2005

Indexing, Mode Definition, and Signal Extraction in Climate Research: Analysis and Applications Involving the MJO, the AO, and ENSO

A. (Anthony) Arguez
INDEXING, MODE DEFINITION, AND SIGNAL EXTRACTION IN CLIMATE RESEARCH: ANALYSIS AND APPLICATIONS INVOLVING THE MJO, THE AO, AND ENSO

By

ANTHONY ARGÜEZ

A Dissertation submitted to the Department of Meteorology in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Degree Awarded: Fall Semester, 2005
The members of the Committee approve the dissertation of Anthony Argüez defended on November 7, 2005.

James J. O’Brien  
Professor Directing Dissertation

James B. Elsner  
Outside Committee Member

Fei-Fei Jin  
Committee Member

Kwang-Yul Kim  
Committee Member

Guosheng Liu  
Committee Member

Sharon E. Nicholson  
Committee Member

The Office of Graduate Studies has verified and approved the above named committee members.
ACKNOWLEDGMENTS

Several sources provided funding for this compilation of research projects: the NOAA Office of Global Programs, NASA through Oregon State University’s Seawinds project, the Florida State University’s Office of Graduate Studies (University and Delores Auzenne Fellowships), NASA’s Earth Science Enterprise initiative, and the National Climatic Data Center. Additional recognition goes to Remote Sensing Systems for their role in QuikSCAT data processing, as well as NCEP for providing the Reanalysis data. Computer programming codes were provided by Drs. Kim, Krishnamurti, and Ahlquist of Florida State University’s Department of Meteorology.

Several individuals offered valuable advice and support throughout the dissertation process: Peng Yu, Shawn Smith, Khara Santorelli, Justin Brolley, Justin Schoof, Hilawe Semunegus, Jorge Zavala-Hidalgo, Ajaya Mohan-Ravindran, Anthony Barnston, Robert Livezey, and anonymous reviewers of the MJO paper. Drs. Bourassa, Kim, and Jin offered valuable guidance in the course of the manuscript’s preparation. I am very thankful to the entire committee, Drs. O’Brien, Elsner, Jin, Kim, Liu, and Nicholson, for agreeing to participate in this memorable journey. I would particularly like to thank Dr. Kim for his assistance and encouragement. I am especially indebted to my major professor, Dr. O’Brien, for his support, guidance, for always looking out for me, and for helping to stimulate my “intellectual curiosity.”
# TABLE OF CONTENTS

List of Figures .......................................................................................................................................................v
Abstract ...................................................................................................................................................................viii

INTRODUCTION .........................................................................................................................................................1

1. CYCLO-STATIONARITY OF THE ARCTIC OSCILLATION .......................................................... 5

2. DETECTION OF THE MJO FROM QUIKSCAT ........................................................................ 38

3. AN IMPACT-BASED PERSPECTIVE OF ENSO ........................................................................... 50

4. TIME SERIES FILTERING NEAR ENDPOINTS ....................................................................... 79

CONCLUSION .......................................................................................................................................................... 98

REFERENCES ..........................................................................................................................................................101

BIOGRAPHICAL SKETCH ..................................................................................................................................104
LIST OF FIGURES

1.1: The Arctic Oscillation Mode ................................................................. 17
1.2: The Second EOF ................................................................................. 18
1.3: The Third EOF .................................................................................. 19
1.4: Monthly Variances ........................................................................... 20
1.5: Leading CSEOF Using 20 EOF Modes ............................................. 21
1.6a: Leading CSEOF Using 10 EOF Modes ............................................ 22
1.6b: Leading CSEOF Using 30 EOF Modes ............................................ 23
1.6c: Leading CSEOF Using 40 EOF Modes ............................................ 24
1.6d: Leading CSEOF Using 50 EOF Modes ............................................ 25
1.7a: Leading CSEOF Using 7 EOF Modes ................................................ 26
1.7b: Leading CSEOF Using 8 EOF Modes ................................................ 27
1.8: The Eighth EOF ................................................................................. 28
1.9: The Tenth EOF ................................................................................ 29
1.10: CSEOF of the AO .......................................................................... 30
1.11: AO PC Time Series .......................................................................... 31
1.12: CSEOF PC Spectra .......................................................................... 32
1.13: Regressed Winds ............................................................................ 33
1.14: Regressed Air Temperatures ............................................................ 36
1.15: Regressed Precipitation ................................................................. 37
2.1: The Fourth Leading EXEOF .............................................................. 45
2.2: The Fourth Leading EXEOF PC ....................................................... 46
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3</td>
<td>NCEP Reanalysis MJO PC</td>
<td>47</td>
</tr>
<tr>
<td>2.4</td>
<td>Intraseasonal Energy</td>
<td>48</td>
</tr>
<tr>
<td>2.5</td>
<td>Longitude-Time Plots</td>
<td>49</td>
</tr>
<tr>
<td>3.1</td>
<td>Pacific SST</td>
<td>64</td>
</tr>
<tr>
<td>3.2</td>
<td>SST Correlations</td>
<td>65</td>
</tr>
<tr>
<td>3.3</td>
<td>The Impact Perspective – Air Temperature</td>
<td>66</td>
</tr>
<tr>
<td>3.4</td>
<td>The Impact Perspective – Precipitation</td>
<td>67</td>
</tr>
<tr>
<td>3.5</td>
<td>The Indicator Perspective – Air Temperature</td>
<td>68</td>
</tr>
<tr>
<td>3.6</td>
<td>The Indicator Perspective – Precipitation</td>
<td>69</td>
</tr>
<tr>
<td>3.7</td>
<td>EOF Decomposition over US – Air Temperature</td>
<td>70</td>
</tr>
<tr>
<td>3.8</td>
<td>EOF Decomposition over US – Precipitation</td>
<td>71</td>
</tr>
<tr>
<td>3.9</td>
<td>EOF Decomposition over Florida – Air Temperature</td>
<td>72</td>
</tr>
<tr>
<td>3.10</td>
<td>EOF Decomposition over Florida – Precipitation</td>
<td>73</td>
</tr>
<tr>
<td>3.11</td>
<td>EOF Decomposition over the Southeast – Air Temperature</td>
<td>74</td>
</tr>
<tr>
<td>3.12</td>
<td>EOF Decomposition over the Southeast – Precipitation</td>
<td>75</td>
</tr>
<tr>
<td>3.13</td>
<td>EOF Decomposition over the Upper Midwest – Air Temperature</td>
<td>76</td>
</tr>
<tr>
<td>3.14</td>
<td>EOF Decomposition over the Pacific Southwest – Precipitation</td>
<td>77</td>
</tr>
<tr>
<td>3.15</td>
<td>EOF Decomposition over the Pacific Northwest – Air Temperature</td>
<td>78</td>
</tr>
<tr>
<td>4.1</td>
<td>AO Filter Series</td>
<td>89</td>
</tr>
<tr>
<td>4.2</td>
<td>ENSO Filter Series</td>
<td>90</td>
</tr>
<tr>
<td>4.3</td>
<td>MJO Filter Series</td>
<td>91</td>
</tr>
<tr>
<td>4.4</td>
<td>AO RMS Errors</td>
<td>92</td>
</tr>
<tr>
<td>4.5</td>
<td>Typical AO Results</td>
<td>93</td>
</tr>
</tbody>
</table>
ABSTRACT

There are two objectives of the present study. The primary objective is to undertake the following research projects involving the Arctic Oscillation (AO), the El Niño Southern Oscillation (ENSO), and the Madden Julian Oscillation (MJO): (1) an assessment of the utility of using Cyclo-stationary empirical orthogonal function (CSEOF) analysis to define the AO, (2) an empirical analysis of ENSO impacts based on varying indicator and impact regions, (3) detection and extraction of the MJO signal from QuikSCAT, and (4) the development of a general algorithm for determining optimal filter weights for time series endpoints. A secondary objective is to enumerate the statistical and analytical treatments of the AO, ENSO, and the MJO. This will include comparisons of how these three modes are defined (including their indices) and extracted from geophysical data sets.

The AO is defined using empirical orthogonal function (EOF) analysis of sea level pressure north of 20°N. The resulting spatial pattern and time series captures the regional influence of its precursor, the North Atlantic Oscillation (NAO), which is a measure of mid-latitude zonal winds over the North Atlantic. ENSO was originally defined as the pressure difference between Tahiti and Darwin, Australia: the Southern Oscillation Index. Scientists now primarily use sea surface temperature (SST) anomalies averaged over one of the Niño regions as ENSO indices. The MJO was originally observed using spectral analysis of zonal wind time series in the Indian Ocean and Western Pacific. Present day researchers use extensions of EOF analysis to construct MJO time series. For all three climate modes, the creation of high quality space-time data sets has allowed for more sophisticated indices, supplanting the simpler point-based metrics.

For the AO project, the cyclo-stationarity of Northern Hemisphere sea level pressure variability is considered. CSEOF analysis is an extension of EOF analysis that allows multiple spatial maps per mode. It accomplishes this by cyclically extending the covariance matrix based on a parameter called the nested period. By using a nested period of 12, a climate mode can be decomposed into a series of 12 monthly maps and an associated time series. Unlike EOF PC time series, which typically have larger amplitudes during winter months, CSEOF PC time series do not favor a particular season because the physical evolution of the climate mode is posited in the loading vectors (the maps) rather than the time series. This is impossible to accomplish with
regular EOF analysis because it relegates each mode to one single map. A compelling case is made for a cyclo-stationary interpretation of AO variability. The leading CSEOF mode includes AO-type variability during a winter regime, as well as a summer regime characterized by pressure anomalies centered over Mongolia and associated with rainfall variability in the vicinity of the Ganges delta and eastern China. EOF modes that contribute to the resulting maps of the leading CSEOF mode are identified, including the eighth mode, which is deemed responsible for the summertime Asian pattern. CSEOF analysis of the AO mode only exemplifies the power of CSEOF analysis with regard to transferring a mode’s physical evolution from a PC time series to a series of loading vectors.

For the ENSO project, traditional ENSO impact analysis was recast to investigate the teleconnections between U.S. climate and varying indicator regions of SST anomalies in the tropical Pacific. This serves the dual purpose of finding a targeted indicator region for a particular impact zone (i.e. a localization of the teleconnection pattern) and indirectly assessing the viability of well-established ENSO indices (i.e. the Niño indices). Based on a selection of impact grid points with known ENSO responses, it appears that the most appropriate indicator region often varies from one impact grid point to another, as well as from warm SST phase to cold SST phase. In addition, air temperature composites behave differently than precipitation composites. In order to simultaneously consider the “impact perspective” detailed above with the typical “indicator perspective” (in which climate impacts are computed based on the well-established Niño indices), EOF analysis of composited climate fields, conditioned on SST phase, as functions of indicator region and impact zone was performed. The resulting modes represent indicator-impact pairs. Each mode has an impact amplitude function (a spatial temperature or precipitation anomaly signature over the impact region) and an associated indicator weighting function, which modulates the impact amplitude function based on the location of the indicator region. Based on this approach, the unusual yet well-established La Niña air temperature impact over the U.S. when using the Niño 1+2 region is accounted for as the superposition of two EOF modes. In addition, a teleconnection between tropical Pacific SST and Southeastern U.S. temperature anomalies is documented that is not related to ENSO.

For the MJO project, wind data from the SeaWinds instrument on the QuikSCAT satellite are investigated to ascertain how well the surface manifestation of the MJO can be resolved. The MJO signal is detected in non-filtered gridded data using Extended EOF analysis of the zonal
wind field, overshadowed by annual, semi-annual, and monsoon-related modes. After bandpass filtering with Lanczos weights, MJO signals are clearly detected in several kinematic quantities, including the zonal wind speed, the zonal pseudostress, and the velocity potential. Extraction of the MJO using QuikSCAT winds compares favorably with extraction using NCEP Reanalysis 2, except that the QuikSCAT signal appears to be more robust.

For the filtering project, least squares techniques are utilized to retain endpoint intervals that are normally discarded due to filtering with convolutions in the time domain. The techniques minimize the errors between the pre-determined frequency response function (FRF) of interior points with FRF’s that are to be determined for each position in the endpoint zone. The least squares techniques are differentiated by their constraints: (1) unconstrained, (2) equal-mean constraint, and (3) an equal-variance constraint. The equal-mean constraint forces the new weights to sum up to the same value as the pre-determined weights. The equal-variance constraint forces the new weights to be such that, after convolved with the input values, the expected variance is identical to the expected variance of the interior points. These 3 least squares methods are tested under three separate filtering scenarios and compared to each other as well as to the spectral filtering method, which is the standard of comparison. The results indicate that all 4 methods (including the spectral method) possess skill at determining suitable endpoints estimates. However, both the unconstrained and equal-mean schemes exhibit bias toward zero near the terminal ends due to problems with appropriating variance. The equal-variance and spectral techniques do not show evidence of this attribute and were never the worst performers. The equal-variance method showed great promise in the ENSO project involving a 5-month running mean filter, and performed at least on par with the other methods for virtually all time series positions in all three filtering scenarios.
INTRODUCTION

The term ‘climate’ is used in many ways in the atmospheric sciences. It can refer to both the average atmospheric conditions as well as the range of conditions for a particular region. Atmospheric scientists actively differentiate between climate and weather, which refers to the current state of atmospheric conditions, whereas climate is concerned with the distribution of weather conditions. Hence climate, by its very nature, is inextricably linked with statistics. In fact, climate can be defined as the statistics of weather. An important aspect of ‘climate’ is that the distribution of atmospheric conditions depends on the time period over which the distribution is constructed.

Climatologists came to realize that these distributions could change with time, leading to an explosion of interest in climate change. The field of climate change includes the study of climate trends (which includes the controversial tenets of global warming) and climate variability. Climate variability is usually thought of in terms of distinct modes that fluctuate, or oscillate, about a mean climate state. These are referred to as climate modes, or climate oscillations. Climatologists are charged with detecting these modes and evaluating their impacts to the mean climate state. Therefore, the study of climate variability relies heavily on statistics as well.

A fundamental issue climatologists have to contend with is how to represent, or define, climate modes. These are not conventional dictionary definitions; climate modes are defined using mathematical equations. How we define climate modes has mathematical consequences that can impair our ability to assess climate impacts, as well as distort our understanding of the underlying physics involved. In practice, climate mode definitions are limited by the data that are available. Therefore, it is quite common for climate definitions to evolve based on improvements in the spatial and temporal coverage of large-scale datasets.

