left(\frac{Im(b_{xxx})}{Re(b_{xxx})}\right) = \phi_1 + \phi_2 - \phi_3. \quad (4.11)$$

It was noted by Maccarone (2013), however, that the biphase should be defined on the full 2π interval and thus in this paper the four-quadrant inverse tangent is computed and not the inverse tangent as shown above. By doing so, statistically significant autobicoherence detected together with the biphase can be used to quantify cycle geometry. A biphase of 0° indicates positive skewness and a biphase of 180° indicates negative skewness (Maccarone, 2013). An example of a skewed oscillation time series with biphase close to 0° is shown in Fig. 4.2a. Mathematically, the time series is written as

$$X(t) = \sum_{j=1}^{40} \frac{1}{j} \cos[0.1jt + a(j-1)], \quad (4.12)$$

where $a = 0$ (Maccarone, 2013). The time series is skewed because the positive spikes are not accompanied by negative spikes of equivalent magnitude and therefore the distribution of the time series would be positively skewed, with the right tail being larger than the left tail.

For asymmetric waveforms, a biphase of 90° indicates that the time series is linearly rising but rapidly falling as shown in Fig. 4.3, whereas a biphase of -90° indicates that the time series rises rapidly and falls linearly. A purely asymmetric time series will have a biphase of 90° or -90°, as shown in Fig. 4.3, where the saw-toothed time series obtained by setting $a = \pi/2$ in Eq. (4.12) rises more slowly than it falls. In a physical setting, asymmetric cycle geometry implies that phase transitions occur at different rates, as observed in the QBO time series.

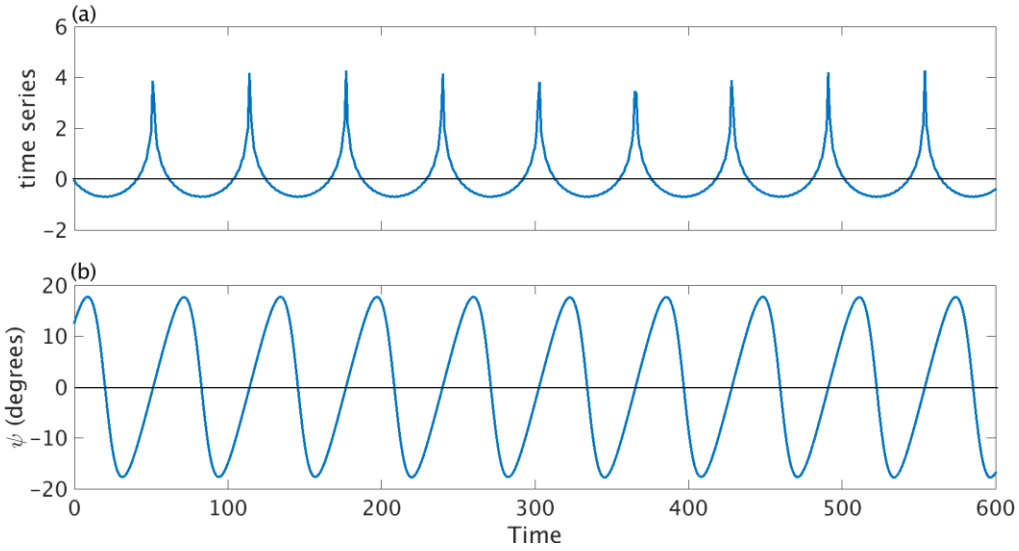

Figure 4.2. (a) a skewed time series and (b) its corresponding local biphase. The biphase close to zero indicates a nonlinear interaction resulting in a skewed oscillation. The biphase was calculated from the first three cosines in the summation described in the text. The large deviations from zero at the edges are the result of edge effects.

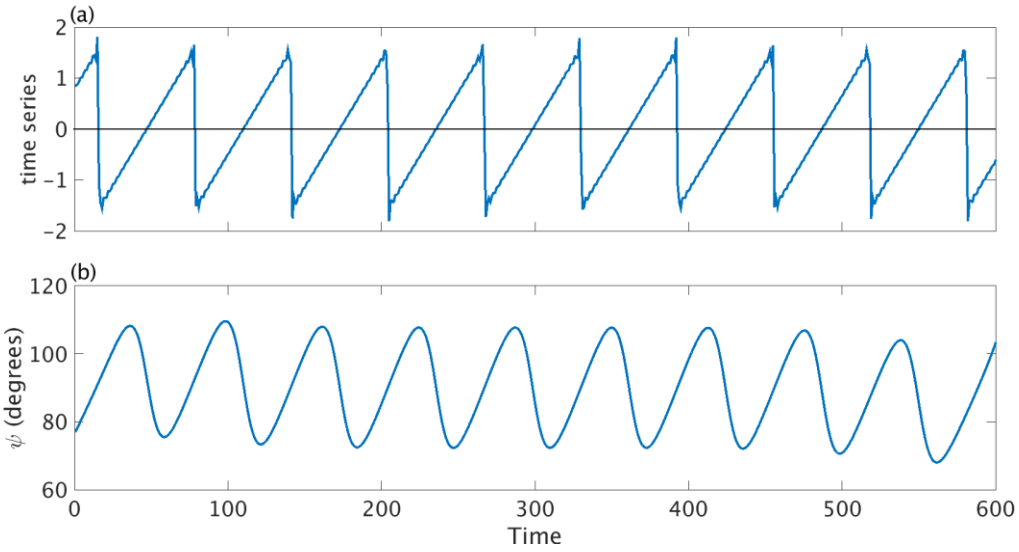

Figure 4.3. (a) A saw-toothed time series and (b) its corresponding local biphase. The biphase close to 90° indicates a nonlinear interaction resulting in an asymmetric waveform. The biphase was calculated from the first three cosines in the summation.

According to Elsayed (2006), the wavelet-based autobicoherence is defined as

$$b^{w}_{xxx}(s_1, s_2) = \frac{|B^{w}_{xxx}(s_1,s_2)|^2}{\left(\int_T |W_x(s_1,t)W_x(s_2,t)|^2 dt\right)\left(\int_T |W_x(s,t)|^2 dt\right)}, \tag{4.13}$$

where

$$B^{w}_{xxx}(s_1, s_2) = \int_T W^{*}_{x}(s,t)\, W_x(s_1,t) W_x(s_2,t) dt\ , \tag{4.14}$$

$$\frac{1}{s_1} + \frac{1}{s_2} = \frac{1}{s}, \tag{4.15}$$

T is a time interval, $W_x(s,t)$ is the wavelet transform of a time series X at scale s and time t, and $W^{*}_{x}(s,t)$ denotes the complex conjugate of $W_x(s,t)$. The wavelet-based autobicoherence measures the degree of quadratic phase coupling, where a peak at (s_1, s_2) indicates an nonlinear interaction between the scale components s_1, s_2, and s.

In practice, the autobicoherence is computed discretely so that Eq. (4.13) can be written as

$$\overline{W_b}(s_1, s_2) = \frac{|B^{w}_{xxx}(s_1,s_2)|^2}{\left(\sum_{n=n_1}^{n_2} |W^{X}_{n}(s_1) W^{X}_{n}(s_2)|^2\right)\left(\sum_{n=n_1}^{n_2} |W^{X}_{n}(s)|^2\right)}, \tag{4.16}$$

where

$$B^{w}_{xxx}(s_1, s_2) = \sum_{n=n_1}^{n_2} W^{*X}_{n}(s) W^{X}_{n}(s_1) W^{X}_{n}(s_2)$$

$$= \sum_{n=n_1}^{n_2} B^{w}_{n}(s_1, s_2), \tag{4.17}$$

$n_1 \geq 1$, and $n_2 \leq N$. Note that if $n_1 = 1$ and $n_2 = N$, then Eq. (4.15) represents the global autobicoherence spectrum.

The Monte Carlo approach to pointwise significance testing is adopted in this paper and is similar to that used in wavelet coherence (Grinsted et al., 2014). To estimate the significance of wavelet-based autobicoherence at each point (s_1, s_2), Monte Carlo methods are used to (1) generate a large ensemble of red-noise processes with the same lengths and lag-1 autocorrelation coefficients as the input time series and (2) compute for each randomly generated red-noise process the autobicoherence spectrum. From the ensemble of autobicoherence spectra, the p = 100(1 - α_p) percentile of the autobicoherence estimates is computed for every point (s_1, s_2), where p corresponds to the critical level of the test

and α_p is the pointwise significance of the test. Given the symmetry of the autobicoherence spectrum, the critical level of the test can be computed using only half of the autobicoherence estimates, reducing computational costs.

4.3.2 Multiple Testing

Let α_p be the significance level of the pointwise significance test as described above and let *K* denote the number of autobicoherence estimates being tested, then there will be on average $\alpha_p K$ false positive results. A similar problem occurs in traditional wavelet analysis (Maraun et al., 2007; Schulte et al., 2015). In the case of simultaneously testing multiple hypotheses, the number of false positive results can be reduced by applying, for example, the Bonoferonni correction (Lehmann, 1986). However, this simple correction often results in many true positives being rejected and is especially permissive in the case of autocorrelated data (Maraun et al., 2004). Other procedures also exist, including the Walker *p*-value adjustment procedure, which has more statistical power than the Bonferonni correction. An even more powerful method is the Benjamini and Hochberg (1995) procedure, which controls the false discovery rate (FDR), where the FDR is the expected proportion of the false rejections that are actually true. An advantage of this method, in addition to its statistical power, is that it takes into account the confidence with which local hypotheses are rejected and is robust even in the case of autocorrelated data (Wilks, 2002). Benjamini and Yekutieli (2001) developed a modified version of the Benjmini and Hochberg (1995) procedure that works for any dependency structure among the local test statistics and thus this procedure will be used in this paper to control the FDR.

The procedure can be described as follows: Suppose that *K* local hypotheses were tested. Let $p_{(i)}$ denote the smallest of the *K* local *p*-values, then, under the assumption that the *K* local tests are independent, the FDR can be controlled at the *q*-level by rejecting those local tests for which $p_{(i)}$ is no greater than

$$p_{FDR=} \max_{j=1,\dots,k} \left[p_{(j)} : p_{(j)} \leq q(j/K)\right]$$

$$= \max_{j=1,\dots,k} \left[p_{(j)} : p_{(j)} \leq \alpha_{global}(j/K)\right] \tag{4.18}$$

so that the FDR level is equivalent to the global test level. For a local *p*-value to be deemed significant using this procedure, it must be less than or equal to the largest *p*-value for which Eq. (4.18) is satisfied. If no such local *p*-values exist, then none are deemed insignificant, and, therefore, the global test hypothesis cannot be rejected. If the test statistics have an unknown dependency structure, q can be replaced with $q/\sum_{i=1}^{K}\frac{1}{i}$, though this substitution makes the procedure less powerful (Reiner et al., 2002). This modified method will be applied to autobicoherence spectra at the 0.05 level throughout this paper.

4.3.3 Wavelet-based Autobicoherence of an Idealized Time Series

To demonstrate the features of a time series that can be extracted using higher-order wavelet analysis, an idealized nonstationary time series will first be considered. Consider the quadratically nonlinear time series

$$X(t) = \cos(2\pi f t + \phi) + \gamma(\mathrm{t})\cos(4\pi f t + 2\phi) + w(t), \tag{4.19}$$

where *f* is frequency, $w(t)$ is Gaussian white noise, and $\gamma(t)$ is a time-dependent nonlinear coefficient given by

$$\gamma(t) = 0.001t. \tag{4.20}$$

Note that Eqs. (4.3) and (4.4) are satisfied because $f_1 + f_2 = 2f_1 = 2f_2$ and similarly for ϕ. The sinusoid with frequency $2f_1$ is said to be the harmonic of the primary frequency component with frequency f_2, where the amplitude of the harmonic depends on $\gamma(t)$, the strength of the quadratic nonlinearity. $X(t)$ and the corresponding wavelet power spectrum for the case when $f_1 = 0.03$ is shown in Fig. 4.4. The Gaussian white noise was set to 1 decibels. The primary frequency component results in a large region of 5% pointwise significance at $\lambda = 30$, whereas its harmonic only results in a few small significance regions located from $t = 700$ to $t = 1000$. It also noted that the appearance of the significance power at $\lambda = 15$ from $t = 700$ to $t = 1000$ is accompanied by large positive spikes in the time series that result in the time series favoring positive values. Prior to the emergence of the significant power at $\lambda = 15$, the time series varied smoothly in the sense that negative phases were accompanied by positive phases of similar amplitude.

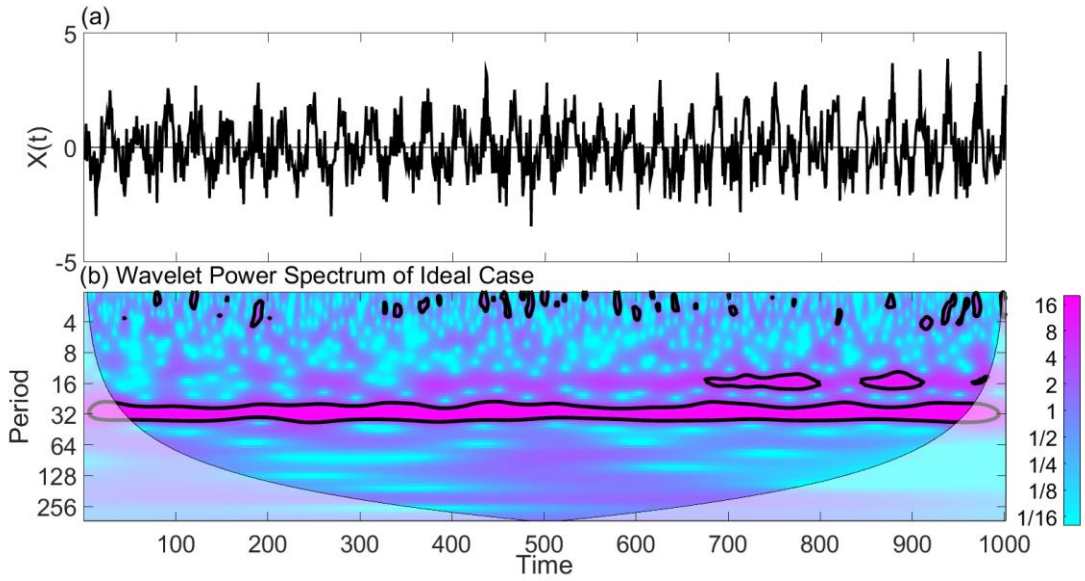

Figure 4.4. (a) Time series corresponding to Eq. (4.19). (b) Corresponding wavelet power spectrum.

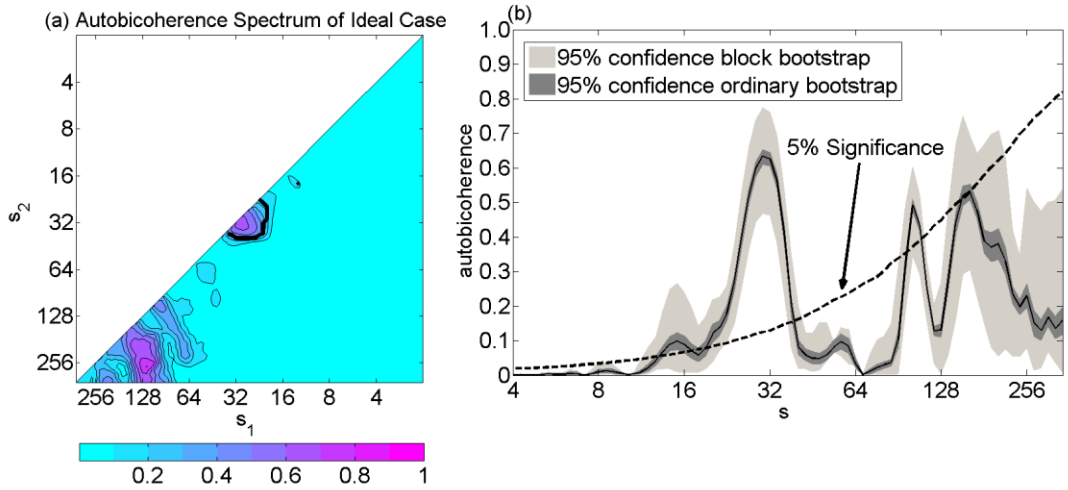

Figure. 4.5. (a) Wavelet-based autobicoherence spectrum of the ideal time series. Thick contours enclose regions of 5% pointwise significance after controling the FDR. The diagonal line separates the spectrum into two symmetric regions. (b) The diagonal slice of the autobicoherence spectrum at $s_1 = s_2 = s$. The critical level for the test represented by the dotted line was calculated using Monte Carlo methods.

To determine if the oscillations are quadratically interacting, the autobicoherence of $X(t)$ was computed (Fig. 4.5). The significant peak centered at (30, 30) indicates that

an oscillation with period 30 is phase-coupled to an oscillation with $\lambda = 15$. The result implies that the variability at $\lambda = 15$ is partially due to the interaction between the two modes. The fraction of variability is determined by the autobicoherence value corresponding to the significant peak. In the present case, $\overline{W_b}(s_1, s_2) = 0.5$ so about half of the variability at $\lambda = 15$ is due to the nonlinear interaction. Note that no other peaks were found to be significant.

4.3.4 Wavelet-based Autobicoherence of Geophysical Time Series

Shown in Fig. 4.6 is the wavelet-based autobicoherence spectrum for the QBO time series. A large region of significance was identified, which contained the local maximum at (28, 28) months. The peak represents the phase coupling of the primary frequency component with its harmonic with a period of 14 months. The power at $\lambda = 14$ months therefore partially resulted from the interaction between its primary frequency component and its harmonic. The significance and magnitude of the autobicoherence in the QBO spectrum is consistent with how the QBO does not vary smoothly, shifting to the easterly phase more quickly than to the westerly phase and with the westerly phase tending to be stronger than the easterly phase. The asymmetry in both phase transition and magnitude are suggestive of nonlinearities.

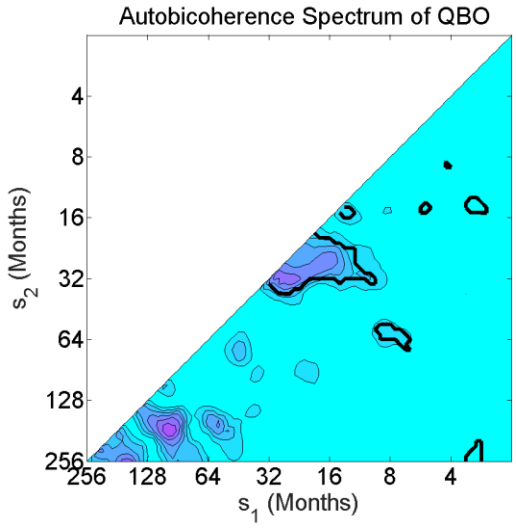

Figure. 4.6. The wavelet-based autobicoherence spectrum of the QBO index for the period 1950-2013. Thick contours enclose regions of 5% pointwise significance.

4.3.5 Local Wavelet Autobicoherence

It may also be desirable to see how autobicoherence along slices of the full autobicoherence spectrum changes with time. To compute local autobicoherence, apply a smoothing operator $S(W) = S_{scale}\left(S_{time}\left(W_n^X(s)\right)\right)$ (Grinsted et al., 2004) to each term in Eq. (4.13) instead of summing in time, i.e.,

$$b_n^w(s_1, s_2) = \frac{\left|S({s_1}^{-1} B_n^w(s_1, s_2))\right|^2}{S({s_1}^{-1}|W_n^X(s_1) W_n^X(s_2)|^2) \cdot S(s^{-1}|W_n^X(s)|^2)}. \tag{4.20}$$

The smoothing operator for the Morlet wavelet is given by

$$S_{time}(W)|_s = \left(W_n^X(s) * {c_1}^{\frac{-t^2}{2s^2}}\right)|_s \tag{4.21}$$

and

$$S_{scale}(W)|_n = \left(W_n^X(s) * c_2 \Pi(.6s)\right)|_n, \tag{4.22}$$

where c_1 and c_2 are normalization constants determined numerically and Π is the rectangular function.

It is important to mention that the numerator of Eq. (4.20) contains a term with wavelet coefficients at two different scales so that the choice of smoothing is not as straightforward as for wavelet coherence. Smoothing autobicoherence estimates with respect to $s_{min} = \min(s_1, s_2)$ was found to result in larger autobicoherence estimates, whereas smoothing the autobicoherence with respect to $s_{max} = \max(s_1, s_2)$ resulted in smaller autobicoherence estimates. Given that the autobicoherence estimates are influenced by the choice of smoothing, it is inevitable that the significance of the autobicoherence estimates is also impacted. In particular, smoothing the autobicoherence spectrum with respect to s_{max} allowed extrema to be smoothed out, eliminating spuriously large autobicoherence. For this reason, all local autobicoherence spectra in this paper will be computed by smoothing with respect to s_{max}.

The advantage of using Eq. (4.20) is that transient quadratic nonlinearities can now be detected and the need for choosing an integration time interval has been eliminated. If $s_1 = s_2$, then $(t, s_1, s_1) = (t, s_2, s_2) = (t, s)$ and thus, in the case of this diagonal slice, the local wavelet-based bicoherence spectrum is a two-dimensional representation of the

degree of local quadratic nonlinearity. The vertical axis corresponds to the primary frequency and the horizontal axis corresponds to time. As a concrete example, a peak at (64, 64) would indicate that at time index $t = 50$ the oscillation with a fundamental period $\lambda = 1.03s \approx 64$ is locally coupled to an oscillation with period $\lambda \approx 32$.

One can also compute a local biphase from the smoothed bispectrum by taking the four quadrant inverse tangent of the smoothed imaginary part divided by the smoothed real part. The local biphase, for example, was computed for the skewed time series shown in Fig. 4.2a. As expected, the biphase fluctuates regularly around 0° and the mean is 2°. The local biphase for the saw-toothed time series is shown in Fig. 4.3b. The biphase fluctuates about 90° and the mean biphase is 90° as expected.

The procedure for the estimation of the statistical significance of local autobicoherence is the following: generate red-noise time series with the same lag-1 autocorrelation coefficients as the input time series and use the local autobicoherence estimates outside the COI to generate a null distribution of $b_n^w(s_1, s_2)$. Note that the calculation only needs to be performed at a fixed time outside of the COI because red-noise is a stationary process, which produces a stationary background spectrum.

4.3.6 Local Wavelet-based Autobicoherence of an Idealized Time Series

The local autobicoherence spectrum of $X(t)$ for (30, 30) is shown in Fig. 4.6b. Initially, there is no local autobicoherence that exceeds the 5% significance level. At *t* = 250 and *t* = 500, on the other hand, small regions of 5% significant autobicoherence emerge, indicating a transient nonlinear interaction. At *t* = 500 the nonlinearity is strong and results in a large region of significant local autobicoherence extending from *t* = 500 to

the edge of the wavelet domain.

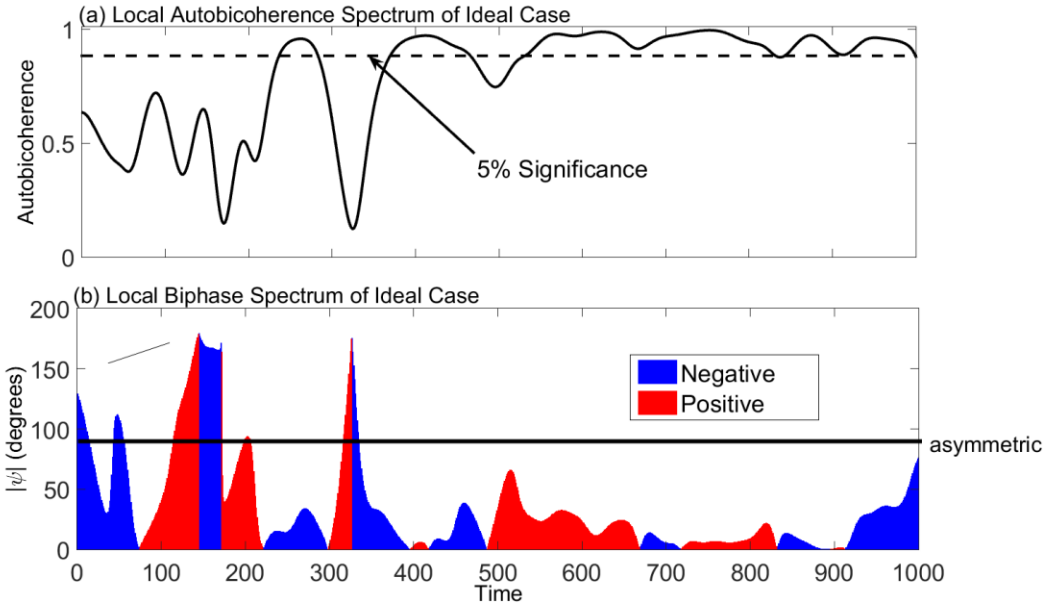

Figure 4.7. (a) The local autobicoherence and (b) local biphase corresponding to (30, 30) in the full autobicoherence spectrum. Biphases differing from 90° indicate that the nonlinear interaction resulted in a waveform with skewness.

In order to determine if the peaks in autobicoherence are associated with a quadratic nonlinearity, it is important to compute the biphase, which is shown in Fig. 4.7b. From $t = 0$ to $t = 400$ there is an unstable phase relationships between the phase of the primary frequency component and its harmonic. Such a lack of phase coherence indicates a weak nonlinear interaction, which is consistent with how the autobicoherence is lower before $t = 400$. In contrast, after $t = 400$, the biphase becomes stable, changing little with time, indicating a consistent phase relationship between the primary frequency mode and its harmonic. It also noted that the biphase during this time fluctuates near 0°, which implies that the phase relationships arise from a quadratic nonlinearity. The near zero biphase is consistent with how $X(t)$ was constructed from the sum of two cosines with zero phase and also suggests that the interaction results in skewed cycle geometry, where positive values of the time series are preferred. Indeed, by inspection of Fig. 4.4a the oscillations initially appear to be sinusoidal, varying smoothly, whereas after $t = 400$ spikes begin to appear and $X(t)$ favors positive values.

4.3.7 Local Wavelet-based Autobicoherence of the QBO Time Series

The local autobicoherence spectrum of the QBO index at the point (28, 28) in the full autobicoherence spectrum is shown in Fig. 4.8. From 1950 to 1970 the magnitude of the autobicoherence fluctuated and consisted of one local significant peak at 1965. Significant autobicoherence was also found from 1975 to 1998, contrasting with the autobicoherence after 1998, which was not found to be significant until 2010.

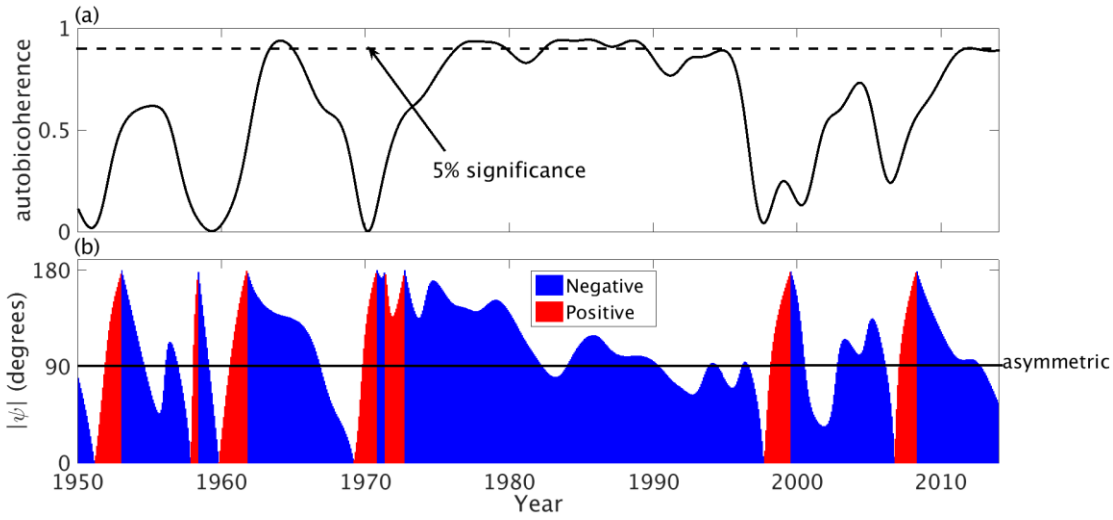

Figure. 4.8. Same as Fig. 4.7 except at (28, 28) in the autobicoherence spectrum of the QBO index Biphases differing from 90° indicate that the nonlinear interaction resulted in a waveform with skewness.

To determine if the peaks indicated in the autobicoherence are associated with a quadratic nonlinearity, the local biphase was computed. Fig. 4.8a shows the local biphase for the autobicoherence peak at (28, 28). For most of the study period, the biphase was found to vary considerably, particularly during the 1950-1970 and 1995-2013 periods. On the other hand, the biphase varied smoothly from 1970 to 1995, consistent with how the autobicoherence during that period was large and stable (Fig. 4.8a). Also, during that period the biphase was nonzero; in fact, the mean biphase during the period was -100°, suggesting that the phase coupling is not the result of a quadratic interaction. A biphase of -100° indicated asymmetric geometry, which physically represents how phase transitions of the QBO occurred at different rates. Recall that it has already been discussed in the introduction that the QBO transitions from easterly phases to westerly phases more rapidly than from westerly to easterly phases (Lu et al., 2009). Another interesting feature is the general increase in the biphase from 1970 to 1995. In the beginning of the time period, the biphase was -180° and after 1980 the biphase switched to -90°.

The local autobicoherence and biphase corresponding to the peak (16, 26) was also computed (Fig. 4.9). The mean of the absolute value of the biphase for the period 1950-2013 was 130°, indicating that the interaction among the modes with periods 10, 16, 26 resulted in skewed waveforms. In fact, because the biphases were close to 180° the waveforms should have been skewed to negative values (Maccarone, 2013) and such skewness is evident by inspecting Fig. 4.1. Also note that some of the largest negative phases of the QBO occurred from 1995 to 2010, which coincided with the period of most significant autobicoherence as shown in Fig. 4.9a.

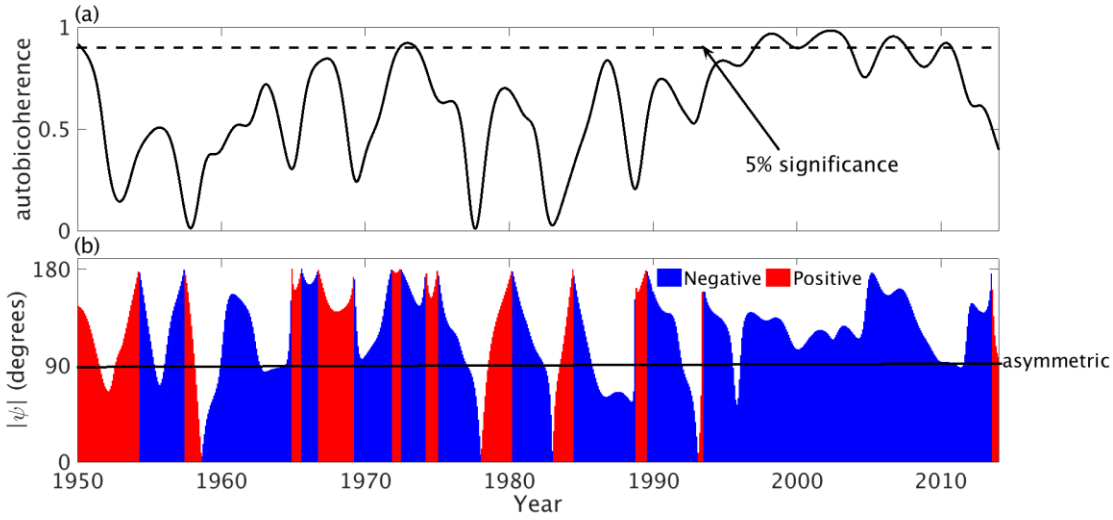

Figure. 4.9. Same as Fig. 4.8 except at the point (16, 26).

4.4. Block Bootstrapping Methods

4.4.1 Block Bootstrapping Autobicoherence

Bootstrapping is a widely used technique to estimate the variance or uncertainty of a sample estimate. For independent data one samples with replacement individual data points (Efron, 1979); for dependent data one must sample with replacement blocks of data to preserve the autocorrelation structure of the data (Kunsch, 1989). The latter technique is called block bootstrapping and should be used for variance estimation of global wavelet quantities, as wavelet coefficients are known to be autocorrelated in both time and scale. The use of traditional bootstrapping techniques would result in confidence intervals that are too narrow. It is expected, however, that the choice of the bootstrapping technique is

more critical at larger scales, as the decorrelation length of the mother wavelet increases with scale.

A brief overview of the procedure is provided below but a more detailed discussion can be found in Schulte et al. (2015). To find the approximate $100(1-\beta)\%$ confidence interval of an autobicoherence estimate, divide the set of wavelet coefficients at each scale into over-lapping blocks. The lengths of the blocks at each scale should be the same and the randomly resampled blocks chosen should be the same at each scale to avoid randomizing the data. The concentration of the blocks then results in a synthetic set of wavelet coefficients at each scale. The synthetic set of wavelet coefficients can then be used to calculate a bootstrap replicate of the autobicoherence. The iteration of the procedure 1000 times results in a distribution of bootstrap replicates from which a 95% confidence interval can be obtained.

As noted by Schulte et al. (2015), the appropriate block length to use can be determined by Monte Carlo methods. Using Monte Carlo methods, it was determined that a block length of $N^{0.6}$ was found to produce accurate confidence bounds while also producing the widest confidence intervals at all scales.

4.4.2 Application to Ideal and Climatic Time Series

Figure 4.5b shows the application of the block bootstrap procedure to the diagonal slice $s_1 = s_2 = s$ of the autobicoherence for the ideal case. The 95% confidence intervals were also obtained using the ordinary bootstrap. A pronounced peak at $s = 30$ was identified and represents the interaction between the primary frequency and its harmonic. By inspection of Fig. 4.5b, there is a clear difference between the widths of the confidence intervals obtained from the two bootstrapping procedures. For the ordinary bootstrap, the confidence intervals are narrow and the width of the confidence intervals appear to be only weakly dependent on scale. On the other hand, the confidence intervals obtained using the block bootstrap procedure are wide, especially at large scales, and the width of the confidence intervals depends strongly on scale, increasing from small scales to large scales. It is also noted that, whereas the block bootstrap procedure has deemed no spurious peaks as significant, the ordinary bootstrap procedure deemed two the spurious peaks at s = 14 and s = 100 as significant. The implementation of the block bootstrap procedure can

therefore enhance confidence in results, facilitating the investigation of a deeper physical understanding.

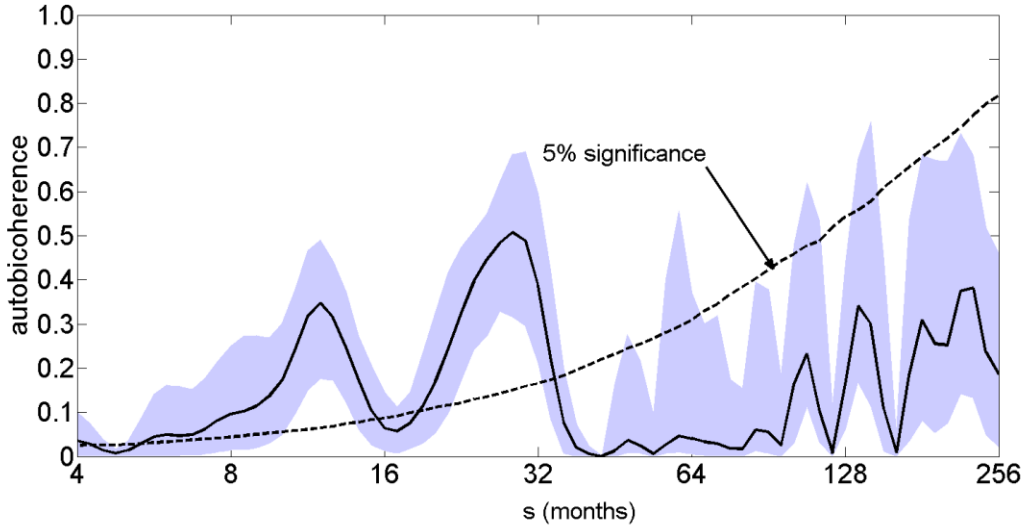

Figure. 4.10. Same as Fig. 4.5b except for the QBO index for the period 1950-2013.

The application of the block bootstrap procedure to the diagonal slice $s_1 = s_2 = s$ of the full autobicoherence spectrum of the QBO index is shown in Fig 4.10. The 95% confidence intervals corresponding to the peaks (14, 14) and (28, 28) do not cross the 5% significance bound and thus one has more confidence that those peaks are significant. All other peaks have been deemed insignificant.

4.5. Summary

Higher-order wavelet analysis together with significance testing procedures were used to detect nonlinearities embedded in an ideal time series and the QBO time series. The autobicoherence spectrum of the QBO index revealed phase coupling of the 28 month mode with a higher frequency mode with period 14 months. A local autobicoherence spectrum of the QBO index showed that the strength of the nonlinearities varied temporally. Furthermore, the local biphase spectrum indicated that the nonlinear interaction resulted in waveforms that were both skewed and asymmetric, indicating that the strength of negative QBO events were stronger than positive events, and that transitions between events occurred at different rates.

Chapter 5

The influence of Climate Modes on Streamflow in the Mid-Atlantic Region of the United States

5.1. Introduction

The Susquehanna, Delaware, and Hudson River Basins (SRB, DRB, and HRB) drain to three important estuaries of the Mid-Atlantic region of the United States (US, Figure 5.1), which have experienced substantial climate change and are likely to continue do so with increases in greenhouse gas concentrations (Najjar et al., 2009). This projected climate change is likely to render more difficult efforts to restore these estuaries, which have been stressed by anthropogenic activities, including pollution (e.g., toxic metals, polychlorinated biphenyls, and excess nitrogen), dredging, conversion of wetland habitats, and overfishing (Najjar et al., 2010; Kreeger et al., 2010; Steinberg et al., 2004).

Climate change is likely to manifest itself through changes in existing climate modes, which are recurring and often oscillatory patterns of climate variables, such as sea level pressure (SLP) and sea surface temperature (SST), that operate on timescales ranging from weekly to multi-decadal. For example, Ning et al. (2012) found increases in projected wintertime precipitation in the Northeast US to be consistent with a projected decrease in the positive phase of one prominent climate mode, the North Atlantic Oscillation (NAO). The NAO and other important climate modes, such as the Atlantic Multi-decadal Oscillation (AMO), the Pacific Decadal Oscillation (PDO), El Niño-Southern Oscillation (ENSO), the Pacific Northern-American Teleconnection (PNA), and the North Pacific Oscillation (NPO), have regional- to global-scale impacts on climate and weather (Philander, 1983; Trenberth and Hurrell, 1994; Mantua et al., 1997; Thompson and Wallace, 1998; Hurrell et al., 2003).

An understanding of the historical impacts of climate modes on regional climate variability can enhance our understanding of future changes in that region. Furthermore, an understanding of climate-mode impacts on regional meteorological, hydrological, and ecological characteristics will improve monthly and seasonal forecasts, which are of economic importance. With that in mind, the goal of this study is to analyze the variability in streamflow of the Susquehanna, Delaware, and Hudson Rivers, three large rivers in the Mid-Atlantic region of the US, in the context of climate modes.

There are relatively few studies on the impacts of climate modes on streamflow variability in the Mid-Atlantic region as compared to precipitation-climate mode and temperature-climate mode studies. Dettinger and Diaz (2000) found associations between the Dec-Feb Southern Oscillation (SO) and Oct-Sept streamflow across the Northeast, where El Nino years are associated with wetter-than-normal conditions. Futhuremore, Xu et al. (2012) identified relationships between North Pacific SSTs and Northeast US streamflow and Barlow et al. (2001) found North Pacific SSTs to have been possible drivers of the 1960s drought, a major hydrometeorological event that strained water-management agreements between New York City and Philadelphia (USDA, 2000). The relationships may be the result of prevailing synoptic regimes that set up during certain NAO and ENSO phases (Miller et al., 2006). Barlow et al. (2001) and Miller et al. (2006), however, only considered climate mode-streamflow simultaneous relationships. It is not clear if relationships hold on longer timescales. Labat (2008) and Whitney (2010) investigated streamflow variability across the Mid-Atlantic Bight and found multi-decadal variability in the flows of many rivers in the region, which Whitney (2010) hypothesized was related to the NAO.

The streamflow-climate mode relationships discussed above are associated with changes in precipitation, temperature, snow cover, and evapotranspiration, all of which have been investigated in the context of climate modes. Leathers et al. (1991) found that positive PNA phases are accompanied by colder and drier conditions across the US on monthly timescales. Serreze et al. (1998) noted increased snowfall in the Mid-Atlantic region during positive phases of the PNA, which was found to be associated with below-normal maximum temperature on precipitation days. Barlow et al. (2001) noted ENSO,

NPO, and PDO influences on Northeast US precipitation and drought conditions. Pattern et al. (2003) found associations between ENSO and winter snowfall across the Northeast US, with El Niño years being accompanied by more frequent major snow events. Archambault et al. (2007) found cool seasons under positive NAO and negative PNA regimes to be wettest. Eichler and Higgins (2005) found increased spring precipitation during El Niño years as a result of more frequent East Coast storms. Similarly, Seager et al. (2010) found strong winter snowfall-NAO and snowfall-ENSO linkages, both of which were related to the frequency of East Coast storms. Despite the impact of ENSO on Northeast precipitation, previous research did not relate ENSO to the 1960s drought in the Northeast US; in fact, it has been argued that the drought (and the subsequent wet period that continues to the present) resulted from internal atmospheric variability because global climate models with prescribed SSTs did not reproduce the drought (Seager et al., 2012).

This paper presents a comprehensive investigation into climate-mode impacts on Mid-Atlantic streamflow by considering all timescales ranging from months to decades. To understand the proximate forcing of Mid-Atlantic streamflow, data sets on mean watershed temperature and precipitation are analyzed as well. Relationships among streamflow, precipitation, temperature, and climate indices are investigated first through a standard linear correlation analysis at monthly, seasonal, and annual timescales. A more general understanding of the variability of Mid-Atlantic streamflow and its linkages to the proximate climate and climate modes is obtained via wavelet analysis and wavelet coherence analysis.

The main advantage of wavelet analysis is that it can decompose a complex time series, such as streamflow, into a two-dimensional (time and frequency) representation, from which inferences about the time series can be drawn. Wavelet analysis can, in particular, detect important features embedded in the time series, such as modes of dominant variability and their temporal behavior, which may be linked to some physical mechanisms, facilitating further scientific investigation. The previously noted studies by Labat (2008) and Whitney (2010) detected decadal streamflow variability but did not quantitatively link it to physical mechanisms, underscoring the need to evaluate the relationship between multi-decadal streamflow variability and large-scale climate indices.

The relationship between streamflow and a climate index at a variety of timescales can be quantified using wavelet coherence analysis, which has the advantage of minimizing aliasing while also eliminating the choice of averaging window or filter. Climate modes are typically most energetic at certain periods so that climate mode-streamflow relationships may be used to identify particular timescales. The PDO, as an example, is most energetic at periods of 15-25 years and 50-70 years and therefore one might expect its influence to be strongest at those timescales (Mantua, 2010).

Another advantage of wavelet coherence analysis is that phase relationships at a given timescale can be quantified as a function of time. The goal of a traditional cross-correlation analysis is to determine the time delay for which the association between two time series is greatest, but one cannot determine how the relationship changes over time. In a wavelet coherence analysis, on the other hand, phase relationships are calculated such that the degree to which two time series are positively or negatively related can be measured as both a function of time and period. Such a decomposition is important in this study because the temporal variations of ENSO teleconnections result from differences in the atmospheric basic state during ENSO events and are modulated by other teleconnection patterns (Gershunov and Barnett, 1998).

The remainder of the paper is organized as follows. Section 5.2 describes the data sets used (streamflow, temperature, precipitation, and climate indices) and the methods employed (correlation and wavelet analysis). Section 5.3 first provides a brief description of Mid-Atlantic streamflow and precipitation variability, and then presents the main results: the outcomes of the linear correlation and wavelet analyses. Section 5.3 also presents an analysis of climate-mode contributions to decadal-scale Mid-Atlantic streamflow anomalies. Section 5.4 provides a brief discussion of the possible teleconnection patterns that link climate modes to Mid-Atlantic hydrology and Section 5.5 summarizes the analysis with some concluding remarks.

5.2. Data and Methods

5.2.1 Streamflow, Temperature, and Precipitation

Mean monthly streamflow data were obtained from the United States Geological Survey (USGS) for the Delaware, Hudson, and Susquehanna Rivers at the Trenton, Waterford, and Harrisburg gauging stations, respectively (Figures 5.1a and 5.1b), which have corresponding USGS station numbers of 01463500, 01335754, and 01570500. These gauging stations were chosen because of their relatively large drainage areas and long records. The period of record used in this study is 1900-2010 for the Susquehanna and Hudson Rivers, and 1913-2010 for the Delaware River. These records are complete except for Waterford, which has a few missing months in 1976 (0.9% of the record). These gaps were filled by linearly regressing streamflow from the nearby Green Island gauging station (USGS station number 01358000) with Waterford streamflow from 1950-2010 (Pearson correlation coefficient equals 0.99). All streamflow data were converted to anomalies by subtracting the climatological mean monthly value for each month from the monthly values, thereby removing the annual cycle.

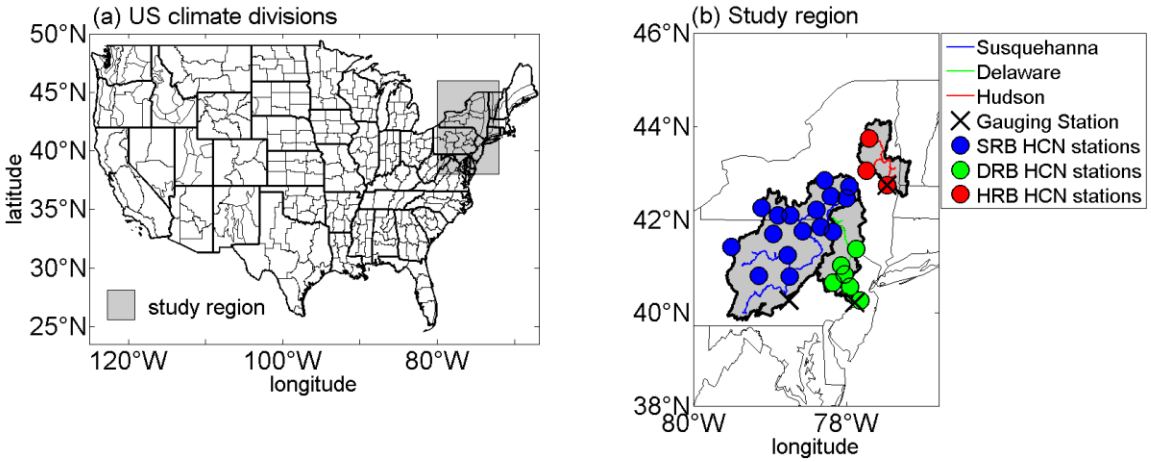

Figure 5.1. (a) Location of the US climate divisions delimited with thin black lines. State boundaries are thick black lines and the study region is indicated by the gray box. (b) Location of Harrisburg, Trenton, and Waterford gauging stations with corresponding rivers and Historical Climate Network stations. Thick black lines represent the boundaries of the Harrisburg, Trenton, and Waterford drainage basins and thin lines represent state boundaries.

Mean monthly maximum temperature and mean monthly precipitation data for 1900-2010 were obtained from the Historical Climate Network (HCN; Menne et al., 2009); station locations are shown in Figure 5.1. The reason for using maximum temperature data

is that precipitation type in the Northeast US is impacted by mean maximum temperature on precipitation days (Serreze et al., 1998). Compared to other precipitation and temperature data sets, the HCN data are of relatively high quality and long record length, which facilitates analysis on long timescales. To emphasize the basin-wide impacts of climate modes, the station-based precipitation and temperature data were averaged without any weighting to obtain a time series for each of the three river basins upstream of the gauging station. Like streamflow, the precipitation and temperature time series were converted to anomalies.

A few stations used in the averaging are located just outside the drainage basins. These stations were included in order to improve the significance of the basin-wide averages. Although precipitation at these stations will not directly contribute to streamflow at the gauging station in question, the distances between the boundaries of the drainage basins and the stations are small so it is expected that even mesoscale convective precipitation events contributed to precipitation at the drainage basin boundaries and the stations similarly.

Monthly climate divisional data were also used to understand spatial variability in climate-mode precipitation relationships (Figure 5.1a). The data, extending back to 1895, are available for 344 climate divisions, regions within states that have a uniform climatology, eight to ten for each state (Guttman and Quayle, 1998). In the present study, only the precipitation data for the period 1900-2010 were used and converted to anomalies like the previously mentioned data sets. The benefit of using the divisional data versus the station-based HCN data is that local climatological effects can be smoothed out, allowing better detection of climate signals.

5.2.2 Climate Indices

For this study, eight climate indices were selected (Table 5.1) based on studies that have identified physical relationships between climate modes and precipitation and streamflow in the Mid-Atlantic region (see Introduction). We first briefly describe the indices before discussing the sources of the data.

Perhaps the most well-known climate mode is ENSO, whose evolution and strength can be monitored using two metrics, the Southern Oscillation Index (SOI) and the Niño 3.4 index (Trenberth, 1984; 1997), which capture the atmospheric and oceanic components, respectively, of ENSO. The SOI is calculated as the difference of SLP anomalies between Tahiti and Darwin, Australia. The Niño 3.4 index is defined as the average SST in the region 5°N-5°S, 170°W-150°W.

The climate modes most closely linked to the North Pacific region are the NPO, PDO, and PNA. The North Pacific Index, which describes the NPO, is defined as the area-weighted SLP over the region 30°N-65°N, 160°E-140°W (Trenberth and Hurrell, 1994). The PDO index is constructed from the leading mode of an un-rotated Empirical Orthogonal Function (EOF) analysis of monthly residual SST anomalies in the North Pacific poleward of 20°N, where the monthly residual is the difference between the observed anomaly and the global-mean SST anomaly (Mantua et al., 1997; Mantua and Hare, 2002). The PNA index is constructed from a rotated EOF analysis of daily 500-hPa height anomalies in the region bounded by 20°N and 90°N in the Northern Hemisphere (Barnston and Livezey, 1987).

Table 5.1. Climate indices, data sources, record lengths, and relevant publications.

Climate index	Source	Record length used	References
NAO	NCAR	1900 - 2010	Hurrell (1995)
AO	CPC	1950 - 2010	Thomas and Wallace (1998)
AMO	CPC	1900 - 2010	Enfield et al. (2001), Trenberth and Shea (2006)
Nino 3.4	NCAR	1900 - 2010	Trenberth (1997)
SO	NCAR	1900 - 2010	Trenberth (1984)
NPO	NCAR	1900 - 2010	Trenberth and Hurrell (1994)
PNA	CPC	1950 - 2010	Wallace and Gutzler (1981)
PDO	UW	1900 - 2010	Mantua et al. (1997), Mantua and Hare (2002)

Metrics of climate variability related to the North Atlantic Ocean are the AMO and NAO indices. The Atlantic basin exhibits SST variability with a preferred multi-decadal timescale. This multi-decadal oscillation of SST has been termed the AMO (Kerr, 2000),

whose index is defined as the detrended average SST in the North Atlantic basin from 0° to 70°N (Enfield et al., 2001; Trenberth and Shea, 2006). The NAO consists of a dipole SLP pattern established by the Azores High and the Icelandic Low and the NAO index is defined as the normalized SLP difference between these pressure centers (Hurrell, 1995). A climate mode related to the NAO is the AO, an oscillation of the polar vortex. The AO index is constructed by standardizing the first principle component time series of northern-hemisphere SLP for all months of the year (Thomas and Wallace, 1998).

Climate index data used in the analysis (Table 5.1) were obtained from the Climate Prediction Center (CPC), the National Center for Atmospheric Research (NCAR, http://www.cgd.ucar.edu/cas/catalog/climind/), and the University of Washington (UW, http://jisao.washington.edu/pdo/PDO.latest). All data are monthly averages and, when necessary, were converted to anomalies by removing the mean annual cycle. All of the indices were available for the period 1900-2010, except those for the AO and PNA, which were available for the period 1950-2010. For consistency, the correlation analysis was applied to all eight indices for the period 1950-2010. The wavelet analysis was applied to all indices, except those for the AO and PNA, for the period 1900-2010.

5.2.3 Correlation Analysis

Linear Pearson correlation coefficients between streamflow and climate indices, between precipitation and climate indices, and between temperature and climate indices were computed using monthly, seasonal, and annual averages for the period 1950-2010. Means were identified as DJF (December, January, February) for winter, MAM (March, April, May) for spring, JJA (June, July, August) for summer, and SON (September, October, November) for fall. The calendar year was used for the annual means. Seasonal averages were computed because many climate modes are preferentially expressed in certain seasons. For example, it is in the Northern Hemisphere winter that the NAO and the AO often reach their maximum amplitudes and have the strongest influence. Significance was tested using the non-parametric bootstrap method (Efron, 1979) as follows: the data were resampled 10,000 times, correlation coefficients from the resampled data were computed, and then 95% and 99% confidence intervals of the resulting

distribution of synthetic correlation coefficients were computed. Correlation coefficients were rejected at the 5% significance level if the confidence interval contained a correlation coefficient of zero.

Correlation analysis was also used to quantify the impact of precipitation and temperature on streamflow. Linear Pearson correlation coefficients were computed for each calendar month between streamflow and precipitation for 1950-2010. Temperature can influence streamflow variability through evapotranspiration, snowfall, and snowmelt. In fact, in the New England region, the significant storage of water in snow results in temperature explaining up to 30% of streamflow variability (Bradbury et al., 2002). In the SRB, Najjar (1999) found that annual precipitation minus streamflow (a proxy for annual evapotranspiration) was positively correlated with temperature. Nevertheless, precipitation dominates the streamflow signal so that temperature-streamflow relationships may appear to be non-existent or weak. Therefore, the partial correlation coefficient (Mattson, 1981) between temperature (t) and streamflow (q), controlling for precipitation (p), was computed for each calendar month for 1950-2010, which allows temperature impacts on streamflow to be assessed with the precipitation-streamflow dependence removed:

$$r_{tq.p} = \frac{r_{tq} - r_{tp} \cdot r_{qp}}{\sqrt{(1 - r_{tp}^2)(1 - r_{qp}^2)}}, \tag{5.1}$$

where r_{xy} represents the simple correlation coefficient between x and y. The partial correlation coefficient can be equal to the simple correlation coefficient or it can be very different depending on how strong the influence of the third variable is on the relationship between the other two variables. The statistical significance of the correlation coefficients among temperature, precipitation, and streamflow was computed using the bootstrap method (Efron, 1979) in the same way as for the Pearson correlation coefficient.

5.2.4 Wavelet Analysis

Local and global wavelet power spectra were computed for streamflow, precipitation, and temperature. In this study we adopt the Morlet wavelet, which is given by

$$\psi_0(\eta) = \pi^{-1/4} e^{i\omega\eta} e^{-\frac{1}{2}\eta^2}, \tag{5.2}$$

where ω is the dimensionless frequency and η is the dimensionless time. Providing a balance between time and frequency localization, the Morlet wavelet with ω = 6 is recommended for identifying features of geophysical time series (Grinsted et al., 2004). To find the local wavelet power spectrum of a time series (x_n; n = 1, ... , N), such as streamflow or precipitation, one must take the wavelet transform of the time series, which is defined as

$$W_n^X(s) = \sqrt{\frac{\delta t}{s}} \sum_{n'=1}^{N} x_{n'} \psi_0[(n' - n)] \frac{\delta t}{s}], \tag{5.3}$$

where δt is a uniform timestep (one month in this case), s is the scale of the Morlet wavelet, and $\eta = s \cdot t$. The more traditional Fourier period λ is approximately related to the wavelet scale by $\lambda = 1.03s$. The wavelet power at a given scale and time is then given by $|W_n^X(s)|$. Averaging $W_n^X(s)$ over the time index results in the global wavelet power spectrum. The significance of both global and local wavelet power at a given frequency and time can be tested against a red-noise background. The original time series can be reconstructed by taking the inverse wavelet transform of the wavelet coefficients (Torrence and Compo, 1998). In particular, to reconstruct the signal at a particular frequency, all wavelet coefficients are set to zero except for those corresponding to the frequency components of interest. Taking the inverse wavelet transform of the modified wavelet coefficients will produce the signal at the desired frequency or period with all other frequency components removed. The reader is referred to Torrence and Compo (1998) and Grinsted et al. (2004) for a more detailed discussion of the theory of wavelet analysis and significance testing used in this paper.

To quantify the relationships between climate modes and Mid-Atlantic streamflow, precipitation, and temperature as a function of frequency and time, a wavelet coherence analysis was conducted. Following Grinsted et al. (2004), the wavelet coherence between two time series *X* and *Y* is given by

$$R_n^2(s) = \frac{\left|S(s^{-1}W_n^{XY}(s)\right|^2}{S(s^{-1}|W_n^X(s)|^2)\cdot S(s^{-1}|W_n^Y(s)|^2)}, \quad (5.4)$$

where $W_n^{XY}(s)$ is the cross-wavelet transform, defined as the product of the wavelet transform of *X* and the complex conjugate of the wavelet transform of *Y*, and *S* is a smoothing operator defined by $S(W_n(s)) = S_{scale}\left(S_{time}(W_n(s))\right)$. S_{time} represents smoothing in time and S_{scale} is smoothing along the wavelet scale axis. Eq. (5.4) resembles the definition of the correlation coefficient and, in fact, can be regarded as such. That is, a coherence value of 0 signifies that the two time series are unrelated, whereas a coherence value of 1 indicates the two time series are linearly related at the given frequency and time. Using Monte Carlo methods, the statistical significance of wavelet coherence was found by generating a large number of synthetic data pairs with the same lag-1 autocorrelation coefficients as the input time series, calculating the wavelet coherence for each pair, and then estimating the significance level at each scale using values outside the cone of influence (Grinsted et al., 2004). A more detailed discussion of wavelet coherence can be found in Grinsted et al. (2004).

Characteristic timescales reported in this study were identified using global wavelet power spectra, the time-averaged representations of the sample wavelet power spectra. Periods of maximum time-averaged power were considered the dominant timescale of variability. For time-averaged wavelet coherence, however, an alternative definition was used, which is given by

$$G_C(s) = \frac{\left|W^{XY}(s)\right|^2}{\left(\sum_{n=1}^{N}|W_n^X|^2\right)\left(\sum_{n=1}^{N}|W_n^Y|^2\right)}, \quad (5.5)$$

where

$$W_n^{XY}(s) = \sum_{n=1}^{N} W_n^X(s)W_n^{Y*}(s), \quad (5.6)$$

with the asterisk denoting the complex conjugate (Elsayed, 2006). Eq. (5.5) measures the coherence between two time series in the entire study period at a scale *s*. Statistical

significance of $G_C(s)$ was computed using Monte Carlo methods as follows: red-noise time series with the same lengths and autocorrelation coefficients as the two input data series were generated and $G_C(s)$ was computed for each pair of red-noise time series. The resulting distribution of $G_C(s)$ at each scale was then used to estimate the significance of the global coherence estimates.

To quantify the impacts of climate modes on streamflow and precipitation variability, the quantity

$$F_n^X(s) = R_n^2(s) W_n^X(s), \tag{5.7}$$

representing the fraction of the power of *X* at *n* and *s* related to *Y*, was computed. In the present case, *Y* is a climate mode, the input into the system, and *X* is streamflow or precipitation, the linear response to *Y*. Before the above quantity was calculated, wavelet coherence values that were not statistically significant at the 5% level were set to zero, the idea being that insignificant wavelet coherence means that the null hypothesis cannot be rejected and the two time series at a given time and scale are considered independent (i.e., $R_n^2(s) = 0$). To obtain physical estimates of climate-mode impacts on streamflow and precipitation, the inverse transform of $F_n^X(s)$ was computed. If X^{mode} denotes the inverse transform of $F_n^X(s)$ for all time indices and scales, then *X* can be decomposed as

$$X = X^{mode} + X^{residual}, \tag{5.8}$$

where $X^{residual}$ is the component of *X* unrelated to the climate mode. If $X = X^{mode}$, for example, then all of the time-series variability is due to forcing from the climate mode. It is noted that $X^{residual}$ may not be purely noise, which would be the case if two independent climate modes were driving streamflow or precipitation.

5.3. Results

5.3.1 Observed Variability of Mid-Atlantic Streamflow, Precipitation, and Temperature

The observed annually averaged Susquehanna, Delaware, and Hudson River streamflow anomalies are shown in Figures 5.2a, c, and e, respectively. The variability in the time series from 1900 to 1940 was characterized by high-frequency oscillations, whereas low-frequency variability as well as high-frequency variability occurred during 1940-1980. The low-frequency variability is evident by examining the five-year running mean of the time series (shown in black), which highlights the decadal variability of streamflow during the period 1950-1980. The 1960s drought and the pluvials (very wet periods) of the 1970s and 2000s are evident in all three basins. The 1960s drought was deepest in the DRB whereas the pluvials appear to have been greatest in the HRB. The mean annual cycles of streamflow and precipitation in the three basins are shown in Figures 2b, d, and f. The streamflow annual cycles are characterized by large maxima in spring, minima in summer, and a secondary maximum in late fall, extrema that are caused by the annual cycle in evapotranspiration and snowmelt—not precipitation, which has a weak annual cycle (e.g., Najjar, 1999). Unlike Susquehanna and Delaware River streamflow, Hudson River streamflow exceeded precipitation during March, possibly due to snowmelt during that month.

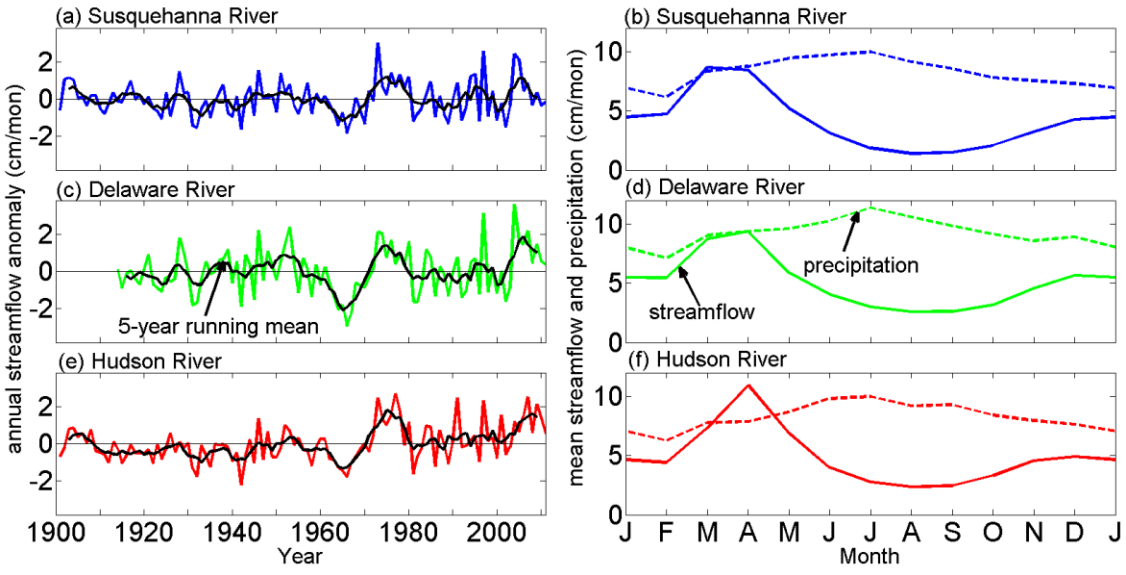

Figure 5.2. Observed annually averaged (a) Susquehanna, (c) Delaware, (e) Hudson River streamflow anomalies and 5-year running mean of the observed time series for 1900-2010. Mean annual cycles of (b) Susquehanna River streamflow and SRB precipitation, (d) Delaware River streamflow and DRB precipitation, and (f) Hudson River streamflow and HRB precipitation for 1900-2010.

5.3.2 Streamflow/precipitation-climate mode Correlation Analysis

Panels a and b of Tables 5.2, 5.3, and 5.4 show the linear correlation analysis of streamflow-climate index and precipitation-climate index for the Susquehanna, Delaware, and Hudson Rivers, respectively. The Niño 3.4 index was found to be weakly correlated with Delaware River streamflow and DRB precipitation in April, accounting for approximately 16% of the variance of precipitation. A similar relationship was found for SRB precipitation and no relationships were identified for the HRB. A possible explanation for the increased April streamflow during positive Niño 3.4 phases is the increased January-March US East Coast storm frequency and subsequent increase in precipitation (Eichler and Higgins, 2005). The significant correlation coefficients with the SO and Niño 3.4 indices were generally similar and of opposite sign. There are a few notable exceptions: for example, November Susquehanna streamflow and HRB precipitation were not correlated with the Niño 3.4 index but were correlated with the SO index. Differences may be the result of how SSTs lag changes in the atmosphere.

Numerous statistically significant PNA-precipitation and PNA-streamflow relationships were found, consistent with the fact that the PNA tends to be in a positive phase during positive Niño 3.4 events (Feldstein, 2000). The April PNA index was found to be significantly and positively correlated with April precipitation for the DRB and SRB, though the PNA index was only significantly correlated with Delaware River streamflow. The streamflow-PNA and precipitation-PNA relationships may be the result of a southeastward-shifted trough that provides more favorable conditions for coastal storm development (Leathers, 1991). The PNA and the related NPO also had seasonal influences: the PNA and NPO indices were significantly correlated with Delaware River streamflow and precipitation at the 1% level in spring and were also found to be significantly correlated with Susquehanna and Hudson River streamflow in fall.