The simplest climate mode definition is a time series index. These are time series that are constructed (not simply measured) to quantify the state, or intensity, of a climate mode. A meteorological example is the zonal index, which strives to measure the intensity of mid-latitude westerlies for a particular point in time. The Southern Oscillation Index (SOI) and the North Atlantic Oscillation (NAO) index are examples of indices used to quantify the intensities of ENSO and the NAO, respectively. Both use sea level pressure measurements at 2 stations to
estimate horizontal pressure gradients, in order to infer variations in zonal winds (ENSO over the tropical West Pacific and NAO over the mid-latitudes of the North Atlantic). These fluctuations in zonal winds, in turn, are associated with climate impacts that can be thousands of kilometers away from where they are measured.

More complicated climate mode definitions consider the spatial domain as well as the temporal domain. The most common such type of definition involves empirical orthogonal function (EOF) analysis. This tool is often termed factor analysis or principal component analysis (PCA). This method yields a set of maps that have associated time series. The time series represents the changes in amplitude, or intensity, of the climate mode. The map represents the mode’s spatial signature. The Arctic Oscillation (AO), for example, is defined using EOF analysis. ENSO and the Madden Julian Oscillation (MJO) can be defined in this way as well. Several extensions to EOF analysis, most of which accommodate multiple spatial maps per mode, are also utilized to define climate modes.

Climate modes have characteristic regions over which they are defined. For example, ENSO is defined in the equatorial Pacific. In addition, climate modes have characteristic timescales. In the case of the MJO, this is the defining characteristic. With ENSO and the AO, the variability is spread across a broad range of timescales, but we can choose to focus on a certain band only. This means that part of defining climate modes deals with what timescales we want to extract. Therefore, time series filtering is an integral part of climate mode definition.

In the present study, the way in which climate modes are defined, their indices constructed, and their timescales extracted is investigated. This is accomplished by performing individual projects on the AO, MJO, and ENSO. The AO project involves using an alternate definition than the one currently used. The MJO project deals with detecting its signature in a new data set. The ENSO project revolves around using different time series indices for different impact zones. A fourth project introduces a new technique for filtering near time series endpoints, and utilizes AO, MJO, and ENSO time series as example climate indices. While the primary goal is to undertake four distinct projects, a secondary goal is to use the results as evidence of the ramifications that climate mode definitions, indexing, and signal extraction can have in the field of climate variability. In the remainder of this section, additional background on the AO, MJO, ENSO, and time series filtering in general is provided.
The Arctic Oscillation

The AO is heralded as the leading mode of variability in the extratropical Northern Hemisphere. It is customarily defined as the leading empirical orthogonal function (EOF) of sea level pressure (SLP) in the Northern Hemisphere north of 20°N. However, the AO has also been defined in a similar manner using other variables, such as wind and geopotential height (Thompson and Wallace 2000; hereafter TW). The primary promoters of the AO are David Thompson and John Wallace, who initially coined the term “Arctic Oscillation,” but now prefer the phrase “Northern Hemisphere Annular Mode.” This proposed name change was intended to draw attention to the geographical structure of the AO pattern and point out the existence of the AO’s counterpart in the Southern Hemisphere, which is aptly named the “Antarctic Oscillation” or the “Southern Hemisphere Annular Mode.” An important, albeit controversial, attribute of the AO paradigm is that it encompasses the regionalized variability of the North Atlantic Oscillation (NAO) due to the AO’s hemisphere-wide nature. In fact, Wallace (2000) forcefully argues that the NAO is merely the Atlantic manifestation of the broader AO. As is true with the NAO, the AO plays an important role in regulating European and North American winters (e.g. Higgins et al. 2000). Moreover, unlike the MJO and to some extent ENSO, the AO varies across numerous timescales (perhaps an artifact of the way it is defined).

The Madden Julian Oscillation

The MJO’s defining characteristic is its intraseasonal timescale. In fact, the MJO’s namesakes originally called it the 40-50 day oscillation (Madden and Julian 1971). MJO definitions usually require prior removal of timescales outside of the intraseasonal range. In addition to its timescale, another important and unique attribute of the MJO is its eastward, equatorially-trapped propagation. The MJO’s defining variable is outgoing longwave radiation (OLR) because MJO activity is seen as a metric of intraseasonal convective anomalies. However, the signal is also conspicuous in zonal winds. Although the MJO is technically present near the equator around the globe, it is considerably more conspicuous in the Indian and Western Pacific Oceans.
El Niño Southern Oscillation

No other climate mode has received as much attention as ENSO. Originally known as the Southern Oscillation, because it was first identified in the southern region of the tropical Pacific Ocean, it was later called ENSO due to the hysteria surrounding the El Niño phenomenon, which happens to be only one of three phases of ENSO. Thus, the strange fusion of El Niño and Southern Oscillation leads us to the unlikely acronym of ENSO (as opposed to, say, ITPO: the Interannual Tropical Pacific Oscillation). ENSO plays an important role in worldwide climate variability. The other 2 phases of ENSO (La Niña and Neutral), despite lacking in notoriety, are oftentimes associated with the most extreme conditions in a region. An example is the well-documented link between La Niña (the phase opposite of El Niño) and North Atlantic hurricane activity (Gray 1984). Hence, the term ENSO can be considered a misnomer, since its name has more historical meaning than intrinsic relevance to the actual phenomenon it describes. ENSO is normally defined in terms of box-averaged sea surface temperature (SST) anomalies in the eastern Pacific Ocean. In addition to hurricane activity, ENSO has well-documented impacts on interannual variability of temperature and precipitation in Latin America, the United States, Africa, Australia, and elsewhere (see Hanley et al. 2003).

Time Series Filtering

Filtering is an essential tool for analyzing geophysical time series. Common applications include extraction of salient timescales, suppression of high-frequency noise, and smoothing spectra in order to increase statistical confidence in spectral estimates. The defining characteristic of a filter is its frequency response function (FRF), which dictates the portion of variability for each timescale that gets passed to the output time series (as well as the timescales that are ‘filtered out’). Time series filtering can be done in the frequency or time domains. The problem with filtering in the time domain is that points are lost at the endpoints because a convolution is necessary. In the present study, variable filter weights in the endpoint intervals are determined such that the FRF optimally (in the least squares sense) reproduces the predetermined FRF of the interior points, thus allowing endpoints to be retained.
CHAPTER 1

CYCLO-STATIONARITY OF THE ARCTIC OSCILLATION

Background

The need to consolidate geophysical data into the most energetic modes prompted the application of EOF analysis to geophysical space-time data sets. EOF analysis is a variance decomposition technique that reduces an M by N (where N ≥ M) data matrix \(D\) into a linear combination of M spatial maps (assuming space is collapsed into 1 dimension) and M associated time series:

\[
D(\lambda, \phi, t) = \sum_{i=1}^{M} E_i(\lambda, \phi) P_i(t)
\]

where \(\lambda\) is the longitude, \(\phi\) is the latitude, \(t\) represents time, \(E_i\) is the ith EOF pattern, and \(P_i\) is the ith principal component time series. Note that each EOF pattern is a function of space only. As a result, a geophysical mode is forced to be represented as a fixed spatial pattern whose amplitude undulates in time. In this manner, TW defined the Arctic Oscillation.

A new definition of the AO, in which the spatial pattern is allowed to change, may be a valuable alternative to the TW paradigm. Several extensions to EOF analysis that accommodate propagating patterns and multiple modal spatial patterns have been proposed and applied to geophysical data sets. These include Extended EOF analysis (see Chapter 2) and Complex EOF analysis (e.g. Horel 1984). However, a promising new technique is Cyclo-stationary EOF (CSEOF) analysis. CSEOF analysis involves a covariance matrix that is periodically extended (see Kim 2002). The periodicity parameter is called the nested period, \(d\). The decomposition of the space-time data set under CSEOF analysis involves a linear combination of cyclic spatial patterns and a time evolution:
\[ D(\lambda, \phi, t) = \sum_{i=1}^{M} E_i(\lambda, \phi, t) P_i(t) \] \hspace{1cm} (1.2)

Notice that the spatial pattern, \( E \), is now time dependent. It is now referred to as a Bloch function and has a nested period of \( d \). By setting \( d \) equal to 12 and applying CSEOF analysis to monthly data, each mode will have a unique spatial pattern associated for each month. By allowing the AO’s spatial pattern to change from month to month, it is hoped that new insight into AO variability can be attained, and that the consequences of the TW definition can be assessed.

**Data and Methods**

Fifty-five years of NCEP Reanalysis SLP data are used to compute a new CSEOF representation of the AO. The data consist of monthly-averaged SLP anomalies (the 55-year climatology has been removed) on a 2.5° by 2.5° grid from 1948-2002. Following the TW definition, only points north of 20°N are utilized. Initially, the nested period is set to 12. Borrowing from Kim (2002), an exercise will be conducted to check the viability of using 12 for the nested period. Using conventional EOF analysis, the TW definition of the AO will be reproduced, which only utilizes winter values. The EOF-type AO definition will also be recomputed using the full-year anomalies, as well as the ‘seasonal anomalies’ employed by Thompson and Wallace, which are succinctly described by Deser (2000).

NCEP Reanalysis wind, temperature, and precipitation (precipitation is on a 1.875° by 1.875° grid) will be used to find patterns that are physically consistent with the CSEOF definition of the AO. This is accomplished using regression analysis of the PC time series that result from CSEOF analysis of wind, temperature, and precipitation (see Kim 2002). These three variables are also projected onto my version of the EOF-based AO. The CSEOF and EOF definitions; including their co-evolution with wind, temperature, and precipitation; will be compared and the practicality of using the CSEOF definition versus the TW definitions will be addressed.
EOF Analyses

AO Definition: The Leading Mode

The EOF-based AO spatial pattern using SLP anomalies for the full year is shown in Fig. 1.1a. Projections of temperature, winds, and precipitation are presented in Fig. 1.1b-d, respectively. The AO signature accounts for 18% of the variance decomposition and is consistent with the pattern defined by Thompson and Wallace. The AO spatial pattern is marked by a dominant action center covering the Arctic and opposite signed anomalies stretching zonally in the midlatitudes of the Atlantic and Pacific Oceans. The Arctic and Atlantic centers are most intense when using winter anomalies, whereas the Pacific center is most intense when using the “intraseasonal” anomalies. This is not expected since, as described in Deser (2000), the intraseasonal anomalies were utilized to minimize influence from the Pacific (i.e. ENSO). Despite this oddity, using wintertime or intraseasonal anomalies do not alter the spatial pattern in a materially significant way. Henceforth, only the full-year AO definition will be considered.

The temperature anomalies associated with the AO spatial patterns can be accounted for by advective and radiative effects. Over the North Atlantic, changes in maritime influence (i.e. anomalous geostrophic advection in the vicinity of a large-scale land-ocean interface) are associated with warmer (colder) conditions in onshore (offshore) wind regimes. Similar reasoning can account for the thermal effects in the eastern two-thirds of the United States. However, strong temperature anomalies in Northern Asia appear to be a radiative effect, with colder (warmer) temperature anomalies associated with a strengthening (weakening) of the climatological Siberian High during winter (although it must be remembered that this region continues to be the coldest place on earth during winter, even when there are anomalously warm conditions).

Precipitation anomalies are most coherent in the North Atlantic sector. Between 50°N and 70°N over the North Atlantic, where the strongest winds associated with the AO reside, precipitation anomalies appear to be related to moisture flux, as maintained by Thompson and Wallace (2001). Further south (in the vicinity of the North Atlantic’s subtropical action zone), precipitation anomalies are closely aligned with the pressure anomalies, suggesting a radiative impact.
In order to summarize the EOF-based AO pattern and its associated climate, three regions are considered: Labrador (easternmost Canada), Northern Spain, and Siberia. In much of Labrador, the pressure anomalies are fairly weak; however, the large-scale pressure gradient is fairly potent. Anomalous winds blow offshore (onshore) during the high (low) phase, which reduces (increases) maritime influence and results in colder (warmer) temperatures and drier (wetter) conditions. Northern Spain comprises part of the subtropical Atlantic action center. The winds are very close to climatology, and there is just a slight positive correlation between pressure and temperature. However, there does exist a negative correlation between rainfall and pressure. In Siberia, the EOF pattern represents a weakening (strengthening) of the Siberian High during the high (low) phase. There is a great deal of cross-isobaric flow here, with flow directed toward the low-pressure anomaly centers. Warmer (colder) temperatures during the high (low) phase are related to a weaker (stronger) Siberian High and anomalous southerly (northerly) flow.

**EOF Modes 2 and 3**

While the TW definition of the AO receives a great deal of attention because it is the leading mode, subsequent modes of Northern Hemisphere SLP also play a role in dictating climate anomalies. The second and third modes, for instance, account for 10.0% and 9.4%, respectively, jointly accounting for a greater amount of variance in Northern Hemisphere SLP variability than the 18.0% explained by the AO. In this section, these two under-touted modes are considered.

Mode 2 is characterized by a primary action center in the vicinity of the climatological Aleutian Low (which is a wintertime feature), as well as a weaker action center of the same sign located near Spitsbergen (see Fig. 1.2a). An opposite-signed zone is located in the mid-latitudes of the North Atlantic, centered just west of France. Temperature anomalies are largely controlled by anomalous advection. In Northwestern North America, cold (warm) anomalies result due to anomalous northerly (southerly) flow on the eastern side of the Aleutian action center’s positive (negative) pressure anomaly (see Figs. 1.2b-c). A similar, albeit less intense, scenario occurs for temperatures in the vicinity of the Spitsbergen action center. The North Atlantic center does not have a significant temperature impact. Precipitation anomalies are
negatively correlated with the pressure anomalies near all three action zones for the most part, except for the western quadrant of the Aleutian action zone where anomalous flow near the coastlines overcomes the radiative effect (Fig. 1.2d).

For Mode 3, the primary feature is a contrast between the high latitudes of the Eastern and Western Hemispheres (Fig. 1.3a). A secondary action center is located over the North Atlantic centered at 40°N, having the same sign as the zone over Russia (in the Eastern Hemisphere). Much warmer (colder) conditions are experienced in the vicinity of Spitsbergen during the high (low) phase of mode 3, thanks to intense anomalous southerly (northerly) advection between the primary action centers (Figs. 1.3b-c). Precipitation anomalies are negative (positive) over the oceans near anomalous high (low) pressure anomalies (Fig. 1.3d). Drier (wetter) conditions are experienced near the center of the Russian action zone during the high (low) phase, likely reflecting a wintertime radiation impact. Adveotive maritime effects are apparent in the extreme North Atlantic.

Cyclo-stationary EOF Analysis

The Leading CSEOF

Using a nested period of 12 months, Cyclo-stationary EOF analysis of monthly data produces a series of 12 spatial patterns per mode, each corresponding to a different month. As a result, the mode is allowed to reveal its spatial attributes more freely than the strict, one map per mode artifact of EOF analysis. EOF analysis of monthly data typically results in PC time series where the monthly variance of winter months is greater than summer months, regardless of whether the monthly climatology is removed before submission to the EOF routine. The 3 modes outlined above are no different (see Fig. 1.4). Cyclo-stationary EOF analysis with $d=12$ attempts to extract this physical evolution during the year from the PC time series and posit it in the Eigenfunctions. Although $d=12$ is a natural choice for monthly data, there are situations where other nested periods may be more appropriate (e.g. see Kim 2002). Based on the nested period check outlined in Kim (2002), 12 appears to be a suitable choice for the extratropical Northern Hemisphere SLP field, and it is the only nested period considered in the present work.
Cyclo-stationary EOF analysis can be done in EOF space, i.e. using an EOF representation of the data using all or a subset of the leading EOF modes in lieu of the full data matrix values. If a subset of the EOF modes is used, the noise captured by higher-order modes is removed. This can be interpreted loosely as spatial and temporal smoothing of the original data set. In the current investigation, the leading 20 EOF modes (accounting for 87% of the total variance) were passed to the Cyclo-stationary EOF routine. The sensitivity to mode retention is addressed in the next section.

The leading mode of the CSEOF analysis is presented in Fig. 1.5. The most conspicuous feature is the existence of dual regimes: a summer regime and a winter regime. The winter regime (November to March) is clearly reminiscent of AO variability. The most fundamental difference between the winter regime and the AO is the absence or inverted sign of the Aleutian Low action center (except for December). It is noteworthy that the pattern changes noticeably from month to month (see, for example, the Hudson Bay region during February). During summer months (May to September), we see a persistent anomaly centered squarely over Mongolia.

**Sensitivity of EOF Mode Retention**

As mentioned above, Cyclo-stationary EOF can be accomplished in EOF space. This effectively removes higher order EOF noise from consideration and improves computational efficiency. However, the number of EOF modes retained is a significant parameter of the problem for two fundamental reasons. The first is rather intuitive: the addition of a new EOF mode can alter the prevailing CSEOF modes by altering the 12 spatial maps and the PC time series. In the next section, this phenomenon will be explored in the context of identifying EOF modes that alter the CSEOF loading vectors in a conspicuous way. The second consequence of the retention parameter is that a new mode can alter the order of the CSEOF modes since, as with any EOF-based method, the modes are ranked according to percent variance explained after the covariance manipulation is completed. Therefore, what we consider the leading mode may be relegated to a higher order mode based on the sole inclusion of one additional EOF mode. In practice, higher order modes rarely change the fundamental structure of the leading CSEOF modes or their rankings, but it is imperative to verify this nonetheless.
The top 10, 30, 40, and 50 EOF modes account for 73%, 92%, 95%, and 96% of the total variance of the SLP data set, respectively. The CSEOF patterns associated with each representation of the SLP data are shown in Figure 1.6(a-d). When 10 EOF modes are used, the AO pattern is clearly present from November to January. However, the February and March loading vectors have deviated somewhat from the AO, particularly in the Northern Pacific where the sign of the action center has changed. In addition, February also displays a breakdown in the Arctic action center over Canada (this also shows up when 20 modes are retained). From retaining 10 modes to retaining 20 modes (recall Fig. 1.5), the main difference is that in February an action center develops over the North Atlantic with the same sign as action centers in the Northeast Pacific and over Russia. This feature persists through the 50-mode CSEOF. Besides a diminishing of the relative spatial amplitudes of action centers with increasing noise (although the total variance increases when additional EOF modes are retained), the patterns do not undergo any significant changes as the retention parameter increases. In addition, the correlation coefficients between the leading 20-EOF-mode PC and the 10, 30, 40, and 50 mode PC time series are 0.92, 0.91, 0.84, and 0.75, respectively. Therefore, we are confident that we have isolated the true leading mode and that utilizing 20 EOF modes is an acceptable representation for interpretation purposes.

Contributions of Individual EOF Modes

The above exercise regarding mode retention begs a very important question: is it possible to determine the EOF mode that is (most) responsible for the existence of an action center in the loading vectors? For example, is there an EOF mode that imparts the summertime Asian action center in the CSEOF analysis? It is clear that the AO pattern (the leading EOF) contributes to the resulting CSEOF patterns during winter. As we incrementally consider an additional EOF in the CSEOF routine, the resulting changes to the CSEOF can be interpreted in terms of changes imparted by the new addition. We can then analyze the EOF pattern to verify that the spatial amplitude is indeed in the appropriate location.

In the case of the summertime Asian signal, this variability can be attributed to the inclusion of the eighth EOF. The resulting CSEOF patterns when the leading 7 and 8 EOF modes are retained are presented in Figs. 1.7a and 1.7b, respectively. It is evident the difference
a single mode can make. The spatial pattern of the 8th EOF mode follows in Fig. 1.8, making it clear that the mode is indeed contributing to the action zone over Asia. The uniqueness of the February map is related to the 10th mode, as can be seen in Fig. 1.9. The PC time series associated with this mode has maximum monthly variance in February (Fig. 1.4). The changing sign of the North Pacific action center from early winter to late winter cannot be ascribed to any single mode; modes 2 and 8 both make contributions to this feature. Cross-correlations between PC time series associated with the first, eighth, and tenth EOF’s are not significant for any lags, indicating that these 3 contributors to the leading CSEOF are not related to each other.

Cyclo-stationary EOF of the Leading EOF Mode Only

In addition to identifying which EOF modes make large contributions to the resulting CSEOF modes, as was done in the previous section, it is also possible to conduct CSEOF analysis with just one EOF mode. In this case, we consider CSEOF of the leading EOF (the Arctic Oscillation). In this case, the data matrix has been replaced with what are essentially a single spatial amplitude map and a time series that serves as the amplitude of the fixed spatial pattern. Since CSEOF attempts to extract the physical evolution from the stochastic component of the temporal evolution, positing the physical evolution in the loading vectors, we can make additional inference on the practicality of using the EOF-based definition of the AO.

The CSEOF loading vector breakdown of the AO is presented in Figure 1.10. It is noteworthy that the pattern itself is unadulterated from one month to another – the only change is to the spatial amplitude and not the spatial orientation. The main difference between the CSEOF representation of the AO and the TW-defined AO lies in the resulting time series (see Fig. 1.11). Whereas the AO PC time series is noisier, with monthly variances clearly larger during winter months, the CSEOF PC time series is smoother with what can be considered the true amplitude of the AO, as the seasonality of the amplitude function has been extracted (see Fig. 1.4). It should be emphasized that the nested period is not the same as the characteristic timescale, or return frequency. The spectrum of the leading CSEOF of the AO mode shows a strong peak in the vicinity of a 3-year period (Fig. 1.12a), with a considerable amount of spectral energy at longer timescales as well. The leading CSEOF PC when 20 EOF modes are used (Fig. 1.12b),
on the other hand, has spectral energy concentrated across the inter-annual range between 3 and 7 years.

**Physically Consistent Temperature, Wind, and Precipitation Anomalies**

As with the EOF analysis described in Section 3, it is possible to compute the temperature, wind, and rainfall anomalies associated with the leading CSEOF mode. In this case, since we used a nested period of 12, there will be a regression map for each month. However, the presence of 12 spatial maps instead of one per mode precludes the use of the simple projection method used in Section 3. With the projection method, the PC time series of an indicator variable (e.g. pressure) can be matrix multiplied with a space-time data array of an impact variable (e.g. temperature), resulting in a spatial pattern that evolves with the same time series as the indicator variable. However, with CSEOF analysis, we have extracted the physical evolution over the nested period from the PC time series. Since there are 12 distinct spatial patterns, there is no straightforward way of utilizing the projection method. Say we used only February temperature values and projected it onto the February pressure CSEOF. We would be using a target variable that has not had its physical evolution separated from its stochastic evolution. Therefore, it is necessary to use regression analysis after first performing CSEOF analysis of the target variables. It can be shown that, for EOF analysis, the regression and projection methods are identical if all EOF modes are utilized (Kim, personal communication); otherwise, the regression method will exclude higher-order modes of both the target and indicator variables. Following Kim (2002), the following regression analysis is performed:

\[
S(t) = \sum_{i=1}^{30} \alpha_i T_i(t) + \epsilon^T(t) \quad (1.3)
\]

\[
S(t) = \sum_{i=1}^{30} \beta_i W_i(t) + \epsilon^W(t) \quad (1.4)
\]

\[
S(t) = \sum_{i=1}^{30} \gamma_i R_i(t) + \epsilon^R(t) \quad (1.5)
\]
Here, S represents the PC time series of the leading SLP CSEOF, whereas T, W, and R represent the 30 leading PC time series of CSEOF analysis of temperature, wind, and rainfall, respectively. All PC time series are normalized. The goal is to determine the coefficients of $\alpha$, $\beta$, and $\gamma$ while minimizing the error terms. Once these coefficients have been determined, they are linearly combined with their corresponding Bloch functions ($B_i^*$) to arrive at the patterns ($X_i'$) of temperature, wind, and precipitation that are physically consistent with the leading CSEOF mode of SLP:

\[
T'(\lambda, \phi, t) = \sum_{i=1}^{30} \alpha_i B_i^T(\lambda, \phi, t) \tag{1.6}
\]

\[
W'(\lambda, \phi, t) = \sum_{i=1}^{30} \beta_i B_i^W(\lambda, \phi, t) \tag{1.7}
\]

\[
R'(\lambda, \phi, t) = \sum_{i=1}^{30} \gamma_i B_i^R(\lambda, \phi, t) \tag{1.8}
\]

The regressed winds are presented in Fig. 1.13. As before, we observe counterclockwise (clockwise) flow around the periphery of, as well as cross-isobaric flow into (away from), the low (high) pressure anomalies. The regressed temperatures are presented in Fig. 1.14. Wintertime temperatures anomalies appear to be mostly related to advective effects. During wintertime of the high (low) phase in Eastern North America, there is a tendency for onshore (offshore) flow along the US coast and offshore (onshore) flow along the Canadian, resulting in warmer (cooler) conditions in the US east coast and cooler (warmer) conditions along Canada’s east coast. There is also a tendency for Northern and Western Europe to receive onshore (offshore) flow during high (low) phase winters, resulting in warmer (cooler) temps. In Eastern Europe, we find colder (warmer) conditions on the eastern flank of the high (low) pressure anomaly during the high (low) phase, due to the anomalous northerly (southerly) flow there. There is also a fairly consistent signal over Russia, which appears to be related to a superposition of radiative and advective effects. Advective maritime effects are responsible for Alaskan anomalies. Summertime temperature impacts are weaker and not as easy to interpret.

Physically-consistent precipitation anomalies are presented in Figure 1.15. Anomalous winter rainfall in the Southeastern US, especially over the Florida peninsula, is clearly associated
with anomalous maritime influence (via advection). As before, we find a negative correlation between open ocean pressure anomalies and precipitation. Year-round impacts along the western coast of North America and over Europe are consistent with shifts in maritime influence. Impacts in summertime precipitation over Asia are concentrated in two regions: northeastern China and the Indian subcontinent (particularly eastern India and Bangladesh). In NE China, there is a strong meridional wind on the eastern flank of the summertime action center, resulting in wetter conditions during southerly regimes and vice versa. The impact near Bangladesh is not as clear cut. Drier conditions are observed in May and June and wetter conditions in July and August. This suggests a link to the onset of the summer monsoon, consistent with the findings of Kakade and Dugam (2000).

Summary

In the current project, the primary goal was to put the Arctic Oscillation in perspective by considering the leading Cyclo-stationary EOF of extratropical Northern Hemisphere SLP variability. In addition, the AO was considered alongside the second and third leading EOF modes of NH SLP variability. The AO has received a great deal of attention in the scientific literature; however, it is the author’s contention that the TW AO definition has not been fully vetted by the climate community, prompting tacit acceptance of the definition without a thorough scientific dialogue. The AO paradigm is indeed an important research avenue and its contributions to our understanding of global climate and impacts are indisputable. However, the same can be said about the NAO, which TW cast as a regional phenomenon better interpreted inside of the broader AO paradigm. In the same spirit, we propose that the current AO definition may be more adequately represented within the context of the cyclo-stationarity of extratropical Northern Hemisphere SLP Variability.

The leading Cyclo-stationary EOF mode is characterized by two regimes: summer and winter. The winter regime can be considered a slightly modified version of the AO. Note, for example, that in the leading CSEOF a consistent pressure gradient exists between Iceland and Western Europe between November and March, hallmarks of the AO and NAO. In addition, the associated temperature, wind, and rainfall anomalies of the leading CSEOF mode can be interpreted using regression analysis and the concepts of geostrophy, maritime influence, and
advection (just as the projection method is used with EOF modes). The summertime regime is characterized by a fairly static pressure anomaly centered over Mongolia, influencing summertime precipitation in China and India. Recent studies have attributed rainfall anomalies in the regions with AO/NAO variability, and the results suggested therein are consistent with those presented here. Additional study should be dedicated to determining a more comprehensive understanding of this link.

Other facets of Cyclo-stationary EOF analysis were considered. The retention parameter was examined in order to determine the consequences of representing the SLP data set as a subset of EOF modes. It was determined that retaining 20 EOF modes was an acceptable balance between isolating a consistent leading mode and removing higher-order EOF noise. This led naturally to an examination of the role individual EOF modes have on the resulting CSEOF patterns, revealing that EOF modes 8 and 10 (in addition to the obvious role played by the leading EOF mode, i.e. the AO) contributed to the Asian summertime regime and to the uniqueness of the February map, respectively. The Cyclo-stationarity of the AO mode (the leading EOF mode) was examined, highlighting the ability of CSEOF analysis to transfer monthly variance from the EOF’s PC time series to the CSEOF’s loading vectors, resulting in a CSEOF PC time series that is not partial to a particular season.