Table 5.2. Linear Pearson correlation coefficients between climate indices and (a) Susquehanna River streamflow, (b) SRB precipitation, and (c) SRB temperature for 1950-2010. Only correlation coefficients significant at the 5% level are displayed, with correlation coefficients significant at the 1% level shaded in gray.

(a) Streamflow

	J	F	M	A	M	J	J	A	S	O	N	D		DJF	MAM	JJA	SON	Annual
NAO													NAO					
AO									0.34				AO					
AMO						-0.26				0.26			AMO			-0.26		
Nino 3.4		-0.30											Nino 3.4					
SO		0.27									-0.30		SO					
NPO						0.26			0.26		0.41		NPO				0.35	
PNA											-0.34		PNA				-0.30	-0.26
PDO													PDO					

(b) Precipitation

	J	F	M	A	M	J	J	A	S	O	N	D		DJF	MAM	JJA	SON	Annual
NAO													NAO					
AO													AO					
AMO						-0.26				0.49	-0.30		AMO					
Nino 3.4				0.38									Nino 3.4					
SO				-0.30							-0.30		SO					
NPO													NPO					
PNA				0.27									PNA					
PDO				0.33									PDO					

(c) Temperature

	J	F	M	A	M	J	J	A	S	O	N	D		DJF	MAM	JJA	SON	Annual
NAO	0.38	0.43	0.34	0.33						0.31		0.50	NAO	0.49				
AO	0.42	0.37	0.32	0.28						0.35	0.36	0.58	AO	0.53				
AMO													AMO				0.28	
Nino 3.4										-0.28			Nino 3.4				-0.25	
SO													SO					
NPO	0.29										0.32	0.33	NPO					
PNA	-0.30				-0.26				0.34		-0.29		PNA					
PDO	-0.30									-0.38	-0.39		PDO				-0.43	

Table 5.3. Same as Table 5.2 but for the DRB.

(a) Streamflow

	J	F	M	A	M	J	J	A	S	O	N	D		DJF	MAM	JJA	SON	Annual
NAO										-0.32			NAO					
AO									0.34				AO					
AMO						-0.30				0.44			AMO				0.31	
Nino 3.4		-0.28		0.27									Nino 3.4					
SO					0.30								SO					
NPO				-0.34					0.25		0.28		NPO		-0.37		0.36	
PNA				0.39									PNA		0.36			
PDO				0.27				-0.25					PDO					

(b) Precipitation

	J	F	M	A	M	J	J	A	S	O	N	D		DJF	MAM	JJA	SON	Annual
NAO								0.25					NAO					
AO													AO					
AMO						-0.31				0.45			AMO				0.26	
Nino 3.4				0.37									Nino 3.4					
SO				-0.26	0.33								SO					
NPO				-0.29									NPO		-0.31		0.30	
PNA				0.38									PNA		0.34			
PDO				0.29									PDO					

(c) Temperature

	J	F	M	A	M	J	J	A	S	O	N	D		DJF	MAM	JJA	SON	Annual
NAO	0.37	0.41	0.32	0.30						0.28		0.53	NAO	0.50				
AO	0.41	0.36	0.31							0.39	0.36	0.56	AO	0.53		0.28		
AMO									0.30				AMO				0.30	
Nino 3.4										-0.26			Nino 3.4					
SO													SO					
NPO	0.25										0.33	0.27	NPO					
PNA	-0.25								0.30		-0.31		PNA					
PDO	-0.24									-0.39	-0.36		PDO				-0.42	

Table 5.4. Same as Table 5.3 but for the HRB.

(a) Streamflow

	J	F	M	A	M	J	J	A	S	O	N	D		DJF	MAM	JJA	SON	Annual
NAO	0.26												NAO					
AO			0.30						0.32				AO					
AMO										0.31			AMO					
Nino 3.4													Nino 3.4					
SO													SO					
NPO									0.29				NPO				0.29	
PNA									-0.27				PNA					
PDO													PDO					

(b) Precipitation

	J	F	M	A	M	J	J	A	S	O	N	D		DJF	MAM	JJA	SON	Annual
NAO		0.26											NAO					
AO					0.32								AO					0.26
AMO					-0.31					0.48	-0.39		AMO					
Nino 3.4		-0.32											Nino 3.4					
SO		0.27									-0.30		SO					
NPO													NPO					
PNA													PNA					
PDO		-0.28											PDO					

(c)Temperature

	J	F	M	A	M	J	J	A	S	O	N	D		DJF	MAM	JJA	SON	Annual
NAO	0.38	0.40					0.32			0.30		0.49	NAO	0.51				
AO	0.41	0.27					0.37			0.42	0.40	0.53	AO	0.52		0.37	0.31	
AMO									0.28				AMO				0.28	
Nino 3.4			0.29					-0.27					Nino 3.4					
SO			-0.30										SO					
NPO										0.27	0.34	0.26	NPO					
PNA									0.37	-0.25	-0.30		PNA					
PDO										-0.40	-0.37		PDO				-0.47	

It is noted that the NP, PDO, and PNA indices were not found to be correlated with summer streamflow for all three rivers, though Barlow et al. (2001) found North Pacific SSTs, the so-called North Pacific mode (NPM), to be related to summer streamflow and drought conditions across the Northeast US. On the other hand, the modes were generally related to spring and fall streamflow, which were correlated with the SO and Niño 3.4. Differences between the results from Barlow et al. (2001) may be due to how the NPM, using EOF analysis, was statistically constrained to be independent from the PDO that was simultaneously derived from the same data set. In the present case, the NPO and PDO indices are not orthogonal but strongly correlated so that no independent information was gained.

In agreement with Xu et al. (2012), some PDO-streamflow associations were identified, particularly in April. The lagged relationship between Susquehanna streamflow and the PDO index found by Xu et al. (2012) may be an artifact of the autocorrelation inherent in the PDO time series and the PDO's relationship with the PNA and NPO. For example, the winter PDO index may be correlated with the subsequent fall PNA index,

which was related to fall streamflow, resulting in the PDO index being correlated with future streamflow.

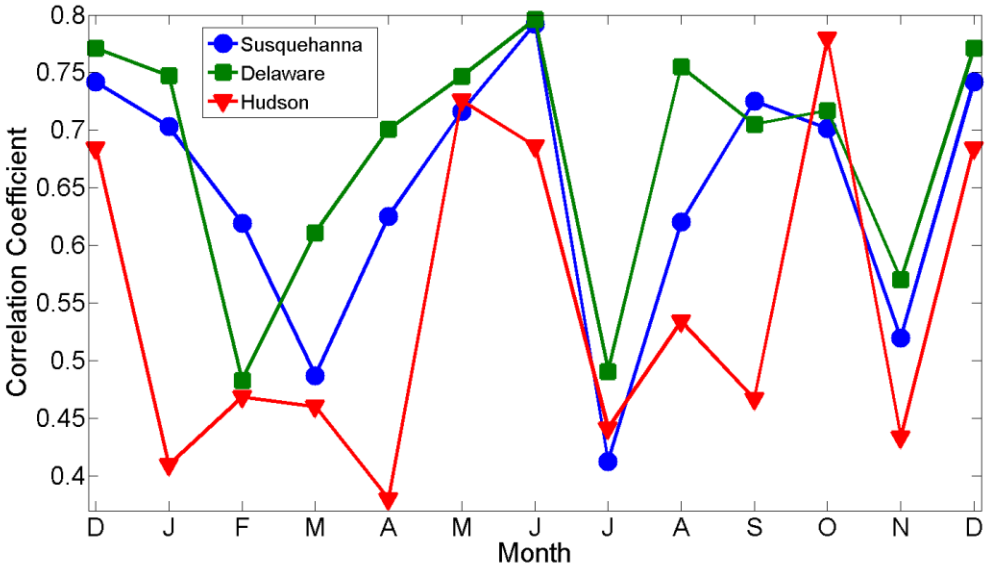

Figure 5.3. Pearson correlation coefficients between mean monthly streamflow and precipitation for the Susquehanna, Delaware, and Hudson River basins for the period 1950-2010. All correlation coefficents are signifcant at the 5% level.

Correlation coefficients between streamflow and precipitation are shown in Figure 5.3, where all the correlation coefficients are significant at the 5% level. For all three river basins, the streamflow-precipitation correlation coefficients reached local minima in February-April, July, and November; local maxima were found in May-June, August-October, and December. As shown in Figure 5.2, precipitation equals streamflow in February and March for the DRB and SRB and yet the correlation coefficients reach local minima in those months, suggesting that some of the precipitation is falling as snow. The same arguments for the HRB hold in January and February but streamflow exceeds precipitation in March so that snowmelt may contribute to the low correlation coefficients in March and April. The minima in July for all three river basins may be due to large base-flow contributions and large evapotranspiration rates. Base-flow has a longer time-scale so that streamflow from prior months may be contributing to streamflow variability in subsequent months. Perhaps the most interesting minima occurred in November, when there also may be a base-flow contribution in this month given the relatively low mean

streamflow. Precipitation, in addition, may also be falling as snow, contributing little to streamflow. Perhaps it is the superposition of the base-flow contribution and snowfall that results in lower correlation coefficients in November.

In contrast to the seasonal time scale, streamflow and precipitation are highly correlated on the annual timescale, with the correlation coefficients between annual-mean streamflow and precipitation for 1950-2010 being 0.92, 0.95, and 0.85 for the SRB, DRB, and HRB, respectively, similar to previous analyses (Najjar, 1999; Najjar et al., 2009).

5.3.3 Temperature-climate mode Correlation Analysis

Partial correlation coefficients (with the dependence of precipitation removed) between mean maximum temperature and streamflow are shown in Figure 5.4. In general, temperature was positively and significantly partially correlated with streamflow during the cold season for all three river basins, consistent with the fact that winter precipitation type is strongly dependent on daily mean maximum temperature on precipitation days so that precipitation on cold days will fall as snow and contribute little to streamflow (Serreze et al., 1998).

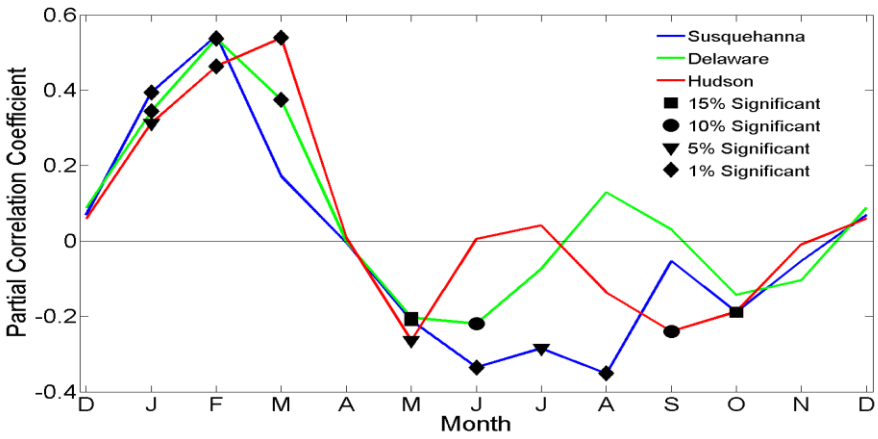

Figure 5.4. Partial correlations coefficients between mean monthly maximum temperature and streamflow for the Susquehanna, Delaware, and Hudson Rivers for the period 1950-2010. Markers indicate the significance of the partial correlation coefficients.

The negative relationships between temperature and streamflow during the summer are physically consistent with an increase in temperature leading to more

evapotranspiration and less streamflow. Having been correlated with temperature, streamflow may also have been influenced by climate modes through temperature effects in addition to precipitation effects, motivating the temperature-climate mode analysis discussed next.

Tables 5.2c, 5.3c, and 5.4c show the results of the temperature-climate index (simple) correlation analysis. The NAO index was found to be positively correlated with temperature, with the strongest influence occurring in the winter, when the atmosphere is most dynamically active and an enhanced jet and surface southerly flow associated with the positive NAO phase advects positive temperature anomalies across the Northeast US (Notaro et al., 2006). The positive correlation coefficients between the AO and temperature are a result of a positive AO phase being associated with fewer polar air outbreaks across North America (Thompson and Wallace, 2001), whereas the moderate temperature-PNA and temperature-NPO connections are the result of a deep Eastern-US trough during positive PNA and negative NPO phases (Leathers et al., 1991). A moderate correlation was identified between temperature and the PDO index in the fall, though no significant correlation coefficients were found with the PNA and NPO indices.

September Hudson River streamflow was negatively correlated to the PNA index but no precipitation-PNA association was identified. The PNA, however, was positively correlated with temperature during that month, and therefore the streamflow variability could possibly be related to PNA-related changes in temperature. The PNA-temperature relationships may also explain why, in November, the PNA index was significantly related to Susquehanna River streamflow and but not to precipitation. A positive PNA index, being associated with lower-than-normal temperatures, may have resulted in more precipitation falling as snow, reducing November monthly streamflow. Note that the PNA index needs to only be related to precipitation type to affect streamflow. On the other hand, the precipitation-climate mode relationships could explain the April streamflow-climate mode associations found for the Delaware River because April streamflow was strongly correlated with April precipitation (Figure 3) and was uncorrelated with April maximum temperature (Figure 5.4). The ENSO-induced Delaware River streamflow variability is also associated with the concurrent changes in precipitation. Despite the moderate

correlation between temperature and the indices for the AO and NAO, the NAO and AO were not found to be significantly correlated with winter streamflow.

5.3.4 Wavelet Analysis

In order to detect the major frequencies of oscillation in the hydrometeorology of the Mid-Atlantic Region, we now present wavelet power spectra of SRB, DRB, and HRB streamflow and precipitation along with the corresponding global wavelet spectra (Figures 5.5, 5.6, and 5.7).

The local power spectra revealed significant streamflow variability at a period of roughly two decades in all three river basins from roughly 1940 to 2010 (Figures 5.5a, 5.6a, and 5.7a). The global power spectra of streamflow also identified statistically significant wavelet power (Figures 5.5b, 5.6b, and 5.7b). For the Hudson and Susquehanna Rivers, 1% significant peaks at λ = 26 years were identified, while a 0.1% significant peak at λ = 26 years was identified for the Delaware River (Figures 5.5b, 5.6b, and 5.7b). Secondary peaks at λ = 18 years were also identified in the global wavelet spectrum of Hudson (1% significance) and Susquehanna (5% significance) River streamflow. The wavelet spectrum for Susquehanna River streamflow also contains a 5% significant peak at ~10 years, which was not found for the other two rivers.

The wavelet power spectra of precipitation (panels c and d of Figures 5.5, 5.6, and 5.7) are similar to the wavelet power of streamflow, suggesting that the streamflow variability was primarily driven by regional precipitation changes. This is not surprising given the strong correlation between annual-mean streamflow and annual-mean precipitation noted earlier (Section 5.3.2). The statistical significance of the peaks, however, are lower, especially for HRB and DRB precipitation, where the peaks at λ = 26 years only exceed the 5% significance level (Figures 5.5d, 5.6d, and 5.7d). The peak at λ = 26 years for Susquehanna River streamflow is nearly significant at the 1% level. There is also a secondary peak at ~18 years for the DRB not found for the SRB or HRB, though a similar peak was found in the global spectra for Hudson River streamflow.

The results from the wavelet analysis of temperature (not shown) identified no significant decadal variability for all three river basins and support the idea that streamflow variability at the decadal timescale was mainly driven by precipitation changes.

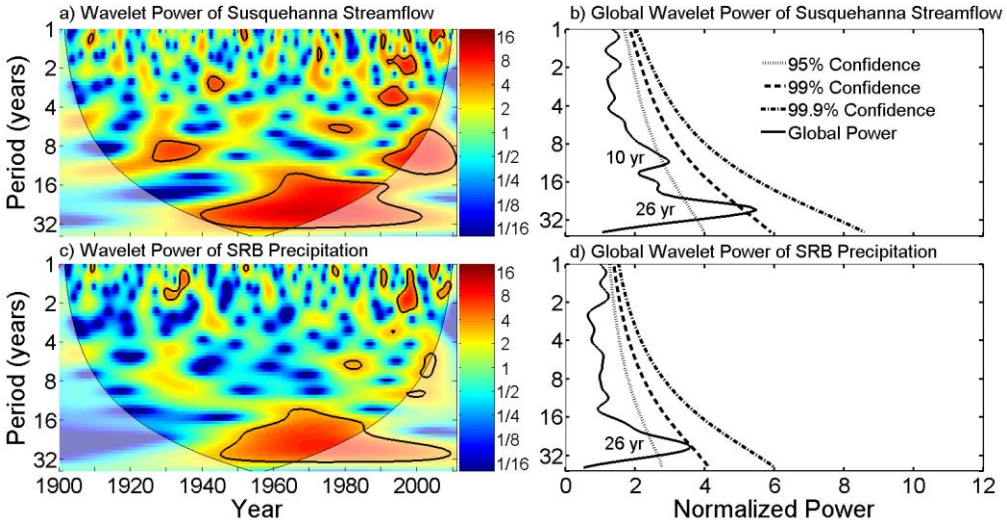

Figure 5.5. (a) Normalized wavelet power spectrum of Susquehanna River streamflow from 1900 to 2010 together with the global wavelet power spectrum (b). (c) The normalized wavelet power spectrum of SRB precipitation together with the global wavelet power spectrum (d). Thick black contours in wavelet power spectra enclose areas of 5% significance against a red-noise background. Light shading represents the cone of influence. In the global power spectra, thick black lines represents the global wavelet power estimates and thin dashed black lines are the 95%, 99%, 99.9% confidence bounds against red-noise background spectra. Periods of peaks in the global power spectra exceeding 95% confidence are indicated.

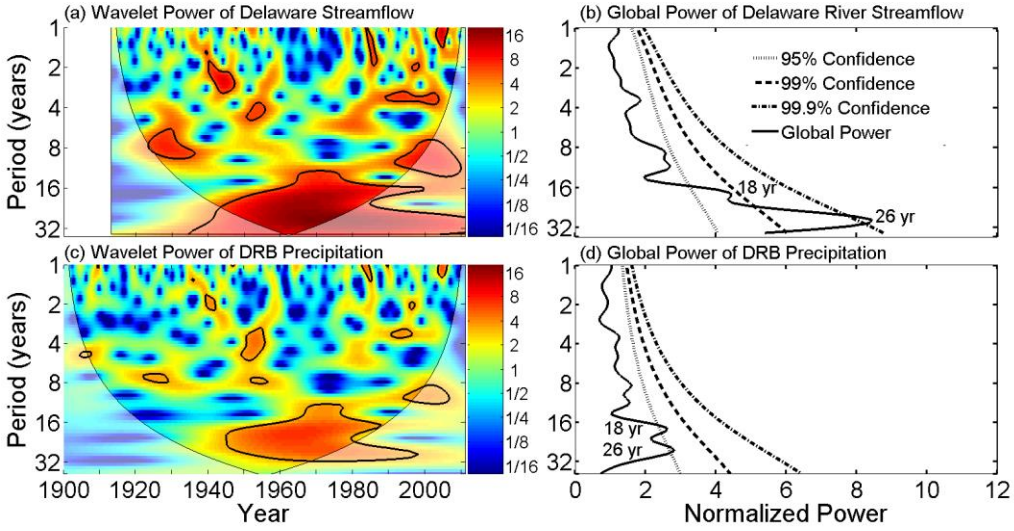

Figure 5.6. Same as Figure 5.5 but for the Delaware River Basin. Note that the streamflow record is shorter (1913-2010).

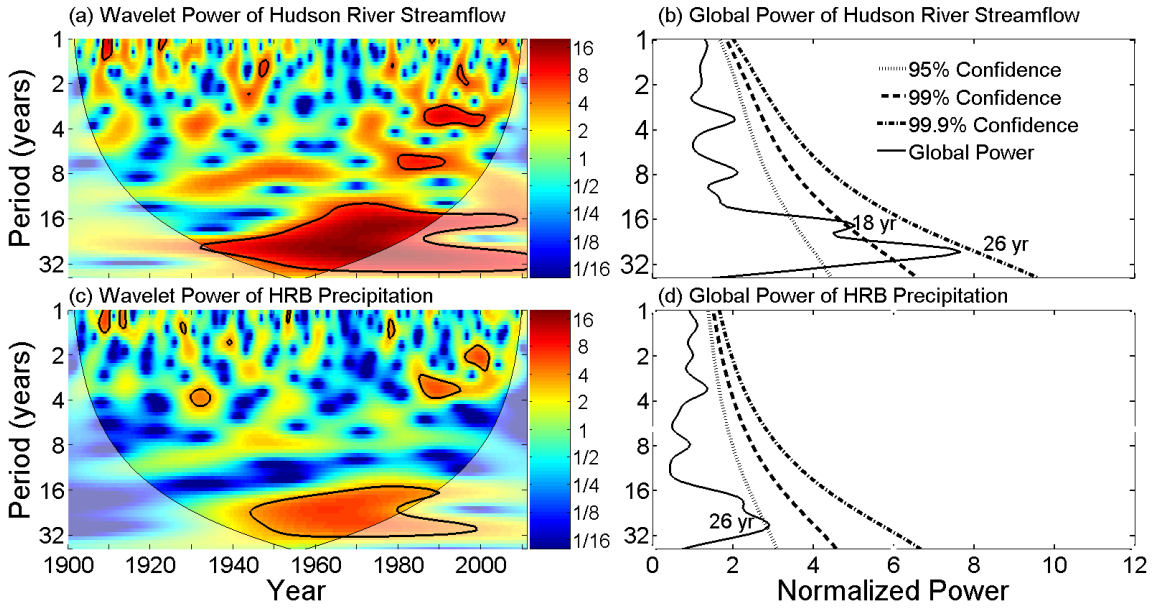

Figure 5.7. Same as Figure 5.5 but for the Hudson River Basin.

5.3.5 Wavelet Coherence Analysis

Smoothed time series of the PDO index, SOI, and Delaware River anomalies are shown in Figure 5.8. From 1940 to 2000, the SOI and streamflow time series fluctuated coherently. For example, the streamflow decline associated with the 1960s drought was accompanied by a negative SOI and the subsequent wet period in the 1970s was accompanied by positive values of the SOI. Also note that the 1960s drought occurred during a strongly negative PDO phase. The relationship of the climate indices with streamflow of the Hudson and Susquehanna Rivers were similar (not shown). A wavelet coherence analysis was used to determine precisely what frequency components of the time series fluctuated coherently.

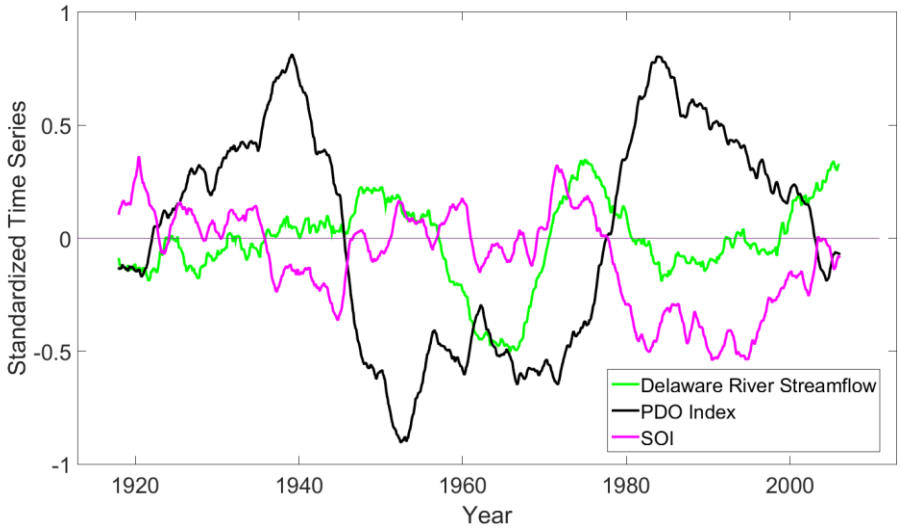

Figure 5.8. 10-year running mean of the PDO index, SOI, and Delaware streamflow anomalies. The smoothed time series were standardized by dividing by their respective standard deviations.

Local and global wavelet coherence spectra show that the PDO, AMO, NAO, and SO are coherent with Susquehanna River streamflow at a variety of timescales (Figure 5.9). Wavelet coherence spectra for the Niño 3.4 index and NPO index are not shown for any of the basins because the Niño 3.4 results are nearly identical to those for the SO and the NPO results reveal no significant coherence. The PDO, NAO, and SO were found to be coherent with Susquehanna River streamflow at multi-decadal timescales, mainly in the latter half of the record. The wavelet coherence analysis detected a 7-year lagged (i.e., streamflow leads) relationship with the PDO at λ = 23 years, an in-phase relationship between streamflow and the SO at λ = 25 years, and a 2-year lagged relationship with the NAO at 20 years, where the peaks were inferred from the significance of global coherence as shown in Figure 5.9e and the phase relationships were determined by finding the circular mean of the phase in the local coherence spectra at each period. The AMO has a multidecadal peak (λ = 23 years in the global spectra), which reflects local coherence in the 1950s and 60s. The secondary peak in SO global coherence significance at λ = 18 years suggests that the region of local coherence in Figure 5.9c is actually the merger of two separate peaks. The SO and the PDO also show peaks in the global spectra at around 10 years, which are presumably related to local spectra significance since approximately 1990. The phases are

roughly opposite to those corresponding to the ~20-year periods, the implications of which will be discussed in Section 5.4.6.

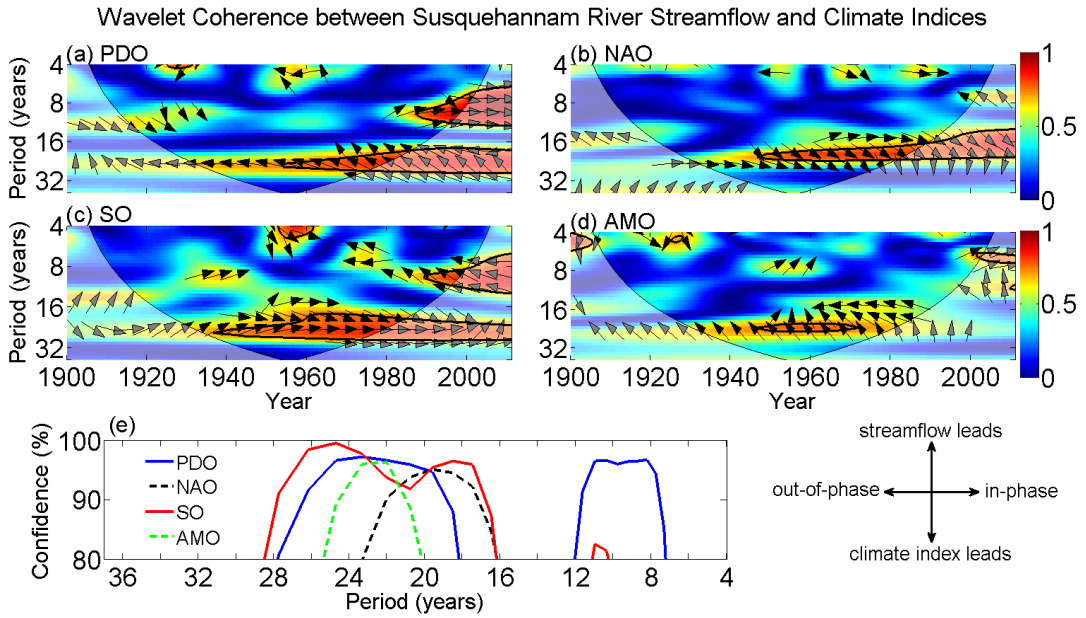

Figure 5.9. (a) Wavelet coherence of monthly Susquehanna River streamflow and the (a) PDO, (b) NAO, (c) SO, and (d) AMO indices during 1900-2010. Thick black contours indicate significance at the 5% level and arrows indicate relative phase relationships (see legend in the lower right). Light shading represents the cone of influence. The wavelet coherence spectra were truncated to four years for clarity. Arrows are only plotted for those wavelet coherence values exceeding 0.55. (e) Significance of the 1900-2010 time-averaged wavelet coherence between streamflow and the climate indices.

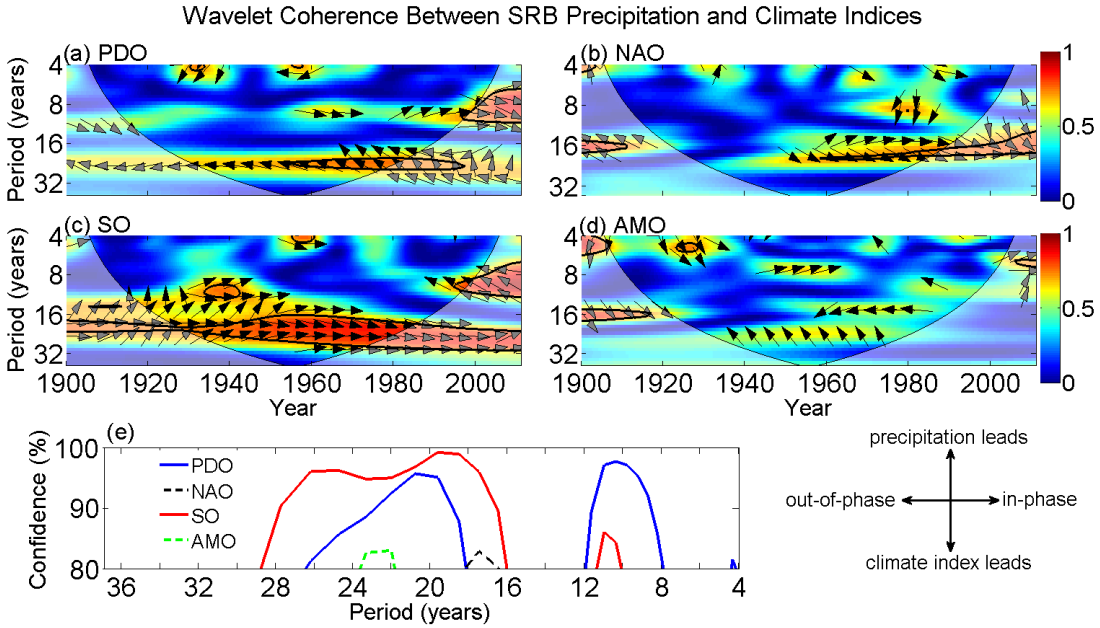

Figure 5.10. Same as Figure 5.9 but for Susquehanna River Basin precipitation

Wavelet coherence between SRB precipitation and climate modes (Figure 5.10) are broadly similar to spectra between Susquehanna River streamflow and climate modes, with both coherence spectra having significant (95% confidence) peaks for the SO at approximately 18 and 26 years and for the PDO at about 10 and 20 years. However, the AMO and NAO peaks, significant in the streamflow global coherence spectra, have less than 85% confidence in the corresponding precipitation spectra.

Wavelet coherence results for the DRB (Figures 5.11 and 5.12) were broadly similar to those for the SRB in terms of dominant periods, phase relationships, and temporal variability, except that the significance in the local spectra for the PDO and SO at ~20 years extended earlier into the 20th century. As with the SRB, the coherence between streamflow and climate modes was similar to that between precipitation and climate modes, except that precipitation generally had a greater response to the PDO and SO at ~20 years, suggesting that another mechanism may be weakening the response of streamflow to the climate modes.

The local coherence significance region extending from 1980 to 2010 corresponding to local coherence between the SO and streamflow at ~10 years for the SRB is also present for the DRB, with the same out-of-phase relationship as the SRB. The PDO streamflow relationships at ~10 years were found to be identical. In particular, the local coherence spectra for SRB and DRB precipitation contain significance regions at ~10 years (extending from 1980 to 2010), which are similar and smaller than those for streamflow. The differences indicate that another mechanism enhanced the streamflow response to the PDO, unlike the responses at ~20 years. The difference between the streamflow and precipitation response to the PDO is not, however, evident in global coherence spectra for the SRB, unlike for the DRB, where the confidence of peak for precipitation is approximately 5% less. Statistically significant global coherence between Delaware River streamflow and the NAO index at λ = 17 years was also identified, though locally little statistical significance was found. No global or local statistically significant NAO-precipitation relationships were detected. At both 10 years and 20 years the precipitation-PDO and precipitation-SO relationships were similar to the streamflow-climate mode

relationships, suggesting that precipitation was the primary driver of the multi-decadal streamflow variability, which is influenced by the PDO and SO.