In summary, identifying and defining climate modes is not a trivial task. The NAO definition was relegated to simple two point differences due to data availability issues. Nevertheless, its discovery was a breakthrough in climate variability studies. The AO paradigm also advanced the science, incorporating the NAO as a regional manifestation. Advances in science often manifest themselves as a series of such integrative innovations, exemplified by Einstein’s incorporation and extension of Newtonian mechanics in offering the theory of relativity. By using a promising new technique, Cyclo-stationary EOF analysis, it is the author’s contention that a more complete understanding of extratropical Northern Hemisphere SLP variability has been attained that assimilates both the AO and NAO paradigms. This new tool should be exploited to make advances in other avenues of climate variability.
Figure 1.1: The Arctic Oscillation Mode. (a) Leading EOF of extratropical Northern Hemisphere SLP anomalies. Commonly referred to as the Arctic Oscillation. (b-d) projections of temperature, wind, and precipitation, respectively.
Figure 1.2: **The Second EOF.** As in Figure 1.1, but for the second leading EOF.
Figure 1.3: **The Third EOF.** As in Figure 1.1, but for the third leading EOF.
Figure 1.4: **Monthly Variances.** Monthly Variances of the PC time series from the leading EOF (purple), EOF 2 (blue), EOF 3 (green), EOF 10 (yellow), the CSEOF of AO EOF (orange), and the leading CSEOF using 20 EOF modes (red). All functions have been scaled by their means. In addition, CSEOF PC time series were offset in opposing directions in order to distinguish them in the figure, since they are both close to unity for all months.
Figure 1.5: Leading CSEOF Using 20 EOF Modes.
Figure 1.6a: Leading CSEOF Using 10 EOF Modes.
Figure 1.6b: Leading CSEOF Using 30 EOF Modes.
Figure 1.6c: **Leading CSEOF Using 40 EOF Modes.**
Figure 1.6d: **Leading CSEOF Using 50 EOF Modes.** Note that the sign of the values should be flipped when comparing to previous figures.
Figure 1.7a: **Leading CSEOF Using 7 EOF Modes.**
Figure 1.7b: **Leading CSEOF Using 8 EOF Modes.**
Figure 1.8: The Eighth EOF. As if Figure 1.1, but for the eighth leading EOF.
Figure 1.9: The Tenth EOF. As in Figure 1.1, but for the tenth leading EOF.
Figure 1.10: **CSEOF of the AO.** CSEOF of the leading EOF mode only. In other words, a cyclo-stationary representation of the AO mode.
Figure 1.11: **AO PC Time Series.** The AO time series (the PC time series of the leading EOF; black) and the CSEOF PC time series when only the AO mode is used (red). Both time series have been normalized. The time series have a correlation coefficient of 0.39.
Figure 12: **CSEOF PC Spectra.** (a) Spectrum of the leading CSEOF when only the leading EOF is used. (b) Spectrum of the leading CSEOF when 20 EOF modes are used. The spectra are plotted in the variance-preserving frequency-times-spectrum versus log frequency portrayal. The 95% confidence bands based on the application of 15 consecutive 1-2-1 hannings are shown as dashed lines.
Figure 1.13: **Regressed Winds.** Regressed winds of the leading CSEOF (using 20 EOF modes).
Figure 1.13, Continued.
Figure 1.13, Continued.
Figure 1.14: Regressed Temperatures. As in Figure 1.13, but for regressed temperatures.
Figure 1.15: **Regressed Precipitation.** As in Figure 1.13, but for regressed precipitation.
CHAPTER 2

DETECTION OF THE MJO FROM QUIKSCAT

Background

The Madden Julian Oscillation (MJO) is the primary mode of intraseasonal variability in the tropics, although the signal is often more conspicuous in the Indian and Western Pacific Oceans than it is over the Eastern Pacific and the Atlantic. The oscillation exhibits a relatively broad frequency range, repeating anywhere from 20 to 100 days. However, much of the signal’s spectral energy is concentrated in the 40-60 day band. First observed by the mode’s namesakes in 1971, the MJO is somewhat unique among large-scale climatic signals in that the MJO’s frequency range is its defining characteristic (Madden and Julian 1971). In fact, Madden and Julian originally dubbed this signal the 40-50 Day Oscillation. A secondary attribute of the MJO is its eastward, equatorially trapped propagation.

The MJO signal is fairly conspicuous in outgoing longwave radiation (OLR) data, which is a proxy for convective activity in the tropics (Lau and Chan 1985). MJO convective anomalies in the Indian and western Pacific Oceans account for a large portion of precipitation variability on intraseasonal timescales. However, the MJO signal is found in other variables: wind (Madden and Julian 1994), pressure (e.g., Kayano and Kousky 1998), and even sea surface temperature (Krishnamurti et al. 1988). Given the clear link between the MJO and tropical convection, the divergent component of the wind field plays a more visible role than the rotational component, although a growing body of evidence suggests that rotational effects cannot be discounted (e.g., Raymond 2001).

In the present study, the ability of the MJO to be identified in QuikSCAT data is investigated. QuikSCAT is a relatively new tool that can be utilized to address outstanding MJO questions near the surface, such as the physics involved in the propagation and re-initiation mechanisms. First, it is necessary to assess the level to which QuikSCAT is able to capture the MJO signature. It will be shown that the MJO is indeed conspicuous in a pre-filtered gridded QuikSCAT product, and that clear signals can be seen in the following variables after bandpass
filtering over the MJO timescales: zonal wind, velocity potential, and zonal pseudostress at the surface. Moreover, we show that the average propagation speed over the Indian and western Pacific Oceans is on the order of 4 m/s. The novelty of the present work is the ability to clearly detect the MJO in a gridded QuikSCAT dataset. The QuikSCAT dataset can be utilized in the future to improve our understanding of MJO physics at the surface, a research avenue that has so far relied on reanalysis data.

**QuikSCAT Data**

The gridded scatterometer fields are produced using the variational approach outlined in Pegion et al. (2000), resulting in daily values of pseudostress ($\tau$) on a 1° X 1° grid. Observation accuracy of the scatterometer is described in Bourassa et al. (2003). From the zonal ($\tau_x$) and meridional ($\tau_y$) components of pseudostress, the zonal (u) and meridional (v) wind components are computed. From u and v, the velocity potential ($\chi$) is derived via a spectral method (see Krishnamurti 1998). Only results of analyses of u, $\tau_x$, and $\chi$ are considered. The data span from 40°E to 80°W and from 10°N to 10°S on a 1° X 1° grid, covering the deep tropics of the Indian and Pacific Oceans. The data are temporally averaged into 392 consecutive pentads, from late July 1999 through early December 2004.

**Detection of the MJO from Pre-filtered QuikSCAT Data**

**Methodology**

Obtaining a representative time series of the MJO is not a trivial task. The MJO is assumed to be present in the time series of every grid point in our domain. A variance decomposition technique is required to obtain a representative time series of the MJO for the entire domain. The need to consolidate geophysical data into the most energetic modes prompted the application of principal component analysis, or empirical orthogonal function (EOF) analysis, to geophysical space-time data sets. Kessler (2001) and Wheeler and Hendon (2004) are among the many works that have successfully utilized EOF analysis to define the MJO. A common problem with EOF analysis is that the spatial modes are fixed, an unattractive
attribute when dealing with a propagating feature like the MJO. In addition, proper detection of
the MJO with EOF analysis typically requires prior removal of low-frequency variability.

To overcome the shortcomings of EOF analysis on non-filtered data, we utilize Extended
EOF analysis, which uses a covariance matrix extended by cross-covariances at different lags
(e.g., see Weare and Nasstrom 1982). This allows the spatial pattern to move in time because
propagating patterns (assuming the propagation speed is fairly uniform) will be spatially
correlated at a given distance over a given time interval. Extended EOF analysis has been
widely used to represent MJO variability (e.g., Lau and Chan 1985, Kayano and Kousky 1998,
Myers and Waliser 2003). Consecutive plots of the lagged Eigenfunctions clearly demonstrate
propagating modes (a practical feature that EOF analysis lacks by its very nature). Moreover,
since Extended EOF analysis is able to differentiate between a propagating intraseasonal pattern
(the MJO) and similar spatial signatures that are associated with low-frequency variability,
Extended EOF analysis is preferable over EOF analysis for non-filtered data in the MJO domain.

The justification for extracting the MJO timescale from space-time data is often
misconstrued by some in the scientific community. A frequent concern is that this act constitutes
statistical voodoo, observing a 40-60 day oscillation only after bandpass filtering over the 40-60
day timescale. Based on the superposition principle, all space-time data can be considered
infinite combinations of spatial patterns and associated time series. Therefore, bandpass filtering
over any frequency band would yield modes that vary primarily across the timescale that was
isolated. This begs the following question: why is the MJO significant if it appears as though it
must be extracted via filtering before it can be observed?

By performing Extended EOF analysis on the pre-filtered zonal velocity data, it will be
clear that the MJO, while not as energetic as the annual cycle of solar radiation or the Indian
monsoon mode, is indeed embedded in the data and does account for a non-trivial portion of the
variance. A lag covariance parameter of 15 pentads is used, which corresponds to a maximum
lag of 75 days. Varying this parameter between 5 and 21 pentads did not materially affect the
results. Annual and semi-annual variations associated with the seasonal march of solar radiation
are not removed. Given its propagating nature, it is anticipated that MJO variability will be
represented by 2 Extended EOF modes in temporal quadrature with each other. Cross-
correlation and hodograph analysis will be utilized to test for quadrature. Spectral and wavelet
analyses will be conducted on the leading PC time series of MJO variability to ascertain the
energetic timescales captured by QuikSCAT, as well as how the timescale fluctuates over the record length. In addition, the above analysis will be repeated with the NCEP Reanalysis 2 zonal wind data (see Kanamitsu et al. 2002), provided by the NOAA-CIRES Climate Diagnostics Center directly from their web site at http://www.cdc.noaa.gov/. This data set covers the same time period (late July 1999 through early December 2004) and was linearly interpolated and re-gridded to the 1° by 1° grid of the QuikSCAT data in order to compare extraction of the MJO using these two independent data sources.

Results

Our results indicate that the primary modes of zonal velocity over our domain are associated with the annual cycle, the transitions between summer and winter Indian monsoons, and semi-annual variations. These three phenomena are associated with the three leading modes of pre-filtered zonal wind, accounting for 41.5, 10.1, and 3.7% of the total variance, respectively. However, modes 4 and 5, in quadrature with each other, are clearly associated with the MJO. This is evident by examining the lagged Eigenfunctions of the fourth mode (Fig. 2.1a) as well as the associated PC time series (Fig. 2.1b). Based on the Extended EOF maps, the pattern propagates eastward at approximately 4 m/s. This compares favorably with previous reports (Hendon and Salby 1994; Jones and Weare 1996; Shinoda et al. 1998; Rui and Wang 1990). Together, modes 4 and 5 account for about 4.5% of the total variance, a respectable portion considering the analysis was performed on the actual zonal velocity. EOF analysis of the non-filtered zonal velocity data was also performed (not shown). As expected, EOF analysis was not capable of separating the MJO’s intraseasonal, eastward-propagating variability from lower frequency modes.

The spectral characteristics of the PC time series in Figure 2.1b are shown in Figure 2.2a. The spectral peak occurs near the 50-day timescale, in general agreement with previous observations of the MJO timescale (see Madden and Julian 1994). Wavelet analysis provides an idea of the variation in timescale (Fig. 2.2b). Clearly, the timescale fluctuates around 50 days, with considerable energy occurring between 35 and 85 day periods in certain instances.
Comparison with NCEP Reanalysis

Extended EOF analysis of the non-filtered zonal velocity from NCEP Reanalysis 2 reveals a clear MJO pair as modes 6 and 7, garnering a combined 3.2% of the variability. The PC time series of mode 6 is presented in Figure 2.1c. Its spectrum and a wavelet analysis of this time series are presented in Figure 2.3. The PC time series has a maximum lagged correlation (at 1 pentad) of 0.96 with the fourth leading mode PC from QuikSCAT shown in Figure 2.1b. In addition, the raw and wavelet spectra in Figure 2.3 for the reanalysis data compares remarkably well with the corresponding QuikSCAT plots in Figure 2.2, showing both clear spectral peaks at a timescale of 50 days and evincing common fluctuations in spectral energy over time.

The difference in percent variance accounted for by the MJO modes in QuikSCAT (4.5%) and NCEP Reanalysis 2 (3.2%) is noteworthy. This represents an increase of relative importance of about 40% from one data set to another. If we account for the fact that the QuikSCAT data overall contains a 9% larger total variance over the MJO domain than the reanalysis data prior to interpolation (20% greater after interpolation), the disparity in intraseasonal energy between the two datasets is even greater. Spectral energy over the MJO domain in the 30-60 day range is about 30% lower on average in the Reanalysis zonal wind data compared to QuikSCAT data (Fig. 2.4). Therefore, it appears that the major difference between observed MJO’s in QuikSCAT and NCEP Reanalysis 2 lies in the signals’ amplitudes, whereas strong agreement exists regarding the observed timescales as well as temporal fluctuations in timescale.

Bandpass Filtered Analyses

Methodology

Based on the spectrum in Fig. 2.2a, Lanczos weights were determined for bandpass filtering $u$, $\chi$, and $\tau^v$. The filter consists of 31 weights, symmetric about the central point, with half power at about 33 and 75 days. Time-longitude plots are computed from the bandpass filtered values of $u$, $\chi$, and $\tau^v$. From these charts, propagation speeds are estimated. Endpoints are retained using the equal-variance least squares filtering method (see Chapter 4).
Results

Time-longitude plots of bandpass filtered $u$, $\chi$, and $\tau^*$ are shown in Figure 2.5. Clear eastward propagation is evident for each variable. Approximate estimations from Figure 2.5 suggest a propagation speed across the longitudinal extent of our domain of about 5 m/s. However, it is clear that the propagation speed is not at a constant rate, but ranges from about 1 to 10 m/s. This compares favorably with previous studies (e.g., Lau and Chan 1985). Closer inspection reveals that the propagation speed is considerably higher in the eastern Pacific (~10 m/s), in accordance with previous research (Hendon and Salby 1994, Maloney and Hartmann 1998).

The time-longitude plots portray fairly regular MJO events, although the MJO signal was either absent or diminished during several brief episodes. These include disruptions in early 2000 and late 2003 when a westward propagated occurred, as well as a brief absence in late summer of 2002. In general, it appears that the MJO is not prone to favor a particular season, although recent studies maintain that the signal is most conspicuous during springtime because of reduced interaction with the Indian monsoon (e.g., Myers and Waliser 2003). Similar time-longitude analyses were conducted with other kinematic variables, such as the meridional velocity and rotational quantities such as the vorticity and streamfunction. As expected from previous works, clear MJO signals were not apparent in these parameters.

Summary

Gridded data from the SeaWinds instrument on the QuikSCAT satellite were inspected to ascertain how well the instrument could resolve MJO variability at the surface. Extended EOF analysis of non-filtered surface zonal velocity over the tropical Indian and Pacific Oceans, with a lag parameter of 15 pentads, reveals that MJO variability is captured by modes 4 and 5. These two modes are in quadrature with each other and explain a combined 4.5% of the total variance. The first 3 modes explain the majority of the variance, and are associated with annual, semi-annual, and monsoon-related variations. Although the MJO is not the dominant mode in the non-filtered data set, it does represent the dominant mode of intraseasonal variability,
overshadowed by modes associated primarily with the annual cycle of solar radiation. The QuikSCAT results compare favorably with results from NCEP Reanalysis 2, with the main difference being the considerably larger amplitude of the MJO’s surface signal in QuikSCAT data.

After confirming the existence of MJO variability in the data, bandpass filtering was conducted to isolate MJO timescales. Clear MJO patterns (with appropriate spatial scales and eastward propagations) were found in zonal velocity, velocity potential, and zonal pseudostress. Time-longitude plots and Extended EOF analysis reveals typical propagation speeds ranging from 1-10 m/s, with a mean value in the vicinity of 4 m/s.

Our findings highlight the usefulness of a gridded QuikSCAT dataset for investigation of the MJO. As the steady stream of satellite data, such as those available from QuikSCAT, further improve the spatial and temporal coverage of available data sets, more detailed descriptions of weather phenomena will be possible. In particular, the potential of improved understanding of the MJO due to increased coverage, rather than relying on reanalysis data, is promising, especially for winds.
2.1: The Fourth Leading EXEOF. (a) Lagged Eigenfunctions of the fourth leading EXEOF of non-filtered zonal velocity from QuikSCAT. The lag parameter for the EXEOF analysis was 15 pentads. Negative contours are dashed and the zero line is bold. Land areas, including the Maritime Continent and Africa in the west, are shown in black. The lag number is shown in the upper right hand corner of each plot. Note that plots 9 to 10 lags apart are similar, indicating an MJO timescale of 45 to 50 days. (b) The associated PC time series. (c) The PC time series of the sixth leading mode of EXEOF analysis using NCEP Reanalysis Version 2. The time series in (b) and (c) have a maximum lag correlation coefficient of 0.96.
2.2: The Fourth Leading EXEOF PC. (a) Smoothed spectral estimates of the time series in 2.1b. The 95% $\chi^2$ confidence bands are shown as dashed lines. (b) Wavelet analysis of the time series in 2.1b. The values have been scaled to range between 0 and 7.
2.3: NCEP Reanalysis MJO PC. As in Figure 2.2, but for the PC time series of the sixth leading mode of zonal wind using NCEP Reanalysis Version 2.
2.4: **Intraseasonal Energy.** Local percent deviation of spectral energy in the 30-60 day band for zonal winds between QuikSCAT and NCEP Reanalysis 2. Areas with light shading indicate regions where the spectral energy is at least 25% greater in reanalysis than QuikSCAT. Darker shading indicates that QuikSCAT has at least 25% more spectral energy than reanalysis. It is clear that QuikSCAT zonal wind data displays considerably larger spectral energy within MJO timescales, especially near the Maritime Continent and in the western Pacific Ocean.
2.5: **Longitude-Time Plots.** (a) Hovmoller diagram of bandpass filtered $u$. (b) Same as 2.5a, but for velocity potential. (c) Same as 2.5a, but for zonal pseudostress. For all three charts, the magnitudes are arbitrary and dimensionless.
CHAPTER 3

AN IMPACT-BASED PERSPECTIVE OF ENSO

Background

Several decades ago, the reemergence of interest in the Southern Oscillation sparked a continuous flurry of activity in the climate community to unravel the nature of the El Niño Southern Oscillation (ENSO) that continues to this day. During this time, many aspects of ENSO have been addressed, including its timescale, impacts, and physical mechanisms. For many of these applications, ENSO indices have been used. As with any climate index, ENSO indices have been constructed as tools that quantitatively describe the state and intensity of (what we perceive to be) the ENSO phenomenon (see Hanley et al. 2003; Goddard et al. 2002).

The earliest indicator of ENSO is the Southern Oscillation Index, a standardized pressure difference (usually) between Tahiti and Darwin, Australia (see Walker and Bliss 1932; van Loon and Madden 1981; Horel and Wallace 1981). However, as climate scientists recognized the pivotal role played by the ocean in ENSO variability, the variable of choice for defining ENSO time series became (and continues to be) SST (SCOR 1983). The five major Niño regions used for box-average SST indices cover the longitudinal extent of the equatorial Pacific Ocean (see Barnston et al. 1997). Niño 1+2 (they are combined into one region) sits just off the coast of South America, Niño 3 just to its west out to 150°W, and Niño 4 west to 160°E. The Niño 3.4 region, as its name suggests, is comprised of the western part of Niño 3 and the eastern part of Niño 4 (see Barnston et al. 1997).

It is well known that the ENSO response across the equatorial Pacific is not monolithic. The ENSO lifecycle includes warm and cold anomalies that evolve temporally and spatially, although the direction seems to be one of many observed changes attributed to the climate shift of 1976 (Rasmusson and Carpenter 1982). Rasmusson and Carpenter (1982) write that ENSO “episodes cannot be adequately described in terms of standing oscillations.” However, by relying on a box-averaged SST time series over a fixed region, we ignore their admonition and utilize the Niño region indices as authoritative representations of the ENSO physical process.
The purpose of the present study is to determine targeted indicator regions for a particular impact zone. In doing so, focus is shifted from analyzing a unified ENSO formulation in favor of identifying practical indicator-impact pairs. In essence, the work complements the ENSO indicator region analysis of Hanley et al. (2003). Although the Niño regions are not tested directly, the above procedures indirectly assess their viability for capturing ENSO-related remote impacts. Alternatively, the results can be interpreted as highlighting which Niño regions are best capable of capturing a remote response for a given impact zone. Analyses are performed at each impact grid point, for the United States as a whole, and for select regions within the U.S. It will be shown that there are indeed conspicuous differences in teleconnection patterns between the tropical Pacific and the United States. For example, the Niño 1+2 temperature impacts over the U.S. during “La Niña” (anomalies that differ greatly from the temperature anomalies associated with cold phases of the other Niño regions) are reconciled with the prevailing ENSO paradigm via EOF analysis of composited indicator-impact pair anomalies. In addition, a previously undocumented wintertime teleconnection between equatorial Pacific SST and the Southeastern U.S. is identified.

**Data Sets**

Monthly Pacific SST data are extracted from the ERSST data set (Smith and Reynolds 2003), which is constructed on a 2° X 2° grid, from 151°E to 81°W and 15°S to 15°N. The record length is 55 years, from 1946 through 2000. Trends and monthly climatologies are removed at each grid point. Monthly U.S. impact data consists of mean temperature and precipitation data from the USHCN (see Karl 1990), gridded onto a 2° X 2° grid by simple averaging over available data in each box. The resulting U.S. grid has 235 total grid points with data values; however, four interior points are completely devoid of data. For temperature, the grid points are virtually complete. The precipitation values are about 90% complete; missing values are not replaced. This rather coarse gridding is essential for efficient computations for an empirical study such as this, and is adequate because large-scale patterns are of particular interest. Local effects, especially for precipitation, are important as well, but are not a focus of the current investigation. However, the target indicator regions can also be computed per station or grid point if a more localized assessment is desired.
Several important points regarding tropical Pacific SST are worth mentioning. First, the total (Fig. 3.1a) and interannual variability (Fig. 3.1b) of SST across the Pacific are not uniform. The most spectral energy is concentrated along the equator in the eastern two-thirds of the Pacific, especially in the cold tongue region. This is an important point to keep in mind when determining thresholds for warm and cold SST phases. A substantial portion of SST variability is concentrated between 3 and 6 years (Fig. 3.1c). It is also paramount to assess the correlation structure of the tropical Pacific (Fig. 3.2). SST is to a large degree spatially auto-correlated across the tropical Pacific, except in the far western portion of the domain, which we term the negative correlation region (NCR).

Methodology

SST indicator regions are varied across the equatorial Pacific Ocean. Based on preliminary results, a 6° X 6° degree indicator region was settled on. This provides sufficient removal of small-scale noise while maintaining the desire to pinpoint the best indicator region for a particular impact zone. On rare occasion, larger boxes seemed to account for larger ENSO-related impacts over the US, but we assume that the constraint of a 6° X 6° box would select a region within the larger box (or sufficiently close to it) to justify this computational efficiency. The 6° X 6° setting means that there are 806 (62 by 13) indicator regions that overlap each other.

An additional timesaver is that impacts are only considered for DJF. Air temperature deviations are computed using composite analysis based on a warm and cold phase, which is defined as when the local ENSO index is above or below symmetric threshold values. For precipitation anomalies, percent deviations are computed to account for the disparities between rainfall amounts in different regions of the US, especially the sharp contrast between the dry west and the moist east. As a result, the minimum precipitation deviation is –100% while there is no upper bound to a precipitation increase. By focusing on DJF, we can disregard the extraordinary precipitation anomalies associated with tropical cyclone activity.

We utilize local thresholds of ±1 standard deviation to define warm and cold phases for each box. Since SST anomalies are normally distributed, this means that the warm and cold phases will each represent approximately 17% of the distribution. In climate parlance, this would be considered extreme cases, excluding those “events” that are slightly warmer or colder.
than normal (i.e. neutral) and concentrating on the rare, large events. A uniform threshold, such as 0.5°C for the entire basin, is not adequate because of the spatial variations of SST variance across the equatorial Pacific (Fig. 3.1a)

For each SST indicator region, the monthly mean temperature anomaly composite is computed for both warm and cold events. These data arrays are denoted by $T_w(r,s)$ and $T_c(r,s)$, where w stands for the warm phase, c for the cold phase, r for the indicator location, and s for the impact grid point. A similar calculation is performed on the monthly precipitation, except that these are reported as composite percent deviations from the monthly climatology: $P_w(r,s)$ and $P_c(r,s)$. It is important to note that the data arrays are functions of indicator and impact locations only; the temporal dimension has been eliminated by creating composites conditioned on SST phase.

The fields are inspected from two perspectives: impact grid point and indicator region. The impact perspective is defined as the vantage point in which the anomalies for a given impact region are quantified as a function of indicator region. For each impact grid point, a contour map is created that denotes the composite climate anomaly (temperature or precipitation) for all 806 indicator regions for both warm and cold phases. Although each contour map is plotted over the Pacific Ocean, it must be emphasized that the values are composited air temperature anomalies or precipitation deviations for a given impact grid point. This highlights the variations in impact from varying indicator locations. We focus on four grid points for temperature that are centered on North Florida, Eastern Pennsylvania, Southwest Washington state, and Northeast Montana. For precipitation, we use grid points in Southern California, Southwest Florida, Northwest Kansas, and Southeast Montana. These points are subjectively determined based on well-established ENSO-related impacts.

In the alternate perspective, temperature and precipitation impacts are plotted as a function of indicator region location, allowing an inspection of the climate impacts associated with a single indicator region (much like the ubiquitous El Niño and La Niña impact maps). Therefore, the indicator perspective is defined as the vantage point from which climate anomalies across a large impact zone are computed as a function of a given indicator region. Five indicator regions are highlighted for both precipitation and temperature. All five are centered on the equator. The first four are centered on the midpoints of the longitudinal ranges
of the Niños 1+2, 3, 3.4, and 4 boxes. The fifth region is located at the western extreme of our SST domain, i.e. the NCR.

In order to simultaneously interpret both the impact and indicator perspectives described above, Empirical Orthogonal Function (EOF) analysis of the space-by-space data arrays is employed. Typically in geophysical applications, EOF analysis is used as a variance decomposition tool applied to a space by time field, reducing a cumbersome data array into a much smaller set of spatial patterns and associated time series. The time series is considered to be the loading vector, i.e. the time-varying weight (or amplitude) associated with a particular spatial signature. In the present case, the EOF modes represent temperature or precipitation signatures that are modulated by the location of the SST indicator region. In other words, the leading modes represent the most energetic impact-indicator pairs.

This type of EOF analysis is particularly sensitive to the impact zone used. For example, the temperature impacts in Florida may not be associated with a similar SST “loading pattern” as the temperature impacts of the Upper Midwest. Therefore, the EOF analyses are repeated for five regions with known ENSO impacts: Florida, the Southeast, the Upper Midwest, California/Nevada, and the Pacific Northwest.

Results of Grid Point Analyses

Impact Perspective of Temperature

Warm and cold temperature anomalies as a function of indicator region are presented in Fig. 3.3a for North Florida. For warm SST phases, conditions in North Florida are up to 1°C colder, with the action center sitting squarely in the Niño 3.4 region. Positive anomalies occur when the index is constructed in the NCR. For cold phases, the main effect over North Florida are warmer conditions (>1°C) for indicator regions over most of the SST domain, except for the NCR and the East Pacific Upwelling Zone (EPUZ) where associated cooling can be colder than -1°C.

Warm phase conditions in Eastern Pennsylvania (Fig. 3.3b) are typically warmer than normal (up to 2°C warmer) regardless of the indicator location. The peak anomalies occur when the index is calculated near the EPUZ. During the Eastern Pennsylvania cold phases, a pattern
similar to North Florida cold phases emerges, with colder conditions observed with an EPUZ index and warmer conditions if an indicator region further west is used. It is interesting to note that, for Eastern Pennsylvania, open ocean SST anomalies are associated with warmer air temperatures for both warm and cold SST phases.

For Southwest Washington (Fig. 3.3c), warm phases largely bring warmer conditions, with anomalies exceeding 1°C if an EPUZ index is used. A relative weak spot is observed in the Niño 3-3.4 overlap region, suggesting that Niños 3 and 3.4 may underestimate the warm phase temperature deviations for Southwestern Washington. For cold phases, the prominent feature is associated with colder conditions in Southwestern Washington, up to 1°C colder. The cold phase action center sits squarely in the Niño 3.4 region, with opposite-signed anomalies in the NCR to the west.

Warm phase conditions for Northeast Montana (Fig. 3.3d) typically involve much warmer conditions than normal, with anomalies exceeding 2.5°C for most of the Eastern Pacific indicator regions. For cold phases, colder conditions are felt in Northeast Montana, although deviations do not go below -2°C. The anomalies are coldest when the indices are chosen in the Central Pacific or in the EPUZ. For both warm and cold phases, opposite-signed anomalies occur in the NCR.

**Impact Perspective of Precipitation**

Warm and cold precipitation deviations as a function of indicator region are presented in Fig. 3.4a for Southern California. During the warm phases, increased precipitation occurs in Southern California when indices are chosen anywhere in the Central or Eastern Pacific, punctuated by deviations greater than a 75% increase when the indicator region is picked in the southwestern corner of the SST domain. During cold phases, precipitation is reduced 15%-40% for indicator regions in the CPAC and EPAC. For both warm and cold phases, the peak deviations occur near the Niño 1+2 region as well as the Niños 3.4-4 overlap region. Opposite-signed deviations also occur in the NCR. For Southwestern Florida (Fig. 3.4b), the precipitation deviations are comparable to Southern California’s patterns. The only major difference is that in the cold phase, the driest conditions occur when the indicator region is located in the Niños 3-3.4 overlap region.
For Northwest Kansas (Fig. 3.4c), wetter conditions (up to 40% more precipitation) occur during the warm phase, with action centers in the EPUZ and just north of the Niño 4 region. For the cold phases, dry conditions occur with precipitation being reduced by as much as 60%. The largest cold phase deviations occur when the index is selected in the EPUZ, but significant deviations occur for Niño 3 as well. Once again, opposite-signed anomalies occur in the west.