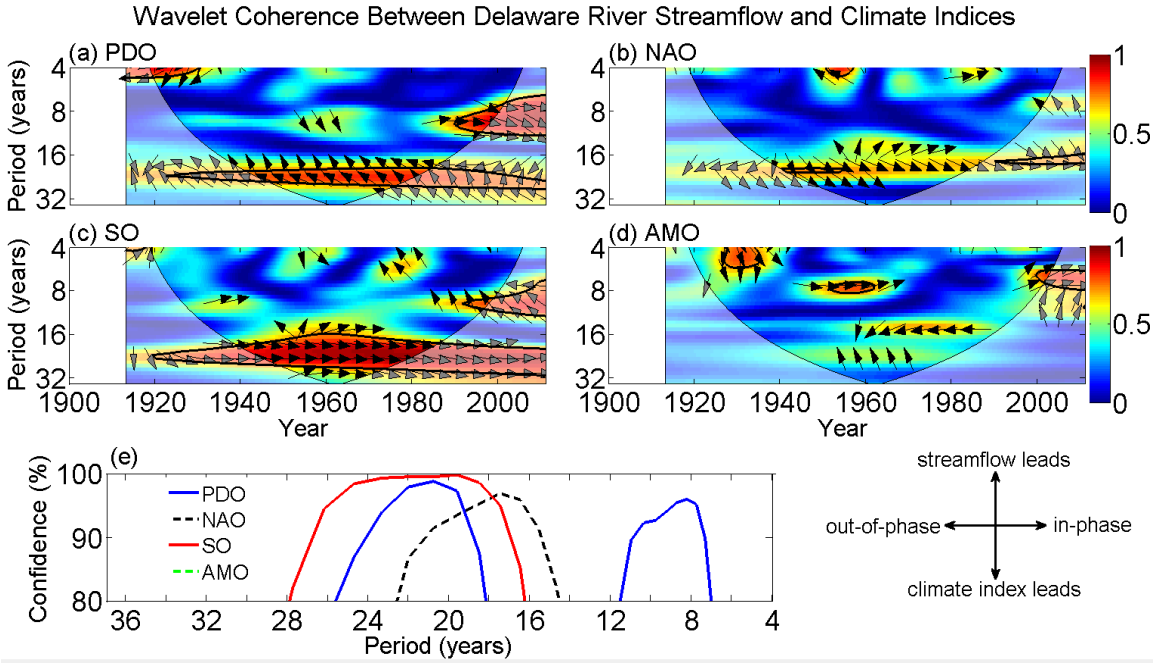

Figure 5.11. Same as Figure 5.9 but for Delaware River streamflow for 1913-2010.

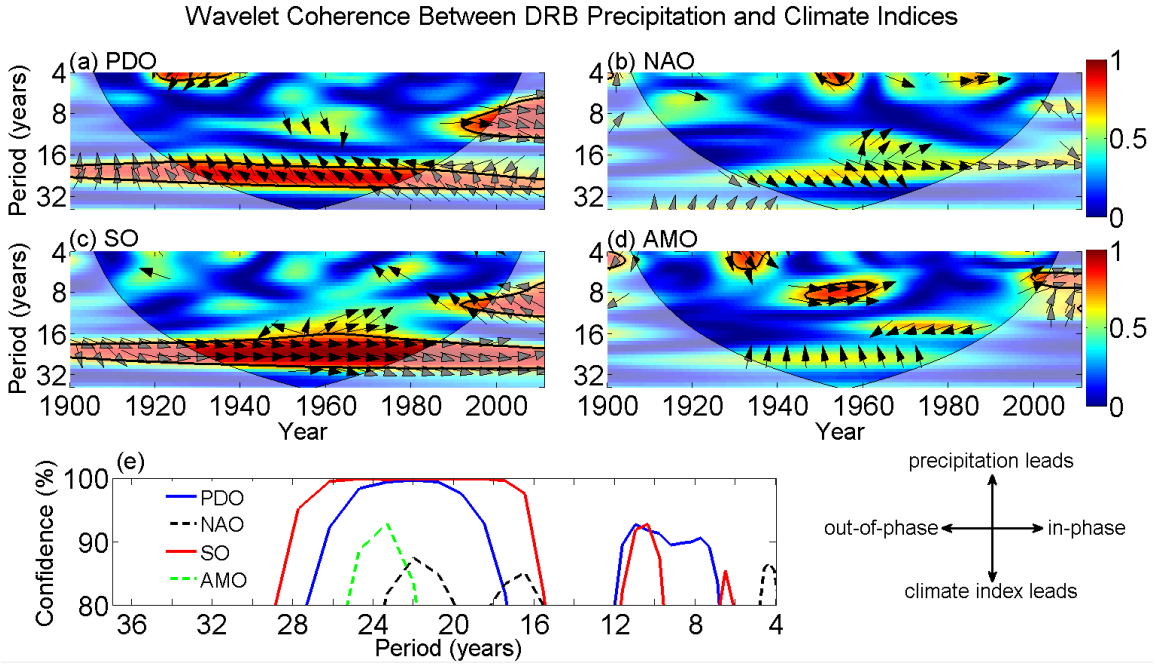

Figure 5.12. Same as Figure 5.9 but for Delaware River Basin precipitation.

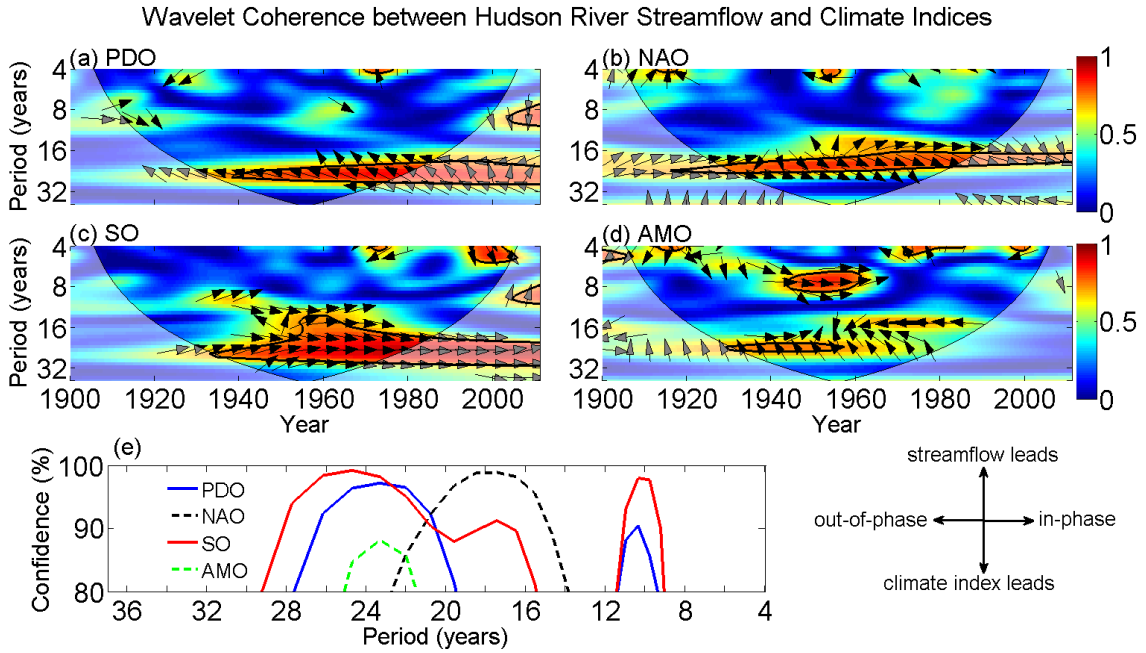

Figure 5.13. Same as Figure 5.9 but for Hudson River streamflow.

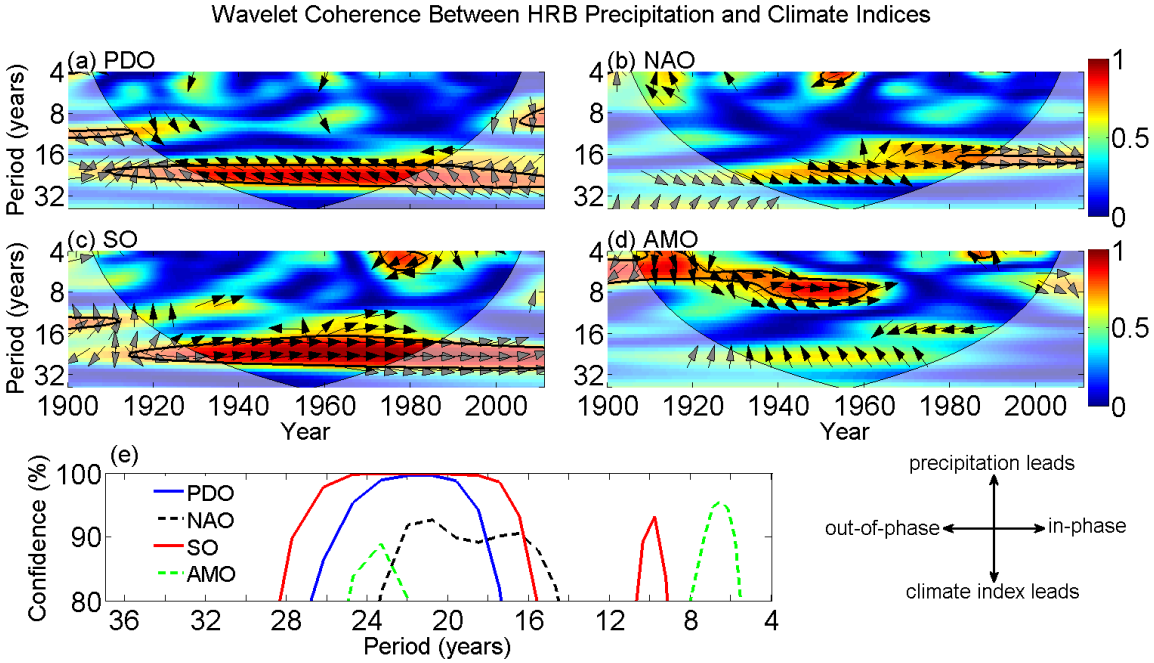

Figure 5.14. Same as Figure 5.9 but for Hudson River Basin precipitation.

The wavelet coherence between the four climate modes and Hudson River streamflow is shown in Figure 5.13. The wavelet coherence analysis detected a statistically significant coherent and in-phase SO-streamflow relationship at λ = 25 years, as indicated

by the peak in global coherence (Figure 5.13e). A period of significant coherence between the PDO and streamflow extending from 1940 to 2010 was also identified at $\lambda = 23$ years, with streamflow lagging by 7 years. The results were similar to those for the other rivers except that the Hudson River streamflow relationship with the SO was stronger than the Susquehanna-SO relationship. The global coherence spectra indicate that the SO was more coherent with Hudson River streamflow than with streamflow of the other rivers at $\lambda = 10$ years. It is also noted that the HRB precipitation response to the PDO and SO is stronger than the streamflow response, implying that another mechanism is operating to weaken the streamflow response.

One commonality between the HRB and SRB is the NAO-streamflow relationships; little statistically significant local coherence was found for the DRB, whereas both Hudson and Susquehanna River streamflow were both significantly coherent with the NAO at ~18 years. The commonality also exists in the local coherence spectra for precipitation but is less evident. A transient relationship between precipitation and the AMO in the HRB—not existing in the local coherence spectra for the DRB and SRB—was also found in the period band of 6-8 years from 1900 to 1960. A corresponding peak in the global coherence was also identified, which was also not found for the DRB and SRB.

It is possible that the observed streamflow variability at a period of 20 years was partially due to temperature variability. The wavelet coherence between climate indices and temperature for each river basin was therefore computed (not shown). For the SRB no significant wavelet coherence with temperature and the SO and AMO indices at a period of 20 years was found. On the other hand, significant out-of-phase coherence was found with the PDO index at a period of 20 years, suggesting that the positive phase of the PDO contributes to drier and cooler conditions across the Northeast US at that timescale. Namias (1966) also found the 1960s drought to be accompanied by below-normal temperatures associated with a prevailing northerly wind component. Identical relationships with the PDO index were found for the DRB and HRB and little to no significant coherence was found with the other climate indices. The out-of-phase relationship between the PDO and temperature could explain the weaker response of streamflow to the PDO for the DRB and

HRB; cooler conditions would favor more streamflow on longer-scales, opposing the drying effects caused by the lack of rainfall.

5.3.6 Impact of Climate Modes on Low-frequency Variability of Streamflow

The PDO and SO components of mean annual anomalies of Delaware streamflow, estimated using Equation (5.7), are shown in Figure 5.15 from 1920 to 2010, when large, low-frequency fluctuations were present. These climate modes appear to have contributed substantially to the 1960s drought and the pluvials of the 1970s and 2000s. To quantify the contributions of these two climate modes to the three aforementioned hydrometeorological events, the mean streamflow anomaly for the periods 1963-1968, 1971-1979, and 2002-2007 was computed for the SO component, the PDO component, and the observed time series for the Susquehanna, Delaware, and Hudson Rivers (Table 5.5). For the 1960s drought, the SO and PDO contributions were greatest for the Hudson and lowest for the Susquehanna, being roughly a factor of two larger for the Hudson; the SO contributions were about twice those of the PDO. Note that the contributions of the SO and PDO may not be independent because of the close relationship between these two climate modes (Newman et al., 2003).

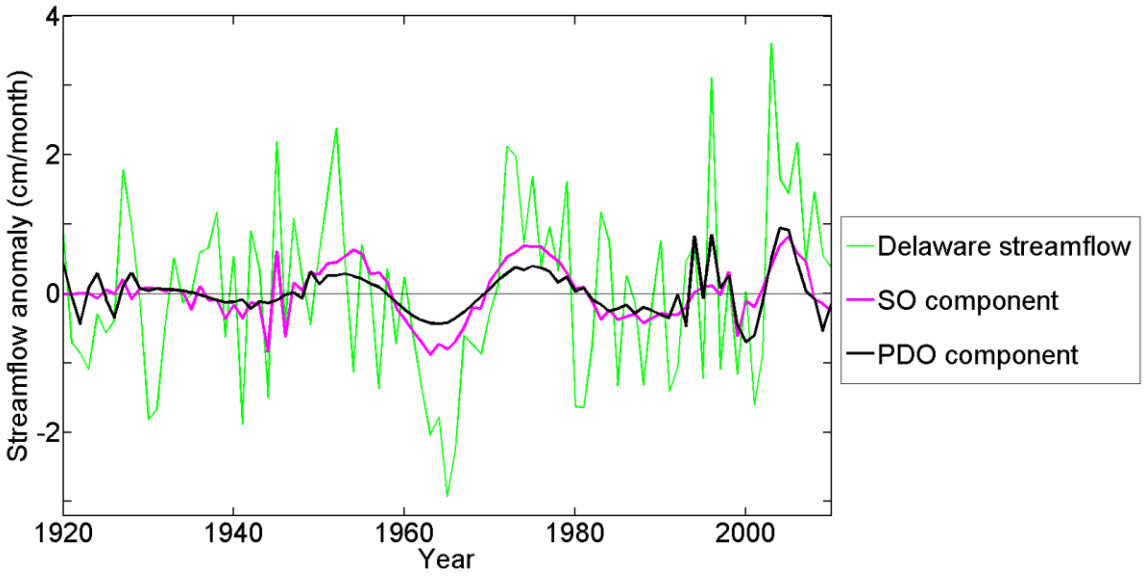

Figure 5.15. The observed annually averaged Delaware River streamflow anomalies and the PDO and SO components of the observed time series for 1920-2010.

The impact of the PDO and SO on the subsequent wet period 1971-1979 was similar for all three rivers but was greatest for Delaware River. The PDO's contribution was also smaller compared to the SO contributions for all three rivers. Similarly, the SO and PDO contributed approximately equally to the pluvial of the 2000s, with the Susquehanna being most influenced and the Hudson the least. In fact, the contributions differed by more than a factor of two, opposite to the climate-mode contributions to the 1960s drought. Results for precipitation (not shown) were similar to those for streamflow, indicating that the climate modes influenced streamflow during these hydrometeorological events mainly through precipitation. In addition to the SO and PDO, the analyses presented in Figure 5.15 and Table 5.5 were conducted for the NAO, NPO, AMO, and Niño 3.4 indices. No substantial contributions to the 1960s drought and pluvials of the 1970s and 2000s were found for the NAO, NPO, and AMO; Niño 3.4 results were very similar to those for the SO.

5.3.7. Teleconnection Analysis

Among all of the climate modes, the Southern Oscillation appears to have the greatest influence on Mid-Atlantic streamflow. Further, the SO influence is greatest at a period of approximately two decades. To investigate a possible teleconnection influence of the SO on Mid-Atlantic streamflow, a wavelet coherence analysis was conducted between the SO index and climate-division precipitation. The phase of the SO-precipitation relationship at a period of 20 years for each climate division is indicated by black arrows in Figure 5.16 when the global coherence exceeds the 5% significant level. The SO-precipitation relationship was found to be in-phase across the eastern US, with statistically significant global coherence limited to the Northeast. A remote region of significant global coherence is also present in the Western and Southern US but the phase relationship is opposite to that of the Northeast.

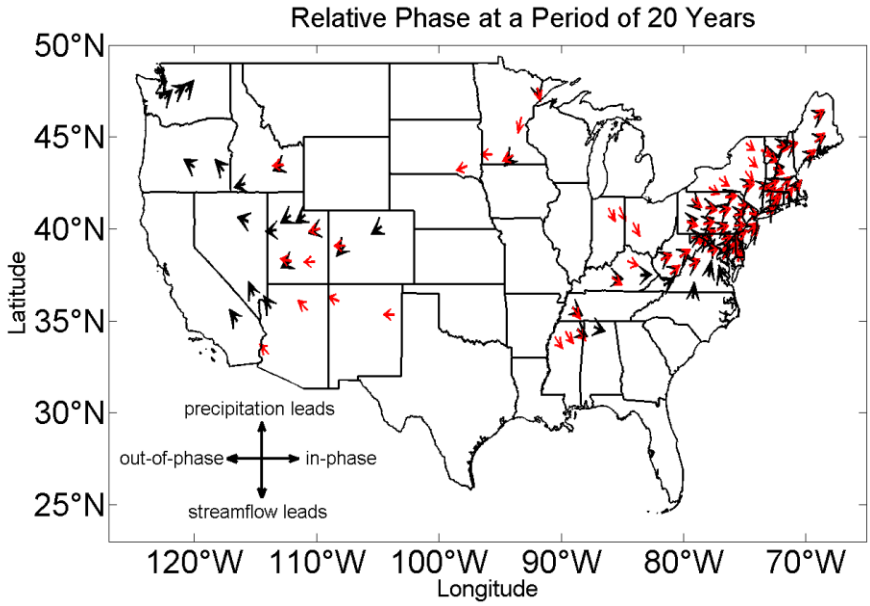

Figure 5.16. Phase relationships of the SO (black arrows) and Delaware River streamflow (red arrows) with climate divisional precipitation anomalies at a period of 20 years for 1913-2010. Arrows are located at the centroids of the climate divisions for which the global coherence was significant at the 5% level

A similar analysis was conducted between Delaware River streamflow and climate-division precipitation (red arrows). Remarkably, it was found that the global coherence spatial pattern was nearly identical to the SO-precipitation associations. The remote region of significant global coherence in the West is suggestive of a teleconnection pattern, with wet conditions in the Western US being accompanied by drier conditions in the East. In fact, Namias (1966) noted that while the Northeast was dry during the period 1962-1966, the Southwest and Northern plains were wet, generally consistent with the results presented in Figure 5.16. Seager et al. (2005) found, using a general circulation model, that tropical SSTs can produce the Dust Bowl in the 1930s for the Southwestern US, but Seager et al. (2012) showed that the models cannot produce the 1960s drought in the Northeastern US. However, Figure 5.16 implies that Delaware River streamflow is phase-locked to Southwestern US precipitation and so must also be phase-locked to tropical SSTs that forced precipitation variability in that region.

5.4. Discussion

The AMO, NAO, PDO, and the SO were found to be contributors to decadal to multi-decadal variability of streamflow in the Mid-Atlantic region of the US. The wavelet

coherence analysis identified the PDO and the SO as being the main regulators of the enhanced 26-year streamflow and precipitation variability identified in the wavelet power analysis. As noted by Mantua (2010), the PDO is most energetic at periods of 15-25 years and 50-70 years, which could explain why the PDO is most coherent with precipitation at a period of 22-24 years. Whitney (2010) speculated that the NAO could be responsible for the 26-year periodicity. However, our analysis clearly shows that the NAO plays a minor role while the PDO and the SO play a major role. Variability of Susquehanna River streamflow at low-frequencies was also noted by Labat (2008), who revealed statistically significant periodicities at periods of 3, 12, and 27 years using a global wavelet analysis. Though Labat (2008) and Whitney (2010) did not analyze precipitation, the findings from this study suggest that the 26-27 year streamflow periodicities identified in these studies were driven by precipitation changes related to the SO and PDO, and to a lesser extent, the NAO.

Labat (2010), using a cross-wavelet analysis, determined that the PDO was driving North American continental freshwater discharge variability at a period of 22 years, a result similar to that found in this study (despite the very different study regions considered). However, this study added the phase relationship. Labat (2010) found the SO to have a weaker relationship at 21 years with North American continental freshwater discharge. It is important to note that, while the PDO and ENSO are regarded as distinct phenomena, the PDO is dependent on ENSO. According to Newman et al. (2003), the PDO may be regarded as the reddened response of atmospheric noise and ENSO, which could be problematic when separating pure ENSO and PDO signals in North American drought and climate proxies. Numerous studies have analyzed the connection between North Pacific SSTs and ENSO (Bjerknes, 1969; Luksch and Storch, 1992; Trenberth and Hurrell, 1994; Nath and Lau, 1996), further supporting the similarity between the streamflow-PDO and streamflow-SO relationships as identified in our wavelet coherence analysis. According to Latif and Barnett (1996), SST anomalies in the North Pacific undergo a 20-year oscillation involving the interaction between the Aleutian Low and the subtropical gyre. This model suggests that a positive SST anomaly in the North Pacific resulting from an anomalously strong subtropical gyre weakens the Aleutian Low, which alters air-sea heat fluxes, reinforcing the initial SST anomaly set up by the subtropical gyre. Changes in the wind

stress curl associated with the atmospheric response to the SST anomalies, however, act to spin down the subtropical gyre, reversing the initial temperature anomaly, with the oceanic response lagging the changes in wind stress curl. It is the interaction between the Aleutian Low and subtropical gyre that generates a 20-year oscillation, which coincides with the period at which the PDO is most coherent with Northeast US streamflow and precipitation, providing a possible physical mechanism for the PDO-streamflow teleconnection.

The similarity between the streamflow-SO and the streamflow-PDO relationships can be interpreted using the "atmospheric bridge" concept (Alexander, 2002). The essential component of the atmospheric bridge is the anomalously strong Aleutian Low during El Niño phases (Bjerknes, 1969), which alters heat fluxes, Ekman transport, and the ocean mixed layer depth, resulting in negative SST anomalies in the central North Pacific (Nath and Lau, 1996). Low SSTs in the central North Pacific contribute to a positive phase of the PDO and thus similarities between the streamflow-PDO and streamflow-SO relationships are expected.

A common theme from the wavelet analysis is the relationship between streamflow and four dominant modes on the ~ 20 year timescale. Even the less-coherent streamflow-NAO and streamflow-AMO relationships have a preferred timescale similar to the SO and PDO. Perhaps this is due to the link of the AMO and NAO with ENSO. For example, several studies have related NAO variability to SO variability and Atlantic Ocean SSTs to ENSO through Ekman pumping, ocean dynamics, and surface heat fluxes (e.g., Alexander et al., 2002). Furthermore, some studies suggest that the AMO may also influence the NAO and the PDO (Higuchi et al., 1999; Czaja and Frankignoul, 2002; Zhang and Delworth, 2007). Finally, solar activity may be responsible for the similarity in the wavelet coherence spectra. For example, Velasco and Mendoza (2008) found that solar activity influences the AMO, PDO, SO and the NAO at a 22-year timescale. Hence it is not surprising that a similar timescale was found for all climate indices in the wavelet coherence analysis. However, the SO and PDO have much stronger coherence with Mid-Atlantic streamflow than the NAO and AMO do.

Our finding on the connection between the SO and the 1960s drought contrasts with that of Seager et al. (2012), who attributed it to internal atmospheric variability. The

contrasting results may be due to differences in statistical approaches; Seager et al. (2012) only considered the correlation between Northeast US seasonal precipitation and North American precipitation and concluded that Northeast precipitation was only locally correlated so that teleconnections were unlikely during any season. In contrast, in the present study, a correlation pattern suggestive of a teleconnection pattern was identified at a period of 20 years, where precipitation in the Southwestern US fluctuated coherently with Northeast precipitation. Another potential problem in Seager et al. (2012) is the exclusive use of standard correlation coefficients. A standard correlation coefficient quantifies the linear relationship between two time series on all timescales simultaneously, ignoring the fact that the phase relationships between two time series at one timescale can be opposite to a phase relationship at a another timescale. Figure 5.12, as an example, shows that at a period of 20 years the SO is in-phase with DRB precipitation, whereas at period of 10 years it is out-of-phase with precipitation. Effects at the two timescales could cancel each other, resulting in a low overall correlation coefficient, even though the relationships are strong at particular timescales.

Seager et al. (2012) also base their conclusions on atmospheric global circulation model (GCM) simulations with prescribed SSTs, but uncertainties in the representation of convection in GCMs can manifest as false simulations of precipitation, particularly in terms of spatial patterns and variance (Peters et al., 2013). The uncertainties in convection and cloud feedback were also noted by Randall et al. (2007) and represent some of the largest problems in the latest GCMs. Given that the Northeast US given is relatively small, an accurate determination of the spatial pattern of precipitation is critical; small errors in the locations can produce large differences among GCM ensembles. The 1930s drought, for example, can be reproduced by forcing the GCM with SSTs during that period but the drought center is too far south (Cook et al., 2008). The statistical analysis of historical data in this study provides compelling evidence that the 1960s drought was externally forced, with precipitation anomalies fluctuating coherently with those across the Southwestern US, a region where tropical influences on drought have already been documented (Seager et al., 2005).

5.5. Conclusions

Streamflow, precipitation, and temperature variability in the Mid-Atlantic region of the US was analyzed and linked to dominant modes of climate variability at annual to multi-decadal timescales. The influence of climate modes on precipitation, temperature, and streamflow were found to vary from month to month and from season to season. Moreover, correlation coefficients were generally similar for the three river basins considered here. The streamflow-climate mode relationships, precipitation-climate mode relationships, and temperature-climate mode relationships can be physically attributed to alterations in storm tracks, jet-stream positions, and prevailing winds, all of which affect the air mass characteristics across the Mid-Atlantic region. Both streamflow and precipitation showed significant variability at a period of 26 years as revealed by wavelet power spectra. Moreover, such periodicities could not be attributed to red-noise; rather, it was the result of ENSO and the PDO. The SO explains 37%-54% of the streamflow decline across the Mid-Atlantic region during the 1960s drought, 33-49% of the 1970s pluvial, and 19- 50% of the 2000s pluvial. It is therefore important for hydrological and climate studies to include such oscillations when assessing future impacts of climate change on the hydrology and ecology of the Hudson, Susquehanna, and Delaware watersheds and their receiving estuaries. It is hoped that the results from this study will aid hydrological and climate forecasts so that predictions about the future state of hydrological systems such as rivers and estuaries under a changing global climate system can be more accurately assessed.

Chapter 6

Salinity and Streamflow Variability in the Mid-Atlantic Region of the United States and its Relationship with Large-scale Atmospheric Circulation Patterns

6.1. Introduction

Estuaries are complex, diverse aquatic systems of ecological, economic, industrial, and recreational importance. Estuaries contain seas ports and shipping channels through which commodities are transported. Local economies are supported through fishing activities in the estuaries, which contain diverse aquatic species. Increased human development around the estuaries, however, has negatively impacted the health of estuaries. For example, overfishing has threatened many species residing in estuaries but efforts have been made to mitigate overfishing. Other anthropogenic impacts include: pollution, agricultural waste, and dredging (Najjar et al., 2010; Kreeger et al., 2010). Human waste, as an example, resulted in low levels of dissolved oxygen in the Delaware River in the early 1900s but has since recovered as a result of substantial cleanup efforts (Albert, 1998). Agricultural waste also harms estuaries because it often contains nitrogen, which increases the prevalence of toxic algal blooms.

The health of estuaries also responds to natural fluctuations. Streamflow, for example, impacts many estuarine parameters, such as the circulation, vertical stratification, eutrophication, and hypoxia. The primary role of streamflow is to drive the exchange circulation, whereby fresher water at the surface flows seaward and saltier water flows at depth flows landward (MacCready and Geyer, 2010). If the circulation becomes strong, then a vertical salinity gradient can be generated, which may lead to hypoxia. Streamflow, for example, has been shown to cause summer vertical salinity stratification and hypoxia in the Chesapeake Bay (Hagy et al., 2004) and to promote eutrophication in the Hudson River (Howarth et al., 2000). Hypoxia and eutrophication are serious problems in estuaries throughout the world and have been related to declines in numerous aquatic species (Diaz and Rosenberg, 2008), such as zooplankton in the Chesapeake Bay (Kimmel and Roman, 2004; Miller et al., 2004). Thus, to understand changes in many historical estuarine

parameters demands knowledge of how streamflow varied historically. If the causes of streamflow variability are better understood, then assessments about the future health of estuaries can be made.

Streamflow in the Northeast region of the United States has fluctuated substantially, resulting in, for example, droughts, but the causes of such fluctuations are subject to debate. Historical streamflow variability has been related to large-scale climate indices such as the El-Niño/Southern Oscillation (ENSO), where the streamflow decline during the 1960s drought, for example, may have been partially forced by the Southern Oscillation (SO; Schulte et al., 2015). Seager et al. (2012), on the other hand, argued that the decline was the result of intrinsic atmospheric variability. During the 1960s drought, salinity in the Delaware River was anomalously high and created conflicts between New York City and Philadelphia water management agencies. The potential for salinity intrusions in the Philadelphia area water supply also led to the monitoring of the salt front position in Delaware River (Hull and Titus, 1986). Najjar et al. (2010), using climate models, determined that warming in response to greenhouse gas forcing is likely across the Delaware, Hudson, and Susquehanna River basins. The climate model results, however, showed disagreement with precipitation projections and suggested that precipitation increases are still in the range of interannual variability. The uncertainties in the precipitation projections suggest that climate factors impacting historical precipitation variability need to be better understood.

The ocean also impacts estuaries through changes in sea level (Ross et al., 2015) and meandering of the Gulf Stream (Lee and Lwiza, 2008). Through the analysis of long-term salinity observations and numerical modeling, sea-level rise was found to lead to measureable increases in salinity in the Chesapeake (Hilton et al., 2008) and Delaware (Ross et al., 2015) Bays. These studies suggested that continued sea-level rise will lead to substantial salinity increases in these estuaries throughout the 21st century. Lee and Lwiza (2008) found fluctuations in the Gulf Stream position to influence Chesapeake Bay salinity, with a more poleward location being associated with higher salinity. Taylor et al. (1998) found shifts in the Gulf Stream position to follow ENSO events, with the SO leading by two years. The ENSO-Gulf Stream linkage would then suggest that historical salinity

variability was influenced by ENSO-related changes in mid-latitude atmospheric circulation patterns. However, Lee and Lwiza (2008) found no link between the SO and Chesapeake Bay bottom salinity.

In this study, monthly salinity variability in the Hudson River, Delaware, and Chesapeake Bay estuaries will be investigated using wavelet analysis, which identifies the dominant frequencies of variability and how they change with time. The use of daily streamflow data in addition will allow salinity variability to be linked to daily weather events occurring on synoptic timescales. The previous work of Whitney (2010) and Lee and Lwiza (2010) used monthly data, which may have precluded the extraction of important physical phenomena occurring on daily timescales. In particular, the mid-latitude response to tropical convection occurs on daily timescales and this fact underscores the need for an analysis of daily streamflow data.

The paper is organized as follows. The data used are described in Section 6.2 and a brief review of wavelet analysis is presented in Section 6.3. In Section 6.4, results from a composite analysis of daily streamflow anomalies with atmospheric fields are presented and results from the wavelet analysis are also presented. The section concludes with an analysis of fluctuations in the Gulf Stream position and linkages with tropical convection and the overlying atmospheric conditions. Implications for historical salinity variability are also discussed. Concluding remarks are presented in Section 6.5.

6.2. Data

In this study, streamflow and estuarine salinity data are analyzed to determine the main climate drivers of their variability. Several common climate indices are considered in this study: the North Atlantic Oscillation (NAO), Pacific-North American teleconnection pattern (PNA), the Gulf Stream, and the Niño 4 indices. Meteorological fields such as mean sea level pressure (MSLP) and 300-hPa streamfunction are also analyzed to investigate their relationships with streamflow and salinity conditions. The upper tropospheric streamfunction was chosen for this study because circulation anomalies in the mid-latitudes often arise from the propagation of Rossby waves from the tropics (Hoskins and Karoly, 1981) and the Rossby waves are most prominent in the upper troposphere.

6.2.1. Salinity and Salt Front Position

Chesapeake Bay salinity data (Figure 6.1) for 1985-2013 were obtained from the Chesapeake Bay Program (CBP), which samples the Bay approximately twice monthly. Tidal fluctuations and weather events can produce rapid changes in salinity, but these are damped in the mesohaline region of Bay, making the infrequent sampling suitable for estimating the monthly average in this region (Gibson and Najjar, 2001). We thus limit ourselves to the analysis of salinity data in the upper 5 m of the water column of the central bay as shown in Figure 6.1 (stations CB 4.1C, CB 4.1E, CB 4.1W, CB 4.2C, CB 4.2E, CB 4.2W, CB 4.3C, CB 4.3E, CB 4.3W, and CB 4.4), which were spatially and temporally averaged for each month of the time series.

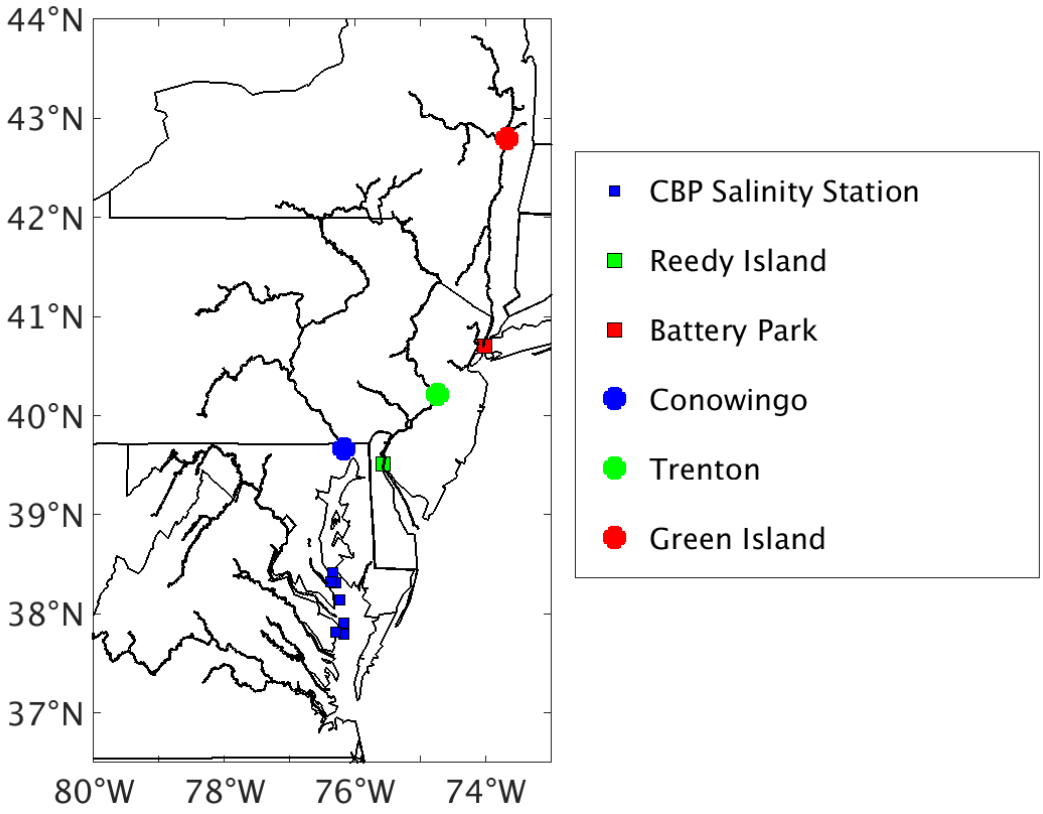

Figure 6.1. The locations of CBP stations, USGS gauging stations, Reedy Island, and The Battery, New York.

Daily Reedy Island conductance data for 1985-2013 were obtained from the United States Geological Survey (USGS). The daily conductance was converted into practical salinity units (psu) as describe by Lewis and Perkin (1981) and the daily values were

converted into monthly means. For 6% of the months, there were greater than 15 days with missing data and so monthly means were not computed for those months. However, after the monthly means were calculated the months with missing values were filled using a linear interpolation.

Daily Hudson River salt front position data for 1991-2013 were obtained from the USGS and were converted into a monthly anomaly time series by first computing monthly averages and then removing the mean annual cycle. Salt front position was calculated using a threshold of 100 mg L^{-1} chloride concentration and represents the location, in river kilometers, above The Battery, New York City (Figure 6.1). A daily value for the time series is the most upstream location at high slack tide for the previous day. The Hudson salt front time series was extended back to 1985 by linear regressing the monthly anomaly time series with monthly Waterford streamflow anomaly (see below). The extended time series allows for a more consistent comparison with salinity variability of the Delaware and Chesapeake Bay estuaries and also reduces the impact of edge effects inherent in wavelet analysis (Section 6.3).

6.2.2. Streamflow

Daily streamflow data for the Susquehanna River at the Conowingo Dam (01578310), the Delaware River at Trenton (01463500), New Jersey, and the Hudson River at Waterford, New York (01335754) were obtained from the USGS for the period 1985-2014. The data were converted to daily anomalies by subtracting off the mean annual cycle.

6.2.3. Meteorological Data

Daily European Center for Medium Range Forecast reanalysis (ERA) 300 hPa streamfunction and mean sea-level pressure data were used. The data, on a $2^{\circ} \times 2^{\circ}$ grid, were converted into daily anomalies by removing the mean annual cycle.

6.2.4. Climate Indices

Daily indices for the North Atlantic Oscillation (NAO) and the Pacific-North American teleconnection pattern (PNA) were obtained from the Climate Prediction Center (CPC). The NAO represents the coherent fluctuations in sea-level pressure (SLP) over the

North Atlantic Ocean. Its index is defined as the difference between the normalized SLP anomalies at Lisbon, Portugal and Stykkisholmur, Iceland (Hurrell, 1995). During the positive phase of the NAO, the Jetstream over the Atlantic Ocean is stronger than normal and results in warmer-than-normal conditions across the Northeast US during the winter. The PNA index is defined as leading Empirical Orthogonal Function of Northern hemisphere 500 hPa geopotential height between 20°N and 90°N (Barnston and Livezey, 1987). A positive phase is associated with a ridge of high pressure over the western US and a trough of low pressure over the eastern US.

The monthly Niño 4 index data were obtained from the CPC and represent the strength and evolution of ENSO. The index represents the average sea surface temperatures in the region bounded by 5°S-5°N and 160°E-150°W. The Gulf Stream Index (GSI) data for 1985-2013 were obtained from http://www.pml-gulfstream.org.uk/. The index captures the meandering of the Gulf Stream about its mean position, where a positive value indicates that the north wall of the Gulf Stream is poleward of its mean position and a negative value indicates that it is equatorward of its mean position.

6.3. Methods

6.3.1 Wavelet Analysis

To compute the wavelet transform of time series, one needs to choose a wavelet function. The Morlet wavelet was chosen for this study and is given by

$$\psi_0(\eta) = \pi^{-1/4} e^{i\omega\eta} e^{-\frac{1}{2}\eta^2}, \tag{6.1}$$

where ω is the dimensionless frequency and η is the dimensionless time. The reason for choosing the Morlet wavelet is that it provides a balance between time and frequency localization and is thus recommended for identifying features of geophysical time series (Grinsted et al., 2004). The wavelet transform of a time series $X = (x_n; n = 1, \ldots, N)$ is given by

$$W_n^X(s) = \sqrt{\frac{\delta t}{s}} \sum_{n'=1}^{N} x_{n'} \psi_0[(n' - n)] \frac{\delta t}{s}], \tag{6.2}$$

where δt is a uniform timestep (one month or one day in this study), s is the scale of the Morlet wavelet, and $\eta = s \cdot t$. The more traditional Fourier period λ is approximately related to the wavelet scale by $\lambda = 1.03s$. The wavelet power at a given scale and time is then given by $|W_n^X(s)|^2$. The pointwise significance of each wavelet power coefficient can be test invidually against a red noise background, where it is assumed that wavelet power coefficients are independent. The reader is referred to Torrence and Compo (1998) and Grinsted et al. (2004) for a more detailed discussion of the theory of wavelet analysis and pointwise significance testing used in this paper.

The relationships between two time series was quantified using a wavelet coherence analysis. The (local) wavelet coherence between two time series *X* and *Y* is given by

$$R_{n,xy}^2(s) = \frac{\left|S(s^{-1}W_n^{XY}(s)\right|^2}{S(s^{-1}|W_n^X(s)|^2)S(s^{-1}|W_n^Y(s)|^2)}, \tag{6.3}$$

where $W_n^{XY}(s)$ is the cross-wavelet transform, defined as the product of the wavelet transform of *X* and the complex conjugate of the wavelet transform of *Y*, and *S* is a smoothing operator defined by $S(W_n^X(s)) = S_{scale}\left(S_{time}\left(W_n^X(s)\right)\right)$. S_{time} represents smoothing in time and S_{scale} is smoothing along the wavelet scale axis. A coherence value of 0 signifies that the two time series are independent and a coherence value of 1 indicates that the two time series are linearly related at the given frequency and time. The statistical (pointwise) significance of wavelet coherence is obtained using Monte Carlo methods by: (1) generating a large number of synthetic data pairs with the same lag-1 autocorrelation coefficients as the input time series, (2) calculating the wavelet coherence for each pair, and (3) estimating the significance level at each scale using values outside the cone of influence (Grinsted et al., 2004). A more detailed discussion of wavelet coherence can be found in Grinsted et al. (2004).

It will also be useful to calculate the wavelet coherence between two time series when their common dependence is removed. Such a calculation requires partial wavelet coherence (Ng and Chan, 2012), which is defined analogously to a partial correlation

coefficient; for two time series *X* and *Y* that are both related to a time series *Z*, the partial wavelet coherence is defined as

$$PR^2_{n,Z}(s) = \frac{\left|R^{xy}_n(s) - R^{xy}_n(s)R^{xy}_n(s)^*\right|^2}{\left(1 - R^2_{n,xz}(s)\right)\left(1 - R^2_{n,yz}(s)\right)}, \tag{6.4}$$

where

$$R^{xy}_n(s) = \frac{S(s^{-1}W^{XY}_n(s))}{\sqrt{S(s^{-1}W^{X}_n(s))S(s^{-1}W^{Y}_n(s))}}., \tag{6.5}$$

and the asterisk denotes the complex conjugate. If the influence of *Z* on either *X* or *Y* is strong, then the wavelet partial coherence will differ from the wavelet coherence. Statistical significance can be estimated using Monte Carlo methods in the same way as wavelet coherence.

It is known that many pointwise significant results found in local spectra may be artifacts of multiple testing, arising from the large number of wavelet coefficients to which the pointwise significance test is applied (Maraun et al., 2007; Schulte et al., 2015). The problem can be remedied by applying an areawise test (Maraun et al., 2007), a geometric test (Schulte et al., 2015), or a cumulative areawise test Schulte (2015). In this study, the cumulative areawise test (referred to as the areawise test, hereafter) was applied because it was found to have greater statistical power than the geometric test. The areawise test assesses the cumulative areas of so-called pointwise significance patches, which are contiguous regions of pointwise significance. In other words, the area of a patch is tracked as the pointwise significance level increases from α_{min} to α_{max}. The cumulative areas of patches corresponding to the input time series are then compared to that of a null distribution of cumulative areas associated with red-noise processes. In the present study, $\alpha_{min} = 0.82$, $\alpha_{max} = 0.98$, and the spacing between pointwise significance levels was 0.02. These choices of parameters were used because they were found to be associated with high statistical power (Schulte, 2015).

An additional topological method will also be applied to wavelet power spectra. The method was introduced by Schulte et al. (2015) and can provide additional insight into

the time series in question. The idea of the procedure is to examine holes located in patches, which are similar to those found in annuli. These holes can arise from time series containing nonlinearities and can also arise when two nearby frequency components of a time series have associated wavelet power exceeding a red-noise background (Schulte et al., 2015). For pointwise significance levels of 0.1%, 1%, and 5%, holes are infrequently present in patches generated from red-noise processes so that patches containing holes in the wavelet power spectrum under consideration are less likely to result from stochastic fluctuations. In fact, according to Schulte et al. (2015), the probability of finding a hole at the 5% significance level is 3.4×10^{-4} and virtually zero at the 0.01 and 0.001 pointwise significance levels. It is therefore useful to project the centroids of holes existing at the 0.1%, 1%, 5% significance levels onto the wavelet domain, resulting in a topological wavelet diagram.

Characteristic timescales reported in this study were identified using global wavelet spectra, the time-averaged representations of the sample local wavelet spectra. For global power, periods of maximum time-averaged power were considered the dominant timescale of variability, where global power is calculated by averaging $|W_n^X(s)|^2$ over all time indices for a fixed scale. For time-averaged coherence, an alternative definition to Eq. (3) was used, which is given by (Elsayed, 2006)

$$G_C(s) = \frac{|W^{XY}(s)|^2}{\left(\sum_{n=1}^{N}|W_n^X|^2\right)\left(\sum_{n=1}^{N}|W_n^Y|^2\right)}, \tag{6.6}$$

where

$$W_n^{XY}(s) = \sum_{n=1}^{N} W_n^X(s) W_n^{Y*}(s), \tag{6.7}$$

Equation (6.6) measures the coherence between two time series in the entire study period at a scale s. Statistical significance of $G_C(s)$ was computed using Monte Carlo methods as follows: red-noise time series with the same lengths and lag-1 autocorrelation coefficients as the two input data series were generated and $G_C(s)$ was computed for each pair of red-noise processes. The resulting distribution of $G_C(s)$ at each scale was then used to estimate the significance of the global coherence estimates. A global partial coherence spectrum can also be defined in the same way as global wavelet coherence by replacing the smoothing

operators in Eq. (6.5) by summations. Statistical significance was also computed using Monte Carlo methods by generating 2000 triples of red-noise processes with the same lag-1 autocorrelation coefficients as the input time series.

One can also sum the wavelet coherence at each time over scale. The result is scale-averaged coherence measuring the linear correlation between two time series as function of time. The temporal resolution of the correlation time series is determined by the Nyquist frequency, which in this case corresponds to a period of two months for the monthly data. Statistical significance can be estimated by generating red-noise time series with lag-1 autocorrelation identical to the input time series and calculating the scale-averaged coherence for each time.

6.3.2 Estimation of Confidence Intervals

The significance of global spectral peaks was further assessed by developing a block bootstrap procedure (Kunsch, 1989), which was chosen because adjacent wavelet coefficients are correlated and the correlation becomes more pronounced for large periods. The application of the traditional bootstrap would result in confidence bounds that are too narrow and thus statistical significance of quantities may be overestimated. The basic idea behind the procedure is to divide the wavelet coefficients at each period into overlapping blocks or segments and sample with replacement the blocks to create a synthetic set of wavelet coefficients. The synthetic set of wavelet coefficients should have the same number of elements as the original set of wavelet coefficients and can be used to calculate a bootstrap replicate of the desired wavelet quantity. If the procedure is performed a large number of times, a confidence interval can be estimated from the distribution of bootstrap replicates.

Mathematically, to find the approximate $100(1-\beta)\%$ confidence interval of a global coherence estimate, the set of wavelet coefficients $W^X(s) = \{W_1^X(s), W_2^X(s)\ , \dots, W_N^X(s)\}$ at each scale is divided into $M = N - l + 1$ overlapping blocks of length l. For example, block 1 for $W^X(s)$ is $B_1(s) = \{W_1^X(s), W_2^X(s), \dots, W_l^X(s)\}$, block 2 is $B_2(s) = \{W_2^X(s),\ W_3^X(s), \dots, W_{l+1}^X(s)\}$, and the ith block is $B_i(s) = \{W_i^X(s), W_{i+1}^X(s), \dots, W_{l+i-1}^X(s)\}$. The overlapping blocks form a

collection of M blocks $B(s) = \{B_1(s), \ldots, B_M(s)\}$ at each scale. To obtain a bootstrap sample of wavelet coefficients, k blocks from each collection at a given scale are randomly sampled with replacement. It is noted that there is a probability $1/M$ of being drawn from their respective collections. The concatenation of the randomly sampled blocks $B_1'(s)$, $B_2'(s)$, … , $B_k'(s)$ forms a new set of wavelet coefficients $W_1'(s),\ W_2'(s), \ldots, W_m'(s)$ with $m = k\ l$.

A known problem with the block bootstrap procedure is the selection of the block length l, though it has been suggested that $l \to \infty$ and $l/n \to 0$ as $n \to \infty$. (Wilks, 1997). In this study, the optimal block length was estimated using Monte Carlo methods as follows: 2000 surrogate pairs of red-noise processes with N = 1000 and lag-1 autocorrelation coefficients equal to 0.5 were generated, (2) the global coherence between each of the 2000 pairs was calculated at all scales, and (3) the 95% confidence intervals of the global coherence estimates at each scale were calculated. The widths of the 2000 confidence intervals at each scale were then computed to generate a distribution of widths at each scale and the mean of the widths was also calculated. The goal of the calculations was to find the largest c such that $l = N^c$ produced the widest confidence intervals. The experiment was repeated for different values of c and it was determined that c = 0.7 generally produced the widest confidence intervals for all scales in both global power and coherence spectra. If $c < 0.7$, then the blocks were not large enough at large periods to account for the large auto-correlation among the wavelet coefficients. On the other hand, if $c > 0.7$, the blocks were too large for all periods and thus the randomization of the bootstrap method was compromised. Thus, c = 0.7 was used in this study. It may happen that N^c is not integer-valued, in which case the block lengths were chosen to be the greatest integer less than N^c. Moreover, if the concatenation of blocks of length l resulted in a set of wavelet coefficients whose length was $L < N$, a randomly sampled block with length $N - L$ was added to the set to make the length equal to N.

6.4. Results

6.4.1 Salinity and Streamflow Time Series

The monthly streamflow and salinity time series are shown in Figures 6.2a, 6.2c, and 6.2e. Periods of high streamflow were associated with periods of low salinity for all three estuaries. Moreover, Figures 6.2b, 6.2d, and 6.2e indicate that salinity and streamflow were highly coherent, though the strength of the relationship fluctuated. For example, from 1985 to 1990, Chesapeake Bay salinity and Susquehanna streamflow were less coherent relative to subsequent periods. Note that the large scale-averaged coherence for the Hudson salt front position from 1985 to 1991 (not shown) is an artifact of how the Hudson River salt front position data were extended back to 1985 using a linear regression.

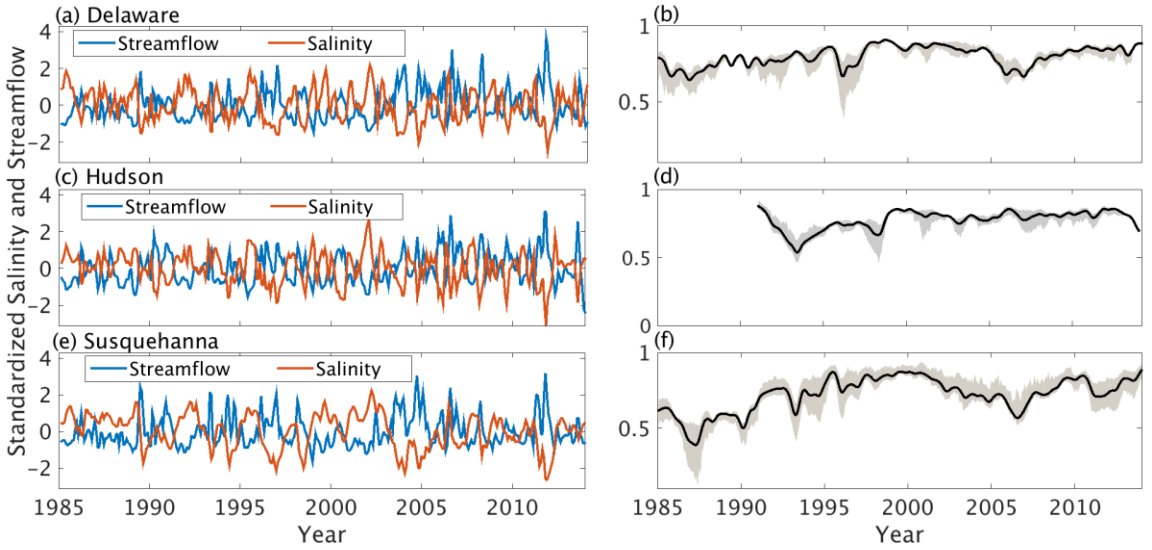

Figure 6.2. (a) Standardized mean monthly Delaware Streamflow and Reedy Island salinity and (b) the scale-averaged coherence between the two time series. Gray shading represents the 95% confidence interval. (c) – (d) Same as (a) – (b) except for Hudson River streamflow and Hudson salt front position. (e) – (f) Same as (a) – (b) except for Susquehanna River streamflow and Chesapeake Bay salinity. Time series were standardized by dividing the monthly means by the standard deviation for that month.

The wavelet power spectra of daily streamflow for the Delaware, Hudson, and Susquehanna Rivers are shown in Figure 6.3. Statistically significant power was found in the 0.1-0.9 month period band, representing streamflow variability resulting from weather events. However, not all weather events produced streamflow variability exceeding a red-noise background. For example, there is a lack of significant results in the 0.1-0.9 month period band around 1993 and 2002 for the Delaware and Susquehanna Rivers, which coincide with two dry periods, as shown in Figure 2. For the Delaware and Hudson Rivers, a region of areawise significance was also located at a period of 41 months, which extended nearly throughout the entire study period for the Delaware River. The global power spectra

indicate that streamflow for all three rivers exhibited significant variability at a period of 41 months. The low-frequency peaks may have resulted from the modulation of daily weathers events, as investigated in Section 6.4.2.

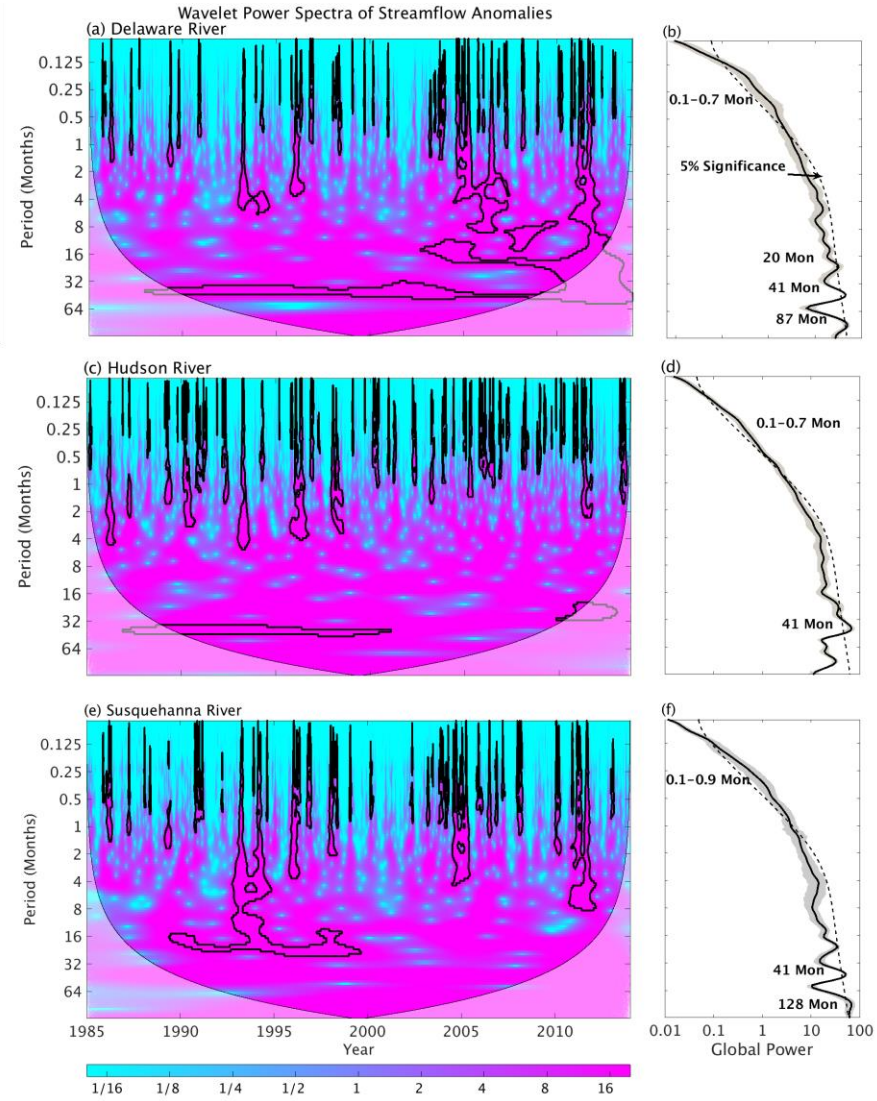

Figure 6.3. (a) Wavelet power spectrum of Delaware River streamflow. Contours enclose regions of 5% areawise significance. Light shading represents the Cone of Influence. (b) The corresponding global wavelet power spectrum. Graying shading represents the 95% confidence interval. (c) – (d) Same as (a) – (b) except for Hudson River streamflow. (e) – (f) Same as (a) – (b) except for Susquehanna River streamflow.

The topological wavelet diagrams for the three rivers also suggest that there was significant power in the 0.1-0.9 month period band (Figure 6.4). For all three diagrams, the holes were

mainly located at periods less than 2 months, which is consistent with the wavelet power spectra showing most of the significant features at the same periods. Therefore, the features in the 0.1-0.9 period band are likely to be deterministic, recalling that the holes at the 0.01 and 0.001 pointwise significance levels indicate with virtual certainty that the wavelet power did not arise from red noise (Section 6.3.1). The results from both the wavelet power analysis and topological analysis suggest with high confidence that the variability in the 0.1-0.9 period band was forced and likely due daily weather events (Section 6.4.2).

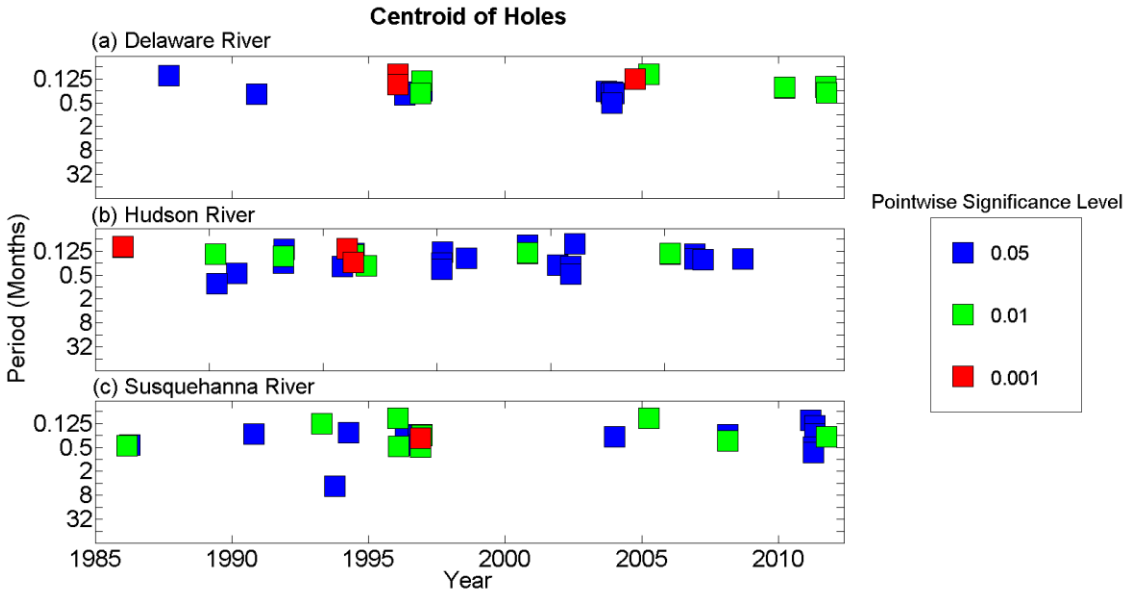

Figure 6.4. Topological wavelet diagrams corresponding to the wavelet power spectra of streamflow for the (a) Delaware, (b) Hudson, and (c) Susquehanna Rivers. Blocks indicate the centroids of holes at the pointwise significance level corresponding to the color.

The wavelet power spectra of monthly salinity anomalies are shown in Figure 6.5. For the Delaware and Hudson Rivers, a few significant results were found from 2005 to 2010 in the 4-16 month period band. The wavelet power spectrum of Chesapeake Bay salinity indicates that significant power was present in the 8-16 month period band, though the global wavelet power spectrum indicated that the power was insignificant. The regions of enhanced power were found to generally correspond to the regions of enhanced power in the wavelet power spectra of streamflow. The similarity was particularly evident by inspecting the global wavelet spectra, where enhanced but not significant power was identified at periods of 18, 41 months, and 88 months.

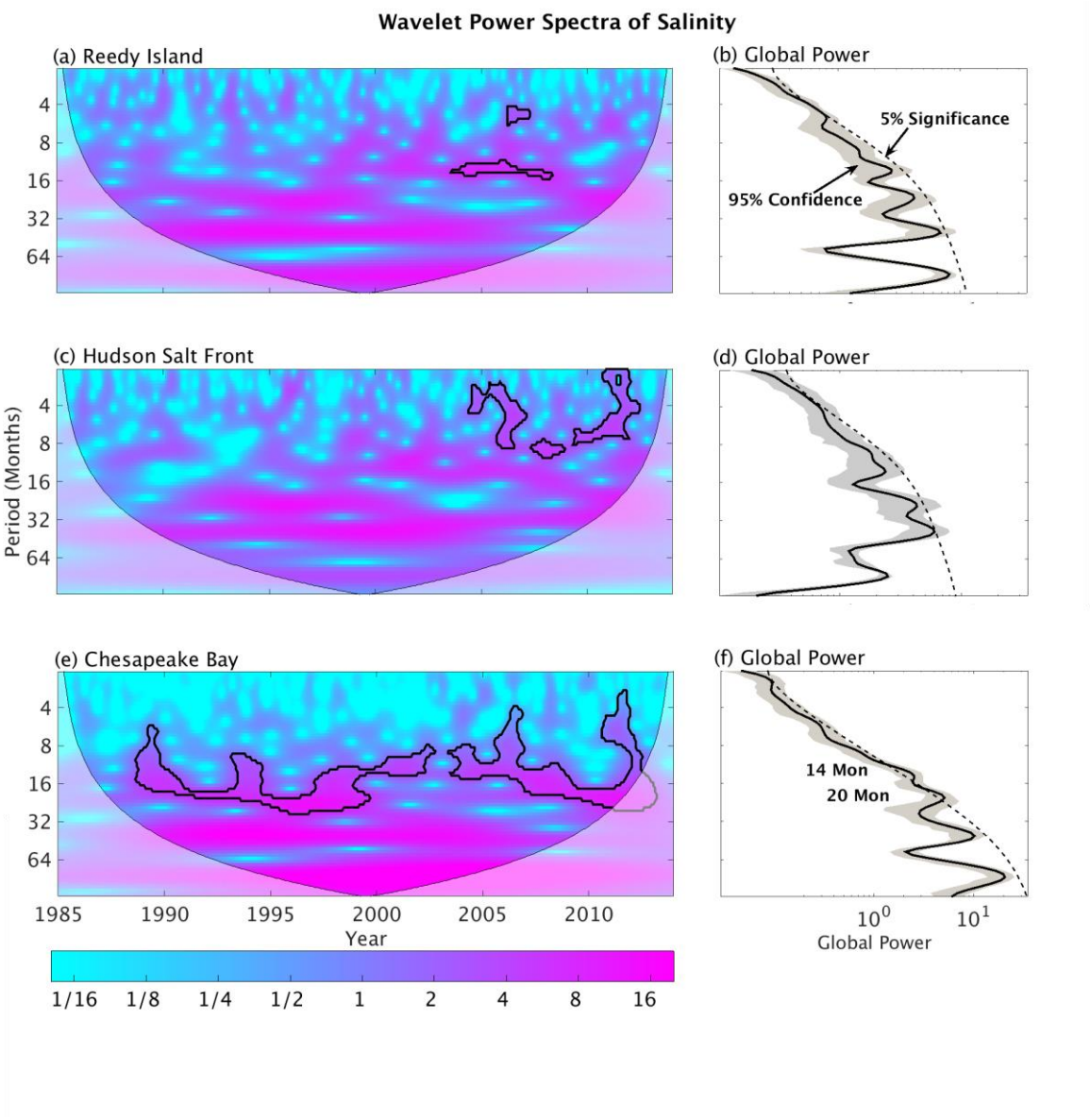

Figure 6.5. (a) Wavelet power spectrum and (b) corresponding global wavelet spectrum of Reedy Island salinity. (c) – (d) Same as (a) – (b) but for the Hudson River salt front position. (e) – (f) Same as (a) – (b) except for Chesapeake Bay surface salinity. See Figure 6.3 for details of features found in the wavelet spectra.

The low-frequency peak at 88 months may be due to Gulf Stream position variability, as noted by Lee and Lwiza (2008). This idea is supported by how the wavelet power at 88 months is greatest for Chesapeake Bay salinity, which is most sensitive to changes in the Gulf Stream position. The Chesapeake Bay was located closer to the mean Gulf Stream position relative to the Delaware and Hudson estuaries so that small changes in the Gulf Stream position would have led to greater changes in Chesapeake Bay salinity. On the other hand, the global wavelet power spectra of Delaware and Susquehanna River

streamflow contained peaks at approximately the same timescales. Therefore, the enhanced variance found for salinity may have resulted from atmospheric forcing. It is therefore necessary to investigate the relationship between atmospheric fields and streamflow.