Warm phase precipitation deviations for Southeast Montana (Fig. 3.4d) are negative over most of the equatorial Pacific. Peak deviations reach a 40% decline in scattered regions, which are typically displaced from the equator somewhat. For the cold phases, wetter conditions (up to 65% more precipitation) occur with indices in the far eastern part of the SST domain, with the largest deviations occurring near the Niño 1+2 region.

**Indicator Perspective of Temperature**

The indicator region perspective for air temperature and precipitation are presented in Figures 3.5 and 3.6, respectively. The top four rows in each figure are well-documented plots used to determine the varying impacts of the Niño regions, while the bottom row represents the impacts of SST in the NCR. For warm phase temperature analyses, it is evident that the Niño regions (Fig. 3.5a-d) impart similar impacts over the United States. Much warmer conditions are experienced in the upper Great Plains and Midwest, whereas somewhat cooler conditions are felt in Florida. The main difference between the Niño region impacts is that the warm anomalies are more intense for Niño 1+2. The cold phases of the Niño regions are not as similar. The most striking impact occurs during the cold phase of Niño 1+2, when the Eastern two-thirds of the United States is greater than 1°C below normal. In the Ohio River Valley, the departure below normal is greater than 3°C. For Niños 3, 3.4, and 4, cold phase deviations bring warmer conditions to the Southeast. For the NCR, warm SST anomalies are associated with warmer conditions in the eastern half of the U.S. (Fig. 3.5e), with anomalies of up to 2°C in the Ohio River Valley. Somewhat cooler conditions are felt in the Pacific coastal states as well as Montana and North Dakota. The pattern bears resemblance to both the cold phases of Niños 3 and 3.4 as well as the inverse of the Niño 1+2 cold phase. The cold phase of the NCR looks a great deal like the warm phase for the Niño regions, except that the Florida anomalies are weaker.
Indicator Perspective of Precipitation

Warm phase precipitation deviations are also comparable for the Niño regions (Fig. 3.6a-d). Once more, the larger intensities when using Niño 1+2 as the indicator region is the main difference. The “Sun Belt” region, stretching along the southern US from California to Florida, receives up to twice as much precipitation during the warm phase. Pockets of drier conditions occur over the Ohio River Valley, Montana, and Minnesota. Cold phase deviations for Niño 1+2 bring drier conditions to the West Coast, the Atlantic Seaboard, the Upper Midwest, and the Great Plains. Wet anomalies are concentrated in Arizona, New Mexico, and Texas. In contrast, cold phase anomalies for Niños 3, 3.4, and 4 are similar to each other, with drier conditions experienced along the Sun Belt and modestly wetter conditions in the Midwest and Pacific Northwest. In the NCR (Fig. 3.6e), warm phases are associated with drier conditions in Florida and Arizona and moister weather in the Midwest and the Pacific Northwest. In the cold phase, positive precipitation deviations stretch along the Sun Belt, while drier conditions occur in the Northern Rockies and the Midwest. As with temperature, the cold phase negative correlation region anomalies resemble the warm phase impacts of the Niño regions.

EOF-based Indicator-Impact Decomposition for the United States

Temperature

In order to incorporate both indicator and impact region perspectives, EOF analysis was conducted on $T_w$, $T_c$, $P_w$, and $P_c$. The 2 leading modes of $T_w$ and $T_c$ are presented in Figure 3.7. The leading mode of $T_w$ accounts for 86% of the total variance. The impact pattern bears a strong resemblance to the warm phase anomalies for all four Niño regions. The indicator weighting function shows that the amplitude of the impact-indicator pair is maximized in the Eastern Pacific, in agreement with the finding in Figure 3.5 that showed much more intense temperature impacts for Niño 1+2 than for the others. It is important to note that the sign of indicator weighting function does not determine the sign of the SST anomalies in the region. Since $T_w$ is composited over warm phases, its indicator weighting function always represents
warm SST anomalies. A zone that has a negative indicator weighting function (such as in the northwest of the SST domain for the leading mode of $T_w$) indicates that the associated temperature impacts are reversed in sign.

The second leading mode of $T_w$ accounts for $\sim$6% of the variance only. The temperature impacts denote an opposing anomaly field between the Eastern half of the U.S. and the Pacific Northwest. The indicator weighting field contains smaller spatial scales than the leading mode’s field. Although the percentage of variance accounted for by this mode is dwarfed by that of the leading mode, it is noteworthy that the impact amplitudes are most intense in the Southeast. For the leading mode, the impacts in the Southeast are relatively weak. Thus, although the leading mode of $T_w$ garners such a high portion of the total variance and captures much of what we know of El Niño type impacts over the U.S., mode 2 will be shown to play a bigger role for the Southeast.

For cold phase temperatures ($T_c$), the variance decomposition is not as one-sided. The two leading modes account for about 53% and 32%, respectively. Varimax rotation of the EOF’s did not alter the variance percentages or the impact and indicator patterns in a significant way. The leading mode compares favorably to the cold phase impacts of Niños 3 and 3.4, and to a lesser extent Niño 4. The dipole-like impact structure also compares favorably with that of the second mode of $T_w$. The indicator weighting has a large swath of same-signed values for most of the equatorial Pacific, centered on the Niños 3-3.4 overlap region. Opposite-signed values are associated with the northwest of the SST domain and the EPUZ. Mode 2 of $T_c$ involves an impact in the northern two-thirds of the U.S., peaking over North Dakota. The indicator weighting function is dominated by a zonal contrast between the NCR and the EPUZ. The impact footprint is comparable to the typical El Niño type response.

**Precipitation**

The leading modes of $P_w$ account for about 82% and 6% of the total variance, respectively (Fig. 3.8). For Mode 1, a strong wet signal is observed is the Sun Belt, as well as the Great Plains and the Southeast coast. Drier conditions occur in Montana. This pattern is associated with a broad signature in the indicator weighting plot, with peak intensity in the eastern Pacific. The wet and dry impacts are reversed in this mode for indicators in the NCR.
The second mode of $P_w$ is much less energetic, but plays a role locally in Texas and California. For example, for indicator regions near and to the south of Niño 1+2, the wet anomalies associated with $P_w$ modes 1 and 2 are additive for California, whereas in Texas mode 2 moderates the wet conditions associated with mode 1.

The leading mode of $P_c$ accounts for $\sim 72\%$ of the variance. For a large swath of the equatorial Pacific, it is associated with significantly drier conditions in the Sun Belt, and modestly wetter conditions in the Midwest and Pacific Northwest. Once again, the opposite climate impacts occur in the NCR. The indicator weighting is weak in the EPUZ. For mode 2, the indicator weighting function is that significant.

**EOF-based Indicator-Impact Decomposition for Selected Regions in the U.S.**

**Florida**

The leading mode of $T_{w,FL}$ (Fig. 3.9) generally brings colder conditions to Florida, accounting for $\sim 93\%$ of the variance. The peak positive indicator region is centered near the Niño 3.4 region. However, the most intense “warm phase” anomalies occur in the. Based on a local indicator function near the maximum, warm phase Florida winters would be about 2°C warmer on average. The second mode ($\sim 5\%$ of the variance) modulates the signal in the Niño 3 and Niño 4 regions by adjusting the variation between impact in Northern and Southern Florida.

For cold phase temperatures in Florida, the primary pattern ($\sim 97\%$ of the variance) is characterized by warmer conditions (about 1°C) over Florida when an index over the open ocean of the equatorial Pacific is used. However, cold conditions are experienced if the indicator region is chosen in the northwest or in the EPUZ. The second mode of $T_{c,FL}$ does not play a critical role.

The primary mode of $P_{c,FL}$ ($\sim 90\%$ of variance) is characterized by wetter conditions across Florida (Fig. 3.10). The indicator weighting is positive for most of the equatorial Pacific, but is largest in the Eastern Pacific. Mode 2 accounts for 7% of the variance and has action centers in the southeast corner of the SST domain as well as in the NCR. The southeastern indicator function’s action center interacts constructively with mode 1 for eastern and central Florida (effectively filling the whole in central Florida in the leading mode’s impact map) and
destructively in Northwest Florida. During cold phases, the Florida peninsula is largely drier, thanks to a positive indicator region for most of the Pacific, except in the NCR. The leading mode accounts for 81% of the variance. The second mode of $P_{c,FL}$ ($\sim$16%) is primarily a response associated with the NCR.

**Southeast**

The primary mode for $T_{w,SE}$ (Fig. 3.11) is associated with the second-leading mode of temperature variability for the entire United States. It accounts for 68% of the variance and involves a uniform signed response throughout the Southeast. Mode 2 accounts for $\sim$28% of the variance and appears to be the typical El Niño-type pattern. However, since the Southeast represents a transition zone between El Niño anomalies in the Northern U.S. and Florida, this effect relegates it to mode 2 status when we look squarely on the Southeast.

During the cold phases, the leading mode in temperature for the Southeast is the typical La Niña pattern, accounting for over 96% of the variance. Although the cold phase temperature field for the entire U.S. gets decomposed into two modes with about 50% and 30% each, in the Southeast, only the first is a primary participant. Again, this mode is primarily the same sign for the vast majority of the Pacific and brings warm conditions to the Southeast in these regions of the Pacific. However, the opposite impacts occur with this mode if the indicator region is selected in the northwest of the SST domain or in the EPUZ.

For precipitation, the primary warm phase pattern in the Southeast is the typical El Niño type impact, garnering about 85% of the total variance of $P_{w,SE}$ (Fig. 3.12). The second and remaining modes do not play a big role, especially when considering only indicator weightings along the equator. During the cold phases, the leading mode (75% of variance) represents the Southeast’s distinction as a transition zone between drier conditions in the Deep South and wetter conditions in the Midwest (recall Fig. 3.8). The second-leading mode ($\sim$10%) has a strong indicator weighting in the Niño 1+2 region. Here, it interacts constructively with the first mode of $P_{c,SE}$ to create a pattern that is dry throughout the Southeast (recall Fig. 3.6).
Upper Midwest

Warm phase temperature anomalies over the Upper Midwest (Fig. 3.13) are dominated by the El Niño type pattern (~99% of the variance) that brings very warm conditions to the region, with anomalies of up to 6°C. The index weighting is the same pattern seen over and over again with a peak in the EPAC and opposite signs in the NCR. The leading cold phase mode accounts for about 89% of the variance. The indicator weightings have action centers in the EPUZ and the NCR, with opposing signs in the two regions. For the EPUZ, colder conditions are experienced in the Upper Midwest, whereas warmer temperature anomalies when the indicator is chosen in the NCR. Precipitation effects in the Upper Midwest are not considered.

Pacific Southwest

For the Pacific Southwest, only precipitation is considered. The leading mode of $P_{w,PSW}$ resembles the El Niño footprint and accounts for 89% of the variance (Fig. 3.14). During this phase, the Pacific Southwest receives a significant increase in winter precipitation. Since this region has a Mediterranean climate, such an increase in precipitation in what is already the “rainy season” can have disastrous effects on agriculture in the region. Mode 2 (~8%) modulates mode 1 by accentuating differences between the Northern and Southern parts of California and Nevada.

During the cold phase, the Pacific Southwest is generally drier. The indicator weighting of $P_{c,PSW}$ (~73% of the variance) is mostly the same sign throughout the Pacific, peaking near the Niño 1+2 region, but with opposite signs in the NCR. The indicator weighting of mode 2 (~16%) indicates that it acts destructively with mode 1 in its southeastern action center, but constructively with its northeast action center.

Pacific Northwest

The leading mode of warm phase temperature anomalies in the Pacific Northwest represents the typical El Niño type pattern, accounting for 94% of the variance (Fig. 3.15). The warm anomalies reach up to 2°C. The indicator weighting is most intense in the EPUZ. Mode 2
only accounts for ~4% of the variance. The leading mode of $T_{c,PNW}$ accounts for 87% of the variance and is associated with cold conditions in the Pacific Northwest when in an indicator in the Central Pacific is used. However, opposite-signed anomalies occur in the NCR and near the Niño 1+2 region. Precipitation deviations are not considered for the Pacific Northwest.

Summary

In the current investigation, an attempt was made to recast the ENSO impact problem to incorporate the obvious notion that a single ENSO index may not be the most appropriate indicator for all impact zones. Composite air temperature and precipitation anomalies (Figs. 3.3 and 3.4, respectively) were computed as a function of indicator region for select stations with known ENSO responses. Not only are different indicator regions better suited for different impact grid points, but the same impact point may have different regions that capture the largest anomalies based on SST phase (warm or cold) or variable (temperature or precipitation). Furthermore, the sign of the impact anomaly may vary in different parts of the Pacific for a given impact point, SST phase, and variable. This phenomenon typically occurs when the NCR and/or the EPUZ differs from the interior ocean (e.g. the cold SST phase air temperature anomalies for North Florida and Eastern Pennsylvania).

While quantifying ENSO impacts based on the impact perspective is useful, rather than the customary focus on the indicator perspective only, a natural next step would be to develop a method to simultaneously describe both perspectives. The method offered to accomplish this is EOF analysis of compositied temperature or precipitation anomalies, conditioned on SST phase, as functions of indicator and impact positions. While EOF analysis is certainly not new to climate research, it is customarily performed on a space-time data matrix to arrive at series of spatial patterns and associated time series. In the current examination, the results of the EOF analysis are impact amplitude maps with associated indicator weighting maps, with the leading modes representing energetic indicator-impact modes. The impact amplitude map represents a recurring signature in temperature or precipitation over the United States, whereas the indicator weighting map depicts a modulation of the impact amplitude based on location of an indicator region. This method is applicable to other teleconnection patterns as well, and can also be utilized as a domain identification technique to determine the appropriate impact and indicator
regions for additional analysis such as conventional composite analysis or canonical correlation analysis.

Of the many findings presented in the current investigation, two stand out. The first is the EOF decomposition of cold SST phase air temperature anomalies over the U.S. Even after varimax rotation, the first and second leading modes account for about 53% and 32%, respectively. This breakdown stands in sharp contrast to the warm SST phase air temperature decomposition, and both phases of the precipitation decomposition, in which there were leading modes with a commanding share of the variability and clearly related to “El Niño” and “La Niña.” Especially noteworthy is the indicator weighting functions near the vicinity of the Niño 1+2 region in the two leading cold phase SST patterns. Combining the effects of these two figures for the Niño 1+2 indicator region, the Niño 1+2 U.S. air temperature impact (Fig. 3.5) is eloquently accounted for.

The other noteworthy finding is the warm SST phase breakdown of air temperature over the Southeastern U.S. (Fig. 3.11). It is well established, based on analyses such as that in Fig. 3.5, that there is no meaningful “El Niño” impact in the Southeastern U.S. (excluding Florida). However, the EOF decomposition of warm SST phase air temperature anomalies for the Southeast reveals a leading mode (68% of the variance) whose impact amplitude is uniform across the Southeast. This is part of the same pattern that is the second leading warm SST phase mode for the entire United States (Fig. 3.7). Based on the indicator weighting function, this rather clandestine mode appears to be totally independent from the conventional ENSO phenomenon, suggesting that there is indeed a potent, non-ENSO teleconnection between the Southeast U.S. and the tropical Pacific Ocean.

The findings presented herein are a step forward for our understanding of ENSO and teleconnections between the Pacific and the U.S. in general. However, a substantial amount of additional analysis needs to be performed. This includes expanding the study for all seasons and determining the role of lagged responses in the findings. The study can also be repeated using pre- and post-1976 to compare and contrast the teleconnective behavior before and after the 1976 climate shift. That being said, the method outlined here can be utilized by stakeholders and decision makers to respond to a localized ENSO impact assessment for their region of interest. In closing, it is imperative that we investigate ENSO-related variability with both impact and indicator perspectives in mind, and not be restricted to using only the established ENSO indices.
Figure 3.1: **Pacific SST.** (a) Local SST standard deviation for each SST grid point. The range of the standard deviations is from about 0.05 to 1.55. (b) Spectral energy in the 2-7 year period range for each SST grid point. (c) Spectral estimates of latitude-averaged equatorial Pacific SST time series. Minimal filtering was applied to the plots. All three plots were also scaled to range between 0 and 8, rendering the fields dimensionless.
Figure 3.2: **SST Correlations.** Correlations between SST grid point time series and (a) Niño 1+2, (b) Niño 3, (c) Niño 3.4, and (d) Niño 4. The box represents the corresponding Niño region.
Figure 3.3: **The Impact Perspective – Air Temperature.** Mean winter temperature anomalies (°C) for impact grid points representative of (a) North Florida, (b) Eastern Pennsylvania, (c) Southwest Washington, and (d) Northeast Montana. The anomalies are contoured as a function of warm and cold phases of each 6° X 6° SST indicator box across the tropical Pacific. Anomalies are contoured at the box center. Although the values are plotted across the Pacific and the reader may be tempted to assume the contoured values are SST anomalies, it must be reiterated that the plotted values are air temperature anomalies over the indicated regions conditioned on SST indicator zone.
Figure 3.4: **The Impact Perspective – Precipitation.** Percent deviations in winter precipitation (from climatology) for impact grid points representative of (a) Southern California, (b) Southwest Florida, (c) Northwest Kansas, and (d) Southeast Montana.
Figure 3.5: The Indicator Perspective – Air Temperature. Warm and cold phase temperature deviations for $6^\circ$ by $6^\circ$ indicator regions centered on the equator. The longitudinal ranges straddle (a) $275^\circ$E, (b) $240^\circ$E, (c) $215^\circ$E, (d) $185^\circ$E, and (e) $155^\circ$E. The five SST regions are representative of Niños 1+2, 3, 3.4, 4, and the extreme Western Pacific, respectively.
Figure 3.6: **The Indicator Perspective – Precipitation.** As in Figure 3.5, but for percent deviations in monthly precipitation.
Figure 3.7: EOF Decomposition over US – Air Temperature. EOF analyses of warm and cold phase temperature anomalies as a function of indicator and impact regions for the entire United States. The left column represents the impact spatial signatures and the right column denotes the associated loadings for each 6° by 6° indicator region. The product of each mode pair is in degrees Celsius.
Figure 3.8: **EOF Decomposition over US – Precipitation.** As in Figure 3.7, but for percent deviations in monthly precipitation. The product of each mode pair is expressed as a percentage.
Figure 3.9: **EOF Decomposition over Florida – Air Temperature.** As in Figure 3.7, but for Florida.
Figure 3.10: **EOF Decomposition over Florida – Precipitation.** As in Figure 3.8, but for Florida.
Figure 3.11: **EOF Decomposition over the Southeast – Air Temperature.** As in Figure 3.7, but for the Southeastern United States.
Figure 3.12: **EOF Decomposition over the Southeast – Precipitation.** As in Figure 3.8, but for the Southeastern United States.
Figure 3.13: **EOF Decomposition over the Upper Midwest – Air Temperature.** As in Figure 3.7, but for the Upper Midwest.
Figure 3.14: **EOF Decomposition over the Pacific Southwest – Precipitation.** As in Figure 3.8, but for California and Nevada.
Figure 3.15: **EOF Decomposition over the Pacific Northwest – Air Temperature.** As in Figure 3.7, but for the Pacific Northwest.
CHAPTER 4

TIME SERIES FILTERING NEAR ENDPOINTS

Background

Filtering is an essential tool for analyzing geophysical time series. Common applications include extraction of salient timescales, suppression of high-frequency noise, and smoothing spectra in order to increase statistical confidence in spectral estimates. Time series filtering can be done in the frequency or time domains. Consider an input time series, \( x(t) \), that is to be filtered to produce \( y(t) \), the output time series. Filtering in frequency space is accomplished by (1) applying a Fourier Transform (FFT) to the time series \( X \), (2) multiplying by the frequency response function (FRF, \( H \)), and (3) back-transforming into time space:

\[
X(f) = \sum_{t=0}^{N-1} x(t)e^{i2\pi ft/N}, \quad \text{where} \quad f = \frac{0,1,2,\ldots,N-1}{N} \quad (4.1)
\]

\[
Y(f) = H(f)X(f) \quad (4.2)
\]

\[
y(t) = \sum_{f=0}^{(N-1)/N} X(f)e^{-i2\pi ft} \quad (4.3)
\]

As can be seen in (4.2), the FRF regulates the transfer of “spectral energy” from the input to the output series (the raw spectrum is proportional to the FFT of a time series multiplied by its conjugate). The FRF is defined, using filter weights denoted as \( h(\tau) \), as follows:

\[
H(f) = \sum_{\tau=a}^{\tau=b} h(\tau)e^{i2\pi f\tau} \quad (4.4)
\]

Typically, however, time series filtering is computed in time space. This requires a convolution between the input time series and the filter weights:
\[ y(t) = \sum_{\tau=a}^{\tau=b} h(\tau) x(t + \tau) \] (4.5)

The parameters \(a\) and \(b\) are integers that are usually chosen such that \(a\) is the negative of \(b\). Thus, the number of filter weights becomes \(2b+1\). The parameter \(\tau\) represents a time lag. Filtering in the time domain results in lost points in the left and right endpoint intervals; this is a consequence of the convolution. Specifically, the convolution cannot be defined at the first \(b\) points and the last \(b\) points of the time series. In these regions, at least one of the lags is associated with an unavailable point in the time series (points beyond the terminal values). These points where the full convolution cannot be computed are customarily dropped from consideration. In the present study, variable filter weights in the endpoint intervals are determined such that the FRF optimally (in the least squares sense) reproduces the predetermined FRF of the interior points.

Data and Methods

Sample time series

The present investigation uses 3 climate mode time series: an AO time series, an ENSO time series, and an MJO time series. The AO is defined using empirical orthogonal function (EOF) analysis (see Thompson and Wallace 1998). The AO time series is defined as the leading PC time series of NCEP/NCAR reanalysis SLP north of 20°N from 1948-2002 (Fig. 4.1a). Since the AO definition utilizes EOF analysis and does not make any assumptions on the timescale, the time series has spectral energy spread across the frequency domain. Therefore, filtering of the AO time series is utilized to extract the component of AO variability within a particular frequency range. In the present investigation, our AO filter will be a low-pass filter comprised of 24 consecutive passes of the 1-2-1 filter. This results in a 49-point Gaussian filter with an FRF shown in Fig. 4.1b.

ENSO indices are computed in numerous different ways. However, this is ordinarily accomplished by averaging SST anomalies over a given domain box in the tropical Pacific. A commonly used indicator region is the Niño 3.4 region (120°W-170°W, 5°S-5°N). Monthly
ERSST values (see Smith and Reynolds 2003), binned into 2° by 2° grid boxes, from 1946-2000 are utilized to compute the box averaged time series for the Niño 3.4 region (Fig. 4.2a). The most common filters applied to SST-based ENSO indices are running means of 3 or 5 months. In the present investigation, a 5-month running average is used (FRF shown in Fig. 4.2b).

Unlike the AO and ENSO, the MJO is unique because its timescale is its defining characteristic. Therefore, a fundamental difference between time series of the AO and ENSO is that MJO time series are customarily defined post-filtering. Alternatively, the MJO can be defined using EOF-type analysis of non-filtered data (see the MJO section). However, that does not assist in the present investigation. The MJO was first observed in station time series (Madden and Julian, 1971) using spectral analysis. In the present investigation, a time series must be obtained that reveals intraseasonal variability after bandpass filtering. Following the pioneering work of Madden and Julian (1971), this is accomplished by using a point source in the MJO domain. The time series is obtained from a gridded QuikSCAT zonal velocity field (see Pegion et al. 2000) at the equator and 80°E (Fig. 4.3a). The time series is comprised of 392 pentad averages from late July 1999 through early December 2004. Based on the results from the MJO study, 31 Lanczos weights are utilized (see FRF in Fig. 4.3b).

**Generating Multiple Samples**

Investigating the impact of the endpoint scheme on a particular climate time series may provide anecdotal evidence of its efficacy, but a significantly larger sample size is necessary to increase confidence in the filtering method. On the other hand, using a large number of purely random time series may help in terms of increased confidence, but it would also be beneficial to create daughter time series that, in a spectral average sense, are acceptably similar to the mother time series. In order to accomplish both goals, a manipulation of white noise is utilized (see Schoof et al. 2005). For each of the three climate indices, the fast Fourier transform (FFT) is computed. The resulting function will be utilized as the FRF to be imposed on white noise time series.

Ten thousand randomly generated, uniformly distributed time series are created, representing white noise. Each new sample is passed through an FFT and multiplied by the FRF before inverse Fourier transforming back into the time domain. Finally, the ten thousand
daughter time series are normalized using each individual mean and standard deviation. Note that this protocol is analogous to filtering in spectral space except that, here, the FRF is used to impose, rather than regulate, the spectral characteristics of the output time series. Therefore the only difference is in scientific intent; the mathematics is completely identical.

The Least Squares Filtering Schemes

As stated earlier, the problem with filtering in the time domain is that the full convolution cannot be computed at the endpoints of $x(t)$. This begs the following question: can variable-length filter weights be determined in the endpoint intervals such that the signal extraction replicates that observed in the interior points? For this project, a penalty function is constructed that minimizes the squared error between the FRF in the interior and a new FRF in each slot of the endpoint intervals. Following (4.4) the FRF of interior points are denoted as follows:

$$G(f) = \sum_{k=-b}^b \alpha(k)e^{i2\pi kf}$$

(4.6)

where $\tilde{\alpha}$ represents the new filter weights to be determined. The value of $b$ is incrementally decreased to force a symmetric filter at each point (i.e. the number of filter weights for a particular point in the time series will always equal twice the distance to the terminal end plus one). The result is a set of filter weights for each point in the endpoint interval. It can be shown that, unconstrained, this minimization technique will simply truncate the filter weights used in the interior (see Bloomfield 2000). A practical constraint is to force the new weights to sum to 1 (see Argüez et al. 2005), thereby preserving the expectation of the mean of the output series. The cost function, $J$, for a given point $L$ in the endpoint interval is a function of $\tilde{\alpha}$ and $\lambda$ (the Lagrange multiplier that imposes our constraint):

$$J(\tilde{\alpha}, \lambda) = \sum_f [H(f) - G(f)]^2 + \lambda \left[ \sum_{\tau=-b}^b h(\tau) - \sum_{k=-L}^L \alpha(k) \right]$$

(4.7)
The minimum of this penalty function is obtained by taking the partial derivatives with respect to each \( \alpha \) and \( \lambda \) and setting them equal to zero:

\[
\frac{\partial J}{\partial \alpha_j} = -2 \sum_f [H(f) - G(f)]e^{i2\pi f} + \lambda = 0
\]

\( (4.8) \)

\[
\frac{\partial J}{\partial \lambda} = \sum_{r=a}^{b} h(r) - \sum_{k=L}^{L} \alpha(k) = 0
\]

\( (4.9) \)

Equations (4.7) and (4.8) are solved simultaneously using simple matrix operations to arrive at the coefficients.

An alternative, albeit more complicating, constraint is to force the total variance to be preserved. Variance is related to the spectrum via the following relation:

\[
\sigma^2_y = \int S_{yy}(f) df
\]

\( (4.10) \)

Combining (4.10) with (4.2) and discretizing, the following equality will be imposed to insure that total variance in the interior will match the total variance of each point in the endpoint interval:

\[
\sum_f H(f)H^*(f)X(f)X^*(f) = \sum_f G(f)G^*(f)X(f)X^*(f)
\]

\( (4.11) \)

The superscript asterisk represents complex conjugates. Note that the constraint is dependent on the input time series. Modifying (4.7) to impose the new constraint, and incorporating the definition in (4.6), results in the new cost function:

\[
J(\tilde{\alpha}, \tilde{\lambda}) = \left\{ \sum_f \left[ H(f) - \sum_{k=L}^{L} \alpha(k)e^{i2\pi f} \right]^2 \right\} + \lambda \left\{ T - \sum_{f} \left[ S_{xx}(f) \left[ \sum_{j=L}^{L} \sum_{m=-L}^{L} \alpha(j)\alpha^*(m)e^{i2\pi f(j-m)} \right] \right] \right\}
\]

\( (4.12) \)
where $T$ is the variance of interior points and is proportional to the LHS of (4.11). As before, the minimum is found by partial differentiation and setting to 0:

$$\frac{\partial J}{\partial \alpha_n} = -2 \sum_f [H(f) - \sum_{k=L}^L \alpha_k e^{2\pi fk} e^{2\pi fn} + \lambda \sum_f \left( 2S_{xx}(f) \sum_{j=L}^L \alpha_j \cos(2\pi[f(n-j)]) \right)] = 0 \quad (4.13)$$

$$\frac{\partial J}{\partial \lambda} = T - \sum_f \left( S_{xx}(f) \left[ \sum_{m=-L}^L \sum_{p=-L}^L \alpha_m \alpha_p e^{2\pi f(m-p)} \right] \right) = 0 \quad (4.14)$$

These equations are not solvable with simple matrix manipulation because of nonlinearities between the Lagrange multiplier and the coefficients in (4.13) and between the coefficients in (4.14). The equations are solved using the Newton method available in the IDL programming language.