6.4.2 Composite Analysis

Figure 6.6 shows the composite 300-hPa streamfunction anomaly pattern prior and during anomalously high streamflow events. The anomalous high streamflow events were defined as those events exceeding the 75^{th} percentile of all daily flows, whereas anomalously low events were defined as those falling below the 25^{th} percentile. The statistical significance of the composite means were calculated using the two-sample student-*t*, where the confidence that the composite mean for the high streamflow event was statistically different from that of the low streamflow events was assessed. The test was applied at the 5% significance level in all cases.

Twenty days before the anomalous event (lag = -20 days), positive streamfunction anomalies were preferred over the East Coast of the US, with the greatest anomalies being centered around 60°W and 47°N. The anomaly center persisted through lag = 0. The Rossby wave train first emerged at lag = -16, where a series of alternating anomaly centers appeared, representing a series of upper-level high- and low-pressure systems. The pattern appears to be consistent with a typical mid-latitude upper atmospheric response to tropical convection (Horel and Wallace, 1981; Hoskins and Karoly, 1981). The coherent patterns of streamfunction anomalies were generally most pronounced at lag = -4 days. The reason why the most pronounced atmospheric anomalies preceded the anomalous streamflow events by 4 days was because streamflow typically lagged precipitation by 4 days. This lag relationship was corroborated by conducting a lag composite analysis of daily ERA precipitation with daily Delaware River streamflow. Similar lag composite analyses were conducted for the Hudson and Susquehanna Rivers and the results were nearly identical. The similarity among the results was not surprising, however; the atmospheric patterns analyzed had large spatial extents compared to that of the study region.

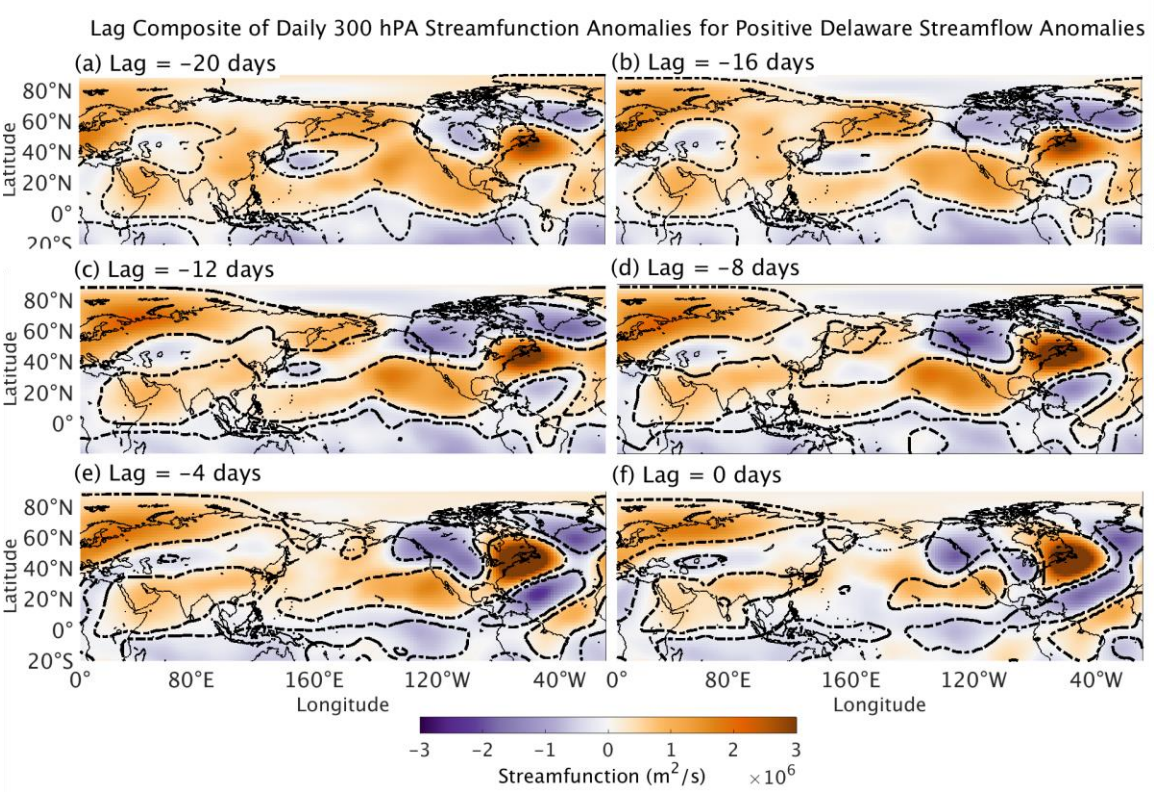

Figure 6.6. (a) – (f) Mean 300 hPa streamfunction anomalies for 75th percentile daily Delaware River streamflow events (a) 20 days, (b) 16 days, (c) 12 days, (d) 8 days, (e) 4 days, and (f) 0 days before the anomalous streamflow events. Contours enclose regions of 5% statistical significance as determined by the two-sample student-*t* test.

Figure 6.7 shows a lag composite analysis of daily MSLP anomalies associated with the high streamflow events. Similar to the 300-hPa streamfunction composite, a region of positive MSLP anomalies were centered around 60°W and 47°N and the feature persisted at all lags. A negative anomaly center was also located over the Southeastern US and also persisted at all lags. Another interesting feature is the negative MSLP anomalies over the polar region from lags -20 to -4 days. The negative anomalies imply that anomalously high streamflow events are related to a stronger-than-normal Arctic Oscillation days in advance of the anomalous events. The Arctic Oscillation represents the surface signature of fluctuations in the polar vortex aloft (Thomas and Wallace, 1999). A similar composite analysis was conducted for the Hudson and Susquehanna Rivers and the results were found to be nearly identical.

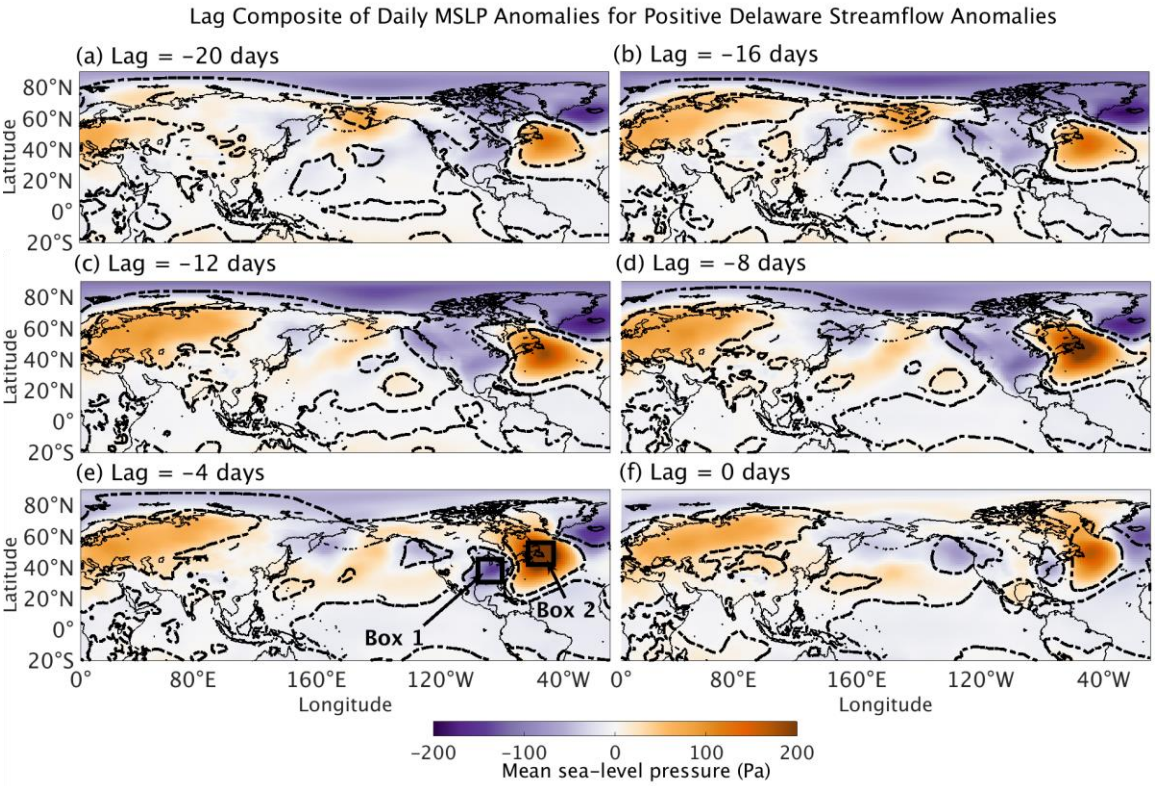

Figure 6.7. Same as Figure 6.6 except for MSLP from 1985-2013. The region used to calculate the Atlantic MSLP index is labeled Box 2 and the region used to calculate the Southeast MSLP index is labeled Box 1. Box 1 represents the region bounded latitudinally by 30°N and 40°N and longitudinally by 80°W and 95°W, while Box 2 represents the region bounded latitudinally by 40°N and 50°N and longitudinally 45°W and 60°W.

6.4.3 Correlation Analysis

A prominent feature in Figure 6.7 is the MSLP dipole situated over the Eastern US during and prior to anomalous streamflow events. This MSLP dipole-streamflow relationship is physical because the dipole pattern is conducive to precipitation over the Eastern US. Therefore, an index representing the strength and evolution of the dipole pattern may be a better indicator of the streamflow intensity than the standard teleconnection indices such as NAO and PNA. A MSLP index was thus constructed as follows: (1) Daily MSLP in Box 1 and in Box 2 shown in Figure 6.7e were spatially averaged; (2) the time series were normalized by dividing them by their respective standard deviations; and (3) the Eastern North American dipole index (ENA) was calculated using the formula

$$p_{dipole} = (MSLP)_{box2} - (MSLP)_{box1}, \tag{6.8}$$

where $(MSLP)_{box1}$ is the normalized MSLP anomaly corresponding to Box 1, and the same for Box 2. The larger the value of p_{dipole}, the greater the pressure difference between the anomaly centers. For the remainder of the paper $(MSLP)_{box2}$ will be referred to as the Atlantic MSLP index and $(MSLP)_{box1}$ will be referred to as the Southeast MSLP index.

The daily ENA index together with its wavelet power spectrum is shown Figure 6.8. The time series is rather noisy but the wavelet spectrum indicates that the index exhibited significant variability at a period of 7 days, timescales often associated with synoptic-scale weather. An interesting feature in the wavelet power spectrum is the alternating pattern of significant power and non-significant power. The significant power occurred during the winter when the atmosphere was more dynamically active and the deficits coincided with the summer when the atmosphere was less active. The wavelet power spectrum therefore suggests that the MSLP dipole is mainly a winter phenomenon. The global wavelet power spectra also detected enhanced power at a period of 41 months.

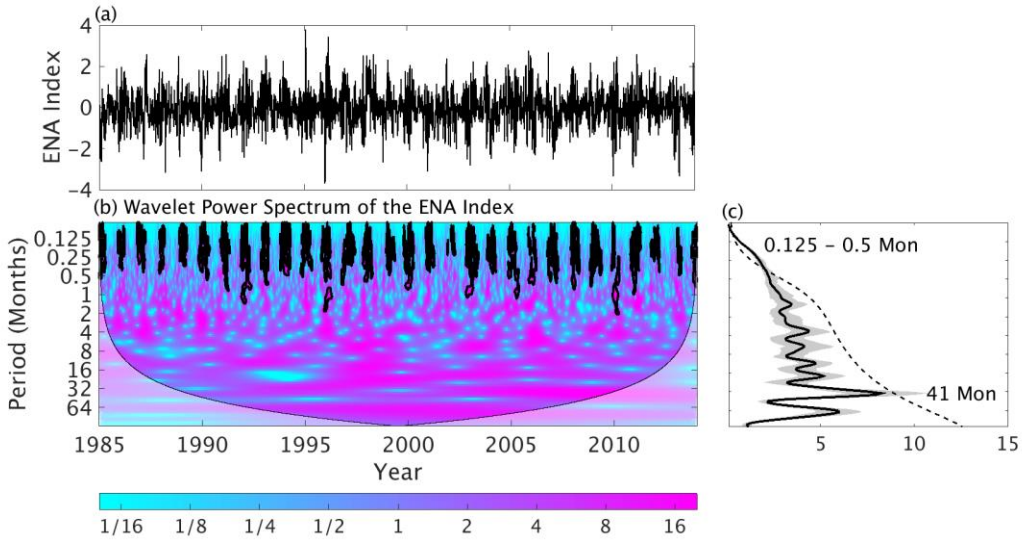

Figure 6.8. (a) 7-day running mean of the ENA index, (b) the wavelet power spectrum of the raw ENA index, and (c) the corresponding global wavelet power spectrum. See Figure 6.3 for details of features found in the wavelet power spectra.

Figure 6.9 shows the cross-correlation of streamflow for the Delaware, Hudson, and Susquehanna Rivers with the daily Southeast MSLP, Atlantic MSLP, ENA, and NAO indices. With the exception of the NAO index, the cross-correlation coefficients with streamflow peaked at lags of 2-4 days because streamflow lags precipitation by 2-4 days, as mentioned earlier. It is also noted that the ENA index was a better predictor of

streamflow than the other three indices. The results suggests that the daily ENA index should be monitored and forecasted in the same way as the NAO index is at the CPC. A similar cross-correlation analysis with the PNA index was also performed and the cross-correlation coefficients were found to be weaker than those associated with the NAO index.

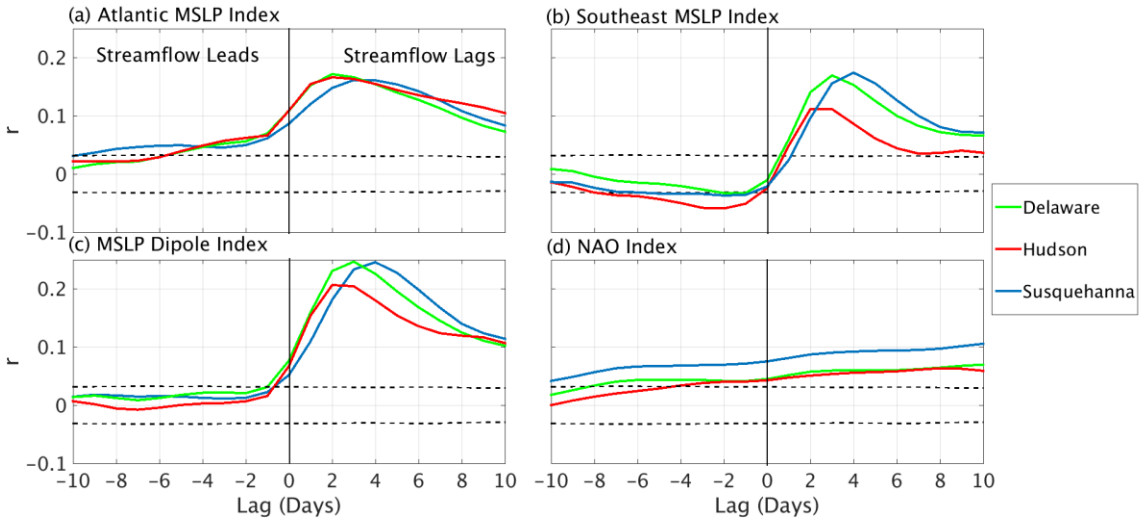

Figure 6.9. Cross-correlation of streamflow for the Delaware, Hudson, and Susquehanna Rivers with (a) Atlantic MSLP index, (b) Southeast MSLP index, (c) ENA index, and (d) the NAO index for 1985-2013. Dotted lines represent the 5% significance bounds.

The wavelet coherence between the ENA index and the daily flows for the Delaware, Hudson, and Susquehanna Rivers are shown in Figure 6.10. Numerous areawise significant results were found in the 0.8-5 month period band for all three rivers. For all three rivers, at least some areawise significant coherence was identified in the 11-41 month period band and thus the significant power shown in Figure 6.3 may have partially resulted from atmospheric forcing. There are also corresponding areawise significance regions in the wavelet coherence spectra for salinity (Figure 6.11). For example, enhanced coherence with both salinity and streamflow at a period of 92 months was identified and the coherence for Chesapeake Bay salinity and Susquehanna streamflow was areawise significant. The coherence at that timescale implies that the Chesapeake Bay salinity variability may have been due to atmospheric forcing, contrasting with the result from Lee and Lwiza (2008), who argued that the quasi-decadal variability of Chesapeake Bay bottom salinity was due to the meandering of the Gulf Stream.

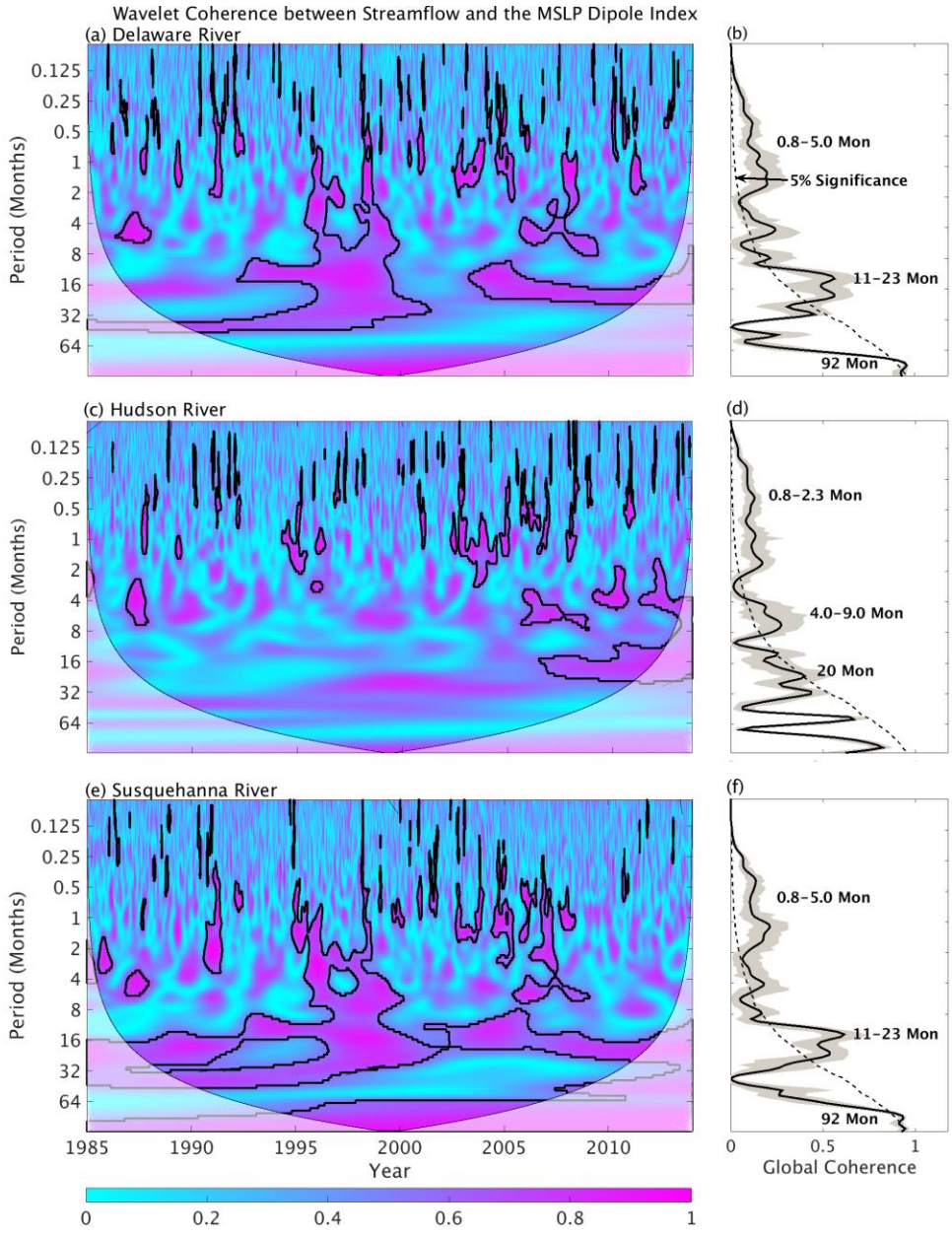

Figure 6.10. (a) Wavelet coherence between the ENA index and daily Delaware River streamflow. Contours enclose regions of 5% areawise significance and arrows indicate relative phase relationships. (b) The corresponding global coherence spectrum. (c) – (d) Same as (a) – (d) except for Hudson River streamflow. (e) – (f) Same as (a) – (b) except for Susquehanna River streamflow.

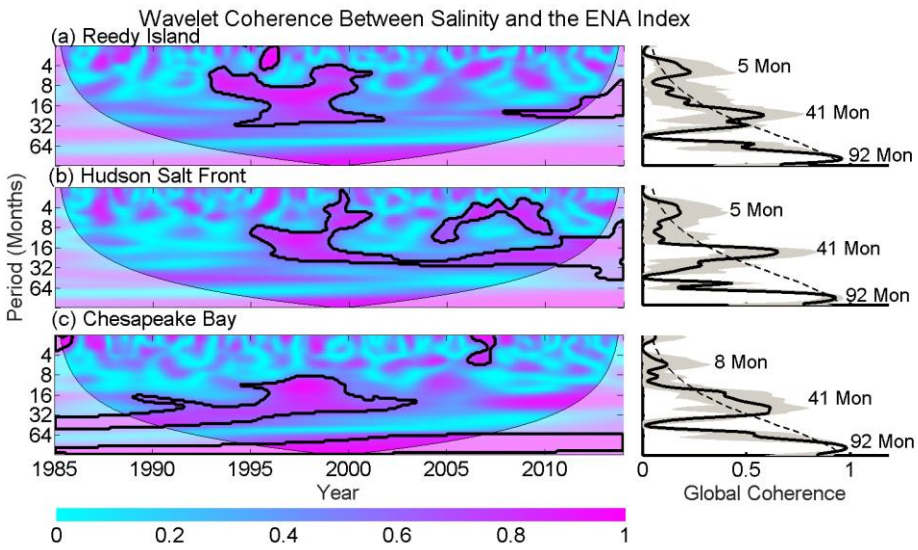

Figure 6.11. (a) Wavelet coherence between the monthly ENA index and Reedy Island salinity and (b) the corresponding global coherence spectra. (c) – (d) Same as (a) – (b) except for the Hudson River salt front position. (e) – (f) Same as (a) – (b) except for Chesapeake Bay surface salinity. For details of features found in the coherence spectra see Figure 6.10.

6.4.4 Tropical and Oceanic Connections

The fact that the arching Rossby wave train (Fig. 6.6) is concurrent with the MSLP dipole (Fig. 6.7) suggests that the ENA index may be preceded by tropical convection. A lag composite analysis of the ENA index with 300-hPa streamfunction anomalies was conducted. As shown in Figure 6.12, a positive ENA index event was generally preceded by negative 300-hPa streamfunction anomalies situated over the Southeast US. The anomaly centers at lag = 0 days were shifted eastward relative to their initial position at lag -20 days. A coherent pattern of alternating anomaly centers emerged at lag = -4 days and the pattern became most pronounced at lag = 0 days. The coherent pattern may have been due to Rossby waves generated in the tropics that propagated to the mid-latitudes, arced over North America, and finally propagated equatorward. Furthermore, the results of the composite analysis indicate that the ENA index may have been modulated by Rossby wave activity from the tropics. However, a more direct linkage between tropical convection and the ENA index would require an additional investigation of its relationship with outgoing longwave radiation or tropical precipitation. The wave pattern, nevertheless, resembles the pattern of a typical mid-latitude upper-atmospheric response to tropical forcing (Feldstein, 2000).

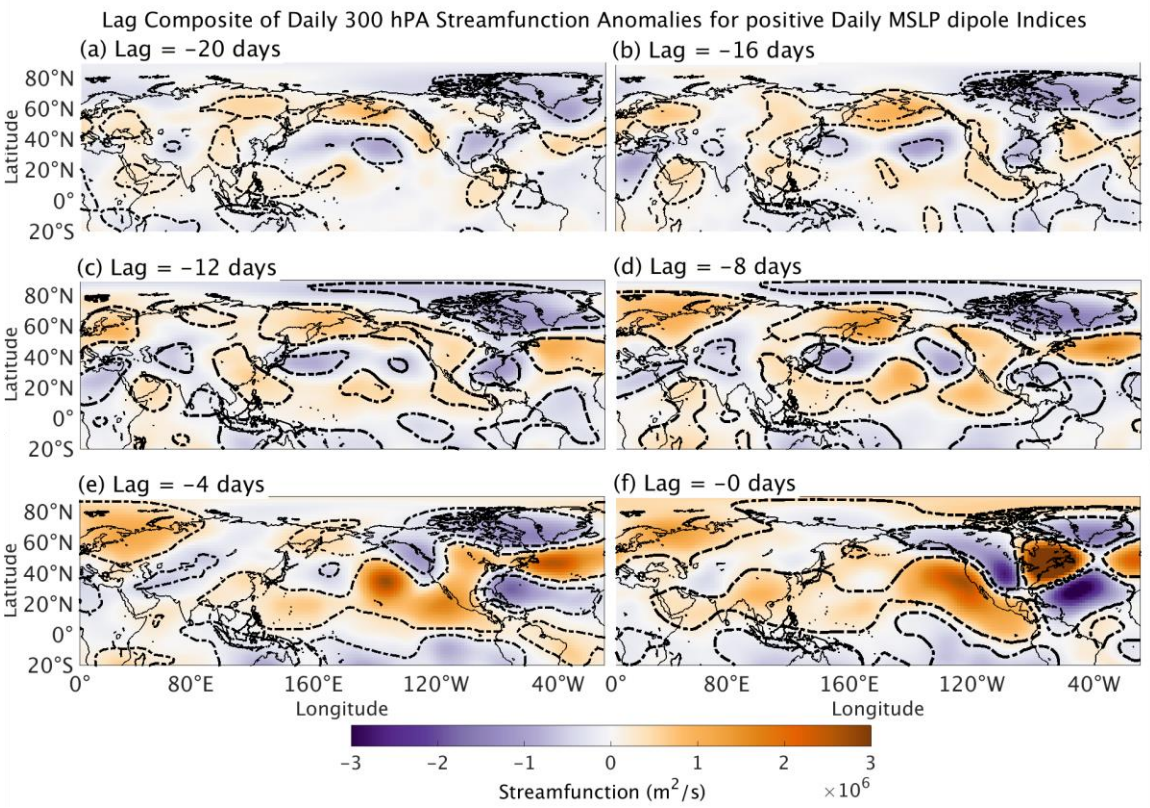

Figure 6.12. Same as Figure 6.7 except for the ENA index composited with 300 hPa streamfunction anomalies for the period 1985-2013.

To determine if ENSO was related to the ENA index, wavelet coherence between the Niño 4 index and the monthly ENA index was computed (not shown). Significant coherence was found between the two indices at a period of 74 months. Furthermore, a composite analysis was conducted between a filtered version of the ENA index and daily 300-hPa streamfuction anomalies. The filtered ENA index was calculated by taking the wavelet transform of ENA index, setting all wavelet coefficients to zero except those associated with the period of 74 months, and taking the resulting inverse wavelet transform of the new set of wavelet coefficients. The result of the composite analysis is shown in Figure 6.13. Remarkably, the results are similar to those shown in Figure 6.7, with Rossby waves emanating from the tropics, arcing over North America, and finally moving equatorward. The tropically forced Rossby waves operated on timescales of days so that the role of the low-frequency ENSO modes was to modulate the high-frequency variability of the ENA index. The modulation of the high-frequency variability would have then led

to low-frequency variability of the ENA index, inducing the low-frequency linkages with streamflow in the 24-48 month period band shown in Figure 6.10.

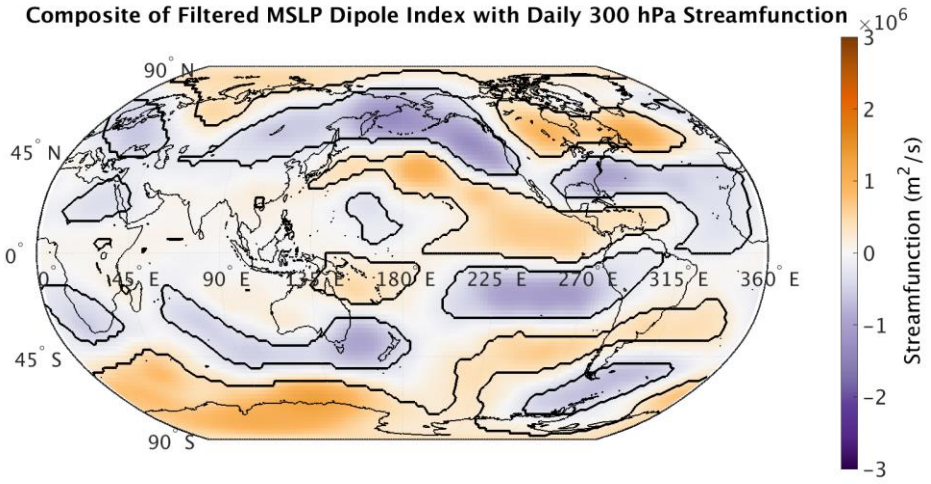

Figure 6.13. Mean 300 hPa streamfunction anomalies for the positive phase of the quasi-decadal mode with period of 74 months of the ENA index. Contours enclose regions of 5% significance as determined by the two-sample student-*t* test by comparing the mean for the positive phase with that of the negative phase.

The corresponding composite plot for MSLP shown in Figure 6.14 revealed a negative MSLP signature in which negative MSLP anomalies were located over the Southeastern US and positive anomalies were located over northern Canada and extended over the North Atlantic. The negative anomalies situated over the Southeast US also extended over the subtropical Atlantic Ocean, which would have spun down the subtropical gyre resulting in the equatorward movement of the Gulf Stream. It was thus hypothesized that at a period of 74 months, the Gulf Stream index was coherent with the ENA index. Figure 6.15a shows the 2-year running mean of the ENA index and GSI, and shows how the time series varied coherently with opposite phase. Moreover, Figure 6.15b shows that the coherent, out-of-phase fluctuations were most pronounced at a period of 74 months. The global coherence spectrum also indicates that the relationship at a period of 74 months was statistically significant. The out-of-phase relationship between the ENA index and the GSI can be partially explained by ENSO: during warm ENSO phases, according to Figure 6.12, daily positive extremes of the ENA index are more likely to occur, which are also associated with greater negative MSLP anomalies over the subtropical Atlantic Ocean. The atmospheric forcing would then have been integrated by the ocean and the integration

would have resulted in the spin down of the subtropical gyre and the equatorward shift of the Gulf Stream.

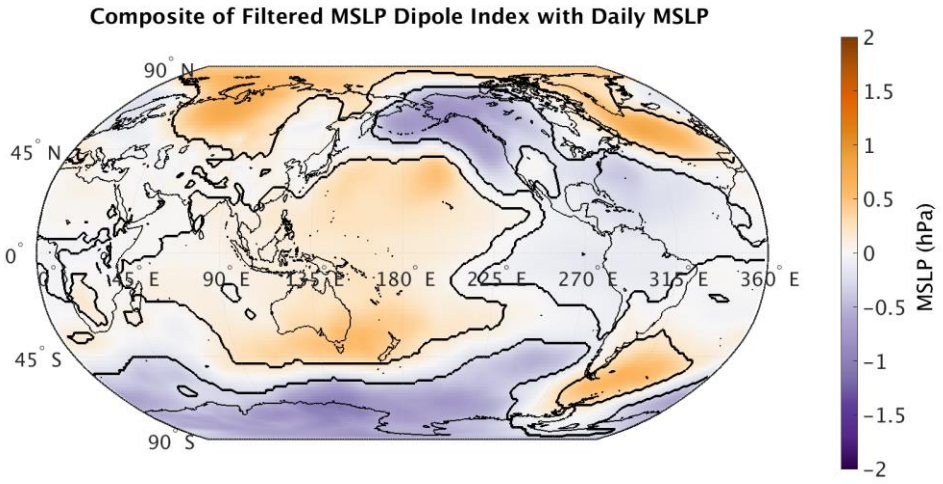

Figure 6.14. Same as Figure 6.13 except for daily MSLP anomalies.

The impact of ENSO on the GSI-ENA index relationship is supported in the partial wavelet coherence diagram shown in Figure 6.15d. The removal of the common dependence of the GSI and ENA index with the Niño 4 index resulted in weaker coherence between the time series. The reduction in coherence suggests that the observed relationship between the GSI and ENA index would have been weaker had it not been for ENSO-related changes in daily weather. The results are in agreement with Figure 6.13 that showed how a quasi-decadal mode of the ENA index was associated with Rossby wave trains from the tropics. Consistent with Taylor et al. (1998), the results also show that changes in the Gulf Stream were partially related to ENSO. The phase-locked behavior between the GSI and ENA index may explain why both salinity and streamflow exhibited enhanced variability at quasi-decadal timescales. During positive GSI phases, shelf salinity tends to be greater and concurrent with positive GSI phases were also negative ENA phases, which corresponded to low-flow conditions. The low-flow conditions resulted in increased estuarine salinity and thus the similarity between streamflow and salinity wavelet power spectra. The GSI-ENA index co-variability will make it difficult to separate pure Gulf Stream influences on estuarine salinity from precipitation variability associated with evolving atmosphere patterns.

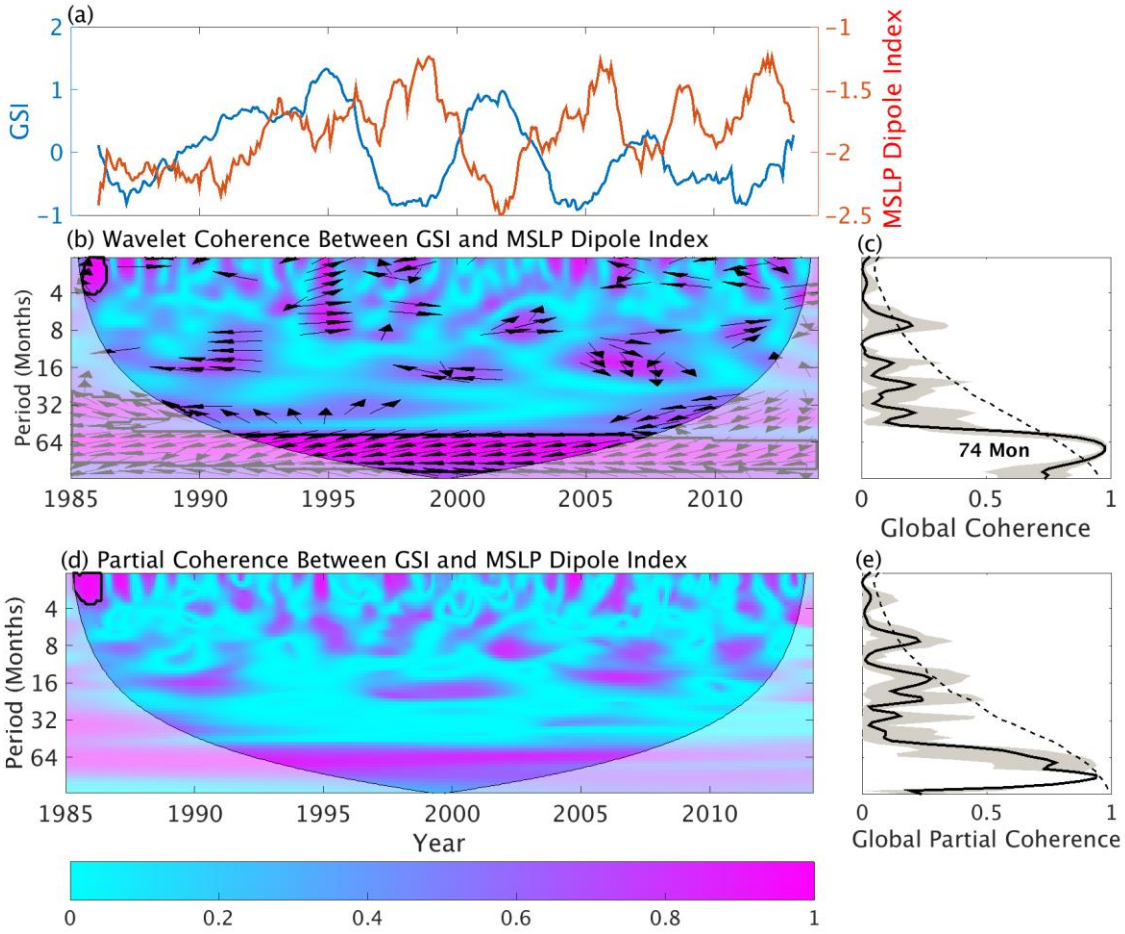

Figure 6.15. (a) 2-year running mean of the GSI and the ENA index, (b) wavelet coherence between the GSI and the ENA index, and (c) the corresponding global coherence spectra. For details of features found in the coherence spectra see Figure 6.10. (d) Partial wavelet coherence between the GSI and ENA index with their common dependence with the Niño 4 index removed. Contours enclose regions of 5% significance. (e) The global partial wavelet coherence spectra.

6.5. Conclusion

The relationship of large-scale atmospheric patterns with daily streamflow and monthly salinity was investigated on daily to decadal timescales. A composite analysis determined that anomalous high streamflow events are associated with Rossby waves emanating from the tropics. Coincident with the Rossby wave trains were dipole signatures of MSLP in which negative anomalies were located over the Southeastern US and positive anomalies were located over the northwestern North Atlantic Ocean. An ENA index was created to measure the strength and evolution of the pressure difference between the anomaly centers. A wavelet coherence analysis revealed that the index fluctuated

coherently with streamflow at timescales of 4-64 days and also at a period of 41 months. Estuarine salinity was also determined to have varied coherently with the ENA index at the same timescales. A wavelet coherence analysis together with a composite analysis of 300-hPa streamfunction anomalies suggested that the MSLP dipole was modulated by tropical convection and the resulting relationship generated phase-locked behavior with the GSI. The common forcing mechanism of both the GSI and the ENA index may make it difficult to separate pure Gulf Stream influences on mid-Atlantic estuaries from those associated with fluctuations in the ENA index.

Chapter 7

Additional Topics

7.1. Introduction

In previous chapters either monthly streamflow data or a portion of the daily streamflow time series were examined. Moreover, the powerful cumulative areawise test was not applied to the full streamflow time series because it was developed after Chapter 5 was submitted publication. Thus, in this chapter previous results from earlier chapters are re-examined and new results are presented.

7.2. Wavelet Analysis of Streamflow and Precipitation

Shown in Figure 7.1a is the daily Delaware River streamflow time series from 1913-2014. The most salient features were the 1930s and 1960s droughts, which were marked by a suppression of flow events exceeding the 95^{th} percentile. In fact, there was a 2-year period in the 1960s in which not a single 95^{th} percentile flow event occurred. Other notable features include the highest flow event that occurred in 1955 and the recent wet period or pluvial following the drought of 2002. The 1955 extreme flow event was caused by Hurricanes Diane and Connie, which traversed the region only a few days apart.

The wavelet power spectrum of the time series revealed numerous significant periodicities, where the significance was assessed from the application of the cumulative areawise test at the 5% level (Figure 7.1b). Numerous significant results were found in the 0.625-2 month period and are presumably associated with weather events. Perhaps more interesting is the lack of significance in the 1930s, 1960s, and around 2002. The lack of significance implies that extreme events were suppressed during those periods of drought, as shown in Figure 7.1a. Another prominent feature was the region of significance located in the 128-512 month period band. Those low-frequency periodicities resulted from the suppression of high-flow events during the 1930s drought and 1960s droughts and the relatively wet intermediate period. Such a relationship between high-frequency and low-frequency events is inconsistent with a red-noise process: for a red-noise process, statistical moments do not change with time, being a stationary process. Daily Delaware River streamflow, on the other hand, was marked by a decrease in variance during the 1930s

drought, the 1960s droughts, and other smaller droughts. This change in variance suggests that the behavior of streamflow was not simply governed by a red-noise process but related to mechanisms with nonstationary statistics.

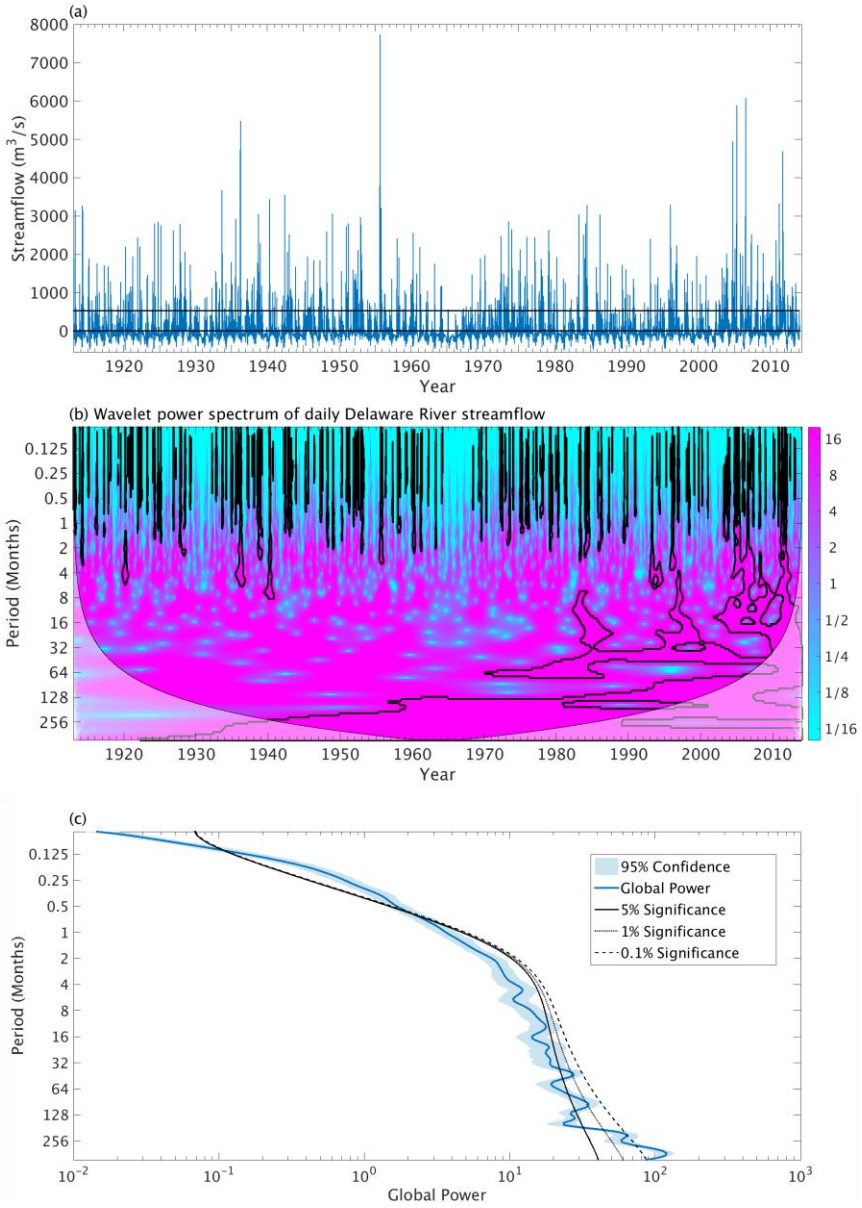

Figure 7.1. (a) Daily Delaware River streamflow anomalies and (b) the corresponding wavelet power spectrum. Contours enclose regions of 5% cumulative areawise significance. (c) Global wavelet power spectrum corresponding to (a). Blue shading represents the 95% confidence interval obtained from the block bootstrap procedure.

To see whether streamflow may have been related to external forcing from tropics, the wavelet coherence between monthly Delaware River Basin precipitation and the monthly Southern Oscillation index was computed (Figure 7.2). Unlike in Chapter 5, the cumulative areawise test will be applied to the wavelet coherence spectrum, providing more confidence in the results. To be consistent with Chapter 5, the analysis was only conducted for the period 1913-2013. Figure 7.2 shows cumulative areawise significant coherence at a period of 22 years. Moreover, the global coherence spectrum also provides evidence that the two time series evolved coherently from 1913 to 2013. The result suggests that the significant power found in the wavelet power spectrum of Delaware streamflow was the result of forcing from the tropics.

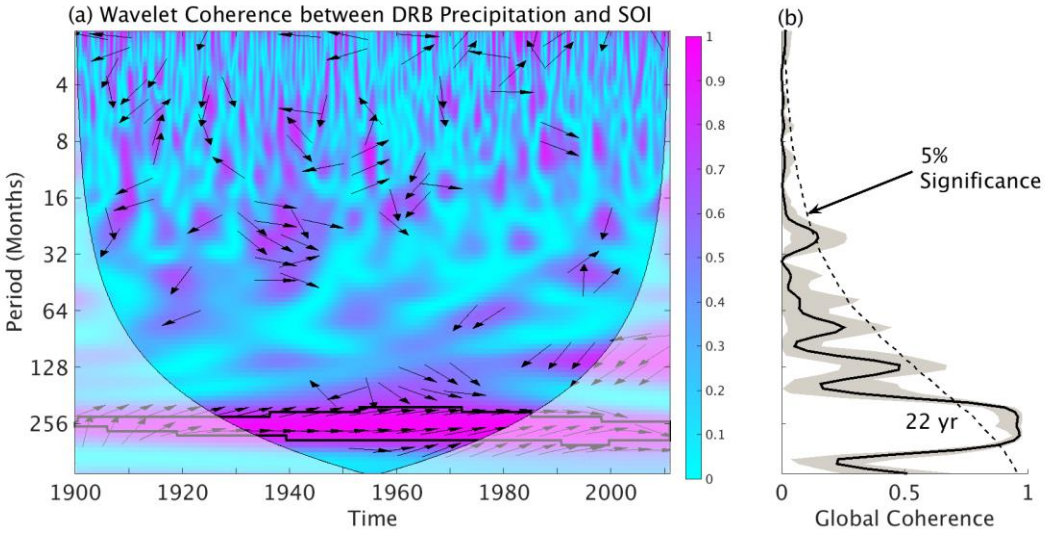

Figure 7.2. (a) Local wavelet coherence between monthly DRB precipitation and SOI. Arrows indicate the relative phase relationship. (b) Global coherence corresponding to (a). Graying represents the 95% confidence interval obtained from the block bootstrap procedure.

7.3. Higher-order Wavelet Analysis of Streamflow

It is not known what, if any, nonlinear properties the Delaware River streamflow time series possesses. The nonlinear properties are therefore investigated using high-order wavelet analysis and topological methods. The topological wavelet diagram of Delaware River streamflow for 1900-2015 is shown in Figure 7.3. The plot shows the centroids of holes located in significance patches at the 0.1%, 1%, and 5% levels. There were numerous holes located in the 0.625-1 month period band.

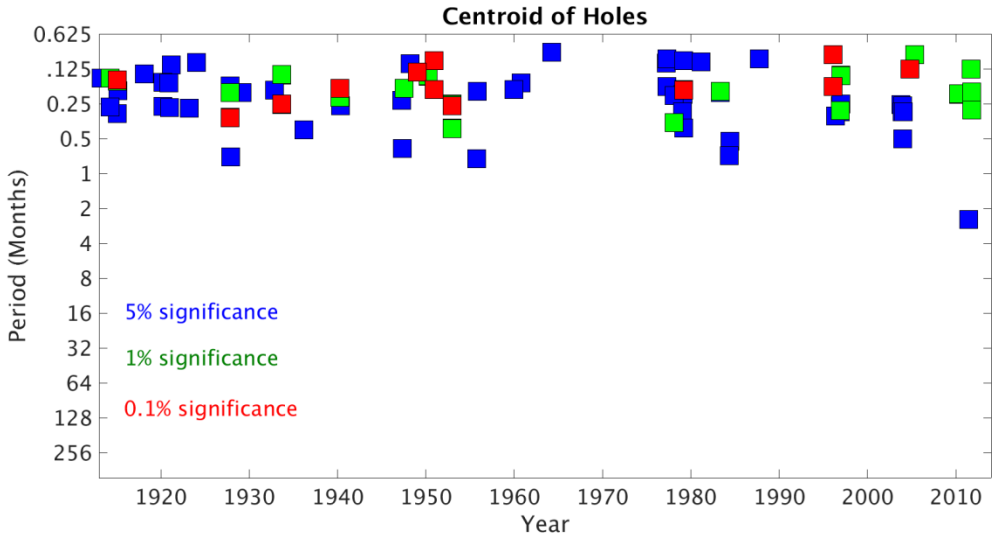

Figure 7.3. Topological wavelet diagram corresponding the wavelet power spectrum of daily Delaware River streamflow.

As shown by Schulte et al. (2015) the presence of holes may indicate the presence of nonlinearities. A notable feature is the void in holes from 1965 to 1975, which corresponds partially with the 1960s drought. To determine if the holes were related to nonlinearities, the autobicoherence spectrum of Delaware River streamflow was computed (Figure 7.4). As indicated by the contours, a large region of significance was found, with a local maximum in autobicoherence located at (6.5, 16) months. Recall that the autobicoherence spectrum is the decomposition of skewness so that skewness of the streamflow time series was attributed to high-frequency oscillations in the time series. Indeed, an inspection of Figure 7.1a shows that the time series is skewed, where positive anomalies are generally greater than negative anomalies. Skewness is almost always observed in streamflow and precipitation time series because the distributions of streamflow and precipitation are bounded by zero. However, if the time series are smoothed, say, using a 365-day running mean, then the time series become less skewed. The decrease in skewness could explain why only a few peaks were found in the low-frequency portion of the autobicoherence spectrum. It is noted that nonlinearities can be of higher order and such nonlinearities would not be detected by a bispectral analysis.

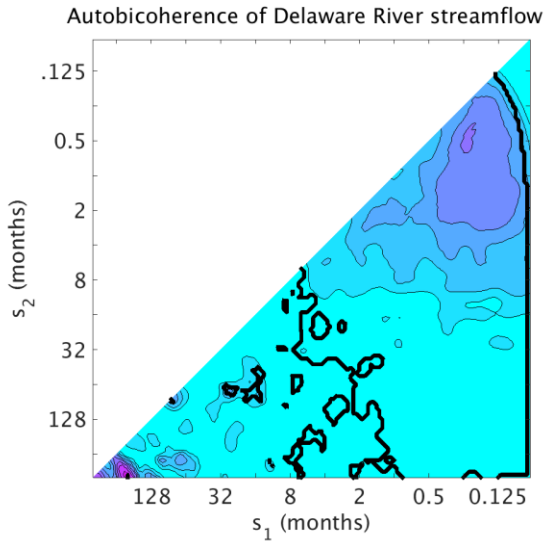

Figure 7.4. Autobicoherence spectrum of daily Delaware River streamflow anomalies for 1913-2015. Contours enclose regions of significance after controlling the false discovery rate at the 5% level.

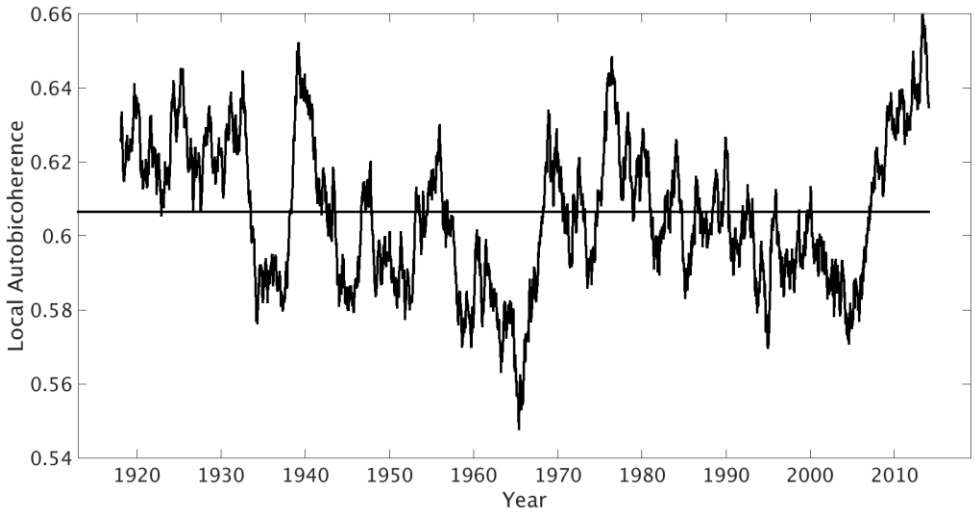

Figure 7.5. 5-year running mean of local autobicoherence corresponding to the point (10, 10) in the full autobicoherence spectrum shown Fig. 7.4.

The autobicoherence spectrum shown in Figure 7.4 does not show time behavior of the nonlinearities. The local autobicoherence at (6.5, 16) in the autobicoherence spectrum of Delaware River streamflow was therefore computed and is shown in Figure 7.5. The strength of the nonlinearity at (6.5, 16) varied considerably and also implies that the skewness of the time series also evolved temporally. A comparison of the streamflow time series and the autobicoherence time series reveals that the declines in streamflow coincided with the declines in local autobicoherence. During droughts, therefore, streamflow time

series becomes less skewed and the positive anomalies become similar in magnitude to negative anomalies. During high-flow periods the opposite occurred; high-skewness or greater bicoherence was identified.

7.4 Conclusions

A wavelet power spectrum of daily Delaware River streamflow detected significant periodicities in the period bands 0.0625-1, 128, 4-8, and 128-512 months. The inspection of the time series with its power spectrum showed that variance changes with time, which is inconsistent with a red-noise process. In particular, the declines in variance at high-frequencies were associated with droughts of various magnitudes and the absence of high daily flow events resulted in low-frequency oscillations in the time series. The implication of the temporally changing variance is that statistical analyses assuming stationarity are inappropriate. The skewness of the time series was also found to be nonstationary, with the lowest skewness having occurred during droughts. Together, the wavelet power and autobicoherence spectra showed that not only is the mean streamflow changing with time, but also the variance and skewness. Droughts are associated with large changes in all three of the statistical moments and therefore analyses assuming stationary may not be the most appropriate approach to understanding them.

Chapter 8

Summary and Discussion

Of interest to many climate scientists is the understanding of historical time series. Insight into the past behavior of time series is useful from a prediction point of view. If, for example, the known causes of drought are identified it may be possible foresee future droughts. However, the caveat is that the physical mechanisms governing a time series behavior must be themselves predictable. The understanding of time series requires the use of various time series and statistical tools. A particularly powerful tool is wavelet analysis, which allows non-stationary features of time series to be quantified as a function of time and frequency. Like in Fourier analysis, the results must be tested against a null hypothesis because random processes such as red-noise can produce seemingly deterministic features. The significance test developed by Torrence and Compo (1998), which first put wavelet analysis in a statistical framework, was found to have low statistical power by Maraun et al. (2004). Thus, in the present work, new statistical procedures were developed that addressed the limitations of existing procedures. The cumulative areawise test, in particular, was found to have greater statistical power than the geometric test but at the expense of slightly worse computational efficiency.

The cumulative areawise test in the present work was only applied to a nested sequence of two-dimensional objects but there are natural generalizations. In global wavelet spectra, the test could be applied to a sequence of one-dimensional objects, namely, arcs, whose test statistics would be cumulative arc lengths. Analogously, the test could also be applied to a sequence of three-dimensional objects and the test statistic would be cumulative volume. Thus, the cumulative areawise test could be applicable to *n*-dimensional wavelet analysis.

The application of cumulative areawise testing extends beyond the realm of wavelet analysis. In the physical sciences, one often applies statistical procedures to autcorrelated fields, where the data at one point are temporally correlated with data at another point. Thus, the implementation of a statistical procedure would result in clusters of spurious results arising from multiple-testing problems. However, according to the cumulative

areawise test, one can distinguish clusters of true positives from clusters of false positives by examining how their areas change as the significance level of the test is changed. The research presented here suggests that future work should pursue the development of a more generalized cumulative areawise test and an evaluation of its statistical power relative to existing procedures.

The topological methods developed in this work were limited in scope and the results determined from their application were empirically derived. In Chapter 2, for example, the number of holes and patches were computed in the wavelet domain. There are, however, natural extensions to higher dimensions, where one would be examining the behavior *n*-dimensional holes. For example, 2-dimensional holes are voids in a topological space such as that found in the 2-dimensional manifold $S^2 = \{(x_1, x_2, x_3): x_1^2 + x_2^2 + x_3^2 = 1\}$. With higher dimensional analogs, it may be possible to apply well-known theorems of algebraic topology such as Poincare Duality and Alexander Duality (Hatcher, 2002) to the understanding of statistical significance testing of *n*-dimensional wavelet analysis. The application of the generalized cumulative areawise test will also have some theoretically constrained results: according to the Bokum-ulam theorem (Hatcher, 2002), for a continuous function onto a sphere, at least one pair of antipodal points must obtain the same value of the function. Viewing *p*-values on a spherical domain as the range of a continuous function, the theorem dictates there are at least two anti-podal points such that the test statistics at the points are equal and the associated *p*-values are also equal. Thus, significance patches, as constrained by the theorem, may appear at anti-podal points. The uninvestigated connections between algebraic topology and field significance testing merits further investigation. Thus, future work in statistical significance testing could include the establishment of links between topological theorems and statistical procedures.

The 1960s drought in the Northeast US well-exemplifies the issue of predictability because different studies have produced contrasting results. Whereas climate models with prescribed SSTs cannot reproduce the 1960s drought (Seager et al., 2012), the investigation of observed data determined that La-Niña like conditions favored the abnormally dry conditions (Ning and Bradley, 2015). In the present work, statistical evidence points toward the 1960s drought being externally influenced from the tropics as opposed to internally

forced from the chaotic mid-latitude atmosphere. The wavelet coherence analysis, in particular, determined that streamflow of the Delaware, Hudson, and Susquehanna Rivers evolved coherently with the Southern Oscillation (SO) from 1900 to 2013 at a period of 22 years. It remains unclear, however, how the teleconnection operates; the SO is a phenomenon with a characteristic timescale of 2-7 years, far outside the period band of 22 years in which it is phase-locked to streamflow.

A wavelet analysis of monthly streamflow determined that streamflow exhibited a significant 26-year periodicity and the region of significance in the wavelet power spectrum was co-located with the 1960s drought. The results imply that the 1960s drought was a low-frequency phenomena that was forced by the higher-frequency SO. A comparison of the wavelet power spectra of streamflow and precipitation showed that the drought was related to precipitation variability. Precipitation events are high-frequency events and therefore do not contain characteristic time scales exceeding days. Then, the natural question arises: how do low-frequency features appear in precipitation time series? The low-frequency periodicities are, in fact, produced through the modulation of daily weather events. It is therefore necessary to examine daily streamflow data to obtain a more complete understanding of the physical processes governing streamflow variability. In the previous works of Labat (2008), Whitney, (2010), Seager et al. (2012), and Chapter 5 (Schulte et al., 2015a) only monthly data were considered, though the deficiency was addressed in Chapter 6. A composite analysis of 300 hPa streamfunction anomalies with daily Delaware River streamflow showed that anomalous streamflow events in the 1985-2013 period were associated with Rossby waves emanating from the tropics. The result suggests that the recent pluvial to some extent has resulted from wave activity from the tropics. This result contrasts with that from Seager et al. (2012) who argued that the pluvial is the result of intrinsic atmospheric variability.

One may wonder what insights about the 1960s drought can be gained through the examination of daily data. The application of wavelet analysis to the daily Delaware River streamflow time series from 1913-2015 revealed that the historical drought were coincident with the absence of high daily streamflow events. This result supports the idea that the droughts arose through the modulation of daily weather events. In future work, daily

streamflow data will be composited with daily atmospheric fields to gain insight into the 1960s drought.

Appendix A

Reproducing Kernel Property of Wavelets

Let $F(s, t)$ be the continuous wavelet transform of a function $f(t)$ such that

$$F(s, t) = \iint K(s, t;\ s', t'')F(s', t'')ds'dt''. \tag{A1}$$

Then the reproducing kernel is given by

$$K = \frac{1}{C_\psi \sqrt{ss'^{5/2}}} \int \left[\psi\left(\frac{t'-t''}{s'}\right)\psi^*\left(\frac{t-t'}{s}\right)\right] dt', \tag{A2}$$

where

$$C_\psi = \int_0^\infty \frac{|\Psi(\omega)|^2}{\omega} d\omega < \infty, \tag{A3}$$

and $\Psi(\omega)$ is the Fourier transform of ψ, and the asterisk denotes the complex conjugate. The reproducing kernel captures the structure of wavelet coefficients whereby the wavelet coefficient at any point contains information about a nearby wavelet coefficient weighted by K (Tropea, 2007).

Appendix B

Independence of Test Statistics from Chosen Reproducing Kernel

Let $A^N_{patch}(C_t, C_s)$ be the test statistic associated with a significance patch whose centroid is (C_t,C_s) and let $A^N_{\alpha_g}$ be the value of the test statistic corresponding to the $1-\alpha_g$ significance level of the geometric test. Writing

$$A^N_{\alpha_g} = \frac{A_{\alpha_g}}{A_R} \tag{B1}$$

and

$$A^N_{patch}(C_t, C_s) = \frac{A_{patch}}{A_R}, \tag{B2}$$

it follows that

$$\frac{A^N_{patch}(C_t,C_s)}{A^N_{\alpha_g}} = \frac{A_{patch}}{A_{\alpha_g}}, \tag{B3}$$

where is A_{patch} the area of the significance patch and is the A_{α_g} the area of a typical patch under the null hypothesis corresponding to the $1-\alpha_g$ significance level. Since Eq. (B3) no longer contains A_R, the relationship between $A^N_{patch}(C_t, C_s)$ and $A^N_{\alpha_g}$ no longer depends on P_{crit}.

Appendix C

Proof of Green's Theorem for a Polygon

Recall that Green's Theorem in the plane states that

$$\int_C (Pdx + Qdy) = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA\,, \tag{C1}$$

where *C* is a positively oriented, piecewise smooth curve, bounding a region *D*; $\boldsymbol{F} = \langle P, Q \rangle$ is a vector field on *D*; and *x* and *y* are the usual Cartesian coordinates (Baxandall and Liebeck, 2008). Note that if one sets

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1, \tag{C2}$$

then the right-hand side of Eq. (C1) can be used to calculate the area of a region *D*. Thus, let $Q = x/2$ and $P = -y/2$ so that

$$\tfrac{1}{2}\int_C (xdy - ydx) = A(D), \tag{C3}$$

where $A(D)$ denotes the area of *D*. Let (x_0, y_0),…,(x_{m-1}, y_{m-1}) be *m*-1 vertices of a polygon. If C_0 is a line segment from (x_0, y_0) to (x_1, y_1), then

$$\int_{C_0}(xdy - ydx) = x_0\, y_1 - x_1 y_0. \tag{C4}$$

More generally, denote by C_k the segment from (x_k, y_k) to (x_{k+1}, y_{k+1}), recalling that $x_m = x_0$ and $y_m = y_0$. Since $C = C_0 \cup C_1, \ldots, \cup C_{m-1}$, we can write

$$A(D) = \frac{1}{2}\int_C (xdy - ydx)$$

$$= \tfrac{1}{2}\int_{C_0}(xdy - ydx) + \tfrac{1}{2}\int_{C_1}(xdy - ydx) + \cdots + \tfrac{1}{2}\int_{C_{m-1}}(xdy - ydx) \tag{C5}$$

and thus

$$A(D) = \frac{1}{2}(x_0 y_1 - x_1 y_0) + \frac{1}{2}(x_1 y_2 - x_2 y_1) + \cdots + \frac{1}{2}(x_{m-1} y_0 - x_0 y_{m-1})$$

$$= \tfrac{1}{2}\textstyle\sum_{k=0}^{m-1}(x_k\, y_{k+1} - x_{k+1} y_k). \tag{C6}$$

Thus, Eq. (2.11) follows from Green's Theorem in a plane.

Appendix D

Definition of Path-Component

A path in a set X is defined as a continuous function $f: I \rightarrow X$. A set X is said to be path-connected if any two points x and y in X can be joined by a path. The path-component of a topological space X is the maximal path-connected subset of a set. Intuitively, one can think of a path-component as the largest isolated piece of the set. For example, the set could be the disjoint union of a square and a disc, in which case both the square and the disc are path-components.

Appendix E

Proof of Inequality of Normalized Areas

Let P_1 and P_2 be two subsets of a patch P with area A such that P_2 is the set complement of P_1. Let A_1 and A_2 denote the areas of P_1 and P_2, respectively. One can thus write

$$A = A_1 + A_2, \tag{E1}$$

$$A = r_1 A + r_2 A, \tag{E2}$$

and

$$r_2 = 1 - r_1, \tag{E3}$$

where $r_1, r_2 \in [0,1]$. The centroid of P can be written as

$$\frac{1}{A}\iint_P sdsdt = \frac{1}{A}\iint_{P_1} sdsdt + \frac{1}{A}\iint_{P_2} sdsdt \tag{E4}$$

$$= \frac{1}{A}\iint_P sdsdt = \frac{r_1}{A_1}\iint_{P_1} sdsdt + \frac{r_2}{A_2}\iint_{P_2} sdsdt \tag{E5}$$

or

$$C^s = r_1 C_1^s + r_2 C_2^s, \tag{E6}$$

so that

$$\frac{C^s - r_2\, C_2^s}{r_1} = C_1^s, \tag{E7}$$

where C_1^s and C_2^s are the scale coordinates of the centroids for P_1 and P_2. The equation implies that

$$C^S - r_2\, C_2^S > 0 \tag{E8}$$

because C_1^s is always positive. The normalized areas of P and P_1 are given by

$$A^N = \frac{A}{(C^s)^2} \tag{E9}$$

and

$$A_1^N = \frac{A_1}{(C_1^S)^2}. \tag{E10}$$

Thus,

$$r_{norm} = \frac{A_1^N}{A^N} = \frac{r_1^3 (C^S)^2}{(C^S - (1 - r_1) C_2^S)^2}. \tag{E11}$$

At $r_1 = 0$, $r_{norm} = 0$ because P_1 has no area. At $r_1 = 1$, $r_{norm} = 1$ because $A_1 = A$. Moreover, the function is monotonically increasing for $r_1 \in [0\ 1]$ so that $r_{norm} \leq 1$. The same arguments hold for P_2 except that r_{norm} decreases monotonically.

Bibliography

Albert, R. C.: The Historical Context of Water Quality Management for the Delaware Estuary. Estuaries, 11, 99-107, 1988.

Alley, R. B., Marotzke, J., Nordhaus, W. D, Overpeck, J. T., Peteet, D. M., Pielke Jr., R. A., Pierrehumbert, R. T., Rhines, P. B., Stocker, T. F., Talley, L. D., Wallace, J. M.: Abrupt Climate Change. Science, 299, 2005-2010, 2003.

Alexander, M. A., Bladé, I., Newman, M., Lanzante, J. R., Lau, N., Scott, J. D.: The Atmospheric Bridge: The Influence of ENSO Teleconnections on Air–sea Interaction over the Global Oceans. J. Climate, 15, 2205-2231, 2002.

Archambault, H. M., Bosart, L. F., Keyser, D., Aiyyer, A. R.: Influence of Large-scale Flow Regimes on Cool-season Precipitation in the Northeastern United States. Mon. Wea. Rev., 136, 2945-2963, 2008.

Barber, C. B., Dobkin, D. P., and Huhdanpaa, H.: Quickhull Algorithm for Convex Hulls. ACM T. Math. Software, 22, 469-483, 1996.

Barlow, M., Nigam, S., Berbery, E. H.: ENSO, Pacific Decadal Variability, and US Summertime Precipitation, Drought, and Streamflow. J. Climate, 14, 2105-2128, 2001.

Barnston, A. G., Livezey, R. E.: Classification, Seasonality and Persistence of Low-Frequency Atmospheric Circulation Patterns. Mon. Wea. Rev., 115, 1083-1126, 1987.

Baxandall, P. and Liebeck, H.: Vector Calculus, Dover Publications, INC., Mineloa, New York, 550 pp., 2008.

Benjamini, Y., Hochberg, Y.: Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. Royal Statistical Society, 57, 289-300, 1995.

Benjamini, Y., Yekutieli, D.: The Control of the False Discovery Rate in Multiple Testing under Dependency. Ann Statist., 29, 1165-1188, 2001.

Beyer, W. H.: CRC Standard Mathematical Tables. 28th Ed. CRC Press, Boca Raton, Florida, 674 pp., 1987.

Bjerknes, J.: Atmospheric Teleconnections from the Equatorial Pacific. Mon. Weather Rev., 97, 163-172, 1969.

Bradbury, J. A., Dingman, S. L., Keim, B. D.: New England Drought and Relations with Large Scale Atmospheric Circulation Patterns. J Am. Water Resources Assoc., 38, 1287-1299, 2002.

Cook, B., Miller, R. L., Seager, R.: Amplification of the North American “Dust Bowl” Drought through Human-induced Land Degradation. Proc. Natl. Acad. Sci. USA, 106, 4997–5001, 2008.

Collis, W. B., White, P. R. and Hammond, J. K.: Higher-order Spectra: The Bispectrum and Trispectrum. Mech. Syst. Signal Process. 12, 375–394, 1998.

Curtis, S.: The Atlantic Multidecadal Oscillation and Extreme Daily Precipitation over the US and Mexico during the Hurricane Season. Clim. Dyn., 30, 343-351, 2008.

Dettinger, M., Diaz, H.: Global Characteristics of Streamflow and Flow Seasonality and Variability. J. Hydrometeorol., 1, 289–310, 2000.

Edelsbrunner, H. and Harer, J.: Computational Topology: an Introduction, Amer. Math. Soc., Providence, Rhode Island, 241 pp., 2009.

Edelsbrunner, H. and Harer, J.: Persistent homology-a survey, Cotemp. Math., 12, 1-26, 2010.

Efron, B.: Bootstrap Methods: Another Look at the Jackknife. Ann. Statist., 7, 1–26, 1979.

Eichler, T and Higgins, W.: Climatology and ENSO-related Variability of North American Extratropical Cyclone Activity. J. Climate, 19, 2076-2093, 2006.

Elgar, S. and Chandran, V.: Higher-order Spectral Analysis to Detect Nonlinear Interactions in Measured Time Series and an Application to Chua’s Circuit. Internat. J. Bifurcat. Chaos, 3, 19–34, 1993.

Elsayed, M. A. K.: Wavelet Bicoherence Analysis of Wind–wave Interaction. Ocean Eng., 33, 458–470, 2006.

Enfield, D. B., Mestas-Nuñez, A. M., Trimble, P. J.: The Atlantic Multidecadal Oscillation and its Relation to Rainfall and River Flows in the Continental US. Geophys. Res. Lett. 28, 2077-2080, 2001.

Feldstein, S. B.: The Time Scale, Power Spectra, and Climate Noise Properties of Teleconnection Patterns, J. Climate., 13, 4430-4440, 2000.

Folland, C. K., Knight, J., Linderholm, H. W., Fereday, D., Ineson, S., Hurrell, J. W.: The Summer North Atlantic Oscillation: Past, Present, and Future. J. Climate, 22, 1082-1103, 2009.

Gan, T. Y., Gobena, A. K., and Wang, Q.: Precipitation of Southwestern Canada: Wavelet, Scaling, Multifractal Analysis, and Teleconnection to Climate Anomalies, J. Geophys. Res., 112, D10110, DOI: 10.1029/2006JD007157, 2007.

Gershunov, A., Barnett, T. P..: Interdecadal Modulation of ENSO Teleconnections. Bull. Amer. Meteor. Soc., 79, 2715-2725, 1998.

Gilman, D. L., Fuglister, F. J., and Mitchell J. M. Jr.: On the Power Spectrum of "Red Noise", J. Atmos. Sci., 20, 182–184, 1963.

Ghrist, R.: Barcodes: the Persistent Topology of Data, Bull. Amer. Math. Soc., 45, 61-75, 2008.

Grinsted, A., Moore, J. C. and Jevrejeva, S.: Application of the Cross Wavelet Transform and Wavelet Coherence to Geophysical Time Series. Nonlinear Process. Geophys., 11, 561–566, 2004.

Guttman, N. B., Quayle, R. G.: A Historical Perspective of US Climate Divisions. Bull. Amer. Meteor. Soc., 77, 293–303, 1996.

Hagelberg, T., Pisias, N., Elgar, S.: Linear and Nonlinear Couplings between Orbital Forcing and the Marine δ^{18}O Record during the Late Neocene. Paleoceanography, 6, 729-746, 1991.

Hagy, J. D., Boynton, W. R., Keefe, C. W., Wood, K. V.: Hypoxia in Chesapeake Bay, 1950–2001: Long-term Change in Relation to Nutrient Loading and River Flow. Estuaries, 27, 634-658, 2004.

Hamiliton, K. and Hsieh, W. W.: Representation of the Quasi-biennial Oscillation in the Tropical Stratospheric Wind by Nonlinear Principal Component Analysis. J. Geophys. Res., 107, 4232, doi:10.1029/2001JD001250, 2002.

Hanna, E., Cropper, T. E., Jones, P. D., Scaife, A. A., and Allan, R.: Recent Seasonal Asymmetric Changes in the NAO (a Marked Summer Decline and Increased Winter variability) and Associated Changes in the AO and Greenland Blocking Index. Int. J. Climatol., 35, 2540-2554, 2014.

Haskin, H. H. and Ford, S. E.: Haplosporidium Nelsoni (MSX) on Delaware Bay Seedoyster Beds: a Host-parasite Relationship Along a Salinity Gradient. J. Invertebr. Pathol., 40, 388-405, 1982.

Hasselmann, K.: Stochastic Climate Models Part I. Theory. Tellus, 28, 473-485, 1976.

Hatcher, A.: Algebraic Topology, Cambridge University Press, New York, 544 pp., 2001.

Hayes, M. H.: Statistical Digital Signal Processing and Modeling, John Wiley & Sons, New York, 624, 1996.

Higuchi, K., Huang, J., Shabbar, A.: A Wavelet Characterization of the North Atlantic Oscillation Variation and its Relationship to the North Atlantic Sea Surface Temperature. Int. J. Climatol., 19, 1119-1129, 1999.

Hilton, T. W., Najjar, R. G., Zhong, L., and Li, M.: Is There a Signal of Sea-level Rise in Chesapeake Bay Salinity?. J. Geophys. Res., 113, C09002, DOI: 10.1029/2007JC004247 2008.

Horel, J. D. and Wallace, J. M.: Planetary-Scale Atmospheric Phenomena Associated with the Southern Oscillation. Mon. Wea. Rev., 109, 813–829, 1981.

Hoskins, B. J. and Karoly, D. J.: The Steady Linear Response of a Spherical Atmosphere to Thermal and Orographic Forcing. J. Atmos. Sci., 38, 1179–1196, 1981.

Howarth, R. W., Swaney, D., Butler, T. J., and Marino, R.: Climatic Control on Eutrophication of the Hudson River Estuary. Ecosystems, 3, 210–215, 2000.

Hubeny, J. B., King, J. W., Reddin, M.: Northeast US Precipitation Variability and North American Climate Teleconnections Interpreted from Late Holocene Varved Sediments. Proc. Natl. Acad. Sci. USA, 108, 17895-17900, 2011.

Hull, C. H. J. and Titus, J. G.: Greenhouse Effect, Sea-level Rise, and Salinity in the Delaware Estuary. US EPA 230-05-86-010, US Environmental Protection Agency and the Delaware River Basin Commission, Washington, DC, 1986.

Hurrell, J. W.: Decadal Trends in the North Atlantic Oscillation Regional Temperatures and Precipitation. Science, 269, 676–679, 1995.

Hurrell, J. W. and van Loon, H.: Decadal Variations in Climate Associated with the North Atlantic Oscillation. Climate Change, 36, 301–326, 1997.

Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M. (Eds.): The North Atlantic Oscillation: Climatic Significance and Environmental Impact. Geophys. Monogr. Ser., 134, American Geophysical Union, 279 pp., 2003.

Jacobs, G. A., Hurlburt, H. E., Kindle, J. C., Metzger, E. J., Mitchell, J. L., Teague, W. J., Wallcraft, A. J.: Decade-scale Trans-Pacific Propagation and Warming Effects of an El Niño Anomaly. Nature, 370, 360-363, 1994.

Jenkins, G. W. and Watts, D. G.: Spectral Analysis and its Applications. Holden-Day, San Francisco, California, 541 pp., 1968.

Jevrejeva, S., Moore, J. C., Grinsted, A.: Influence of the Arctic Oscillation and El Nino-Southern Oscillation (ENSO) on Ice Conditions in the Baltic Sea: The wavelet Approach. J. Geophys. Res., 108, 4677, doi:10.1029/2003JD003417, D21, 2003.

Kay, S. M.: Modern Spectral Estimation: Theory and Application, Prentice Hall, Englewood Cliffs, NJ, 560 pp., 1988.

Kemp, W. M., Boynton, W. R., Adolf, J. E., Boesch, D. F., Boicourt, W. C., Brush, G., Cornwell, J. C., Fisher, T. R., Glibert, P. M., Hagy, J. D., Harding, L. W., Houde, E. D., Kimmel, D. G., Miller, W. D., Newell, R. I. E., Roman, M. R., Smith, E. M., Stevenson, J. C.: Eutrophication of Chesapeake Bay: Historical Trends and Ecological Interactions, 2005.

Kerr, R. A.: A North Atlantic Climate Pacemaker for the Centuries. Science, 288, 1984-1985, 2000.

Kimmel, D. G. and Roman, M. R.: Long-term Trends in Mesozoo-plankton Abundance in Chesapeake Bay, USA: Influence of Freshwater Input. Mar. Ecol. Prog. Ser., 267, 71–83, 2004.

King, T.: Quantifying Nonlinearity and Geometry in Time Series of Climate. Quat. Sci. Rev., 15, 247–266, 1996.

Kreeger, D., Adkins, J., Cole, P., Najjar, R., Velinsky, D., Conolly, P., Kraeuter, J.: Climate Change and the Delaware Estuary: Three Case Studies in Vulnerability Assessment and Adaptation Planning. Tech. rep., Partnership for the Delaware Estuary, Wilmington, DE, 2010.

Kunsch, H. R.: The Jackknife and the Bootstrap for General Stationary Observations. Ann. Statist., 17, 1217–1241, 1989.

Labat, D.: Wavelet Analysis of the Annual Discharge Records of the World's Largest Rivers. Adv. Water Resour., 31, 109-117, 2008.

Labat, D.: Cross Wavelet Analyses of Annual Continental Freshwater Discharge and Selected Climate Indices. J. Hydrol., 385, 269-278, 2010.

Latif, M. and Barnett, T.: Causes of Decadal Climate Variability over the North Pacific and North America. Science, 266, 634–637, 1994.

Latif, M. and Barnett, T. P.: Decadal Climate Variability over the North Pacific and North America: Dynamics and Predictability. J. Climate, 9, 2407-2423, 1996.

Latif, M. and Keenlyside, N. S.: El Nino/Southern Oscillation Response to Global Warming, Proc. Natl. Acad. Sci. USA, 106, 20578–20583, 2008.

Lau, K. M. and Weng, H.: Climate Signal Detection using Wavelet Transform: How to Make a Time Series Sing. Bull. Amer. Meteor. Soc., 76, 2391–2402, 1995.

Leathers, D. J., Yarnal, B., Palecki, M. A.: The Pacific/North American Teleconnection Pattern and United States Climate. Part I: Regional Temperature and Precipitation Associations. J. Climate, 4, 517-528, 1991.

Lee, Y. J. and Lwiza, K. M.: Factors Driving Bottom Salinity Variability in the Chesapeake Bay. Cont. Shelf Res., 28, 1352-1362, 2008.

Lehmann, E.: Testing Statistical Hypothesis, Springer Texts in Statistics, John Wiley and Sons, New York, New York, 369 pp, 1986.

Linkin, M. E. and Nigam, S.: The North Pacific Oscillation–west Pacific Teleconnection Pattern: Mature-phase Structure and Winter Impacts. J. Climate, 21, 1979-1997, 2008.

Luksch, U. and von Storch, H.: Modeling the Low-frequency Sea Surface Temperature Variability in the North Pacific. J. Climate 5, 893-906, 1992.

Lu, B.W., Pandolfo, L., and Hamilton, K.: Nonlinear Representation of the Quasi-Biennial Oscillation. J. Atmos. Sci., 66, 1886–1904, 2009.

MacDonald, G. M., Case, R. A.: Variations in the Pacific Decadal Oscillation over the Past Millennium, Geophys. Res. Lett., 32, L08703, doi:10.1029/2005GL022478, 2005.

Maccarone, T. J.: The Biphase Explained: Understanding the Asymmetries in Coupled Fourier Components of Astronomical Timeseries, Mon. Not. R. Astron. Soc., 435, 3547, doi: 10.1093/mnras/stu1824, 2013.

Mantua, N. J. and Hare, S. R.: The Pacific Decadal Oscillation. J. Oceanogr., 58, 35-44, 2002.

Mantua, N. J., Hare. S. R., Zhang, Y., Wallace, J. M., Francis, R. C.: A Pacific Interdecadal Climate Oscillation and its Impact on Salmon Production. Bull. Amer. Meteor. Soc., 76, 1069–1079, 1997.

Maraun, D. and Kurths, J.: Cross Wavelet Analysis: Significance Testing and Pitfalls, Nonlin. Processes Geophys., 11, 505-514, 2004.

Maraun, D., Kurths, J., and Holschneider, M.: Nonstationary Gaussian Processes in Wavelet Domain: Synthesis, Estimation, and Significance Testing, Phys. Rev. E, 75, doi: 2 10.1103/PhysRevE.75.016707, 2007.

Mattson, D. E.: Statistics: Difficult Concepts, Understandable Explanations. C.V. Mosby, St. Louis, 482 pp, 1981.

McCabe, G. J., Palecki, M. A., Betancourt, J. L.: Pacific and Atlantic Ocean Influences on Multidecadal Drought Frequency in the United States. Proc. Natl Acad. Sci. USA, 101, 4136-4141, 2004.

McCready, P. and Geyer, R. W.: Advances in Estuarine Physics. Annu. Rev. Mar. Sci., 2, 35-58, 2010.

Menne, M. J., Williams, C. N., Vose, R. S.: The US Historical Climatology Network Monthly Temperature Data, Version 2. Bull. Amer. Meteor. Soc., 90, 993-1007, 2009.

Meyers, S. D., Kelly, B. G., and O'Brien, J. J.: An Introduction to Wavelet Analysis in Oceanography and Meteorology: With Application to the Dispersion of Yanai Waves. Mon. Weather Rev., 121, 2858–2866, 1993.

Miller, A. J., White, W. B., Cayan, D. R.: North Pacific Thermocline Variations on ENSO Timescales. J. Phys. Oceanogr., 27, 2023-2039, 1997.

Miller, W. D., Kimmel, D. G., Harding, L. W. Jr.: Predicting Spring Discharge of the Susquehanna River from a Winter Synoptic Climatology for the Eastern United States. Water Resour. Res., 42, W05414, doi:10.1029/2005WR004270, 2006.

Miller, W. and Harding, L. Jr.: Climate Forcing of the Spring Bloom in Chesapeake Bay. Mar. Ecol. Prog. Ser., 331, 11-22, 2007.

Minobe, S.: A 50–70 years Climatic Oscillation over the North Pacific and North America. Geophys. Res. Lett. 24, 683–686, 1997.

Minobe, S. and Mantua, N.: Interdecadal Modulation of Interannual Atmospheric and Oceanic Variability over the North Pacific. Prog. Oceanogr., 43, 163–192, 1999.

Mosier, A. C. and Francis, C. A.: Relative Abundance and Diversity of Ammonia-oxidizing Archaea and Bacteria in the San Francisco Bay Estuary. Environ. Microbiol., 10, 3002–3016, 2008.

Moussas, X., Polygiannakis, J. M., Preka-Papadema, P., Exarhos, G., Solar cycles: A Tutorial. Adv. Sp. Res., 35, 725-738, 2005.

Müller, W. A., Frankignoul, C., Chouaib, N.: Observed Decadal Tropical Pacific–North Atlantic Teleconnections. Geophys. Res. Lett., 35, doi:10.1029/2008GL035901, 2008.

Najjar, R. G.: The Water Balance of the Susquehanna River Basin and its Response to Climate Change. J. Hydrol., 219, 7-19, 1999.

Najjar, R. G., Walker, H. A., Anderson, P. J., Barro, E. J., Bord, R. J., Gibso, J. R., Kenndey, V. S., Knight, C. G., Megonigal, J. P., O'Connor, R. E. , Polsky, C. D., Psuty, N. P., Richards, B. A., Sorenson, L. G., Steele, E. M., and Swanson. R. S.: The Potential Impacts of Climate Change on the Mid-Atlantic Coastal Region. Climate Research, 14, 219–233, 2000.

Najjar, R. G., Pyke, C. R., Adams, M. B., Breitburg, D., Hershner, C., Kemp, M., Howarth, R., Mulholland, M. R., Paolisso, M., Secor, D., Sellner, K., Wardrop, D., Wood, R.: Potential Climate-change Impacts on the Chesapeake Bay. Est. Coast. Shelf Sci., 86, 1-20, 2010.

Najjar, R. G., Patterson, L., Graham, S.: Climate Simulations of Major Estuarine Watersheds in the Mid-Atlantic Region of the US. Climatic Change, 95, 139-168, 2009.

Namias, J.: Nature and Possible Causes of the Northeastern United States Drought during 1962-65. Mon. Wea. Rev., 94, 543-554, 1966.

Neelin, J. D., Weng, W.: Analytical Prototypes for Ocean–atmosphere Interaction at Midlatitudes. Part I: Coupled Feedbacks as a Sea Surface Temperature Dependent Stochastic Process. J. Climate, 12, 697-721, 1999.

Ng, E. K.W. and Chan, J. C. L.: Geophysical Applications of Partial Wavelet Coherence and Multiple Wavelet Coherence. J. Atmos. Oceanic Technol., 29, 1845–1853, 2012.

Newman, M., Compo, G. P., Alexander, M. A.: ENSO-forced Variability of the Pacific Decadal Oscillation. J. Climate, 16, 3853-3857, 2003.

Nidal, K. and Malik, A. S.: EEG/ERP Analysis: Methods and Applications, CRC Press, 334 pp., 2013.

Nikias, C. L. and Raghuveer, M. R.: Bispectrum Estimation: A Digital Signal Processing Framework, IEEE, 75, 869-891, 1987.

Ning, L., Mann, M. E., Crane, R., Wagener, T., Najjar, R. G., Singh, R.: Probabilistic Projections of Anthropogenic Climate Change Impacts on Precipitation for the Mid-Atlantic Region of the United States. J. Climate, 25, 5273-5291, 2012.

Ning, L. and Bradley, R.S.: Winter Climate Extremes over the Northeastern United States and Southeastern Canada and Teleconnections with Large-Scale Modes of Climate Variability. J. Climate, 28, 2475–2493, 2015.

Notaro, M., Wang, W., Gong, W.: Model and Observational Analysis of the Northeast U.S. Regional Climate and its Relationship to the PNA and NAO Patterns during Early Winter. Mon. Wea. Rev, 134, 3479-3505, 2006.

Newman, M., Compo, G. P., Alexander, M. A.: ENSO-forced Variability of the Pacific Decadal Oscillation. J. Climate, 16, 3853-3857, 2003.

Olsen, J., Anderson, J. N., Knudsen, M. F.: Variability of the North Atlantic Oscillation over the past 5,200 years, Nature Geosci., 5, 808-812, 2012.

Özger, M., Mishra, A. K., and Singh, V. P.: Low Frequency Drought Variability Associated with Climate Indices. J. Hydrol., 364, 152–162, 2009.

Paluš, M. and Novotná, D.: Quasi-biennial Oscillations Extracted from the Monthly NAO Index and Temperature Records are Phase-synchronized. Nonlin. Process. Geophys., 13, 287–296, 2006.

Patten, J. M., Smith, S. R., O'Brien, J. J.: Impacts of ENSO on Snowfall Frequencies in the United States. Wea. Forecasting, 18, 965-980, 2003.

Peters, K., Jakob, C., Davies, L., Khouider, B., Majda, A. J.: Stochastic Behavior of Tropical Convection in Observations and a Multicloud Model. J. Atmos. Sci., 70, 3556–3575, 2013.

Philander, S. G. H.: El Niño Southern Oscillation Phenomena. Nature, 302, 295-301, 1983.

Polygiannakis, J. M., Moussas, X., Sonett, C. P. A Nonlinear RLC Solar Cycle Model. Sol. Phys. 163, 193–203, 1996.

Polygiannakis, J., Preka-Papadema, P., Moussas, X.: On Signal–noise Decomposition of Time-series using the Continuous Wavelet Transform: Application to Sunspot Index. MNRAS, 343, 725-734, 2003.

Rial, J. A., Anaclerio, C. A.: Understanding Nonlinear Responses of the Climate System to Orbital Forcing. Quat. Sci. Rev., 19, 1709-1722, 2000.

Ross, A. C., Najjar, R. G., Li, M., Mann, M. E., Ford, S. E., Katz, B.: Sea-level Rise and Other Influences on Decadal-scale Salinity Variability in a Coastal Plain Estuary. Estuarine, Coastal, and Shelf Science, 79-92, 2015.

Rusu, M. V.: The Asymmetry of the Solar Cycle: A result of Non-linearity, Adv. Sp. Res., 40, 1904-1911, 2007.

Rutherford, S., D'Hondt, S.: Early Onset and Tropical Forcing of 100,000-year Pleistocene Glacial Cycles. Nature, 408, 72-75, 2000.

Schimanke, S., Körper, J., Spangehl, T., and Cubasch, U.: Multi-decadal Variability of Sudden Stratospheric Warmings in an AOGCM. Geophys. Res. Lett., 38, doi: 10.1029/2010GL045756, 2011.

Schulte, J. A., Duffy, C., and Najjar, R. G.: Geometric and Topological Approaches to Significance Testing in Wavelet Analysis. Nonlin. Processes Geophys., 22, 139-156, 2015.

Schulte, J. A., Najjar, R. G., and Li, M.: The Influence of Climate Modes on Streamflow in the Mid-Atlantic Region of the United States, submitted to: Journal of Hydrology: Regional Studies, May 2015.

Schulte, J. A.: Cumulative Areawise Testing in Wavelet Analysis and its Application to Geophysical Time Series, submitted to: Nonlinear Processes in Geophysics, May 2015.

Schneider, N. and Cornuelle, B. D.: The Forcing of the Pacific Decadal Oscillation. J. Climate, 18, 4355-4373, 2005.

Seager, R., Pederson, N., Kushnir, Y., Nakamura, J., Jurburg, S.: The 1960s Drought and the Subsequent Shift to a Wetter Climate in the Catskill Mountains Region of the New York City Watershed. J. Climate, 25, 6721–6742, 2012.

Seager, R., Kushnir, Y., Nakamura, J., Ting, M., Naik, N.: Northern Hemisphere Winter Snow Anomalies: ENSO, NAO and the Winter of 2009/10. Geophys. Res. Lett. 37, L14703, doi:10.1029/2010GL043830, 2010.

Seager, R., Kushnir, Y., Herweijer, C., Naik, N., Velez, J.: Modeling of Tropical Forcing of Persistent Droughts and Pluvials over Western North America: 1856–2000. J. Climate, 18, 4065-4088, 2005.

Serreze, M. C., Clark, M. P., McGinnis, D. L., Robinson, D. A.: Characteristics of Snowfall over the Eastern Half of the United States and Relationships with Principal Modes of Low-Frequency Atmospheric Variability. J. Climate, 11, 234-250, 1998.

Sheridan, S. C.: North American Weather-type Frequency and Teleconnection Indices. Int. J. Climatol., 23, 27-45, 2003.

Steinberg, N., Suszkowski, D. J., Clark, L., Way, J.: Health of the Harbor: The First Comprehensive Look at the State of the NY/NJ Harbor Estuary. A Report to the NY/NJ Harbor Estuary Program. Tech. Rep., Hudson River Foundation, New York, 2004.

Taylor, A.H., Jordan, M. B., and Stephens, J. A.: Gulf Stream Shifts Following ENSO Events. Nature, 393, 638, 1998.

Terradellas, E., Soler, M. R., Ferreres, E., and Bravo, M.: Analysis of Oscillations in the Stable Atmospheric Boundary Layer Using Wavelet Methods. Boundary-Layer Meteorol., 114, 489–518, 2005.

Timmermann A.: Decadal ENSO Amplitude Modulations: a Nonlinear Paradigm. Global Planet Change, 37, 135-156, 2003.

Thompson, D. W. J. and Wallace, J. M.: The Arctic Oscillation Signature in the Wintertime Geopotential Height and Temperature Fields. Geophys Res. Lett., 25, 1297–1300, 1998.

Thompson, D. W. and Wallace, J. M., 2001. Regional Climate Impacts of the Northern Hemisphere Annular Mode. Science, 293, 85-89.

Torrence, C. and Compo, G. P.: A Practical Guide to Wavelet Analysis, Bull. Am. Meteorol. Soc., 79, 61–78, 1998.

Torrence, C. and Webster, P. J.: Interdecadal Changes in the ENSO–Monsoon System. J. Climate, 12, 2679–2690, 1999.

Trenberth, K. E.: The Definition of El Niño., Bull. Amer. Meteor. Soc., 78, 2771-2777, 1997.

Trenberth, K. E.: Recent Observed Interdecadal Climate Changes in the Northern Hemisphere. Bull. Amer. Meteor. Soc., 71, 988-993, 1990.

Trenberth, K. E.: Signal Versus Noise in the Southern Oscillation. Mon. Wea. Rev., 112, 326-332, 1984.

Trenberth, K. E. and Hurrell, J. W.: Decadal Atmosphere–ocean Variations in the Pacific. Climate Dyn., 9, 303–319, 1994.

Trenberth, K. E. and Stepaniak, D. P.: Indices of El Niño Evolution. J. Climate, 14, 1697-1701, 2001.

Trenberth, K. E. and Shea, D. J.: Atlantic Hurricanes and Natural Variability in 2005. Geophys. Res. Lett., 33, L12704, doi:10.1029/2006GL026894, 2005.

Tropea, C., Yarin, A. L., and Foss, J. F. (Eds.): Springer Handbook of Experimental Fluid Mechanics. Springer, Berlin, Germany, 1557, 2007.

USDA, 2000. Preparing for Drought in the 21st Century, A Report of the Drought Policy Commission, 48 pp.

Van Milligen, B. P., Sánchez, E., Estrada, T., Hidalgo, C., Brañas, B., Carreras, B., García, L.: Wavelet Bicoherence: A New Turbulence Analysis Tool. Phys. Plasmas, 2, 3017-3032, 1995.

Vautard, R., Yiou, P., Ghil, M.: Singular-spectrum Analysis: A Toolkit for Short, Noisy Chaotic Signals, Physica D: Nonlinear Phenomena, 58, 95-126, 1992.

Velasco, V. M. and Mendoza, B.: Assessing the Relationship between Solar Activity and Some Large Scale Climatic Phenomena, Adv. Sp. Res., 42, 866–878, 2008.

Wallace, J. M., and Gutzler, D. S.: Teleconnections in the Geopotential Height Field during the Northern Hemisphere Winter. Mon. Wea. Rev., 109, 784–812, 1981.

Wang, C. and Picaut, J. (Eds. Wang, C., Xie, S.P. and Carton, J.A.): Understanding ENSO 20 physics—A review, Earth's Climate AGU, Washington, D. C, 2004.

Wanner, H., Brönnimann, S., Casty, C., Gyalistras, D. Lutenbacher, J., Schmutz, C., Stephenson, D. B., Xoplaki, E., 2001. North Atlantic Oscillation - concepts and studies. Surv. Geophys. 22, 321-381.

Watson, P. A. G. and Gray, L. J.: How Does the Quasi-Biennial Oscillation Affect the Stratospheric Polar Vortex?. J. Atmos. Sci., 71, 391–409, 2014.

Weerahandi, S.: Exact Statistical Methods for Data Analysis, Springer, pp. 329, 2003.

Whitney, M. M.: A Study on River Discharge and Salinity Variability in the Middle Atlantic Bight and Long Island Sound. Cont. Shelf Res., 30: 305-318, 2010.

Wilson, M., Meyers S. D., Luther, M. E.: Synoptic Volumetric Variations and Flushing of the Tampa Bay Estuary, Clim. Dyn., 42, 1587-1594, 2014.

Willmott, C. J., Matsuura, K. 2001. Terrestrial Precipitation: 1900-2010 Gridded Monthly Time Series (1900-2010),

http://climate.geog.udel.edu/~climate/html_pages/Global2011/Precip_revised_3.02/README.GlobalTsP2011.html

Wilks, D. S.: Resampling Hypothesis Tests for Autocorrelated Fields. J. Clim., 10, 65–82, 1997.

Wilks, D. S.: On "Field Significance" and the False Discovery Rate, J. Appl. Meteor. Climatol., 45, 1181-1189, 2006.

Worsby, F. M., Duckham, M.: GIS: A Computing Perspective, CRC Press, Boca Raton, FL, 448 pp., 2004.

Xu, J., Long, W., Wiggert, J. D., Lanerolle, L. W., Brown, C. W., Murtugudde, R., Hood, R. R.: Climate Forcing and Salinity Variability in Chesapeake Bay, USA. Estuar. Coast, 35, 237-261, 2012.

Yarnal, B. and Leathers, D. J.: Relationships between Interdecadal and Interannual Climatic Variations and their Effect on Pennsylvania Climate. Ann. Assoc. Am. Geogr., 78, 624-641, 1988.

Zhang, Q., Xu, C., Jiang, T., and Wu, Y.: Possible Influence of ENSO on Annual Maximum Streamflow of the Yangtze River, China, J. Hydrol., 333, 265–274, 2007.

Ziegler, G. M.: Lectures On Polytopes, Graduate Texts in Mathematics, 152 pp., Springer, New York, 370, 1995.

Vita

Justin Alexander Schulte

Justin Schulte was born on May, 16 1988 in Bethlehem, PA. His interest in meteorology started at a young age and was sparked by the Blizzard of 1996. Justin attended Pennsville Memorial High School where he ran cross-country and track and field competitively. Late in his high school career he developed a passion for mathematics and started studying mathematics in preparation for college. After graduating high school in 2006, he attended Kean University and majored in meteorology. It was in college where he developed an interest in fluid dynamics. Besides course work, Justin participated in track and field and cross country for the university and also tutored mathematics and physics. In 2010, he graduated from Kean University magna cum laude with a major in meteorology and a minor in mathematics.

After graduating from Kean University, Justin attended graduate school at The Pennsylvania State University to earn a doctoral degree in meteorology. He developed a passion for estuarine science and time series analysis during graduate school, which led him to pursue fundamental problems in wavelet analysis. In 2014, he was awarded the Hudson River Foundation Fellowship to continue his research on climate impacts on estuaries. Justin presented much of his work at conferences such as the Mid-Atlantic Bight Physical Oceanography and Meteorological Meeting and the Ocean Sciences meeting. He also gave numerous talks and seminars on wavelet analysis and related statistical methods.

On Justin's free time he enjoys running marathons, learning foreign languages, hiking, and relaxing at the beach.