**Test Methodology**

The equal-mean and equal-variance methods will be applied to subsets of time series of the MJO, the AO, and ENSO. These “estimates” will be compared to the “true” filtered values. The “true” values are computed by allowing the filter to access actual data values beyond the subsets’ endpoints. For example, say we have 1000 points in our full time series. We want to use 101 pre-determined filter weights, such that $b$ is 50. The first and last 50 values are removed, leaving a 900-point subset. The subset is filtered using the least squares techniques. At the endpoints of the subset, it is blindly disregarding any points beyond its view (the first and last 50 points of the 1000-point time series). The “true” values are those computed using the same 101 weights over the 1000-point series. Of course, the full convolution cannot be computed at the 1000-point series’ endpoints, so only 900 points will remain. The middle 800 points will be exactly the same as the middle 800 points from the least square simulations. However, the 100 endpoints (50 on each side) of the subset will be different. We term these regions where estimates and true values are compared the “estimation zones.”

To summarize the hypothetical case, the outer 50 points on each side of the 1000-point time series are only used to compute the true values in the estimation zones. The next 50 points on each side comprise the estimation zones where true and estimated values are computed,
leaving 800 points in the middle where the filtered values will be identical. In general, the length of the subset will be $N-2b$, where $N$ represents the length of the particular sample climate series in question. In addition, there will be a total of $2b$ estimated values, with $b$ number of estimates at each end.

In order to assess the viability of both filtering schemes, the root mean square error (RMSE) is computed at each point in the estimation zones. Since the methodology does not impart any preference to the left or the right estimation zones, the RMSE values should be symmetric between the left and right zones. In addition, the RMSE values will also be computed for the unconstrained least squares scheme as well as for the spectral filtering method. A separate RMSE figure is computed for each of the climate series: the AO, the MJO, and ENSO.

The errors due to the least square methods increase closer to the terminal points due to two primary reasons. First, less data are available for the computation (data beyond the terminal ends are blindly disregarded). These errors are not quantifiable, since there are no restrictions on the particular values a time series can take beyond the endpoints. Secondly, the ability to reproduce the frequency response function is hampered by a reduced number of allowed filter weights. The second source of error is especially problematic for bandpass filters, such as the Lanczos filter used in the MJO project, since the weights must include both positive and negative weights to extract salient timescales.

**Results**

**AO**

The RMS errors for the AO filtering project suggests that all 4 filtering schemes work very well in the interior of the estimation zone (Fig. 4.4). The expected RMSE between filtered values of randomly selected input series (using the same procedures as above) is about 0.4, representing a liberal skill threshold. Except for the equal-mean filter, the filters prove to be useful for all positions. The equal-variance, unconstrained, and spectral methods behave in a remarkably similar fashion for the AO filtering scheme.

Typical results are shown in Fig. 4.5. These cases are determined by finding the median in mean absolute error over all the simulations. Several items are worth pointing out. All 3 least
squares methods become erratic toward the end; this is a consistent behavior throughout the 10000 cases. It is possible to empirically determine a smoothing function to make adjustments to the output and decrease the RMSE near the terminal ends. The dashed blue line in Fig. 4.4a represents such an endeavor for the equal variance filter. In essence, the least squares filtering schemes can be considered first guesses than can be modified by an empirically determined adjustment.

In addition, it is important to comment on the relative success and failure of the unconstrained and equal-mean filters near the terminal ends, respectively. The equal-mean filter suffers in part because the sum of the filter weights used to compute all 3 interior FRF functions ($H$) was set to 1 (this can be changed and will be considered in future work). This means the 2 terminal values do not change because they are simply multiplied by 1. As a result, when the input subset ends at a series maximum or minimum, the errors with the equal-mean method are disproportionately large. The unconstrained method, on the other hands, performs very well (it will be shown that this is not a robust result). However, here we have somewhat of the opposite issue occurring. Instead of constraining the weights to equal a certain value or maintaining a variance, these weights are simply truncated. This results in decreased variance close to the terminal ends. In the case of the AO project, this reduction in variance proved to be to an appropriate level to allow strong performance. However, if the level of variance reduction is not appropriate, large errors can result (e.g. in the ENSO project).

**ENSO**

For the ENSO project, the equal variance method performs best, followed closely by the equal mean method (Fig. 4.5). By far the worst performer is the unconstrained method. In this case, domain-averaged SST is used to define the ENSO time series. Unlike other ENSO indicators, such as the noisy Southern Oscillation Index (a pressure difference between Darwin, Australia and Tahiti), SST-based indices have disproportionately lesser amounts of high-frequency variability, attributable to the high heat capacity and inertia of the ocean (see Mizoguchi et al. 1999). Therefore, SST indices have much longer decorrelation times (a widely used metric of serial autocorrelation) than atmospheric indices; longer decorrelation times, in turn, are linked to potential predictability.
In this project, only 2 estimates are computed per times series end for each filtering method tested. For the unconstrained case, all the weights remain at 0.2 (recall the 5-month running mean whose weights sum to 1) since this method simply truncates the weights. For the equal mean case, the weights become 1/3 when computing the 2nd and penultimate points, and 1 for computing the terminal ends. Therefore, the unconstrained case tends to result in values that veer toward 0 near the terminal ends, as exemplified in Fig. 4.7. On the other hand, the other methods remain closes to the original pre-filtered values.

MJO

Filter performance for the MJO project is relatively uniform amongst the different methods (Fig. 4.8). The liberal skill threshold in this case is about 70, suggesting that all methods are useful for all time series positions. The unconstrained and equal-variance methods are the leading performers in the inner two-thirds of the estimation zone, while the spectral method outperforms near the terminal points. Looking at typical results (Fig. 4.9), the aforementioned variance issue at the terminal ends is problematic for the equal mean and unconstrained methods, with both evincing outputs that veer toward zero at the terminal points. In other words, both methods are plagued by reduced variance toward the terminal points. The equal-variance method’s RMSE are similar near the endpoints to the equal-mean and unconstrained methods, but it does not exhibit the veering to zero characteristic: the equal-variance method results are neither partial to overshooting nor overshooting the true values.

Summary

In the present investigation, different options were explored for retaining time series endpoints when filtering. In addition to the spectral method, three variations of a least squares technique were also utilized. As one would expect, the resulting errors are typically greater for the positions closest to the terminal ends. The most consistent least squares technique was the equal-variance method, which is also the most computationally expensive. By imposing the variance in the estimation zones, this method was not prone to over- or under-shooting the true output, an attribute that was present to certain degrees with both the unconstrained and equal-
mean methods. In addition, the equal-variance method is unique because not only does it depend on the filter weights (like the other methods), it also is customized to a particular time series input through the constraint. Overall, the equal-variance scheme was on par with the others (including the spectral method) for the AO and MJO projects, and was the best performer for the ENSO study. It is the author’s contention that the equal-variance method is a suitable alternative to simply dropping endpoints, a common practice in the geosciences.

Wheeler and Hendon (2004) point out that the major obstacle of real-time monitoring and prediction of the MJO, for example, is the loss of endpoints due to bandpass filtering. In the case of real-time monitoring of any climate time series, the equal-variance method provides a useful “first guess” without having to wait for future values to be recorded. In this regard, the equal-variance filtering scheme is an attractive alternative to the more indirect methods used for real-time monitoring (e.g. see Wheeler and Hendon 2004). However, it should be noted that, in the case of the MJO, the spectral method is also a suitable alternative. How the methods described in the present investigation compare to Wheeler and Hendon’s (2004) EOF-based monitoring algorithm is worthy of a detailed inspection.

It is important to note that the techniques described herein do not involve time series prediction in any manner. Future work may involve combining the least squares technique with a statistical prediction method that takes on increasing weight closer to the terminal ends. Simple autoregressive forecasting (and hindcasting) showed promise near the endpoints in preliminary analyses toward that end, both when extending the output (filtered) series with a forecast, as well as forecasting the subset (input) series in order to compute the full convolution.
Figure 4.1: AO Filter Series. AO time series utilized (top) and the FRF of the filter used on it (bottom).
Figure 4.2: **ENSO Filter Series.** As in Fig.4.1 but for ENSO.
Figure 4.3: **MJO Filter Series.** The MJO time series (top), the filter weights used (middle), and the associated FRF (bottom).
Figure 4.4: **AO RMS Errors.** Root mean square errors as a function of position for the unconstrained (purple), equal-mean (green), equal-variance (blue), and spectral (red) methods for the AO project (top). A liberal skill threshold based on random filter output is an RMSE value of about 0.4. The empirical adjustment to the equal-variance filter is shown as a blue dashed line. The bottom panel shows the least square techniques’ RMSE values divided by the spectral RMSE values.
Figure 4.5: **Typical AO Results.** Typical results (see text) for the AO case.
Figure 4.6: **ENSO RMS Errors.** As in Fig. 4.4, but for the ENSO project (skill threshold equals 0.63).
Figure 4.7: Typical ENSO Results. As in Fig. 4.5, but for the ENSO case.
Figure 4.8: **MJO RMS Errors.** As in Fig. 4.4, but for the MJO case (skill threshold equals 71.2).
Figure 4.9: **Typical MJO Results.** As in Fig. 4.5, but for the MJO case.
CONCLUSION

In the current investigation, four independent projects are presented. All four deal with aspects of climate mode definitions, time series indexing, and signal extraction. The first project dealt with a cyclo-stationary definition for the AO. Several important issues were considered, including the retention parameter (the number of EOF modes used in the CSEOF analysis), CSEOF analysis of a single EOF mode only, and the contributions that individual EOF modes make to a CSEOF pattern. CSEOF analysis of the AO mode (the leading EOF) produced a set of 12 spatial maps that were similar to the AO map, except that the amplitude changed from one month to another. The resulting PC time series had a peak in the vicinity of 2.5 to 3 years, and the monthly variances were virtually the same. In contrast, the AO mode has only one spatial pattern, and the PC time series is substantially noisier, with monthly variances clearly larger in the winter months. This exemplifies the power of CSEOF analysis to transfer a physical evolution from a PC time series to the series of maps. A compelling case is made for a cyclo-stationary definition for the AO, rather than the current EOF definition.

The second project dealt with the detection of the MJO signal in QuikSCAT, a relatively new satellite-based dataset. The MJO is not apparent by inspecting unfiltered spatial maps. The need to extract intraseasonal timescales before observing the MJO signature has prompted some degree of skepticism in the scientific community regarding the MJO’s significance. Extended EOF analysis was conducted on the non-filtered zonal wind data, revealing a clear MJO signal in modes 4 and 5, which are in quadrature with each other. EOF analysis of the non-filtered zonal wind was not successful at uncovering an MJO signal, since EOF analysis (due to being relegated to one map per mode) is not capable of distinguishing between the MJO spatial pattern and other signatures (such as ENSO) that have common spatial features. No other multivariate method has been shown to detect the MJO signal in non-filtered data in the literature, to the best of the authors’ knowledge. This project shows the power of Extended EOF’s for extracting propagating features (in non-filtered data) that would otherwise be inseparable from other modes.

The third project involved incorporating the ‘impact perspective’ into our approach to ENSO impact studies. This resulted in practical, targeted indicator regions for particular impact
grid points or regions. Climate composites were computed as functions of local SST indicator mode, indicator region, and impact zone. EOF analysis of the warm phase and cold phase functions revealed the most energetic indicator-impact pairs for a particular variable and phase. Precipitation patterns were shown to be much better behaved than the air temperature EOF patterns. This unique method of EOF decomposition accounts for the La Niña air temperature pattern over the United States when the Niño 1+2 region is used in conventional composite analysis: the two leading modes superpose over the Niño 1+2 region to explain the significant cooling over the eastern two-thirds of the United States. In addition, the EOF method revealed a heretofore unobserved teleconnection between the tropical Pacific “warm phases” and Southeast U.S. air temperatures. This project underscores the difficulty involved in selecting indicator and impact domains and constructing indices. It also reveals the usefulness of using old techniques (EOF analysis) in new ways to enhance our understanding of climate modes.

The fourth and final project considered the development of a least squares technique to retain endpoints that are normally discarded when filtering in the time domain due to the use of a convolution. Sample time series were created by manipulating random numbers that represent white noise. Three constraints were used with the least squares method: equal-variance, equal-mean, and no constraint. The equal-mean and unconstrained versions had problems appropriating variance at the terminal ends, resulting in very large RMS errors. The equal-variance version, however, performed very well across time series positions and for all 3 time series considered. It outperformed all others in the ENSO test, which involved a five-month running mean. It performed as good as or better than the others, including the spectral method, in the AO and MJO cases for virtually all time series positions. It is the author’s contention that the equal-variance method should be utilized (up to the uncertainty tolerance of a given user) when filtering instead of dropping points needlessly. The method can also be combined with predictive tools to improve real-time monitoring, for example. This project exemplifies the application of new tools to combat the side effects of statistical techniques.

The results of the four projects clearly demonstrate the utility of an assortment of statistical tools to improve our understanding of climate modes. The projects also illustrate the importance of index construction (the ENSO project), mode definition (the MJO and AO projects), and signal extraction (the filtering project) in climate research. Climate researchers have a tremendous array of statistical techniques in their arsenals. This abundance of tools must
be exploited in order to continue to increase our level of understanding of climatic phenomena. As the present work strives to portray, applying new techniques, or simply applying old techniques in new ways, can provide additional insight into old problems.
REFERENCES


BIOGRAPHICAL SKETCH

Anthony Argüez was born in Miami, FL in 1978. The son of Angela and Antonio Argüez, who fled Cuba in the early 1960s to escape persecution, Anthony is the youngest of five children. He was the first member of his family to attend a university. Anthony knew he wanted to study meteorological phenomena when Hurricane Andrew devastated his community in 1992. Four years later, he moved to Tallahassee, FL and began his journey at the Florida State University. During his undergraduate years, Anthony began working at the Center for Ocean-Atmospheric Prediction Studies, under the supervision of Dr. James J. O’Brien, who would later become his major professor in graduate school. Anthony graduated summa cum laude and with honors in 2000 with a BS in Meteorology and a BS in Geography. He continued at FSU and COAPS and earned his MS in Meteorology in 2002. Anthony recently accepted a federal position in the Climate Analysis Branch of NOAA’s National Climatic Data Center, located in Asheville, NC.

Publications:

